



# **Balancing of Associated Attributes for a Viable Indoor Real-time Location System with Obstruction Detection**

Submitted in fulfilment of the requirements of the degree of

**Doctorate in Information and Communications**

**Technology**

Faculty of Accounting and Informatics

Durban University of Technology

Jeebodh Pancham

Date Submitted: 1<sup>st</sup> November 2022

Supervisor: Prof. R. C. Millham (PhD)

Date: 01/11/2022

Co-supervisor: Prof. S. J. Fong (PhD)

Date: 01/11/2022

## DECLARATION

I, Jeebodh Pancham, declare that this dissertation is a representation of my own work in both conception and execution.

Student

1 November 2022

Date

## Approval for examination

Supervisor

Prof R. C. Millham (PhD)

1 November 2022

Date

\_\_\_\_\_  
Supervisor

Prof S. J Fong (PhD)

1 November 2022

Date

## LIST OF PUBLICATIONS

1. Pancham, J., Millham, R. C., & Fong, S. J. (2021). Determining the Impact of Perspex Obstructions on Bluetooth Transmission Paths Within a Simulated Office Environment for More Accurate Location Determination. *In: Lecture Notes in Computer Science*, (pp. 410-420). Springer, Cham. doi: [10.1007/978-3-030-87013-3\\_31](https://doi.org/10.1007/978-3-030-87013-3_31)
2. Pancham J., Millham R., Fong S.J. (2020) Analysis of Bluetooth Low Energy RSSI Values for Use as a Real-time Link Quality Indicator for Indoor Location. *In: Lecture Notes in Computer Science*, vol 12254. Springer, Cham. [doi.org/10.1007/978-3-030-58817-5\\_69](https://doi.org/10.1007/978-3-030-58817-5_69)
3. Pancham, J., Millham, R., & Fong, S. J. (2018). Analysis of Bluetooth Low Energy Detection Range Improvements for Indoor Environments. *In: Lecture Notes in Computer Science* (pp. 598-609). Springer, Cham. doi: [10.1007/978-3-319-93713-7\\_56](https://doi.org/10.1007/978-3-319-93713-7_56)
4. Pancham, J., Millham, R., & Fong, S. J. (2018). Investigation of Obstructions and Range Limit on Bluetooth Low Energy RSSI for the Healthcare Environment. *In: Lecture Notes in Computer Science*. (pp. 261-274). Springer, Cham. doi: [10.1007/978-3-319-95171-3\\_21](https://doi.org/10.1007/978-3-319-95171-3_21)
5. Pancham, J., Millham, R., & Fong, S. J. (2018). A Scalable Bluetooth Low Energy Design Model for Sensor Detection for an Indoor Real-time Location System. *In: Lecture Notes in Computer Science* (pp. 317-330). Springer, Cham. doi: [10.1007/978-3-319-95171-3\\_25](https://doi.org/10.1007/978-3-319-95171-3_25)
6. Pancham, J., Millham, R. and Fong, S. J. 2017. Assessment of feasible methods used by the health care industry for real-time location. *FedCSIS (Position Papers)*, 129–134.
7. Pancham, J., Millham, R. and Fong, S. J. (2017). Evaluation of real-time location system technologies in the health care sector. In *Proceedings of the 2017 17th International Conference on Computational Science and its Applications (ICCSA)*. IEEE. Trieste Italy. doi: [10.1109/iccsa.2017.7999645](https://doi.org/10.1109/iccsa.2017.7999645)

**Note:** Many of the expressions used in my papers to explain principles and descriptions have been integrated into this thesis. Consequently, there may be phrases in the thesis that are similar to those in published articles.

## ACKNOWLEDGEMENTS

Professor Richard C. Millham of the Department of Information Technology at Durban University of Technology provided invaluable direction and assistance throughout this research. This journey has been made interesting by Professor Millham's emphasis on practicality and industrial applicability, as well as his continual guidance in the right direction. I would also like to place on record my sincere appreciation to my co-supervisor, Professor Simon Fong from the University of Macau, for his advice, direction and recommendations throughout this study.

Thank you to Ms. Sara Bibi Mitha for her continual assistance in correcting the citations and formatting of the thesis.

Lastly, I must express my deepest appreciation to my wife (Nasha) and children (Rivajh, Shivash and Shreeya) for their unwavering support and constant encouragement during my years of education and the research and writing process. Without them, this achievement would not have been possible. Thank you.

## ABSTRACT

Outdoor location determination is often achieved by using GPS, but indoor location determination is not possible with GPS due to the limited link to satellites from indoor environments. Research has indicated that indoor location determination is applicable in a variety of domains, including asset location, location of people, emergency evacuation, participant attendance in a venue, tracking of people's mobility, and obstruction location. Cost, energy consumption, interference, coverage, detection range and form factor are some of the constraints reported in the literature. The attributes for further research were derived from these constraints. Bluetooth Low Energy was identified as the most suitable technology with which to design a model that would achieve an optimal balance between the identified attributes. The research used Unified Modelling Language to document the model and design science methodology to design, test and validate the model.

The introduction of obstructions in the path of transmission often affects the received signals and hence affects the location determination. The connection quality indicator was used in this model to determine location instead of the widely used fingerprint method, whose data becomes unusable over time as it becomes stale and inaccurate. The design was tested with a variety of obstructions, including drywall partitions, glass, solid brick walls, metal sheets and Perspex, all of which were utilised to resemble a typical office environment. The received signal strength indicator measurements from low power nodes were filtered and smoothed using mean, median and mode statistics. This received signal strength indicator data was then used by support vector machine, k-nearest neighbour and artificial neural network machine learning models to determine the location and impact of the obstructions in the path of Bluetooth Low Energy transmission.

The results obtained from machine learning and prediction revealed that the location of obstructions was determined to be within an acceptable level of accuracy. In particular, k-nearest neighbour performed the best compared to support vector machine and artificial neural network using the mean squared error, mean absolute error, root mean square error and R-squared score metrics. In particular, mean squared error and mean absolute error metrics revealed the best results. The study indicates that machine learning can therefore be used to determine positions of semi-fixed obstructions within a select indoor environment.

## TABLE OF CONTENTS

DECLARATION .....	ii
LIST OF PUBLICATIONS .....	iii
ACKNOWLEDGEMENTS .....	iv
ABSTRACT.....	v
TABLE OF CONTENTS.....	vi
LIST OF FIGURES .....	xi
LIST OF TABLES .....	xiv
LIST OF ACRONYMS .....	xvi
CHAPTER ONE: INTRODUCTION.....	1
1.1 Background.....	1
1.2 Problem statement.....	2
1.3 Aim and objectives .....	2
1.4 Research questions.....	3
1.5 Contributions .....	3
1.6 Success criteria .....	3
1.7 Significance of the study.....	4
1.8 Scope.....	4
1.9 Limitations .....	5
1.10Chapter summaries .....	5
CHAPTER TWO: LITERATURE REVIEW .....	6
2.1 Introduction.....	6
2.2 Real-time location systems .....	6
2.3 Typical use cases of indoor real-time location systems.....	7
2.4 Constraints of indoor real-time location systems .....	8
2.4.1 Interference and range .....	9

2.4.2	Coverage and scalability .....	10
2.4.3	Power consumption .....	11
2.4.4	Cost .....	12
2.5	Indoor real-time location system technologies .....	13
2.5.1	Wi-Fi .....	14
2.5.2	Radio-frequency identification .....	15
2.5.3	Near-field communication .....	17
2.5.4	Device-free localisation.....	17
2.5.5	Bluetooth Classic.....	18
2.5.6	Bluetooth Low Energy.....	18
2.5.7	Hybrid technologies .....	21
2.6	Evaluation of technologies.....	22
2.7	Design models.....	23
2.8	Location determination techniques .....	24
2.8.1	Angle of arrival, time of arrival or time of flight .....	25
2.8.2	Trilateration and Kalman filter.....	25
2.8.3	Fingerprint.....	27
2.8.4	Received signal strength indicator .....	27
2.9	Machine learning .....	29
2.9.1	Artificial neural network .....	30
2.9.2	Bayesian network.....	31
2.9.3	Naïve Bayes.....	31
2.9.4	Decision tree .....	31
2.9.5	Random forest .....	32
2.9.6	k-nearest neighbour.....	32
2.9.7	Support vector machine.....	33
2.9.8	Metrics for evaluation.....	34

2.9.9 Results of use cases of machine learning for location determination.....	36
2.10 Effect of obstructions.....	38
2.11 Research gap.....	38
2.12 Summary.....	39
CHAPTER THREE: Methodology.....	41
3.1 Introduction.....	41
3.2 Paradigm and methods.....	41
3.2.1 Research paradigm.....	41
3.2.2 Research design.....	43
3.2.3 Research methods.....	44
3.3 Application of methodology.....	45
3.3.1 Design science methodology.....	45
3.3.2 Application of design science methodology.....	47
3.3.3 Design process.....	48
3.4 Methods.....	50
3.5 Technologies.....	51
3.6 Environment.....	51
3.7 Experimental setup.....	52
3.7.1 Bluetooth Low Energy.....	52
3.7.2 Received signal strength indicator.....	53
3.7.3 Hardware.....	53
3.7.4 Location determination.....	56
3.8 Design processes and configuration.....	56
3.8.1 Data collection.....	56
3.8.2 RTLS model design notations.....	58
3.8.3 Machine learning.....	58
3.9 Summary.....	62

CHAPTER FOUR: RESULTS .....	63
4.1 Introduction.....	63
4.2 iRTLS model.....	63
4.2.1 Requirements of an indoor real-time location system .....	63
4.2.2 Analysis .....	67
4.2.3 Design.....	69
4.3 Experiments .....	79
4.3.1 Throughput and range determination .....	79
4.3.2 Transmission through obstructions .....	84
4.3.3 Real-time quality link indicator .....	90
4.4 Effect of obstructions.....	94
4.4.1 Environment and obstructions .....	94
4.4.2 Fingerprinting.....	101
4.4.3 Trilateration .....	101
4.5 Machine learning .....	101
4.5.1 Experimental layout.....	102
4.5.2 Training of the model and prediction of obstruction location .....	104
4.6 Summary.....	115
CHAPTER FIVE: CONCLUSION.....	116
5.1 Introduction.....	116
5.2 Summary of the analysis of results .....	116
5.2.1 iRTLS technologies .....	116
5.2.2 Throughput and range .....	116
5.2.3 Transmission through obstructions .....	118
5.2.4 Real-time quality link indicator .....	119
5.2.5 Improvement of location determination in a dynamic environment.....	120
5.2.6 Indoor real-time location system model .....	120

5.2.7 Machine learning for obstruction location determination.....	120
5.3 Addressing the study objectives .....	122
5.4 Addressing the research questions.....	122
5.5 Success criteria and contribution of the research.....	123
5.6 Limitations.....	123
5.7 Future work.....	123
5.8 Summary.....	124
REFERENCES .....	126
APPENDICES .....	144
Appendix A - Use Cases .....	144
Appendix B - Measurement of Throughput.....	148
Appendix C – Results from Perspex Obstruction.....	150
Appendix D – Results from Metal Sheet Obstruction .....	158
Appendix E – Results from Machine Learning .....	170

## LIST OF FIGURES

Figure 2.1: Technologies for localisation .....	13
Figure 2.2: BLE channels .....	19
Figure 2.3: Indoor localisation techniques.....	24
Figure 2.4: Trilateration .....	26
Figure 2.5: Commonly used techniques in machine learning.....	29
Figure 3.1: Nominal design science process.....	46
Figure 3.2: Application of design science.....	47
Figure 3.3: Design process for an iRTLS model .....	49
Figure 3.4: IoT BLE architecture.....	53
Figure 3.5: Nordic nRF52840 PDK.....	54
Figure 3.6: SKB501-LPN top view .....	55
Figure 3.7: SKB501-LPN rear view .....	55
Figure 3.8: Architecture .....	56
Figure 3.9: Layout of nodes and gateways .....	57
Figure 3.10: Application user interface .....	58
Figure 3.11: Sample machine learning training network.....	60
Figure 3.12: Sample machine learning prediction network.....	61
Figure 4.1: Legend of entities in detect node scenario .....	64
Figure 4.2: Node use case diagram.....	66
Figure 4.3: Gateway use case diagram .....	66

Figure 4.4: Activity diagram and scenario processing.....	68
Figure 4.5: Node and gateway messaging .....	68
Figure 4.6: Subset domain model diagram .....	69
Figure 4.7: High-level BLE mesh architecture .....	70
Figure 4.8: Open-plan office.....	81
Figure 4.9: Boardroom.....	82
Figure 4.10: Partitioned office .....	82
Figure 4.11: RSSI levels .....	83
Figure 4.12: Throughput of 1 Mb of data .....	84
Figure 4.13: Adjacent offices.....	85
Figure 4.14: Meeting room .....	85
Figure 4.15: RSSI levels received from position TX1.....	87
Figure 4.16: RSSI levels received from position TX2.....	87
Figure 4.17: RSSI levels between two adjacent offices.....	88
Figure 4.18: RSSI levels between TX1 and RX1 across single partition .....	88
Figure 4.19: RSSI levels between TX2 and RX1 across double partition.....	89
Figure 4.20: RSSI levels between TX2 and RX2 double partition.....	89
Figure 4.21: RSSI levels between TX3 and RX2 across double partition.....	90
Figure 4.22: RSSI levels between TX4 and RX2 across double partition.....	90
Figure 4.23: LPN and gateway layout .....	91
Figure 4.24: LPN002 messages received at GW-001 .....	92

Figure 4.25: LPN002 messages received at GW-002.....	92
Figure 4.26: LPN002 messages received at GW-003.....	92
Figure 4.27: LPN002 messages received at GW-004.....	93
Figure 4.28: Mode at GW-001 from LPN002 .....	93
Figure 4.29: Mode at GW-002 from LPN005 .....	94
Figure 4.30: Office.....	94
Figure 4.31: Fabricated metal obstruction .....	95
Figure 4.32: Received RSSI through Perspex.....	95
Figure 4.33: Metal obstruction at (X, Y) position (0.5, 2).....	97
Figure 4.34: Metal obstruction at (X, Y) position (3.5, 2).....	98
Figure 4.35: Effect of fabricated obstruction.....	100
Figure 4.36: Gateways and node layout for machine learning .....	102
Figure 4.37: SVM Parameters setup for training and prediction.....	105
Figure 4.38: KNN parameters setup for training and prediction .....	107
Figure 4.39: ANN parameters setup for training and prediction .....	108

## LIST OF TABLES

Table 2.1: Technology attributes .....	22
Table 3.1: Four worldviews (Creswell 2013) .....	42
Table 4.1: Business use case Determine Location.....	65
Table 4.2: Use case LPN Detect Motion .....	71
Table 4.3: Use case LPN Detect Inactivity Timeout Interval.....	71
Table 4.4: Use case LPN Acknowledge Receipt .....	72
Table 4.5: Use case Gateway Received Message .....	72
Table 4.6: Use case Gateway Detect Time Interval.....	73
Table 4.7: Configurable parameters for LPN .....	73
Table 4.8: Configurable parameters for LPN .....	74
Table 4.9: Receive Message definition at gateway.....	75
Table 4.10: Main test case list.....	78
Table 4.11: Throughput .....	80
Table 4.12: Extract from detailed RSSI values.....	96
Table 4.13: Results from Perspex obstruction .....	97
Table 4.14: Extract of results from metal obstruction .....	99
Table 4.15: Classification of different levels .....	99
Table 4.16: Results with fabricated obstruction .....	100
Table 4.17: Extract of data captured for machine learning.....	103
Table 4.18: Results of test and score parameters for SVM.....	104

Table 4.19: Results of different SVM kernels .....	104
Table 4.20: Results of different parameters with KNN .....	106
Table 4.21: Results of different parameters with ANN .....	107
Table 4.22: Experiment 1, training on dataset 1 .....	109
Table 4.23: Experiment 1: Training on dataset 1, prediction on dataset 1 and dataset 2 .....	109
Table 4.24: Experiment 1: Training on dataset 1, prediction on dataset 2 and dataset 3 .....	110
Table 4.25: Experiment 1: Training on data set 1, prediction on dataset 3 .....	110
Table 4.26: Experiment 2: Training on dataset 1 .....	111
Table 4.27: Experiment 2: Training on dataset 1, prediction on dataset 1 and dataset 2 .....	111
Table 4.28: Experiment 2: Training on dataset 1, prediction on dataset 2 and dataset 3 .....	111
Table 4.29: Experiment 2: Training on dataset 1, prediction on dataset 3 .....	112
Table 4.30: Experiment 3: Training on dataset 1 .....	112
Table 4.31: Experiment 3: Training on dataset 1, prediction on dataset 1 and dataset 2 .....	113
Table 4.32: Experiment 3: Training on dataset 1, prediction on dataset 2 and dataset 3 .....	113
Table 4.33: Experiment 3: Training on dataset 1, prediction on dataset 3 .....	113
Table 4.34: Experiment 4: Training on dataset 1 .....	114
Table 4.35: Experiment 4: Training on dataset 1, prediction on dataset 1 and dataset 2 .....	114
Table 4.36: Experiment 4: Training on dataset 1, prediction on dataset 2 and dataset 3 .....	115
Table 4.37: Experiment 4: Training on dataset 1, prediction on dataset 3 .....	115
Table 5.1: Number of successful predictions per trained dataset .....	121

## LIST OF ACRONYMS

Acronym	Description
<b>ANN</b>	Artificial neural network
<b>BLE</b>	Bluetooth Low Energy
<b>dBm</b>	Decibel-milliwatts
<b>GPS</b>	Global Positioning System
<b>IoT</b>	Internet of Things
<b>iRTLS</b>	Indoor real-time location system
<b>ISM</b>	Industrial, scientific and medical
<b>ISO</b>	International Standards Organization
<b>KNN</b>	k-nearest neighbour
<b>LPN</b>	Low power node
<b>MAE</b>	Mean absolute error
<b>MSE</b>	Mean squared error
<b>PDK</b>	Preview development kit
<b>RFID</b>	Radio-frequency identification
<b>RMSE</b>	Root mean square error
<b>RSSI</b>	Received signal strength indicator
<b>RTLS</b>	Real-time location system
<b>R<sup>2</sup></b>	R-squared score
<b>SVM</b>	Support vector machine
<b>UML</b>	Unified Modelling Language
<b>Wi-Fi</b>	Wireless fidelity
<b>WLAN</b>	Wireless local area network

# CHAPTER ONE: INTRODUCTION

## 1.1 Background

The concept of indoor location is referred to as real-time location systems (RTLS), indoor positioning systems and location-based systems. In the past several years, this has been a significant field of study, garnering considerable interest from the networking research community (Wang *et al.* 2015).

Indoor localisation has reached a turning point due to the increasing demand for location services provided by smart devices and related technology (Thaljaoui *et al.* 2015). With the development of integrated, compact and power-efficient mobile sensors, indoor positioning systems constitute one of the most actively explored fields in the domain of mobile sensing (Chandel *et al.* 2016). Such indoor location is important for many applications, such as location in large conference venues, complex hospitals, large shopping malls (Chandel *et al.* 2016) and smart libraries or workplaces where objects need to be easily located (Dwiyasa and Lim 2016). Other applications include health monitoring, intrusion detection, inventory management (Wu, Tan and Xiong 2016), navigation and tracking, marketing entertainment location-based information retrieval, and security (Basiri *et al.* 2017). Patient location is another important application of RTLS, especially in cases where patients need to be easily located (Tănase and Vulpe 2022).

Indoor positioning and navigation are currently well-researched topics as they are valuable for industry and location-based system applications (Mendoza-Silva, Torres-Sospedra and Huerta 2017). In the case of healthcare, for example, the timeous location of medical equipment will result in speedy delivery of treatment to patients. Proper asset management will lead to maximising asset usage and, hence, return on investment. Asset location and tracking can also be used to deter theft as well as to locate missing items. Tracking pathways can lead to the gathering of data for more efficient service delivery, such as using proximity sensors to manage congestion in a museum (Germak and Khan 2017).

Constraints such as cost and detection range (Tsang *et al.* 2015), energy consumption (Yu, Xu and Li 2012), size and accuracy (Deng *et al.* 2013) were identified through a systematic literature review and the discussions with industry members. Other constraints did exist but, for the purposes of this study, were not identified as the most common by researchers and

industry. In order to overcome the constraints, the proposal is to enhance methods for improved indoor real-time location system (iRTLS) research.

This improved iRTLS model must balance the selected constraints while minimising costs. For indoor location determination, the most viable technologies were employed to develop a model that provides the optimal balance between throughput, load balancing and energy consumption.

## **1.2 Problem statement**

Due to real-world constraints, many of the proposed indoor iRTLS solutions are feasible only in a laboratory environment rather than a real-world domain such as an office environment. There are several key demands of indoor location systems, viz. accuracy, real time, cost reliability and robustness (Roy and Chowdhury 2021). Zhu, Mukhopadhyay and Kurata (2012) found that many models based on academic research are not implementable and, as a result, do not aid practitioners. Organisations need accurate tracking and location of equipment and people, but such solutions are hindered by factors such as cost (Wirola, Laine and Syrjärinne 2010), form factor and electromagnetic interference (Bai *et al.* 2012; Gikas *et al.* 2016; Behboodi *et al.* 2017). Consequently, there is a need for a viable, cost-effective iRTLS model that can function optimally within the constraints of a selected domain, such as providing real-time detection of obstructions.

## **1.3 Aim and objectives**

### **Aim**

The aim of this study is to design a viable, cost-effective iRTLS model that is able, by using machine learning, to detect the location of obstructions within a simulated office environment.

### **Objectives**

The objectives of the research are:

1. to determine the most viable iRTLS technology which balances the identified real-world constraints of small form factor, low cost, reasonable accuracy and low power consumption

2. based on this technology, to determine a workable range for the purposes of location determination
3. based on this range, to develop a cost-effective model for a suitable layout of various pieces of equipment that would meet the needs of location determination within the constraints identified
4. based on this layout, to identify suitable methods to accurately determine the location of obstructions within the identified environment.

## **1.4 Research questions**

1. What is the most suitable technology which balances selected constraints that can be incorporated within a viable model of an iRTLS that provides accurate location determination within an acceptable range?
2. What is the most feasible method for obstruction location determination that is compatible with the newly developed model of an iRTLS?

## **1.5 Contributions**

The thesis aims to establish the most viable technology that satisfies the identified constraints. A model was developed using this technology and the constraints. The research culminates in the most appropriate machine learning metrics based on the machine learning method identified for obstruction location determination. The selected method was validated through relevant metrics. This machine learning method can be used to improve indoor real-time location of objects and people. Although an indoor setting like an office was considered for this research, other environments such as a hospital, an open-plan work space and partitioned offices could also be simulated. Similar constraints will be present in these environments, and there will be obstructions that interfere with signal transmission.

## **1.6 Success criteria**

A variety of attributes may be considered for incorporation within the design of an indoor location system. These attributes were reduced to six, based on a review of the relevant literature in order to design a model with which to produce a feasible indoor location system. The feasibility of this system is constrained by the characteristics of this particular domain.

According to the title of the thesis, the primary success criterion was to achieve a balance between the majority of the selected attributes.

## **1.7 Significance of the study**

As observed in the literature review, indoor location systems focus on precisely determining the location of a selected entity. However, research does not explore the location of obstructions in the path of transmission between the transmitter and the receiver. These obstructions often affect the received signal and hence affect the location accuracy. The contribution of the study is that it designs and validates a feasible, cost-effective iRTLS model with enhanced techniques for improved location of stationary obstructions. The model is accurate for the selected domain and requirements, and meets the identified constraints of indoor environments. The model is also able to balance the identified attributes of power consumption, cost and form factor.

The model mitigates some of the limitations of existing systems, such as high cost, high power consumption, high form factor, practicality and applicability. The formula that determines location uses a factor (path loss exponent) that ranges between 2 and 4 in free space, depending on the environmental conditions. However, previously this factor could not dynamically adapt to a changing environment. The model designed in this study is able to calculate this factor in real time and in such a changing environment. This model identifies the location of obstructions in a given environment by employing feasible machine learning techniques and the most pertinent metrics. Once obstructions are located, this information can be used to improve an indoor location system.

## **1.8 Scope**

Scalability was not considered because the monitored environment could grow in size. Also, scalability and determination of the shortest path could not be incorporated within the scope of this work. The shortest path from source to destination is not considered, as this process would consume more power. Therefore, simple devices were used to reduce cost and increase feasibility of a practical solution. Accuracy was traded off against speed whilst using the available resources. The focus of this research was on stationary obstructions rather than moving entities. The researcher also wanted to achieve the goal of low power consumption, that is, of minimising resource consumption.

## 1.9 Limitations

A moving entity would consume high resources. Therefore, the researcher wanted to ensure low power consumption, thus minimising resource consumption. The environment was a simulated office that represented the real world. However, this may not have displayed all the characteristics of a real-world environment as there are many variables that need to be taken into account. There was no equipment that could be used to determine the effect of transmission for location determination.

## 1.10 Chapter summaries

**Chapter 1:** This chapter presents the background, the research questions, the aims, the significance of the study and the structure of the thesis.

**Chapter 2:** This chapter reviews the relevant literature in order to analyse the current technologies and techniques applied for iRTLS. Furthermore, the attributes for consideration in designing an iRTLS are identified. Machine learning techniques are considered for location of obstructions in the Bluetooth transmission path.

**Chapter 3:** This chapter discusses the methodology used in designing the iRTLS, as well as the techniques applied to determine the position of obstructions. Design science was used to develop the required model and outline the steps followed to meet the objectives of the research.

**Chapter 4:** This chapter presents the results obtained from the research and experiments conducted, demonstrating the viability of the model. In addition, the results of the machine learning experiments are presented, and an analysis of the results is provided.

**Chapter 5:** This chapter concludes the research and addresses how the objectives were achieved. The chapter further highlights the limitations of the study and presents new opportunities for further research in this domain.

## CHAPTER TWO: LITERATURE REVIEW

### 2.1 Introduction

This chapter begins with the definitions, concepts and domains related to iRTLS. The literature review also includes an evaluation of the technologies and methods used for iRTLS, location techniques used in iRTLS, and modelling use cases for experimentation. iRTLS can be used in various domains that have their own constraints. The key attributes that need to be considered in designing and implementing an iRTLS solution are covered in this chapter. Based on one or more use cases/scenarios that require the detection of obstructions within the defined environment, the subsequent main machine learning techniques used for these detections are explored.

Although commercial systems that use various indoor location technologies are available, such systems contain vague details and are surrounded by marketing hype. This lack of technical detail on such proprietary systems limits one's ability to make an informed and detailed assessment. Attempts to contact companies with solutions have been unsuccessful. This study is therefore based on academic literature from recognised journals and conference publications. For example, Guimarães *et al.* (2016) conclude that the search for a robust and accurate indoor positioning system still remains open due to the issues with fingerprinting and the high cost of infrastructure. Bai, Gu and Hu (2016) have also identified a lack of mature technologies for highly accurate indoor location and tracking.

### 2.2 Real-time location systems

The terms *location*, *localisation* and *positioning* are used synonymously to indicate where a person or an object is in an environment. This can be further divided into two broad areas, namely outdoors and indoors. According to the International Standards Organization (ISO) (2007) definition, an RTLS is a

combination of hardware and software that is used to continuously determine and provide the real-time position of assets and resources equipped with devices designed to operate with the system. (ISO 2007)

Indoor location refers to locating people or objects within a building such as a house, mall, or hospital. Although Global Positioning System (GPS) technology makes it easy to detect people

and objects outdoors, it is also important to establish the precise location of objects indoors in real time. Due to infrastructure and equipment interference, indoor GPS monitoring is not possible (Alexander and Kusuma 2019; Bai *et al.* 2020; Mahfouz, Nader and Abi-Char 2020). This is because indoor GPS tracking requires a connection to both the GPS device and the positioning satellites. Additionally, due to severe object occlusion and multipath effects of signal transmission, satellite-based location technologies face numerous challenges (Li *et al.* 2019). Consequently, new technologies and techniques are needed for iRTLS. The Internet of Things (IoT) is one of these new technologies that is increasing in popularity and becoming more cost-effective to implement.

### **2.3 Typical use cases of indoor real-time location systems**

There are a variety of domains where iRTLS can be used. Some of these include coal mining (Song and Qian 2016), navigation systems for blind and visually impaired people, tracking expensive equipment, tracking children in crowded places (Alarifi *et al.* 2016), disaster relief (Lee, Jeon and Jeong 2017), robotic communication (Alsamhi, Ma and Ansari 2019), smart cities (Alsamhi *et al.* 2019) and indoor intelligent vehicle navigation (Alsamhi, Ma and Ansari 2020). In addition, other significant advantages of employing RTLS include managing assets to prevent loss, readily locating equipment (Wamba, Anand and Carter 2013) and determining utilisation of equipment (Perry and Kocakülâh 2010).

The healthcare environment is an important domain where a multitude of applications can be identified. An important one, the tracking of patients, personnel and assets (Khan *et al.* 2021), is important in case of an emergency. Cost-effective and feasible solutions must be designed for such environments. Researchers are therefore studying feasible solutions for patient and asset location within the healthcare context. Well-managed implementations of iRTLS systems in the healthcare industry can provide enormous value for patient and staff location (Potisek *et al.* 2007) as well as improved service delivery (Attarha 2011). Yazici (2014) surveyed the largest public health institution in Florida and discovered that real-time patient tracking was one of the top two goals, followed by the location and tracking of assets like medical equipment. The only difference between these two use cases is the technique of attaching the locating device to the asset or patient. Such an application is also applicable in an office environment where the location of assets is equally important and the location of personnel,

for example, in the case of a fire, is critical. Therefore, the office domain could be used to illustrate the concepts, technologies and constraints of indoor location.

Fisher and Monahan (2012) examined the indoor location systems used by twenty-four hospitals and found, at the time of the survey, that only eleven hospitals had fully working systems. Although this could be attributable to a number of factors, one of the suppliers' claims was that the systems did not provide the promised precision. Future study is thus required to confirm the consistency of detection accuracy claims. Furthermore, the above-mentioned authors asserted that research attempts to improve RTLS have not progressed beyond the pilot phase or laboratory testing (Shirehjini, Yassine and Shirmohammadi 2012).

## **2.4 Constraints of indoor real-time location systems**

As mentioned above, there are a variety of domains where iRTLS can be applied. Each of the domains can have different and sometimes common requirements and constraints within their environment. In the context of the present research and type of domain, these attributes were deemed significant and had to be balanced when determining an indoor location. Given the domain chosen for this study, the power consumption of the device had to be minimal in order to reduce the maintenance and replacement of its power source. Also, in an environment containing equipment sensitive to electromagnetic interference, the transmission level had to be reduced to prevent interference. These devices must be portable and therefore had to have a small form factor. Some of these characteristics may be contradictory, such as power consumption and maximising range, which necessitates greater power consumption. Therefore, it was necessary to achieve an optimal combination of these attributes for a viable indoor location system. A combination of technologies and techniques have been used in other studies to mitigate these constraints and determine accurate location. However, a number of issues are documented in the research that restrict cost-effective, viable, large-scale deployment. The constraints are discussed in the sections that cover the identified technologies as these factors are dependent on those technologies.

The primary concerns of research on this topic are the enhancement of localisation factors like precision, computational complexity, deployment simplicity and cost (Thaljaoui *et al.* 2015). Five main attributes were identified, namely cost and detection range (Tsang *et al.* 2015), energy consumption (Yu, Xu and Li 2012), form factor and accuracy (Deng *et al.* 2013).

Depending on the example, the scalability attribute is also a crucial consideration. Identified criteria were used to compare wireless fidelity (Wi-Fi), radio-frequency identification (RFID), near-field communication, device-free localisation, Bluetooth Classic, Bluetooth, Bluetooth Low Energy (BLE) and hybrid technologies with respect to their status as the four principal technologies. These are examined along with their benefits and drawbacks.

There are numerous limitations in healthcare, with price and transmission power being two of the most prominent (Alemdar and Ersoy 2010). The office environment, which is comparable to that of healthcare, can serve as a model for the RTLS needs. Both environments have additional constraints such as maximising battery life, using wearable devices with a compact form factor, and the venue's size and variable layout.

A total relative rating was adopted because of the difficulties of rating the overall cost, form factor and energy usage. The range and precision of detection measurements depend on the type of tag and technology employed. Consequently, measurements have a broad range and vary based on the infrastructure employed. The selected technologies are the most suitable for RTLS in an office context. The following are important factors to consider for the various technologies: cost, because of the constrained budgets of institutions; size of the tag, since the tag will be worn by individuals; energy consumption, to ensure long battery life and to enable the battery to be as compact as possible; accuracy, to precisely locate a person or object; and detection range, to reduce infrastructure.

Due to the nature of the chosen environment and its limitations, a number of different technologies that can be utilised for localisation, including ultra-wide band (Deng *et al.* 2013), were not considered. Furthermore, for a variety of reasons, including cost, they are not suitable for large-scale iRTLS implementation.

#### **2.4.1 Interference and range**

Healthcare is one of the sectors that need iRTLS. An important constraint in healthcare is the prohibition of interfering with the operation of any medical equipment (Yao, Chu and Li 2012). Technical issues such as metallic object distractions and electromagnetic interference in reading tags, impacting detection range and accuracy, were identified by Yazici (2014) as the primary obstructions to RFID implementation. Despite the great potential of RFID for positioning, its effective range is limited and its signal may be blocked by surrounding objects

(Faruqui and Tayab 2008). Objects in motion will cause further interference, resulting in abrupt changes in received signal strength indicator (RSSI), which, in turn, results in inaccurate locations (Gikas *et al.* 2016).

Wi-Fi signals can be seriously affected by access points in the surrounding environment as well as obstructions such as walls, floors and other devices (Bai, Gu and Hu 2016). In addition, Wi-Fi-based positioning continues to experience inaccuracy due to unreliable signals and multipath interference (Moghtadaiee and Dempster 2012). As the density of wireless devices grows, wireless networks and services are more susceptible to interference from unintentional wireless transmissions from other devices, resulting in packet loss (Behboodi *et al.* 2017).

It was also observed by Behboodi *et al.* (2017) that the higher power interferences causes RSSI levels to change, and when a certain threshold is reached the packet reception is no longer possible. Causes of interference identified by Mainetti, Patrono and Sergi (2014) include errors by multipath and non-line-of-sight conditions, presence of moving people that modify the indoor propagation channel and greater density of obstructions. Mahfouz, Nader and Abi-Char (2020) and Jain, Sashank and Markkandan (2021) also recognised that obstructions cause an interference, loss of signal strength and fluctuation of signal strength between the transmitter and receiver. Atashi *et al.* (2019) studied the impact of a wooden obstruction in the path of BLE signal. They found that the RSSI levels received with the obstruction in the path were lower than those received without the obstruction.

#### **2.4.2 Coverage and scalability**

A number of challenges plague the industry, preventing the wide-scale rollout of RFID solutions. Wi-Fi access points are installed for optimal communication coverage at minimal cost. However, these are not positioned for location-based systems and are not optimised to scan for positioning (Powar, Gao and Harle 2017). In order to achieve a high degree of accuracy for localisation systems using both fingerprinting and radio propagation techniques, the areas to be covered need to be highly populated with access points (Le and Havinga 2017). Such solutions to these problems are prohibitively expensive, preventing their widespread application. Also, scalable systems incur a cost, but suffer from resilience difficulties since they rely on opportunistic signals over which they have no control (Powar, Gao and Harle 2017).

Bluetooth Classic has signal attenuation and interference issues in busy areas. BLE is an upgraded and later version of Bluetooth that offers various advantages, including a smaller form factor, lower cost and greater coverage. Research in iRTLS aims to produce a solution that is robust, accurate and scalable. However, such systems rely on extensive infrastructure installation and support that hinders scalability (Powar, Gao and Harle 2017). Furthermore, in order to use RSSI values, a fingerprint needs to be referenced. However, such a fingerprint will need continuous calibration when the environment changes (Le and Havinga 2017), which can be very tedious if done manually. The alternative is to use additional infrastructure, but the system complexity must be considered for scalability (Dwiyasa and Lim 2016). In addition, there is noticeable variation in RSSI levels between devices of the same models and brands, let alone those of different models and brands (Bai, Gu and Hu 2016), leading to uncertainty about reliable coverage and scalability.

### **2.4.3 Power consumption**

RFID uses passive and active tags for identification, and passive tags have a limited range. Therefore, active tags powered by a battery would be used for location. However, despite the great potential of RFID for positioning, there are limitations to battery life (Faruqui and Tayab 2008). Also, using Wi-Fi on low-powered devices will consume a lot of battery power (Powar, Gao and Harle 2017), and battery life will therefore be limited.

Another technology used for indoor location is Bluetooth. Bluetooth Classic can transfer a large amount of data, but has a shorter battery life and is more expensive than BLE or other indoor localisation technologies (Zaim and Bellafkih 2016). Methods used to increase location accuracy, for instance the Kalman filter, will increase form factor and power consumption due to their increased storage requirements (Tsang *et al.* 2015). Hence, a solution will need to balance the accuracy with power consumption due to the increased form factor. One such technique, used by Barsocchi *et al.* (2017), is to transmit the signals at different levels, resulting in a variation of power consumption at the different levels. The trade-off will be to transmit at the lowest level to determine the acceptable accuracy. A localisation approach using fingerprinting should be self-organising, lightweight and accurate (Le and Havinga 2017). A solution for maximising battery power is to use the lowest levels of signal transmission level (Pancham, Millham and Fong 2018), a finding synonymous with (Barsocchi *et al.* 2017).

#### 2.4.4 Cost

Cost reduction is a significant factor when implementing indoor RTLS (Wirola, Laine and Syrjärinne 2010). RFID is a frequently examined technology in the realm of indoor location. Although passive tags are inexpensive (Deak, Curran and Condell 2012), they require more readers because they must be identified at a close range, hence raising the total cost. In addition, systems require expensive and specialised equipment that must be installed and maintained (Mainetti, Patrono and Sergi 2014). Active RFID tags are more expensive than passive RFID tags, but their accuracy increases to between 1 and 2 m (Deng *et al.* 2013). Analysis by Yazici (2014) suggests that the high cost of infrastructure and tags is one of the obstructions to a widespread implementation of finding and monitoring both patients and assets. This conclusion is consistent with prior research conducted by Reyes (2012), who discovered that cost was a significant obstruction for both non-implementers and future implementers of RTLS, specifically RFID.

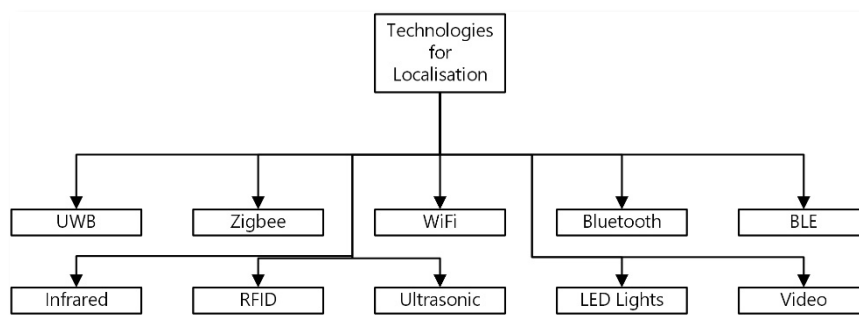
As mentioned previously, RSSI requires continuous calibration for accurate detection (Le and Havinga 2017). To reduce the amount of continuous calibration required entails using additional infrastructure, but such infrastructure comes at an increased cost for large-scale deployment (Dwiyasa and Lim 2016). Improvements in accuracy are possible, though, as demonstrated by Moreira *et al.* (2017) who obtained an improvement in accuracy of 40% by using five Wi-Fi interfaces in a real-world experiment. However, this will add to the infrastructure cost, which will need to be balanced based on the requirements.

Several researchers and practitioners have utilised RTLS for localisation in various industries for a number of years. Although much research has been conducted, a feasible solution that mitigates the constraints identified, has not been determined (Dwiyasa and Lim 2016). Consequently, new technologies and techniques for location and their improvements for iRTLS need to be evaluated. In their research, (Adame *et al.* 2016) proposed investigating further solutions to reduce energy usage and enhance the precision of localisation tracking in the healthcare industry. They also identified challenges such as optimising the size, battery life and memory capabilities of the device for an optimal RTLS.

## 2.5 Indoor real-time location system technologies

Although iRTLS systems have been implemented in certain healthcare facilities in several countries and utilising a variety of technologies, large-scale rollout faces a range of challenges. (Pancham, Millham and Fong 2017) compared the most prominent iRTLS technologies and their characteristics to the constraints highlighted in recent peer-reviewed studies and within a healthcare context. The characteristics of the example are representative of many work environments. Wi-Fi, Bluetooth, RFID, vision, ultrasound and infrared technologies were evaluated. In addition to the example, hospital survey data from twenty-three United States hospitals (Fisher and Monahan 2012) was utilised in the evaluation procedure. ZigBee is a low-power wireless personal area network that uses the 2.4 GHz band. This technology employs low-power, low-cost, compact radios that provide self-organised, multi-hop and reliable mesh networking over a distance of 10 m along with a long battery life (Shukla and Bergmann 2004). Although ZigBee uses low power and is available at low cost, it is not used for indoor location due to unavailability of end user devices (Zafari, Gkelias and Leung 2019).

Near-field communication is another wireless technology used for short-distance data transfer. In fact, the tag and the sensing device need to be in close proximity for communication to take place. The most popular technologies such as Wi-Fi, RFID, Bluetooth and BLE are discussed in more detail below. Several technologies such as ultra-wide band (Deng *et al.* 2013), ultrasonic, infrared, standard television signals, physical contact and computer vision (Deak, Curran and Condell 2012) were not considered for indoor iRTLS because they did not meet the minimum criteria of the essential characteristics such as reasonable cost. Two main technologies used in indoor RTLS are RFID and, much more recently, BLE. These technologies are depicted in figure 2.1.



**Figure 2.1: Technologies for localisation**

### 2.5.1 Wi-Fi

Wireless local area network (WLAN) uses the 2.4 GHz and 5 GHz bands for their transmission. Wi-Fi is a popular technology studied for RTLS (Kjærgaard *et al.* 2013; Mathisen *et al.* 2016); Wang *et al.* (2016) that connects a large number of devices indoors. The advantage of using Wi-Fi is the use of installed infrastructure. The core architecture includes a wireless access point and a Wi-Fi device with a Wi-Fi module for connectivity. This architecture is designed primarily for data and information exchange using carrier-sense multiple access with collision avoidance. The process is to check if the channel is clear to transmit a packet of information. However, if the channel is busy, it waits a random amount of time and checks again. If the channel is busy the system will wait, causing delays, and this will be an issue for real-time location (Wang *et al.* 2013).

According to (Wang *et al.* 2013), “The concept behind Wi-Fi RTLS is that this technology can utilise the existing Wi-Fi infrastructure to communicate with Wi-Fi tags.” Even after increasing the number of Wi-Fi access points required for indoor location systems, however, determining a precise location remains challenging (Paiva, Brito and Leiva-Marcon 2018). The limitations of using RSSI from Wi-Fi networks are the volatility of signal propagation (Mathisen *et al.* 2016) and the uncertain fluctuation of measurements (Li *et al.* 2016). Another issue with using Wi-Fi is that Wi-Fi on mobile devices for location will have a high energy consumption (Neishaboori and Harras 2013). To use Wi-Fi for RTLS, multiple Wi-Fi access points must be deployed for adequate detection in order to achieve real-time tracking of assets or persons.

Using a single wireless access point, time difference of arrival, and RSSI, (Wang *et al.* 2013) determined the position of Wi-Fi tags. The detection accuracy is approximately 3 m (Shirehjini, Yassine and Shirmohammadi 2012; Deng *et al.* 2013). However, this method adds to the complexity of the hardware, as all devices must be accurately synchronised, thereby increasing the cost. The value of devices that connect to the wireless access point will vary depending on the capabilities offered. However, the battery life will be extremely short with added features. Examples of solutions using Wi-Fi are Ekahau, Microsoft Research RADAR, Aero Scout, Intel Place Lab and Pinpoint 3D (Deak, Curran and Condell 2012).

Wi-Fi location accuracy can be specified at the zone, room or sub-room level. Even in some instances, Wi-Fi tags might be improperly recognised in a room, and moving assets cannot detect signals due to barriers between the tag and access points (Wang *et al.* 2013). Wi-Fi

networks are not designed for RTLS and, due to the increased traffic, the network performance will therefore be compromised. In order to build a bespoke solution, the optimal compromise between performance characteristics, user needs and environmental circumstances must be established (He and Chan 2016). Wi-Fi scanning and algorithmic support should be based on accuracy requirements, user experience and battery constraints (Neishaboori and Harras 2013).

