Department of Electrical and Electronic Engineering

MSc in Communication Networks and Signal Processing

Design of Coverage Digital Twin for Wireless 5G/6G Networks

Author:
JASLEEN KAUR SETHI
YW23759

Supervisor:
DR. RASHEED HUSSAIN

SEPTEMBER 2024

Abstract

Radio-wave propagation in enclosed environments can be highly complex due to a variety of factors. Obstacles such as walls, furniture, and other structures can cause phenomena like reflection, diffraction, scattering, and refraction. Additionally, environmental effects such as multi-path propagation, shadowing, and fading contribute to signal degradation, leading to increased path loss and a reduction in signal quality at the receiver. To determine signal characteristics of an indoor environment, it is thus imperative to perform propagation modelling for determining reliable factors such as prediction of optimal placement of transmitter and network topology, calculation of signal strength, network performance, capacity, path loss, and quality of the signal on the basis of parameters such as frequency, transmit power, wavelength, and characteristics of antennas. This model can be crucial for various sectors such as telecommunications, healthcare and industrial automation in smart cities.

For an accurate design and simulation of a wireless environment, we create a coverage digital twin for the modelling of this environment to produce accurate and reliable results.

The creation of a twin network involves three fundamental components. First, a computer-generated model must be developed that accurately replicates the physical environment. This model should account for factors such as the material properties, obstacles within the setting, and their dimensions. The second critical component is the acquisition of accurate, reliable, and consistent data, which is essential for precisely predicting signal propagation characteristics. Data can be collected through various methods, including software measurement, training machine learning models, or transmission via Wi-Fi networks. This data is then integrated and trained to predict signal properties beyond the existing data, another vital part of the digital twin. The twin network is subsequently utilised for applications such as allocation of resources and it's management, predictive maintenance, capacity planning, service quality monitoring, and network segmentation of the real-time system.

This thesis in collaboration with Toshiba Europe Ltd. thus explores an approach to building a coverage digital twin for single and multi-floor environments, addressing the challenge of fusion of real-time and synthetic data from ray-tracing simulations and empirical models such as 3GPP and multi-floor path-loss models to create a virtual replica of the environment showcasing transmitter placement on a coverage map to maximise coverage. The digital twin framework is further employed to predict path loss and signal strength from the integrated data, thereby enhancing the accuracy of coverage planning in multi-floor structures.

The results of data fusion by the measurement and ray-tracing, along with the data generated from the 3GPP model, are gathered, and the difference between the real-time and the synthetic data is compared in a single-room environment, with the difference being between 10-15 dB, proving the ability of data integration. Coverage maps are thus simulated to determine the maximum coverage with the minimum number of transmitters by calculating the maximum average received power at each of the transmitter locations. A neural network is then trained using various material properties, such as permittivity and reflectivity, along with transmitter and receiver positions, and thus provides the path loss in the environment. The model achieves a mean absolute error of 2.52 and a mean square error of 1%. This approach demonstrates the benefit of machine learning over classical methods by expanding the dataset by merging the data, paving the way for more reliable and optimised 5G network deployments.
Keywords: Radio-Wave Propagation, multi-path, propagation modelling, Digital-Twin, Data fusion, Coverage Map, Deep Learning

Dedication and acknowledgement

I would like to express my gratitude to **Toshiba Europe Ltd.** for giving me the incredible opportunity to undertake my thesis under their esteemed guidance. In addition to their mentorship, I am immensely thankful for their generous provision of office resources, staff support, and lab appliances, which were critical to the successful completion of this project. This experience has been invaluable to my academic and professional growth.

I also extend my sincere thanks to my project supervisors, **Dr. Adnan Aijaz and Dr. Peizheng Li**, for their regular assistance, innovative ideas, and invaluable suggestions. Their project briefs and strategic direction have significantly shaped the course of this work and provided me with a clear way forward.

I am profoundly thankful to my academic supervisor, **Dr. Rasheed Hussain**, whose constant support, insightful advice, and regular meetups have been instrumental in guiding me throughout this research. His motivation and unwavering follow-up have been crucial in helping me stay focused and driven.

A special mention goes to **Sajida Gufran** and my colleagues at Toshiba, whose help in the labs was indispensable. Sajida's assistance in measurements, hardware configurations, the 5G setup, and her guidance in antenna configuration was invaluable. My colleagues at Toshiba provided ongoing support and collaboration, which greatly contributed to the practical success of this project. Their expertise and willingness to explain complex concepts enhanced my understanding of the technical aspects, making this a truly collaborative effort.

I am deeply grateful to my **family** for their regular check-ins, constant encouragement, and understanding. Their unwavering support has been a pillar of strength throughout this journey.

Lastly, I would like to thank my friends, especially **Sharzeel Saleem**, for their support, ideas, and motivation. Their belief in me has been a great source of inspiration and has helped me persevere through the challenges of this research.

Declaration and Disclaimer

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Taught Postgraduate Programmes and that it has not been submitted for any other academic award.
Except where indicated by specific reference in the text, this work is my own work. Work done in collaboration with, or with the assistance of others, is indicated as such. I have identified all material in this dissertation which is not my own work through appropriate referencing and acknowledgement. Where I have quoted from the work of others, I have included the source in the references/bibliography.
Any views expressed in the dissertation are those of the author.
The author confirms that the printed copy and the electronic version of this thesis are identical.

SIGNED: Jasleen Kaur Sethi
DATE: 5/09/2024

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1 Background

With the advent of exponential growth in data traffic, due to higher user engagement and the proliferation of diverse user equipment, there is a need to significantly enhance the capabilities of wireless communication networks. [2] These enhancements focus on improving critical performance parameters such as latency, quality of service (QoS), connectivity, and the power received in various environments. The emergence of 5G and 6G technologies is set to revolutionize the landscape of wireless communications by offering substantial advancements in data rates, reducing end-to-end latency, enabling processing of data in a real-time environment, and fostering the development of energy-efficient, secure wireless networks. While earlier generations of wireless technologies primarily operated below 6 GHz, 5G is designed to function across a much broader spectrum, beginning from 0.6 GHz, with 6G projected to operate in the 95 GHz to 3 THz range—potentially providing data transmission speeds up to 1000 times faster than those of existing 4G technologies. [3] This leap in technology necessitates the need for a robust infrastructure to support the next wave of connectivity demands.

To fuel these advancements and unlock the full potential of these next-generation networks, there must be a seamless integration between the physical and digital worlds. This integration is essential for the creation and eventual prediction of real-time environments that accurately reflect the complexities of the physical world. It is not feasible to perform every operation and test in the real world due to practical constraints. Thus, virtual simulations allow for rigorous testing and analysis in controlled settings, enabling the refinement and optimization of network configurations before deployment in the physical world. Such an integrated approach is vital for various applications, including network optimization, ensuring seamless connectivity, advancing virtual and augmented reality experiences, and enabling a wide array of Internet of Things (IoT) implementations. [4] Achieving these objectives requires the deployment of advanced software platforms, sophisticated tools, and high-performance computing systems, ultimately managing the increased complexity and scale of modern wireless networks.

To continually improve the design, simulation, management, and optimization of these networks, ensuring optimal performance and reliability, the building of a digital twin is proved essential. [5] A digital twin is a virtual model that mirrors the entire life cycle of a real network, including elements like the core network, radio access network, servers, databases, and security systems. By replicating these systems in a virtual environment, Digital Twins allow for extensive testing and optimization without impacting real-world operations.

Radio wave propagation is crucial for modern wireless communication systems, driving extensive research in the field as it is fundamental to the reliable transmission of data across vast distances. Radio waves are a type of electromagnetic radiation utilized extensively to carry information through the air. However, their propagation is subject to a variety of influencing factors, including the power of the transmitted signal, obstacles within the environment, antenna characteristics, and the various forms of losses that can occur as the signal travels through different media. [6] These losses, which include reflection, diffraction, scattering, absorption, and insertion losses, can significantly degrade the quality and strength of the signal that is received.

