Indoor Radio Tomographic Imaging with Sparse Device Topologies

W.D.N.C Weerasiri

2025

Indoor Radio Tomographic Imaging with Sparse Device Topologies

Nipun Weerasiri

Index No : 20002033

Supervisor: Dr. Asanka P. Sayakkara

June 2025

Submitted in partial fulfillment of the requirements of the
B.Sc in Computer Science Final Year Project (SCS4224)

Declaration

I certify that this dissertation does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any university and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, be made available for photocopying and for interlibrary loans, and for the title and abstract to be made available to outside organizations.

Candidate Name : W. D. N. C. Weerasiri

Signature of Candidate **Date:** 2025.06.30

This is to certify that this dissertation is based on the work of Mr. W.D.N.C Weerasiri under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Supervisor Name : Dr. Asanka P. Sayakkara

Signature of Supervisor **Date:** 2025-06-30

Abstract

Device-Free Localization (DFL) using Wi-Fi signals has emerged as a compelling solution for indoor tracking without requiring individuals to carry any device. This thesis investigates the use of Channel State Information (CSI) for Radio Tomographic Imaging (RTI) in sparse network deployments. By developing a custom round-robin protocol over ESP-NOW and a lightweight data processing pipeline, it was shown that effective localization can be achieved with as few as four ESP32 nodes. CSI-based models significantly outperformed traditional Received Signal Strength Indicator (RSSI) methods, particularly in non-line-of-sight (NLOS) and multipath-rich environments. Feature importance analysis highlighted the stability and informativeness of CSI amplitude features for localization tasks. Although limitations such as environmental sensitivity and payload constraints were identified, the study proposes future improvements including phase calibration, domain adaptation, protocol enhancements, and the use of advanced machine learning models. This research contributes to the development of practical, low-cost, and scalable DFL systems suitable for indoor environments.

Preface

This document has been produced for the partial fulfilment of the requirements of the B.Sc. in Computer Science (Hons) Final Year Project in Computer Science (SCS4124).

This dissertation presents a study on DFL using CSI to enhance RTI in sparse network deployments. A novel data collection protocol and lightweight processing pipeline were developed to enable efficient, real-time localization using a minimal number of ESP32 nodes.

The dissertation is organized as follows: Chapter 1 presents the introduction and background of the study, along with a brief overview of the entire dissertation. Chapter 2 discusses related work and existing techniques in the field. Chapter 3 describes the design and methodology of the proposed system, as well as the implementation process and technical challenges encountered. Chapter 4 provides a thorough evaluation of the system's performance. Finally, Chapter 5 concludes the research and outlines future directions for improvement.

This dissertation represents original work that I, under the guidance of my supervisor, have carried out. Any material drawn directly from other sources has been properly referenced; all other content is my own contribution.

Acknowledgement

I would like to express my sincere gratitude to my supervisor, Dr. Asanka P. Sayakkara, Senior Lecturer at the University of Colombo School of Computing for his expertise, enthusiasm, patience, motivation and encouragement. I would also like to thank Ms. Nethmi Rodrigo for her introduction to the field and for giving me valuable advice and assistance. Their guidance helped me in every aspect of my undergraduate research study. Without it, I would not have been able to complete my work and dissertation.

I also extend my sincere thanks to the entire staff and my colleagues at the University of Colombo School of Computing (UCSC) for their constructive input, support, and collaboration, which have significantly contributed to the success of this research.

This dissertation is also dedicated to my loving family and friends who has been an immense support to me throughout this journey of life. Without them, my effort would have been worth nothing. Their love, support, and patience inspire me to overcome all the obstacles in life and achieve goals with success.

Contents

List of Figures viii

List of Tables xi

Acronyms xii

1 Introduction 1

1.1 Background 1

1.2 Research Gap 2

1.3 Research Questions 3

1.4 Motivation 4

1.5 Aim and Objectives 4

1.5.1 Project Aim 4

1.5.2 Objectives 4

1.6 Research Scope 5

1.6.1 In Scope 5

1.6.2 Out of Scope 5

1.7 Significance of the Research 5

1.7.1 Concerning Related Theories and Similar Works 5

1.7.2 Research Contribution and Benefits to Society 5

2 Literature Review 6

2.1 Radio Frequency Signal Metrics 7

2.1.1 Received Signal Strength Indicator 7

2.1.2 Channel State Information 8

2.1.3 Other Metrics . . . 9
2.2 Radio Tomographic Imaging . . . 9
2.3 RTI Topology . . . 12

3 Methodology **13**
3.1 Preliminary Study . . . 13
3.1.1 Evaluation Plan . . . 15
3.1.2 Results and Discussion . . . 16
3.2 Custom Protocol for Data Collection . . . 18
3.2.1 Protocol Overview . . . 19
3.2.2 Firmware . . . 20
3.2.3 Data logging . . . 21
3.2.4 Limitations . . . 23
3.3 Data Collection . . . 23
3.3.1 Collection Setup . . . 23
3.3.2 Collection Methodology . . . 26
3.4 Data Preparation . . . 26
3.4.1 Data Preprocessing . . . 26
3.4.2 Feature Extraction . . . 27
3.5 Data Analysis and Model Development . . . 28
3.5.1 Classification-Based Localization . . . 28
3.5.2 Regression-based Approach . . . 30

4 Results and Evaluation **33**
4.1 Classification-based Approach . . . 33
4.2 Regression-based Approach . . . 39
4.2.1 Using 9 coordinates . . . 39
4.2.2 Using 12 coordinates . . . 48

5 Conclusion **52**
5.1 Conclusion on the Research Questions . . . 52

5.1.1 How can we reduce the number of nodes required for effective RTI be reduced? 52
5.1.2 Does CSI outperform RSSI for RTI tasks in sparse network deployments? 52
5.2 Contributions 53
5.3 Limitations 53
5.4 Future Work 54

List of Figures

1.1 Experimental setup used by Dang et al. (2019) 2
1.2 Weighing matrix in multipath-RTI (Kim et al. 2022) 3

2.1 Shadowing of a link caused by objects in the line-of-sight (LOS) of the two nodes (Wilson & Patwari 2010) 10
2.2 Experimental setup used by Patwari & Wilson (2010) 12

3.1 Schematic diagram of the area of interest, showing the placement of the two transmitters and the floor layout. 14
3.2 Testbed setup in bedroom environment. 15
3.3 Evaluation chart flowchart 16
3.4 Accuracy of subcarriers for activity "Standing-Stationary" on different voxels. 17
3.5 Behavior of devices at a time index in the custom round-robin protocol. 19
3.6 Schematic diagram of the Area of Interest (AoI) with four devices. . 24
3.7 Testbed setup in the empty 5th floor environment (Loc 1). 25
3.8 Testbed setup in the empty 5th floor environment (Loc 2). 26
3.9 Testbed setup in the 4th-year laboratory environment. 27
3.10 Voxel-to-coordinate mapping used for regression. Each square represents a voxel in the 3×3 grid, labeled by its (x, y) coordinate. Additional coordinates at the corners of the center voxel indicate extended positions used in one specific setup. 31

4.1 Confusion Matrix — Empty Floor (using CSI). 33

4.2 Confusion Matrix — Laboratory (using CSI). 34
4.3 Confusion Matrix — Bedroom (using CSI). 34
4.4 Overall feature importance within the same session across all environments. 35
4.5 Confusion Matrix — Cross-session in same part of empty floor (using CSI). 36
4.6 Confusion Matrix — Cross-session in different part of empty floor (using CSI). 37
4.7 Confusion Matrix — Empty Floor (using RSSI). 38
4.8 Confusion Matrix — Laboratory (using RSSI). 38
4.9 Confusion Matrix — Bedroom (using RSSI). 39
4.10 Confusion Matrix — Empty Floor Cross-Session (using RSSI). . . . 39
4.11 Scatterplots for voxels 1-9 with just the actual/expected location. . 40
4.12 Scatterplots with time heatmap for voxels 1-9 in the empty floor environment (train-test split). 41
4.13 Scatterplots with time heatmap for voxels 1-9 in the laboratory environment (train-test split). 42
4.14 Scatterplots with time heatmap for voxels 1-9 in the bedroom environment (train-test split). 43
4.15 Overall feature importance within the same session across all environments. 44
4.16 Scatterplots with time heatmap for voxels 1-9 in the empty floor environment (Loc 1) (cross-session). 45
4.17 Scatterplots with time heatmap for voxels 1-9 in the empty floor environment (Loc 2) (cross-session). 46
4.18 Scatterplots with time heatmap for voxels 1-9 in the empty floor environment (Loc 2) taken at a different time (cross-session). 47
4.19 Scatterplots for the 12 coordinates with just the actual/expected location. 49

4.20 Scatterplots with time heatmap for the 12 coordinates in the empty floor environment (Loc 2) (train-test split). 50
4.21 Scatterplots with time heatmap for the 12 coordinates in the empty floor environment (Loc 2) (cross-session). 51

List of Tables

2.1 Comparison of RSSI and CSI 9
2.2 Overview of Techniques and References by Signal Type 11

4.1 Mean Absolute Error (MAE) and R^2 Score for train-test split evaluation. 40
4.2 MAE and R^2 Score for cross-session evaluation. 44
4.3 MAE and R^2 Score for 12 coordinates. 48

Acronyms

AoA Angle of Arrival

AoI Area of Interest

CSI Channel State Information

CT Computed Tomography

LDPL Log-normal Distance Path Loss

LOS line-of-sight

MAC Medium Access Control

MIMO Multiple Input and Multiple Output

NIC Network Interface Card

NLOS non-line-of-sight

OFDM Orthogonal Frequency Division Multiplexing

PHY Physical

RADAR RAdio Detecting And Ranging

RF Radio Frequency

RFID Radio Frequency Identification

RSS Received Signal Strength

RSSI Received Signal Strength Indicator

RTI Radio Tomographic Imaging

SWAT Special Weapons And Tactics

ToA Time of Arrival

ToF Time of Flight

UWB ultra-wideband

Chapter 1

Introduction

1.1 Background

The resurgence and proliferation of context-aware computing and location-based services have made the task of localization of people and objects within a given area more important than ever. It is the process of estimating or predicting the location of an individual relative to a collection of anchor points in a given AoI. Localization has seen potential use cases and applications in various fields including but not limited to disaster management, pervasive computing, health care, surveillance, home security, and elderly care (Zafari et al. 2017, Gholamhosseini et al. 2019).