### **2.5.2 Radio-frequency identification**

RFID was one of the earliest technologies to be utilised in the production of RTLSs. It is a technology that identifies and tracks objects using electromagnetic fields. An RFID system typically consists of three components: an RFID tag, a scanning antenna and a transceiver. The antenna scans for tags, and when a tag senses an activation signal, it transmits the data (Paiva, Brito and Leiva-Marcon 2018). There are three classifications for RFID tags: passive, semi-passive and active. Both active and semi-passive RFID tags use batteries to power their circuits. The active tag transmits radio waves to a reader using its battery, whereas a semi-passive tag relies on an external power source, such as the reader, to transmit radio waves. An antenna coil plus a silicon chip with basic modulation circuitry and non-volatile memory constitute a passive tag. These tags are entirely powered by the reader, which subsequently communicates the tag's unique identifier (Wang *et al.* 2013). Tag collision (in which readers simultaneously read multiple tags and are unable to determine their individual identities) and reader collision (in which multiple readers simultaneously read the same tag, resulting in duplicate locations) are two additional technical issues with RFID that must be resolved (Paiva, Brito and Leiva-Marcon 2018). These technical difficulties significantly impair detection accuracy. Anti-collision algorithms that address some of these issues increases the system's complexity and at the same time decrease the efficiency of the location system (Su *et al.* 2018).

Additional concerns that have been raised include privacy and signal interference (Yao, Chu and Li 2012). Wamba, Anand and Carter (2013) conducted a comprehensive analysis of 215 academic articles pertaining to RFID applications. According to that study's conclusions, the total benefit of implementing RFID technology will depend on numerous aspects, including cost. RFID does not have sufficient bandwidth for large-scale implementation of indoor locating systems, despite the low cost of RFID tags (Terán *et al.* 2017). Notwithstanding the substantial academic research conducted over the years of RFID's existence to aid in the design and improvement of RFID systems, there are still challenges that need to be overcome.

Another issue, identified by Zhu, Mukhopadhyay and Kurata (2012), is that many models that are based on academic research are not implementable and so are of no use to practitioners. Even when implemented, the industry faces a variety of challenges that preclude a widespread adoption. According to (Yazici 2014), the primary challenges to RFID implementation include technical issues such as interference from metallic objects and electromagnetic interference when reading tags, which affects detection range and accuracy, as well as the high cost of infrastructure and tags. The high cost of resolving these challenges makes large-scale implementation impractical.

Yazici (2014) notes that the cost of RTLS for patient and asset monitoring is a major impediment. This is consistent with prior research by Reyes (2012), which noted that cost was a significant obstruction for both non-implementers and future implementers of RTLS, specifically RFID. A further shortcoming of RFID, as identified by Gulcharan *et al.* (2013), is its inability to obtain an effective and precise data transfer to support continuous data transmission. Lim, Bahr and Leung (2013) reviewed literature from 1991 to 2005 of RFID implementations in warehousing. The study identified several limitations in these implementations, including uncertain return on investment and poor performance, privacy and security. This is consistent with the findings of (Yazici 2014).

Although passive tags are relatively inexpensive (Deak, Curran and Condell 2012), they necessitate more readers because they must be identified at close range, hence raising the total cost. Active RFID tags are more expensive than passive RFID tags, but their accuracy increases to between 1 and 2 m (Deng *et al.* 2013). Therefore, the precision of the detection as well as the detection range will depend on the type of tag employed. Some batteries have a lifespan of a few months to a year, resulting in moderate energy usage. Once the battery and holder are attached, the device becomes relatively large. RFID is not suitable for indoor location since it cannot pinpoint the exact location of a tag; it can only indicate the location of an object within a defined area. Moreover, there are significant expenses associated with implementing an RFID system because a large number of antennae are required (Paiva, Brito and Leiva-Marcon 2018).

Hence, because of these disadvantages, RFID tags do not meet the constraints of cost, low battery consumption and form factor. Passive tags require very close proximity to the readers for detection. Active tags, in contrast, can be detected from a distance but would also require a

power source. As previously observed from research, because this technology requires a direct link between a tag and a reader, any obstruction in this path will affect the signal and hence the location. This technology is limited to the identification of the tag and therefore cannot be extended for data transmission. Although RFID has good coverage, it cannot be used for location due to its poor accuracy. Passive RFID, meanwhile, has no battery – resulting in a small form factor – but has a limited range. Therefore, RFID is not a recommended technology for indoor location.

### **2.5.3 Near-field communication**

A more recent technology based on near-field communication is a passive one-way identifying technique for very short distances that does not require batteries for its tags because power is induced electromagnetically when a near-field communication reader is within range. This method is only effective for data transfer over small distances, perhaps a few centimetres (Cheng *et al.* 2016). Consequently, it cannot be utilised in an RTLS.

### **2.5.4 Device-free localisation**

Device-free location does not utilise a dedicated location-determining device. This method utilises existing infrastructure to pinpoint the location of an object or person. Indoor device-free localisation uses various electronic means such as IoT and cameras to recognise human behaviour (Lee, Park and Kim 2022). However, one of the main disadvantages of using cameras is the invasion of privacy. Another method is to use the changes in the radio map generated from the channel state information of the Wi-Fi signals (Gao *et al.* 2017). Using the data from the radio maps, this technology applies deep learning to detect the location of a human. However, according to Ding *et al.* (2022), whilst this technology can detect the location of a human and the related movement, it cannot be used as it does not give the unique identification information required. Furthermore, wireless signals that are used to determine location are also affected by variation in the environment.

Changes in the patterns of received RSSI levels can also be used to determine the location or movement of an object, such as a human, by analysing the patterns of received RSSI levels. (Abdull Sukor *et al.* 2020) used a deep multi-layer perceptron neural network to train and classify RSSI sensor data from twelve Wi-Fi receivers. Their results showed that the deep multi-layer perceptron had a higher accuracy of 84% compared to other models like a support

vector machine (SVM) at 64% and a traditional multi-layer perceptron at only 55%. However, the receivers in this case were only 1 m apart. This research concluded that more research would be needed with receivers and sensors placed further away from each other. Device-free localisation is useful, for example when searching for humans in the case of a disaster. However, for the purposes of indoor location described in the present research, this method will not apply due to the privacy issues of using cameras, the complexity of the system and the inability to uniquely identify entities.

### **2.5.5 Bluetooth Classic**

Bluetooth is a wireless technology that operates in the 2.4 GHz unlicensed industrial, scientific and medical (ISM) band (Woolley 2019). Basic rate is one of the forms of Bluetooth, which includes an optional enhanced data rate and alternate media access control and physical extensions. These protocols are typically referred to as Bluetooth Classic and are used primarily to connect a plethora of electronic devices such as headsets, cell phones, etc., and for audio streaming at close range. There are both synchronous and asynchronous connections with data rates of 721.2 Kb/s for basic rate and 2.1 Mb/s for enhanced data rate (Core Specification Working Group 2021). Frequency hopping and time division after clock synchronisation of devices are used to share the radio channel or data transmission and avoid interference. Bluetooth Classic uses 79 channels with 1 MHz spacing from 2 402 MHz to 2 480 MHz, and the channels are numbered 0 to 78.

Bluetooth Classic uses pairing between devices, and once paired they become “undiscoverable”, hence raising questions regarding the early tracking approaches used in previous research (Cominelli *et al.* 2020). (Cominelli *et al.* 2020) further conclude that Bluetooth Classic is inadequate for ensuring user identity and location privacy and hence called for revising the Bluetooth specification. Furthermore, Bluetooth Classic can transfer large quantities of data, but consumes battery life quickly and is more costly than BLE or other indoor localisation technologies (Zaim and Bellafkih 2016).

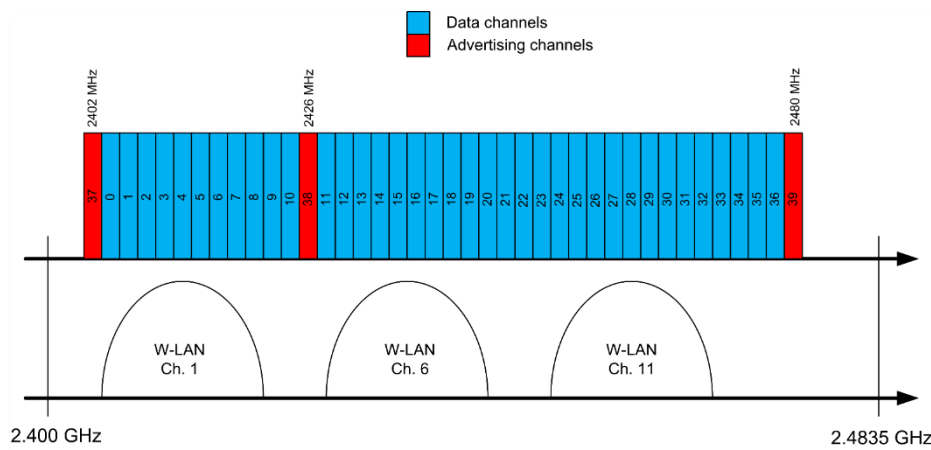
### **2.5.6 Bluetooth Low Energy**

The second Bluetooth standard is BLE, which incorporates characteristics aimed at enabling products with lower current consumption, reduced complexity and lower cost than basic

rate/enhanced data rate. This Bluetooth variant is targeted at applications and use cases with lower data rates and duty cycles (Core Specification Working Group 2021).

BLE utilises forty channels separated by 2 MHz in the unlicensed 2.4 GHz ISM frequency band. BLE operates between 2 402 MHz (RF channel 0) and 2 480 MHz (RF channel 39). The centre frequencies of the channels are defined as  $f_k = (2402 + k * 2)$  MHz, with  $k \in [0; 39]$ . Three channels (channels 37, 38 and 39) are referred to as advertising channels, while channels 0 to 36 are data channels (Pang *et al.* 2021). The advertising channels are positioned so as to be unaffected by the non-overlapping WLAN channels 1, 6 and 11 in the ISM band (see figure 2.2).

Compared to Bluetooth Classic’s 1 MHz channels, BLE’s greater channel width of 2 402 MHz enables it to provide faster transmission speeds. In addition, these channels may send more data due to the higher transmission frequency.



**Figure 2.2: BLE channels**

BLE has various advantages over Bluetooth Classic, including a smaller form factor, a lower price and an increased range. Typically, a BLE system comprises a stationary anchor for detecting tags, a tag and a location engine for calculating location (Han *et al.* 2015). BLE is the IoT-specific power version of Bluetooth Classic, making it ideal for devices that run for extended durations on power sources such as coin cell batteries or energy-harvesting devices. Also, BLE transmits small packets of data whilst consuming significantly less power than the previous version of Bluetooth Classic (Collotta *et al.* 2018). Because it consumes less energy, the battery life will be extended (Taşkan and Alemdar 2021).

Bluetooth Local Infotainment Point is an example of a Bluetooth system (Kolodziej and Hjelm 2017). It is a controlled network that provides access to local area network/wide area network over Bluetooth (Deak, Curran and Condell 2012). Due to its restricted range, such a network will require a number of Bluetooth Local Infotainment Point nodes to which Bluetooth devices can connect. These Bluetooth nodes then grant access to the local area network/wide area network.

Lee, Dong and Sun (2015) compared BLE and ZigBee technologies using a single fixed distance of 1 m, but they were unable to determine conclusively which technology is superior due to the fact that wireless transmission is greatly affected by real-world situations, such as interferences from the actual environment. In addition, this experiment did not offer measurements of features such as RSSI or throughput beyond the set distance, which are required for a proper evaluation of network technology at the fixed distance. However, Collotta *et al.* (2018) noted in their research that BLE offers a higher range and throughput than Zigbee.

BLE combined with beacon technology can cover up to 50 m (Cheng *et al.* 2016). Raza *et al.* (2016) determined that the present BLE nodes' point-to-point communication has limited range and coverage over a short distance. One of the proposed solutions is a wireless mesh multi-hop network with many nodes capable of inter-node communication to enable packet routing in order to extend this limited coverage. This distance can be increased further with the combination of existing technologies and more efficient algorithms.

The goal of work by (Han *et al.* 2015) was to evaluate if people with BLE tags are reliably tracked and detected. The system comprises a tag that transmits 2.4 GHz signals, a fixed anchor that receives signals and a local engine that collects data and performs calculations. The benefit of employing BLE tags is that they consume little energy and are inexpensive. Nevertheless, Han *et al.*'s results demonstrated that detection was inaccurate for boundary circumstances, particularly in permeable walls, whereas detection was more reliable when tags were closer to anchors. Therefore, they determined that additional effort was necessary to increase accuracy. Taşkan and Alemdar (2021) also observed a significant loss of signal through walls. In their research, (Adame *et al.* 2016) proposed that industries, such as the healthcare industry, should investigate alternate methods with lower power consumption and increased accuracy of location tracking. Additionally, they recognised improving the size and battery life as impediments to an ideal RTLS.

In their research, Raza *et al.* (2016) determined that the point-to-point communication of the present BLE nodes had limited range and coverage. As a potential solution, they proposed employing a wireless mesh multi-hop network with numerous nodes capable of inter-node communication to enable packet routing and extend this limited coverage. This distance can be increased with the integration of current, more efficient technology.

BLE is one of the most common wireless technologies consuming low energy for short range communications (Zafari, Gkelias and Leung 2019). Because of these advantages BLE has become the de facto standard for IoT entities and is currently integrated into a multitude of wearable technology (Bertuletti *et al.* 2016). This technology is well designed for location determination using a wide range of techniques such as angle of arrival, received signal strength and time of flight among others (Zafari, Gkelias and Leung 2019).

### **2.5.7 Hybrid technologies**

A hybrid method uses multiple technologies to determine location. Because of the limitations of Wi-Fi, Baniukevic, Jensen and Lu (2013) have proposed a hybrid method to determine location, comprising of Wi-Fi and Bluetooth. Kao *et al.* (2017), in their research, also used Wi-Fi to locate mobile devices using fingerprinting, and those devices were in turn used to locate beacons using trilateration of BLE signals. Trilateration employs just distances and coordinates derived from the geometry of circles to determine the location of an unknown point (Paiva, Brito and Leiva-Marcon 2018). This hybrid method uses the cell phones of users of hospital equipment to perform the location. This experiment achieved a 93% accuracy within 3.16 m for Wi-Fi and 90% accuracy within 1.21 m for BLE. It should also be noted that battery consumption will be fairly high, especially when using both Wi-Fi and Bluetooth. Bai *et al.* (2020), in their research in old-age homes, noted that the battery of a sensor is critical in indoor location, with battery life needing to last in excess of a month. Several methods, such as those applied in old age homes (Bai *et al.* 2020), sometimes use personal devices for location, which is not practical. Users are likely to be reluctant to use their own devices for such an application for location. Also, these devices are onerous to carry and have applications that the elderly would consider complex to use. A fingerprint database will become outdated as a result of the changing environment, resulting in inaccurate locations.

According to Shang and Wang (2022), who conducted a survey of publications on wireless networks between 2017 and 2020, the amount of research on Bluetooth networks was greater

than that on Wi-Fi, ultra-wide band, RFID and Zigbee. They also highlighted that the advantages of Bluetooth networks are low power consumption and ease of deployment. Also, Basiri *et al.* (2017) found Bluetooth to have a better accuracy range, better coverage in indoor environments and lower cost than Wi-Fi, ultra-wide band and RFID. Although wireless transmission is a better option than the other technologies, these signals are affected by attenuation, fading, shadowing, human movement and presence of interfering devices and obstructions (Roy and Chowdhury 2021).

## 2.6 Evaluation of technologies

RFID properties are dependent on the tag type, either active or passive, and the associated hardware (Tsang *et al.* 2015). RFID has a number of disadvantages, such as limited bandwidth of RFID for large-scale implementation, interference, high cost, poor performance, very short range of passive tags and inability to extend this technology for data transmission. These have been identified in section 2.5.2. The price of Bluetooth Classic and BLE modules is inexpensive, whereas the price of Wi-Fi modules is moderate to high (Deak, Curran and Condell 2012). Bluetooth has a detection range of between 50 and 100 m. Wi-Fi accuracy ranges from 10 to 20 m (Deng *et al.* 2013), whereas Bluetooth accuracy ranges from 10 cm to 10 m (Tsang *et al.* 2015). BLE, on the other hand, is accurate to within 3 cm to 5 m (Faragher 2014). Scalability is generally advantageous for all technologies (Deak, Curran and Condell 2012). Table 2.1 uses the following legend due to space constraints:

EC: Energy consumption
DR: Detection range
Cost and energy consumption constraints: L – Low, M – Medium, H – High
Size: S – Small, M – Medium, L – Large

**Table 2.1: Technology attributes**

Tech	Cost	EC	DR	Size	Accuracy	Scalability
<b>RFID</b>	H	M	Depends on tag type	L	Depends on tag type	Good with minimum of 2 tags
<b>BT</b>	M	M	100 m for class 1	S	10 cm to 10 m	Good
<b>BLE</b>	L	L	High (50 m)	S	3 cm to 5 m	Good
<b>Wi-Fi</b>	M	H	up to 200 m	L	10 m to 20 m	Good

The ground-breaking update to Bluetooth Classic, Bluetooth 5, was released in December 2016, and significantly improved the Bluetooth applications' range, speed and broadcast messaging capabilities. This update quadrupled low energy connection range and doubled connection speed and throughput (Schultz 2016; Collotta *et al.* 2018).

BLE was found to be the most promising for use in indoor location when compared to other technologies (Pancham, Millham and Fong 2017). Currently, indoor positioning is one of the most important BLE applications. For instance, 75% of the leading American stores have implemented BLE beacons for various uses (Nikoukar *et al.* 2018). Low-cost and simple-to-deploy technologies, such as BLE, for indoor location can also be used to monitor the location of people, such as the elderly in a nursing home (Thaljaoui *et al.* 2015). BLE is used to mitigate the high energy consumption demanded by Wi-Fi networks (Terán, Carrillo and Parra 2018). Moreover, BLE-based infrastructure provides several advantages over Wi-Fi, such as low power consumption and higher sample rate (Roy and Chowdhury 2021).

Various iRTLS technologies have been discussed above, indicating the wide applicability, success and issues of these technologies for indoor location determination. These technologies have been used in the domain of indoor location determination of objects and people. The most appropriate technology that would facilitate a balance between the identified attributes was selected for the present study.

## **2.7 Design models**

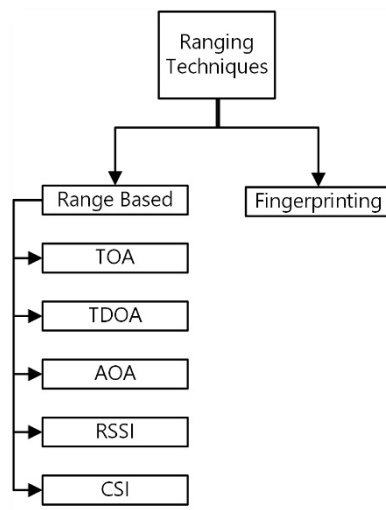
Formal techniques and Unified Modelling Language (UML) are among the languages/models that can be used to describe requirements, analysis and design of artifacts. Literature highlights the reasons formal methods were not explored, including the challenge of comprehending formal methods and the suitability only for large projects. UML is a standard, object-oriented, visual language for modelling software-intensive systems (Milicev 2009). Furthermore, UML provides many perspectives of the system for diverse stakeholder groups (Booch, Rumbaugh and Jacobson 1999). UML includes capabilities for all phases of the system development life cycle. Consequently, models from one process can be readily moved to the following workflow. For instance, use cases developed during requirements elicitation can be converted to activity diagrams and then to system sequence diagrams used in detailed design. The end

users can easily comprehend the scenarios and activity diagrams, whereas the developers will use the use cases to proceed with the design and implementation.

## 2.8 Location determination techniques

Several algorithms are used to determine the location of a node in an indoor environment. Methods such as time difference of arrival, time of arrival and angle of arrival can be classified as deterministic algorithms. Due to the multipath effect of an indoor environment, deterministic methods are difficult to use (Sun *et al.* 2020). RSSI, trilateration (D'Aloia *et al.* 2016) or fingerprinting are some of the methods used to determine positioning. The RSSI technique is the most popular method in distance estimation for wireless systems, with an average accuracy error of 1–2 m (Bal *et al.* 2010). RSSI is also widely used for indoor location, mainly through use of a fingerprint database.

Probabilistic algorithms use RSSI to determine locality. However, due to the varying nature of the received RSSI level, preprocessing is required. One of the methods that can be used is to average the levels received so that it can be used in further analysis by different algorithms. Several techniques are used to improve the accuracy of systems (Mainetti, Patrono and Sergi 2014). Indoor localisation techniques are summarised in figure 2.3.



**Figure 2.3: Indoor localisation techniques**

There are a number of methods for increasing accuracy, with the most prominent being the RSSI technique, which improves accuracy to 1–2 m (Bal *et al.* 2010).

### **2.8.1 Angle of arrival, time of arrival or time of flight**

Angle of arrival and time of arrival or time of flight can be used for location determination (Sadowski and Spachos 2018). The angle of arrival technique uses the phase difference of the received signals at the antennas to calculate the location of the target. Angle of arrival can provide accurate location estimation when the distance between the transmitter and receiver is small. The accuracy decreases when the distance between the transmitter and receiver increases because a small error in the angle of arrival calculation results in a magnified error in the location determination. Therefore, angle of arrival is not suitable for indoor locations, considering the constraints identified (Kumar *et al.* 2014).

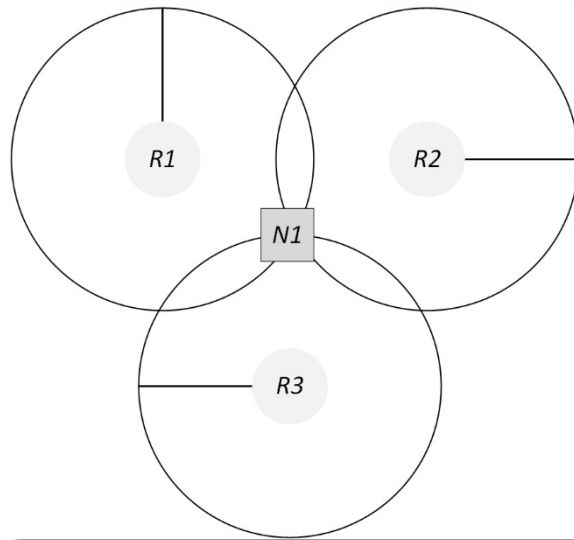
With time of arrival/time of flight, the propagation time of a signal between the transmitter and receiver is measured and used to calculate the distance between the transmitter and receiver (Nessa *et al.* 2020). This distance is then used for trilateration of the receiver's position. In time of arrival/time of flight, all devices must be time synchronised. This technology also requires only one antenna per device and requires very high clock accuracy to obtain reasonable positioning accuracies. Signal bandwidth and sampling rate are key factors that contribute to location accuracy. Although improvements can be made by improving these factors, localisation errors cannot be eliminated when obstructions are in the path of transmission (Zafari, Gkelias and Leung 2019).

### **2.8.2 Trilateration and Kalman filter**

The primary function of signal processing filters is to reduce wireless signal noise and distortion. Invented by the Hungarian mathematician Rudolph E. Kalman in 1960 (Kalman 1960), the Kalman filter is characterised by a sequence of recursive mathematical expressions. The formulas give an effective and computable strategy for estimating the state of the process and minimising the estimated mean square error.

Trilateration is a method used to determine the location of a node using the three ranges or distances between a node and three points of reception. In the case of an indoor location, the point of reception is the receiver that receives the RSSI signal. The attributes of the signal are processed to determine the location depending on the technique chosen, such as the level of the RSSI signal or angle of arrival of the signal, among others. Figure 2.4 below is a simplified representation of the layout of the transmitter and the receivers. The three circles represent the

reception areas of the three receivers  $R1$ ,  $R2$  and  $R3$ , and the point of intersection of the three circles represents the node  $N1$  that needs to be located. This representation is an ideal situation where the three circles intersect at a single point. In the real world, circles can intersect over an area, or one or two circles might not intersect.



**Figure 2.4: Trilateration**

Cantón Paterna *et al.* (2017) used Kalman filtering on RSSI measurements to improve accuracy of locations and then used weighted trilateration as an improvement on the basic trilateration. The weighted triangulation method requires more than three receivers and gives more weight to the receivers closest to the node being located. This combined method yielded an improvement of 43.47% and 38.33% in a medium-sized room and a large-sized room respectively.

Lee *et al.* (2019) proposed an improved filtering algorithm, Kalman-LULU, to reduce the variation in RSSI levels for subsequent trilateration. The overall performance of their newly proposed filter achieved 22.42% error when compared with the effective range, whilst reducing error by 30% when compared with non-filtered results. Huang, He and Du (2019) used trilateration, dead reckoning and a hybrid of the two methods for location determination. The root mean square errors (RMSEs) of the three methods reported were 2.33 m, 0.82 m and 0.76 m respectively, indicating that the hybrid method was the best. Consequently, as demonstrated by the studies above, the Kalman filter and then trilateration can be used to

improve the accuracy of location. However, this adds to the complexity to the location methodology.

### 2.8.3 Fingerprint

Another method used is the fingerprinting-based method to locate the BLE beacon in a defined grid and then to limit these areas to a location of interest (Bai *et al.* 2020). It is used to determine the position of sensors on people around key areas/stay points and obtain a good level of accuracy (> 90%). However, the process of creating a fingerprint is a time-consuming and tedious task, and it is even more difficult to keep such a database updated (Roy and Chowdhury 2021). This causes an increase in the discrepancy between the observed fingerprint and the test fingerprints, which may result in lower accuracy. BLE fingerprinting promises high accuracy provided that the training data is densely populated. However, this process is time-consuming and does not adapt to environmental changes (Naghdi and O’Keefe 2020). Fingerprinting was used by Taşkan and Alemdar (2021) in their multilateration algorithm to identify the best matching positions in the given RSSIs. They noticed that fingerprinting takes into consideration the environmental factors, thereby helping improve location determination. Furthermore, the fingerprints database will need to be maintained and updated in real time when the environment changes, for instance when obstructions are placed in the path of transmission.

### 2.8.4 Received signal strength indicator

The received signal strength is the actual signal power strength received at the receiver, measured in decibel-milliwatts (dBm) or milliwatts. RSSI, meanwhile, is the received signal strength indicator, a relative measurement of the received signal strength as defined by chipset manufacturers (Zafari, Gkelias and Leung 2019). The distance  $d$  between the transmitter and the receiver can be estimated from the following equation (1):

$$\text{RSSI} = 10n\log_{10}(d) + A, \quad (1)$$

where  $n$  is the path loss exponent (which varies from 2 for free space to 4 for indoor environments) and  $A$  is the RSSI value at the receiver. The presence of the RSSI in the vast majority of commercial off-the-shelf radio transceivers has led to the development of many RSSI-based ranging algorithms that may be utilised without additional hardware (Thaljaoui *et al.* 2015). The proliferation of IoT devices that make use of RSSI signalling makes this

technology the most promising with respect to cost and simplicity, since no additional hardware is needed (Mahfouz, Nader and Abi-Char 2020).

One major drawback of using an RSSI fingerprint database is the overhead involved in maintenance of the database (Yoo 2019). RSSI requires a fingerprint with continuous calibration, which can be tedious to update and could also provide incorrect data between calibrations (Le and Havinga 2017). An available improvement of RSSI involves a Kalman filter, which increases Bluetooth accuracy to 0.47 m (Tsang *et al.* 2015).

Another method is the inertial measurement unit, which uses additional hardware on the node being located. These include an accelerometer, a step sensor and a gyroscope. Measurements from these instruments are used the direction and location of pedestrian dead reckoning. This method introduces errors as the algorithm would need calibration at regular intervals (Wu *et al.* 2016). Therefore, this method is infeasible in many environments due to its high level of maintenance.

The Kalman filter is a powerful method used to improve accuracy, but requires a large number of RSSI measurements, which is therefore a limitation of this algorithm in terms of processing and storage capacity (Tsang *et al.* 2015). Tsang *et al.*'s simple prototype with a rapid and dynamic approach without detail calibrations delivered a fixed-point positioning error of 0.47 m.

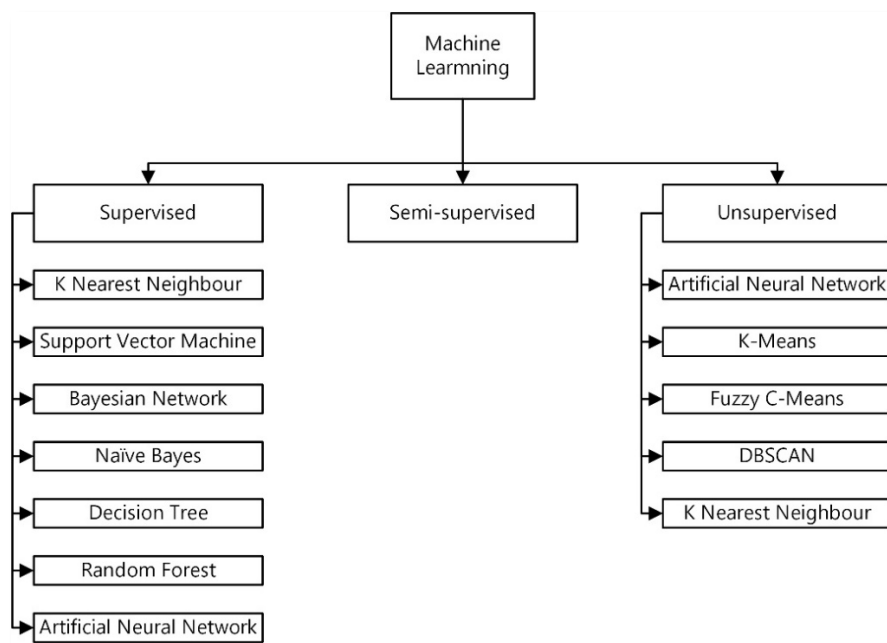
The RSSI-based method of determining location is simple and cost effective, but the accuracy is affected by signal attenuation resulting from obstructions such as walls and other large obstructions, received signal strength fluctuation, indoor noise and multipath fading (Xiao *et al.* 2013; Yang, Zhou and Liu 2013). Accuracy can be improved by using filters and complex algorithms, making RSSI a feasible technique for location determination.

Pancham, Millham and Fong (2020) and Cottone *et al.* (2016) used a link quality indicator to form a footprint of RSSI levels between sensors and readers at known distances. This information was then used to calculate the location of a sensor by using the RSSI levels of the sensor and the footprint previously established. The advantage of this method is that the RSSI levels are known in real time and therefore do not need to be stored in a database.

## 2.9 Machine learning

Machine learning can be classified into supervised learning, where a machine learns from a huge amount of training data, and unsupervised learning, where no training is required (Reddy, Viswanath and Reddy 2018). Classification is a supervised learning technique where data is classified into previously learned categories, while clustering is an unsupervised learning technique in which data is partitioned into groups with comparable qualities (Erl, Khattak and Buhler 2016). Although unsupervised learning can cluster data points, it is extremely difficult to uncover underlying structures in unlabelled data (Dike *et al.* 2018). Semi-supervised learning attempts to use minimal labelled data together with the majority of unlabelled data for optimising machine learning (Roy and Chowdhury 2021). Consequently, as the present research domain involves the determination of obstruction location, using machine learning with specific coordinates that form labels along with their nodes is a preferred choice over unsupervised and semi-supervised learning.

A properly labelled fingerprint dataset requires intensive data collection efforts (Roy and Chowdhury 2021). However, the degeneration in performance limits use and large-scale deployment (Li and Liang 2019). Nessa *et al.* (2020) identified the most popular machine learning algorithms widely used for location determination as SVM, k-nearest neighbour (KNN) and artificial neural networks (ANNs). Other algorithms include Naïve Bayes and decision tree (Sabanci *et al.* 2018).



**Figure 2.5: Commonly used techniques in machine learning**

### 2.9.1 Artificial neural network

An ANN model comprises three layers, namely an input layer that is responsible for the input data, a hidden layer to process that data and an output layer to produce the predictive results (Alexander and Kusuma 2019). An ANN uses backpropagation to iteratively correct its predictive errors (Nessa *et al.* 2020).

An ANN has the advantages of being capable of managing a large number of features, categorising features efficiently and identifying and eliminating redundant features. The training duration is relatively lengthy, and excessive noise has a negative effect on the results. (Hwang and Chen 2017).

(Iqbal *et al.* 2018) researched the tracking of patients and staff in a clinical environment. They used a unique combination of a convolutional neural network and an ANN to improve accuracy in location. Their results indicated that the combined use of the machine learning methods outperformed the convolutional neural network, a thresholding model and a triangulation classifier mode. Although the combination of methods produces improved results, the process of determining the location of obstructions in a real time and in an indoor environment is time consuming.

During the transmission of the radio signal, the signal experiences loss as determined by the medium. (Giuliano *et al.* 2020) performed indoor localisation taking into consideration obstructions due to humans, but used no architectural obstructions. They used a neural network approach to convert RSSI levels to distances and reported error values around 1 m to 6 m, depending on the scenario. In some cases, the signal level increased due to reflection of the signal. Such changes to the signal strength depended on the type of obstruction and the position of the obstruction between the transmitter and receiver.

ANNs are utilised for indoor location systems due to their robustness with regard to noise and interference, which is one of the most significant elements affecting the accuracy of indoor location systems (Nessa *et al.* 2020). Although this domain will always have some outliers in RSSI levels, which an ANN is able to manage and still produce precise results, there will not be a large number of outliers. Consequently, an ANN is suitable to be used for indoor location.

### **2.9.2 Bayesian network**

A Bayesian network is a probabilistic model based on the Bayes theorem. The network is a directed acyclic graph in which nodes represent variables and edges represent conditional relationships between variables. The Bayesian network approach is simple to apply and provides a training model with less training data and reduces the cost of data collection in an indoor setting. A Bayesian network assumes that all the features are mutually independent, but in real-world scenarios this is very difficult to attain (Roy and Chowdhury 2021). Furthermore, a Bayesian network is not accurate as it depends on probability.

### **2.9.3 Naïve Bayes**

The Naïve Bayes algorithm is a probabilistic classifier similar to the Bayesian network algorithm. The expectation is that each feature is independent and therefore makes an equal contribution to the end result. Similar to the Bayesian network algorithm, this algorithm requires a small number of training samples and is also easy to implement. However, Naïve Bayes assumes that the features are independent and contribute equal, independent weighting to the end result (Roy and Chowdhury 2021). For indoor location the RSSI levels captured depend on related layout and environmental conditions. Consequently, these conditions are dependent features. Therefore, Naïve Bayes is not suitable for the present research due to its presumption of independence of features.

### **2.9.4 Decision tree**

The decision tree algorithm uses a tree structure for the classification model. Each node in the tree signifies a decision for an attribute to be made. Each path from the node connects to another node or a leaf. The leaf node is used to predict the outcomes or a class label (Uddin *et al.* 2019). Although this technique is easy to understand and implement, the algorithm generally results in overfitting, which leads to a major disadvantage (Roy and Chowdhury 2021): each path to a decision is not based on evenly weighting the different paths. Therefore, the split may not be balanced and optimal, and this can be influenced by noise. This technique works well with training data. However, if there is noise due to external factors in the indoor environment, this process can result in bad tree construction, as the process does not take noise into account (Xia 2019). Majority classes are overrepresented whilst minority classes are underrepresented. Complex decision trees are difficult to comprehend, which has a significant impact on overall

accuracy (Maimon and Rokach 2005). Therefore, the decision tree algorithm is not suitable for indoor real-time location.

### **2.9.5 Random forest**

This algorithm uses a large number of decision trees to form a learning model. The final prediction result is determined by the most common class predicted by each of the trees. Each decision tree is used to determine a class, and the model's final prediction is based on the most frequent class. The easy-to-implement random forest model works by training itself and thereafter dynamically adapting through continuous training (Gomes, Ahsan and Denton 2018). Due to the large positioning area and fingerprints of random forests, this algorithm requires large amounts of resources, such as memory and processing power, to execute the algorithms (Shang and Wang 2022). Hence, as it is time consuming to obtain a result with this method due to the number of permutations, random forest is not a suitable algorithm for indoor real-time location.

### **2.9.6 k-nearest neighbour**

The KNN model creates groupings depending on the proximity of the data points. The prediction of data points is achieved by comparing the proximity of new data points to training data and identifying  $k$  neighbours with the highest degree of proximity to the new data (Nessa *et al.* 2020).

(Terán *et al.* 2017) used two different algorithms, a KNN classifier using Euclidean distance metrics and a  $k$ -means clustering classifier, to estimate the zone position of a BLE tag using RSSI. The  $k$ -means algorithm yielded better results than KNN. The  $k$ -means clustering algorithm can provide good localisation performance, but it cannot remove noisy points (Shang and Wang 2022). (Sabanci *et al.* 2018) analysed the Wi-Fi signal strength in four distinct rooms with numerous machine learning methods. ANNs, KNN, decision trees, Naïve Bayes, extreme learning machines (ELM), and support vector machines (SVMs) were among these methods. Their findings demonstrated that KNN performed the best.

KNN has the advantages of making no assumptions regarding the distribution of the data collection and allowing multidimensional data (Hwang and Chen 2017). Some of its disadvantages include slow classification speed, the requirement that all training sets be

retained in memory and sensitivity to noise. KNN does not perform well with huge and high-dimensional datasets (Nessa *et al.* 2020).

The characteristics of the present research domain include that no prior information is available on the data collected, and a low-dimensional and relatively small dataset, all of which fit well with the advantages of KNN. Consequently, KNN was deemed one of the most suitable machine learning methods to be evaluated for location determination.

### **2.9.7 Support vector machine**

The foundations of SVMs have been established and have gained popularity as a result of their many promising characteristics, including improved empirical performance. SVM is one of the supervised machine learning algorithms. (Cortes and Vapnik 1995) initially defined SVM to resolve discrimination issues such as classification and regression analysis. This classification is based solely on the observation of a sample of the form  $[(x_i; y_i), i: 1, \dots, n]$ , where  $x_i$  is the  $i^{\text{th}}$  input and  $y_i$  is the corresponding output. The objective of SVM is to find a hyperplane or a set of planes in the N-dimension (where N is the number of features) to distinctly classify the sample points (Roy and Chowdhury 2021). The most prominent kernels in SVM use for indoor location determination include the polynomial, Gaussian, radial basis function and sigmoid kernels (Salamah *et al.* 2019).

Chriki, Touati and Snoussi (2017) found that it was difficult to use only RSSI to perform location determination due to the unpredictable and dynamic nature of signals. Therefore, they used SVM with polynomial and Gaussian kernels with RSSI to perform localisation. However, this was based on a zone, such as a room or corridor, of the target rather than the exact geographical location. This degree of accuracy is acceptable in some scenarios. They also compared the results obtained from ANN and found that SVM yielded better results for localisations. SVM has also been used by other researchers (YuFeng *et al.* 2014); (Wu, Liu and Liu 2016) for location determination.

(Janicka and Rapinski 2016) used RSSI-based fingerprinting together with machine learning methods such as SVM with radial basis function kernel and nearest neighbour for indoor location. They noted that when the environment is not affected by disturbances, the choice of classification method was less important than in environments that are affected by disturbances. In their experiment the best result was obtained with SVM.

SVM scales well with high-dimensional data, performs well with unstructured and semi-structured data, and small changes do not affect the hyperplanes (Hernández *et al.* 2016). Due to the issues of Wi-Fi received signal strength, such shadowing attenuation and multipath propagation (Chriki, Touati and Snoussi 2017) used SVM to determine the zone of a target. They used SVM as it can easily adapt to multiclass classification problems like indoor localisation. Such an environment has varying numbers and types of obstructions at discrete locations. SVM is able to detect and categorise nodes into different classes based on their features. Other advantages include fast convergence speed, strong generalisation ability, multiple adaption methods and easy integration (Shang and Wang 2022), ability to manage high-dimensional data, and limitation of the effect of outliers (Otchere *et al.* 2021). The disadvantages include high resource requirements, longer training times, required selection of a suitable kernel function for high-dimensionality problems (Ahmadi and Bouallegue 2017) and poor performance given noisy data (Otchere *et al.* 2021).

For SVM, choosing the most suitable kernel is critical to obtain optimal results (Cortez *et al.* 2009). Various sigmoid kernel parameters need to be selected carefully to avoid errors in the results. The polynomial kernel is appropriate when all training samples are normalised. The radial basis function kernel, however, performs best in a low-dimensional dataset (Al-Mejibli, Alwan and Abd Dhafar 2020). The radial basis function kernel was therefore selected because the present research domain has a low-dimensional dataset, the training samples were not normalised since the data capture was already in a uniform scale, and parameter selection such as that of sigmoid is outside the scope of the thesis.

Given SVM's fast convergence speed and its effect on outliers, SVM was chosen as a suitable method, particularly in the present domain where there will be several outliers resulting from RSSI levels captured from the environment.

### **2.9.8 Metrics for evaluation**

The most common metrics used to measure machine learning models include classification (accuracy, precision, recall, F1-score, receiver operating characteristic curve and area under the receiver operating characteristic curve) and regression metrics (mean squared error (MSE) and mean absolute error (MAE)). Classification accuracy is the number of overall correct predictions, whilst precision is the performance within a class. Classification metrics are used for binary identification of entities. True positive and false positive are a viable metric in

classification metrics but this results in a binary decision (Choudhary and Gianey 2017). Such a binary decision will determine whether an obstruction exists at a given location or not. The objective, however, is to determine how precisely the location of an obstruction may be determined. Hence, classification, with metrics, is not deemed suitable for this study.

Regression models are another family of machine learning and statistical models that are used to predict a continuous target value. In the case of the location of an obstruction, the aim is to establish how accurately the position can be determined. However, this research is focused on determining the error of prediction versus the actual location rather than determining if an obstruction is at an exact location or not. This again indicates that classification, with metrics, is not suitable for the goals of this study. This study is more suited to regression models.

In some instances, RSSI levels vary depending on environmental factors, causing the regular occurrence of outliers. These outliers need to be minimised in order for the data to be utilised in subsequent location determination algorithms for accurate predictions (You and Wu 2019). Recent research shows that MSE is one of the most crucial metrics for regression problems such as indoor location determination (Roy and Chowdhury 2021). MSE finds the average squared error between the predicted and actual values, which are shown in equation 2.

$$\text{MSE} = 1/N \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (2)$$

where

$y_i$  is the predicted value and

$\hat{y}_i$  is the actual value.

In indoor location experiments it was determined that MSE outperformed other metrics, such as MAE, in achieving prediction accuracy (Zhang, Qu and Wang 2020). Consequently, MSE was chosen as the predominant metric for determining prediction accuracy in the present research.

Some predictions use the root mean squared error (RMSE), which is the square root of the MSE. However, the RMSE metric is prone to outliers, and these outliers influence the error (Chicco, Warrens and Jurman 2021).

MAE or mean absolute deviation is another metric that finds the average absolute distance between the predicted and target values and is more tolerant to outliers. MAE is defined in equation 3.

$$\text{MAE} = 1/N \sum_{i=1}^N |y_i - \hat{y}_i| \quad (3)$$

It is important to note that MAE and MSE are frequently utilised metrics for indoor fingerprint-based location determination (Alhomayani and Mahoor 2020). Moreover, according to Hyder *et al.* (2021), MAE, MSE, RMSE and the R-squared score (R2) are the most popular metrics for evaluating indoor location methods. MAE calculates the absolute distance between the predicted and actual value, whereas MSE significantly penalises larger errors or outliers by squaring the error distance (Santos and Krishnan 2020). This study seeks to identify the method that produces a majority of predictions that are close to the actual value (actual indoor location), while minimising outliers caused by environmental conditions affecting RSSI signals, as opposed to identifying the method that produces the overall average error distance, which may contain undesirable outliers. MSE is thus the preferred evaluation metric for the purposes of this research.

### **2.9.9 Results of use cases of machine learning for location determination**

Machine learning techniques are often used for iRTLSs (Khatab, Hajihoseini and Ghorashi 2017; Luo *et al.* 2019). Juri, Arslan and Wang (2016) used an algorithm to improve location accuracy in line-of-sight and non-line-of-sight environments. In non-line-of-sight environments, they used concrete walls as obstructions between the transmitter and receiver. Their research identified the two environments using standard deviation. Based on this observation, the algorithm was applied, showing an average improvement of 1.13 m compared to positioning without the algorithm.

(Ahmadi and Bouallegue 2017) note that machine learning accuracy is obtained with no additional hardware requirements and a simple implementation, thereby reducing the cost and energy consumption. In their performance evaluation, they observed that the Naïve Bayes algorithm outperformed the decision tree, SVM and KNN, whilst decision tree and SVM presented a low localisation error compared to the KNN algorithm. It was noted that the main disadvantage of SVM methods is the parameter calibration, which requires a high computational cost.

(Terán, Carrillo and Parra 2018) used SVM and KNN classifiers for indoor location with Wi-Fi and BLE networks. Their results indicated an accuracy level of 75% for a combination of Wi-Fi and BLE using KNN, 68% for a BLE network using KNN and 60% for a BLE network using SVM.