In enclosed environments, the complexity of radio wave propagation increases substantially. This is due to the presence of numerous obstacles and the materials they are made of, such as walls, furniture, doors, windows, and the overall architectural layout of the building. These obstacles cause the signal to experience multiple interactions, leading to phenomena such as multi-path propagation, where signals take multiple paths due to reflections, and variations in signal strength due to small-scale and large-scale fading. These factors collectively make it challenging to predict signal behavior accurately in indoor environments, where the dynamics of radio

wave propagation are highly intricate.
To address these challenges, we use propagation modeling as a means to predict signal coverage and to assess and optimize the parameters crucial for effective network planning, optimization, and management. This includes tasks such as managing interference, determining path loss, ensuring quality of service, and maintaining network security.
The project is thus focused on integrating the concept of digital twins with propagation modeling, such that networks can be further optimized in real-time. Digital twins offer a powerful tool for simulating different propagation environments and scenarios, allowing network operators to predict and mitigate potential issues before they impact the real network. This integration enhances the ability to manage interference, optimize signal quality, and ensure that networks meet the stringent requirements of next-generation wireless communication.

1.1 Aim and Objectives

The project at Toshiba aims to develop a coverage digital twin to predict the path loss and the received signal strength for a multi-floor indoor environment using fully connected deep learning neural network by integrating the measured real-time data and data that is simulated using ray-tracing techniques and empirical models.

The objectives of the project are:

- Collecting real-time data and simulating the physical environment using ray-tracing methods to generate synthetic data, followed by a comparative analysis of the accuracy of it's integration in terms of closeness to the real-time data, to validate the effectiveness of integrating these two data sources
- Developing a virtual propagation model by constructing a coverage map to optimize signal coverage by determining the ideal placement and number of transmitters by analyzing maximum average received power across all points and selecting the configuration that provides the highest overall signal strength while considering factors such as multi-path propagation, channel interference, and path loss
- Development of a neural network model to predict signal strength by training on unified data, including material properties and simulated measurements and validate the trained model's accuracy for a multi-floor environment by assessing key performance indicators such as measured absolute error, mean squared error, and the comparison between the actual and predicted path loss using visualization tools to ensure the effective integration of real and synthetic data within the Digital Twin framework

Chapter 2 Literature Review

2 Overview of existing approaches: Literature Review

Numerous studies have employed various methods to analyze signal characteristics, such as path loss and signal strength, for indoor environments.

2.1 Empirical Models

Empirical models are one of the traditional approaches to predicting the path loss in an environment. These models can be statistical or heuristic. Statistical modeling is effective for site-specific environments with known propagation properties and sufficient measurement data. [7] However, these deterministic models increase computational complexity as they require recalculation with environmental changes. Additionally, if the data doesn't fit the model well, accuracy decreases, leading to prediction errors. Some of the statistical path loss models are Log Distance, ITU Path Loss Model, 3GPP Model, and Cost 231 Multi-Wall Model.
A considerable amount of research has focused on propagation modeling through the use of statistical methods. For instance, in [8], a path loss model examines the impact of humidity and temperature on signal attenuation at 2.4 GHz. Nevertheless, it lacks validation, which affects its accuracy and limits its applicability in real-world scenarios. Similarly, [9] uses simulations to study wave propagation with different antenna positions and material properties, though it requires significant memory for storing solutions. In [10], the ITU indoor path loss model predicts path loss in an office environment with low errors. However, its accuracy is limited to the specific environment for which it was designed and thus lacks generalization. The other mathematical method we use for the modeling of path loss is the heuristic model. In this model, the determination of path loss is based on free space models derived from practical empirical observations and approximations. [11] The free-space Path Loss Model (FSPL) is given by:

$$\text{FSPL(dB)} = 20\log_{10}(d) + 20\log_{10}(f) + 20\log_{10}\left(\frac{4\pi}{c}\right) \tag{1}$$

This model is suitable for unobstructed environments as it doesn't account for multi-path propagation or environmental factors. FSPL can be refined with real-world data for better accuracy. Heuristic models are generally quicker and less dependent on specific environmental conditions, but their accuracy can be limited by the need for precise parameter adjustments and rules.
Utilizing the heuristic model, there are some studies that are conducted. The authors of [3] have proposed a tool to predict path loss of an indoor environment in the office buildings of Belgium and Ghent using the dominant path loss model. The signal strength validation was done using a wireless test bed. In another study, the authors in [12] used a heuristic approach to build an indoor propagation model, taking corners of a room into consideration. It deviates from the real-time data by 2.2 dB. However, it's accuracy is limited to environments with a lesser effect of diffraction. Also in [13], indoor path loss propagation is designed using the dominant path model, which ensures quick and accurate propagation with a deviation of 3 dB. The study is carried out in multiple buildings in Belgium, and also tested on a wireless test bed. While it can be generalized, no tuning is performed on the model. The heuristic model discussed in [14] is more advanced than the previous one, incorporating multiple parameters to study factors like minimum attenuation path, wave-guide effects, and losses due to multi-path propagation. However, its complexity is also a limitation, as the model is specifically tailored to a particular building and hasn't been validated for use in other structures.

2.2 Ray-Tracing

The limitations of utilizing mathematical models for taking into account the categorical aspects of an environment can be minimized using the ray-tracing model. In ray-tracing, the geometrical features of the setting are implemented along with the material properties for accurate propagation modeling of the environment. [15] It is therefore crucial to offer a vector-based representation of the environment to the model to accurately assess multi-path propagation characteristics from both the edges and surfaces. Ray-tracing accounts for all possible propagation paths at individual multi-path arrival times to estimate angle of departure and arrival, path loss, and change of phase on the basis of parameters such as antenna characteristics and properties of transmitter and receiver, such as coordinate system, transmit power, and transmit frequency. Ray-tracing can be performed by two methods. The Shooting and Bouncing Ray method (SBR) is more computationally efficient and versatile, modeling multi-path propagation with reflections, diffraction, and scattering, whereas the Image method, though less complex for reflections, does not account for diffraction, refraction, or diffuse scattering.

Propagation modeling utilizing ray-tracing has a lot of research going on. In [16], the author compares the results of a rectangular room environment with measured data and the data simulated from the ray-tracing environment. It concludes that ray-tracing is an accurate way to determine the losses due to reflections in the walls and the scattering method since it resonates with the measured data. Similarly, the authors of [17] present the simulation of the wave propagation inside the building, achieving good agreement with field measurements and modeling complex 2D and 3D environments using the ray-tracing method. However, it does not fully capture scattering effects near metal structures or complex diffraction scenarios. The study carried out in [18] analyzes multi-path propagation in urban environments, revealing the purpose of diffuse scattering to predict the characteristics such as path-loss, delay spread, and direction of arrival accurately. The model's performance may vary with the complexity of the environment and the accuracy of its representation.

2.3 Deep-Learning

This approach uses neural network models to analyze and forecast the attenuation of the signal while learning complex patterns and relationships from measured or simulated signal characteristics and providing accurate and adaptive predictions of path loss in diverse scenarios. These models enhance the ability to optimize network design and improve signal coverage in real-world environments.

There has been a lot of work done on the propagation modeling utilizing the deep learning techniques. For example, the authors in [19] evaluate the received signal strength with the help of input variables such as real-world data and the radio map images at 2.4 GHz. They create a convolutional neural network for the indoor radio propagation optimization using loss functions and an optimizer in an office environment to get a discrepancy of 4.25 dB as compared to the real-time measured values. In another research, the authors propose a machine learning (ML) approach for radio propagation modeling, improving accuracy and efficiency in the received signal strength estimation. The Deep Neural Network model enhances prediction accuracy by 25% over traditional models and significantly reduces prediction time compared to ray tracing [20]. The study in [21] represents efficiency in characterization of the channel for a line-of-sight scenario using the deep-learning approach. The model is trained using data from one environment to validate with multiple other environments, not a part of the learning dataset. The accuracy achieved by this model is 2% proving the ability to determine the signal characteristics in LOS scenarios.