Active localization involves tracking objects to carry a transmitter, receiver, or transmitter (Zhang et al. 2007). In outdoor environments, GPS tracking can be used to track objects, such as smartphones, accurately within a 5-meter radius away from obstructions. However, their accuracy degrades with the presence of buildings, bridges, etc (*GPS.gov: GPS Accuracy* 2022). This makes them impractical for use in an indoor setting since the potential location of the user indoors could theoretically span the entire AoI. Radio Frequency Identification (RFID) technology offers the capability to monitor objects, contingent upon the presence of an affixed tag, necessitating premeditation. This makes it ineffective for localizing subjects in an emergency. Moreover, RFID tags can only operate over a short range and are relatively expensive.

Passive localization (device-free) techniques do not require the objects to carry any tracking device. Instead, they may use localization techniques that include acoustic, Bluetooth, infrared, pressure, ultrasound, Wi-Fi, ZigBee, etc (Yang et al. 2013). Techniques using video, pressure, and infrared often have a higher cost of deployment, suffer from occlusion and short range, and may cause privacy concerns. Moreover, video-based systems are ineffective in low-light conditions.

1.2 Research Gap

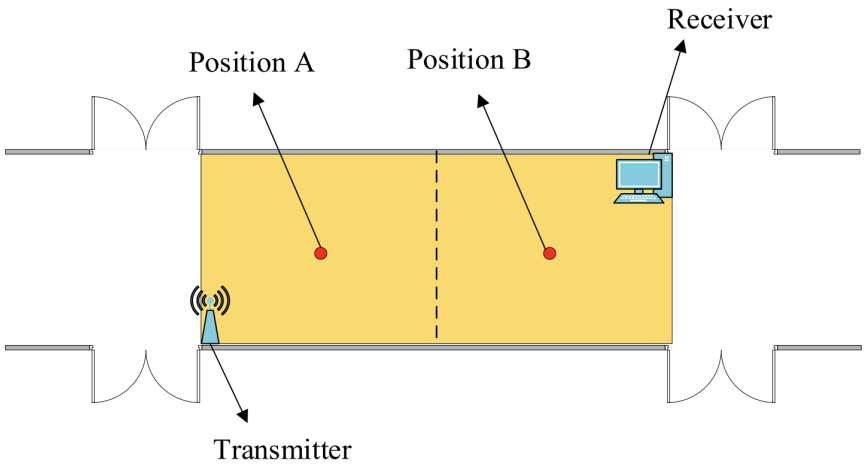


Figure 1.1: Experimental setup used by Dang et al. (2019)

One initial motivation for introducing indoor localization methods using RTI is for use in emergencies where information about a particular AoI is limited and time is of the essence.

Literature surrounding RTI methods have leaned towards improving performance by increasing number of nodes or employing learning methods. While most existing work has demonstrated high accuracy with several proposed methods, the setups are too complex or require a fingerprinting database to work well. Majority of papers using non-learning approaches use more than 20 nodes in their setup. There have been no comparisons made for different devices used as nodes in the same RTI setup.

For example, the setup used by Dang et al. (2019) (Figure 1.1) looks ideal

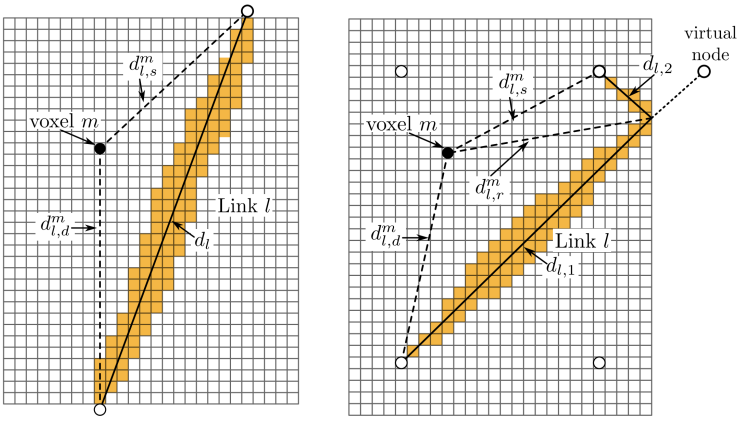


Figure 1.2: Weighing matrix in multipath-RTI (Kim et al. 2022)

for emergencies since it only requires a single transmitter and receiver on either end. However, collecting data to build the fingerprinting database requires a lot of effort. There is a need to improve on setups similar to these but with little to no access to the site.

The recent development of multipath-RTI for localization using virtual nodes (Kim et al. 2022, Ikegami & Kim 2022) looks promising for sparse topologies (Figure 1.2). Here, the voxel size does not need to depend on the density of nodes, increasing its resolution. While it requires fewer nodes, it uses millimeter-waves of frequencies greater than 50 GHz, making it infeasible.

1.3 Research Questions

1. **How can we reduce the number of nodes required to effectively perform RTI using Wi-Fi?**

   Hypothesis: There should be a way to reduce the number of nodes required to effectively perform RTI using Wi-Fi by leveraging more advanced signal processing techniques and richer data sources such as CSI. Reducing node density while maintaining accuracy could make RTI systems more scalable and cost-effective.

2. **Which of the parameters, CSI or RSSI, would perform better in**

sparse networks?

Hypothesis: While previous work such as Kim et al. (2022) utilized RSSI measurements in their setup, no direct comparison was made between RSSI and CSI performance in sparse networks. This research hypothesizes that incorporating both RSSI and CSI measurements may enhance localization accuracy, with CSI expected to perform better due to its finer-grained information.

1.4 Motivation

Imagine an emergency rescue situation in which a building has collapsed and it is critical to find survivors inside the debris or in a section of the collapsed building. It is nearly impossible to implement a standard configuration under such circumstances. Furthermore, maintaining a regular topology surrounding the region could be fatal in time-sensitive scenarios, particularly in Special Weapons And Tactics (SWAT) and military operations.

RTI has a big impact on real-world applications. Despite its potential, the current state of research has not progressed enough for commercial and practical usage. The goal of this study is to improve the application of RTI in practical settings.

1.5 Aim and Objectives

1.5.1 Project Aim

This project focuses to exploring the ability to localize objects in an AoI using RTI when the setup is deployed in a sparse network using multipath-RTI.

1.5.2 Objectives

- Exploring the ability to form an image/detect the occupancy in an enclosed area when the setup used is sparse and deployed in a non-uniform manner.

- Evaluating the behaviour of RSSI and CSI parameters under sparse network conditions.

1.6 Research Scope

1.6.1 In Scope

- Developing a scalable protocol of collecting CSI and RSSI data from multiple devices.
- Investigating the capability of RTI systems in randomly deployed sparse setups using Wi-Fi.

1.6.2 Out of Scope

- The project's feasibility will not be tested in large or outdoor environments.

1.7 Significance of the Research

1.7.1 Concerning Related Theories and Similar Works

There was not enough research done on the issue of when the setup is nonuniform, according to the relevant material that was examined during the preliminary literature review. The findings of this study will make a substantial contribution to the eventual development of a setup that can self-localize in the best possible way to produce a realistic image of an enclosed region.

1.7.2 Research Contribution and Benefits to Society

The results of this study would greatly enhance the applicability of the RTI system in targeted real-world scenarios. Deploying RTI systems around an area of interest in difficult situations could provide significant insight without requiring direct access to compromised regions, thereby saving the lives of military and security personnel.

Chapter 2

Literature Review

Some of the shortcomings of the techniques mentioned above alleviated the importance of employing Radio Frequency (RF) in passive localization. It is often cost-effective, can work over a large area, and can pass through non-metallic walls and obstacles. In the late 20th century, as the field of wireless communication began to flourish, researchers recognized the potential of RF signals beyond communication. Bahl & Padmanabhan (2000) developed an RF-based system to track users inside buildings using signal propagation models and physical measurements to estimate and track user locations. RF-based systems have been used to estimate locations using ultra-wideband (UWB). However, they require specialized military equipment that is costly and not available to the general public Zhao et al. (2013). Zhang et al. (2007) devised a RF-based localization technique that utilized RSSI measurements of an AoI covered by a wireless sensor grid. This system could also keep count of the number of objects that are not close to each other by tallying the resulting number of clusters.

2.1 Radio Frequency Signal Metrics

Various metrics have been used to characterize and analyze the properties of RF signals. These metrics provide valuable insights into signal strength, quality, and propagation characteristics which vary in granularity, stability, resolution, and level of detail.

2.1.1 Received Signal Strength Indicator

RSSI is a metric used to gauge the strength or intensity of the RF signal received by a device's antenna. This metric provides details about the signal's strength at the receiving end of a communication link. It is a function of the length of the link between the transmitting and receiving node i.e. as the propagation distance increases, the signal power decreases and is the basis of the Log-normal Distance Path Loss (LDPL) propagation model. Using this model, the RSSI can be estimated using the following equation (Kumar et al. 2009):

$$P(r) = P_0 - 10n \log_{10} \left(\frac{d}{d_0} \right) \tag{2.1}$$

where:

$P(d)$: Received Signal Strength (RSS) at distance d
P_0 : reference RSS
n : path loss exponent
d : distance from transmitter
d_0 : reference distance from the transmitter where RSS is P_0

A key limitation of RSSI is its susceptibility to temporal fluctuations in elaborate indoor environments due to multipath effects (Farahani 2008). When signals take multiple paths due to reflection, diffraction, and scattering, the phase of these signals can change rapidly. This fast-changing phase can lead to the superimposition of multipath signals, making it challenging for RSSI measurements to accurately capture the characteristics of each signal component, making RSSI a coarse-grained feature.