Mahfouz, Nader and Abi-Char (2020) recognise that in some use cases it is adequate to identify the location of a sensor in a zone, for example a room, instead of determining the exact coordinates. They used KNN and SVM on received RSSI levels and fingerprint data and obtained good classification accuracy to locate sensors in the zones. The classification of a sensor to an area rather than an exact location is supported by Janicka and Rapinski (2016).

Hsieh *et al.* (2019) experimented with BLE equipment in a lab environment, collecting RSSI levels at several points averaging 100 values, and then preprocessing this data with a Kalman filter. They then used three different machine learning algorithms, namely KNN, SVM and random forest, to train on the data collected and thereafter predict the position. Results showed that accuracy of the KNN algorithm was marginally higher than SVM and random forest.

Atashi *et al.* (2019) used a fusion framework consisting of a particle filter together with KNN machine learning to improve location. They observed an improvement in location accuracy. Jain, Sashank and Markkandan (2021) noted that an indoor environment is affected by multipath propagation and interference. They employed machine learning classifiers to compare the performance of the random forest, XGBoost, decision tree, KNN, and support vector classifier machine learning algorithms for location determination. Their results showed that random forest machine learning performed better than the others.

Although Wi-Fi demands more resources, (Terán, Carrillo and Parra 2018) used a combination of Wi-Fi and BLE RSSI levels with KNN and SVM machine learning classifiers to estimate location. Their results showed that the KNN classifier of four nearest neighbours delivered better results. Sun *et al.* (2021) proposed a novel convolutional neural network architecture for determining the position of a moving user in a complex indoor environment. They compared the convolutional neural network with KNN, decision tree, logistic regression, SVM, ANN and DNN and found that their newly formulated method yielded the best results.

(Alexander and Kusuma 2019) improved location using Bluetooth and ANN, multiple linear regression, random forest regression and support vector regression machine learning models.

Their results indicated that the support vector regression model performed the best out of the selected models. Yuan *et al.* (2019) found that the KNN, SVM, random forest and multi-layer perceptron together with a particle filter improve the accuracy of indoor positioning. Their results showed that random forest displayed the highest improvement of the selected machine learning models. Yin and Lin (2022) also noted that the most common machine learning algorithms for indoor positioning are KNN, SVM, decision tree and random forest.

## **2.10 Effect of obstructions**

To comprehend the indoor RTLS technologies, one must examine the first advancements in this field, such as the RADAR system. This system was one of the first to utilise radio beacons and RSSI data for localisation purposes (Thaljaoui *et al.* 2015). Several researchers have employed RSSI for indoor location in conjunction with other techniques, such as triangulation, trilateration and fingerprinting, to improve its accuracy. Obstructions, such as partitions and walls, degrade and, in some cases, totally block the signal intensity, necessitating the use of these distinct techniques to determine and increase the accuracy (Abdullah *et al.* 2015). Other obstructions include furniture (Mahfouz, Nader and Abi-Char 2020) and the movement of people and equipment (Hazar *et al.* 2020) in an indoor location. Obstructions such as metal will cause reflection of the signal (Sthapit, Gang and Pyun 2018).

An indoor environment is characterised by many fixed obstructions (such as furniture and equipment) and moving obstructions (such as people) (Hazar *et al.* 2020) between the transmitter of a Bluetooth signal and the wireless access point.

(Naghdi and O’Keefe 2020) note that signal attenuation arising from obstructions, such as from shadowing, is difficult to model and therefore this effect is neglected in literature. They further note that to adequately model RSSI propagation, all obstructions need to be considered within this complicated environment.

## **2.11 Research gap**

Researchers have employed diverse technologies and strategies for indoor localisation. However, the lack of research that considers obstructions in the path of transmission provides an opportunity for new methods, technologies, algorithms and location methods to be defined (Naghdi and O’Keefe 2020). The location of obstructions can be used to improve indoor

location determination. Since machine learning has the ability to learn from input data, this technology has the advantage of improving positioning accuracy, particularly where the variation in measured RSSI levels in indoor positioning systems pose a challenge (Nessa *et al.* 2020). Machine learning is proposed to be used for identification of obstructions in a noisy environment.

From the literature reviewed above, SVM, KNN and ANN were among the most commonly used methods in the field of indoor positioning in the last six years. Other common methods such as Bayesian networks, Naïve Bayes and decision trees, among others, are impractical to be used in the real-time environment, especially for indoor location. Although other methods do yield good results, the choice of method is based on the feasibility of the method, such as for high resource requirements and complexity. Furthermore, the advantages and disadvantages of the various methods employed in indoor location were reviewed. The methods that yield the best results were considered. Other machine learning methods are applicable, but they do not meet this study's identified criteria and constraints, notably cost, form factor, time and practical implementation. Other considerations were training time, transfer learning time and the difference between the training environment and the live environment.

## **2.12 Summary**

There is no optimal solution for indoor location systems, and the best choice of technologies for these systems depends on many factors that have to be carefully balanced: complexity of the environment, economic costs, user requirements and technical parameters, among many others (Moreira *et al.* 2017). Aspects such as accuracy, robustness, scalability and coverage must also be considered when selecting the optimal technology for the construction of an indoor localisation system. There are also constraints in the environment, such as power consumption, low transmission level and form factor. These constraints have been explored together with various indoor location methods. The techniques for determining the location of these technologies were researched and BLE was found to be best suited to meet the criteria identified.

Several machine learning techniques have been employed to enhance location determination, according to the reviewed literature. For indoor location, SVM, ANN and KNN are the most common algorithms and MSE, MAE and RMSE were frequently employed metrics for these

machine learning models. Research on locating obstructions has not been carried out by researchers, providing an opportunity for a new research domain. The complexity of location increases severalfold when obstructions are in motion, and this area will therefore be outside the scope of the present research.

## **CHAPTER THREE: METHODOLOGY**

### **3.1 Introduction**

A methodology in research relates to the underlying approach and justification within the domain of study. It entails studying the methodologies employed in the domain and the theories or principles underlying them in order to design an approach that meets the study's objectives. The methodology outlines the paradigm, design and methods to be used during the research. A description and justification of the chosen methodology, methods, technologies and settings form the basis for conducting the research. This chapter presents how the research was conducted in the collection and analysis of data, and provides a justification of the method employed.

### **3.2 Paradigm and methods**

The required strategy for conducting research involves the interaction of philosophy, study designs and particular procedures (Creswell 2013). The selection of research philosophy, design strategy and techniques will depend on the researcher's field, beliefs and experience. This will serve as the basis for the study's methodology and design.

#### **3.2.1 Research paradigm**

The research strategy for the construction of an iRTLS model must incorporate the researcher's philosophical viewpoint, study designs and research methods. Other authors refer to these perspectives as "paradigms" or distinctions in epistemology or ontology. According to Creswell (2013), worldview is a general philosophical perspective on the world and the essence of the researcher's contribution to the study. As summarised in table 3.1, Creswell explains that these worldviews can be classified as constructivist, positivist, post-positivist, pragmatist or transformative methods.

**Table 3.1: Four worldviews (Creswell 2013)**

<b>Post-positivism</b>	<b>Constructivism</b>
<ul style="list-style-type: none"> <li>• Determination</li> <li>• Reductionism</li> <li>• Empirical observation and measurement</li> <li>• Theory verification</li> </ul>	<ul style="list-style-type: none"> <li>• Understanding</li> <li>• Multiple participant meanings</li> <li>• Social and historical construction</li> <li>• Theory generation</li> </ul>
<b>Transformative</b>	<b>Pragmatism</b>
<ul style="list-style-type: none"> <li>• Political</li> <li>• Power and justice oriented</li> <li>• Collaborative</li> <li>• Change oriented</li> </ul>	<ul style="list-style-type: none"> <li>• Pluralistic integration of two or more worldviews to solve a real-world issue</li> </ul>
<b>Positivism</b>	
<ul style="list-style-type: none"> <li>• Experimental</li> <li>• Reliance on current theories and methods</li> <li>• Reliance on established procedures for achieving quality</li> </ul>	

According to Guba (1994), when defining a paradigm, the following three questions must be addressed:

1. What is the nature of the discussed reality (ontology)?
2. What are the aspects of knowledge (epistemology)?
3. What is the most efficient approach (methodology) for acquiring the necessary knowledge and comprehension?

Positivism is based on objectivity that applies established procedures to examine phenomena, generalise reality to develop a model, and assess the quality of the resulting output (Lincoln 2011). Positivism entails utilising current methods or developing new methods by adapting existing methods (Creswell, 2013). This version of positivism also incorporates a reductionist approach, as the goal is to reduce concepts into a discrete collection of ideas in order to analyse the variables comprising hypotheses and research questions. Post-positivism, like positivism, posits that an underlying theory can be replaced with a derived theory. If there is no verified theory, this worldview is irrelevant. A transformative worldview incorporates studies on political strategy and transformation (Mertens 2015). This viewpoint is inapplicable if research

is unrelated to political transformation. Another viewpoint, constructivism, focuses on the different opinions of community members. The constructivist worldview is incompatible with research that depends solely on objective facts and disregards participant viewpoints.

After analysing the paradigms presented in the literature and comparing them to the research environment, positivism was deemed the most suitable research paradigm for the study. This research environment was restricted to collecting empirical data from machines, which is consistent with the positivist worldview, as opposed to using research based on the perspectives of research environment participants (constructivism), validating an existing theory (post-positivism) or advocating for change on behalf of these participants (transformativism). Since this research is aligned with the positivist worldview, there is no need to investigate other worldviews (such as pragmatism). Due to its use of mathematical formalism, reductionism and reliance on proven methodologies and theories for both development and quality, positivism was chosen as the most acceptable paradigm.

### **3.2.2 Research design**

The research design was based on a positivist experimental methodology in order to provide an objective, tangible iRTLS model. The research approach used for this research was design science with iterative experiments involving development, simulation and testing. The primary objective was to use machine learning to determine the position of obstructions in the path of Bluetooth transmission. In order to achieve the primary objective, the environment needed to be established through research, experimentation and design.

The main phases are as follows:

1. Identify the most suitable technologies that can be used for an iRTLS.
2. Identify the range of the selected location technology through experimentation and analysis.
3. Establish the constraints, including the range, under which this system needs to operate.
4. Evaluate the most feasible methods for location determination within the constraints.
5. Develop a scalable design model based on initial fingerprinting of the environment and subsequent quality link indicator to establish expected values in an obstruction free environment.

6. Based on this model, investigate changes in received values using the selected technology that indicate possible obstructions in the path of transmission within the given environment.
7. Based on the changes of these received values, train the selected machine learning models. Based on the trained machine learning models, predict the location of these obstructions, and evaluate the results.

As workplace in this study was simulated, the constraints that would be experienced in a typical office environment needed to be considered. Such constraints included obstructions, cost, form factor, energy consumption, low transmission levels, location accuracy and detection range. Cost is a factor that sometimes limits technologies' ability to be location-based. Consequently, location-based technologies needed to be scaled up or down to provide for resizing of the network as well as the varying number of devices in the network. An important factor was energy consumption, which needed to be minimised to ensure that battery life was maximised. Efficient algorithms had to be designed to balance energy consumption and the availability of accurate real-time location. In certain sensitive environments, the possible electromagnetic interference would have to be minimised whilst balancing transmitted signal strength with maximum detection range. In some instances, obstructions prevent a signal from reaching its intended target entirely. Therefore, it was essential to identify the exact location of these obstructions.

### **3.2.3 Research methods**

Within the positivist worldview and its corresponding methodological framework, information technology researchers have historically deployed a variety of research approaches that can be generally divided into two categories: qualitative and quantitative (Mehrad and Zangeneh (2019)). The purpose of qualitative research methods is to comprehend problems by studying the perspective and behaviour of persons in these settings, as well as the context in which they function, resulting in a profound understanding of the meaning and context of the phenomena studied in the actual world (Kaplan and Maxwell 2005). The qualitative approaches usually rely on opinions, case studies or complex phenomena and are sensitive to local settings, specific circumstances and the needs of stakeholders. They have the disadvantage of being difficult to generalise and forecast, as well as being exceedingly time-consuming. Nonetheless, qualitative research methods can be utilised to obtain rich insights into restricted issues. Quantitative

research methods, meanwhile, focus heavily on the collecting of quantitative data, the identification of variables and constants, and the examination of their correlations (Johnson and Christensen 2019). Using quantitative research methodologies, one can generalise and make predictions based on the exact data collected. This form of research is beneficial for large samples and reduces prejudice among researchers. However, quantitative research makes it challenging to examine issues in depth.

### **3.3 Application of methodology**

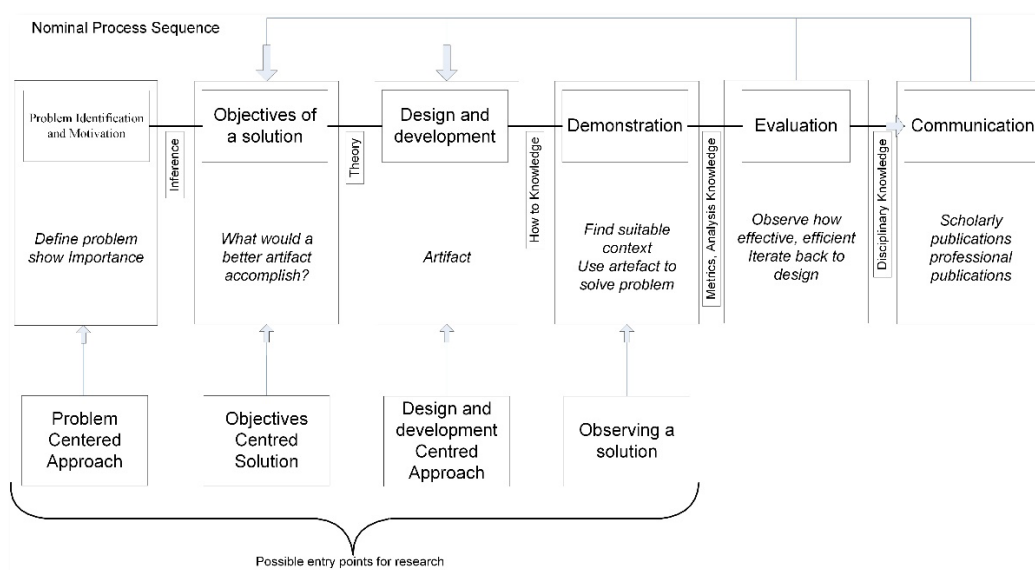
A methodology is a set of guidelines followed in a discipline to achieve an outcome and ensure that nothing is overlooked in the process (Davis and Yen 2019). These guidelines use systematic processes or instruments to acquire and analyse the research data. The three primary types of research design are quantitative, qualitative and mixed. Quantitative research designs (data collection, analysis and generalisation) differ from qualitative designs (understanding human experience and meaning within a given context) (Petty, Thomson and Stew 2012). In the present study, no new software or new theories are being developed. Instead, this research uses existing evaluation methodologies to determine the optimal option. The experimental design was adopted as a methodology. This experimental design made it possible to control numeric attributes and measure causal effects (Creswell 2013). This research manipulated factors such as obstructions and assessed their causative impacts, including RSSI levels. On the basis of experimental methodology, design science offers a more structured approach to investigation and evaluation, among other aspects (Nunamaker Jr, Chen and Purdin 1990).

#### **3.3.1 Design science methodology**

Systems development as a research methodology is a complex strategy that includes an in-depth examination of a problem, the construction of effective concrete solutions and the empirical evaluation of those answers (Nunamaker Jr and Chen 1990). From the perspective of research methodology, the steps of systems development procedures can be summarised as follows: construct a system, derive theories and principles from observing behaviour, and incorporate software tool expertise to boost availability. In turn, these tools can be utilised to aid in the development of new systems (Kim 2013). This methodology begins with an emphasis on design and development, similar to the design science process.

Design science has multiple entry points based on research or experimental requirements. This study seeks to explain observable occurrences, resulting in a positivist approach to research. The generated artifact can collect data about phenomena and apply numerous techniques to this data in order to achieve the desired outcome. Design science is more adaptable to research on software and systems with minimally specified requirements. As experiments are done, parameters and requirements change, and this necessitates tool adaptation. Design science, with distinct phases, beginning with problem identification and ending with the communication of results, was thus considered ideal for this study. In addition, evaluation and communication are not included in the preceding methodologies. In fact, other research methodologies, such as Waterfall and Agile, fit within the design, development and demonstration phases of design science.

Design science provides a model for the design of a model. There are two applications for empirical research: first, the validation of a created artifact before its implementation and, second, the evaluation of a design's performance once it has been implemented (Wieringa 2010). According to Peffers *et al.* (2006), design science is a prominent process model for conducting research using the following six steps: problem definition and motivation, solution objectives, design and development, demonstration, evaluation and communication, as shown in figure 3.1 below. This process intends to generate and offer action instructions that enable the design and operation of information technology and novel concepts within information systems, resulting in artifacts, models, methods and instantiations (Hevner 2004).



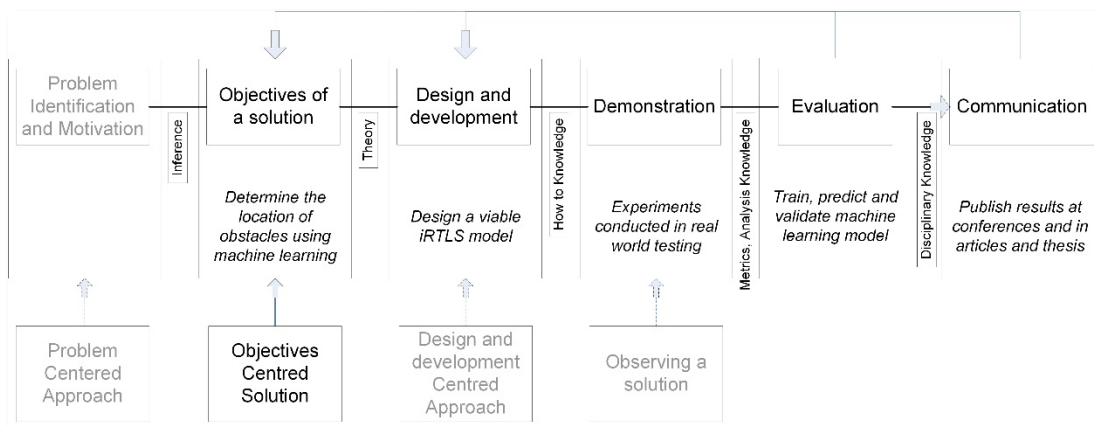
**Figure 3.1: Nominal design science process**

### 3.3.2 Application of design science methodology

The application of the design science methodology to this research provides for the feedback loops and iterations that are required for modelling the iRTLS, as the requirements and scenarios are not well defined. Therefore, the process required several opportunities for confirmation, reworking and redefinition in order to optimise location determination as well as location of obstructions.

The entry point for this research was an objective-centered approach, as illustrated in figure 3.2 below. The objectives that follow the main phases in the research design are:

1. Identify the range of the selected location technology through experimentation and analysis.
2. Develop a scalable design model based on initial fingerprinting of the environment and subsequent quality link indicator to establish expected values in an obstruction-free environment.
3. Based on this model, investigate changes in received values using the selected technology that indicate possible obstructions in the path of transmission within the given environment.
4. Based on the changes in these received values, apply and evaluate selected machine learning methods to most accurately determine the location of these obstructions.



**Figure 3.2: Application of design science**

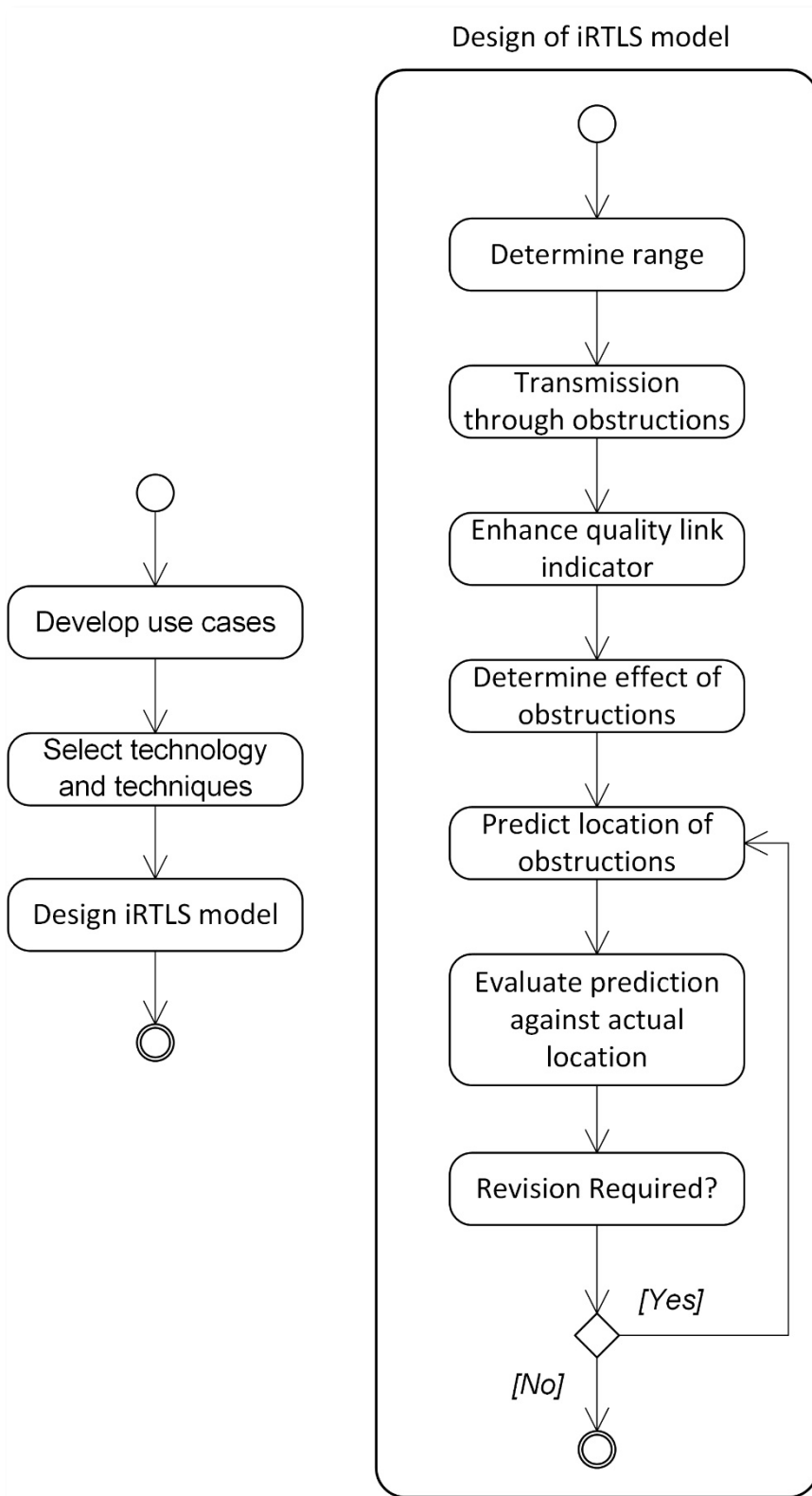
As per the design science methodology, experiments were conducted iteratively, the results evaluated and the methods refined until the desired outcomes were achieved. This included

experiments to determine the best possible machine learning method for obstruction location within the constraints that would mimic a typical office environment.

### **3.3.3 Design process**

As an extension of design science, the process to design an iRTLS model can be depicted in more detail using a UML activity diagram, as indicated in figure 3.3 below. The process commenced with identification of the most suitable technologies that can be used for an iRTLS through a detailed literature review and was followed by the determination of the range of the most suitable technology through experimentation and analysis. Once the constraints were established, the most feasible methods for location determination within these constraints were evaluated. This process led to the development of a design model using initial fingerprinting and, subsequently, using a link quality indicator to establish the expected RSSI levels in an obstruction-free environment. Obstructions were introduced and changes were observed. Based on the changes observed, machine learning was used to most accurately determine the location of obstructions in the path of transmission.

The steps of evaluation included conducting experiments and evaluating whether the objectives were met. If they were not met, the previous processes were repeated until a finalised model was achieved, as indicated the loop in figure 3.3.



**Figure 3.3: Design process for an iRTLS model**

This study followed the steps outlined below to achieve its objectives:

1. Develop use cases.
2. Select technology and techniques.
3. Develop the model.

In order to develop the model, the process followed was to:

- 3.1 determine the range and throughput
- 3.2 determine the transmission through obstructions
- 3.3 enhance the real-time quality link indicator
- 3.4 determine the effect of obstructions
- 3.5 determine the location of obstructions using machine learning.

### **3.4 Methods**

The most common method used to determine the location of BLE nodes and obstructions is RSSI level received at the receiver. The implementation of such a method is very cost effective as BLE RSSI is readily available in most IoT devices. Furthermore, the RSSI levels can be easily measured by cost-effective receivers that easily integrate into a WLAN environment for onward transmission. Fingerprints of RSSI levels for nodes and obstructions were recorded during the initial phase. These were then used in the subsequent phase for location determination in real time. Another method employed was a link quality indicator, where the RSSI levels were recorded for gateways at known locations. These measurements were taken regularly and used for location determination.

Although there are many techniques, the following were not adopted for the given reasons. Angle of arrival requires specialised reception hardware, which increases cost and form factor. Arrival time needs network-wide clock synchronisation and high precision, which increase overall cost. For correct time of arrival determination, a centralised server must establish and maintain the accurate time across the entire network. Pedestrian dead reckoning was not suggested because it required extra hardware and regular calibration, both of which increase node cost and form factor.

### 3.5 Technologies

BLE is the latest technology that consumes lower power, (Yang *et al.* 2020), is cost-effective and has a small form factor (Jeon *et al.* 2018). Most IoT devices have BLE signals as an option. As such, BLE was the most suitable technology for further experimentation as it met the objectives of low power consumption, low interference, long battery life and availability in IoT devices. This makes it a feasible solution for large-scale implementation. Once the location of an obstruction is determined, the location of nodes can be determined more easily and accurately. Testing was conducted in a simulated use case environment from the use cases identified. Although BLE is gaining importance in the area of indoor location, there is still a lack of open datasets (Baronti *et al.* 2018), hence the need to generate new datasets for experimentation and analysis. Therefore, experiments were set up to capture the required data in the environment and setting, as determined by the objectives. However, open datasets often contain cleansed data and there was therefore no need for data preprocessing. This research was designed to investigate a real-world environment and determine the most suitable data preprocessing method(s).

The following are some of the technologies deemed infeasible for achieving the objectives outlined, along with a brief description of why they were not selected:

- Wi-Fi requires an IP address to connect to a WLAN network. As more devices connect, the network becomes overloaded.
- Cost was cited as the key factor for the slow adoption of RFID, which involves nodes and receivers placed at strategic locations. Furthermore, the RFID range depends on the receiver's quality.
- Near-field communication is short-range technology, so nodes without batteries must be placed close to the receiver for communication.

### 3.6 Environment

There are several important use cases that require location of people and assets in an indoor environment. This research is restricted to use cases in a typical office environment that comprises people, office furniture made of different types of material, electronic equipment and office walls made of different types of material, amongst other items. Because Bluetooth

transmission is affected by obstructions, the researcher simulated an office environment by using steel metal shelves arranged vertically, as this would represent a steel filing cabinet in the path of Bluetooth transmission.

This experiment was set up to establish the impact of such an obstruction in the path of Bluetooth transmission and to be able to determine where such an obstruction is located. Other common use cases in a real-world scenario include location of people and assets. Both will need to have a BLE tag associated with them. For the purposes of this study and experimentation, people and assets entities were represented by nodes equipped with a BLE transmitter and receiver. The office environment's content reflected signal transmission between the node and the receiver. Such an environment can be altered at any time by the movement of people or objects, resulting in a modified radio map. The complexity, resource requirements and computational power required to continually update to determine the obstruction location via the chosen process was not explored due to the extended scope this would have entailed. Using items and materials typically seen in office environments, a simulated environment was set up. Because electromagnetic waves move in a straight line, the measurements were recorded in a straight line. The objective of the experiment was to determine the location of the obstruction in the transmission path, leading to the determination of the location of the target.

### **3.7 Experimental setup**

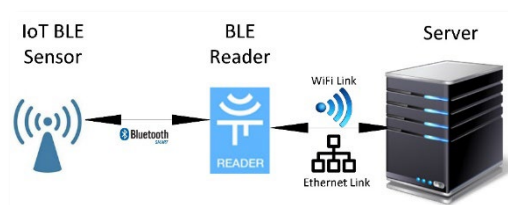
#### **3.7.1 Bluetooth Low Energy**

The most appropriate technology identified from literature that satisfies the constraints is BLE. The strategy in this study was to use a combination of BLE tags represented by nodes, receivers to receive the BLE data, and appropriate algorithms to demonstrate the viability of an RTLS that overcomes these obstructions. The benefits of BLE devices is that it has lower cost, lower power consumption and longer battery life (Lee, Dong and Sun 2015) than other technologies. Both readers and nodes had unique identifiers to identify the data. The readers had a fixed position whilst the nodes moved within the environment to collect the required data.

### 3.7.2 Received signal strength indicator

Fingerprinting is a pattern-based indoor localisation approach used to characterise an indoor environment by matching measurement data with a previously determined pattern (Jachimczyk, Dziak and Kulesza 2016). The signal strength together with other properties are measured and stored at predefined locations in advance (Kim *et al.* 2015). In this study, the power level readings calculated in real time, together with a combination of other methods, were used to approximate the location of the node. After a detailed fingerprinting exercise, readers were placed at strategic points to cover the pilot area under study.

Using the RSSI power levels, messages received from the nodes, triangulation as well as the fingerprint impressions and/or the quality link indicator, a location of the node was determined with improved accuracy. The objective was to evaluate the lowest transmission level of -20 dBm to study the range through obstructions for practical application, so that the solution may be implemented in an office setting. Broadcasting at the lowest level will reduce interference, for example on medical or other equipment, and permit the use of additional devices within the same frequency ranges. The high-level architecture of the node, reader and server and their proposed communication mediums is depicted in figure 3.4 below.

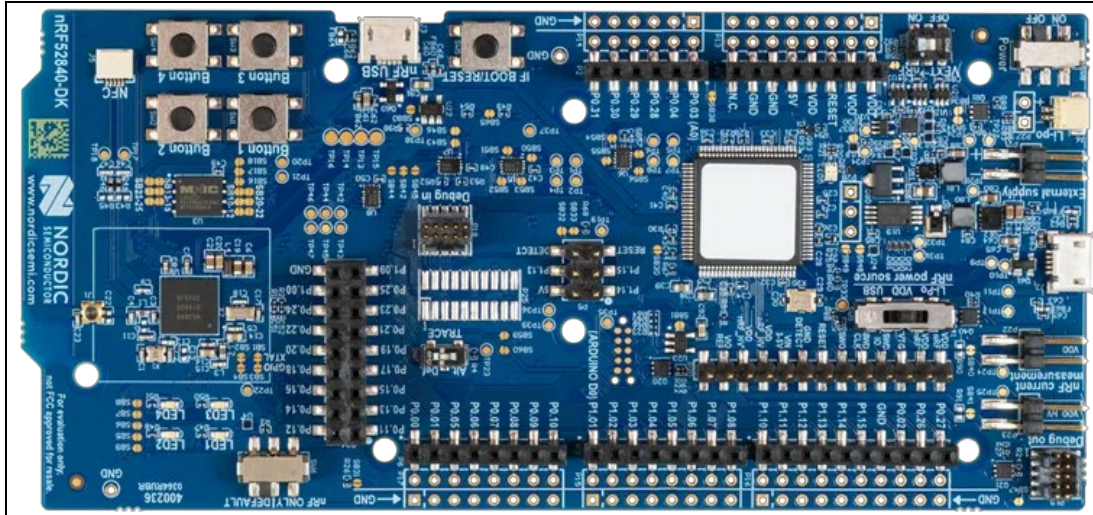


**Figure 3.4: IoT BLE architecture**

### 3.7.3 Hardware

The hardware selected for preliminary testing were two Nordic nRF52840 Preview Development Kits (PDKs), which was also used in the study by Terán, Carrillo and Parra (2018) for their indoor experiments. The nRF52840 BLE System on Chip was deployed using Segger Embedded Studio for application creation, testing and deployment. The selection of this most recent BLE System on Chip was based on the numerous benefits outlined by Collotta *et al.* (2018). The latest BLE V5 was selected for its mesh network capability to offer advanced scalability. At the time of commencement of the research, the two main suppliers identified,

namely Nordic and Texas Instruments, had Bluetooth 5 PDKs available for testing and evaluation. Nordic was chosen as the preferred vendor because of its price advantage, feature set and availability over Texas Instruments' comparable offerings. A picture of the Nordic nRF52840 PDK is shown in figure 3.5.



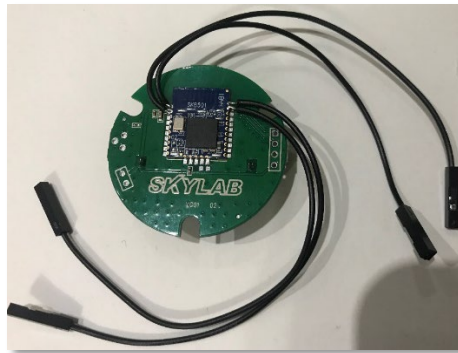
**Figure 3.5: Nordic nRF52840 PDK**

(Nordic 2018)

The initial experimentation conducted with this hardware included investigation of the range limit, assessment of feasible methods used in RTLS and analysis of BLE detection range improvements. Although initial hardware was identified during the preliminary work, additional hardware for the readers and BLE nodes were configured to provide the connectivity between the readers, nodes and wireless access points. The selected hardware employs the most recent Bluetooth technology, Bluetooth 5, and the module was powered by two AAA batteries. The firmware has been designed to transmit the device's Universally Unique Identifier and the message identifier at regular intervals. The hardware and obstructions were set up at different positions, and the data was captured from the nodes and transmitted to a server for processing.

As depicted in figures 3.6 and 3.7, the Skylab SKB501 solution was used as the low power node (LPN) in this experiment. This solution utilised the Bluetooth 5 System on Chip technology from Nordic's nRF52840. It is designed to benefit from the advanced features of Bluetooth 5 and its enhanced performance capabilities, which include modes with a long range

and high throughput. The Bluetooth 5 specification empowers the nRF52840 and other similar Systems on Chip to take advantage of BLE V5's significant performance improvements.

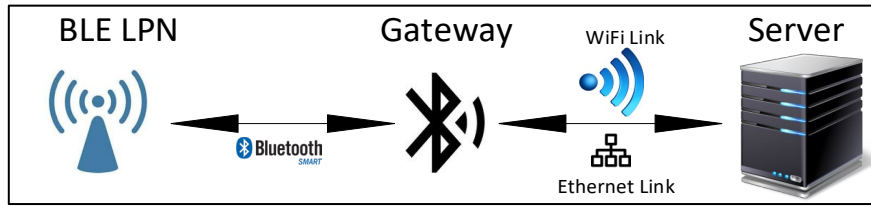


**Figure 3.6: SKB501-LPN top view**



**Figure 3.7: SKB501-LPN rear view**

A Raspberry Pi 4 running Ubuntu V18.04 served as the gateway connecting the LPN and the server. BlueZ provided support for the fundamental Bluetooth layers and protocols on the Raspberry Pi. It was chosen as the best option due to its adaptability, efficiency and modular implementation. The gateway software monitors the Bluetooth interface for broadcasts and transforms messages for transmission via the Wi-Fi network port to the server. This C# program uses User Datagram Protocol (UDP) to transfer messages from the gateway to the server. The server's operating system is Ubuntu and its database is PostgreSQL. The program that receives the message from the gateway performs a continual scan for incoming UDP communications. Using C#-built software, these messages are written to a PostgreSQL database's data table. Figure 3.8 illustrates this architecture at its highest level.



**Figure 3.8: Architecture**

### 3.7.4 Location determination

The known locations and parameters of the nodes were recorded. Based on this data, a machine learning algorithm was used to determine a footprint of the predetermined area. This footprint was updated at regular intervals, and the machine learning process improved the accuracy of the location coordinates by using previously calculated data at real time. The location was determined and compared to the location determined from the footprint. The results of the different methods were also compared to each other in terms of their accuracy.

Once the location was determined to the identified level of accuracy, the processes of optimisation and balancing of attributes were conducted. Consideration was given to the identified attributes, weights of attributes, routing paths, transmission frequency and a scalable model of the solution. The model provided a solution with which to modify the different factors so that an optimal design could be obtained for the predetermined use case.

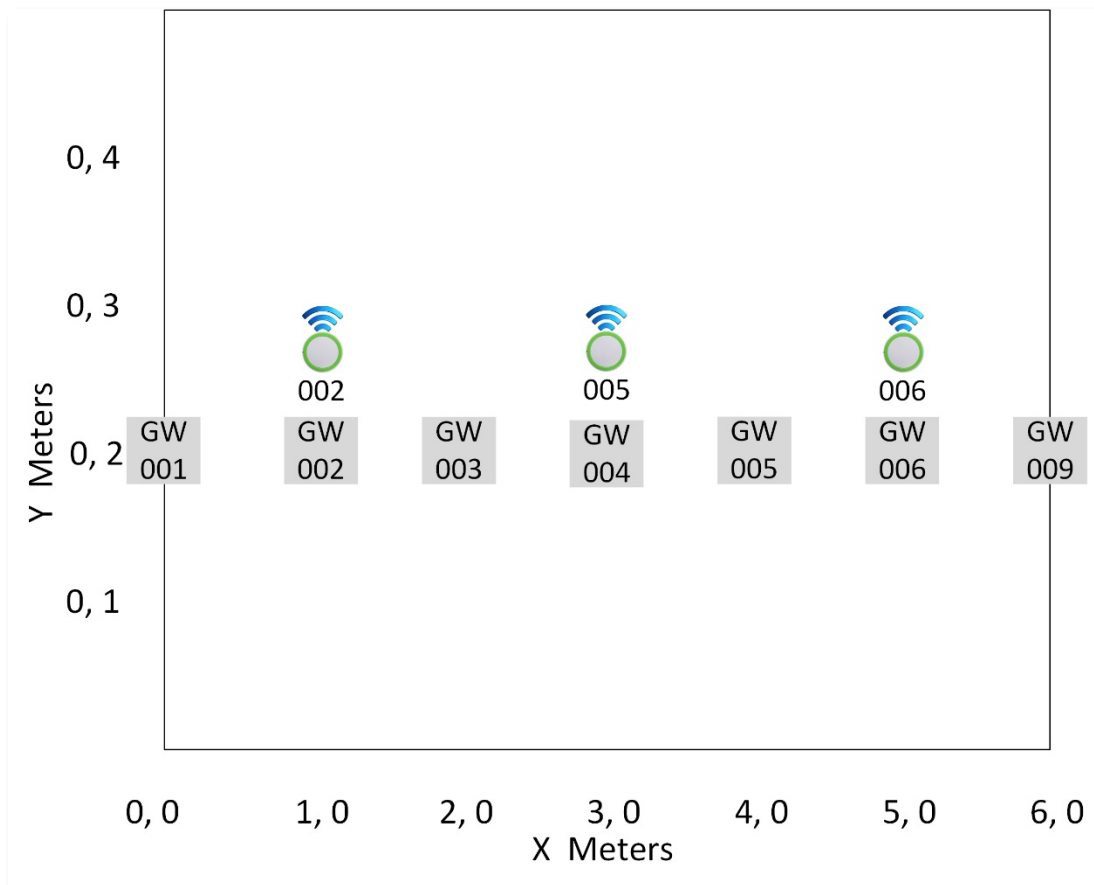
## 3.8 Design processes and configuration

In this section, the data gathering process and its associated equipment configuration are described. The methodology for the iRTLS model design is then outlined. Lastly, the implementation procedure for machine learning models is explained.

### 3.8.1 Data collection

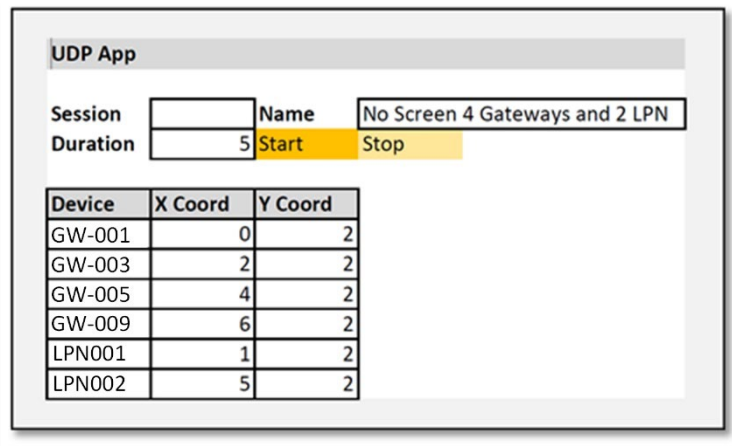
The nodes' firmware was programmed to broadcast the transmit RSSI level, node identifier and unique message identifier every 100 milliseconds. Each gateway transmitted its own identifying information. The gateway and node mac addresses were initially saved so that they could be identified uniquely on the network. The unique message identifier could be utilised to analyse message reception at each gateway. This experiment imitated an office environment with partitions separating workspaces. The obstruction consisted of two 950 cm × 450 cm ×

0.8 mm metal sheets. Each test lasted five minutes, and this procedure was performed three times, that is, test 1, test 2 and test 3. To monitor and analyse the consistency of the measurements taken across tests, the tests were repeated. The GW-001 to GW-006 and GW-009 gateways were spaced 1 m apart at (x, y) coordinates (0, 2; to 6, 2). Nodes labelled 002, 005, and 006 were respectively positioned at (x, y) coordinates (1, 2; 3, 2; and 5, 2). This layout is depicted in figure 3.9.



**Figure 3.9: Layout of nodes and gateways**

An application was used to configure the gateway and node positions, specify the test time and initiate the test. The session name, session ID and automatically produced identifier were used to identify each test's collected dataset. For each test, the positions of the gates and LPNs were saved. Figure 3.10 depicts the program configuration and data capture interface.



**Figure 3.10: Application user interface**

### 3.8.2 RTLS model design notations

As shown in the literature review, UML is a standard used to describe the requirements and the design of systems. Since this study uses a high-level design, detailed documentation tools such as formal methods were not considered as the design would need to be adaptable and hence agile. Also, an object-oriented design lent itself to the subsequent processes, which were object oriented.

The requirement for the model design was described using use cases and scenarios together with use case diagrams. An activity diagram was used to demonstrate the process flow of a particular scenario. A class scope diagram was used to demonstrate how the activity diagram was properly coupled with the domain class model.

### 3.8.3 Machine learning

#### 3.8.3.1 Algorithms

Machine learning methods were used to determine the most likely position of the obstruction. The methods chosen for this research were KNN, SVM and ANN. These machine learning methods, together with other machine learning methods, are also used by researchers to determine location of BLE nodes in a clear (obstruction-free) path of transmission between transmitter and receiver. However, KNN, SVM and ANN yielded better performance and accuracy than other methods such as Naïve Bayes and decision trees. In this research, KNN,

SVM and ANN were used to determine the location of obstructions so that this information could be used in further processing to determine the position of nodes.

### **3.8.3.2 Metrics**

Evaluating model accuracy is a crucial step in the creation of machine learning models, since it describes how well the model performs in terms of its predictions. The evaluation metrics vary based on the problem category. The linear model (regression) is a common example of this sort of problem, and the most distinguishing feature of the regression problem is that the dataset's targets only contain real values. The errors show the extent to which the model predicts incorrectly. The fundamental principle of accuracy evaluation is to compare the actual target to the expected target based on a set of criteria.