Chapter 3 Key Performance Indicators

3 Key Performance Indicators

3.1 Path Loss

As a signal is transmitted from the transmitter to the receiver, the signal often experiences a reduction in power density of the electromagnetic wave, that is, strength of the signal as it travels through the environment. [22] This signal degradation can result from various obstacles, such as buildings, trees, and other physical structures, which cause phenomena like reflection, refraction, and scattering. Additionally, environmental factors such as fading and shadowing contribute to signal loss, leading to complex scenarios like multi-path propagation. In such cases, the signal takes multiple paths to reach the receiver, arriving at different times and potentially causing interference that further diminishes the quality and reliability of the communication. This deterioration of the signal is known as the path loss. This path loss can be determined with the help of 3GPP path loss models, through ray-tracing methods, or using the RAN Intelligent Controller Dashboard, which is the case for measuring real-time data in our project.

Path loss can be quantified as a path loss component, which ranges from 2 to 4. A value of 2 corresponds to free space, while a value of 4 represents more lossy environments. [23]

Path loss can be given by:

$$PL = 10nlog_{10}(d) + C \tag{2}$$

where:
PL = path loss in the environment
n = Path Loss Exponent
d = distance between transmitter and receiver
C = constant that accounts for system losses

3.2 Reference Signal Received Power

The Reference Signal Received Power (RSRP) refers to the strength of the signal received by a device from a transmitter. [24] In the OpenRAN architecture, the User Equipment (UE) measures this RSRP value and sends it to the RAN Intelligent Controller (RIC) software. The RSRP value can be determined using mathematical models, which assist in measuring the strength of the received signal. It is given by:

$$RSRP = P_o - PL \tag{3}$$

where:
P_o = transmit power
PL = path loss in the environment

3.3 Signal to Interference plus Noise Ratio

The signal-to-interference-plus-noise ratio (SINR) is an important metric for assessing wireless signal quality. It compares the power of the desired signal to the sum of interference from other transmitters and background noise in the system. A higher SINR indicates a clearer and stronger signal, leading to better communication performance, while a lower SINR suggests a degraded signal quality due to higher levels of interference and noise. SINR is critical for assessing the reliability and efficiency of wireless communication systems, as it directly impacts data transmission rates and overall network performance. It is given by:

$$SINR = \frac{P}{I + NP} \tag{4}$$

where:
P = Signal Power
I = interference due to other transmitters in the environment
Noise = Background Noise Power in the environment
Noise power can be calculated from noise power density using the following relationship:

$$NP = NPD + BW \tag{5}$$

where:
NP = noise power in the environment
NPD = noise power density (watts/Hz) BW = bandwidth, i.e., the range of frequencies over which the noise power is measured (hertz).

3.4 Mean Absolute Error and Mean Squared Error

We employ a fully connected deep learning model to predict system path loss and enhance the data beyond what is currently available. To evaluate the model's accuracy, we utilize two loss functions: mean absolute error (MAE) and mean squared error (MSE).

Mean Absolute Error: Mean Absolute Error (MAE) measures the average absolute difference between the actual values and the predicted values. [25] It is given by:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{6}$$

Mean Squared Error Mean Squared Error (MSE) is the average of the squares of the differences between predicted and actual values. [25] It is given by:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{7}$$

While the MSE disproportionately contributes more to large errors, MAE treats all errors equally. In our project, we measure MSE and MAE for the comparison of predicted and actual values of the path loss.

Chapter 4 Methodology

4 Fundamental Parts to Build the Digital Twin

There is a notable research gap in creating detailed digital twin frameworks that can effectively manage the complexities of multi-floor structures. These complexities include the effective integration of heterogeneous data sources, the creation of virtual environments that maintain high fidelity, and the prediction of signal characteristics using deep learning models. Challenges such as variability in building materials, interference, and user mobility across different floors further complicate these tasks. This project focuses on overcoming these challenges by developing a coverage digital twin that can seamlessly fuse data to create an accurate virtual replica of the physical environment and predict path loss within enclosed spaces. The system architecture to be followed in this project can be depicted by Figure 4.1.

In our project, we emphasize three key aspects in building a digital twin:

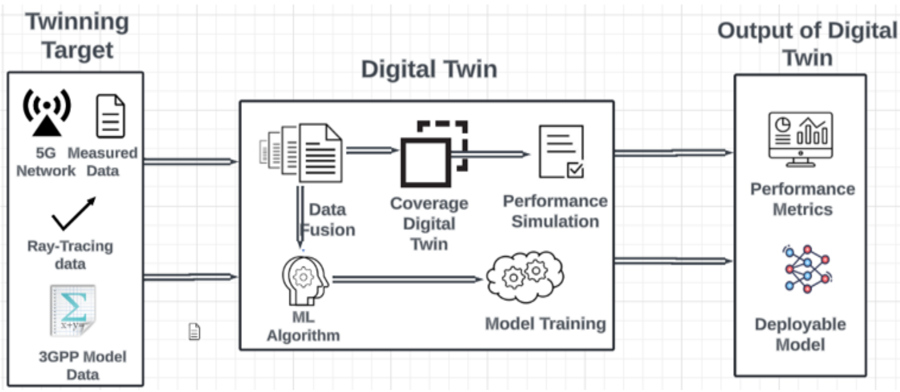


Figure 4.1: System Architecture of a Digital Twin

- Data collection and integration
- Creation of a high-fidelity virtual environment
- Prediction of signal characteristics through Deep Learning model

4.1 Data collection and integration

The initial phase of the project focuses on collecting data, through real-time measurements using a 5G setup in an indoor environment and by generating synthetic data via statistical models and ray-tracing methods. The combination of this data serves as the foundation for creating a comprehensive dataset essential for determining signal specifications. By integrating various types of data, the dataset can be expanded to enhance the accuracy of signal parameter predictions. This enriched dataset will then be utilized to construct a precise virtual environment, which, in turn, will train a deep learning model capable of predicting signal metrics beyond the provided data.

4.1.1 Real-time data

Real-time data was collected using a 5G system setup deployed in the Toshiba Lab. This setup included user equipment (UE), transmitter and receiver antennas, and an OpenRAN architecture comprising the Radio Unit (RU), Distributed Unit (DU), Centralized Unit (CU), RAN intelligent controller (RIC), and the Core Network. The OpenRAN architecture is built to separate the hardware and software elements of the Radio Access Network (RAN), allowing for increased flexibility and scalability. It consists of three main units:

Radio Unit (RU): The RU handles the sending and receiving of radio signals, serving as the interface for the lower functions of the physical layer (Low-PHY). It handles critical tasks such as RF signal processing, precoding, and beamforming. The RU is directly connected to the transmitter antenna and wirelessly communicates with the User Equipment (UE) to ensure effective signal reception.

Distributed Unit (DU): Positioned closer to the network edge, the DU is responsible for real-time baseband processing. It manages Medium Access Control (MAC) and Radio Link Control (RLC) protocols and handles key High-PHY functions, including coding and decoding, modulation and demodulation, and resource block mapping.

Centralized Unit (CU): The CU handles non-real-time baseband processing and manages higher-layer protocols. It hosts critical protocols such as Radio Resource Control (RRC), Next Generation Application Protocol (NGAP), and Packet Data Convergence Protocol (PDCP), which are essential for both control and user plane functions.

In addition to these components, the OpenRAN architecture incorporates the Core Network and the Open RAN Intelligent Controller (RIC). The Core Network oversees functions such as the Access and Mobility Management Function (AMF), User Plane Function (UPF), Policy Control Function (PCF) and Session Management Function (SMF). The RIC, a software-defined component, enhances network intelligence and agility by supporting disaggregation and multi-vendor interoperability.

The RIC is divided into two parts:

Non-Real-Time RIC: This component, within the centralized Service Management and Orchestration (SMO) framework, manages tasks that require more than one second. It is responsible for enforcing SLAs, planning the network offline, and predicting network behavior.