2.1.2 Channel State Information

In Orthogonal Frequency Division Multiplexing (OFDM) systems, data is transmitted and obtained across numerous overlapping frequency bands simultaneously, known as sub-carriers. For this, data transmission and reception involve using multiple antennas through Multiple Input and Multiple Output (MIMO) technology. As the data travels via the channel, it may take various routes, leading to attenuation caused by path loss, fading due to multipath, and fading due to shadowing. It contains details on the properties of the sub-carriers including the amplitude and phase, and different sub-carriers can have different signal strengths. The frequency and spatial diversity provided by CSI can be utilized to mitigate the effects of multipath propagation (Chapre et al. 2015).

Unlike RSSI, which only captures signal strength, CSI records both the signal strength and phase information for each OFDM sub-carrier and between every pair of transmit-receive antennas (Al-qaness et al. 2019). CSI is a more detailed RF measurement than RSSI. It offers a more comprehensive view of the RF channel, encompassing details such as amplitude, phase, and frequency response. CSI measurements provide insights into the characteristics of the propagation medium, including multipath effects, fading, and interference. A tool developed by Halperin et al. (2011) can record CSI information based on the IEEE 802.11 standard. However, while RSSI is almost ubiquitous, CSI is only available in a handful of Network Interface Cards (NICs).

In an OFDM transmission system, the measured signal of the j-th receiving antenna in the frequency domain can be modeled as (Al-qaness et al. 2019):

$$y_j(t) = \sum_{i=1}^{n_t} h_{i,j} x_i(t) + \eta_j(t) \tag{2.2}$$

where:

Metrics	**RSSI**	**CSI**
Level of Detail	Coarse	Fine
Time Resolution	Low	High
Temporal Stability	Lower	Higher
Frequency Resolution	Low	High
Ubiquity	Widely available	Limited

Table 2.1: Comparison of RSSI and CSI

i : index of transmit antenna

j : index of receive antenna

$y_j(t)$: received signal at the jth antenna at time t

$h_{i,j}$: channel fading factor between the ith and jth antenna

$x_i(t)$: signal transmitted by ith antenna at time t

$\eta_j(t)$: additive noise at jth antenna at time t

2.1.3 Other Metrics

Some features that can extracted from RF waves in the time domain include Time of Arrival (ToA) and Time of Flight (ToF). The specific instance when a radio signal sent from a transmitter reaches a remote receiver is known as ToA. The time it takes for this transmission to travel from the transmitter to the receiver is known as the ToF. Angle of Arrival (AoA), the direction in which the signal has been received, introduces an additional dimension orthogonal to the distance for geometric mapping purposes.

2.2 Radio Tomographic Imaging

RTI utilizes the impact of obstacles on the propagation of wireless signal to infer the presence and movement of objects in an AoI (Yang et al. 2013). An RTI system comprises access points or signal emitters, monitoring points for signal metric

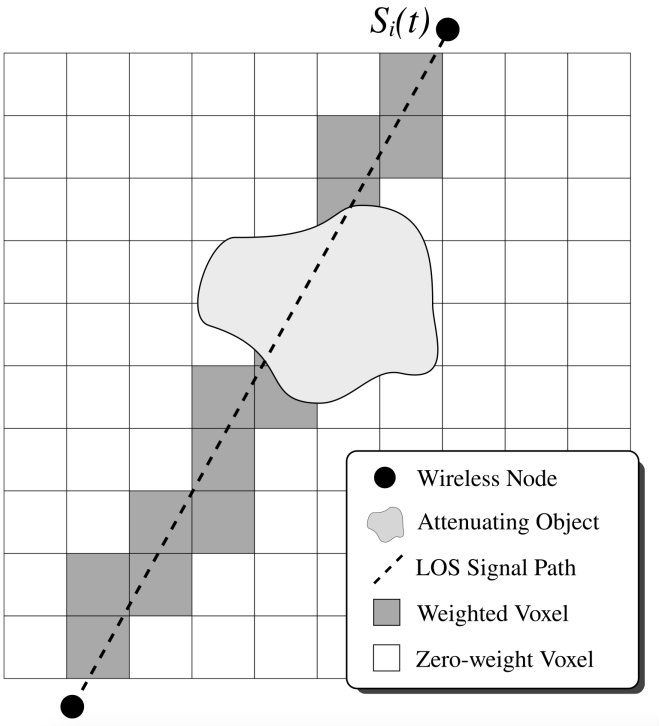


Figure 2.1: Shadowing of a link caused by objects in the LOS of the two nodes (Wilson & Patwari 2010)

measurement, and an application server responsible for processing the measurements to determine the locations of pertinent subjects within an AoI.

It draws concepts from existing Computed Tomography (CT) techniques like X-Ray and RAdio Detecting And Ranging (RADAR) systems. In CT methods, physical measurements are conducted along diverse trajectories to approximate how transmission parameters are distributed spatially across the medium (Kak et al. 1988). RADAR systems transmit RF pulses and images the surrounding area. The distance to the scatter is a function of the delay between transmission and reception (Bahl & Padmanabhan 2000). On the contrary, RTI, which uses lower-bandwidth RF-waves, have to deal with substantial NLOS transmission which complicates the localization process. Conversely, it does not require costly and specialized hardware (Wilson & Patwari 2010).

Various approaches have been proposed, using different signal metrics with statistical, and learning methods. Statistical (non-learning) methods were often used

in approaches where information about the AoI may not be known beforehand, or when the environmental conditions may change. An advantage is its adaptability in various environments as well as fast computational time. Learning methods are often employed in a known AoI with a substantial amount of data collected at various points. It uses data to learn the complex relationships between the signal measurements and the environment. Give enough data, it may be able to achieve higher accuracy than statistical methods, however it requires more computational resources and only works on known environments.

Table 2.2: Overview of Techniques and References by Signal Type

Signal Type	**Technique**	**References**
RSSI	Geometric Mapping	Wilson & Patwari (2010), Wilson & Patwari (2012), Kaltiokallio et al. (2012*b*), Kim et al. (2019), Ikegami & Kim (2022), Kim et al. (2022)
	Fingerprinting	Youssef & Agrawala (2005), Seifeldin & Youssef (2009), Xu et al. (2013), Mager et al. (2015), Mahfouz et al. (2014), Lu et al. (2016), Xu et al. (2016), Subhan et al. (2020), Liu et al. (2019), Hoang et al. (2019), Denis et al. (2020), Poulose & Han (2021), Su et al. (2023)
	Variance	Wilson & Patwari (2011), Kaltiokallio et al. (2012*a*)
	Kernel Distance	Zhao et al. (2013)
CSI	Geometric Mapping	Wu et al. (2012)
	Fingerprinting	Xiao et al. (2012), Sanam & Godrich (2018), Dang et al. (2019), Chapre et al. (2015), Wang et al. (2017), Chen et al. (2017), Hsieh et al. (2019)

2.3 RTI Topology

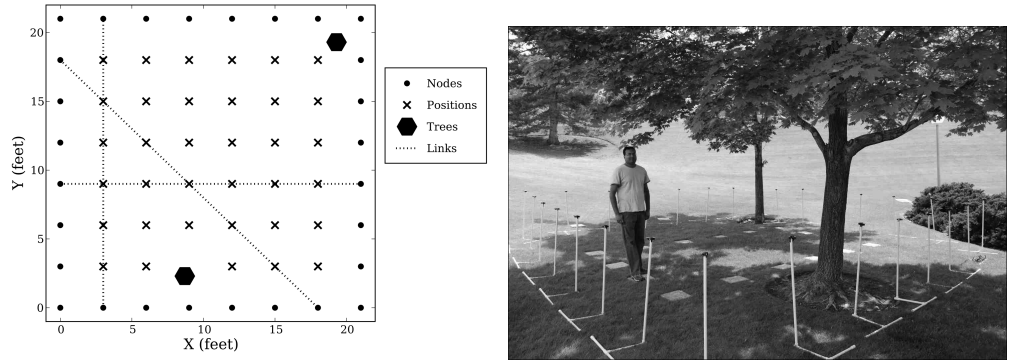


Figure 2.2: Experimental setup used by Patwari & Wilson (2010)

Many setups involves non-learning approaches requires a large number of nodes arranged in an almost perfect grid like pattern (Figure 2.2). The voxels created are a result of the intersection of the LOS of oppositely placed nodes. That means the resolution and accuracy of the image generated would correspond to the number of nodes. Adding more nodes to the RTI network reduces the magnitude of errors (Kaltiokallio et al. 2012*a*). However, there was not a lot of work done to reduce the number of nodes to make it practical in real-life situations.

Chapter 3

Methodology

3.1 Preliminary Study

This study builds upon prior work conducted by Rodrigo & Sayakkara (2024), which, although unpublished, served as the basis for a preliminary investigation into indoor localization using RTI with CSI in sparse wireless networks. The objective of this stage was to replicate Rodrigo's experimental setup in a new environment in order to assess the reproducibility and robustness of the methodology. This approach aims to evaluate the reliability and generalizability of the system under varying spatial and environmental conditions.

- **Experimental Setup:** The testbed consisted of a square area measuring 1.8m×1.8m, subdivided into nine equal square voxels, each measuring 0.6m× 0.6m, as depicted in Figure 3.1. Two ESP32 microcontrollers were employed, with one configured as a transmitter and the other as a receiver. Both devices operated at a tick rate of 100 Hz using the ESP32 CSI Toolkit (Hernandez & Bulut 2020).

  The receiver was connected to a MacOS 15.1 laptop via a serial connection for real-time CSI data acquisition. The transmitter was powered externally and positioned just outside voxel 8, while the receiver was located just outside voxel 2. Both devices were mounted approximately 0.7 meters above the ground (see Figure 3.2).