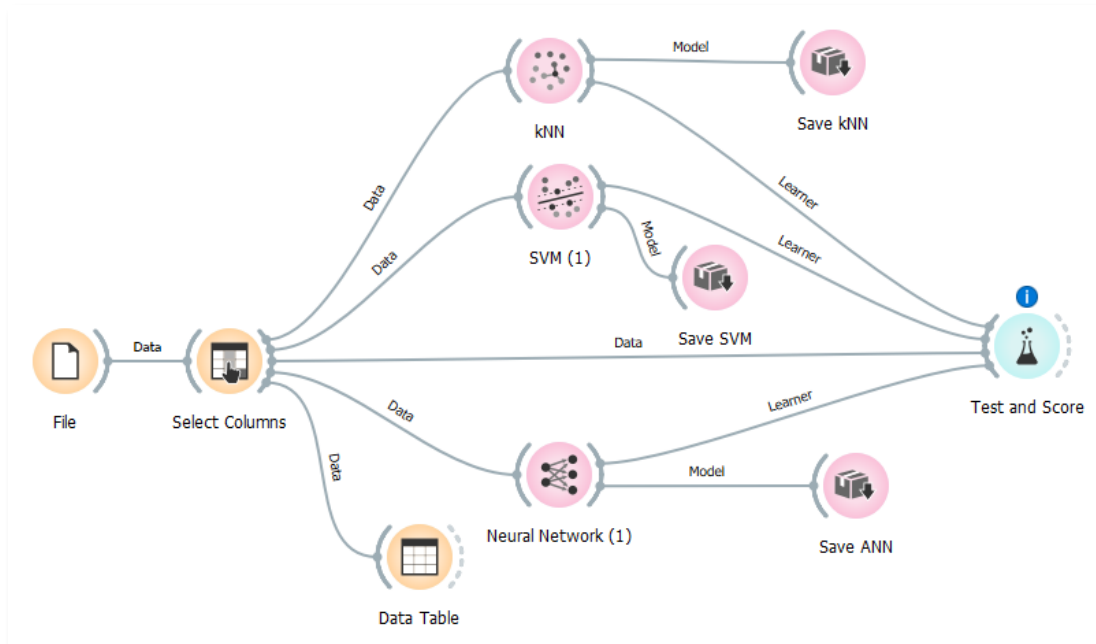
In regression analysis, the MSE, MAE, RMSE and R2 metrics are primarily used to measure prediction error rates and model performance. They were common across the machine learning models and most widely used for these models by researchers (Harishkumar, Yogesh and Gad 2020; Joshua, Priyadharson and Kannadasan 2021). MSE penalises larger prediction errors and is the preferred metric over MAE, which provides a total absolute error distance. This research involved identifying the method where the error distance between the predicted values and the actual location is minimal among the majority of clustered predicted values. In section 2.9, these widely applied location-based metrics were examined. They were identified as being most suitable to accomplish the research goal.

The R2 metric is a measure of the variance of correlation between the variables but it does not focus on the predictive error (Chicco, Warrens and Jurman 2021). As this research is focused on determining error of prediction versus the actual location, the R2 metric was not deemed suitable. Moreover, due to the fixed locations of the nodes and gateways, there was a strong correlation between these variables. Consequently, this metric was not used for prediction and evaluation of methods, but the data is included for completeness.

Once the datasets were obtained, an adequate subset of data was required for training and validating the machine learning model. Research indicates that the percentage used for training affects the estimation level, and that this percentage varies. However, it has been determined that a 70% training and 30% testing (prediction) ratio would yield among the best results (Gholamy, Kreinovich and Kosheleva 2018).

Orange was selected as a tool to implement the selected methods, record the results and perform the analysis using the selected methods and metrics. This tool was selected because it is preprogrammed to use the required methods and is quickly implemented. Furthermore, it has inbuilt analysis tools to determine the most suitable method and metric. Initial experiments were conducted to determine the optimal settings of many of the methods used in the Orange tool during the evaluation phase.

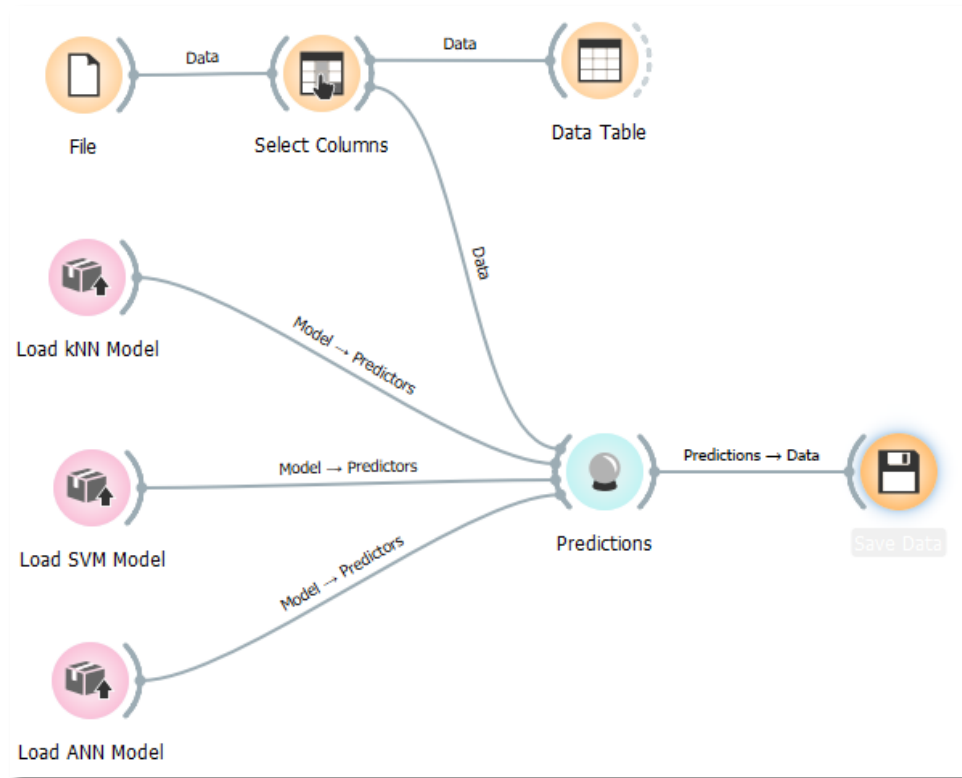
Figure 3.11 below depicts the configuration of a machine learning network in the Orange software. Excel files were utilised to provide input to Orange. The required columns were chosen and configured as “features” to be used in the machine learning process, with the predicted value serving as the “target”. The data was then associated with the machine learning model, which was in turn associated with the training model. After training had been completed, the results were collected for analysis, and the trained model was then saved for use in the prediction phase.



**Figure 3.11: Sample machine learning training network**

Excel was also used to compile the data for the prediction, which was fed into the prediction model. Each of the trained models was loaded into the prediction system and fed data. The model’s output was fed into the predictor in order to generate prediction results. As illustrated

in figure 3.12 after the prediction had been completed, the results were collected for further evaluation.



**Figure 3.12: Sample machine learning prediction network**

The high-level view of the experimental approach for obstruction location determination using machine learning is as follows:

Step 1: Conduct experiment without obstructions as a base case.

Step 2: Place metal obstruction(s), which simulates metal filing cabinets typically found in an office environment, at specified location(s) to evaluate their effect on BLE transmission.

Step 3: Based on these obstructions' locations, observe the RSSI levels and record them in datasets.

Step 4: Using these datasets, apply and evaluate select machine learning methods to detect location of obstruction(s) accurately using chosen metrics.

The placement of obstructions and the observation of the RSSI levels were conducted in four separate experiments, and the RSSI levels were recorded for three five-minute sessions in each experiment, resulting in three datasets for each experiment. The layouts for the four experiments were as follows:

Experiment 1: Obstruction was placed at position (0.5, 2)

Experiment 2: Obstructions were placed at position (0.5, 2) and (1.5, 2)

Experiment 3: Obstruction was placed at position (2.5, 2)

Experiment 4: Obstructions were placed at position (2.5, 2) and (4.5, 2)

Selected machine learning methods were applied to each of the datasets using the following process:

Step 1: Set the obstruction(s).

Step 2: Record the RSSI levels at each of the gateways and record these as datasets. These datasets represent a five-minute recording interval under the same environmental conditions.

Step 3: Train the models with 70% of one dataset. Predict the location of the obstruction and validate it using 30% of the same dataset.

Step 4: Also predict the location of the obstruction using the remaining two datasets captured from the same experiment.

Step 5: Use steps 1 to 4 for each experiment outline above.

### **3.9 Summary**

Using the methodology described above, the research was designed to achieve the objectives of the study. The design science methodology provided the structure to create the model using the objectives and proceed through to the subsequent phases. Wi-Fi and BLE provided the communication infrastructure necessary for conducting the experiments. Using the chosen hardware, data collection methods were designed to produce a framework that facilitated the collection of data for the use cases and scenarios. Appropriate machine learning methods and metrics were identified for training and prediction of the location of obstructions based on the data collected.

## CHAPTER FOUR: RESULTS

### 4.1 Introduction

The iRTLS model proposed in the previous chapter was implemented and tested within a practical office environment. The results of the detailed experiments that determined the throughput and effect of obstructions on BLE transmission were documented. This then led to the application of machine learning to establish the feasibility of location of obstructions. Summarised results are included in this chapter and the detailed results are included in the appendices.

### 4.2 iRTLS model

The iRTLS model was designed using UML to document the model and design science. A user-centred view and the primary use case diagram were used to communicate the model's requirements. A scenario between a node, the gateway and the server was described using an activity diagram during the analysis phase. The design outlines the primary use cases and message configuration. The messages offered the protocol for communication between the nodes, gateway and server. The model comprised the communication protocol, use cases and activity diagram.

#### 4.2.1 Requirements of an indoor real-time location system

The requirement for an iRTLS is that it should be able to locate a person or an asset such as a personal computer, medical equipment, etc. In order for this location to be successful, a hardware device called a node needs to be attached to the person or the asset. The hardware device will then communicate with the server via a gateway. The software residing on the node, gateway and server will enable the system to determine the location of the node and hence the location of the person or asset.

The use case, therefore, is to be able to determine the location of the node. It must also be noted that the node can be in motion and/or change its location. However, the environment is populated with items that are found in a real-world workplace. This environment can also be built with different types of materials. The items found in a workplace and the materials used for construction affect the electromagnetic signals in different ways. The constraints identified

for an iRTLS are cost, form factor, power consumption, interference and range, coverage, and scalability. A model was designed for the requirements and the constraints identified.

In a real-world environment, some of the objects that are being located will be in motion or will change locations. These obstructions can also change locations or be placed in the path of transmission. These obstructions will affect the data captured for processing for location determination of these objects. Therefore, the design considered both the object to be located as well as the obstruction to be in fixed positions.

The user-centred view of the interaction of the node and the events that lead up to location determination of the node are shown in figure 4.1. A node that is used for location determination is associated with the person or an asset. In this case, the node is a wrist band worn by a person. There are two instances that will trigger a broadcast: either a movement detected or an event triggered by an inactivity timeout interval. The receiving gateway receives and processes the message and sends an acknowledgement of receipt. This process ensures that the message sent was acknowledged by the recipient.

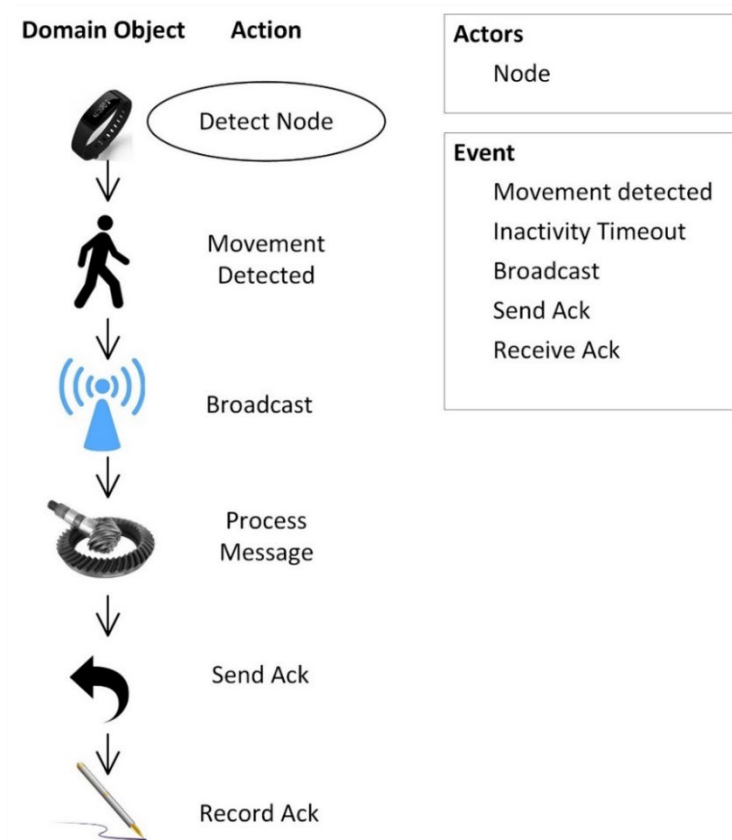


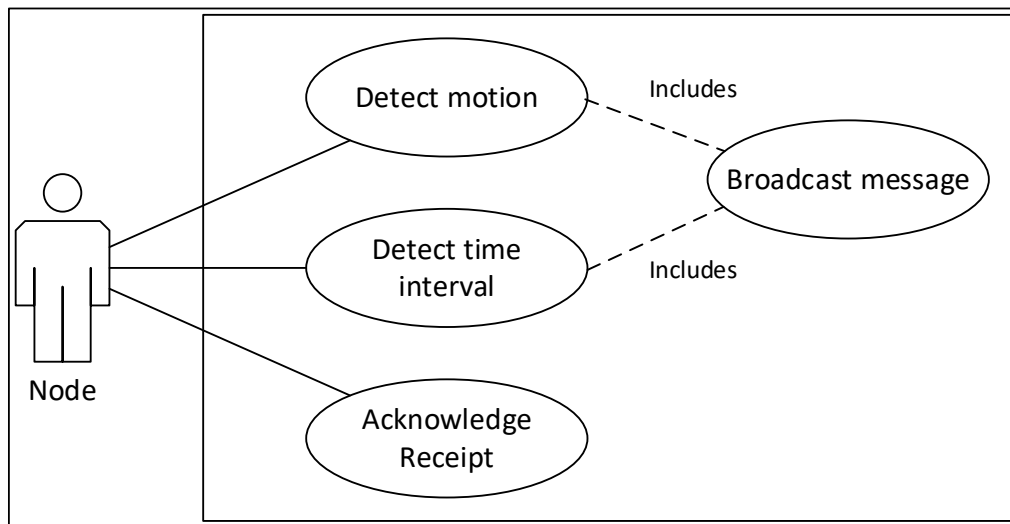
Figure 4.1: Legend of entities in detect node scenario

Table 4.1 below details the main use case, “Determine Location”. The node needs to broadcast a message in order to determine its location, and that message is received by the neighbouring gateways. The message transmitted by the node is triggered by either a node’s motion or the expiration of an inactivity timeout interval. The information contained in the transmitted message, the transmitted RSSI level and the received RSSI level will enable the determination of the location.

**Table 4.1: Business use case Determine Location**

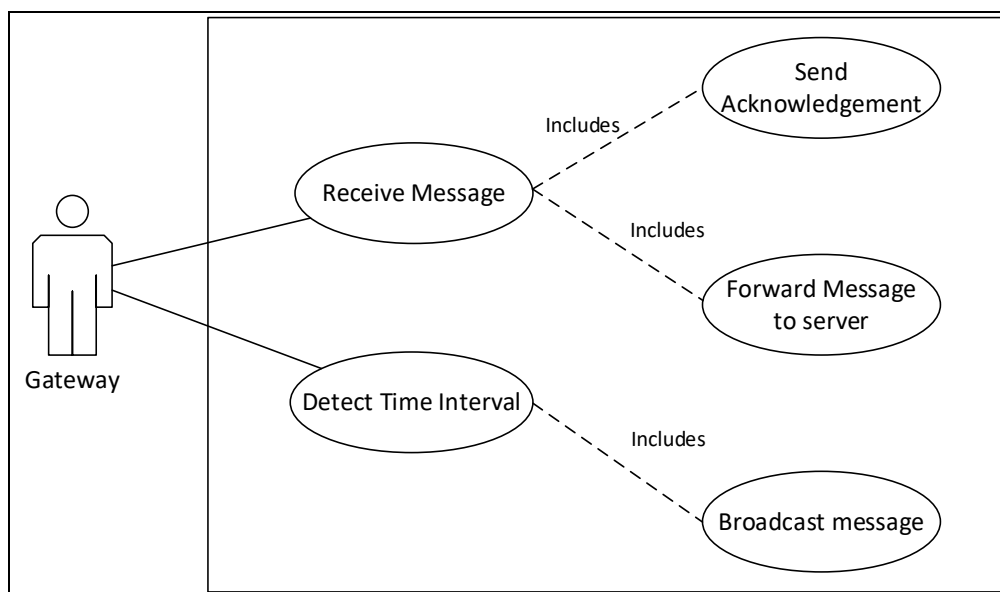
<b>Title</b>	<b>Description</b>
<b>Use case no.</b>	UC01
<b>Name</b>	Determine Location
<b>Related requirements</b>	None
<b>Goal in context</b>	To determine the location of the node
<b>Trigger</b>	Motion detected or inactivity timeout interval reached
<b>Preconditions</b>	
<b>Postconditions</b>	<ol style="list-style-type: none"> <li>1. One or more gateway acknowledgements are received</li> <li>2. Location determined</li> </ol>
<b>Main flow action steps</b>	<ol style="list-style-type: none"> <li>1. Detect motion or broadcast inactivity timeout interval reached</li> <li>2. Broadcast message</li> </ol>
<b>Exceptions</b>	<ol style="list-style-type: none"> <li>1. Timeout reached</li> <li>2. Create a new message and broadcast the message</li> </ol>

The use case diagram depicted in figure 4.2 shows the use cases of the node. A node performs the following: detect when a movement has occurred, detect when an inactivity time interval has been reached, broadcast a message and receive an acknowledgement. “Detect motion” and “detect when a time interval is reached” both include the “broadcast message” use case.



**Figure 4.2: Node use case diagram**

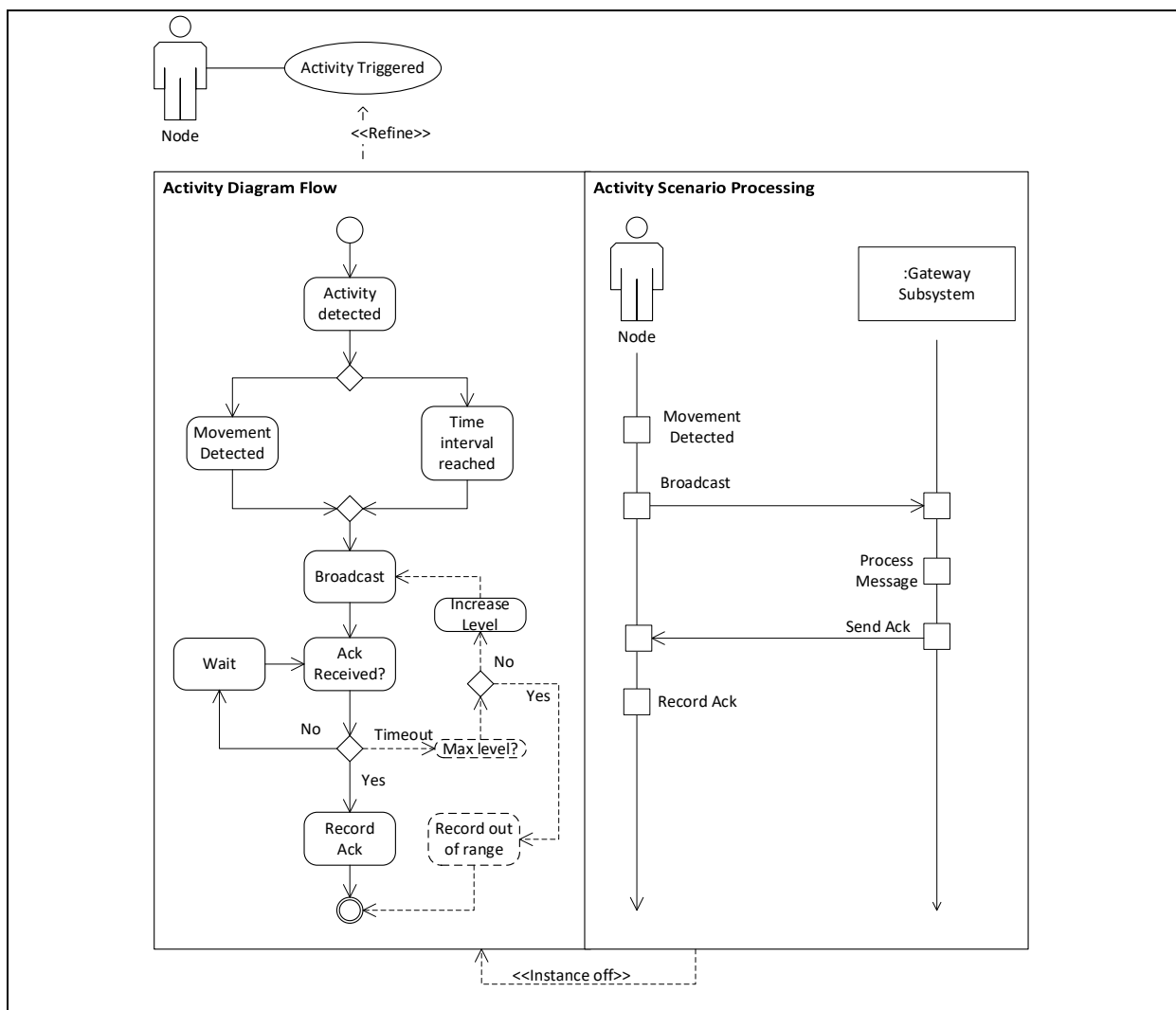
The use case diagram depicted in figure 4.3 shows the use cases of the gateway. These gateways provide the communication between the nodes and the server. A gateway performs the following: receive a message, send acknowledgement of receipt of a message and forward a message to the server. The “receive message” use case includes the “forward message to the server” and “send acknowledgement” use cases. The “detect time to send a message” use case includes the “broadcast message” use case. The gateway also generates its own message based on the time interval to transmit being reached. This is used to record and maintain the RSSI levels received from the gateways. Since the gateways are permanently fixed at a known location, any variation means that there is a change in the environment. The gateway will also be used to send messages, such as configuration, from the server to the node.



**Figure 4.3: Gateway use case diagram**

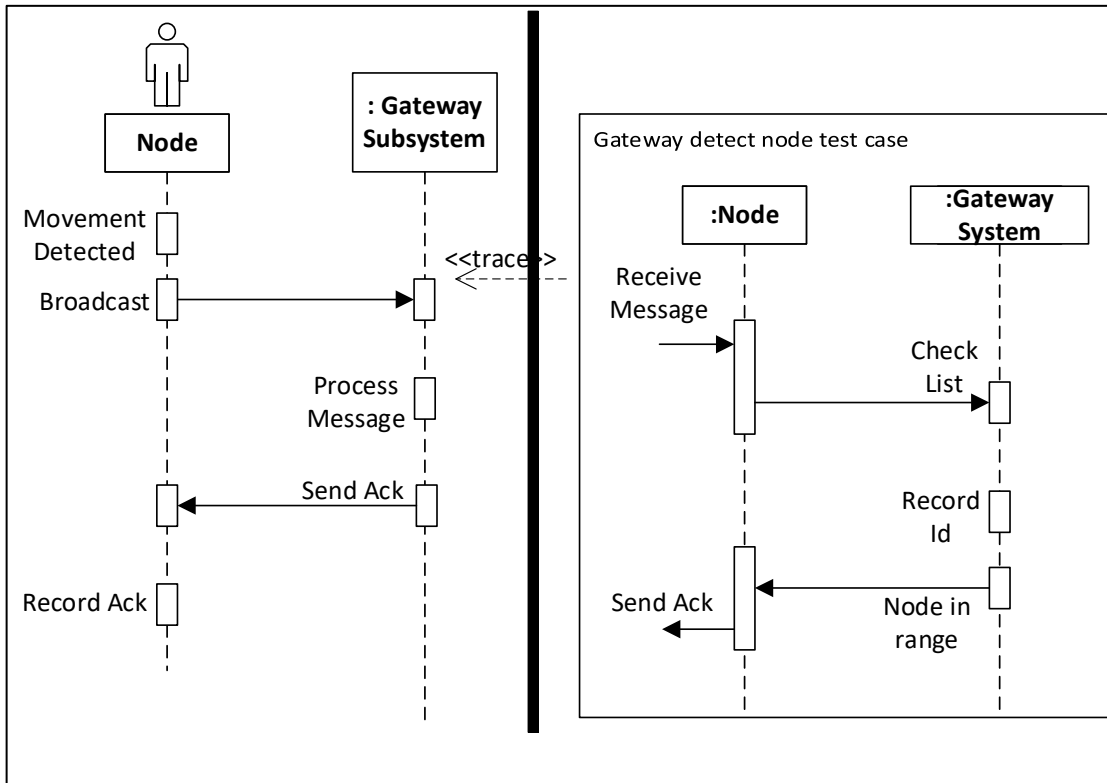
## 4.2.2 Analysis

Figure 4.4 is an activity diagram illustrating the procedure followed by a node when activity is detected. The transmit RSSI level is set to its minimum value, a message is formulated and transmitted, and the node awaits a response from the nearby gateways. When an acknowledgement is received within a predetermined timeout period, the details of the gateway are logged with the acknowledgement. If no acknowledgement is received after a predefined timeout interval, the transmit RSSI level is increased and a new message is broadcast, as indicated by the dashed lines. If no acknowledgement is received after the maximum transmission level has been achieved, no further transmissions will occur. The design ensures that the nodes start broadcasting at the lowest level and then increase to conserve battery power. The activity scenario processing of a successful detection of a neighbouring gateway is depicted as well. The neighbouring gateway is responsible for receiving the message, processing the message and sending an acknowledgement.



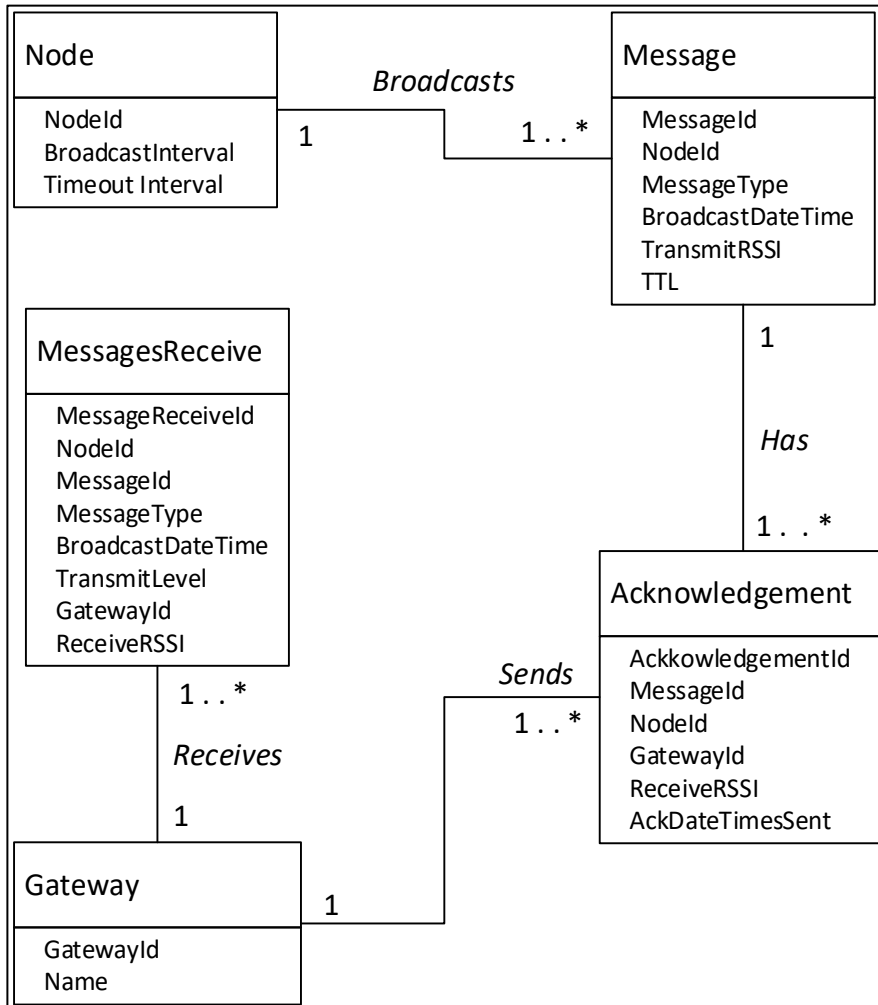
**Figure 4.4: Activity diagram and scenario processing**

Communication between the node and the gateway is achieved via messaging. Figure 4.5 shows the sequence of how a message is initiated, sent from a node, received, and processed by the gateway and an acknowledgement returned.



**Figure 4.5: Node and gateway messaging**

The domain model diagram shown in figure 4.6 shows a subset of the key domain classes, which are the node, message, messages received, gateway and acknowledgement with their attributes. The node broadcasts one or more messages, whilst the gateway receives these messages and sends out the acknowledgements. The received RSSI level is recorded by each gateway, and this is sent to the server for processing. The received RSSI level recorded at the gateway is processed to determine the location of the node.



**Figure 4.6: Subset domain model diagram**

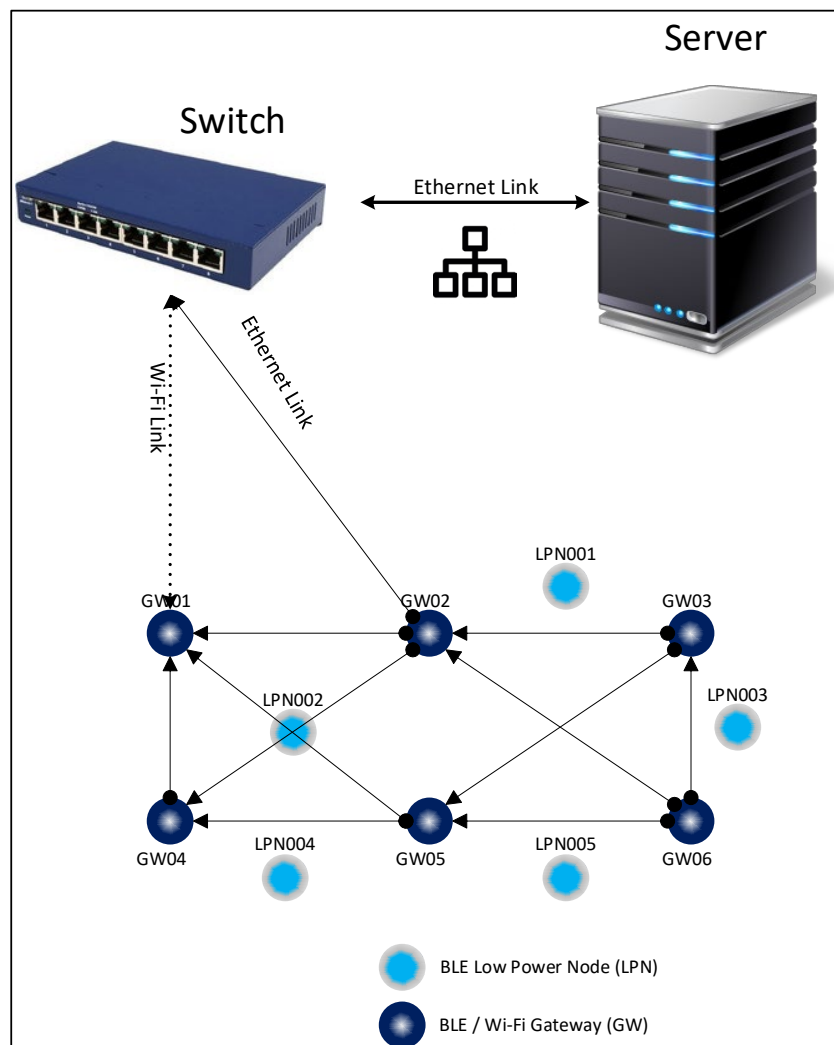
### 4.2.3 Design

The design of the layout of the gateways forms a mesh. This ensures that the entire indoor environment is adequately covered and that sufficient data is captured by the gateways to determine the location. Multiple gateways are positioned to capture the node’s transmissions so as to enable trilateration for location determination. The design of the node is a BLE device that is powered by a small-form-factor coin cell battery that can last more than a year, depending on usage. Therefore, this is labelled as an LPN.

The gateways may even be powered by a battery, electricity or power over Ethernet, enabling mobility, flexibility and consideration of power infrastructure installation. The LPN interfaces with gateways through Bluetooth, and gateways are able to connect to a network via Ethernet or Wi-Fi. The gateways connect over Bluetooth and eventually link back to the switch.

Gateways form a mesh via multiple paths to establish redundancy in the event that a gateway fails. This high-level architecture is represented in figure 4.7. Using this high-level architecture, an LPN is connected to a server via multiple paths.

The gateway houses the complexity of the communication and redundancy, decreasing the complexity of the BLE tag and lowering its power consumption and cost. The gateways are installed in fixed places, allowing for substantially larger batteries and/or connection to an external power source. The BLE tag establishes a connection with the closest gateway, from which their location can be determined using a combination of techniques. The needed precision of the location will vary depending on the application. Therefore, the described arrangement permits a customisable implementation based on the desired precision.



**Figure 4.7: High-level BLE mesh architecture**

The communication between the LPNs and the gateways is defined by use case, as described in tables 4.2 to 4.6. Only the main use cases are defined, together with a high-level process flow of the actions performed for each use case. Additional use cases are included in appendix A.

**Table 4.2: Use case LPN Detect Motion**

<b>Title</b>	<b>Description</b>
<b>Use case no.</b>	UC02
<b>Name</b>	LPN Detect Motion
<b>Related requirements</b>	Broadcast Message
<b>Goal in context</b>	LPN detects motion and broadcasts a message
<b>Trigger</b>	LPN has moved position
<b>Postconditions</b>	Message broadcast
<b>Main flow action steps</b>	<ol style="list-style-type: none"> <li>1. Formulate message and broadcast</li> <li>2. Initialise timer to await acknowledgement</li> </ol>
<b>Exceptions</b>	

**Table 4.3: Use case LPN Detect Inactivity Timeout Interval**

<b>Title</b>	<b>Description</b>
<b>Use case no.</b>	UC03
<b>Name</b>	LPN Detect Inactivity Timeout Interval
<b>Related requirements</b>	Broadcast Message
<b>Goal in context</b>	LPN detects time interval elapsed and broadcasts a message
<b>Trigger</b>	Inactivity timeout interval has been reached
<b>Postconditions</b>	Message broadcast
<b>Main flow action steps</b>	<ol style="list-style-type: none"> <li>1. Formulate message and broadcast</li> <li>2. Initialise timer to await acknowledgement</li> </ol>
<b>Exceptions</b>	

**Table 4.4: Use case LPN Acknowledge Receipt**

<b>Title</b>	<b>Description</b>
<b>Use case no.</b>	UC04
<b>Name</b>	LPN Acknowledge Receipt
<b>Related requirements</b>	Broadcast Message
<b>Goal in context</b>	LPN records acknowledgement of receipt of a message that was broadcast
<b>Trigger</b>	Message was broadcast
<b>Postconditions</b>	Message was acknowledged
<b>Main flow action steps</b>	<ol style="list-style-type: none"> <li>1. Acknowledgement received</li> <li>2. Acknowledgement recorded</li> <li>3. Cancel acknowledgement timer</li> </ol>
<b>Exceptions</b>	<ol style="list-style-type: none"> <li>1. Timeout reached as no acknowledgement received</li> <li>2. Broadcast new message</li> </ol>

**Table 4.5: Use case Gateway Received Message**

<b>Title</b>	<b>Description</b>
<b>Use case no.</b>	UC05
<b>Name</b>	Gateway Received Message
<b>Related requirements</b>	LPN broadcasts message
<b>Goal in context</b>	Send message received from LPN to the server
<b>Trigger</b>	
<b>Postconditions</b>	Message sent to server
<b>Main flow action steps</b>	<ol style="list-style-type: none"> <li>1. Receive message</li> <li>2. Formulate message to be sent to server</li> <li>3. Send message to server</li> <li>4. Send acknowledgement to LPN</li> </ol>
<b>Exceptions</b>	

**Table 4.6: Use case Gateway Detect Time Interval**

Title	Description
<b>Use case no.</b>	UC06
<b>Name</b>	Gateway Detect Heartbeat Interval
<b>Related requirements</b>	
<b>Goal in context</b>	Gateway broadcasts a message when heartbeat interval is reached
<b>Trigger</b>	Time heartbeat interval is reached
<b>Postconditions</b>	Message is broadcast
<b>Main flow action steps</b>	<ol style="list-style-type: none"> <li>1. Formulate message</li> <li>2. Broadcast message</li> <li>3. Reset time out interval</li> </ol>
<b>Exceptions</b>	

Using a domain model, the configurable parameters indicated in table 4.7 are accommodated for the LPN. These need to be set up on the server and be broadcast to the LPNs. The heartbeat interval defines the frequency of transmission of a message. This heartbeat interval is used to set the inactivity timeout interval. The lower the heartbeat, the lower the power utilisation. The transmit level can also be defined. This is set up to the lowest level but will depend on the environment. Setting the level to the lowest level further conserves power, extending the battery life. The broadcast mode can be used to disable and enable the broadcasting of messages.

**Table 4.7: Configurable parameters for LPN**

No	Parameter	Description
1.	Unique name	Unique name of LPN
2.	Heartbeat interval	The advertising frequency of a heartbeat of an LPN
3.	Transmit level	Set the transmit level to broadcast
4.	Broadcast mode	On – Reset counter and start broadcasting Off – Stop broadcasting

The key attributes required for location determination of an LPN are listed in table 4.8. The sender ID and the message ID are used to identify the unique sender and the unique message broadcast. The message type can be one of two types, depending on the trigger that initiates the message: a message triggered by the heartbeat interval reached will have the MessageMode set to LNT, and a message triggered by a motion detected will have the MessageMode set to LNM. Using this data, the application can establish that the LPN has moved its location. The transmitted date and time are sent so that this can be used if it is required for location determination, depending on the algorithm selected.

**Table 4.8: Configurable parameters for LPN**

No	Field	Description
1.	Sender ID	Sender ID
2.	MessageID	Unique message ID – can be reset on power-up of node
3.	MessageType	LNT – location message from LPN triggered by time
4.	MACAddress	ID of sender LPN MAC address – if Unique is used, this can be omitted to shorten the message
5.	TxLevel	Transmit level for LPN
6.	TxDateTime	Date and time of sending (optional)

Table 4.9 defines the key attributes that are required to be saved when a message is received at a gateway. This information will be used for location determination. Therefore, the LPN and message identification are important to determine the source of the message, and these, together with the recorded received RSSI level, are used to determine the location. All messages received at the gateway are sent to the server for storage and calculation of the location. The receiving gateway is also recorded so that it can be used in the process of trilateration.

**Table 4.9: Receive Message definition at gateway**

No	Field	Description
1.	Sender ID	Sender ID
2.	MessageID	Unique message ID from sender
3.	MessageType	1 – location message from LPN triggered by time
4.	MACAddress	ID of sender LPN MAC address – if Unique is used, this can be omitted to shorten the message
5.	TxLevel	Transmit level for LPN
6.	TxDateTime	Date and time of sending (optional)
7.	ReceiveID	Unique ID of gateway receiving message
8.	RxDateTime	Date and time of receiving the message
9.	RxLevel	Level received

Data exchanged between the LPN, the gateway and the server are in JavaScript Object Notation format for ease of processing and readability. All messages are then processed and stored in a PostgreSQL database.

Message definition sent from an LPN:

```
{  
  "txDate":  
  "messageType": "",  
  "senderId": ,  
  "txLevel": ,  
  "txXPos": ' ,  
  "txYPos": ' ,  
  "txVenue": "",  
  "receiverID": ,
```

```

    "rxDate":,
    "rxLevel":,
    "rxXPos":',
    "rxYPos":',
    "TTL":
}

```

Configuration messages are sent from the server to the gateway and/or LPN. The message types are "CSL" or "CSR". The following is a message definition:

```

{
    "date":,
    "messageType":"","
    "parameters":
        {"parameter":"","
    "parameterValue":,}
    "destination":
}

```

Sample message to LPN:

```

{
    "date": "2019/07/10 21:26:15.1486",
    "messageType": "CSL",
    "parameters":
        {"parameter":"TxLevel",
    "parameterValue":-4,
    "parameter":"X",
    "parameterValue":1,
    "parameter":"Y",

```

```
"parameterValue":2,  
"parameter":"Venue",  
"parameterValue":"Lab1",  
"parameter":"Broadcast",  
"parameterValue":"On"},  
"destination": "ab11fbe9"  
}
```

Sample message sent to a gateway:

```
{  
  "date": "2019/07/10 21:26:15.1486",  
  "messageType": "CSR",  
  "parameters":  
    {"parameter":"TxLevel",  
     "parameterValue":-8,  
     "parameter":"X",  
     "parameterValue":5,  
     "parameter":"Y",  
     "parameterValue":5,  
     "parameter":"Venue",  
     "parameterValue":"Lab1", }  
    "destination": "a6eaa1f4"  
}
```

The following key test cases are presented using the main use cases for the LPN and the gateway to ensure the methodology followed addressed all the possibilities. These make provision for the scenarios, the results and the exceptions that may arise.

**Table 4.10: Main test case list**

No	UC no.	Use case	Test scenario	Result	Exception
1.	UC01	Determine Location	Determine location of node within range	Location determined	
2.	UC01	Determine Location	Determine location of node that is out of range		Unable to determine location
3.	UC02	LPN Detect Motion	Motion detected and a message is broadcast	Broadcast message	
4.	UC02	LPN Detect Motion	Motion detected and a message is not broadcast		Message not broadcast
5.	UC03	LPN Detect Inactivity Timeout Interval	Inactivity timeout and a message is broadcast	Broadcast message	
6.	UC03	LPN Detect Inactivity Timeout Interval	Inactivity timeout and a message is not broadcast		Message not broadcast
7.	UC04	LPN Acknowledge Receipt	Acknowledgement received for message sent	Message received	
8.	UC04	LPN Acknowledge Receipt	Acknowledgement not received for message broadcast		Acknowledgement not received
9.	UC05	Gateway Detect Heartbeat Interval	Heartbeat interval reached and a message is broadcast	Broadcast message	
10.	UC05	Gateway Detect Heartbeat Interval	Heartbeat interval reached and a message is broadcast		Message not broadcast

The requirements for the design model are to determine the throughput, range and transmission through obstructions in order to develop a model that is feasible within the identified constraints. The range and throughput are required so that the model can scale to cover a feasible, practical area. Also, gateways at fixed positions need to update the footprint in a dynamic environment so that the link indicator between the gateways can be used for location determination of nodes in the vicinity. The link indicator is used to determine whether there is any change in the path of transmission between gateways. This change is based on an initial footprint of RSSI received values recorded with known locations of the gateways and distance between them.

## **4.3 Experiments**

### **4.3.1 Throughput and range determination**

Developers are primarily familiar with Generic ATtribute Profile, which describes the BLE data transfer protocol. Attributes identify application data, and groups of attributes comprise characteristics that provide permission and interaction rules to unique data. Initial tests were conducted with the PDK boards from Nordic. These were configured as a transmitter node (tester) and a receiver (central). The transmitter node commenced advertising once the test was started. The elapsed time and the throughput were measured and recorded. The detailed interaction between the transmitter and the receiver is recorded in appendix B.

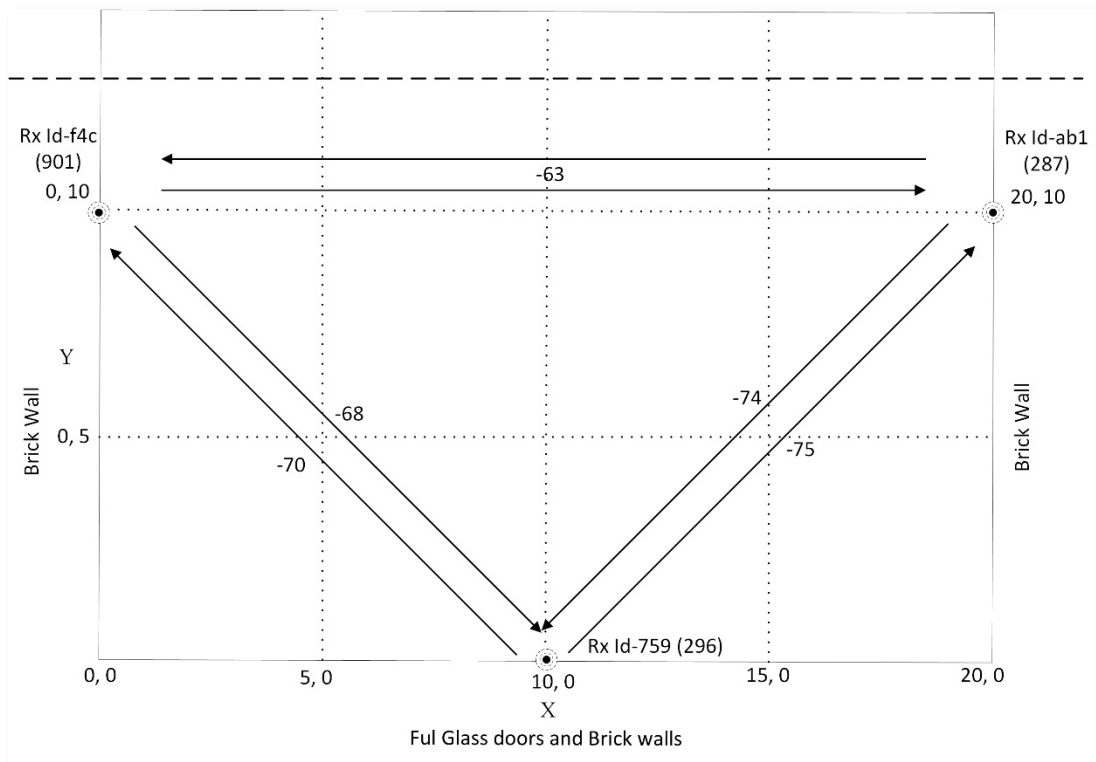
The transmit level was set to -20 dBm and the packet size was configured to 247 bytes. The received RSSI level, time lapse and throughput were measured at a distance of 0 m to 4 m. Then the average RSSI, time lapse and throughput for each distance were calculated, as shown in table 4.11. It can be observed that as the distance increases, the average time lapse for the packet to reach the receiver generally increases. However, the throughput decreases gradually as the distance increases.