Near-Real-Time RIC: This component handles tasks that need immediate responses within one second, including dynamic resource allocation, spectrum sharing and load balancing.

This setup provides a robust foundation for gathering and analyzing real-time data, essential for the development of a highly accurate digital twin in complex indoor environments. The architecture of OpenRAN can be seen in figure 4.2.

In our setup, the OpenRAN functions as the base station, comprising of the Distributed Unit (DU), Centralized Unit (CU), Core Network, and Radio Unit (RU). The Radio Unit is directly connected to the transmitter antenna and establishes a wireless connection with the User Equipment (UE). The Non-Real-Time RAN Intelligent Controller (RIC) Dashboard is a software tool used to monitor and analyze signal characteristics, including

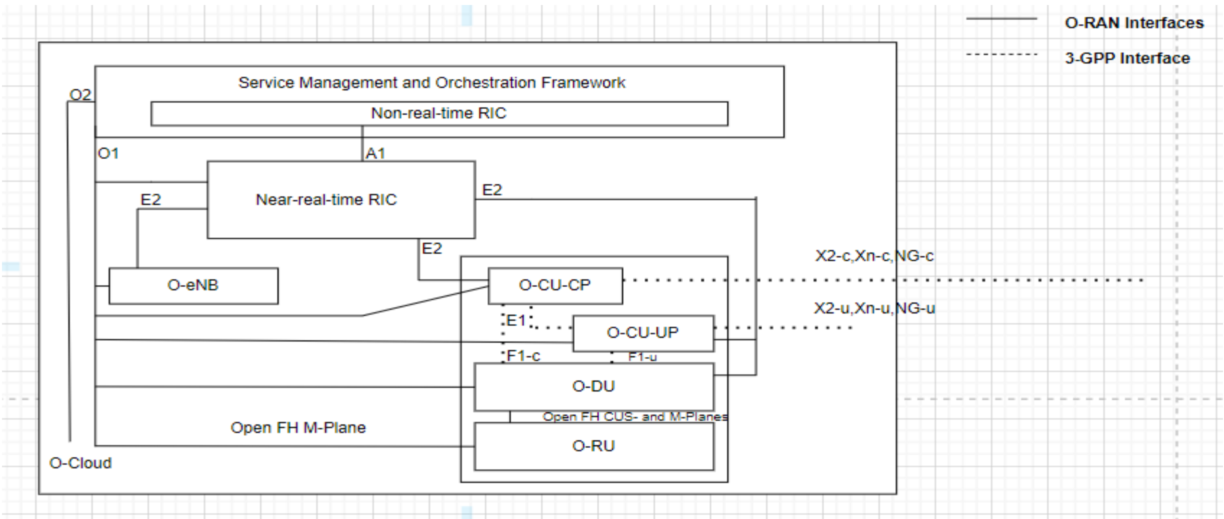


Figure 4.2: OpenRAN Architecture [1]

reference signal received power (RSRP), signal-to-interference plus noise ratio (SINR), and overall signal quality.

The transmitter antenna is omnidirectional and fixed at a specific location. The user equipment is then moved to various points within the environment, where signal metrics and the distance between the transmitter and receiver antennas are continuously monitored and logged. This process enables the collection of real-time data essential for further analysis.We can thus collect real-time data from the specific environment where indoor propagation is needed.

4.1.2 Ray-Tracing data

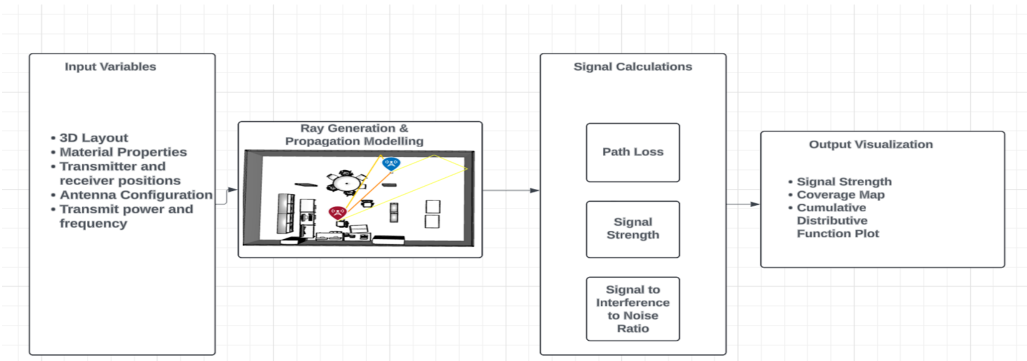


Figure 4.3: Ray-Tracing Method in MATLAB

For generating synthetic data, we use the ray-tracing method on the MATLAB software, depicted in figure 4.3. For the software to perform the ray generation, input variables must be given. The variables that were provided in our project were the transmit power, frequency of the transmitter, and antenna characteristics such as angle, position, and configuration. Tables 4.1, 4.2, and 4.3 list the parameters used as inputs for ray-tracing. Some configurations are kept confidential as per Non-Disclosure Agreement with Toshiba.

A three-dimensional layout of the indoor environment is provided, annotated with the obstacles and material properties of elements such as walls, floors, ceilings, furniture, and other objects, to accurately reflect real-world conditions. Ray generation and subsequent propagation modeling are then carried out. This process involves simulating the propagation of radio waves travelling from the transmitter to the receiver by tracing

Table 4.1: General Parameters

Parameter	**Value**
Channel Model	Indoor Office
Transmit Power	2.36 Watts
Bandwidth	100 MHz
Equivalent Isotropic Radiated Power	38 dB
Noise Power Density	-174 dBm/Hz
Transmit Frequency	3.85 GHz

Table 4.2: Transmitter Antenna Configuration

Parameter	**Value**
Antenna Configuration	2x2 LTE MIMO of 2.4 GHz
Radiation Pattern	Omni-Directional

Table 4.3: Receiver Antenna Configuration

Parameter	**Value**
Antenna Configuration	1x1 Wi-fi MIMO
Radiation Pattern	Omni-Directional

potential paths that the waves might follow. The model considers factors like reflection, refraction, diffraction, and scattering, with the flexibility to specify which factors should be included in the simulation.
There are two primary methods for performing ray tracing in the environment. The image method considers the effects of reflection while excluding diffraction, refraction, and diffuse scattering. In contrast, the Shooting and Bouncing Rays (SBR) method accounts for both reflection and edge diffraction but does not consider corner diffraction, refraction, or diffuse scattering. We employ both methods to determine which more effectively aligns the ray-tracing data with real-time measurements.
The ray-tracing output involves calculating the paths that rays would follow as they propagate from the transmitter to the receiver, factoring in interactions with environmental objects like buildings, walls, and other structures. This analysis helps estimate key signal characteristics, including path loss, direction of arrival, received power and delay spread.