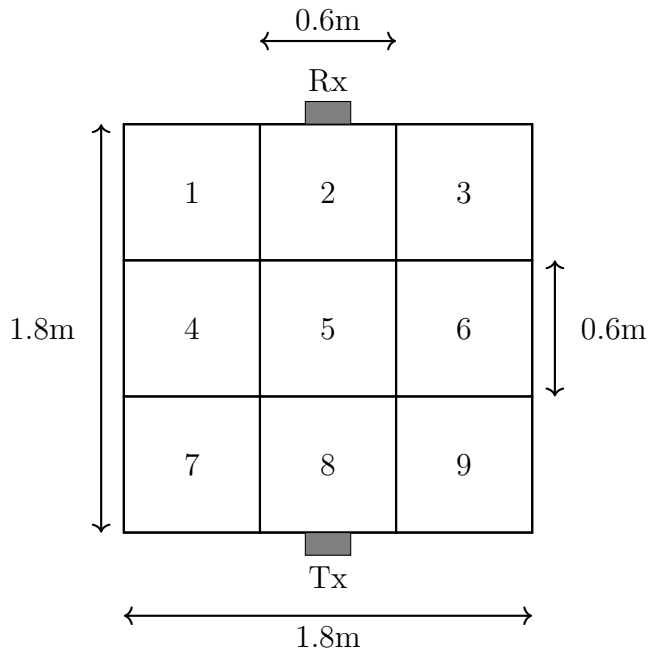


Figure 3.1: Schematic diagram of the area of interest, showing the placement of the two transmitters and the floor layout.

- **Data Collection Protocol:** CSI data were collected for a duration of 30 seconds within each voxel, with the subject performing the following activity types:
  - *Standing-Stationary*: standing motionless.
  - *Standing-Moving*: standing with arm movements.
  - *Sitting-Stationary*: sitting cross-legged without movement.
  - *Sitting-Moving*: sitting cross-legged with arm movements.
  - *Walking*: dynamic movement within the voxel.

  These variations were designed to capture the effects of different postures and motion profiles on the CSI measurements. Data collection was performed in isolation with only one individual present in the room to minimize external interference.

- **Model Development and Evaluation:** A non-learning localization model

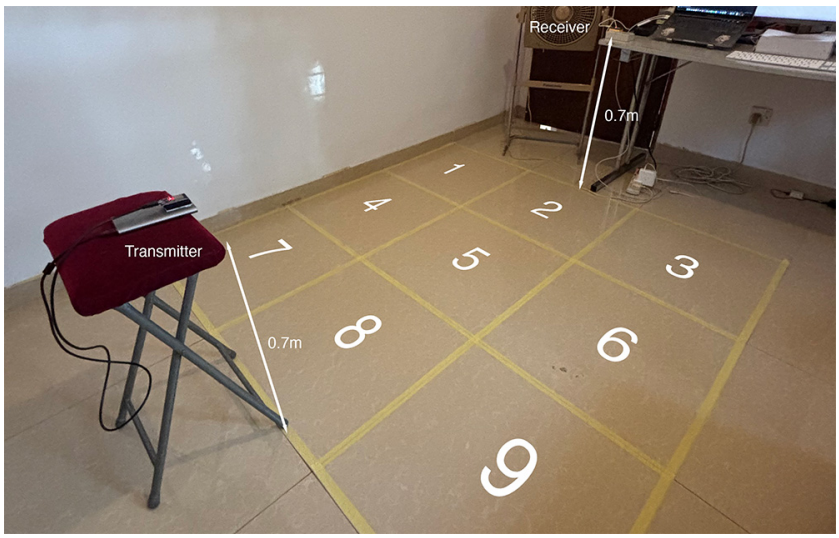


Figure 3.2: Testbed setup in bedroom environment.

was initially developed using the training dataset. Data preprocessing involved the removal of null subcarriers and subsequent storage in a structured database. Separate testing datasets were then collected and used to assess the model's accuracy in voxel prediction.

- **Future Work:** Future experiments will include diverse node placement topologies and environmental scenarios to investigate system performance under more complex conditions. These will encompass multi-user presence, the introduction of physical obstructions, and dynamic activity sequences to further evaluate localization robustness and adaptability.

3.1.1 Evaluation Plan

The evaluation process involved systematic testing of the localization model using both RSSI and CSI data across multiple network topologies (Figure 3.3). Performance was assessed based on the Euclidean distance between the predicted and actual voxel positions.

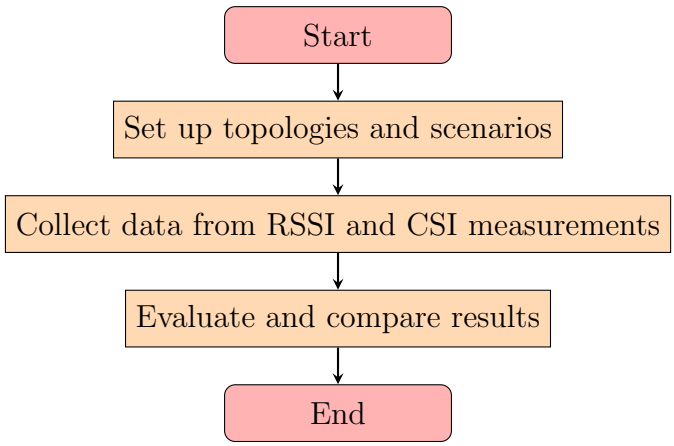


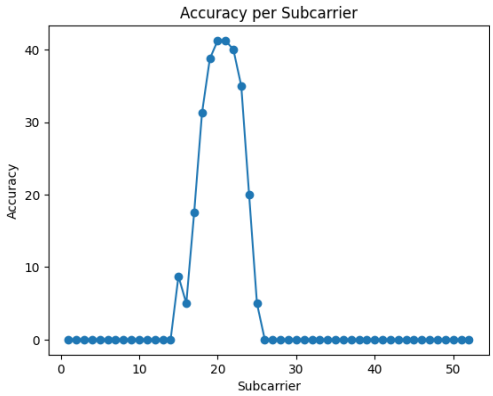
Figure 3.3: Evaluation chart flowchart

3.1.2 Results and Discussion

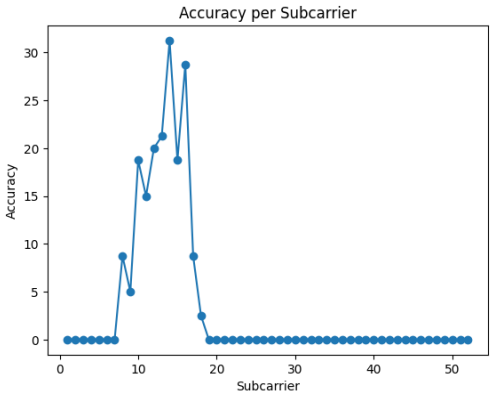
The initial evaluation involved predicting the user's voxel based on collected CSI data under various activity scenarios. A significant amount of effort was dedicated to interpreting legacy code—primarily implemented in Jupyter Notebooks—and adapting it for this analysis. Future iterations of the model will benefit from automation techniques, such as incorporating activity recognition to reduce manual labeling.

One challenge encountered was the inconsistency in subcarrier performance across activities and locations. For instance, Figure 3.4 presents subcarrier accuracy for the activity "Standing-Stationary" across all voxels, revealing significant variability. Even within a single voxel, such as voxel 6, two measurements (Figures 3.4d and 3.4e) demonstrated divergent subcarrier profiles.

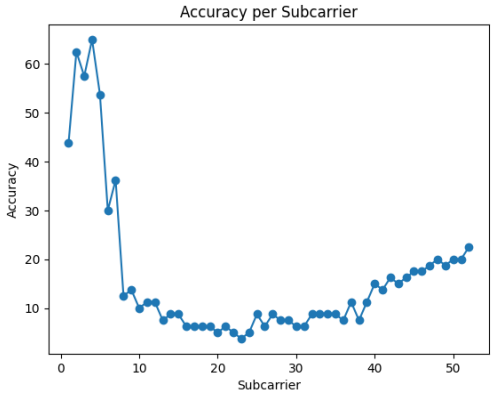
These findings suggest that due to the high variance in subcarrier-level accuracy, relying solely on non-learning methods may be insufficient for robust localization. Future efforts may require the integration of learning-based models or increased node density to enhance spatial resolution and classification reliability.



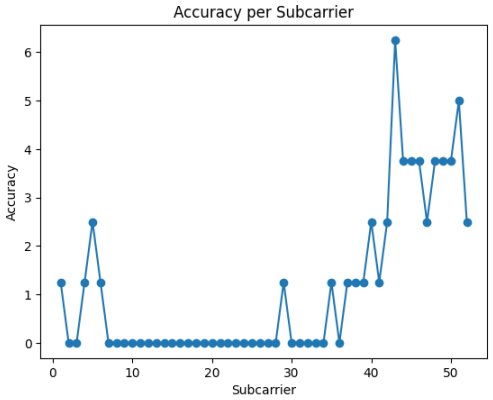
(a) "Standing-Stationary" on voxel 1.



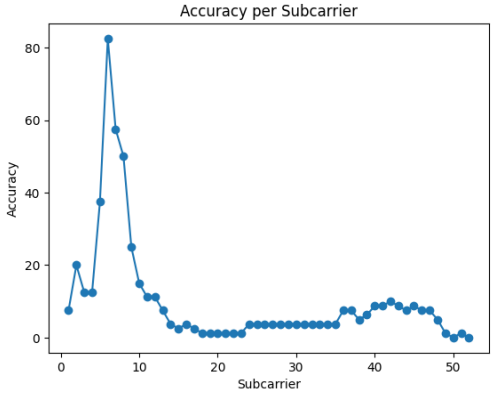
(b) "Standing-Stationary" on voxel 1.



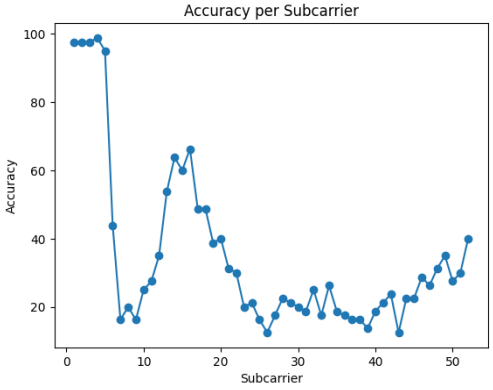
(c) "Standing-Stationary" on voxel 2.