**Table 4.11: Throughput**

	<b>Tx Power (dBm)</b>	<b>Distance (m)</b>	<b>RSSI</b>	<b>Avg. RSSI</b>	<b>Time lapse (sec)</b>	<b>Avg. time lapse (sec)</b>	<b>Throughput (Kbps)</b>	<b>Avg. throughput (Kbps)</b>
<b>1.</b>	-20	0	-35	-36.2	8.7	8.5	1040.0	1036.8
<b>2.</b>	-20	0	-39		8.1		1032.2	
<b>3.</b>	-20	0	-33		8.7		1039.6	
<b>4.</b>	-20	0	-35		8.1		1035.0	
<b>5.</b>	-20	0	-39		8.9		1037.2	
<b>6.</b>	-20	1	-81	-78.2	12.3	12.2	680.1	696.6
<b>7.</b>	-20	1	-81		11.4		734.1	
<b>8.</b>	-20	1	-75		11.2		748.9	
<b>9.</b>	-20	1	-74		12.8		694.7	
<b>10.</b>	-20	1	-80		13.4		625.3	
<b>11.</b>	-20	2	-81	-82.4	15.1	14.9	559.0	569.4
<b>12.</b>	-20	2	-82		14.7		572.2	
<b>13.</b>	-20	2	-81		15.7		556.7	
<b>14.</b>	-20	2	-87		14.4		584.3	
<b>15.</b>	-20	2	-81		14.6		574.8	
<b>16.</b>	-20	3	-87	-86	33.6	32.7	249.4	266.4
<b>17.</b>	-20	3	-83		44.3		189.6	
<b>18.</b>	-20	3	-87		24.7		340.2	
<b>19.</b>	-20	3	-86		31.4		267.5	
<b>20.</b>	-20	3	-87		29.4		285.4	
<b>21.</b>	-20	4	-89	-85.6	16.7	17.0	501.5	495.0
<b>22.</b>	-20	4	-85		17.5		480.7	
<b>23.</b>	-20	4	-86		18.5		454.4	
<b>24.</b>	-20	4	-84		15.7		536.0	
<b>25.</b>	-20	4	-84		16.7		502.1	

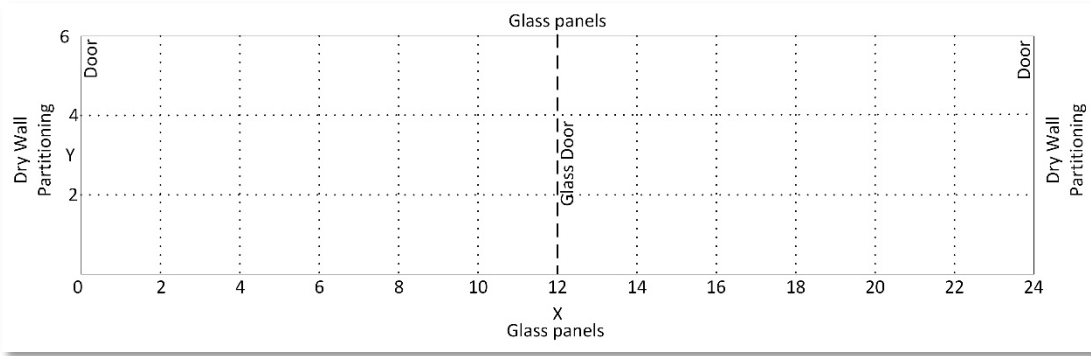
After the initial testing of throughput in open space, further experiments were conducted at Durban University of Technology, South Africa. The locations utilised were an open plan office, a boardroom and an office, as depicted in figures 4.8, 4.9 and figure 4.10, respectively. These locations simulate the typical real-world indoor environment of an office environment.

Figure 4.8 depicts the floor plan of the open-plan office, which was 25 m<sup>2</sup> in size and comprised brick walls and glass doors. The (x, y) distances shown in the diagram represent the distance in metres from the (0, 0) reference point at the bottom left of the diagram. The gateways (GWxx) and nodes were positioned at (0, 10), (10, 0) and (20,10) respectively. The gateways received the messages that each node transmitted.



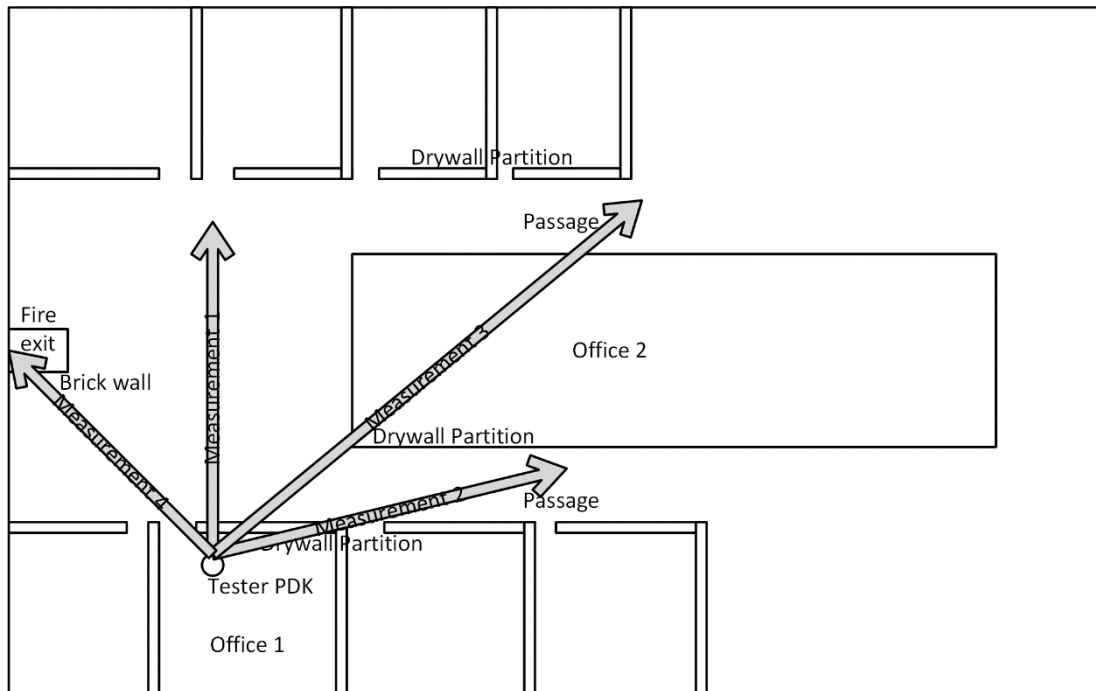
**Figure 4.8: Open-plan office**

The boardroom used was 24 m × 6 m and was populated with office desks and chairs. The walls were constructed of glass and drywall partitioning.



**Figure 4.9: Boardroom**

The office spaces were separated with drywall partitioning through which transmission was tested. The floor plan layout is represented in figure 4.10. The received RSSI levels were measured from the transmit point to four different locations: measurement 1 through one partition in an open space area, measurement 2 through one partition in a passage area, measurement 3 through three partitions and measurement 4 through one partition and a brick wall.

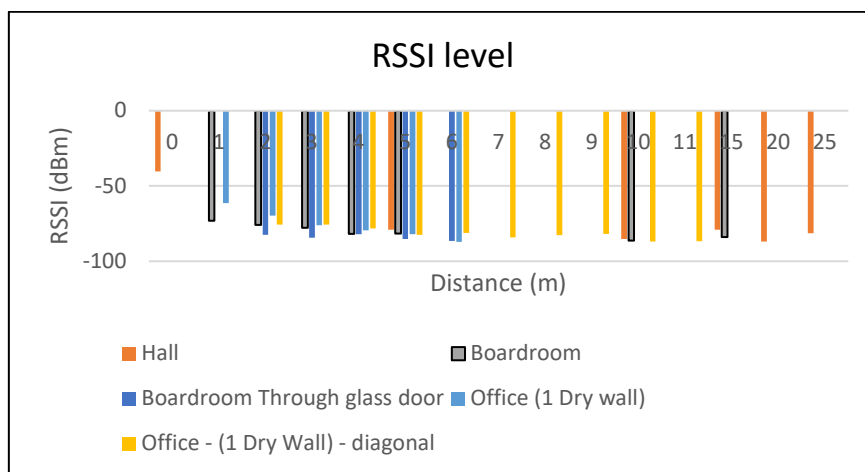


**Figure 4.10: Partitioned office**

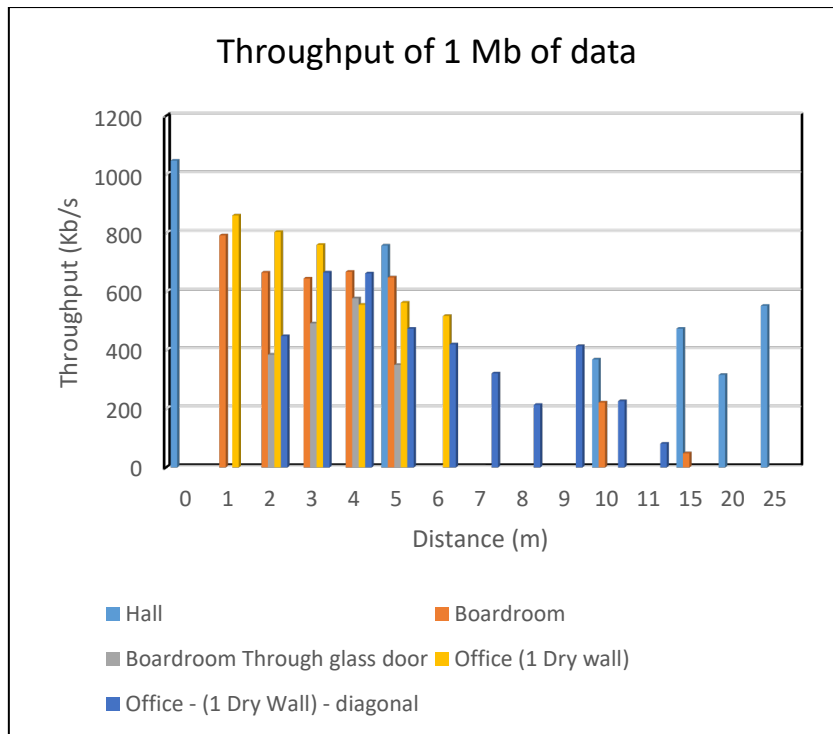
The data was gathered in the open office, the boardroom and the partitioned office. Figures 4.11 and 4.12 illustrate the measured RSSI values and throughput at these locations. Due to space

constraints, data for the numerous attributes of the various venues have been grouped for display on the graphs. The different hues in the legend symbolise the many venues and their respective obstructions. In the open-plan office, measurements were taken at 5-meter intervals until transmission, connectivity and data became unreliable. When the PDK boards were positioned at different angles, the measurements were not significantly different. Consequently, these additional measurements were excluded from the study.

In the boardroom, measurements were made at 1 m intervals from 1 to 5 m, and then at 5 m intervals up to the maximum distance at which the equipment was usable. Subsequently, an obstruction in the form of a glass door was added to divide the boardroom into two sections. For the initial measurement, the PDK boards were placed 1 m from either side of the glass obstruction. The test PDK was then relocated at 1 m intervals up to a distance of 6 m. The boards could not communicate via the glass door at 6.5 m. In the office context, a high measurement throughput was obtained. For measurement 2, good throughput was also seen up to 11 m, but beyond this distance, connectivity was very poor. The transmission distance for measurement 3 was approximately 13 m. At a distance of 5 m, measurement 4 yielded a good throughput, but after that, connectivity was very poor due to the loss of signal strength through the solid brick wall.



**Figure 4.11: RSSI levels**

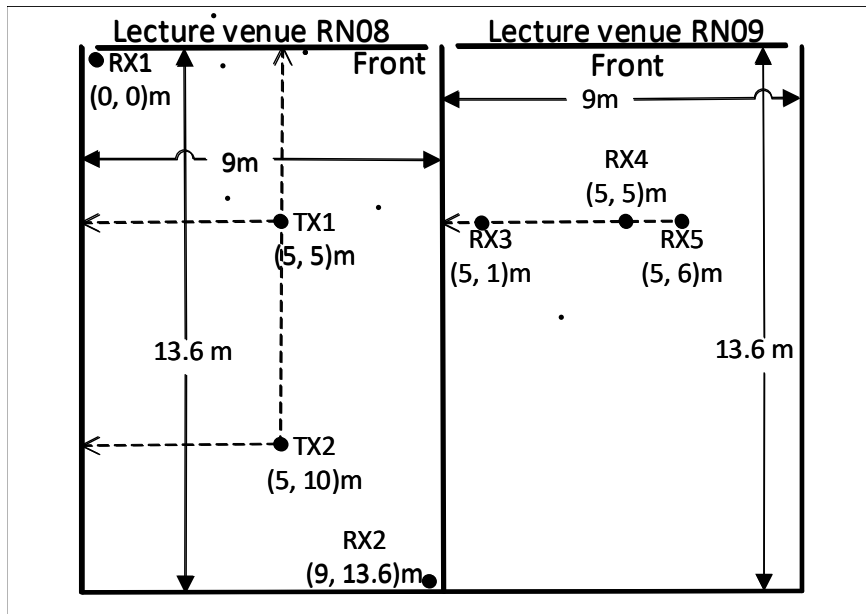


**Figure 4.12: Throughput of 1 Mb of data**

### 4.3.2 Transmission through obstructions

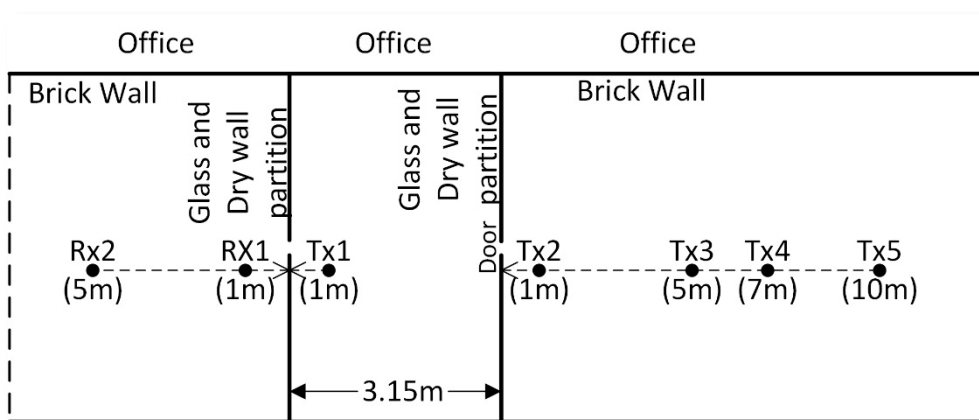
The two different venues were used to measure the effect of obstructions on BLE transmission. A pair of adjacent offices and a pair of meeting rooms were used for these experiments. Figure 4.13 displays the initial venue, which consisted of two identical offices, named RN08 and RN09, that were adjacent to one another. They were built with strong brick walls measuring approximately 21 cm thick and are 9 m by 13.6 m in size.

The locations in figure 4.13 are labelled with Transmit (TX) or Receive (RX) and include the coordinates from the front left corner of the room. For instance, “RX1, (0, 0) m” indicates that the receiver is located in the front left corner, but “TX2, (5, 10) m” indicates that the transmitter is located 5 m away from the left wall and 10 m away from the front wall. For the first test, the two receivers were placed diagonally opposite each other at the corners of venue RN08 (points RX1 and RX2), while the broadcaster was located at points TX1 and TX2 in the same venue (figure 4.13). In the adjacent office, the transmitter and receiver were placed on the fixed desks. For the second test, one of the receivers was positioned at RX1 in venue RN08, while the other was positioned at RX3, RX4, and RX5 in venue RN09. The broadcaster was positioned at TX1 in venue RN08. At the measured sites and across the brick wall obstruction, data was collected.



**Figure 4.13: Adjacent offices**

This second venue consisted of meeting rooms separated by an office with walls that permitted testing through double drywall partitions as well as double glass panels. Figure 4.14 depicts the venue's floor plan arrangement. The wall was constructed of drywall partition from the floor to around midway, and then glass from midway to the ceiling. The receiver was situated at sites RX1 and RX2, which were 1 m and 5 m away from one of the walls, respectively. The transmitter was positioned at TX1, which was 1 m across the first partition, and at TX2 through TX5, which were 1, 5, 7, and 10 m apart from the second partition. These transmit and receive positions in the two labs, along with the office in between, are depicted in figure 4.14. These placements were utilised to collect data across two drywall partitions as well as two glass panels.



**Figure 4.14: Meeting room**

In this experiment, the transmitted power level was the independent variable. The power was adjusted from 4 dBm to -20 dBm to establish the usable ranges through the different obstructions. Range and obstructions were the dependent variables, which led to varied RSSI levels at the receiver. Using these variables, the following experimental stages were conducted:

**Stage 1:** Measure and identify the various predefined distances.

**Stage 2:** Deploy software on the PDKs.

**Stage 3:** Position the PDK receiver in a predetermined starting location.

**Stage 4:** Place the broadcaster PDK at the first measured point.

**Stage 5:** Start measuring the received RSSI value. If the lowest transmit level is reached or no broadcasted packets are received, proceed to step 7.

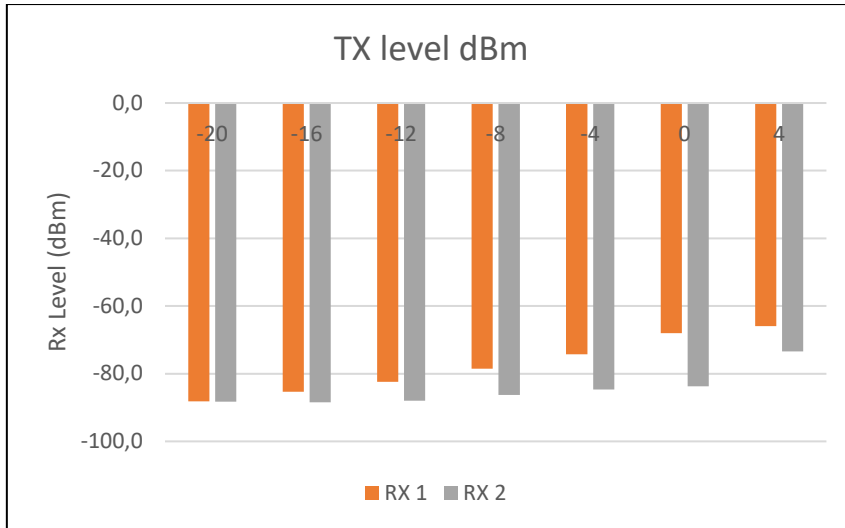
**Stage 6:** Reduce the broadcaster's transmit power and go to step 5.

**Stage 7:** Move the broadcaster PDK to the subsequent measured location, reset the transmit level to 4 dB, and repeat step 5.

**Stage 8:** Stop when the maximum distance is reached and readings are no longer useable, or when the PDKs are disconnected.

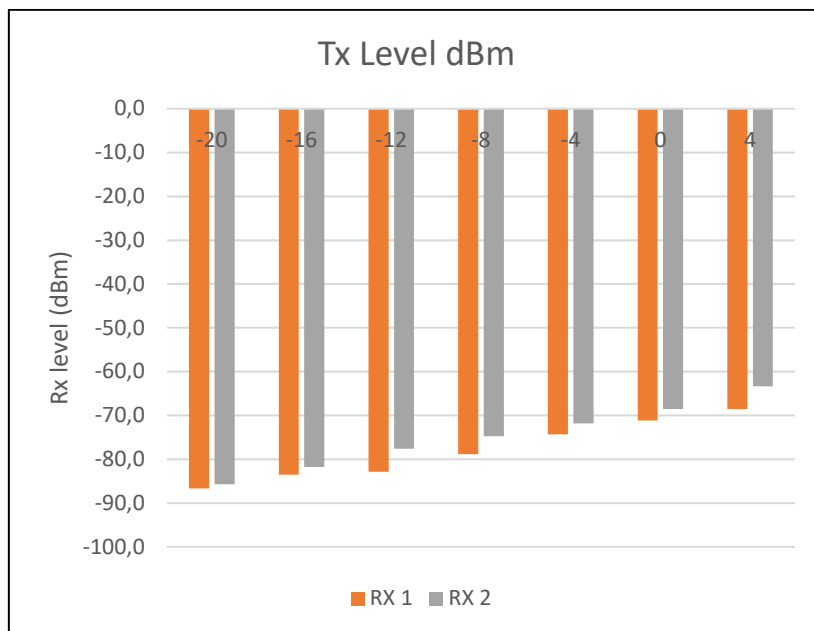
One PDK board was configured as the receiver and the other as the transmitter. The broadcaster's signal strength ranged from 4 to -20 dBm. This receiver PDK was connected to a laptop on which the terminal emulator Putty was used to access the USB interface data. To establish an average reading, a minimum of ten packets were received for each of the various levels and distances. The plan was to repeat the experiment multiple times if the results revealed significant variance. After data collection, the averages were computed and reported. In addition to collecting data in nearby offices, data was also collected from meeting rooms. Figures 4.15 to 4.22 depict the average RSSI levels received by the different impediments at the respective broadcast levels.

In the adjacent offices, measurements were taken until connectivity and transmission were deemed unsuitable and broadcast packets could no longer be received. When the PDK boards were positioned at different angles, the measurements did not differ significantly. As a result, more measures were omitted from the final analysis. Figure 4.15 depicts the data collected in office RN08. As the transmit power was reduced, the received signal level at the receiver decreased steadily. From transmit point TX1 in figure 4.13, the received level indicated by receiver 1 (RX1) is consistently higher than that provided by receiver 2, which was further from the broadcaster. The lowest level of broadcast packets received was -20 dBm.



**Figure 4.15: RSSI levels received from position TX1**

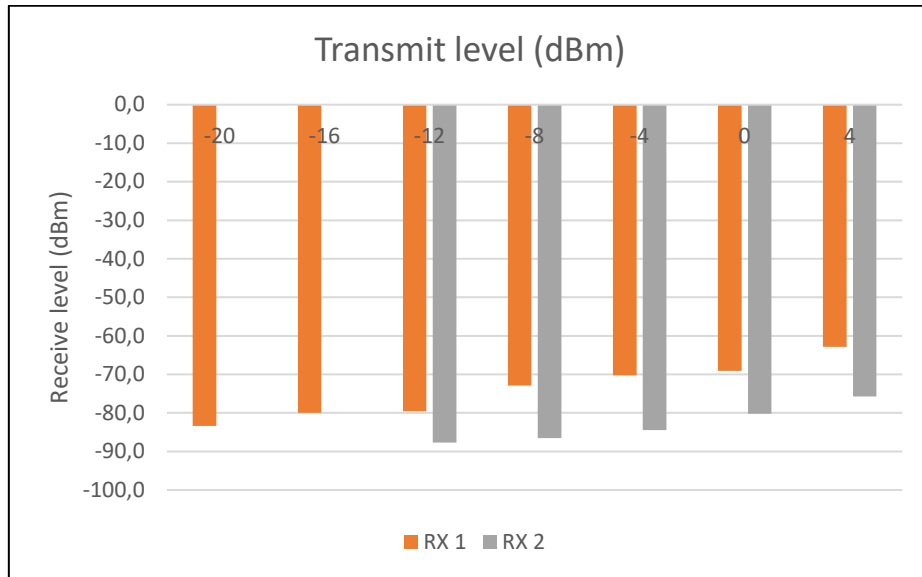
Figure 4.16 illustrates the results obtained after moving the transmitter to position TX2. Due to the larger distance between the two receivers, the lower levels measured by receiver 1 at position RX1 was expected. As the broadcast levels decreased, the receive levels of RX1 and RX2 also decreased steadily.



**Figure 4.16: RSSI levels received from position TX2**

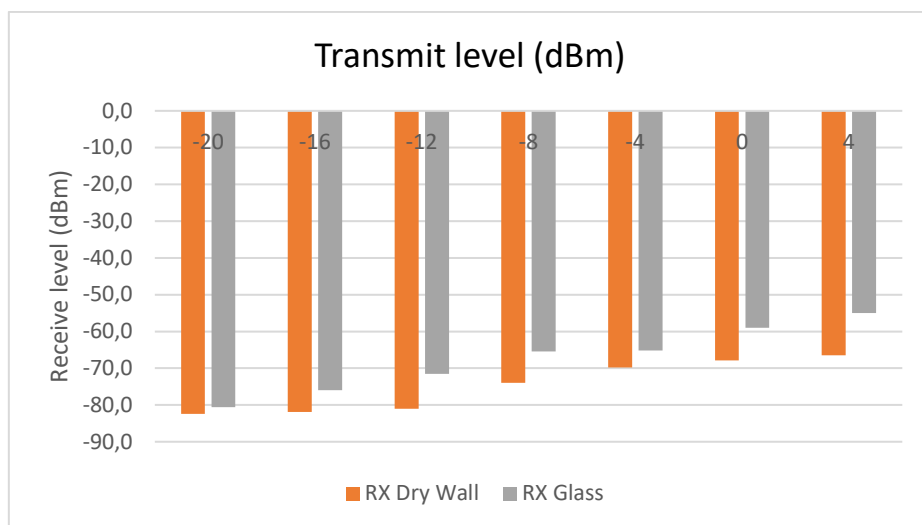
Figure 4.17 depicts the findings obtained between the two adjacent offices. The broadcast packets were received in office RN09 at position RX5, 5 m from the wall (as shown in figure

4.13), with no reception beyond this point. At point RX5, the receive level for a transmit level of 4 dBm was -75.7 dBm, while the receive level for a transmit level of -12 dBm was -87.7 dBm.

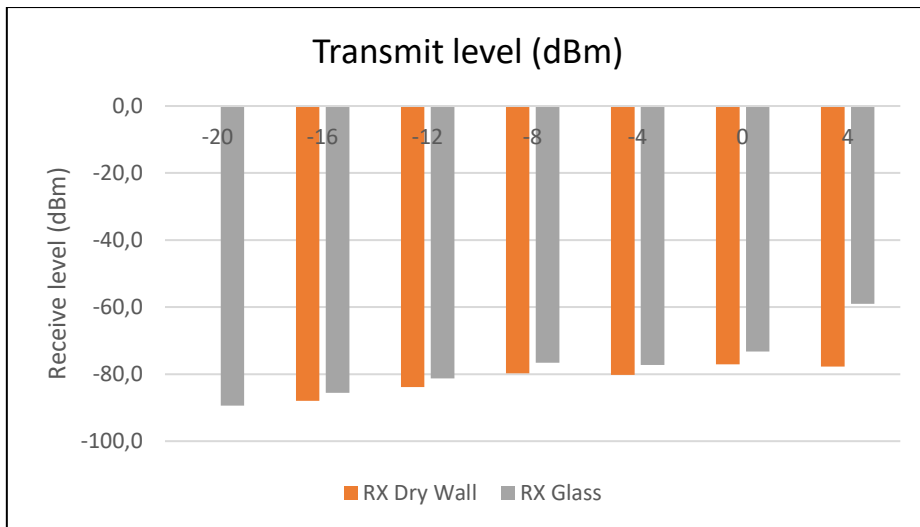


**Figure 4.17: RSSI levels between two adjacent offices**

Figure 4.18 depicts the initial set of measurements obtained between places TX1 and RX1 in the meeting rooms, which are separated by a single partition. The two forms of obstructions that were utilised were drywall partitions and glass, with their corresponding dimensions. The transmitter was subsequently relocated to location TX2 (see figure 4.14), creating two partitions between the transmitter and receiver. Figure 4.19 shows a decrease in receive levels observed for both glass and drywall partitions as the transmit level decreased.

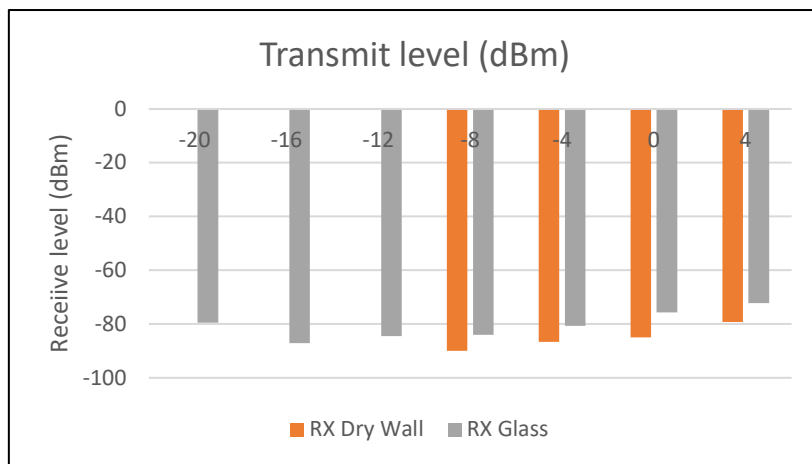


**Figure 4.18: RSSI levels between TX1 and RX1 across single partition**



**Figure 4.19: RSSI levels between TX2 and RX1 across double partition**

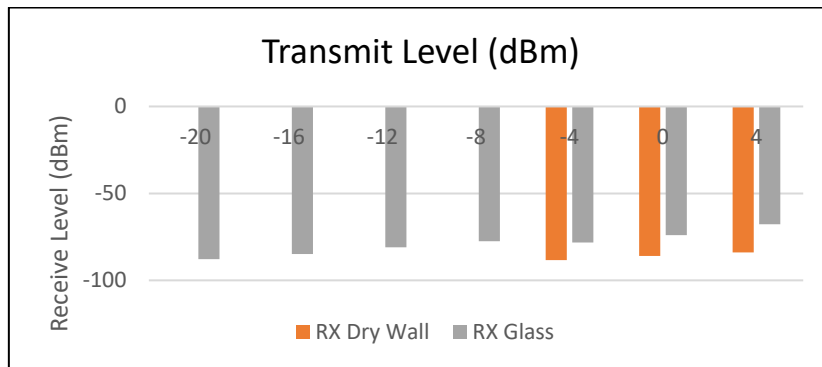
The receiver was subsequently relocated to position RX2, while the transmitter remained at station TX2. Figure 4.20 displays the results of measurements made across the double partitions. The received signal strength decreased from 4 dBm to -8 dBm when the transmit level was 4 dBm. The packets were not received via the drywall barriers when the transmit level was below -8 dB, but they were received through the glass partitions.



**Figure 4.20: RSSI levels between TX2 and RX2 double partition**

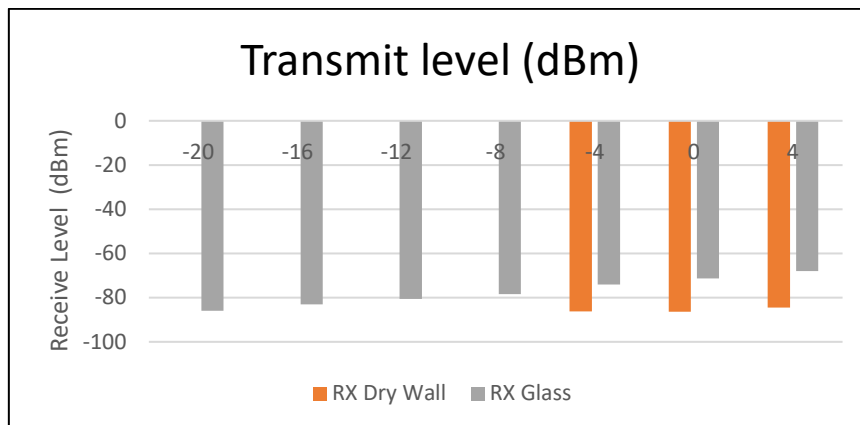
The transmitter was subsequently relocated to TX3, which was 5 m away from the partition. This configuration placed the transmitter and receiver 5 m away from the office partitions. The received levels depicted in figure 4.21 were lower than the transmit levels measured at point

TX2. The signal degraded further due to the drywall partitions, which prevented any level below -8 dB.



**Figure 4.21: RSSI levels between TX3 and RX2 across double partition**

The transmitter was repositioned so that TX5 was 10 m from the partition. Since no packets were received, the transmitter was moved closer to the partition until a signal strength sufficient for use was detected at point RX2. Position RX2 received packets at an acceptable level while the transmitter was located at TX4 and 7 m from the partition. High levels were observed through the drywall partitions for transmit levels between -4 dBm and 4 dBm. The received levels consistently decreased as the transmit level decreased. Figure 4.22 depicts transmission and reception levels.

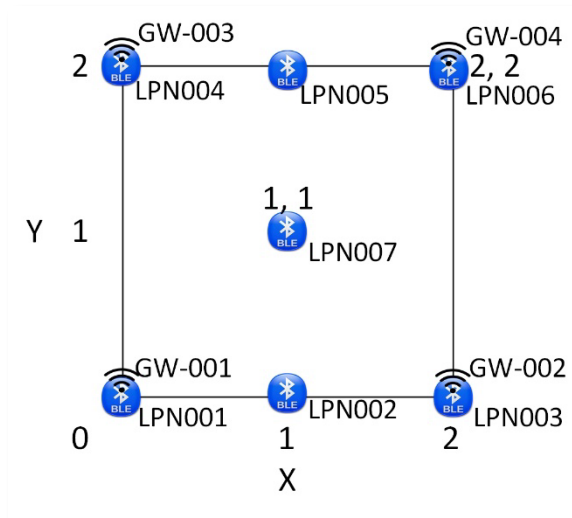


**Figure 4.22: RSSI levels between TX4 and RX2 across double partition**

### 4.3.3 Real-time quality link indicator

The LPN firmware was designed to communicate the unique LPN identifier, RSSI level and unique message identifier every 100 milliseconds. The message identification and RSSI level

were evaluated to establish how each gateway received each broadcast message. The gateways named GW-001 to GW-004 were placed at the corners of a 2 m<sup>2</sup> area, whereas the LPNs labelled LPN001 to LPN007 were placed as depicted in figure 4.23. To determine the initial signal loss, the LPNs 1, 3, 4 and 6 were placed within a centimetre of the gateways. To compare the RSSI received by the gateways, the LPNs 2, 5 and 7 were positioned at identical distances from them. Positions are designated as x, y, where x indicates the experimental area's length in metres and y represents its breadth in metres.



**Figure 4.23: LPN and gateway layout**

Once the LPNs, gateways and server had been configured, the LPNs began transmitting messages every 100 milliseconds. The gateways received these messages through Bluetooth and transformed them for Wi-Fi upload to the server. Sessions of 10 and 30 minutes were utilised as the time restriction for capturing the messages from the LPNs.

Using a value of 20 for n, formula 4.1 was used to compute the mean RSSI in order to find the mean. This meant that a window of 20 RSSI levels was used for each calculation. Utilising the principle of first in, first out, the RSSI stream was managed.

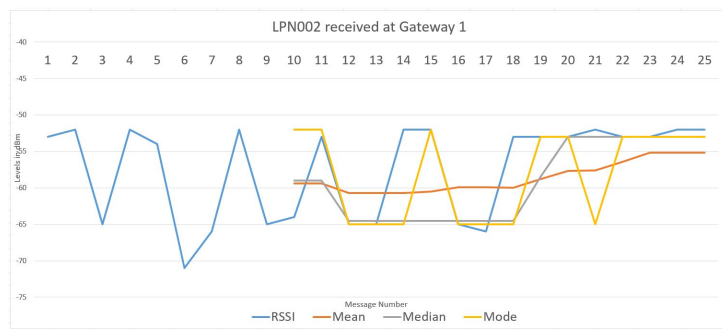
$$\text{Mean RSSI} = 1/n \sum_{i=1}^n x_i \quad 4.1$$

where

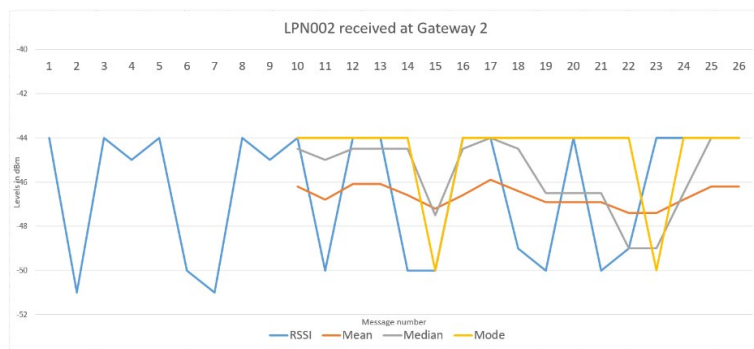
n is the number of RSSI levels and

x is the RSSI level.

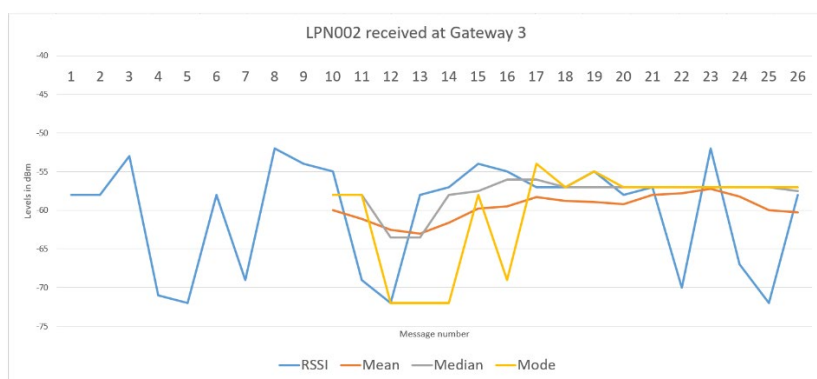
The median of a data set is the middle number when the data set is arranged from highest to lowest. The mode of a dataset indicates the most prevalent element. In addition to defining the mode, the frequency of each distinct RSSI occurrence was determined in this study. Analysing the frequency of occurrences of RSSI values helped uncover a possible pattern for use in subsequent calculations. More than 152 000 messages from LPNs were received. The actual RSSI received at the gateways, as well as the mean, median and mode derived from the RSSI, are depicted in figures 4.24 to 4.27. The RSSIs measured at the gateway and displayed in the graphs are used for illustrative reasons as identical results were obtained for the whole dataset.



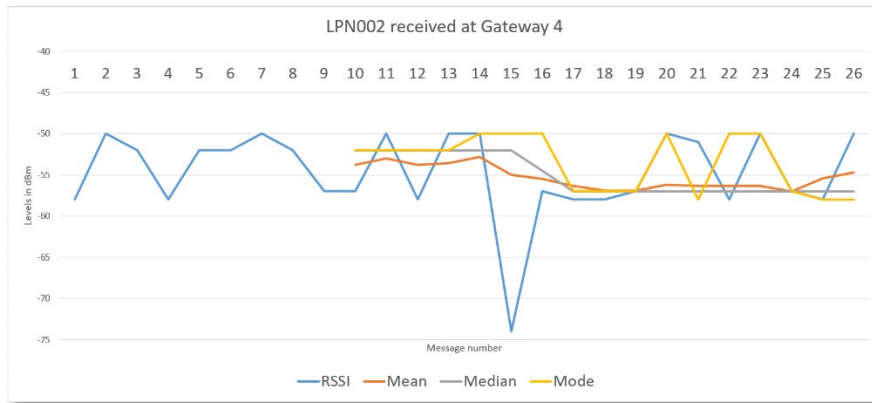
**Figure 4.24: LPN002 messages received at GW-001**



**Figure 4.25: LPN002 messages received at GW-002**

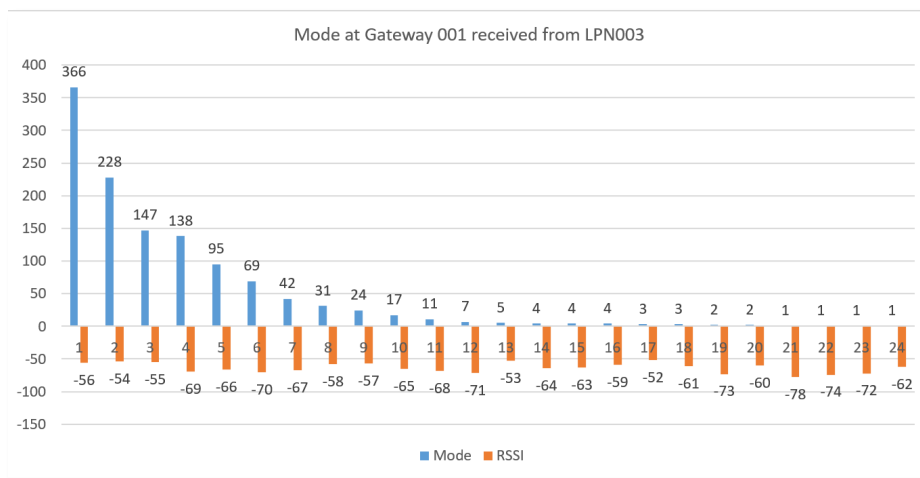


**Figure 4.26: LPN002 messages received at GW-003**



**Figure 4.27: LPN002 messages received at GW-004**

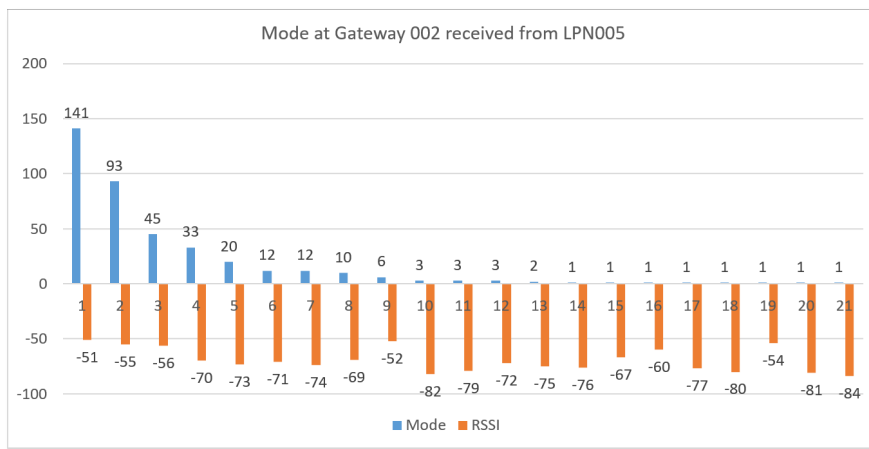
Figure 4.28 displays, for illustrative purposes, the mode of the RSSI received at GW-001 for LPN003 messages. In addition, the graph depicts the number of unique RSSI values, which indicates variation and distribution. This stability is evidenced by the fact that the number of messages received at levels -56 and -54 was 366 and 228, respectively, whilst the remaining levels had a significantly lower number.



**Figure 4.28: Mode at GW-001 from LPN002**

Figure 4.29 illustrates the mode of LPN005 messages received by GW-002. At level -51, 141 messages were received, and at level -55, 93 messages were received. Additionally, the graph depicts the number of unique RSSI values received. In both instances, the remaining messages had a substantially lower RSSI count than the first two. With the LPN and Gateways set in the positions depicted in figure 4.23, all RSSI levels were measured. This pattern was detected in

all of the received and recorded data. Thus, an example of this result is provided for illustrative purposes.

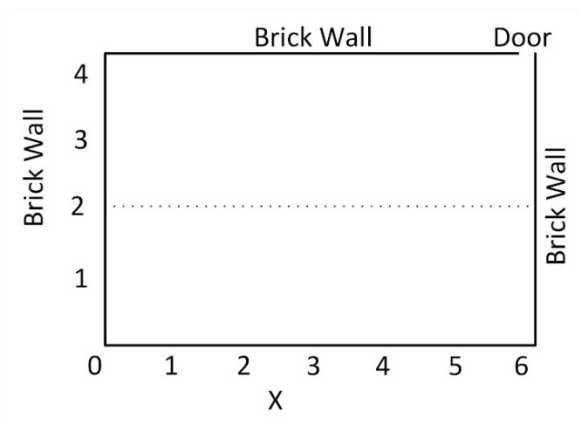


**Figure 4.29: Mode at GW-002 from LPN005**

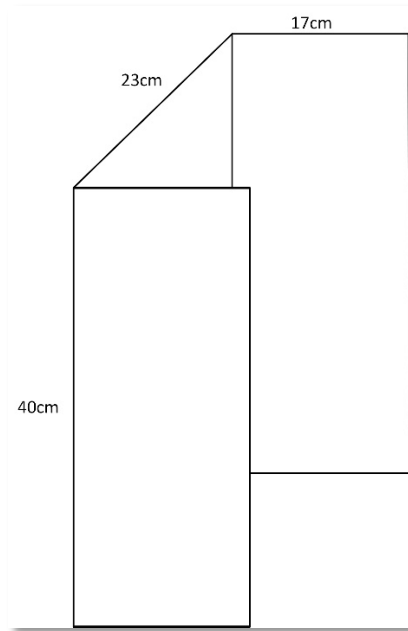
## 4.4 Effect of obstructions

### 4.4.1 Environment and obstructions

Testing through Perspex and metal obstructions was conducted in a small 4 m × 6 m office (see figure 4.30). Only office tables were present in the office, and the equipment was placed thereon. Obstructions used were 1.200 m × 0.6 m × 3 mm Perspex, 1.200 m × 0.6 m × 5 mm Perspex and a 1.200 m × 0.6 m × 0.5 mm metal sheet. A 40 cm × 23 cm × 17 cm piece of stainless-steel metal fabricated in the shape of a U (see figure 4.31) was also used as an obstruction.