4.1.3 Statistical Models

Another way to figure out signal behavior for propagation modeling is by using statistical models. There are different path loss models, among which the best model is selected according to the environment, taking into consideration its own strengths and weaknesses. For instance, the log-distance path loss model suggests that path loss depends on how far the signal travels, with the distance measured in a logarithmic scale. This model shows the effect of distance between the transmitter and receiver to the path loss in the environment. This model is given by:

$$PL(d) = PL(d_o) + 10nlog(d/d_o) + X_\sigma \tag{8}$$

where $PL(d_o)$: Path loss at one meter from the transmitter
X_σ: Gaussian random variable at 0 mean and 1 standard deviation
n: Path loss exponent depending on the surroundings. It is in between 2 and 6.5 for line-of-sight to highly cluttered environments

The limitation of this model is that it is very oversimplified since it does not take into consideration the obstacles such as a wall, floor, or ceiling that can highly affect the path loss of the environment.
The ITU path loss model is another statistical model that is applicable for the frequency ranges between 900 MHz and 100 GHz. [2 on pp] It is given by:

$$PL(d) = 20log(f) + Nlog(d) + L_f(m) - 28 \tag{9}$$

where f: frequency of the transmitter in MHz
N: distance power coefficient
L_f : Floor penetration loss
m: Number of floors between the access point and the terminal
The limitation of this model is that it is very site-specific. Moreover, it can have inaccurate results for very long or very short distances or in the case of a very cluttered environment.
The model we use is the 3GPP (3rd Generation Partnership Project) Model. The model refers to a set of standardized radio propagation models developed by 3GPP for simulating and evaluating the wireless communication system performance. The selection of this model is based on the fact that it takes into account the walls, floor, and obstacles in the environment. Also, these models are designed to represent various real-world environments, such as urban, suburban, rural, and indoor settings. It can thus help in assessing system performance, such as coverage, capacity, interference, and link quality, under standardized conditions. These models also distinguish between scenarios where the signal path is unobstructed (LOS) and where it is obstructed by buildings or other obstacles (NLOS). NLOS conditions generally lead to higher path loss and different propagation characteristics.
According to the recent release of the 3GPP model, it is appropriate for frequencies ranging from 0.5 to 100 GHz. [26] For indoor settings, the model defines two types of environments:
Indoor Hotspot (InH): In this indoor scenario, environments with higher user density, more confined spaces, and complex multipath propagation are considered. It's used for modeling environments like offices, shopping centers, and other public indoor areas where high-capacity wireless communication is needed.
InH-Office:
Line-of-Sight (LOS) :

$$PL_{\text{InH-LOS}} = 32.4 + 20\log_{10}(d_{3D}) + 20\log_{10}(f_c) \tag{10}$$

Non-Line-of-Sight (NLOS):

$$PL_{\text{InH-NLOS}} = \max\left(PL_{\text{InH-LOS}}, 17.3 + 38.3\log_{10}(d_{3D}) + 24.9\log_{10}(f_c)\right) \tag{11}$$

Distance Range:

$$1\,\text{m} \leq d_{3D} \leq 150\,\text{m}$$

Indoor Factory(InF): This indoor scenario is tailored for large industrial environments with different propagation challenges, such as large open areas, high ceilings, and significant metal obstructions. It's designed for robust communication in factories, warehouses, and similar settings, where reliability and coverage are critical.
InF-Factory:
Line-of-Sight (LOS):

$$PL_{\text{InF-LOS}} = 31.84 + 21.5\log_{10}(d_{3D}) + 19\log_{10}(f_c) + 4 \tag{12}$$

Non-Line-of-Sight (NLOS):

Street Level

$$PL_{\text{InF-NLOS}} = \max\left(PL_{\text{InF-LOS}}, PL_{\text{InF-NLOS}}\right) \tag{13}$$

$$PL_{\text{InF-NLOS}} = 33 + 25.5\log_{10}(d_{3D}) + 20\log_{10}(f_c) + 5.7 \tag{14}$$

Dense-Urban

$$PL_{\text{InF-NLOS}} = \max\left(PL_{\text{InF-LOS}}, PL_{\text{InF-NLOS}}\right) \tag{15}$$

$$PL_{\text{InF-NLOS}} = 36.9 + 30.5\log_{10}(d_{3D}) + 20\log_{10}(f_c) + 7.2 \tag{16}$$

Sub-Urban

$$PL_{\text{InF-NLOS}} = \max\left(PL_{\text{InF-LOS}}, PL_{\text{InF-NLOS}}\right) \tag{17}$$

$$PL_{\text{InF-NLOS}} = 18.6 + 37.6\log_{10}(d_{3D}) + 20\log_{10}(f_c) + 5.9 \tag{18}$$

Dense Sub-Urban

$$PL_{\text{InF-NLOS}} = \max\left(PL_{\text{InF-LOS}}, PL_{\text{InF-NLOS}}\right) \tag{19}$$

$$PL_{\text{InF-NLOS}} = 33.63 + 21.9\log_{10}(d_{3D}) + 20\log_{10}(f_c) + 4.0 \tag{20}$$

Distance Range

$$1\,\text{m} \leq d_{3D} \leq 600\,\text{m}$$

For our project, we have used the statistical model for Indoor Hotspot since it satisfies the conditions of our multi-floor office environment.
After collection of data through real-time setup and simulated data, a comparative analysis was conducted to determine if the two datasets are in alignment, representing the accuracy of the computer-generated model to evaluate the signal metrics of the environment.

4.2 Creation of a high-fidelity virtual environment

After gathering and integrating data from various methods, the next step is to develop a computer-generated virtual replica of the physical system. This involves creating a coverage digital twin, which includes coverage maps for single rooms, single floors, and multi-floor environments. The process involves determining the optimal number of transmitters and their placement to maximize coverage while minimizing interference and redundancy.
The model is segmented into the most effective configurations, and potential transmitter and receiver locations are evaluated to identify the best positions for transmitters to achieve maximum coverage. Figure 4.4 illustrates the simulation process, where all potential receiver locations throughout the area are evaluated to assess coverage. This allows for the identification of the optimal transmitter location based on where the received signal strength is highest. We select the transmitter point where the overall signal strength in the area is maximum by calculating the maximum average received power at each of the points.

By plotting received signal strength on the map, the model provides insights into signal strength variations in specific areas, identifies interference zones caused by other transmitters, and highlights coverage gaps. The

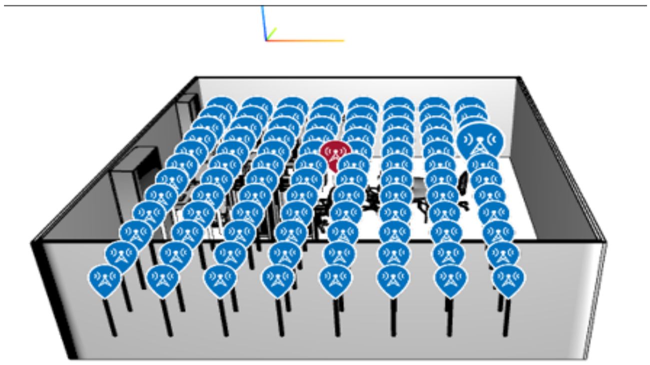


Figure 4.4: Simulation of traversing various receiver locations to find the optimal location of transmitter

simulation also includes iterative testing and adjustments to refine the model, ultimately enhancing the accuracy of the predicted outcome.
When creating a virtual environment, along with simulating the coverage map, it is important to determine the statistical overview of network performance across the coverage area to assess the overall quality of the network and identify areas with weak signal quality or high interference. For this purpose, we plot the cumulative distribution function graph of the signal interference to noise ratio to indicate if the network provides relatively good coverage for the majority of the area. The plot can also reveal how much interference is affecting the network.

4.3 Prediction of signal characteristics through Deep Learning Model

After data collection and integration of real-time and synthetic data and creation of a coverage map, the expanded dataset is used to predict the signal characteristics, such as received signal strength and path loss. This process leverages deep learning models to further enhance the accuracy and efficiency of the prediction.
In our project, we utilize a Fully Connected (FNN) Deep Learning Neural Network (DNN) with one input layer and one output layer, along with two hidden layers in between, as depicted in Figure 4.5, which is built for the prediction of path loss in multiple kinds of environments based on the dataset provided.
Initially, the dataset undergoes cleaning, scaling, and pre-processing to ensure its quality and suitability for

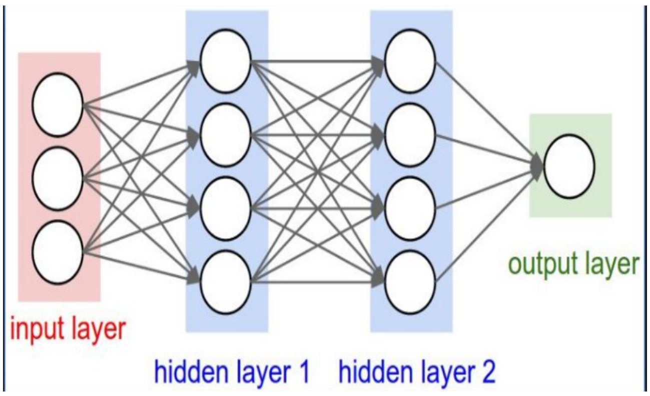


Figure 4.5: Deep Learning model depiction with one input layer, two hidden layers and one output layer [27]

modeling. The neural network is then trained to predict path loss when a signal travels from a transmitter to a receiver, considering obstacles with specific material properties. Key input features for the model include the distance between the transmitter and receiver, their respective locations, material properties of environmental elements, and losses attributed to various environmental factors.