(d) "Standing-Stationary" on voxel 6.



(e) "Standing-Stationary" on voxel 6.



(f) "Standing-Stationary" on voxel 9.

Figure 3.4: Accuracy of subcarriers for activity "Standing-Stationary" on different voxels.

3.2 Custom Protocol for Data Collection

To ensure reliable and interference-free data collection, a round-robin communication protocol was implemented, wherein only one device broadcasts an ESP-NOW packet at a time while other devices obtains the CSI of the packet.

This approach avoids packet collisions inherent in simultaneous transmissions over the shared Wi-Fi medium, which would otherwise degrade the quality and consistency of CSI measurements. Prior work by Halperin et al. (2011) emphasized the sensitivity of CSI to channel conditions and interference, reinforcing the importance of minimizing concurrent transmissions to maintain fidelity in wireless measurements.

By serializing transmissions, the system maintains a clear transmission timeline and provides clear associations between sender-receiver pairs, which is critical for accurate spatial modeling and feature extraction in RTI.

One key advantage of the proposed protocol is that only a single device needs to be connected to a computer to collect data. The other devices in the network only need to be powered on and do not require a direct connection. This setup helps reduce bandwidth usage and makes data collection more straightforward. The connected device is responsible for recording both its own CSI measurements and those received from the other devices, saving all data in a well-organized format for later analysis.

Another advantage is the number of links for a given number of nodes is increased. Most setups involve a sender-receiver pair, where one device sends out packets and another device reads the CSI data obtained from those packets Zhang et al. (2019) Lu et al. (2023) Choi et al. (2022). Only a single link exists between each pair, restricting spacial diversity. Given n devices, there would be $\frac{n}{2}$ links made for each pair. Moreover, data needs to be collected from multiple transceivers, making data collection complex. Some involve a single access point (AP) with multiple stations (STA) connected to it Wang et al. (2014) Schäfer (2022) Suroso et al. (2021) Yin et al. (2024) Brinke et al. (2023). While it makes it easier to collect data if the AP is measuring the CSI data received from the

other STAs, obtaining CSI data collected by the STAs will still be a challenge if spatial diversity is to be improved. If there are n devices, assuming only the AP is collecting data, there would be $n-1$ links in total. Our protocol increases the number of unique links for the given number of nodes to $P(n,2) = n(n-1)$.

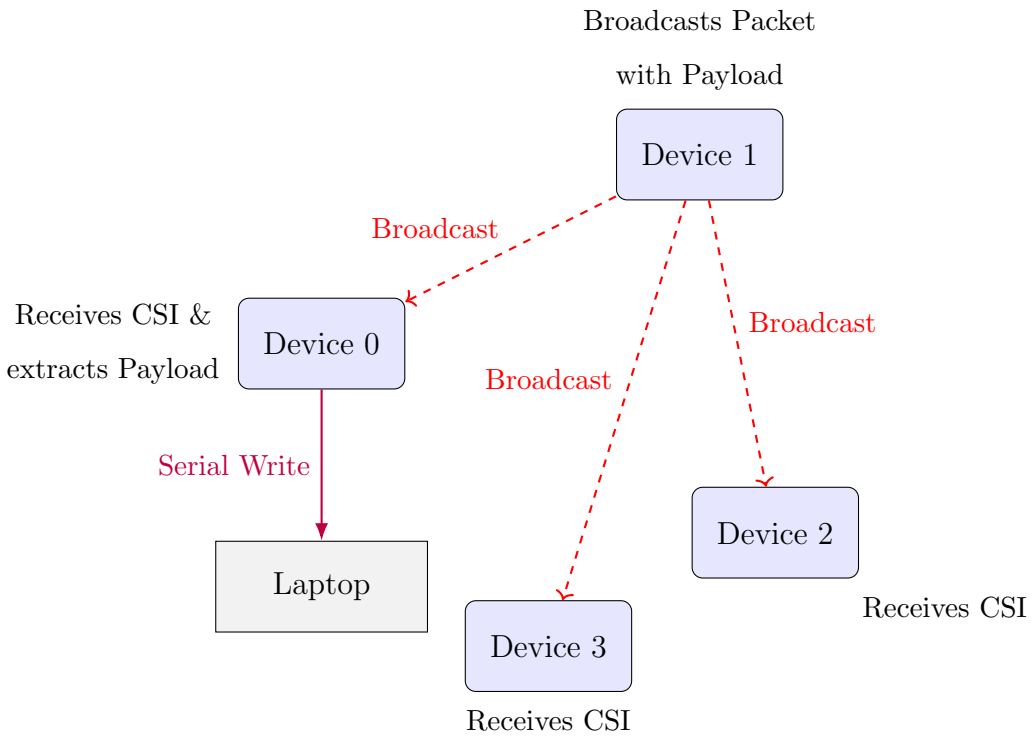


Figure 3.5: Behavior of devices at a time index in the custom round-robin protocol.

3.2.1 Protocol Overview

- The transmission starts from a designated node, referred to as Device 0, which is also connected via USB to a laptop.
- At any given time, one device acts as the sender, while the rest of the devices act as receivers, as shown in Figure 3.5.
- Each device takes turns being the sender across time intervals in a round-robin fashion, depending on the device that has broadcasted the last Wi-Fi packet.

- The receivers collects the CSI data and stores it temporarily in memory. To prevent the receivers from reading CSI data from packets sent from devices outside the network, each packet contains a magic number and an organization ID which are checked before temporarily storing the data.

- When it is the device's turn to send out data, the sender piggybacks the latest CSI data measured from the other devices.

- Device 0 writes to serial the payloads of all the packets received from other devices, as well as its own when sending out data. This data is read by the laptop and stored as CSV files for further processing. The logs include:

  – CSI data for each sender-receiver pair with the time-index.

  – Other metadata associated with the Wi-Fi packets, like RSSI, channel, signal mode, etc.

3.2.2 Firmware

Each ESP32 node operates using identical custom firmware that implements round-robin transmission logic with data piggybacking. When a device is flashed for the first time, it is manually assigned a unique device identifier (ID) in the range $[0, n-1]$ where n denotes the total number of devices in the network.

Nodes transmit payloads containing their own device ID, a time index, Wi-Fi metadata, and CSI from other devices. Transmission is event-driven, such that a node only initiates communication upon receiving a payload from the node with the preceding device ID in the round-robin sequence.

To ensure robustness against packet loss or delays, a timeout mechanism is incorporated. If a node does not receive the expected packet within a certain timeframe, it proceeds to transmit its payload. The timeout duration is proportional to the relative distance in the round-robin sequence between the current node and the node that sent the last received packet. This design ensures that devices wait longer to transmit if the last received packet originated from a node

further ahead in the logical sequence, thereby preserving the intended communication order while maintaining fault tolerance.

3.2.3 Data logging

During the data collection process, Device 0 — connected to a laptop via serial — acts as the central logger. Whenever any device transmits a packet, it includes its own CSI along with measurements from the other devices. All devices receive the broadcast and extract CSI locally, while Device 0 also parses the payload to extract the embedded measurements. These are then written to serial as comma-separated values for subsequent analysis.

An unexpected behavior was observed during payload parsing on Device 0. Specifically, while extracting the `csi_data_arr` from received payloads, an offset of 7 bytes per element had to be manually accounted for, depending on the index of the data being accessed from the `csi_data_arr` array. This discrepancy was not present when accessing the data locally on the transmitting device. The pseudocode for this is as shown:

```
for i from 0 to payload.csi_data_arr.len() - 1 do
        if payload.device_id is not 0 then
                offset = 7 x i
        else
                offset = 0
        end if
        print(address of payload.csi_data_arr[i] + offset)
end for
```

At present, the origin of this offset remains unclear. A potential explanation might involve unintended padding introduced during serialization. It is also possible that structure packing or the transmission process itself results in additional bytes being inserted between elements, which only affects devices receiving the data, not the sender. Further investigation is required to isolate and fully understand this anomaly.

Each logged entry records both metadata and the corresponding CSI trace, structured as follows:

Field	**Description**
`type`	Type of entry, in this case, CSI_DATA.
`time_index`	Index indicating when the packet was captured.
`device_id`	Identifier for the transmitting device.
`recv_device_id`	Identifier for the receiving device.
`mac`	MAC address of the transmitting device.
`rssi`	RSSI (in dBm).
`rate`	PHY rate encoding of the packet.
`sig_mode`	Index of protocol of the received packet (non-HT/HT/VHT).
`mcs`	Modulation Coding Scheme.
`cwb`	Index of Channel width (20 MHz/40 MHz).
`smoothing`	Whether channel estimate smoothing is recommended.
`not_sounding`	Whether the packet (Physical Layer Protocol Data Unit (PPDU)) is not a sounding packet.
`aggregation`	Whether packet aggregation is used.
`stbc`	Spatial Time Block Coding (STBC) (disabled/enabled).
`fec_coding`	Forward Error Correction type (for LDPC 11n packets).
`sgi`	Short Guard Interval (Long GI/Short GI).
`noise_floor`	Measured background noise level.
`ampdu_cnt`	Number of aggregated packets in an AMPDU.
`channel`	Primary channel on which the packet was received.
`secondary_channel`	Secondary channel on which the packet was received (non/above/below)
`timestamp`	Local device timestamp of packet arrival.
`ant`	Antenna index used for reception.
`sig_len`	Length of the signal field.
`rx_state`	State of packet (no errors/error number).

`len`	Length of CSI data buffer.
`csi_data`	CSI (complex values per subcarrier).