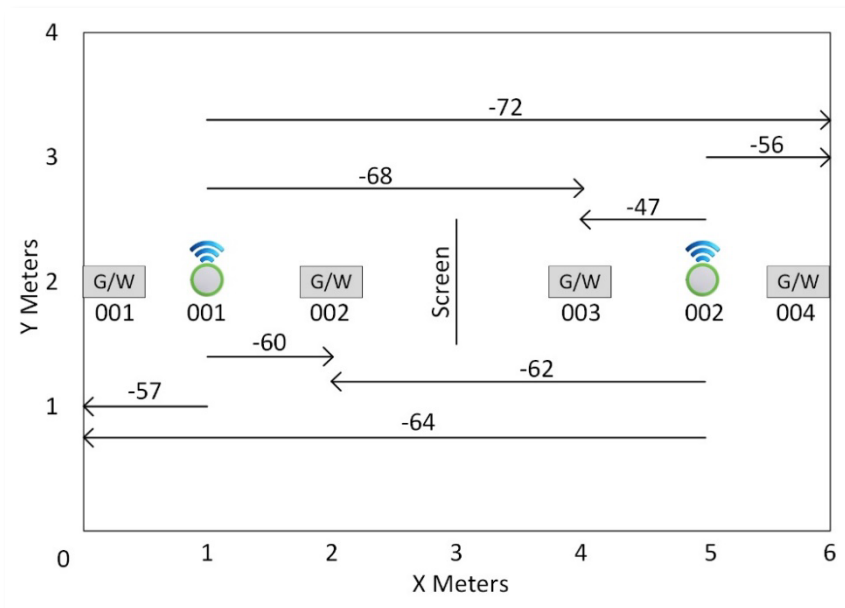
**Figure 4.30: Office**



**Figure 4.31: Fabricated metal obstruction**

### Experiment using Perspex

The first obstruction used was a 3 mm Perspex obstruction, then a 5 mm Perspex screen and finally both the 3 mm and 5 mm Perspex obstructions. The obstructions were placed at location (3, 2), where location is measured in metres from the bottom left of the figure, which indicates a corner of the room as a point of reference. RSSI levels were recorded by each of the four gateways GW-001 to GW-004 placed at locations (0, 0), (2, 2), (4, 2) and (6, 2) respectively. The LPNs were placed at locations (1, 2) and (5, 2).



**Figure 4.32: Received RSSI through Perspex**

The RSSI levels transmitted from the LPNs were recorded at each of the gateways. The recording was done for a 5-minute period and was iteratively tested thrice, viz. test 1, test 2 and test 3. The following process was followed to calculate average RSSI level:

For each iteration

Count the number of the distinct RSSI levels received

Select the top three RSSI levels that have the highest counts

Given the three highest counts of RSSI levels received per iteration, compute the average of these RSSI levels for all three iterations.

Table 4.12 below shows an extract of the data for gateway GW-001 and nodes LPN001 and LPN002. The full set of received levels are listed in appendix C. Using the process described above, the initial experiment contained no obstructions. Subsequent experiments using the same process had obstructions comprising 3 mm Perspex, 5 mm Perspex and a combination of the 3 mm and 5 mm Perspex placed in the path of transmission to observe the effect on the RSSI levels received. These average RSSI levels as per the algorithm, together with short comments on the results, is shown in table 4.13.

**Table 4.12: Extract from detailed RSSI values**

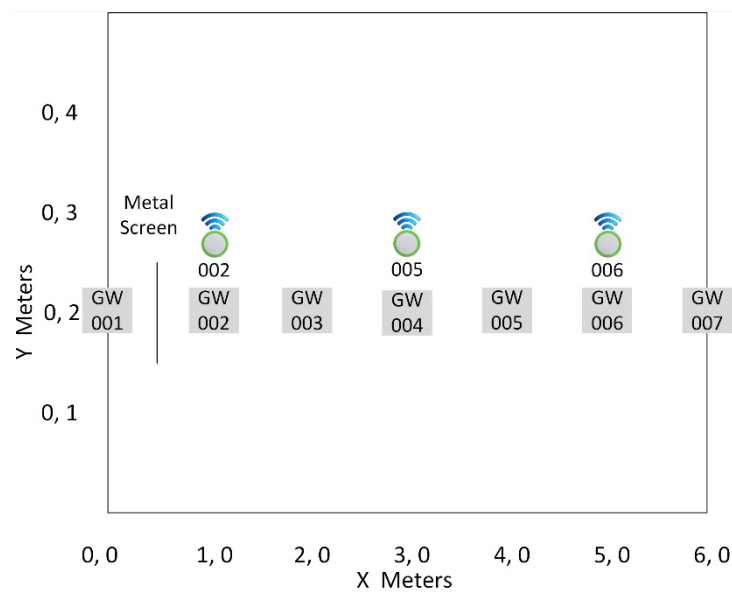
GW	LPN	Test 1		Test 2		Test 3		Avg.
		Count	Level	Count	Level	Count	Level	
GW-001	LPN001	85	-60	111	-59	110	-59	-57
GW-001	LPN001	73	-55	81	-55	86	-55	
GW-001	LPN001	62	-58	78	-58	57	-58	
GW-001	LPN002	83	-64	97	-64	86	-64	-64
GW-001	LPN002	65	-60	55	-61	65	-67	
GW-001	LPN002	52	-67	52	-67	52	-61	

**Table 4.13: Results from Perspex obstruction**

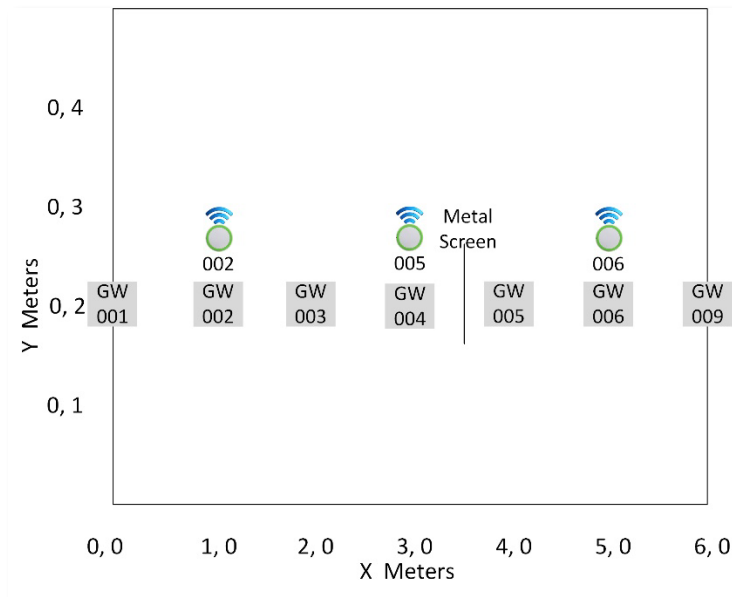
GW	LPN	Avg.	Result
GW-001	LPN001	-55.9	Consistent, expected
GW-001	LPN002	-64.0	Screens have no effect
GW-002	LPN001	-57.9	Consistent, expected
GW-002	LPN002	-62.3	Screens have no effect
GW-003	LPN001	-68.7	Screens have no effect
GW-003	LPN002	-46.7	Consistent, expected
GW-004	LPN001	-73.1	Screens have no effect
GW-004	LPN002	-56.3	Consistent, expected

### Experiment using metal sheet

The 0.5 mm metal obstruction was placed at position (0.5, 2) and (3.5, 2), where location is measured in metres from the bottom left of the figure, which indicates a corner of the room as a point of reference. RSSI levels were recorded by each of the seven gateways GW-001 to GW-007 placed 1 m apart at locations (0, 2), (1, 2), (2, 2), (3, 2), (4, 2), (5, 2) and (7, 2) respectively. LPN002, LNP005 and LPN006 were placed at location (1, 2), (3, 2) and (5, 2) respectively. This is represented in figures 4.33 and 4.34.



**Figure 4.33: Metal obstruction at (X, Y) position (0.5, 2)**



**Figure 4.34: Metal obstruction at (X, Y) position (3.5, 2)**

The RSSI levels transmitted from the LPNs were recorded at each of the gateways. The recording was done for a 5-minute period and was iteratively tested thrice, viz. test 1, test 2 and test 3.

The following process was followed to calculate average RSSI level:

For each iteration

Count the number of the distinct RSSI levels received

Select the top three RSSI levels that have the highest counts

Given the highest count received RSSI level per iteration, compute the average of these RSSI levels for all three iterations.

Table 4.14 shows an extract of the data for gateways GW-001 and GW-002 and nodes LPN002, LPN005 and LPN006. Using the process described above, the initial experiment contained no obstructions. Subsequent experiments using the same process had the obstructions of a 0.5 mm metal sheet placed in the path of transmission to observe the effect on the RSSI levels received. The difference in the average RSSI levels as per the algorithm, together with short comments on the results, is shown in table 4.14 and the detailed results are included in appendix D. The difference in average RSSI level between the absence of the obstruction and the presence of the obstruction is classified using table 4.15 and is shown in table 4.14.

**Table 4.14: Extract of results from metal obstruction**

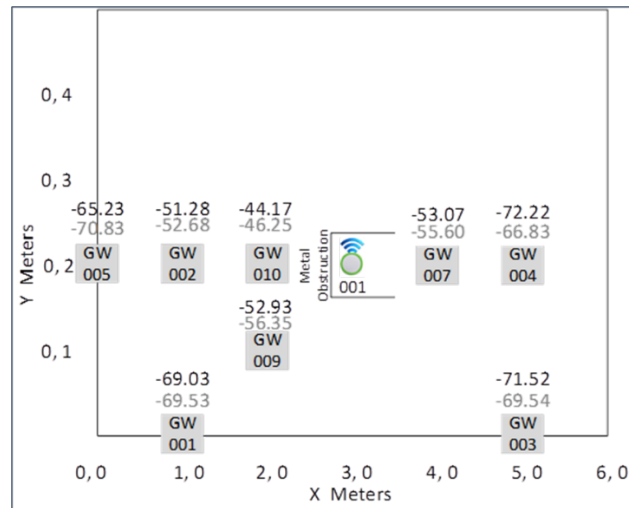
	Gateway	LPN	Difference from first value only	Result – using highest count	Explanation
1	GW-001	LPN002	-1	Negligible	Due to close proximity
2	GW-001	LPN005	1	Negligible	Due to close proximity
3	GW-001	LPN006	-6	Low	Loss from distance
4	GW-002	LPN002	7	High	Possible reflection due to close proximity
5	GW-002	LPN005	-5	Low	Expected

**Table 4.15: Classification of different levels**

Difference in level	Classification category
< -11	High loss
-10	Medium loss
-9	Medium loss
-8	Medium loss
-7	Low loss
-6	Low loss
-5	Low loss
-4	Low loss
-3	Low loss
-2	Negligible change
-1	Negligible change
0	Negligible change
1	Negligible change
2	Negligible change
3	Low gain
4	Low gain
5	Low gain
6	Low gain
7	Low gain
8	Medium gain
9	Medium gain
10	Medium gain
> 11	High gain

### Experiment using fabricated metal shaped obstruction

The third obstruction used was a fabricated steel enclosure placed in the centre of the room at location (3, 2). Eight gateways, GW001 to GW005, GW007, GW009 and GW0010, and node LPN001 were used to capture levels before and after the obstruction was used. This is depicted in figure 4.35.



**Figure 4.35: Effect of fabricated obstruction**

The average RSSI levels computed with and without the obstruction and the difference in RSSI levels are summarised in table 4.16. The RSSI level difference is classified, and a short explanation is given in the table as well.

**Table 4.16: Results with fabricated obstruction**

Gateway	LPN	Avg. without obstruction	Avg. with obstruction	Difference in levels	Result	Explanation
<b>GW-001</b>	LPN001	-69.03	-69.53	-0.51	Negligible	Expected
<b>GW-002</b>	LPN001	-51.28	-52.68	-1.39	Negligible	Expected
<b>GW-003</b>	LPN001	-71.52	-69.54	1.98	Negligible	Expected
<b>GW-004</b>	LPN001	-72.22	-66.83	5.38	Medium gain	Expected
<b>GW-005</b>	LPN001	-65.23	-70.83	-5.60	Medium loss	Expected
<b>GW-007</b>	LPN001	-53.07	-55.60	-2.53	Negligible	Not expected
<b>GW-009</b>	LPN001	-52.93	-56.35	-3.42	Low loss	Expected
<b>GW-010</b>	LPN001	-44.17	-46.25	-2.08	Negligible	Expected

#### 4.4.2 Fingerprinting

Distance is calculated using formula 4.1 for calculation of distance:

$$d = 10^{\left( \frac{\text{TxPower} - \text{RSSI}}{10 * n} \right)} \quad (4.1)$$

where

TxPower is the transmit power level,

RSSI is the receive signal strength indicator at the receiver and

$n$  is a value of between 2 to 4 in free space.

The RSSI levels were captured to form a fingerprint of the target environment. The fingerprints included both the gateways positioned at fixed locations and the nodes that could change positions and which represent the objects in a real-world environment. The initial fingerprint was to find the initial layout. Any change detected from the levels received at gateways as per the initial fingerprint model from the first step, indicated an obstruction in the path of transmission. The model designed used a link quality indicator to refine the value of  $n$  used in the formula for the calculation of the location of a node. RSSI levels recorded during various experiments, which form the basis of a fingerprint, are detailed in section 4.3.

#### 4.4.3 Trilateration

Trilateration is a technique for calculating the intersections of three sphere surfaces given their centres and radii. GPS satellites are used to calculate the location that is equipped with a GPS sensor. In and indoor environment, GPS signals are not always available. In this study, the iRTLS design had multiple fixed gateways that measured the received RSSI levels, from which the trilateration algorithms could be used to determine the location of the nodes.

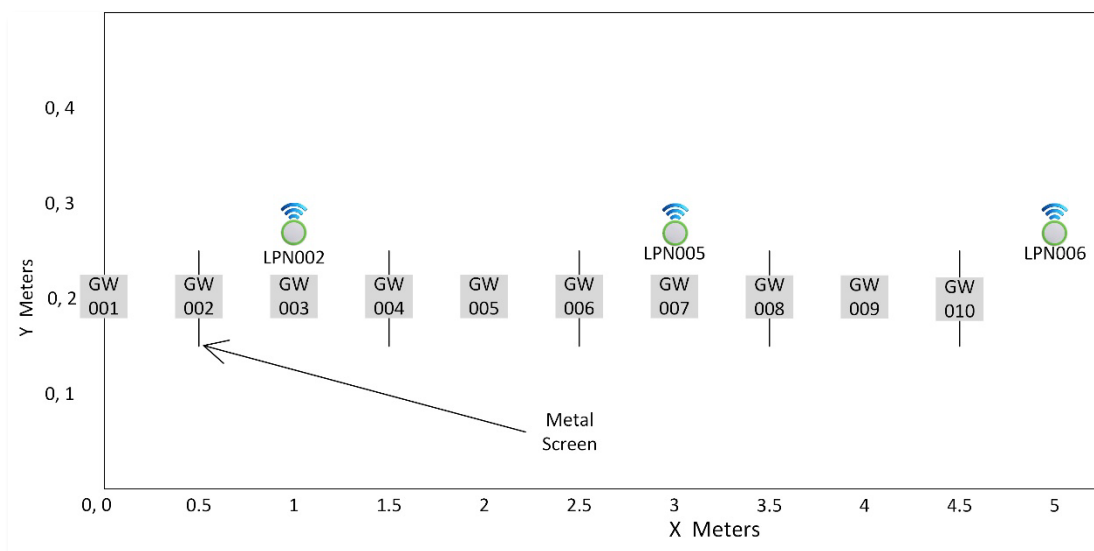
### 4.5 Machine learning

For the use of machine learning in the experiments, KNN, SVM and ANN were chosen in accordance with the methodology. These machine learning methods were applied to the datasets collected using the designed model to determine the location of obstructions. The metrics for these machine learning methods are MAE, MSE, RMSE and R2. The experiments were carried out once the machine learning models' configuration was finalised.

KNN was used for clustering the data, by using a majority voting mechanism. It collected data from a training data set and, based on the clusters discovered, it made a prediction of the obstruction's location. The received RSSI levels coming from the entities in this layout, as well as the Euclidian positions of the gateways and LPNs, were used to categorise the data points into the various classes using SVM. Based on this classification, the SVM model determined the location of the obstructions. ANN received input similar to that of SVM. Based on this input, ANN developed its own internal model with prediction, adjusting the model until the prediction was within an acceptable threshold.

#### 4.5.1 Experimental layout

The experimental layout for the data capture for machine learning used ten gateways labelled GW001 to GW010 and three nodes LPN002, LPN005 and LNP006, as depicted in figure 4.36. The metal obstruction was placed at position (X, Y) where X = x coordinate and Y = y coordinate.



**Figure 4.36: Gateways and node layout for machine learning**

An extract of the data capture for the machine learning is shown in table 4.17. The “Test Name” is used to identify the presence and the location of the obstruction and the “obs\_x\_pos” and “obs\_y\_pos” indicate the position of the obstruction. The “Node\_name” captures the LPN node name, whilst the “Node\_x\_pos” and “Node\_y\_pos” indicate the location of the node. The “Gateway\_name” captures the Gateway name, whilst the “Gateway\_x\_pos” and “Gateway\_y\_pos” indicate the location of the Gateway. The “Session\_id” and “Message\_id”

uniquely identify the message, and the “RSSI” field is used to capture the RSSI level received by the gateway from the node.

**Table 4.17: Extract of data captured for machine learning**

Testname	Obs x_pos	obs_y_pos	session_id	message_id	node_name	node_x_pos	node_y_pos	gateway_name	gateway_x_pos	gateway_y_pos	rsssi
<b>Session 6: no partition</b>	-1	-1	378	3707	LPN002	1	2	GW-004	1.5	2	-47
<b>Session 6: no partition</b>	-1	-1	378	3707	LPN002	1	2	GW-005	2	2	-49
<b>Session 6: no partition</b>	-1	-1	378	3707	LPN002	1	2	GW-007	3	2	-66
<b>Session 6: no partition</b>	-1	-1	378	3706	LPN002	1	2	GW-010	4.5	2	-69
<b>Session 6: no partition</b>	-1	-1	378	3707	LPN002	1	2	GW-010	4.5	2	-55
<b>Session 3: partition at position 0.5</b>	0.5	2	381	6236	LPN005	3	2	GW-009	4	2	-54
<b>Session 3: partition at position 0.5</b>	0.5	2	381	6238	LPN005	3	2	GW-009	4	2	-52
<b>Session 3: partition at position 0.5</b>	0.5	2	381	6235	LPN006	5	2	GW-004	1.5	2	-73
<b>Session 3: partition at position 0.5</b>	0.5	2	381	6236	LPN006	5	2	GW-004	1.5	2	-54
<b>Session 3: partition at position 0.5</b>	0.5	2	381	6237	LPN006	5	2	GW-004	1.5	2	-70

#### 4.5.2 Training of the model and prediction of obstruction location

Initially test datasets were created to be used to validate which model, kernel and parameters to use in subsequent training and prediction models. For SVM the sampling “cross validation, number of folds = 2, 3, 5, 10, 20 and stratified = ‘Yes’, ‘No’” was used. The results of the test for SVM are listed in tables 4.18 and 4.19.

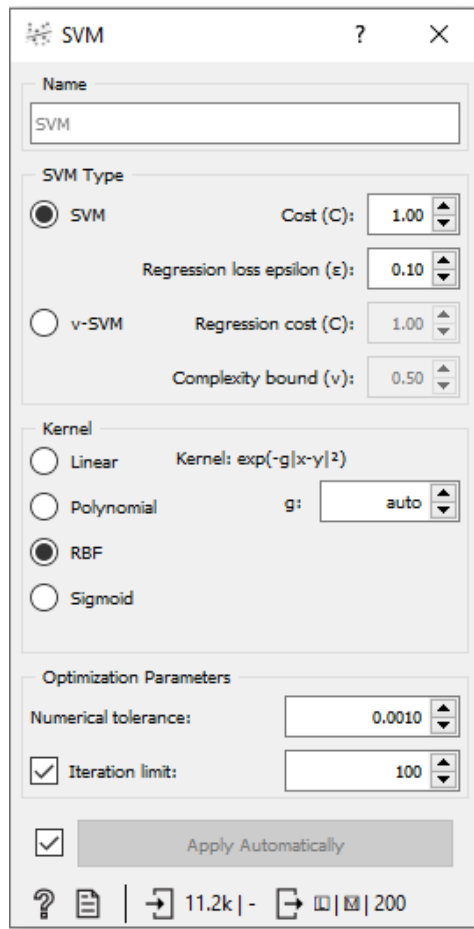
**Table 4.18: Results of test and score parameters for SVM**

Sampling	MSE	RMSE	MAE	R2
<b>Cross validation, Number of folds = 2, 3, 5, 10, 20 Stratified = yes, Stratified = No</b>	4.490	2.119	1.819	- 0.007
<b>Random sampling – repeat train test = 2 Stratified = yes</b>	4.512	2.124	1.828	- 0.006
<b>Random sampling – repeat train test = 10 Stratified = yes</b>	4.497	2.121	1.821	- 0.007

**Table 4.19: Results of different SVM kernels**

Sampling	MSE	RMSE	MAE	R2
<b>Kernel – linear</b>	4.525	2.127	1.828	- 0.015
<b>Kernel – polynomial – G:Auto, C: 1, D:3</b>	4.499	2.121	1.822	- 0.009
<b>Kernel – radial basis function G:Auto</b>	4.490	2.119	1.819	- 0.007
<b>Kernel – sigmoid - G:Auto, C: 1,</b>	4.526	2.128	1.830	- 0.015

The parameters setup for SVM training of the model is indicated in figure 4.37.



**Figure 4.37: SVM Parameters setup for training and prediction**

For KNN the following parameters were tested:

Metrics = “Euclidian” and “Manhattan”,

Weight = “uniform” with number of neighbours 1 to 10, and then multiples of 5 to 30 and

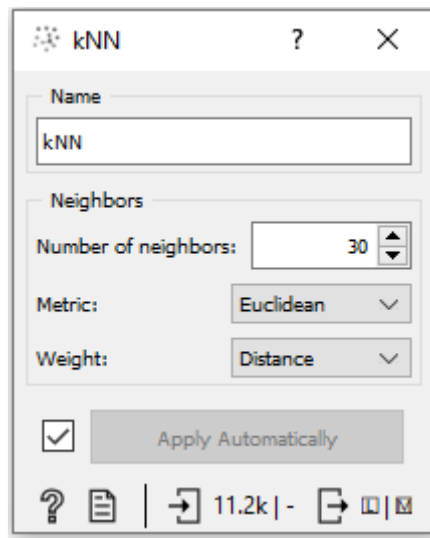
“Distance” with numbers of neighbours in multiples of 10 from 10 to 30.

These results are shown in table 4.20.

**Table 4.20: Results of different parameters with KNN**

<b>Metric</b>	<b>Weight</b>	<b>No. of neighbours</b>	<b>MSE</b>	<b>RMSE</b>	<b>MAE</b>	<b>R2</b>
<b>Euclidian</b>	Uniform	1	2.805	1.675	1.031	0.371
<b>Euclidian</b>	Uniform	2	2.160	1.470	0.960	0.516
<b>Euclidian</b>	Uniform	3	1.949	1.396	0.931	0.563
<b>Euclidian</b>	Uniform	4	1.848	1.359	0.919	0.586
<b>Euclidian</b>	Uniform	5	1.781	1.335	0.909	0.601
<b>Euclidian</b>	Uniform	6	1.742	1.320	0.902	0.610
<b>Euclidian</b>	Uniform	7	1.715	1.310	0.897	0.615
<b>Euclidian</b>	Uniform	8	1.699	1.304	0.896	0.619
<b>Euclidian</b>	Uniform	9	1.683	1.297	0.894	0.623
<b>Euclidian</b>	Uniform	10	1.667	1.291	0.892	0.626
<b>Euclidian</b>	Uniform	15	1.617	1.272	0.885	0.637
<b>Euclidian</b>	Uniform	20	1.595	1.263	0.882	0.642
<b>Euclidian</b>	Uniform	25	1.582	1.258	0.881	0.645
<b>Euclidian</b>	Uniform	30	1.574	1.254	0.881	0.647
<b>Manhattan</b>	Uniform	10	1.669	1.292	0.893	0.626
<b>Manhattan</b>	Uniform	15	1.621	1.273	0.888	0.636
<b>Manhattan</b>	Uniform	20	1.600	1.265	0.885	0.641
<b>Manhattan</b>	Uniform	25	1.588	1.260	0.883	0.644
<b>Manhattan</b>	Uniform	30	1.582	1.258	0.884	0.645
<b>Euclidean</b>	Distance	10	1.668	1.292	0.891	0.626
<b>Euclidean</b>	Distance	20	1.590	1.261	0.879	0.644
<b>Euclidean</b>	Distance	30	1.560	1.249	0.874	0.650

The parameters were set up in the model as indicated in figure 4.38.



**Figure 4.38: KNN parameters setup for training and prediction**

For ANN the following parameters were tested:

Activation = “ReLU”, “Identity”, “Logistic” and “Tanh”,

Solver = “Adam”, “L-BFGS-B” and “SGD” and

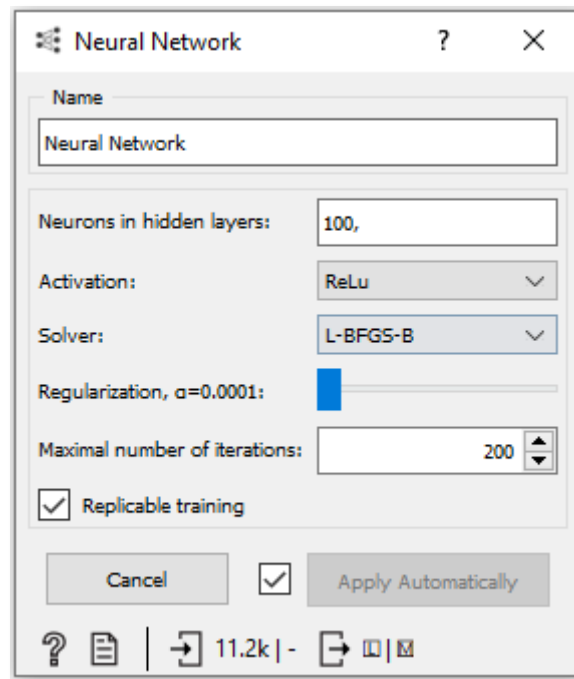
Number for neurons in hidden layers = 100.

The results of these tests are shown in table 4.21.

**Table 4.21: Results of different parameters with ANN**

Activation	Solver	Iterations	Neurons in hidden layers	MSE	RMSE	MAE	R2
ReLU	Adam	200	100	2.009	1.417	1.112	0.550
Identity	Adam	200	100	2.470	1.572	1.282	0.449
Logistic	Adam	200	100	2.433	1.560	1.265	0.455
Tanh	Adam	200	100	2.43	1.429	1.121	0.542
ReLU	L-BFGS-B	200	100	1.907	1.381	1.087	0.572
Identity	L-BFGS-B	200	100	2.462	1.569	1.272	0.448
Logistic	L-BFGS-B	200	100	2.04	1.416	1.122	0.551
Tanh	L-BFGS-B	200	100	1.912	1.383	1.089	0.571
ReLU	SGD	200	100	2.133	1.461	1.155	0.522
Identity	SGD	200	100	2.464	1.57	1.275	0.447
Logistic	SGD	200	100	2.462	1.569	1.273	0.448
Tanh	SGD	200	100	2.189	1.4810	1.177	0.509

The parameters that were set up for training and prediction for ANN are indicated in figure 4.39.



**Figure 4.39: ANN parameters setup for training and prediction**

Once the parameters that yielded that lowest metrics were established, these were used to configure the model for subsequent use. The RSSI levels transmitted from the nodes were recorded at each of the gateways. The training, prediction and validation outcomes are documented within the respective experiments below. For each experiment the results from dataset 1 is included in this section, whilst the results from dataset 2 and dataset 3 are included in appendix E.

### **Experiment 1**

The results of training the KNN, SVM and ANN models for experiment 1 for dataset 1 are displayed in table 4.22. Seventy percent (7 729 rows) of the data from dataset 1 that contained no obstructions and single obstruction at location  $x = 0.5$  were utilised for training the model. The models were then used to predict the locations based on the following datasets and their rows:

1. **Prediction 1:** Dataset 1 – 30% (3 314 rows) of the remaining rows of dataset 1
2. **Prediction 2:** Dataset 2 – 3 407 rows

3. **Prediction 3:** Dataset 2 – 7 110 rows
4. **Prediction 4:** Dataset 3 – 3 130 rows
5. **Prediction 5:** Dataset 3 – 7 325 rows

The KNN and ANN machine learning model performed equally well based on the MSE metric for four predictions, and ANN performed marginally better for one prediction. These results are documented in tables 4.23 to 4.25. Using the trained models on dataset 2, ANN performed the best for the MSE metric in all five predictions. Using the trained model on dataset 3, ANN also performed the best in four of the predictions, and KNN performed the best for one of the predictions. The results for these predictions are listed in appendix E.

**Table 4.22: Experiment 1, training on dataset 1**

<b>Training on dataset 1</b>				
<b>Model</b>	<b>MSE</b>	<b>RMSE</b>	<b>MAE</b>	<b>R2</b>
<b>KNN</b>	0.002	0.039	0.002	0.997
<b>SVM</b>	0.007	0.083	0.080	0.988
<b>ANN</b>	0.000	0.005	0.004	1.000

**Table 4.23: Experiment 1: Training on dataset 1, prediction on dataset 1 and dataset 2**

<b>Prediction 1</b>					<b>Prediction 2</b>				
<b>Model</b>	<b>MSE</b>	<b>RMSE</b>	<b>MAE</b>	<b>R2</b>	<b>Model</b>	<b>MSE</b>	<b>RMSE</b>	<b>MAE</b>	<b>R2</b>
<b>KNN</b>	0.000	0.019	0.000	0.999	<b>KNN</b>	0.000	0.011	0.000	1.000
<b>SVM</b>	0.007	0.082	0.079	0.988	<b>SVM</b>	0.006	0.081	0.078	0.988
<b>ANN</b>	0.000	0.005	0.003	1.000	<b>ANN</b>	0.000	0.005	0.003	1.000

**Table 4.24: Experiment 1: Training on dataset 1, prediction on dataset 2 and dataset 3**

Prediction 3					Prediction 4				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
<b>KNN</b>	0.000	0.021	0.001	0.999	<b>KNN</b>	0.000	0.013	0.000	1.000
<b>SVM</b>	0.007	0.082	0.079	0.988	<b>SVM</b>	0.006	0.080	0.078	0.988
<b>ANN</b>	0.000	0.004	0.003	1.000	<b>ANN</b>	0.000	0.005	0.003	1.000

**Table 4.25: Experiment 1: Training on data set 1, prediction on dataset 3**

Prediction 5				
Model	MSE	RMSE	MAE	R2
<b>KNN</b>	0.001	0.029	0.001	0.998
<b>SVM</b>	0.007	0.082	0.079	0.988
<b>ANN</b>	0.000	0.005	0.003	1.000

## Experiment 2

The results of training the KNN, SVM and ANN models for experiment 2 for dataset 1 are displayed in Table 4.26. Seventy percent (11 479 rows) of the data from dataset 1 that contained no obstructions and two obstructions at locations  $x = 0.5$  and  $x = 1.5$  were utilised for training the model. The models were then used to predict the locations based on the following datasets and their rows:

1. **Prediction 1:** Dataset 1 – 30% (4 929 rows) of the remaining rows of dataset 1
2. **Prediction 2:** Dataset 2 – 4 708 rows
3. **Prediction 3:** Dataset 2 – 10 987 rows
4. **Prediction 4:** Dataset 3 – 4 760 rows
5. **Prediction 5:** Dataset 3 – 11 127 rows

The outcomes of predictions indicate that KNN performed the best for all for predictions based on the MSE metric. These results are presented in tables 4.27 to 4.29. KNN also performed the best based on the trained models on datasets 2 and 3. The results for these predictions are listed in appendix E.

**Table 4.26: Experiment 2: Training on dataset 1**

Training on dataset 1				
Model	MSE	RMSE	MAE	R2
KNN	0.079	0.282	0.142	0.925
SVM	1.057	1.028	0.911	0.007
ANN	0.113	0.335	0.247	0.894

**Table 4.27: Experiment 2: Training on dataset 1, prediction on dataset 1 and dataset 2**

Prediction 1					Prediction 2				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.082	0.286	0.146	0.923	KNN	0.078	0.280	0.152	0.927
SVM	1.065	1.032	0.912	0.000	SVM	1.077	1.038	0.921	-0.006
ANN	0.328	0.573	0.330	0.692	ANN	0.354	0.595	0.355	0.669

**Table 4.28: Experiment 2: Training on dataset 1, prediction on dataset 2 and dataset 3**

Prediction 3					Prediction 4				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.086	0.294	0.151	0.919	KNN	0.076	0.276	0.146	0.930
SVM	1.076	1.037	0.921	-0.006	SVM	1.086	1.042	0.926	-0.002
ANN	0.353	0.594	0.355	0.670	ANN	0.343	0.586	0.345	0.683

**Table 4.29: Experiment 2: Training on dataset 1, prediction on dataset 3**

Prediction 5				
Model	MSE	RMSE	MAE	R2
KNN	0.084	0.290	0.147	0.922
SVM	1.083	1.041	0.925	-0.001
ANN	0.342	0.584	0.344	0.684

**Experiment 3**

Experiment 3 consisted of one obstruction that was placed at location  $x = 2.5$ . The results of training the KNN, SVM and ANN models using 70% (7 892 rows) of data from dataset 1 are displayed in table 4.30. The models were then utilised to predict the locations for the following datasets:

1. **Prediction 1:** Dataset 1 – 30% (3 385 rows) of the remaining data in dataset 1
2. **Prediction 2:** Dataset 2 – 3 250 rows
3. **Prediction 3:** Dataset 2 – 7 583 rows
4. **Prediction 4:** Dataset 3 – 3 249 rows
5. **Prediction 5:** Dataset 3 – 7 576 rows

Results for predictions 1 to 5 showed that KNN performed the best over ANN and SVM based on the MSE metric. These results are shown in tables 4.31 to 4.34. Results from predictions of the trained models from datasets 2 and 3 indicated that KNN outperformed ANN and SVM in all predictions based on the MSE metric. The results of the predictions for datasets 2 and 3 are included in appendix E.

**Table 4.30: Experiment 3: Training on dataset 1**

Training on dataset 1				
Model	MSE	RMSE	MAE	R2
KNN	0.011	0.103	0.005	0.997
SVM	0.006	0.079	0.070	0.998
ANN	0.000	0.005	0.003	1.000

**Table 4.31: Experiment 3: Training on dataset 1, prediction on dataset 1 and dataset 2**

Prediction 1					Prediction 2				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.002	0.040	0.001	0.999	KNN	0.011	0.105	0.004	0.996
SVM	3.287	1.813	1.750	-0.073	SVM	3.377	1.838	1.775	-0.106
ANN	2.000	1.414	1.002	0.347	ANN	2.106	1.451	1.054	0.311

**Table 4.32: Experiment 3: Training on dataset 1, prediction on dataset 2 and dataset 3**

Prediction 3					Prediction 4				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.004	0.061	0.002	0.999	KNN	0.001	0.037	0.001	1.000
SVM	3.370	1.836	1.774	-0.103	SVM	3.283	1.812	1.748	-0.072
ANN	2.099	1.449	1.052	0.313	ANN	1.993	1.412	0.998	0.349

**Table 4.33: Experiment 3: Training on dataset 1, prediction on dataset 3**

Prediction 5				
Model	MSE	RMSE	MAE	R2
KNN	0.007	0.087	0.004	0.998
SVM	3.277	1.810	1.747	-0.070
ANN	1.987	1.410	0.996	0.351

#### Experiment 4

Experiment 4 consisted of obstructions at locations  $x = 2.5$  and  $x = 4.5$ . The results of training the KNN, SVM and ANN models using 70% (11 683 rows) of the data from dataset 1 are

displayed in Table 4.30. The models were then utilised to predict the locations for the following datasets:

1. **Prediction 1:** Dataset 1 – 30% (5 008 rows) of the remaining data in dataset 1
2. **Prediction 2:** Dataset 2 – 4 813 rows
3. **Prediction 3:** Dataset 2 – 11 230 rows
4. **Prediction 4:** Dataset 3 – 4 885 rows
5. **Prediction 5:** Dataset 3 – 11 401 rows

Results for predictions 1 to 5 showed that KNN performed the best over ANN and SVM based on the MSE metric. These results are shown in tables 4.31 to 4.34. Results from predictions of the trained models from datasets 2 and 3 indicated that KNN performed the best for all predictions based on the MSE metric. The results of the predictions for datasets 2 and 3 are included in appendix E.

**Table 4.34: Experiment 4: Training on dataset 1**

Training on dataset 1				
Model	MSE	RMSE	MAE	R2
KNN	0.372	0.610	0.346	0.928
SVM	4.847	2.202	1.996	0.059
ANN	0.509	0.713	0.537	0.901

**Table 4.35: Experiment 4: Training on dataset 1, prediction on dataset 1 and dataset 2**

Prediction 1					Prediction 2				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.377	0.614	0.349	0.927	KNN	0.461	0.679	0.432	0.908
SVM	4.817	2.195	1.990	0.064	SVM	4.711	2.171	1.960	0.058
ANN	0.502	0.708	0.534	0.903	ANN	0.575	0.759	0.607	0.885

**Table 4.36: Experiment 4: Training on dataset 1, prediction on dataset 2 and dataset 3**

Prediction 3					Prediction 4				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.359	0.599	0.329	0.928	KNN	0.378	0.615	0.360	0.927
SVM	4.709	2.170	1.957	0.058	SVM	4.861	2.205	2.006	0.062
ANN	0.505	0.711	0.528	0.899	ANN	0.512	0.716	0.551	0.901

**Table 4.37: Experiment 4: Training on dataset 1, prediction on dataset 3**

Prediction 5				
Model	MSE	RMSE	MAE	R2
KNN	0.393	0.627	0.351	0.924
SVM	4.868	2.206	2.004	0.061
ANN	0.514	0.717	0.539	0.901

## 4.6 Summary

In this chapter the proposed designed model was validated. The model, together with the experiments, took into consideration the practical factors of a real-world environment. The feasible throughput and range for the selected technology were determined. The effect of different obstructions encountered in an indoor environment for BLE was also determined. An improvement to the real-time quality indicator was established. Finally, it was discovered that obstructions could be located with a reasonable degree of accuracy, as demonstrated by the results of machine learning.

## CHAPTER FIVE: CONCLUSION

### 5.1 Introduction

In this chapter, a summary of the research's accomplishments and recommendations for future research within the scope of this thesis are presented. The research focused on finding the best technology, determining the maximum feasible range, assessing the effect of obstructions in the path of transmission, designing a model and using machine learning for determination of obstruction location. The chosen methodology and procedures indicated that the created design could be feasibly implemented. The original objectives were compared against the outcomes to confirm that they were attained.

### 5.2 Summary of the analysis of results

#### 5.2.1 iRTLS technologies

Given the constraints identified in literature, further review of technologies revealed that the most suitable technology for use in an indoor environment is BLE. Recent and continuous enhancements to BLE have made it the preferred technology for location determination. This is supported by improvements in location determination techniques such as fingerprinting, the Kalman filter and the link quality indicator.

#### 5.2.2 Throughput and range

In the throughput and range experiment, the correlation between the time required to transmit 1 Mb of data and the throughput was nonlinear, indicating the presence of potential interference variables. The analysis was unable to identify these interfering variables owing to the limitations.

When broadcasting at a level of -20 dB with a clean line of sight within the open-plan office, a range of 25 m was obtained. The throughput of 759 Kbps at 5 m and 533 Kbps at 25 m suggests that BLE V5 can send data at greater speeds across greater distances. The measured throughput in the boardroom ranged from 793 Kbps at 1 m to 222 Kbps at 10 m. Using a Bluegiga USB dongle as a gateway and a RadBeacon tag transmitting at the same power level, Wang *et al.* (2015) achieved a transmission range of 2 m. This suggests that different environments provide

distinct effects. Possible causes include the difference in size between the boardroom and the open-plan office, as well as the composition of the walls.

When the PDKs were placed at a total distance of 2 m from a glass door, a lower throughput rate of 386 Kbps was measured compared to the maximum throughput rate of 579 Kbps measured when the PDKs were placed at a distance of 4 m apart with the same glass door as the obstruction. At 5 m, the throughput was comparable to that measured at 2 m. However, the throughput reduced to 48 Kbps at a distance of 6 m, and to an unacceptable level at a distance of 6.5 m, and in the majority of cases completely disconnected at this distance. The RSSI measurements did not change significantly from 2 to 6 m. These measurements across a classroom door are noteworthy since the glass may reflect electromagnetic waves.

Figure 4.10 demonstrates that communication is feasible up to a distance of 11 m in the office area and 13 m in the two corridors. This shows that a BLE signal sent at -20 dB can pass through a door and numerous drywall walls. However, as the distance increases, reception will deteriorate due to the signal's weakening through these obstructions. The throughput reduced from 227 Mbps at 10 m to 82 Mbps at 11 m. In the fire escape, connectivity was acceptable behind the closed door, but non-existent at a distance of 5 m, indicating that the brick wall had a substantial impact on signal strength. Signal strength and throughput were excellent when there was a clear line of sight, but when glass, partitions, brick walls, etc. were placed in the signal's path, signal strength decreased rapidly. This loss increased in proportion to the obstruction's material composition, such as a solid brick wall or reflecting glass.

The results of long range and excellent throughput at the lowest power level employed are encouraging for use in environments that are sensitive to high electromagnetic wave levels, such as healthcare. Assuming the BLE node is stationary and a connection has been established, a satisfactory throughput can be attained for pure data transmission. However, multiple gateways and methods must be employed to determine location with a high degree of precision for an iRTLS.

A pattern within some environments and random variations at particular distances were observed, for example at 15 m in the open-plan office, the throughput increased, and the variation in the findings thus represents an actual environment. This suggests that there is a nonlinear relationship as distance rises. Between 5 m and 15 m, the throughput in the boardroom decreased substantially. RSSI results indicate that Bluetooth devices can be

identified with a degree of precision (in an open area, an RSSI level of approximately -80 dBm indicates a distance between 2 and 25 m; in an office environment, an RSSI level of approximately -80 dBm indicates a distance between 2 and 13 m; and in the boardroom, an RSSI level of approximately -80 dBm indicates a distance between 2 and 6 m).

### **5.2.3 Transmission through obstructions**

Generally speaking, the reception of broadcast packets follows a linear pattern. In some situations, however, a nonlinear level is observed, indicating the existence of variables that could potentially interfere with this experiment. Due to restrictions, it was not possible to isolate these conflicting variables in this investigation.

Figure 4.13 showed that there was a variance in the RSSI values of the two receivers, based on the results of tests conducted in the adjacent office. Significant and observable variations existed at the higher transmitted levels, but they were less pronounced at the lower levels. The results of the test conducted between these two locations indicate that BLE signals can pass through solid walls. Even at low transmission levels of -12 dBm, the signal may be received from a distance of around 6 m.

Observations revealed a pattern in some environments as well as and random changes at specific distances. The variation in the results reflects the actual, real-world environment. This indicates that the relationship between the distance and received RSSI levels was nonlinear as distance increased. However, RSSI results indicate that a Bluetooth device can be identified with a certain degree of precision through a solid brick wall, double drywall barriers or double glass partitions. Furthermore, this data can be processed by commonly used algorithms such as the Kalman filter, before being applied to a location determination algorithm.

As a result of the COVID-19 pandemic, the widespread use of Perspex in office environments became a norm to help prevent the spread of the virus. In addition, the material is also used as general protection screens in the medical and office environments. Therefore, Perspex was also considered as an obstruction to be investigated. The results showed that 3 mm, 5 mm and both 3 mm and 5 mm together had no effect on BLE transmission.