The model employs the ReLU activation function and is optimized using the RMSprop algorithm with a learning rate of 0.001 to fine-tune the weights. It is trained over 1000 epochs, with periodic evaluations every 100 epochs to monitor progress and performance. The dataset is further divided into training and validation sets. To train the model, training set is used, while the validation set is utilized to adjust hyper-parameters and prevent over-fitting. The trained model can then predict path loss under different scenarios, including various transmitter placements, environmental conditions, and the presence of obstacles.

To evaluate the accuracy of the DNN model, we use loss functions such as mean squared error (MSE) and mean absolute error (MAE) to measure its prediction performance and ensure the reliability of the results.

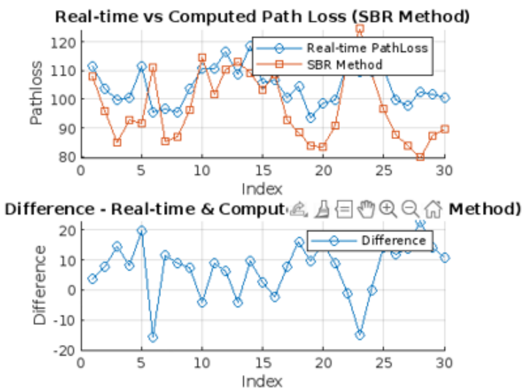
Chapter 5 Results and Analysis

The components of the Digital Twin are initially tested in a single-room environment, followed by evaluations in single-floor and multi-floor settings.

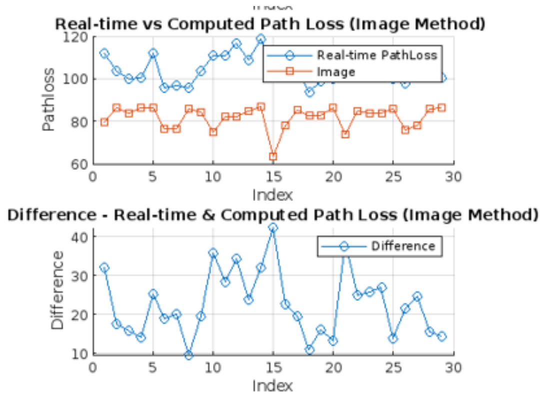
5 Experimental Results and interpretation for various types of environment

5.1 Comparative Analysis: Real-Time vs. Synthetic Data

In the single-room environment, measured data from a 5G system setup is compared to synthetic data generated through ray-tracing methods. This comparison between real-world data and virtual simulations is crucial for validating the accuracy of the Digital Twin. Figures 5.1(a) and 5.1(b) illustrate the comparison between real and synthetic data generated through ray-tracing using the SBR and Image Method, respectively, in MATLAB. The SBR method shows a closer alignment with real-time measurements, with a discrepancy of only 10-15 dB. This small discrepancy can be justified considering that the ray-tracing does not consider the losses due to environmental factors such as fading and shadowing. This close match demonstrates the ability to integrate both the data collection methods.



(a) Comparison of Real and Synthetic Data using SBR Method



(b) Comparison of Real and Synthetic Data using Image Method

Figure 5.1: Comparison of Real and Synthetic Data

5.2 Virtual representation of a single room environment

A coverage map of the single-room lab environment is created using third-party layout building tools. The map is annotated with obstacles and their material properties. Optimal transmitter locations and the required number of transmitters are then plotted on the map to achieve maximum coverage, based on the calculation of maximum average received power at various positions throughout the room to compute the highest overall signal strength at each points. Figure 5.2 shows the virtual layout of the single room with the transmitter and the coverage through the room.

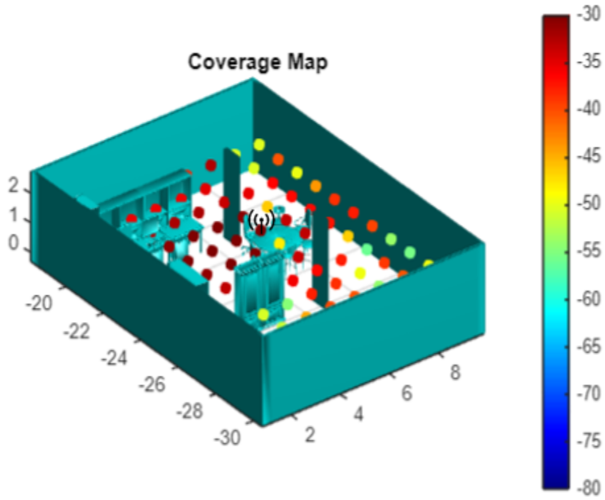
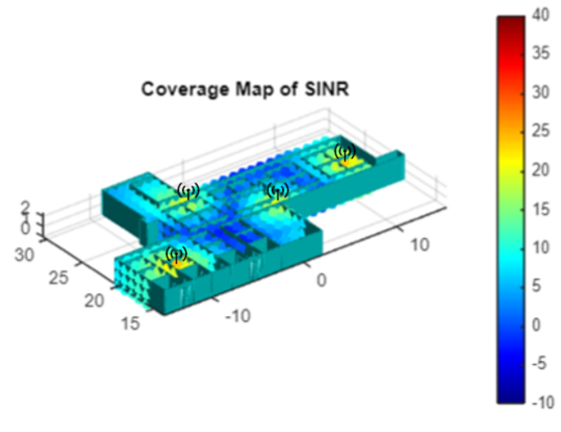


Figure 5.2: Coverage map of a single room environment

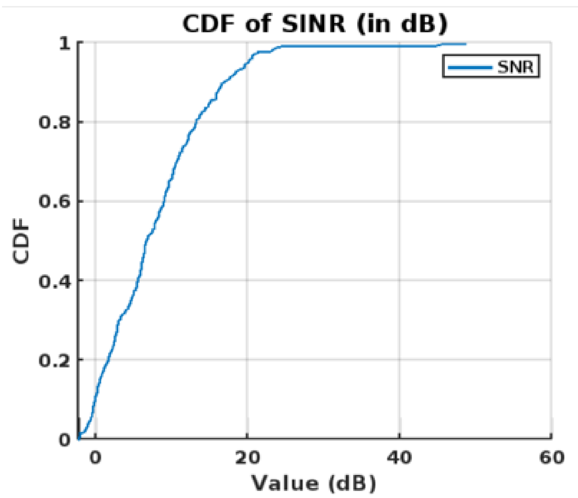
5.3 Virtual representation of a single floor environment

After analyzing the single-room environment, we move on to a single-floor environment. The layout is strategically divided to minimize the number of required transmitters while considering the specific characteristics of the environment. The layout is segmented in a way that optimizes transmitter placement at different channel widths. The maximum average received power at each location is calculated to achieve maximum coverage. As shown in figure 5.3(a), four transmitters are used, with a 100 MHz channel bandwidth split into four 20 MHz segments. This approach helps to determine the optimal transmitter positions while accounting for interference between them, ensuring the best possible coverage.