3.2.4 Limitations

- The payload size of ESP-NOW v2 is limited to 1490 bytes. As a result, only four readings, each containing CSI data and Wi-Fi packet metadata, can be transmitted at once. Consequently, if multiple transmission timeouts occur, older CSI readings stored on each device may be overwritten.
- To simplify memory management and adhere to the ESP-NOW payload size constraint, the maximum length of a CSI data array is limited to 256 bytes. However, devices operating in `HT` mode with a channel bandwidth of 40 MHz produce 384 bytes of CSI data. Under these conditions, the complete CSI array cannot be transmitted unless either the payload capacity of ESP-NOW is increased or dynamic memory allocation is supported in the firmware.
- The protocol does not include a mechanism for precisely controlling the transmission rate of packets over time. This limitation arises from the lack of time synchronization between devices, which was chosen to reduce system complexity. As a result, the amount of CSI data collected may vary depending on environmental conditions and hardware performance.

3.3 Data Collection

3.3.1 Collection Setup

Figure 3.6 presents the schematic layout of the experimental testbed, where four ESP32s (Device 0 to Device 3) are positioned at the corners of a 1.8m×1.8m square grid, which is divided into nine equal regions. All devices are placed directly on the floor level (0m height).

The ESP32 modules utilized in this study were configured to operate on the 2.4 GHz Wi-Fi band, using High Throughput (HT) mode with a channel bandwidth of

20 MHz. The devices were set to operate on channel 11 with a secondary channel positioned below the primary with STBC disabled.

Devices 1, 2, and 3 are powered independently using portable power banks, ensuring untethered operation and consistent power supply. Device 0 is powered via a USB connection to a MacBook Pro M1 laptop running MacOS Sequoia 15.4.1

A FastAPI-based backend Python web server runs on the laptop, continuously reading and processing serial input data from Device 0. The server dynamically labels the collected data based on incoming requests from a React-based frontend application, also hosted on the same laptop. This frontend interface is accessed through a mobile device, allowing for remote control and real-time interaction with the system.

For the purposes of data processing, analysis, and model development, the Python libraries `numPy`, `pandas`, `matplotlib`, and `scikit-learn` were used.

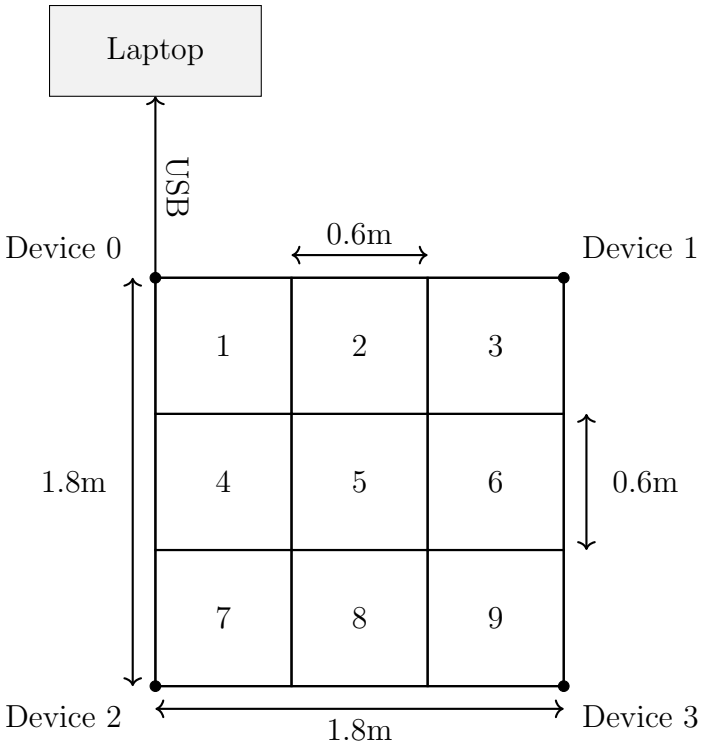


Figure 3.6: Schematic diagram of the AoI with four devices.

Data was collected across three distinct physical environments to capture diverse signal propagation characteristics:

- **Empty Floor:** An open and unoccupied space on the 5th floor of the university building, with only structural pillars located outside the region of interest. The environment was static, with minimal interference and no human activity during the data collection sessions. The testbed was setup in two different locations (see Figure 3.7 and Figure 3.8).

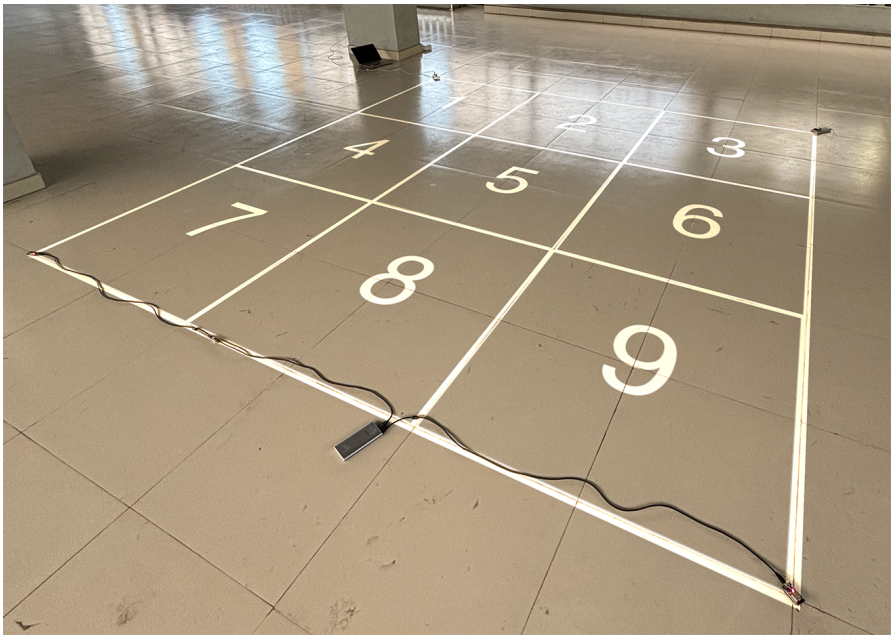


Figure 3.7: Testbed setup in the empty 5th floor environment (Loc 1).

- **Laboratory:** A busy 4th-year student laboratory containing chairs, desks, and other furniture in close proximity to the testbed area. This setting introduced a more cluttered environment with potential for signal multipath (see Figure 3.9).

- **Bedroom:** A domestic bedroom setting with a wall directly adjacent to one side of the grid. This setup provided insight into how typical household elements and nearby obstructions can impact wireless signal characteristics. (see Figure 3.2)

Figure 3.8: Testbed setup in the empty 5th floor environment (Loc 2).

3.3.2 Collection Methodology

An occupant was instructed to remain stationary at the center of each voxel, during which CSI data was collected over a duration of 30 seconds. The number of data points obtained per voxel varied, which may be attributed to random fluctuations, the specific positioning of the occupant within the AoI, or interference from other radio wave sources operating at the same frequency.

3.4 Data Preparation

3.4.1 Data Preprocessing

The preprocessing of CSI data was carried out in accordance with the ESP32 documentation Espressif Systems (2024). As described in the *Data Collection* subsection, CSI data was collected using an ESP32 device configured for Wi-Fi communication. Each CSI measurement consisted of 256 values, representing 128 complex numbers, with each subcarrier encoded as a pair of imaginary and real components.

According to the ESP32's implementation of the IEEE 802.11n standard, the CSI vector includes subcarrier indices ranging from -64 to $+63$, with the center

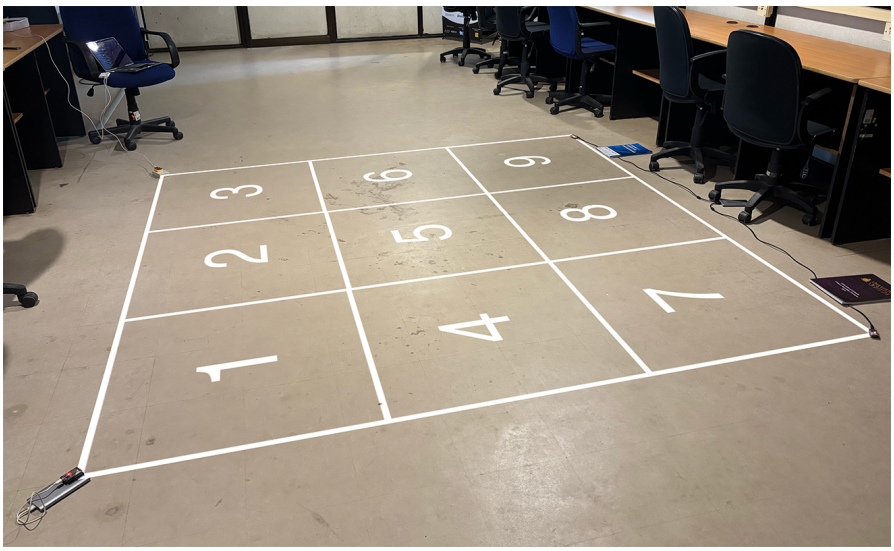


Figure 3.9: Testbed setup in the 4th-year laboratory environment.

frequency at index 0. However, due to the specific Wi-Fi configuration used during data collection—namely, the secondary channel set to `below`, HT mode enabled, and no STBC—only subcarriers corresponding to indices 0 to 63 in the CSI vector were considered for analysis.

Within this range, only the subcarriers from -28 to $+28$ (inclusive) carry usable data, as subcarriers outside this range are designated as null subcarriers and do not contain meaningful channel information. These null subcarriers were excluded from further processing.

Additionally, four pilot subcarriers—located at indices -21, -7, $+7$, and $+21$—are embedded within the usable range. These are reserved for synchronization and channel estimation and were therefore also removed from the analysis. After excluding the null and pilot subcarriers, a total of 52 usable subcarriers remained where were used as the basis for all subsequent tasks.

3.4.2 Feature Extraction

The feature extraction process was designed to transform the preprocessed CSI data into a format suitable for machine learning tasks. For each voxel, time index, sender, and receiver combination, the amplitude and phase of the CSI data were

computed. The amplitude was derived as the absolute value of the complex CSI data, while the phase was obtained as the angle of the complex values.