The material of the metal obstruction used was similar that of a steel filing cabinet. This was considered as such cabinets are typically found in an office. Variation of levels received at the

gateways in close proximity was negligible, whilst gateways that were further apart recorded some medium to high variation. The gain was due to the possible reflection of the signals from the metal screen. The loss was due to the absorption of signals as well as the fact that lower-level signals could not pass through the metal obstruction. The resultant conclusion is that metal sheets in the path of transmission do have an effect on the BLE transmission. Therefore, this will need to be considered for any location determination.

In some cases, an object obstructs the BLE transmission from three sides. A fabricated metal obstruction consisting of three sides was used to simulate this scenario. Gateways GW001, GW002, GW005, GW009 and GW010 showed a decrease in the RSSI level with the obstruction, which was expected as the obstruction was in the path of transmission. Gateways GW003 and GW004 showed relative gains, and this can be attributed to the signal being reflected. However, gateway GW007 showed a negligible loss, and this was not expected. In general, a metal obstruction that blocks the BLE signal in three directions causes relative losses to gateways behind the obstruction whilst causing gains to receive RSSI levels on the open side of the obstruction. Consideration will need to be given to the material of which the obstruction is made and the shape of the obstruction. Such obstructions will need to be carefully considered for location of BLE nodes.

#### **5.2.4 Real-time quality link indicator**

Based on the outcomes of this research, a meaningful measure that is best suited for data smoothing has been established. The moving average (mean) computed from five RSSI levels was less dependable for subsequent use in smoothing the original RSSI. After extensive computation, it was determined that the moving average (mean) derived from ten RSSI levels provided a more accurate basis for future computation. Therefore, a minimum of ten RSSI readings is advised for the most accurate location estimations. The mean provided the best smoothing computation at each of the four gateways, followed by the median.

Jayakody *et al.* (2016) improved the raw RSSI levels by multiplying the current and previous RSSI values by a factor between 0 and 1 to obtain an enhanced RSSI value. Therefore, this new RSSI value depended on the past and current RSSI values. They then generated a smoothed RSSI by applying a 20-value moving average to the newly obtained RSSI stream. The RSSI values are then forwarded to subsequent processing in order to establish the location. Compared to the procedure recommended in the present study, this entails one additional step. Also, the

process requires twice as many RSSI levels for the smoothing step, which increases the difficulty of collecting a data stream for further processing.

According to the mode calculated at each of the four gateways, the RSSI stabilises with time and remains constant. In the majority of instances, a large proportion of RSSI values appear to be within the top three RSSI values, with the mode becoming more apparent when viewed over a longer period. This pattern was discovered across the majority of the data collected. The combination of the mean, mode and RSSI levels close to the mode provides a reliable starting point for the location calculation. As an indoor environment is dynamic, it is unnecessary to collect, save and reference a fingerprint in this instance. The fluctuating indoor environment will result in varying RSSI readings each time. Moreover, because the movement of individuals alters the environment, it becomes more difficult to calculate the location precisely. This will necessitate the use of a more dynamic link quality indicator to detect the position of a node.

### **5.2.5 Improvement of location determination in a dynamic environment**

The formula for location determination includes a factor  $n$  that ranges between 2 and 4. This model, designed with fixed gateways, enables the real-time determination of this factor. The improved source of data as a link quality indicator of the fixed gateways, as demonstrated, is used as an input to determine the factor  $n$  for location determination. The mesh formed by the gateways in the model further strengthens the methodology for the location determination.

### **5.2.6 Indoor real-time location system model**

The designed model uses the technologies and techniques identified, viz. BLE, fingerprinting and machine learning. This model includes the ability to determine the location of a node through obstructions by using a link quality indicator instead of the traditional fingerprint, which will be outdated in a dynamic environment. The most popular machine learning modes (KNN, SVM and ANN) together with the metrics MSE, RMSE, MAE and R2 are used to demonstrate the ability to locate and obstruction.

### **5.2.7 Machine learning for obstruction location determination**

Machine learning models KNN, SVM and ANN, together with the metrics MSE, RMSE, MAE and R2, were used for the prediction of obstructions in an indoor environment and yielded positive results. When the dataset contains fewer locations, predictions of those locations are

typically more accurate. When more data is available for a given experiment, some machine learning methods perform better than others in predicting the location of obstructions. The experiments conducted indicated that KNN performed better than ANN and SVM in predicting obstruction location based on the MSE metric.

Table 5.1 displays the frequency with which each machine learning model achieved the best performance on the trained dataset. To highlight the experiment’s extent of variability, the total number of rows used for all predictions is shown. At the end of the table, the total number of predictions for each machine learning model and their respective percentages are also displayed to show their comparison. It is evident from the findings that the KNN machine learning model performed better 83% of the time, followed by the ANN model at 17%, based on the MSE metric. The difference between ANN and KNN prediction errors was 0.001, 0.002 and 0.004, resulting in a 17% performance advantage over KNN and SVM. This resulting difference is very marginal for the prediction models. SVM did not outperform any of the other prediction methods.

In conclusion, machine learning can be used to predict the location of obstructions in an indoor environment for an iRTLS model utilising BLE. This conclusion is supported by the unprocessed, independent experimental results of the research. The results of location determination can be enhanced through preprocessing.

**Table 5.1: Number of successful predictions per trained dataset**

	KNN	SVM	ANN	No. of rows
<b>Experiment 1</b>				
Dataset 1	4	0	1	24286
Dataset 2	0	0	5	24545
Dataset 3	1	0	4	24330
<b>Experiment 2</b>				
Dataset 1	5	0	0	36511
Dataset 2	5	0	0	37003
Dataset 3	5	0	0	36863
<b>Experiment 3</b>				
Dataset 1	5	0	0	25039
Dataset 2	5	0	0	25348
Dataset 3	5	0	0	25625
<b>Experiment 4</b>				
Dataset 1	5	0	0	37337
Dataset 2	5	0	0	37790
Dataset 3	5	0	0	37619
<b>Total</b>	<b>55</b>	<b>0</b>	<b>5</b>	
<b>Percentage</b>	<b>83%</b>	<b>0%</b>	<b>17%</b>	

### 5.3 Addressing the study objectives

This section details how each of the objectives were achieved.

***Objective 1:** Determine the most viable iRTLS technology which balances the identified real-world constraints of small form factor, low cost, reasonable accuracy and low power consumption.*

A literature survey was conducted to analyse the current technologies and thus determine the most feasible ones, given the constraints, as indicated in section 5.2.1.

***Objective 2:** Based on this technology determine a workable range for the purposes of location determination.*

Detailed experiments were conducted to establish the workable range for an indoor environment. The details of the results are explained in section 5.2.2.

***Objective 3:** Based on this range, develop a cost-effective model for a suitable layout of various pieces of equipment that would meet the needs of location determination within the constraints identified.*

The design science methodology was followed to design and test selected aspects of the model. The details are provided in sections 5.2.3 to 5.2.6.

***Objective 4:** Based on this layout, identify suitable methods to accurately determine the location of obstructions within the identified environment*

Machine learning methods were applied to estimate the most likely position of obstructions in the experimental environment. These methods were evaluated using an experiment, and the most accurate machine learning model was determined via the metrics deemed to be the most suitable for this particular domain. The results are discussed in section 5.2.7

### 5.4 Addressing the research questions

***Research question 1:** What is the most suitable technology which balances selected constraints that can be incorporated within a viable model of an iRTLS that provides accurate location determination within an acceptable range?*

Based on the literature study and the experiments carried out, BLE was identified as the most suitable technology for this type of location determination. This is also illustrated in section 5.2.1.

***Research question 2: What is the most feasible method for obstruction location determination that is compatible with the newly developed model of an iRTLS?***

The results of the experiments carried out indicate that machine learning can be used to determine the location of obstructions in the path of transmission, as indicated in section 5.2.7.

## **5.5 Success criteria and contribution of the research**

The study produced a feasible, cost-effective iRTLS model with enhanced techniques for improved location of stationary obstructions. A review of the pertinent literature led to the reduction of these identified attributes to six. Within the limitations of the domain, the resulting design successfully achieved a balance between the majority of the chosen attributes. The high cost, high power consumption and high form factor of earlier systems were some of the drawbacks that this design managed to reduce. The main contribution of this study is that machine learning can be used to locate obstructions in the path of transmission in a given environment. Once obstructions have been located, this information can be used to improve an indoor location system.

## **5.6 Limitations**

The research was conducted during the COVID-19 pandemic. As a result, the researcher was not able to access a larger real-world testing environment that would have provided for more thorough testing of the model. Only certain aspects of the designed model were tested as a result of the scope of the thesis and the time constraints. As mentioned previously, many aspects of iRTLS were explored. The relevant machine learning methods were applied for the determination of obstruction location. However, a comprehensive investigation of all suitable machine learning methods was not conducted due to time and scope limitations.

## **5.7 Future work**

With the proliferation of IoT and the demand for indoor location, there are several areas for improvement. Some of these are noted here based on the research carried out and the findings.

Current BLE nodes' point-to-point communication has a limited range and restricted coverage. A potential solution could be the design of a wireless mesh multi-hop network with numerous nodes capable of connecting with each other to enable packet routing in order to extend this limited coverage (Raza *et al.* 2016). Using technologies such as mesh node hopping with BLE V5 technology, the model can be scaled up to extend the range using a combination of Wi-Fi and BLE. The ability of nodes to communicate with each other will ensure multiple available paths between an LPN and a gateway, ensuring adequate robustness of the solution. Using intermediate nodes as transmitters, the data being transmitted over the network has the potential to be minimised, leading to more efficient use of power and memory. By minimising the power consumption and the memory usage, the size of the device can be reduced. Communication between LPNs and gateways can be optimised in terms of throughput, load balancing and energy consumption by using a network algorithm. The communication between the many LPNs in a mesh network will enable coverage of the blind spots that are not covered by the readers and will extend the network's coverage, while the availability of several communication pathways will assure scalability.

Nodes that change their location were not considered in this research and hence provide an opportunity to extend these results. Also, some obstructions found in a real-world environment constantly change their location, resulting in an invalidated previously captured footprint. Consequently, future research can expand on these findings, particularly the quality connection indicator and the prediction of the position of barriers using machine learning. The results of machine learning show that both KNN and ANN yield very similar results in terms of precision. Thus, future work may include incorporating both these models in an ensemble combination. Future work can enhance the model so that it can be applied in different domains. Given that technology is constantly evolving, the model developed in this study can be adapted to newer technologies with minimal enhancements.

## **5.8 Summary**

Through the experimental work completed, the combination of methods at each phase of the selected software development methodology produced both a system design and experimental results. The results produced by the chosen methodology demonstrate that the designed model is feasible within this domain and, given the constraints, is able to provide accurate results. The results of the throughput and range experiments and transmission through obstructions using

selected technologies have been discussed. The enhanced link quality indicator serves as a demonstration of how location determination was improved in a dynamic and changing environment. The most commonly used machine learning methods were used in the newly developed model to find obstructions in the path of BLE transmission. It has been shown in this chapter that the objectives were achieved, that the research questions were answered and that this research made significant contributions. This research depended on relevant technologies and, as technology improves, this research can be extended to improve indoor location determination.

## REFERENCES

- Abdull Sukor, A. S., Kamarudin, L. M., Zakaria, A., Abdul Rahim, N., Sudin, S. and Nishizaki, H. 2020. RSSI-Based for Device-Free Localization Using Deep Learning Technique. *Smart Cities*, 3 (2): 444-455.
- Abdullah, M. W., Fafoutis, X., Mellios, E., Klemm, M. and Hilton, G. S. 2015. Investigation into off-body links for wrist mounted antennas in bluetooth systems. In: *Proceedings of 2015 Loughborough Antennas & Propagation Conference (LAPC)*. Loughborough, UK, IEEE, 1-5.
- Adame, T., Bel, A., Carreras, A., Melià-Seguí, J., Oliver, M. and Pous, R. 2016. CUIDATS: An RFID-WSN hybrid monitoring system for smart health care environments. *Future Generation Computer Systems*, 78 (2): 601-615.
- Ahmadi, H. and Bouallegue, R. 2017. Exploiting machine learning strategies and RSSI for localization in wireless sensor networks: A survey. In: *Proceedings of 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*. Valencia, Spain, IEEE, 1150-1154.
- Al-Mejibli, I. S., Alwan, J. K. and Abd Dhafar, H. 2020. The effect of gamma value on support vector machine performance with different kernels. *International Journal of Electrical and Computer Engineering*, 10 (5): 5497.
- Alarifi, A., Al-Salman, A., Alsaleh, M., Alnafessah, A., Al-Hadhrami, S., Al-Ammar, M. A. and Al-Khalifa, H. S. 2016. Ultra wideband indoor positioning technologies: Analysis and recent advances. *Sensors*, 16 (5): 707.
- Alemdar, H. and Ersoy, C. 2010. Wireless sensor networks for healthcare: a survey. *Computer Networks*, 54 (15): 2688-2710.
- Alexander, I. and Kusuma, G. P. 2019. Predicting indoor position using Bluetooth low energy and machine learning. *International Journal of Scientific & Technology Research*, 8: 1661-1667.
- Alhomayani, F. and Mahoor, M. H. 2020. Deep learning methods for fingerprint-based indoor positioning: a review. *Journal of Location Based Services*, 14 (3): 129-200.
- Alsamhi, S. H., Ma, O. and Ansari, M. 2020. Convergence of machine learning and robotics communication in collaborative assembly: mobility, connectivity and future perspectives. *Journal of Intelligent & Robotic Systems*, 98 (3): 541-566.

Alsamhi, S. H., Ma, O. and Ansari, M. S. 2019. Survey on artificial intelligence based techniques for emerging robotic communication. *Telecommunication Systems*, 72 (3): 483-503.

Alsamhi, S. H., Ma, O., Ansari, M. S. and Meng, Q. 2019. Greening internet of things for greener and smarter cities: a survey and future prospects. *Telecommunication Systems*, 72 (4): 609-632.

Atashi, M., Salimibeni, M., Malekzadeh, P., Barbulescu, M., Plataniotis, K. N. and Mohammadi, A. 2019. Multiple model BLE-based tracking via validation of RSSI fluctuations under different conditions. In: *Proceedings of 2019 22th International Conference on Information Fusion (FUSION)*. Ottawa, ON, Canada, IEEE, 1-6.

Attarha, M. M., N. 2011. Focusing on the importance and the role of requirement engineering. In: *Proceedings of the 4th International Conference on Interaction Sciences (ICIS)*. Busan, Korea, 16-18 Aug. 2011. 181-184.

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.

Bai, Y. B., Gu, T. and Hu, A. 2016. Integrating Wi-Fi and magnetic field for fingerprinting based indoor positioning system. In: *Proceedings of 7th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Madrid, Spain, 4-7 October 2016. IEEE, 1-6.

Bai, Y. B., Wu, S., Wu, H. R. and Zhang, K. 2012. Overview of RFID-Based Indoor Positioning Technology. In: *Proceedings of Geospatial Science Research*. Melbourne, Australia, RMIT, 1-10.

Bal, M., Xue, H., Shen, W. and Ghenniwa, H. 2010. A 3-D indoor location tracking and visualization system based on wireless sensor networks. In: *Proceedings of International Conference on Systems Man and Cybernetics (SMC)*. Istanbul, Turkey, IEEE, 1584-1590.

Baniukevic, A., Jensen, C. S. and Lu, H. 2013. Hybrid indoor positioning with Wi-Fi and bluetooth: Architecture and performance. In: *Proceedings of 2013 IEEE 14th International Conference on Mobile Data Management*. Milan, Italy, IEEE, 207-216.

Baronti, P., Barsocchi, P., Chessa, S., Mavilia, F. and Palumbo, F. 2018. Indoor bluetooth low energy dataset for localization, tracking, occupancy, and social interaction. *Sensors*, 18 (12): 4462.

Barsocchi, P., Crivello, A., Girolami, M., Mavilia, F. and Palumbo, F. 2017. Occupancy detection by multi-power bluetooth low energy beaconing. In: *Proceedings of 8th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Sapporo, Japan, 18-21 September 2017. IEEE, 1-6.

Basiri, A., Lohan, E. S., Moore, T., Winstanley, A., Peltola, P., Hill, C., Amirian, P. and e Silva, P. F. 2017. Indoor location based services challenges, requirements and usability of current solutions. *Computer Science Review*, 24: 1-12.

Behboodi, A., Lemic, F., Wolisz, A. and Mathar, R. 2017. Interference effect on the performance of fingerprinting localization. In: *Proceedings of 2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Sapporo, Japan, IEEE, 1-8.

Bertuletti, S., Cereatti, A., Caldara, M., Galizzi, M. and Della Croce, U. 2016. Indoor distance estimated from Bluetooth Low Energy signal strength: Comparison of regression models. In: *Proceedings of 2016 IEEE Sensors Applications Symposium (SAS)*. Catania, Italy, IEEE, 1-5.

Booch, G., Rumbaugh, J. and Jacobson, I. 1999. *The unified modeling language user guide*. Reading, Massachusetts: Addison Wesley Longman, Inc.

Cantón Paterna, V., Calveras Auge, A., Paradells Aspas, J. and Perez Bullones, M. A. 2017. A bluetooth low energy indoor positioning system with channel diversity, weighted trilateration and kalman filtering. *Sensors*, 17 (12): 2927.

Chandel, V., Ahmed, N., Arora, S. and Ghose, A. 2016. InLoc: An end-to-end robust indoor localization and routing solution using mobile phones and BLE beacons. In: *Proceedings of Indoor Positioning and Indoor Navigation (IPIN), 2016 International Conference on*. Madrid, Spain, 4-7 October 2016. IEEE, 1-8.

Cheng, R.-S., Hong, W.-J., Wang, J.-S. and Lin, K. W. 2016. Seamless guidance system combining GPS, BLE Beacon, and NFC technologies. *Mobile Information Systems*, 2016 (1): 1-12.

Chicco, D., Warrens, M. J. and Jurman, G. 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7: e623.

Choudhary, R. and Gianey, H. K. 2017. Comprehensive review on supervised machine learning algorithms. In: *Proceedings of 2017 International Conference on Machine Learning and Data Science (MLDS)*. Noida, India, IEEE, 37-43.

Chriki, A., Touati, H. and Snoussi, H. 2017. SVM-based indoor localization in wireless sensor networks. In: *Proceedings of 2017 13th international wireless communications and mobile computing conference (IWCMC)*. Valencia, Spain, IEEE, 1144-1149.

Collotta, M., Pau, G., Talty, T. and Tonguz, O. K. 2018. Bluetooth 5: A concrete step forward toward the IoT. *IEEE Communications Magazine*, 56 (7): 125-131.

Cominelli, M., Gringoli, F., Patras, P., Lind, M. and Noubir, G. 2020. Even black cats cannot stay hidden in the dark: Full-band de-anonymization of bluetooth classic devices. In: *Proceedings of 2020 IEEE Symposium on Security and Privacy (SP)*. San Francisco, CA, USA, IEEE, 534-548.

Core Specification Working Group. 2021. Bluetooth Core Specification, v5. 3. BlueTooth SIG. Kirkland, WA, USA, 7, pp.1819-1845.

Cortes, C. and Vapnik, V. 1995. Support-vector networks. *Machine learning*, 20 (3): 273-297.

Cortez, P., Cerdeira, A., Almeida, F., Matos, T. and Reis, J. 2009. Modeling wine preferences by data mining from physicochemical properties. *Decision support systems*, 47 (4): 547-553.

Cottone, P., Gaglio, S., Re, G. L. and Ortolani, M. 2016. A machine learning approach for user localization exploiting connectivity data. *Engineering Applications of Artificial Intelligence*, 50: 125-134.

Creswell, J. W. 2013. *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, California: Sage publications.

D'Aloia, M., Cortone, F., Cice, G., Russo, R., Rizzi, M. and Longo, A. 2016. Improving energy efficiency in building system using a novel people localization system. In: *Proceedings of IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS)*. Bari, Italy, IEEE, 1-6.

Davis, W. S. and Yen, D. C. 2019. *The information system consultant's handbook: Systems analysis and design*. Boca Raton, Florida, USA: CRC press.

Deak, G., Curran, K. and Condell, J. 2012. A survey of active and passive indoor localisation systems. *Computer Communications*, 35 (16): 1939-1954.

Deng, Z., Yu, Y., Yuan, X., Wan, N. and Yang, L. 2013. Situation and development tendency of indoor positioning. *China Communications*, 10 (3): 42-55.

Dike, H. U., Zhou, Y., Deveerasetty, K. K. and Wu, Q. 2018. Unsupervised learning based on artificial neural network: A review. In: *Proceedings of 2018 IEEE International Conference on Cyborg and Bionic Systems (CBS)*. Shenzhen, China, IEEE, 322-327.

Ding, X., Hu, C., Xie, W., Zhong, Y., Yang, J. and Jiang, T. 2022. Device-Free Multi-Location Human Activity Recognition Using Deep Complex Network. *Sensors*, 22 (16): 6178.

Dwiyasa, F. and Lim, M.-H. 2016. A survey of problems and approaches in wireless-based indoor positioning. In: *Proceedings of Indoor Positioning and Indoor Navigation (IPIN), 2016 International Conference on*. Madrid, Spain, 4-7 October 2016. IEEE, 1-7.

Erl, T., Khattak, W. and Buhler, P. 2016. *Big data fundamentals: concepts, drivers and techniques*. New Jersey, USA: Prentice Hall Press.

Faragher, R. H., Robert. 2014. An analysis of the accuracy of bluetooth low energy for indoor positioning applications. In: *Proceedings of Proceedings of the 27th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2014), Tampa, FL, USA*. Tampa, FL, USA, September 2014. 2.

Faruqui, R. A. and Tayab, E. 2008. RFID: Importance, privacy/security issues and implementation. In: *Proceedings of International Symposium on Biometrics and Security Technologies ISBAST*. Islamabad, Pakistan, 23-24 April 2008. IEEE, 1-8.

Fisher, J. A. and Monahan, T. 2012. Evaluation of real-time location systems in their hospital contexts. *International journal of medical informatics*, 81 (10): 705-712.

Gao, Q., Wang, J., Ma, X., Feng, X. and Wang, H. 2017. CSI-based device-free wireless localization and activity recognition using radio image features. *IEEE Transactions on Vehicular Technology*, 66 (11): 10346-10356.

Germak, C. and Khan, S. 2017. Interaction design applications for museum spaces. New exhibit paths driven by a Bluetooth sensor's system. *The Design Journal*, 20 (sup1): S3914-S3924.

Gholamy, A., Kreinovich, V. and Kosheleva, O. 2018. *Why 70/30 or 80/20 relation between training and testing sets: a pedagogical explanation*. El Paso, Texas, USA: University of Texas, Department of Computer Science. Available: [https://scholarworks.utep.edu/cs\\_techrep/1209](https://scholarworks.utep.edu/cs_techrep/1209) (Accessed 16 June 2022).

Gikas, V., Dimitratos, A., Perakis, H., Retscher, G. and Ettliger, A. 2016. Full-scale testing and performance evaluation of an active RFID system for positioning and personal mobility. In: *Proceedings of 7th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Madrid, Spain, 4-7 October 2016. IEEE, 1-8.

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): 1055.

Gomes, R., Ahsan, M. and Denton, A. 2018. Random forest classifier in SDN framework for user-based indoor localization. In: *Proceedings of 2018 IEEE International Conference on Electro/Information Technology (EIT)*. Rochester, MI, USA, IEEE, 0537-0542.

Guba, E. G. L., Yvonna S. 1994. Competing paradigms in qualitative research. In: *Handbook of qualitative research*. Thousand Oaks, CA: Sage, 105-117.

Guimarães, V., Castro, L., Carneiro, S., Monteiro, M., Rocha, T., Barandas, M., Machado, J., Vasconcelos, M., Gamboa, H. and Elias, D. 2016. A motion tracking solution for indoor localization using smartphones. In: *Proceedings of 7th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Madrid, Spain, 4-7 October 2016. IEEE, 1-8.

Gulcharan, N. F. B. I., Daud, H., Nor, N. M., Ibrahim, T. and Nyamasvisva, E. T. 2013. Limitation and solution for healthcare network using RFID technology: a review. *Procedia Technology*, 11: 565-571.

Han, G., Klinker, G. J., Ostler, D. and Schneider, A. 2015. Testing a proximity-based location tracking system with Bluetooth Low Energy tags for future use in the OR. In: *Proceedings of 17th International Conference on E-health Networking, Application & Services (HealthCom)*. Boston, MA, USA, 14 - 17 October 2015. IEEE, 17-21.

Harishkumar, K., Yogesh, K. and Gad, I. 2020. Forecasting air pollution particulate matter (PM<sub>2.5</sub>) using machine learning regression models. *Procedia Computer Science*, 171: 2057-2066.

Hazar, M. J., Shaker, B. N., Ali, L. R. and Alzaidi, E. R. 2020. Using Received Strength Signal Indication for Indoor Mobile Localization Based on Machine Learning Technique. *Webology*, 17 (1): 30-42.

He, S. and Chan, S.-H. G. 2016. Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons. *IEEE Communications Surveys & Tutorials*, 18 (1): 466-490.

Hernández, A. C., Gómez, C., Crespo, J. and Barber, R. 2016. Object detection applied to indoor environments for mobile robot navigation. *Sensors*, 16 (8): 1180.

Hevner, V. A., R ; March, Salvatore T; Park, Jinsoo; Ram, Sudha. 2004. Design science in information systems research. *MIS quarterly*, 28 (1): 75-105.

Hsieh, J.-Y., Fan, C.-H., Liao, J.-Z., Hsu, J.-Y. and Chen, H. 2019. Study on the application of indoor positioning based on low power Bluetooth device combined with Kalman filter and machine learning. *EasyChair Preprint*: 1-9.

Huang, K., He, K. and Du, X. 2019. A hybrid method to improve the BLE-based indoor positioning in a dense bluetooth environment. *Sensors*, 19 (2): 424.

Hwang, K. and Chen, M. 2017. *Big-data analytics for cloud, IoT and cognitive computing*. Hoboken, New Jersey: John Wiley & Sons.

Hyder, Z., He, D., Li, D. and Yu, W. 2021. Improved Wi-Fi Indoor Localization Based on Signal Quality Parameters and RSSI Smoothing Algorithm. In: *Proceedings of 2021 10th International Conference on Information and Automation for Sustainability (ICIAfS)*. Negambo, Sri Lanka, IEEE, 140-144.

International Standards Organization (ISO). 2007. *Information technology - Automatic identification and data capture (AIDC) techniques - Harmonized vocabulary - Part 5: Locating systems*, ISO/IEC 19762-5. Geneva: ISO.

Iqbal, Z., Luo, D., Henry, P., Kazemifar, S., Rozario, T., Yan, Y., Westover, K., Lu, W., Nguyen, D. and Long, T. 2018. Accurate real time localization tracking in a clinical environment using Bluetooth Low Energy and deep learning. *PloS one*, 13 (10): e0205392.

Jachimczyk, B., Dziak, D. and Kulesza, W. J. 2016. Using the Fingerprinting Method to Customize RTLS Based on the AoA Ranging Technique. *Sensors*, 16 (6): 876.

Jain, C., Sashank, G. V. S. and Markkandan, S. 2021. Low-cost BLE based Indoor Localization using RSSI Fingerprinting and Machine Learning. In: *Proceedings of 2021 Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*. Chennai, India, IEEE, 363-367.

Janicka, J. and Rapinski, J. 2016. Application of RSSI based navigation in indoor positioning. In: Proceedings of *2016 Baltic Geodetic Congress (BGC Geomatics)*. Gdansk, Poland, IEEE, 45-50.

Jayakody, J. A., Lokuliyana, S., Chathurangi, D. and Vithana, D. 2016. Indoor positioning: Novel approach for Bluetooth networks using RSSI smoothing. *International Journal of Computer Applications*, 137 (13): 26-32.

Jeon, K. E., She, J., Soonsawad, P. and Ng, P. C. 2018. Ble beacons for internet of things applications: Survey, challenges, and opportunities. *IEEE Internet of Things Journal*, 5 (2): 811-828.

Johnson, R. B. and Christensen, L. 2019. *Educational research: Quantitative, qualitative, and mixed approaches*. Thousand Oaks, California: Sage publications.

Joshua, V., Priyadharson, S. M. and Kannadasan, R. 2021. Exploration of machine learning approaches for paddy yield prediction in eastern part of Tamilnadu. *Agronomy*, 11 (10): 2068.

Juri, A. A., Arslan, T. and Wang, F. 2016. Obstruction-aware bluetooth low energy indoor positioning. In: Proceedings of *Proceedings of the 29th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2016)*. Portland, Oregon, September 12 - 16, 2016. ION, 2254-2261.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems.

Kao, C.-H., Hsiao, R.-S., Chen, T.-X., Chen, P.-S. and Pan, M.-J. 2017. A hybrid indoor positioning for asset tracking using Bluetooth low energy and Wi-Fi. In: Proceedings of *2017 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*. Taipei, Taiwan, IEEE, 63-64.

Kaplan, B. and Maxwell, J. A. 2005. Qualitative research methods for evaluating computer information systems. In: *Evaluating the organizational impact of healthcare information systems*. New York: Springer, 30-55.

Khan, A., Khan, A., Bangash, J. I., Subhan, F., Khan, A., Khan, A., Uddin, M. I. and Mahmoud, M. 2021. Cuckoo Search-based SVM (CS-SVM) Model for Real-Time Indoor Position Estimation in IoT Networks. *Security and Communication Networks*, 2021: 1-7.

Khatab, Z. E., Hajihoseini, A. and Ghorashi, S. A. 2017. A fingerprint method for indoor localization using autoencoder based deep extreme learning machine. *IEEE sensors letters*, 2 (1): 1-4.

Kim, S. 2013. iScholar: A mobile research support system. M.Sc, University of Regina.

Kim, S., Ha, S., Saad, A. and Kim, J. 2015. Indoor positioning system techniques and security. In: *Proceedings of Fourth International Conference on e-Technologies and Networks for Development (ICeND)*. Lodz, Poland, 21 - 23 September 2015. IEEE, 1-4.

Kjærgaard, M. B., Krarup, M. V., Stisen, A., Prentow, T. S., Blunck, H., Grønbaek, K. and Jensen, C. S. 2013. Indoor positioning using wi-fi—how well is the problem understood? In: *Proceedings of 4th International Conference on Indoor Positioning and Indoor Navigation*. Montbéliard-Belfort, France, 28-31 October 2013. IEEE,

Kolodziej, K. W. and Hjelm, J. 2017. *Local positioning systems: LBS applications and services*. Boca Raton: CRC Press.

Kumar, S., Gil, S., Katabi, D. and Rus, D. 2014. Accurate indoor localization with zero start-up cost. In: *Proceedings of Proceedings of the 20th annual international conference on Mobile computing and networking*. Maui, awaii, 7-11 September 2014. ACM, 483-494.

Le, D. V. and Havinga, P. J. 2017. SoLoc: Self-organizing indoor localization for unstructured and dynamic environments. In: *Proceedings of 8th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Sapporo, Japan, 18-21 September 2017. IEEE, 1-8.

Lee, C. K. M., Ip, C., Park, T. and Chung, S. 2019. A Bluetooth Location-based Indoor Positioning System for Asset Tracking in Warehouse. In: *Proceedings of 2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. Macao, china, 15-19 December 2019. IEEE, 1408-1412.

Lee, H. W., Jeon, W. S. and Jeong, D. G. 2017. A practical indoor localization scheme for disaster relief. In: *Proceedings of 2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*. Sydney, NSW, Australia, IEEE, 1-7.

Lee, J.-S., Dong, M.-F. and Sun, Y.-H. 2015. A preliminary study of low power wireless technologies: ZigBee and Bluetooth low energy. In: *Proceedings of 10th Conference on Industrial Electronics and Applications (ICIEA)*. Auckland, New Zealand, 15 - 17 June 2015. IEEE, 135-139.

Lee, J., Park, K. and Kim, Y. 2022. Deep Learning-Based Device-Free Localization Scheme for Simultaneous Estimation of Indoor Location and Posture Using FMCW Radars. *Sensors*, 22 (12): 4447.

Li, W., Wei, D., Yuan, H. and Ouyang, G. 2016. A novel method of WiFi fingerprint positioning using spatial multi-points matching. In: *Proceedings of 7th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Madrid, Spain, 4-7 October 2016. IEEE, 1-8.

Li, Y.-F. and Liang, D.-M. 2019. Safe semi-supervised learning: a brief introduction. *Frontiers of Computer Science*, 13 (4): 669-676.

Li, Z., Xu, K., Wang, H., Zhao, Y., Wang, X. and Shen, M. 2019. Machine-learning-based positioning: A survey and future directions. *IEEE Network*, 33 (3): 96-101.

Lim, M. K., Bahr, W. and Leung, S. C. 2013. RFID in the warehouse: A literature analysis (1995–2010) of its applications, benefits, challenges and future trends. *International Journal of Production Economics*, 145 (1): 409-430.

Lincoln, Y. S. L., Susan A; Guba, Egon G. 2011. Paradigmatic controversies, contradictions, and emerging confluences, revisited. In: Denzin, N. K. L., Yvonna S ed. *The Sage Handbook of Qualitative Research*. Los Angeles: Sage, 97-128.

Luo, J., Zhang, Z., Wang, C., Liu, C. and Xiao, D. 2019. Indoor multifloor localization method based on WiFi fingerprints and LDA. *IEEE Transactions on Industrial Informatics*, 15 (9): 5225-5234.

Mahfouz, S., Nader, P. and Abi-Char, P. E. 2020. RSSI-based classification for indoor localization in wireless sensor networks. In: *Proceedings of 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*. Doha, Qatar, IEEE, 323-328.

Maimon, O. and Rokach, L. 2005. *Data mining and knowledge discovery handbook*.

Mainetti, L., Patrono, L. and Sergi, I. 2014. A survey on indoor positioning systems. In: *Proceedings of 22nd International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. Split, Croatia, 17-19 September 2014. IEEE, 111-120.

Mathisen, A., Sørensen, S. K., Stisen, A., Blunck, H. and Grønbaek, K. 2016. A comparative analysis of Indoor WiFi Positioning at a large building complex. In: *Proceedings of 7th*

*International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Madrid, Spain, 4-7 October 2016. IEEE, 1-8.

Mehrad, A. and Zangeneh, M. H. T. 2019. Comparison between qualitative and quantitative research approaches: Social sciences. *International Journal For Research In Educational Studies, Iran*: 1-7.

Mendoza-Silva, G. M., Torres-Sospedra, J. and Huerta, J. 2017. A more realistic error distance calculation for indoor positioning systems accuracy evaluation. In: *Proceedings of 8th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Sapporo, Japan, 18-21 September 2017. IEEE, 1-8.

Mertens, D. M. 2015. *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods*. 4th ed. Thousand Oaks, Calif.: Sage.

Milicev, D. 2009. *Model-driven development with executable UML*. Hoboken, New Jersey: John Wiley & Sons.

Moghtadaiee, V. and Dempster, A. G. 2012. WiFi fingerprinting signal strength error modeling for short distances. In: *Proceedings of 3rd International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Sydney, Australia, 13-15 November 2012. IEEE, 1-6.

Moreira, A., Silva, I. M. M., Meneses, F., Nicolau, M. J., Pendao, C. and Torres-Sospedra, J. 2017. Multiple simultaneous Wi-Fi measurements in fingerprinting indoor positioning. In: *Proceedings of 8th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Sapporo, Japan, 18-21 September 2017.

Naghdi, S. and O'Keefe, K. 2020. Detecting and correcting for human obstacles in BLE trilateration using artificial intelligence. *Sensors*, 20 (5): 1350.

Neishaboori, A. and Harras, K. 2013. Energy saving strategies in WiFi indoor localization. In: *Proceedings of 16th ACM International Conference On Modeling, Analysis & Simulation Of Wireles and Mobile Systems*. Malta, Malta, 13-17 November 2016 ACM, 399-404.

Nessa, A., Adhikari, B., Hussain, F. and Fernando, X. N. 2020. A survey of machine learning for indoor positioning. *IEEE Access*, 8: 214945-214965.

Nikoukar, A., Abboud, M., Samadi, B., Güneş, M. and Dezfouli, B. 2018. Empirical analysis and modeling of Bluetooth low-energy (BLE) advertisement channels. In: *Proceedings of 2018*

17th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net). Capri, Italy, IEEE, 1-6.

Nordic. 2018. Nordic nRF52840 PDK data sheet. Nordic.

Nunamaker Jr, J. F. and Chen, M. 1990. Systems development in information systems research. In: *Proceedings of System Sciences, 1990., Proceedings of the Twenty-Third Annual Hawaii International Conference on.* IEEE, 631-640.

Nunamaker Jr, J. F., Chen, M. and Purdin, T. D. 1990. Systems development in information systems research. *Journal of management information systems*, 7 (3): 89-106.

Otchere, D. A., Ganat, T. O. A., Gholami, R. and Ridha, S. 2021. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. *Journal of Petroleum Science and Engineering*, 200: 108182.

Paiva, S., Brito, D. and Leiva-Marcon, L. 2018. Real Time Location Systems Adoption in Hospitals—A review and a case study for locating assets. *Acta Scientific Medical Sciences*, 2 (7): 02-17.

Pancham, J., Millham, R. and Fong, S. J. 2017. Evaluation of real time location system technologies in the health care sector. In: *Proceedings of 17th International Conference on Computational Science and its Applications (ICCSA)*. Trieste, Italy, 3-6 July 2017. IEEE, 1-7.

Pancham, J., Millham, R. and Fong, S. J. 2018. A scalable bluetooth low energy design model for sensor detection for an indoor real time location system. In: *Proceedings of International Conference on Computational Science and Its Applications*. Melbourne, Australia, 2-5 July 2018. Springer, 317-330.

Pancham, J., Millham, R. and Fong, S. J. 2020. Analysis of Bluetooth Low Energy RSSI Values for Use as a Real Time Link Quality Indicator for Indoor Location. In: *Proceedings of International Conference on Computational Science and Its Applications*. Springer, 980-991.

Pang, B., T'Jonck, K., Claeys, T., Pissoort, D., Hallez, H. and Boydens, J. 2021. Bluetooth low energy interference awareness scheme and improved channel selection algorithm for connection robustness. *Sensors*, 21 (7): 2257.

Peffer, K., Tuunanen, T., Gengler, C. E., Rossi, M., Hui, W., Virtanen, V. and Bragge, J. 2006. The design science research process: a model for producing and presenting information

systems research. In: Proceedings of *the First International Conference on Design Science Research in Information Systems and Technology (DESRIST 2006)*. Claremont, California, 24-25 February 2006. DESRIST, 83-106.

Perry, A. and Kocakülâh, M. 2010. Healthcare supply chain leadership and strategy: how it can help cost containment in the healthcare industry. *Cost Management*, 24 (2): 37.

Petty, N. J., Thomson, O. P. and Stew, G. 2012. Ready for a paradigm shift? Part 1: introducing the philosophy of qualitative research. *Manual therapy*, 17 (4): 267-274.

Potisek, N. M., Malone, R. M., Shilliday, B. B., Ives, T. J., Chelminski, P. R., DeWalt, D. A. and Pignone, M. P. 2007. Use of patient flow analysis to improve patient visit efficiency by decreasing wait time in a primary care-based disease management programs for anticoagulation and chronic pain: a quality improvement study. *BMC health services research*, 7 (1): 8.

Powar, J., Gao, C. and Harle, R. 2017. Assessing the impact of multi-channel BLE beacons on fingerprint-based positioning. In: Proceedings of *8th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Sopparo, Japan, 18-21 September 2017. IEEE, 1-8.

Raza, S., Misra, P., He, Z. and Voigt, T. 2016. Building the internet of things with bluetooth smart. *Ad Hoc Networks*, 57: 19-31.

Reddy, Y., Viswanath, P. and Reddy, B. E. 2018. Semi-supervised learning: A brief review. *Int. J. Eng. Technol*, 7 (1.8): 81.

Reyes, P. M. L., Suhong; Visich, John K. 2012. Accessing antecedents and outcomes of RFID implementation in health care. *International Journal of Production Economics*, 136 (1): 137-150.

Roy, P. and Chowdhury, C. 2021. A survey of machine learning techniques for indoor localization and navigation systems. *Journal of Intelligent & Robotic Systems*, 101 (3): 1-34.

Sabancı, K., Yigit, E., Ustun, D., Toktas, A. and Aslan, M. F. 2018. Wifi based indoor localization: application and comparison of machine learning algorithms. In: Proceedings of *2018 XXIIIrd International Seminar/Workshop on Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory (DIPED)*. Tbilisi, Georgia, IEEE, 246-251.

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

Salamah, A. H., Tamazin, M., Sharkas, M. A., Khedr, M. and Mahmoud, M. 2019. Comprehensive investigation on principle component large-scale Wi-Fi indoor localization. *Sensors*, 19 (7): 1678.

Santos, R. X. M. and Krishnan, S. 2020. Improving Single Reference Indoor Positioning Accuracy Through Machine Learning. In: Proceedings of 2020 IEEE REGION 10 CONFERENCE (TENCON). Osaka, Japan, IEEE, 484-489.

Schultz, B. 2016. From cable replacement to the IoT Bluetooth 5 (White Paper). Rohde & Schwarz. Available: [https://bluetoothworldevent.com/media/PDFs/Rohde--26-Schwarz\\_3e\\_Bluetooth\\_WhitePaper.pdf](https://bluetoothworldevent.com/media/PDFs/Rohde--26-Schwarz_3e_Bluetooth_WhitePaper.pdf) (Accessed 7 January 2018).

Shang, S. and Wang, L. 2022. Overview of WiFi fingerprinting-based indoor positioning. *IET Communications*, 16 (7): 725-733.

Shirehjini, A. A. N., Yassine, A. and Shirmohammadi, S. 2012. Equipment location in hospitals using RFID-based positioning system. *IEEE Transactions on information technology in biomedicine*, 16 (6): 1058-1069.

Shukla, S. and Bergmann, N. W. 2004. Single bit error correction implementation in CRC-16 on FPGA. In: Proceedings of Proceedings. 2004 IEEE International Conference on Field-Programmable Technology (IEEE Cat. No. 04EX921). IEEE, 319-322.

Song, M. and Qian, J. 2016. Improved sequence-based localization applied in coal mine. *International Journal of Distributed Sensor Networks*, 12 (11): 1-11.

Sthapit, P., Gang, H.-S. and Pyun, J.-Y. 2018. Bluetooth based indoor positioning using machine learning algorithms. In: Proceedings of 2018 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia). JeJu, Korea (South), IEEE, 206-212.

Su, J., Sheng, Z., Xie, L., Li, G. and Liu, A. X. 2018. Fast splitting-based tag identification algorithm for anti-collision in UHF RFID system. *IEEE Transactions on Communications*, 67 (3): 2527-2538.

Sun, D., Wei, E., Ma, Z., Wu, C. and Xu, S. 2021. Optimized cnns to indoor localization through ble sensors using improved pso. *Sensors*, 21 (6): 1995.

Sun, H., Zhu, X., Liu, Y. and Liu, W. 2020. Construction of hybrid dual radio frequency RSSI (HDRF-RSSI) fingerprint database and indoor location method. *Sensors*, 20 (10): 2981.

Tănase, C.-A. and Vulpe, A. 2022. Indoor Location Monitoring and Navigation System for Hospitals. In: Proceedings of 2022 14th International Conference on Communications (COMM). Bucharest, Romania, IEEE, 1-6.

Taşkan, A. K. and Alemdar, H. 2021. Obstruction-Aware Signal-Loss-Tolerant Indoor Positioning Using Bluetooth Low Energy. *Sensors*, 21 (3): 971.

Terán, M., Aranda, J., Carrillo, H., Mendez, D. and Parra, C. 2017. IoT-based system for indoor location using bluetooth low energy. In: Proceedings of 2017 IEEE Colombian Conference on Communications and Computing (COLCOM). Cartagena, Colombia, IEEE, 1-6.

Terán, M., Carrillo, H. and Parra, C. 2018. Wlan-ble based indoor positioning system using machine learning cloud services. In: Proceedings of 2018 IEEE 2nd Colombian Conference on Robotics and Automation (CCRA). Barranquilla, Colombia, IEEE, 1-6.

Thaljaoui, A., Val, T., Nasri, N. and Brulin, D. 2015. BLE localization using RSSI measurements and iRingLA. In: Proceedings of 2015 IEEE International Conference on Industrial Technology (ICIT). Seville, Spain, 17 - 19 March 2015. IEEE, 2178-2183.

Tsang, P., Wu, C., Ip, W., Ho, G. and Tse, Y. 2015. A bluetooth-based indoor positioning system: a simple and rapid approach. *Annual Journal IIE (HK)*, 35 (2014/2015): 11-26.

Uddin, S., Khan, A., Hossain, M. E. and Moni, M. A. 2019. Comparing different supervised machine learning algorithms for disease prediction. *BMC medical informatics and decision making*, 19 (1): 1-16.