The Cumulative Distribution Function (CDF) plot of the Signal Interference to Noise Ratio (SINR) is then



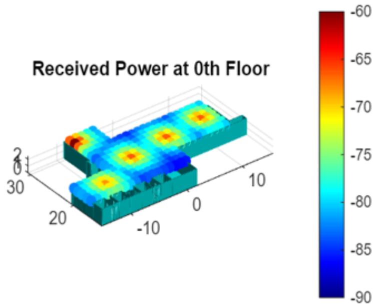
(a) Coverage map with number of Optimal transmitters and it's locations



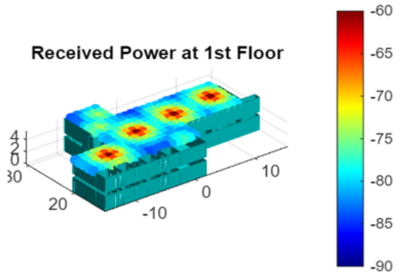
(b) Cumulative Distribution Function plot for Single Interference to Noise Ratio in single floor environment

Figure 5.3: Single floor environment

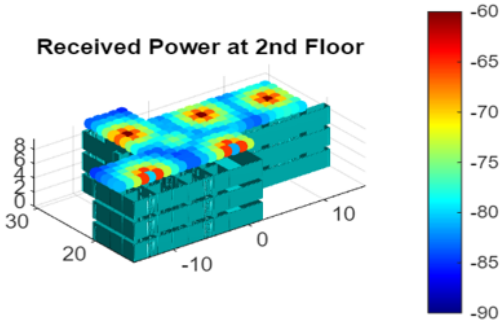
generated, as depicted in figure 5.3(b). The plot reveals that in most areas, the signal strength surpasses the interference, indicating a strong and reliable signal throughout the environment. This suggests that the network is well optimized, with minimal interference affecting the signal quality.



(a) Coverage map with number of Optimal transmitters and it's locations for the ground floor



(b) Coverage map with number of Optimal transmitters and it's locations for the first floor



(c) Coverage map with number of Optimal transmitters and it's locations for the second floor

Figure 5.4: Multi-Floor Environment

5.4 Virtual representation of a multi-floor environment

The study is then extended to a multi-floor building. A coverage map for a three-floor environment is created, highlighting the most effective number of transmitters and their optimal placement based on the received signal strength on each floor, as well as the impact of transmitters on the floors above. The analysis also accounts for penetration loss caused by the ceiling in a multi-floor setting. Moreover, additional transmitters are added to each area of the floor that does not get sufficient coverage, determined by the signal threshold. Figure 5.4 illustrates the transmitter locations on different floors, placed at the dark red points, ensuring optimal coverage across the entire building. For a transmitter on a particular floor, we also observe coverage on the subsequent floors.

To evaluate the effectiveness of transmitter placements concerning interference among all transmitters, we plot the cumulative distribution function (CDF) for the signal interference to noise ratio (SINR) in the environment. Figure 5.5 reveals that interference among transmitters is more significant in the multi-floor setting compared to the single-floor environment, indicating the need for further optimization to address these interference issues.

5.5 Prediction of Path Loss by Deep Learning Model

After creating the virtual environment and gathering data from various methods, we develop a fully connected deep learning network with an input layer, output layer and two hidden layers in between. This model is trained using data such as the distance between the transmitter and receiver, material properties like reflectivity and permittivity, and the angle of incidence to predict path loss values. Following model training and optimization, we evaluate its accuracy using loss functions such as mean absolute error (MAE) and mean squared error (MSE).

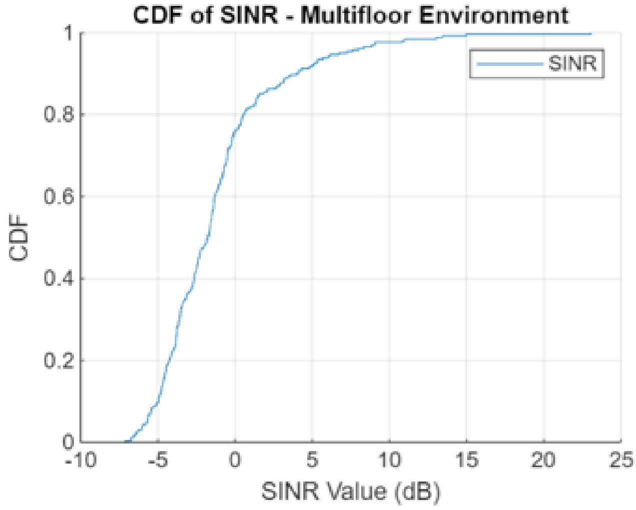
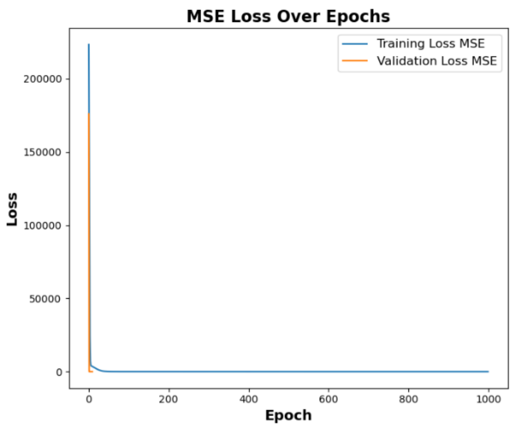
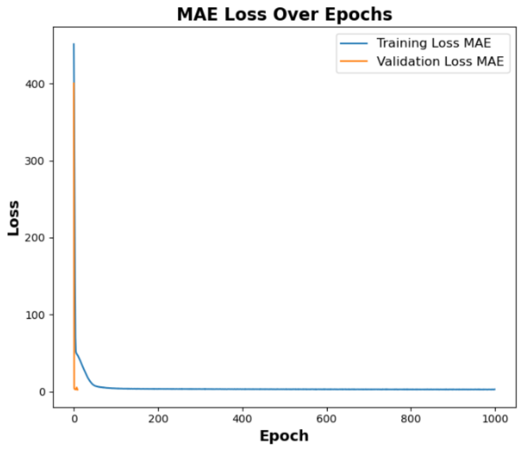


Figure 5.5: Cumulative Distribution Function plot for Single Interference to Noise Ratio in multi-floor environment

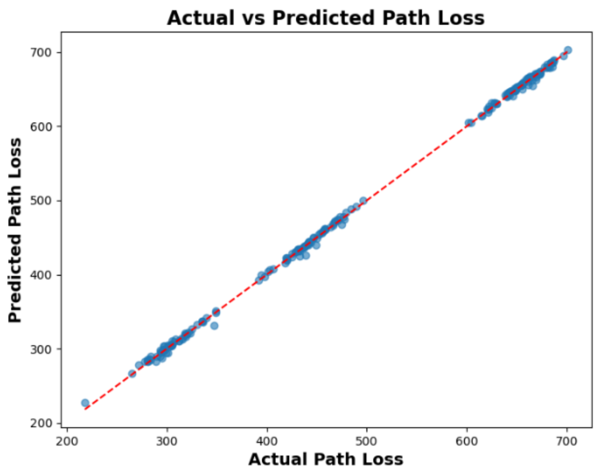
The figures 5.6(a) and 5.6(b) show the MSE and MAE losses over 1000 epochs, which significantly decrease. Also, figure 5.6(c) presents a graph comparing actual versus predicted path loss. The model demonstrates its effectiveness, with an MAE of 2.52 and an MSE of 1%. It demonstrates the model's strong performance in predicting path loss for transmitter and receiver positions with limited data, suggesting its potential for developing an accurate digital twin of coverage.



(a) MSE Loss Over 1000 Epochs



(b) MAE Loss Over 1000 Epochs



(c) Actual vs Predicted Path Loss

Figure 5.6: Fully-Connected Deep Learning Model

Chapter 6 Conclusion and Future Scope

6 Conclusion and Future Scope

6.1 Conclusion

- In this study, we successfully developed and evaluated a Coverage Digital Twin by integrating real-time data with synthetic data generated through ray-tracing and the 3GPP mathematical model. We constructed virtual models for single-room, single-floor, and multi-floor environments to determine the optimal number and placement of transmitters for maximizing coverage. This integrated data was then used to train a deep learning model to predict signal strength in locations beyond the available measurements.
- The comparison between real-time and synthetic data in a single-room environment revealed a minor difference of 10-15 dB, as ray-tracing doesn't consider the environmental factors such as fading and shadowing. It thus validates the effectiveness of combining real and synthetic data for accurate coverage modeling. This result confirms the reliability of training the deep learning model with such integrated data.
- Our approach effectively demonstrates the generation of a coverage map, to create a virtual replica of the physical environment, with optimized the number of transmitters and their placement across various channel widths, allowing us to model and analyze signal characteristics such as received power, path loss, and SINR values.
- With the expanded dataset, the deep learning model demonstrated high accuracy in predicting path loss based on material properties within an environment, achieving a 1% mean squared error (MSE) and a mean absolute error (MAE) of 2.52. These metrics underscore the model's capability to accurately determine signal characteristics and predict indoor environmental conditions.