To reduce noise and smoothen the data, a moving average filter was applied to both the amplitude and phase components. These filtered components were then concatenated to form a feature vector.

This process was repeated for all combinations of voxel, time index, sender, and receiver, resulting in a comprehensive feature matrix (X) and corresponding labels (y), where each label represents the voxel index. The extracted features served as the input for subsequent classification tasks.

3.5 Data Analysis and Model Development

In this work, the objective is to estimate the location of an occupant within an AoI using four ESP32 devices. Two distinct problem formulations were explored: classification and regression. In the classification approach, each voxel is treated as a categorical class. Conversely, in the regression approach, each voxel is represented by a continuous-valued coordinate pair in the range $(-1, 1) \times (-1, 1)$, corresponding to a normalized spatial position. These approaches were applied to analyze the CSI data collected from all three test-bed setups.

3.5.1 Classification-Based Localization

The classification-based formulation aims to determine the voxel in which the occupant is located, treating the problem as a multi-class classification task over a 3×3 spatial grid. Each voxel is assigned a unique integer label from 0 to 8.

Initially, a threshold-based method was considered, similar to the one discussed in Section 3.1. However, while the earlier setup employed only a single transmitter-receiver pair, the updated configuration utilized four ESP32 devices, as illustrated in Figure 3.6. This expansion resulted in 12 directional communication links ($P(4, 2) = 4 \times 3 = 12$), significantly increasing the spatial diversity captured by the system.

Given that CSI is direction-sensitive—meaning the values depend on the specific transmitting and receiving device—each ordered transmitter-receiver pair must be treated as a distinct source of spatial information. Consequently, simple thresholding techniques proved inadequate for generating robust feature representations across all links.

To address this, a data-driven approach was adopted. For each ordered sender-receiver pair (excluding self-links), a moving average filter was first applied across the temporal dimension to smooth CSI measurements and mitigate noise. Feature vectors were then constructed by extracting the magnitude and phase of the 128 subcarriers for each link and appending one-hot encoded vectors to indicate the identities of the sender and receiver devices.

Formally, the feature vector for each time index t and voxel v is defined as:

$$\mathbf{x}_{v,t}^{(s,r)} = \Big[\text{one-hot}(s), \text{one-hot}(r), \text{magnitude}(\mathbf{csi}_{v,t}^{(s,r)}), \text{phase}(\mathbf{csi}_{v,t}^{(s,r)})\Big]$$

where $s \neq r$ represent the sender and receiver device indices, respectively. This process is repeated for all 12 directional links, and the resulting features are aggregated into the dataset X, with corresponding voxel labels stored in y.

A Random Forest Classifier was chosen due to its ability to model non-linear decision boundaries and to handle high-dimensional input data effectively. Moreover, it provides built-in feature importance scores, enabling interpretability of which CSI components are most informative for voxel classification.

The model was implemented using the `scikit-learn` library with 150 decision trees (`n_estimators=150`) and a fixed random seed (`random_state=42`) to ensure reproducibility:

```
RandomForestClassifier(n_estimators=150, random_state=42)
```

The dataset was partitioned into training and testing subsets using an 80:20 split:

```
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42)
```

Here, X contains the flattened input feature vectors constructed from all valid CSI measurements across time, voxels, and device pairs, and y contains the corresponding voxel indices. The trained classifier was then evaluated on the test set using standard metrics such as accuracy and macro-averaged F1-score.

In addition to the CSI data, similar pre-processing steps were applied to the collected RSSI data. The same model parameters were used for consistency. This approach allows for a fair comparison of localization performance between RSSI and CSI data within the same environment, while using a reduced number of nodes.

3.5.2 Regression-based Approach

To formulate the localization task as a regression problem, each voxel in the 3×3 spatial grid was mapped to a continuous-valued coordinate pair in the range $(-1, 1) \times (-1, 1)$, reflecting its normalized position within the AoI. The mapping for the standard 9-voxel layout is defined as follows:

$$\text{voxel\_coord\_map} = \begin{cases} 1 \to (-1, -1), & 2 \to (-1, 0), \\ 3 \to (-1, 1), & 4 \to (0, -1), \\ 5 \to (0, 0), & 6 \to (0, 1), \\ 7 \to (1, -1), & 8 \to (1, 0), \\ 9 \to (1, 1) \end{cases} \tag{3.1}$$

In one particular setup—the empty floor—additional data was collected at four finer-grained locations within the center voxel (voxel 5). These supplementary samples were appended to the dataset and assigned fractional coordinates to reflect their offset positions. This extended mapping was used only for analyzing whether increased granularity affects prediction performance:

$$10 \to (-0.5, -0.5), \quad 11 \to (-0.5, 0.5), \quad 12 \to (0.5, -0.5), \quad 13 \to (0.5, 0.5) \tag{3.2}$$

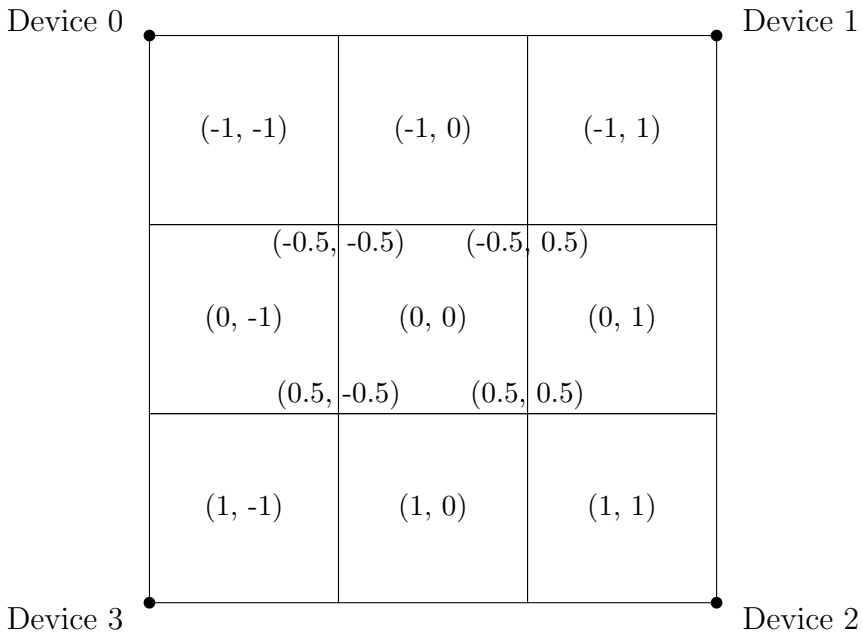


Figure 3.10: Voxel-to-coordinate mapping used for regression. Each square represents a voxel in the 3×3 grid, labeled by its (x, y) coordinate. Additional coordinates at the corners of the center voxel indicate extended positions used in one specific setup.

A multi-output regression model was trained to jointly predict the (x, y) coordinates of the occupant's location based on features extracted from the CSI data. The model used was the `RandomForestRegressor` from the `scikit-learn` library, configured with 50 trees and a fixed random seed to ensure reproducibility:

```
RandomForestRegressor(n_estimators=50, random_state=42)
```

Random Forest was chosen for its effectiveness in modeling non-linear relationships and its interpretability through built-in feature importance scores.

The dataset was randomly split into training and testing subsets using a 60:40 ratio. The same model was used to predict both coordinates simultaneously:

```
X_train, X_test, y_train, y_test, time_indices_train, time_indices_test =
    train_test_split(X, y, time_indices, test_size=0.4, random_state=42)
```

Here, X contains the input features derived from the CSI data, and y consists of the ground truth (x, y) coordinates. The `time_indices` were retained for

visualization purposes, such as temporal color coding in scatter plots.

Model performance was evaluated using MAE and the Coefficient of Determination (R^2 score), providing a detailed view of prediction accuracy across both spatial dimensions.

Chapter 4

Results and Evaluation

4.1 Classification-based Approach

The classification performance across three deployment environments—laboratory, empty floor, and bedroom—shows a similar performance when evaluated using a conventional train-test split from the same data collection session (Figure 4.1, Figure 4.2, Figure 4.3). The empty floor environment achieves the highest accuracy at **96%**, followed by the laboratory at **92%**, and the bedroom at **85%**. This indicates that CSI-based features can encode spatial signatures reliably within a consistent environmental and temporal context.

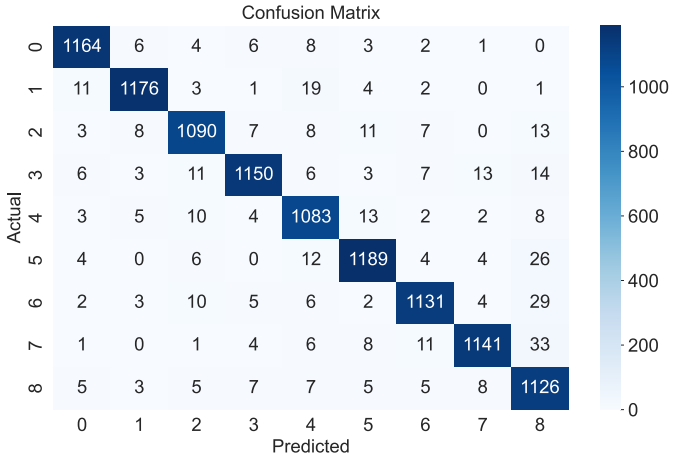


Figure 4.1: Confusion Matrix — Empty Floor (using CSI).