Wamba, S. F., Anand, A. and Carter, L. 2013. A literature review of RFID-enabled healthcare applications and issues. *International Journal of Information Management*, 33 (5): 875-891.

Wang, B., Toobaei, M., Danskin, R., Ngarmnil, T., Pham, L. and Pham, H. 2013. Evaluation of RFID and Wi-Fi technologies for RTLS applications in Healthcare Centers. In: Proceedings of *Technology Management in the IT-Driven Services (PICMET)*, 2013 Proceedings of PICMET'13: . San Jose, CA, USA, IEEE, 2690-2703.

Wang, X., Gao, L., Mao, S. and Pandey, S. 2016. CSI-based fingerprinting for indoor localization: A deep learning approach. *IEEE Transactions on Vehicular Technology*, 66 (1): 763-776.

Wang, Y., Ye, Q., Cheng, J. and Wang, L. 2015. RSSI-based bluetooth indoor localization. In: *Proceedings of 11th International Conference on Mobile Ad-hoc and Sensor Networks (MSN)*. Shenzhen, China, 16 – 18 December 2015. IEEE, 165-171.

Wieringa, R. 2010. Design science methodology: principles and practice. In: *Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering-Volume 2*. Cape Town, South Africa, 01-08 May. New York: ACM, 493-494.

Wirola, L., Laine, T. A. and Syrjärinne, J. 2010. Mass-market requirements for indoor positioning and indoor navigation. In: *Proceedings of International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Zurich, Switzerland, 15-17 September 2010. IEEE, 1-7.

Woolley, M. 2019. *Bluetooth core specification v5. 1*.

Wu, M., Tan, L. and Xiong, N. 2016. Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications. *Information Sciences*, 329: 800-818.

Wu, Z., Jedari, E., Muscedere, R. and Rashidzadeh, R. 2016. Improved particle filter based on WLAN RSSI fingerprinting and smart sensors for indoor localization. *Computer Communications*, 83: 64-71.

Wu, Z., Liu, J. and Liu, B. 2016. Particle filter and support vector machine based indoor localization system.

Xia, Y. 2019. A novel reject inference model using outlier detection and gradient boosting technique in peer-to-peer lending. *IEEE Access*, 7: 92893-92907.

Xiao, J., Wu, K., Yi, Y., Wang, L. and Ni, L. M. 2013. Pilot: Passive device-free indoor localization using channel state information. In: *Proceedings of 2013 IEEE 33rd International Conference on Distributed Computing Systems*. Philadelphia, PA, USA, IEEE, 236-245.

Yang, J., Poellabauer, C., Mitra, P. and Neubecker, C. 2020. Beyond beaconing: Emerging applications and challenges of BLE. *Ad Hoc Networks*, 97: 102015.

Yang, Z., Zhou, Z. and Liu, Y. 2013. From RSSI to CSI: Indoor localization via channel response. *ACM Computing Surveys (CSUR)*, 46 (2): 1-32.

Yao, W., Chu, C.-H. and Li, Z. 2012. The adoption and implementation of RFID technologies in healthcare: a literature review. *Journal of Medical Systems*, 36 (6): 3507-3525.

Yazici, H. J. 2014. An exploratory analysis of hospital perspectives on real time information requirements and perceived benefits of RFID technology for future adoption. *International Journal of Information Management*, 34 (5): 603-621.

Yin, A. and Lin, Z. 2022. Machine Learning aided Precise Indoor Positioning. *arXiv:2204.03990*, 1: 1-8.

Yoo, J. 2019. Change detection of RSSI fingerprint pattern for indoor positioning system. *IEEE Sensors Journal*, 20 (5): 2608-2615.

You, Y. and Wu, C. 2019. Indoor positioning system with cellular network assistance based on received signal strength indication of beacon. *IEEE Access*, 8: 6691-6703.

Yu, B., Xu, L. and Li, Y. 2012. Bluetooth Low Energy (BLE) based mobile electrocardiogram monitoring system. In: Proceedings of *International Conference on Information and Automation (ICIA)*. Shenyang, China, 20-24 August 2012. IEEE, 763-767.

Yuan, Z., Zhang, X., Zhou, P. and Wang, S. 2019. Research on Indoor Position Fingerprint Location Based on Machine Learning combined Particle Filter. In: Proceedings of *2019 2nd International Conference on Safety Produce Informatization (IICSPI)*. Chongqing, China, IEEE, 456-459.

YuFeng, JiangMinghua, LiangJing, QinXiao, HuMing, PengTao and HuXinrong. 2014. An improved indoor localization of wifibased on support vector machines. *International Journal of Future Generation Communication and Networking*, 7 (5): 191-206.

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.

Zaim, D. and Bellafkih, M. 2016. Bluetooth Low Energy (BLE) based geomarketing system. In: Proceedings of *11th International Conference on Intelligent Systems: Theories and Applications (SITA)*. Mohammedia, Morocco, 19-20 October 2016. IEEE, 1-6.

Zhang, Y., Qu, C. and Wang, Y. 2020. An indoor positioning method based on CSI by using features optimization mechanism with LSTM. *IEEE Sensors Journal*, 20 (9): 4868-4878.

Zhu, X., Mukhopadhyay, S. K. and Kurata, H. 2012. A review of RFID technology and its managerial applications in different industries. *Journal of Engineering and Technology Management*, 29 (1): 152-167.

## APPENDICES

### Appendix A - Use Cases

Only the main use cases are detailed.

Category	Description
<b>No</b>	UC01
<b>Name</b>	LPN broadcast message
<b>Related Requirements</b>	
<b>Goal in Context</b>	Send a message to the server
<b>Trigger</b>	A message will be broadcast based on and of the following event. <ol style="list-style-type: none"> <li>1. Movement detected</li> <li>2. Time to broadcast reached</li> </ol>
<b>Preconditions</b>	
<b>Main flow action steps</b>	<ol style="list-style-type: none"> <li>1. If movement detected (LNM) OR time elapsed (LNT) then               <ol style="list-style-type: none"> <li>1.1 Create a new message</li> <li>1.2 Broadcast</li> </ol> </li> </ol>
<b>Exception</b>	
<b>Post Conditions</b>	
<b>Business rules</b>	
<b>System rules</b>	

Category	Description
<b>No</b>	UC02
<b>Name</b>	LPN receives message
<b>Related Requirements</b>	
<b>Goal in Context</b>	Receive and process a message
<b>Trigger</b>	A message will be received from an LPN, Relay or the server.

<b>Preconditions</b>	
<b>Main flow action steps</b>	<p>If message type is (LNM, LNT) or (CPN not for itself) then</p> <ol style="list-style-type: none"> <li>1. On receipt of message check and record Receive level,</li> <li>2. Check TTL</li> <li>3. Update Message</li> <li>4. If TTL &gt; 0 then Broadcast</li> </ol> <p>Else if Message type is CPN for itself</p> <ol style="list-style-type: none"> <li>1. Lookup and setup parameters</li> </ol>
<b>Exception</b>	
<b>Post Conditions</b>	
<b>Business rules</b>	
<b>System rules</b>	<ol style="list-style-type: none"> <li>1. LPN will broadcast messages received from other LPN's provided the TTL is &gt; 0 and the message is not repeated.</li> <li>2. Ignore message broadcast from itself</li> </ol>

<b>Category</b>	<b>Description</b>
<b>No</b>	UC03
<b>Name</b>	Relay broadcast message
<b>Related Requirements</b>	
<b>Goal in Context</b>	Transmit a message to the server
<b>Trigger</b>	<p>A message will be broadcast based on and of the following event.</p> <ol style="list-style-type: none"> <li>1. Time to broadcast reached</li> </ol>
<b>Preconditions</b>	
<b>Main flow action steps</b>	<p>If time to broadcast reached, then</p> <ol style="list-style-type: none"> <li>1. Create a new message</li> <li>2. Broadcast</li> </ol>
<b>Exception</b>	
<b>Post Conditions</b>	
<b>Business rules</b>	
<b>System rules</b>	

<b>Category</b>	<b>Description</b>
-----------------	--------------------

<b>No</b>	UC04
<b>Name</b>	Relay receives message
<b>Related Requirements</b>	
<b>Goal in Context</b>	Receive and process a message
<b>Trigger</b>	A message will be received from the server, relay or LPN
<b>Preconditions</b>	
<b>Main flow action steps</b>	<p>If message type is CRL and is for itself, then</p> <ol style="list-style-type: none"> <li>1. On receipt of message check and record Receive level, and TTL</li> </ol> <p>ELSE for all other messages</p> <ol style="list-style-type: none"> <li>1. Record Receive level and update the TTL</li> <li>2. Send message</li> </ol>
<b>Exception</b>	
<b>Post Conditions</b>	
<b>Business rules</b>	
<b>System rules</b>	

Category	Description
No	UC05
Name	Server send message
Related Requirements	
Goal in Context	Server sends out configuration data or message to check status of a relay or LPN
Trigger	A message is sent on one of the two events <ul style="list-style-type: none"> <li>1. New configuration data setup</li> <li>2. Timeout interval reached for relay or LPN.</li> </ul>
Preconditions	
Main flow action steps	<p>If new configuration data is setup</p> <ul style="list-style-type: none"> <li>1. Create a new message (CSR or CSL)</li> <li>2. Send message</li> <li>3. Function receives config</li> <li>4. Get node ID using the MAC address</li> <li>5. Send config to Node using node ID</li> </ul> <p>If time elapsed, then broadcast own message</p> <ul style="list-style-type: none"> <li>1. Create a new message (CSN or CSR)</li> <li>2. Send message</li> <li>3. Function receives config</li> <li>4. Get node ID using the MAC address</li> <li>5. Send config to Node using node ID</li> </ul>
Exception	
Post Conditions	
Business rules	<p>Parameters to be sent from server to Low power node / or Relay</p> <ul style="list-style-type: none"> <li>1. X – Position decimal one decimal place precision</li> <li>2. Y - Position decimal one decimal place precision</li> <li>3. Venue – character field to indicate the location of the test</li> <li>4. Broadcast – Off: Stop broadcast, On: Begin Broadcast. This parameter is ignored for a Relay</li> </ul>
Systems	
Category	Description
No	UC06
Name	Server receives message

<b>Related Requirements</b>	
<b>Goal in Context</b>	Store message received
<b>Trigger</b>	Server receives a message
<b>Preconditions</b>	
<b>Main flow action steps</b>	If Message type is (LNM, LNT, CSN, CSR)  1. Store message
<b>Exception</b>	
<b>Post Conditions</b>	
<b>Business rules</b>	
<b>System rules</b>	

## Appendix B - Measurement of Throughput

Typical interaction between the PDKs used as transmitter and receiver.

```
[00002546] <info> app: This board will act as tester.
[00002546] <info> app: Type 'config' to change the configuration parameters.
[00002546] <info> app: You can use the TAB key to autocomplete your input.
[00002547] <info> app: Type 'run' when you are ready to run the test.
throughput example:~$ run
[00000004] <info> app: Preparing the test.
[00000006] <info> app: Starting advertising.
[00000011] <info> app: Starting scan.
[00001458] <info> app: Device "Nordic_ATT_MTU" found, sending a connection
request.
[00001466] <info> app: RSSI: -60
[00001466] <info> app: Advertised TX: 100
[00002836] <info> app: Connected as a central.
[00002837] <info> app: Discovering GATT database...
[00000084] <info> app: ATT MTU exchange completed. MTU set to 247 bytes.
[00000330] <info> app: Data length updated to 251 bytes.
```

[00002471] <info> app: PHY update accepted. PHY set to 2 Mbps.  
[00003030] <info> app: AMT service discovered at peer.  
[00003276] <info> app: Notifications enabled.  
[00000172] <info> app: Test started  
[00001552] <info> app: Done.  
[00001553] <info> app: =====  
[00001553] <info> app: Time: 8.933 seconds elapsed.  
[00001553] <info> app: Throughput: 939.18 Kbps.  
[00001554] <info> app: =====  
[00001554] <info> app: Sent 1048712 bytes of ATT payload.  
[00001555] <info> app: Retrieving amount of bytes received from peer...  
[00002212] <info> app: Peer received 1048712 bytes of ATT payload.  
[00002212] <info> app: Disconnecting...

## Appendix C – Results from Perspex Obstruction

Results when no screens are placed

		No Screens						
		Test 1		Test 2		Test3		Avg
GW	LPN	Count	Level	Count	Level	Count	Level	
GW-001	LPN001	85	-60	111	-59	110	-59	-57
GW-001	LPN001	73	-55	81	-55	86	-55	
GW-001	LPN001	62	-58	78	-58	57	-58	
GW-001	LPN002	83	-64	97	-64	86	-64	-64
GW-001	LPN002	65	-60	55	-61	65	-67	
GW-001	LPN002	52	-67	52	-67	52	-61	
GW-002	LPN001	98	-56	102	-56	87	-56	-60
GW-002	LPN001	61	-70	62	-54	80	-54	
GW-002	LPN001	55	-54	37	-69	48	-69	
GW-002	LPN002	106	-63	88	-63	86	-63	-63
GW-002	LPN002	50	-64	60	-61	51	-61	
GW-002	LPN002	50	-61	38	-65	48	-62	
GW-003	LPN001	79	-69	72	-69	67	-68	-68
GW-003	LPN001	43	-68	55	-67	60	-69	
GW-003	LPN001	43	-66	43	-68	57	-67	
GW-003	LPN002	102	-44	117	-46	110	-46	-47

GW-003	LPN002	91	-46	94	-44	104	-44	
GW-003	LPN002	64	-50	53	-50	52	-50	
GW-004	LPN001	61	-68	109	-68	63	-67	-72
GW-004	LPN001	53	-78	68	-78	54	-68	
GW-004	LPN001	48	-67	35	-79	54	-78	
GW-004	LPN002	100	-52	108	-52	105	-52	-56
GW-004	LPN002	66	-57	68	-57	80	-60	
GW-004	LPN002	59	-59	46	-60	67	-57	

Results for 3mm Perspex

		3 mm Screen						
		Test 1		Test 2		Test3		Avg
GW	LPN	Count	Level	Count	Level	Count	Level	
GW-001	LPN001	84	-58	94	-58	109	-58	-57
GW-001	LPN001	75	-55	55	-55	62	-57	
GW-001	LPN001	64	-59	53	-59	57	-55	
GW-001	LPN002	72	-64	64	-64	76	-64	-64
GW-001	LPN002	53	-60	51	-66	55	-61	
GW-001	LPN002	36	-66	50	-61	48	-66	
GW-002	LPN001	100	-56	78	-56	80	-56	-57
GW-002	LPN001	61	-54	53	-54	64	-54	
GW-002	LPN001	41	-69	48	-55	62	-55	
GW-002	LPN002	71	-61	88	-63	81	-63	-63
GW-002	LPN002	68	-63	68	-61	76	-61	
GW-002	LPN002	47	-64	46	-64	52	-64	
GW-003	LPN001	72	-68	95	-68	75	-67	-68
GW-003	LPN001	52	-67	49	-71	67	-68	
GW-003	LPN001	40	-69	40	-67	47	-71	
GW-003	LPN002	112	-46	111	-46	109	-46	-47
GW-003	LPN002	99	-44	97	-44	101	-44	

GW-003	LPN002	54	-51	67	-51	68	-51	
GW-004	LPN001	63	-67	76	-68	56	-68	-71
GW-004	LPN001	51	-68	46	-67	45	-78	
GW-004	LPN001	47	-78	45	-78	37	-67	
GW-004	LPN002	63	-67	75	-52	75	-58	-62
GW-004	LPN002	51	-68	58	-58	63	-52	
GW-004	LPN002	47	-78	43	-60	46	-61	

Results from 5mm Perspex

		5 mm Screen						
		Test 1		Test 2		Test3		Avg
GW	LPN	Count	Level	Count	Level	Count	Level	
GW-001	LPN001	154	-58	149	-58	123	-58	-57
GW-001	LPN001	51	-55	55	-57	74	-57	
GW-001	LPN001	45	-57	53	-55	55	-55	
GW-001	LPN002	80	-64	72	-64	93	-56	-61
GW-001	LPN002	52	-61	65	-66	54	-55	
GW-001	LPN002	48	-66	64	-61	51	-54	
GW-002	LPN001	96	-56	86	-56	93	-56	-57
GW-002	LPN001	69	-54	60	-54	54	-55	
GW-002	LPN001	40	-69	56	-55	51	-54	
GW-002	LPN002	77	-63	77	-63	79	-62	-63
GW-002	LPN002	61	-64	59	-64	68	-61	
GW-002	LPN002	61	-61	58	-61	49	-64	
GW-003	LPN001	60	-67	98	-67	61	-68	-68
GW-003	LPN001	50	-68	52	-68	55	-67	
GW-003	LPN001	45	-66	41	-72	47	-73	
GW-003	LPN002	102	-46	114	-46	106	-46	-47
GW-003	LPN002	81	-44	91	-44	92	-44	

GW-003	LPN002	62	-51	64	-51	48	-51	
GW-004	LPN001	68	-67	80	-67	81	-67	-74
GW-004	LPN001	48	-78	52	-78	42	-77	
GW-004	LPN001	47	-77	31	-77	34	-78	
GW-004	LPN002	70	-52	69	-52	75	-52	-56
GW-004	LPN002	65	-58	65	-58	52	-58	
GW-004	LPN002	43	-53	50	-61	49	-61	

Results from both 3mm and 5mm Perspex

		3 and 5 mm Screen						
		Test 1		Test 2		Test3		Avg
GW	LPN	Count	Level	Count	Level	Count	Level	
GW-001	LPN001	171	-58	157	-58	155	-58	-55.9
GW-001	LPN001	40	-55	48	-55	48	-57	
GW-001	LPN001	33	-54	46	-54	45	-54	
GW-001	LPN002	73	-61	67	-66	77	-65	-64.0
GW-001	LPN002	67	-65	67	-65	74	-61	
GW-001	LPN002	39	-66	54	-61	41	-66	
GW-002	LPN001	104	-56	90	-56	94	-56	-57.9
GW-002	LPN001	58	-68	66	-54	57	-55	
GW-002	LPN001	55	-54	46	-68	46	-54	
GW-002	LPN002	112	-63	93	-63	107	-63	-62.3
GW-002	LPN002	57	-60	58	-64	55	-60	
GW-002	LPN002	49	-64	51	-60	51	-64	
GW-003	LPN001	79	-67	76	-68	80	-67	-68.7
GW-003	LPN001	67	-68	70	-67	67	-68	
GW-003	LPN001	50	-71	67	-71	43	-71	
GW-003	LPN002	110	-46	114	-46	115	-46	-46.7
GW-003	LPN002	89	-44	99	-44	92	-44	

GW-003	LPN002	67	-50	46	-50	52	-50	
GW-004	LPN001	78	-67	64	-67	111	-67	-73.1
GW-004	LPN001	50	-77	50	-77	52	-77	
GW-004	LPN001	34	-78	39	-74	25	-74	
GW-004	LPN002	77	-53	69	-53	56	-52	-56.3
GW-004	LPN002	59	-58	68	-61	56	-58	
GW-004	LPN002	44	-61	63	-58	47	-53	

## Appendix D – Results from Metal Sheet Obstruction

No obstruction in place

		No partition								
	Gateway	LPN	341		342		343		AVG All	
1	GW-001	LPN002	171	-51	153	-51	165	-51	-49.111	
2	GW-001	LPN002	69	-47	53	-47	77	-47		
3	GW-001	LPN002	17	-48	21	-48	12	-52		
1	GW-001	LPN005	73	-63	50	-63	60	-58	-61.556	
2	GW-001	LPN005	56	-58	44	-58	48	-63		
3	GW-001	LPN005	47	-67	33	-57	47	-67		
1	GW-001	LPN006	82	-68	53	-68	67	-64	-68.333	
2	GW-001	LPN006	52	-64	45	-74	47	-68		
3	GW-001	LPN006	32	-74	35	-64	41	-71		
1	GW-002	LPN002	164	-31	140	-31	166	-31	-28	
2	GW-002	LPN002	57	-25	57	-25	39	-25		
3	GW-002	LPN002	19	-28	16	-28	11	-28		
1	GW-002	LPN005	58	-69	43	-70	61	-60	-65.778	
2	GW-002	LPN005	45	-70	36	-69	56	-70		
3	GW-002	LPN005	41	-60	35	-60	30	-64		
1	GW-002	LPN006	58	-73	35	-73	49	-73	-71.778	
2	GW-002	LPN006	38	-74	31	-69	40	-74		

3	GW-002	LPN006	32	-70		31	-70		36	-70	
1	GW-003	LPN002	84	-38		78	-39		92	-38	-39.333
2	GW-003	LPN002	77	-39		73	-38		79	-39	
3	GW-003	LPN002	49	-41		40	-41		38	-41	
1	GW-003	LPN005	106	-55		98	-55		103	-53	-54
2	GW-003	LPN005	91	-53		84	-53		92	-55	
3	GW-003	LPN005	53	-54		45	-54		44	-54	
1	GW-003	LPN006	106	-61		84	-61		77	-61	-56.667
2	GW-003	LPN006	49	-55		63	-54		49	-55	
3	GW-003	LPN006	43	-54		40	-55		47	-54	
1	GW-004	LPN002	130	-50		118	-50		117	-50	-51
2	GW-004	LPN002	67	-51		54	-51		74	-52	
3	GW-004	LPN002	62	-52		54	-52		44	-51	
1	GW-004	LPN005	157	-31		136	-31		146	-31	-19.222
2	GW-004	LPN005	54	-12		33	-12		45	-12	
3	GW-004	LPN005	29-Jan	-8		30	-28		23	-8	
1	GW-004	LPN006	68	-54		59	-59		65	-54	-54.556
2	GW-004	LPN006	51	-59		58	-54		50	-59	
3	GW-004	LPN006	43	-51		40	-51		44	-50	
1	GW-005	LPN002	78	-57		64	-57		80	-61	-55.667
2	GW-005	LPN002	58	-49		49	-61		51	-57	

3	GW-005	LPN002	50	-61		49	-49		51	-49	
1	GW-005	LPN005	97	-45		87	-45		83	-45	-43
2	GW-005	LPN005	72	-44		75	-44		65	-44	
3	GW-005	LPN005	48	-40		37	-40		47	-40	
1	GW-005	LPN006	85	-54		85	-54		105	-54	-50.667
2	GW-005	LPN006	76	-45		66	-53		77	-53	
3	GW-005	LPN006	69	-53		51	-45		53	-45	
1	GW-006	LPN002	72	-59		65	-59		74	-59	-65.111
2	GW-006	LPN002	65	-61		51	-61		45	-61	
3	GW-006	LPN002	32	-62		30	-82		37	-82	
1	GW-006	LPN005	115	-53		111	-53		134	-53	-52
2	GW-006	LPN005	79	-52		74	-52		51	-52	
3	GW-006	LPN005	62	-51		47	-51		41	-51	
1	GW-006	LPN006	97	-31		90	-31		97	-31	-31.333
2	GW-006	LPN006	55	-28		61	-35		58	-28	
3	GW-006	LPN006	55	-35		42	-28		58	-35	
1	GW-007	LPN002	55	-64		47	-64		55	-64	-69.333
2	GW-007	LPN002	37	-69		36	-69		39	-69	
3	GW-007	LPN002	25	-68		23	-79		32	-78	
1	GW-007	LPN005	99	-59		89	-59		103	-59	-56.667
2	GW-007	LPN005	69	-53		72	-58		70	-53	

3	GW-007	LPN005	62	-58		65	-53		45	-58	
1	GW-007	LPN006	180	-46		171	-46		178	-46	-46
2	GW-007	LPN006	68	-45		59	-45		51	-45	
3	GW-007	LPN006	15	-47		15	-47		23	-47	

Steel obstruction placed at location 0.5m

		Partition at position 0.5m									
	Gateway	LPN	346		347		348		AVG All	DIFF for all	
1	GW-001	LPN002	211	-52	177	-52	190	-52	-53.0	-4	
2	GW-001	LPN002	57	-53	64	-53	71	-53			
3	GW-001	LPN002	2	-54	2	-54	2	-54			
1	GW-001	LPN005	83	-61	71	-61	83	-60	-63.2	-2	
2	GW-001	LPN005	58	-60	64	-60	68	-68			
3	GW-001	LPN005	51	-69	33	-69	66	-61			
1	GW-001	LPN006	72	-73	52	-73	78	-73	-70.3	-2	
2	GW-001	LPN006	56	-66	46	-72	40	-66			
3	GW-001	LPN006	30	-72	44	-66	39	-72			
1	GW-002	LPN002	104	-11	100	-31	103	-31	-17.3	11	
2	GW-002	LPN002	96	-31	81	-11	92	-11			
3	GW-002	LPN002	33	-10	40	-10	44	-10			
1	GW-002	LPN005	64	-72	57	-72	50	-71	-67.3	-2	
2	GW-002	LPN005	47	-59	50	-59	48	-73			
3	GW-002	LPN005	40	-70	49	-71	44	-59			
1	GW-002	LPN006	74	-64	70	-64	82	-64	-65.1	7	
2	GW-002	LPN006	56	-63	50	-63	54	-63			

3	GW-002	LPN006	46	-68		37	-69		51	-68		
1	GW-003	LPN002	78	-50		89	-50		106	-50	-47.7	-8
2	GW-003	LPN002	77	-44		72	-49		77	-44		
3	GW-003	LPN002	67	-49		50	-44		56	-49		
1	GW-003	LPN005	66	-53		58	-53		74	-53	-56.2	-2
2	GW-003	LPN005	54	-68		49	-68		64	-67		
3	GW-003	LPN005	45	-48		43	-48		40	-48		
1	GW-003	LPN006	75	-54		83	-54		88	-54	-68.7	-12
2	GW-003	LPN006	29	-75		23	-75		28	-76		
3	GW-003	LPN006	29	-76		22	-77		23	-77		
1	GW-004	LPN002	76	-55		72	-55		67	-56	-56.8	-6
2	GW-004	LPN002	57	-54		57	-56		53	-55		
3	GW-004	LPN002	54	-56		48	-62		53	-62		
1	GW-004	LPN005	89	-31		85	-31		85	-31	-26.6	-7
2	GW-004	LPN005	53	-14		49	-14		72	-14		
3	GW-004	LPN005	48	-33		37	-33		44	-38		
1	GW-004	LPN006	119	-55		104	-55		142	-55	-51.3	3
2	GW-004	LPN006	38	-50		40	-49		47	-49		
3	GW-004	LPN006	38	-49		38	-50		34	-50		
1	GW-005	LPN002	80	-58		78	-58		71	-59	-62.3	-7
2	GW-005	LPN002	68	-59		61	-70		66	-58		

3	GW-005	LPN002	48	-70		42	-59		54	-70		
1	GW-005	LPN005	109	-50		95	-50		89	-44	-49.2	-6
2	GW-005	LPN005	79	-44		88	-44		85	-49		
3	GW-005	LPN005	31	-54		23	-54		34	-54		
1	GW-005	LPN006	92	-48		77	-48		87	-48	-47.9	3
2	GW-005	LPN006	81	-45		71	-45		65	-45		
3	GW-005	LPN006	32	-53		34	-53		35	-46		
1	GW-006	LPN002	90	-59		75	-59		100	-59	-59.0	6
2	GW-006	LPN002	75	-60		66	-58		67	-60		
3	GW-006	LPN002	63	-58		64	-60		56	-58		
1	GW-006	LPN005	81	-48		60	-54		101	-47	-50.9	1
2	GW-006	LPN005	64	-54		59	-48		41	-51		
3	GW-006	LPN005	55	-51		57	-51		39	-54		
1	GW-006	LPN006	65	-33		63	-28		59	-33	-33.3	-2
2	GW-006	LPN006	54	-39		56	-33		53	-39		
3	GW-006	LPN006	50	-28		52	-39		51	-28		
1	GW-007	LPN002	63	-69		57	-69		57	-69	-69.4	0
2	GW-007	LPN002	47	-67		45	-67		40	-67		
3	GW-007	LPN002	27	-81		28	-68		36	-68		
1	GW-007	LPN005	63	-54		74	-52		81	-53	-53.0	4
2	GW-007	LPN005	59	-52		60	-53		60	-52		

3	GW-007	LPN005	52	-53		37	-54		33	-54		
1	GW-007	LPN006	65	-53		52	-52		63	-53	-53.7	-8
2	GW-007	LPN006	58	-52		49	-56		61	-52		
3	GW-007	LPN006	42	-56		47	-53		58	-56		

Steel obstruction placed at location 3.5m

		<b>Partition at position 3.5m</b>							<b>Result</b>	
	<b>Gateway</b>	<b>LPN</b>	<b>355</b>		<b>356</b>		<b>357</b>		<b>DIFF</b>	<b>Use highest count</b>
1	GW-001	LPN002	106	-56	87	-56	75	-56	-5	Low
2	GW-001	LPN002	51	-50	66	-50	43	-50		
3	GW-001	LPN002	50	-54	39	-51	32	-51		
1	GW-001	LPN005	79	-60	61	-59	54	-60	1.67	Negligible
2	GW-001	LPN005	49	-59	52	-60	49	-59		
3	GW-001	LPN005	45	-58	41	-80	32	-80		
1	GW-001	LPN006	119	-65	74	-64	62	-65	2	Negligible
2	GW-001	LPN006	62	-64	73	-65	59	-64		
3	GW-001	LPN006	38	-68	48	-68	33	-68		
1	GW-002	LPN002	101	-31	103	-31	121	-31	0	Negligible
2	GW-002	LPN002	80	-11	65	-11	39	-11		
3	GW-002	LPN002	45	-10	47	-10	25	-10		
1	GW-002	LPN005	92	-63	86	-60	76	-62	4.67	High
2	GW-002	LPN005	84	-60	73	-63	73	-60		
3	GW-002	LPN005	56	-61	38	-61	35	-61		
1	GW-002	LPN006	132	-68	115	-68	87	-68	5	High
2	GW-002	LPN006	53	-69	40	-69	52	-67		
3	GW-002	LPN006	23	-66	23	-65	29	-69		

1	GW-003	LPN002	78	-45	86	-45	73	-45	-6.67	Very low
2	GW-003	LPN002	57	-42	54	-42	50	-42		
3	GW-003	LPN002	41	-43	42	-46	37	-46		
1	GW-003	LPN005	111	-51	95	-51	94	-51	3.33	High
2	GW-003	LPN005	83	-45	71	-45	60	-45		
3	GW-003	LPN005	40	-54	33	-54	26	-54		
1	GW-003	LPN006	90	-54	75	-54	84	-54	7	High
2	GW-003	LPN006	63	-71	57	-71	35	-71		
3	GW-003	LPN006	32	-70	36	-70	31	-70		
1	GW-004	LPN002	100	-61	65	-61	51	-60	-10.7	Very low
2	GW-004	LPN002	65	-56	61	-56	50	-56		
3	GW-004	LPN002	38	-62	47	-62	43	-61		
1	GW-004	LPN005	98	-31	95	-31	82	-31	0	Negligible
2	GW-004	LPN005	71	-15	48	-15	41	-15		
3	GW-004	LPN005	55	-39	41	-39	29	-40		
1	GW-004	LPN006	77	-60	59	-60	80	-60	-4.33	Low
2	GW-004	LPN006	52	-59	59	-51	32	-59		
3	GW-004	LPN006	47	-61	54	-59	31	-61		
1	GW-005	LPN002	71	-60	79	-60	68	-60	-1.67	Negligible
2	GW-005	LPN002	64	-55	48	-55	60	-55		
3	GW-005	LPN002	38	-56	41	-56	26	-56		

1	GW-005	LPN005	102	-51	96	-54	67	-54	-8	Very low
2	GW-005	LPN005	99	-54	76	-51	53	-51		
3	GW-005	LPN005	25	-61	23	-61	33	-53		
1	GW-005	LPN006	80	-46	62	-46	60	-46	8	Very high
2	GW-005	LPN006	62	-49	57	-49	42	-49		
3	GW-005	LPN006	42	-52	48	-50	39	-50		
1	GW-006	LPN002	90	-62	81	-62	61	-59	-2	Negligible
2	GW-006	LPN002	75	-59	51	-59	48	-62		
3	GW-006	LPN002	26	-64	24	-64	20	-64		
1	GW-006	LPN005	97	-47	88	-47	71	-47	6	High
2	GW-006	LPN005	72	-48	65	-48	59	-48		
3	GW-006	LPN005	51	-52	57	-52	50	-52		
1	GW-006	LPN006	78	-34	68	-34	63	-34	-3	Low
2	GW-006	LPN006	47	-39	40	-40	37	-39		
3	GW-006	LPN006	35	-27	38	-35	29	-40		
1	GW-007	LPN002	48	-77	47	-77	49	-71	-11	Very low
2	GW-007	LPN002	42	-72	42	-72	30	-70		
3	GW-007	LPN002	35	-76	38	-71	24	-78		
1	GW-007	LPN005	212	-55	179	-55	153	-55	4	High
2	GW-007	LPN005	21	-67	18	-67	15	-70		
3	GW-007	LPN005	13	-56	16	-70	11	-71		

1	GW-007	LPN006	105	-55	97	-53	76	-53	-7.67	Very Low
2	GW-007	LPN006	86	-53	80	-55	70	-55		
3	GW-007	LPN006	76	-54	62	-54	58	-54		

## Appendix E – Results from Machine Learning

### Experiment 1

Training on dataset 2

Training on dataset 1				
Model	MSE	RMSE	MAE	R2
<b>KNN</b>	0.001	0.030	0.001	0.998
<b>SVM</b>	0.007	0.083	0.080	0.988
<b>ANN</b>	0.000	0.006	0.004	1.000

Training on dataset 2, Prediction on dataset 2 and dataset 1

Prediction 1					Prediction 2				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
<b>KNN</b>	0.004	0.065	0.008	0.992	<b>KNN</b>	0.001	0.034	0.002	0.998
<b>SVM</b>	0.007	0.085	0.082	0.987	<b>SVM</b>	0.007	0.085	0.081	0.987
<b>ANN</b>	0.000	0.007	0.005	1.000	<b>ANN</b>	0.000	0.006	0.004	1.000

Training on dataset 2, Prediction on dataset 1 and dataset 3

Prediction 3					Prediction 4				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
<b>KNN</b>	0.002	0.044	0.003	0.996	<b>KNN</b>	0.004	0.060	0.006	0.994
<b>SVM</b>	0.007	0.084	0.081	0.987	<b>SVM</b>	0.007	0.085	0.082	0.987
<b>ANN</b>	0.000	0.006	0.004	1.000	<b>ANN</b>	0.000	0.008	0.005	1.000

Training on dataset 2, Prediction on dataset 3

<b>Prediction 5</b>				
<b>Model</b>	MSE	RMSE	MAE	R2
<b>KNN</b>	0.001	0.035	0.002	0.998
<b>SVM</b>	0.007	0.084	0.080	0.988
<b>ANN</b>	0.000	0.006	0.004	1.000

## Experiment 1

Training on dataset 3

Training on dataset 1				
Model	MSE	RMSE	MAE	R2
KNN	0.002	0.043	0.003	0.997
SVM	0.007	0.084	0.081	0.987
ANN	0.000	0.006	0.004	1.000

Training on dataset 3, Prediction on dataset 3 and dataset 1

Prediction 1					Prediction 2				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.007	0.081	0.010	0.988	KNN	0.002	0.040	0.002	0.997
SVM	0.006	0.080	0.077	0.989	SVM	0.007	0.083	0.080	0.988
ANN	0.000	0.008	0.005	1.000	ANN	0.000	0.006	0.004	1.000

Training on dataset 3, Prediction on dataset 1 and dataset 2

Prediction 3					Prediction 4				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.003	0.050	0.004	0.996	KNN	0.007	0.087	0.010	0.987
SVM	0.007	0.083	0.080	0.988	SVM	0.006	0.080	0.077	0.989
ANN	0.000	0.006	0.004	1.000	ANN	0.000	0.007	0.005	1.000

Training on dataset 3, Prediction on dataset 2

<b>Prediction 5</b>				
<b>Model</b>	<b>MSE</b>	<b>RMSE</b>	<b>MAE</b>	<b>R2</b>
<b>KNN</b>	0.000	0.020	0.001	0.999
<b>SVM</b>	0.007	0.084	0.081	0.987
<b>ANN</b>	0.000	0.005	0.003	1.000

## Experiment 2

Training on dataset 2

Training on dataset 1				
Model	MSE	RMSE	MAE	R2
<b>KNN</b>	0.069	0.263	0.125	0.935
<b>SVM</b>	1.018	1.009	0.897	0.049
<b>ANN</b>	0.099	0.315	0.218	0.907

Training on dataset 2, Prediction on dataset 2 and dataset 1

Prediction 1					Prediction 2				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
<b>KNN</b>	0.196	0.442	0.265	0.817	<b>KNN</b>	0.106	0.326	0.165	0.900
<b>SVM</b>	0.989	0.994	0.894	0.076	<b>SVM</b>	1.008	1.004	0.895	0.053
<b>ANN</b>	0.354	0.595	0.356	0.669	<b>ANN</b>	0.328	0.573	0.330	0.692

Training on dataset 2, Prediction on dataset 1 and dataset 3

Prediction 3					Prediction 4				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
<b>KNN</b>	0.101	0.317	0.159	0.905	<b>KNN</b>	0.183	0.428	0.249	0.831
<b>SVM</b>	1.013	1.007	0.898	0.048	<b>SVM</b>	0.990	0.995	0.896	0.087
<b>ANN</b>	0.328	0.573	0.330	0.692	<b>ANN</b>	0.344	0.586	0.346	0.683

Training on dataset 2, Prediction on dataset 3

<b>Prediction 5</b>				
<b>Model</b>	MSE	RMSE	MAE	R2
<b>KNN</b>	0.063	0.251	0.117	0.942
<b>SVM</b>	1.050	1.025	0.917	0.029
<b>ANN</b>	0.342	0.585	0.344	0.684

## Experiment 2

Training on dataset 3

Training on dataset 1				
Model	MSE	RMSE	MAE	R2
<b>KNN</b>	0.067	0.259	0.119	0.938
<b>SVM</b>	1.091	1.044	0.934	-0.008
<b>ANN</b>	0.098	0.313	0.213	0.910

Training on dataset 3, Prediction on dataset 3 and dataset 1

Prediction 1					Prediction 2				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
<b>KNN</b>	0.184	0.429	0.252	0.830	<b>KNN</b>	0.105	0.324	0.164	0.902
<b>SVM</b>	1.090	1.044	0.932	-0.006	<b>SVM</b>	1.069	1.034	0.918	-0.004
<b>ANN</b>	0.343	0.586	0.346	0.683	<b>ANN</b>	0.328	0.573	0.330	0.692

Training on dataset 3, Prediction on dataset 1 and dataset 2

Prediction 3					Prediction 4				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
<b>KNN</b>	0.099	0.314	0.157	0.907	<b>KNN</b>	0.195	0.441	0.264	0.818
<b>SVM</b>	1.069	1.034	0.918	-0.005	<b>SVM</b>	1.082	1.040	0.927	-0.011
<b>ANN</b>	0.328	0.573	0.330	0.692	<b>ANN</b>	0.354	0.595	0.356	0.669

Training on dataset 3, Prediction on dataset 2

<b>Prediction 5</b>				
<b>Model</b>	MSE	RMSE	MAE	R2
<b>KNN</b>	0.065	0.255	0.122	0.939
<b>SVM</b>	1.083	1.040	0.927	-0.012
<b>ANN</b>	0.353	0.594	0.355	0.670

### Experiment 3

Training on dataset 2

Training on dataset 1				
Model	MSE	RMSE	MAE	R2
KNN	0.005	0.072	0.004	0.998
SVM	0.007	0.082	0.073	0.998
ANN	0.000	0.006	0.004	1.000

Training on dataset 2, Prediction on dataset 2 and dataset 1

Prediction 1					Prediction 2				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.038	0.194	0.031	0.988	KNN	0.011	0.104	0.009	0.996
SVM	3.239	1.800	1.714	-0.061	SVM	3.243	1.801	1.716	-0.059
ANN	2.106	1.451	1.055	0.310	ANN	2.000	1.414	1.002	0.347

Training on dataset 2, Prediction on dataset 1 and dataset 3

Prediction 3					Prediction 4				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.020	0.142	0.014	0.993	KNN	0.046	0.215	0.035	0.985
SVM	3.246	1.802	1.718	-0.060	SVM	3.138	1.772	1.683	-0.025
ANN	2.000	1.414	1.002	0.347	ANN	1.994	1.412	0.999	0.349

Training on dataset 2, Prediction on dataset 3

<b>Prediction 5</b>				
<b>Model</b>	MSE	RMSE	MAE	R2
<b>KNN</b>	0.010	0.098	0.006	0.997
<b>SVM</b>	3.279	1.811	1.728	-0.071
<b>ANN</b>	1.987	1.410	0.996	0.351

### Experiment 3

Training on dataset 3

Training on dataset 1				
Model	MSE	RMSE	MAE	R2
KNN	0.010	0.098	0.006	0.997
SVM	0.006	0.078	0.071	0.998
ANN	0.000	0.006	0.004	1.000

Training on dataset 3, Prediction on dataset 3 and dataset 1

Prediction 1					Prediction 2				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.034	0.183	0.027	0.989	KNN	0.008	0.092	0.007	0.997
SVM	3.297	1.816	1.749	-0.076	SVM	3.306	1.818	1.751	-0.080
ANN	1.992	1.411	0.998	0.349	ANN	2.001	1.414	1.002	0.347

Training on dataset 3, Prediction on dataset 1 and dataset 2

Prediction 3					Prediction 4				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.013	0.112	0.009	0.996	KNN	0.025	0.159	0.023	0.992
SVM	3.307	1.818	1.751	-0.080	SVM	3.392	1.842	1.776	-0.111
ANN	2.000	1.414	1.002	0.347	ANN	2.105	1.451	1.054	0.311

Training on dataset 3, Prediction on dataset 2

<b>Prediction 5</b>				
<b>Model</b>	MSE	RMSE	MAE	R2
<b>KNN</b>	0.003	0.055	0.002	0.999
<b>SVM</b>	3.394	1.842	1.776	-0.111
<b>ANN</b>	2.101	1.450	1.053	0.312

## Experiment 4

Training on dataset 2

Training on dataset 1				
Model	MSE	RMSE	MAE	R2
KNN	0.377	0.614	0.320	0.925
SVM	4.868	2.206	1.981	0.026
ANN	0.491	0.701	0.513	0.902

Training on dataset 2, Prediction on dataset 2 and dataset 1

Prediction 1					Prediction 2				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.612	0.782	0.499	0.878	KNN	0.417	0.646	0.356	0.919
SVM	4.627	2.151	1.915	0.074	SVM	4.914	2.217	1.998	0.046
ANN	0.775	0.881	0.687	0.845	ANN	0.543	0.737	0.544	0.895

Training on dataset 2, Prediction on dataset 1 and dataset 3

Prediction 3					Prediction 4				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.427	0.653	0.357	0.917	KNN	0.657	0.811	0.496	0.873
SVM	4.928	2.220	2.000	0.043	SVM	4.888	2.211	1.984	0.057
ANN	0.547	0.740	0.545	0.894	ANN	0.855	0.925	0.726	0.835

Training on dataset 2, Prediction on dataset 3

<b>Prediction 5</b>				
<b>Model</b>	MSE	RMSE	MAE	R2
<b>KNN</b>	0.370	0.608	0.323	0.929
<b>SVM</b>	5.007	2.238	2.027	0.034
<b>ANN</b>	0.430	0.656	0.479	0.917

## Experiment 4

Training on dataset 3

Training on dataset 1				
Model	MSE	RMSE	MAE	R2
KNN	0.302	0.550	0.261	0.942
SVM	5.132	2.265	2.070	0.010
ANN	0.378	0.615	0.399	0.927

Training on dataset 3, Prediction on dataset 3 and dataset 1

Prediction 1					Prediction 2				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	1.060	1.030	0.648	0.795	KNN	0.514	0.717	0.366	0.900
SVM	5.051	2.247	2.049	0.025	SVM	5.054	2.248	2.050	0.018
ANN	1.493	1.222	0.965	0.712	ANN	0.695	0.834	0.556	0.865

Training on dataset 3, Prediction on dataset 1 and dataset 2

Prediction 3					Prediction 4				
Model	MSE	RMSE	MAE	R2	Model	MSE	RMSE	MAE	R2
KNN	0.540	0.735	0.379	0.895	KNN	0.905	0.951	0.569	0.819
SVM	5.064	2.250	2.052	0.017	SVM	4.856	2.204	1.996	0.029
ANN	0.712	0.844	0.566	0.862	ANN	1.116	1.056	0.754	0.777

Training on dataset 3, Prediction on dataset 2

<b>Prediction 5</b>				
<b>Model</b>	MSE	RMSE	MAE	R2
<b>KNN</b>	0.402	0.634	0.318	0.920
<b>SVM</b>	4.991	2.234	2.029	0.002
<b>ANN</b>	0.571	0.756	0.507	0.886