6.2 Future Scope

The next phase of this research involves integrating real-time building layouts and environmental data into the deep learning model, significantly enhancing its ability to predict path loss in complex and dynamic environments. A key focus will be on achieving accurate data fusion, which will allow the model to seamlessly combine information from various sources—such as ray-tracing simulations, empirical models, and real-time measurements—resulting in a more comprehensive and precise representation of the environment. This is particularly crucial for multi-floor structures, where the complexity of signal propagation increases due to the presence of multiple obstacles, floors, and varying material properties. A more advanced model is required to handle these complexities effectively.

Moreover, this refined model could be developed into an xApp that continuously adapts to live data streams. By incorporating real-time inputs and achieving precise data fusion, the xApp could instantly update predictions on signal strength and path loss, optimize transmitter placement dynamically, adjust network configurations on-the-fly, and proactively manage coverage gaps or interference issues, particularly in challenging multi-floor environments. This would transform the model into a powerful tool for real-time network management and optimization, ensuring reliable coverage even in the most complex structures.

References

[1] M. Polese, L. Bonati, S. D'oro, S. Basagni, and T. Melodia, "Understanding o-ran: Architecture, interfaces, algorithms, security, and research challenges," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1376–1411, 2023.

[2] C. Gehrmann and M. Gunnarsson, "A digital twin based industrial automation and control system security architecture," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 669–680, 2019.

[3] D. Plets, W. Joseph, K. Vanhecke, E. Tanghe, L. Martens, S. Bouckaert, I. Moerman, and P. Demeester, "Validation of path loss by heuristic prediction tool with path loss and rssi measurements," in *2010 IEEE Antennas and Propagation Society International Symposium*. IEEE, 2010, pp. 1–4.

[4] F. Tao, H. Zhang, A. Liu, and A. Y. Nee, "Digital twin in industry: State-of-the-art," *IEEE Transactions on industrial informatics*, vol. 15, no. 4, pp. 2405–2415, 2018.

[5] Y. Wu, K. Zhang, and Y. Zhang, "Digital twin networks: A survey," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 13 789–13 804, 2021.

[6] "Fundamentals of Radiowave Propagation — link.springer.com," https://link.springer.com/chapter/10.1007/978-94-011-7027-7_2, [Accessed 24-08-2024].

[7] M. E. Diago-Mosquera, A. Aragón-Zavala, and G. Castañón, "Bringing it indoors: A review of narrowband radio propagation modeling for enclosed spaces," *IEEE Access*, vol. 8, pp. 103 875–103 899, 2020.

[8] R. S. de Souza and R. D. Lins, "A new propagation model for 2.4 ghz wireless lan," in *2008 14th Asia-Pacific Conference on Communications*, 2008, pp. 1–5.

[9] T. Zygiridis, E. Kosmidou, K. Prokopidis, N. Kantartzis, C. Antonopoulos, K. Petras, and T. Tsiboukis, "Numerical modeling of an indoor wireless environment for the performance evaluation of wlan systems," *IEEE Transactions on Magnetics*, vol. 42, no. 4, pp. 839–842, 2006.

[10] T. Chrysikos, G. Georgopoulos, and S. Kotsopoulos, "Site-specific validation of itu indoor path loss model at 2.4 ghz," in *2009 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks Workshops*, 2009, pp. 1–6.

[11] G. Wölfle, R. Wahl, P. Wertz, P. Wildbolz, and F. Landstorfer, "Dominant path prediction model for indoor scenarios," in *German Microwave Conference (GeMIC)*, vol. 27. Ulm Germany, 2005.

[12] A. G. Dimitriou, S. Siachalou, A. Bletsas, and J. N. Sahalos, "An efficient propagation model for automatic planning of indoor wireless networks," in *Proceedings of the Fourth European Conference on Antennas and Propagation*. IEEE, 2010, pp. 1–5.

[13] D. Plets, W. Joseph, K. Vanhecke, E. Tanghe, and L. Martens, "Simple indoor path loss prediction algorithm and validation in living lab setting," *Wireless Personal Communications*, vol. 68, pp. 535–552, 2013.

[14] M. B. Khrouf, M. Ayadi, S. B. Romdhane, N. Saghrouni, S. Tabbane, and Z. Belhadj, "Indoor prediction of propagation using dominant path: study and calibration," in *2005 12th IEEE International Conference on Electronics, Circuits and Systems*. IEEE, 2005, pp. 1–4.

[15] "Ray tracing propagation model - MATLAB - MathWorks United Kingdom — uk.mathworks.com," https://uk.mathworks.com/help/comm/ref/rfprop.raytracing.html, [Accessed 24-08-2024].

[16] F. Fuschini, H. El-Sallabi, V. Degli-Esposti, L. Vuokko, D. Guiducci, and P. Vainikainen, "Analysis of multipath propagation in urban environment through multidimensional measurements and advanced ray tracing simulation," *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 3, pp. 848–857, 2008.

[17] C.-F. Yang, B.-C. Wu, and C.-J. Ko, "A ray-tracing method for modeling indoor wave propagation and penetration," *IEEE transactions on Antennas and Propagation*, vol. 46, no. 6, pp. 907–919, 1998.

[18] F. Fuschini, H. El-Sallabi, V. Degli-Esposti, L. Vuokko, D. Guiducci, and P. Vainikainen, "Analysis of multipath propagation in urban environment through multidimensional measurements and advanced ray tracing simulation," *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 3, pp. 848–857, 2008.

[19] C. Pyo, H. Sawada, and T. Matsumura, "A deep learning-based indoor radio estimation method driven by 2.4 ghz ray-tracing data," *IEEE Access*, 2023.

[20] U. Masood, H. Farooq, and A. Imran, "A machine learning based 3d propagation model for intelligent future cellular networks," in *2019 IEEE global communications conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.

[21] K. Kedjar, L. Talbi, and M. Nedil, "Efficient indoor propagation channel prediction based on deep learning approach," *IET Microwaves, Antennas & Propagation*, vol. 15, no. 13, pp. 1671–1687, 2021.

[22] A. Sari and A. Alzubi, "Path loss algorithms for data resilience in wireless body area networks for healthcare framework," in *Security and Resilience in Intelligent Data-Centric Systems and Communication Networks*. Elsevier, 2018, pp. 285–313.

[23] "Path loss - Wikipedia — en.wikipedia.org," https://en.wikipedia.org/wiki/Path_loss#cite_note-1.

[24] "RSRP and RSRQ - Teltonika Networks Wiki — wiki.teltonika-networks.com," https://wiki.teltonika-networks.com/view/RSRP_and_RSRQ, [Accessed 03-09-2024].

[25] V. Trevisan, "Comparing robustness of MAE, MSE and RMSE — towardsdatascience.com," https://towardsdatascience.com/comparing-robustness-of-mae-mse-and-rmse-6d69da870828.

[26] "www.etsi.org - /deliver/etsi_tr/138900_138999/138901/16.01.00_60/ — etsi.org," https://www.etsi.org/deliver/etsi_tr/138900_138999/138901/16.01.00_60/.

[27] admin, "Hidden Layers in Neural Networks | i2tutorials — i2tutorials.com," https://www.i2tutorials.com/hidden-layers-in-neural-networks/, [Accessed 04-09-2024].