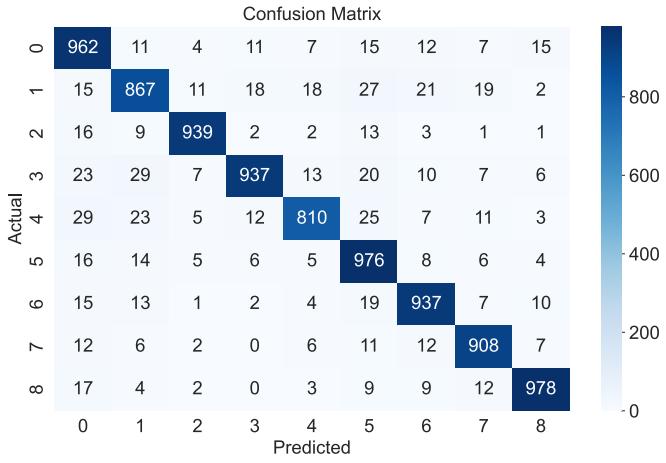


Figure 4.2: Confusion Matrix — Laboratory (using CSI).

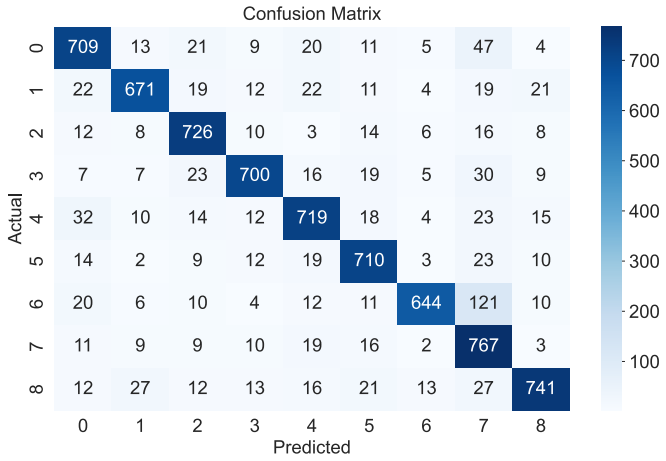


Figure 4.3: Confusion Matrix — Bedroom (using CSI).

Each voxel in the 3×3 grid is classified with high precision and recall across all classes (Figure 4.1, Figure 4.2, Figure 4.3). The classifier benefits from the combined feature vector that includes amplitude and phase information, along with one-hot encoded sender and receiver identities. Feature importance analysis (Figure 4.4) confirms that amplitude features—derived from the first 52 subcarriers—contribute more significantly to the model's performance than phase features. This aligns with prior observations in the literature, such as those by Rao et al. (2020), which highlight the instability of raw phase data due to clock skew, lack of time synchronization, and fixed random offsets between transmitters and receivers.

While calibration techniques exist to mitigate these effects, they are often sensitive to noise and environmental changes. In contrast, amplitude data is inherently more stable and thus plays a more dominant role in classification.

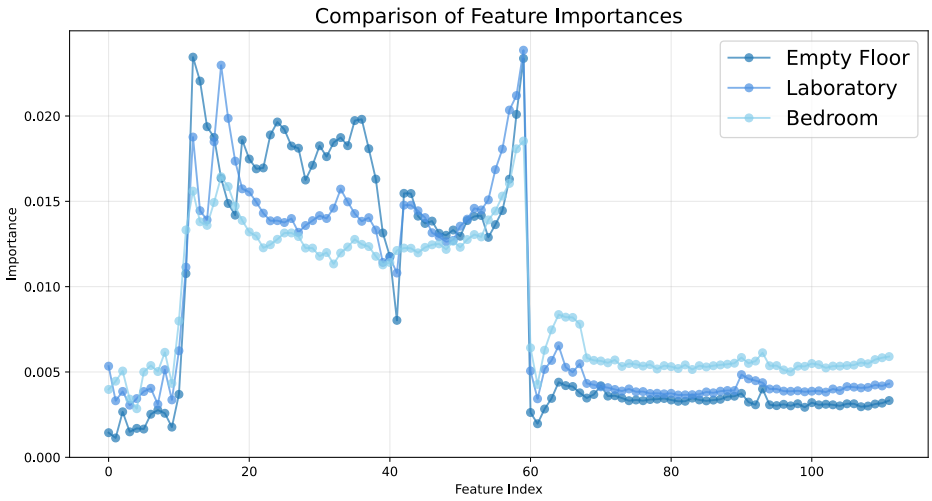


Figure 4.4: Overall feature importance within the same session across all environments.

However, this promising performance does not generalize well when training and testing are performed on data collected during different sessions. In this cross-session evaluation setting, overall accuracy drops significantly—to **46%** in one case and **33%** in another—even though the environment and device placements remain unchanged. Classification performance becomes uneven across voxels, and the model exhibits increased confusion, often defaulting to predictions for a few dominant classes.

Cross-session performance degradation can be attributed to several factors. CSI is sensitive to small environmental changes such as furniture movement, temperature, humidity, and device orientation, which can cause temporal drift in signal patterns even when the spatial location is unchanged. Models trained and tested within the same session may overfit to session-specific noise or artifacts, limiting generalization. Additionally, variations in multipath propagation due to environmental changes can alter channel conditions and degrade classification performance. Differences in calibration between sessions make things even

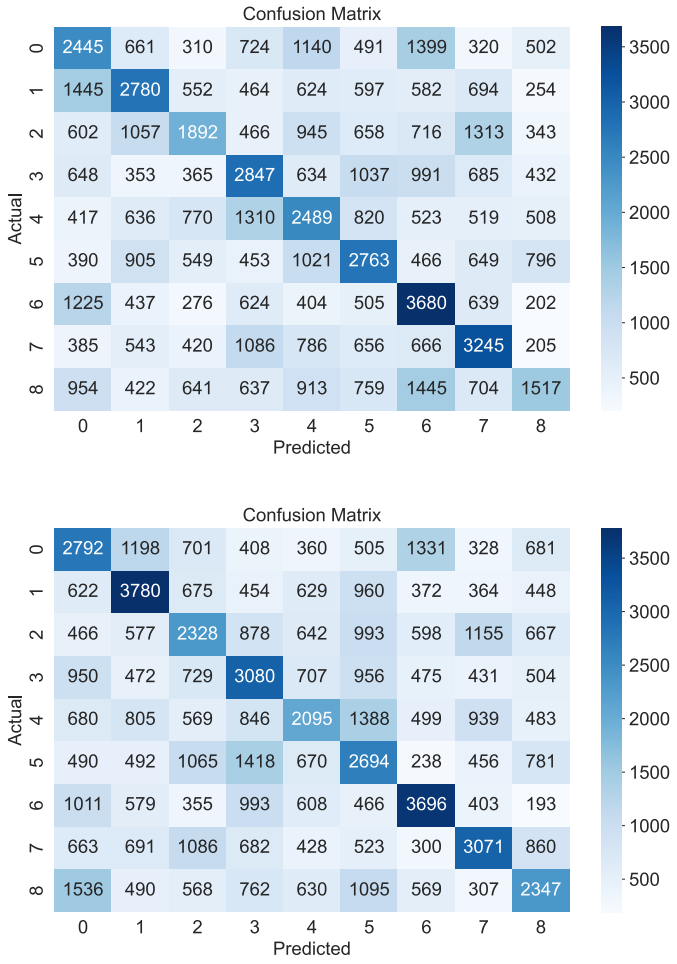


Figure 4.5: Confusion Matrix — Cross-session in same part of empty floor (using CSI).

harder—especially for phase information—and even amplitude values can change due to differences in signal strength, interference, or how the hardware adjusts the signal. The assumption that the wireless channel stays mostly stable—used when applying moving average filters—might not be true across different sessions, which can make the extracted features less reliable. These findings suggest a need for more robust modeling techniques, such as phase calibration, domain adaptation, or session-agnostic representations, to improve generalization across time and changing channel conditions.

The localization performance using RSSI data was also evaluated in the same three environments.

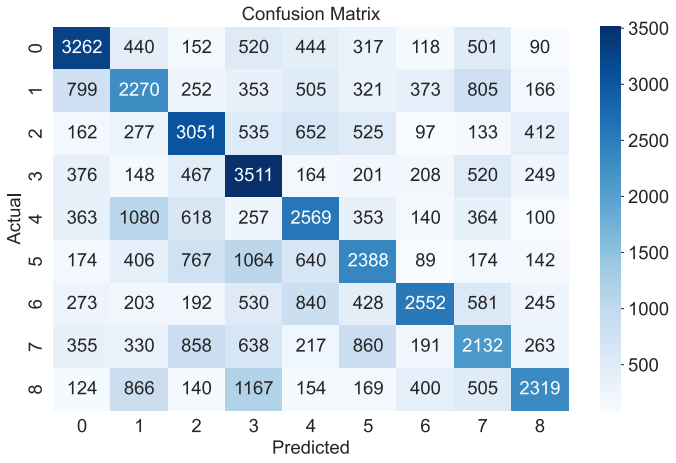


Figure 4.6: Confusion Matrix — Cross-session in different part of empty floor (using CSI).

In the **empty room** environment, the overall accuracy was 0.47. The macro-averaged F1-score was 0.47, indicating moderate performance. Some locations, such as class 2 and class 6, showed relatively higher precision, while others had high recall but low precision, suggesting class imbalance or overlapping signal characteristics.

In the **laboratory**, the accuracy was slightly higher at 0.48. While class 8 achieved high recall and F1-score, several other classes, such as class 0 and class 1, exhibited lower precision and recall, likely due to increased multipath effects and signal interference in a cluttered environment.

The **bedroom** yielded the highest accuracy among the three environments, at 0.54. Notably, class 0 achieved a high recall of 0.93 and an F1-score of 0.67, indicating that this location was more easily distinguishable. Similarly, classes 6 and 7 demonstrated strong performance.

To assess the generalization ability of the RSSI-based localization model, an additional experiment was conducted in the empty room environment, where training and testing data were collected in separate sessions. The model's performance in this cross-session setting significantly decreased, achieving an overall accuracy of only 0.21. The macro-averaged F1-score was also low at 0.21, indicating a

considerable drop in classification effectiveness compared to within-session evaluations.

The per-class results show large variability. For instance, class 2 exhibited relatively high precision (0.56), while classes such as 0 and 6 had very low scores across all metrics. This degradation suggests that the model struggles to maintain performance across temporal changes, likely due to variations in environmental conditions, device orientation, and interference between sessions.

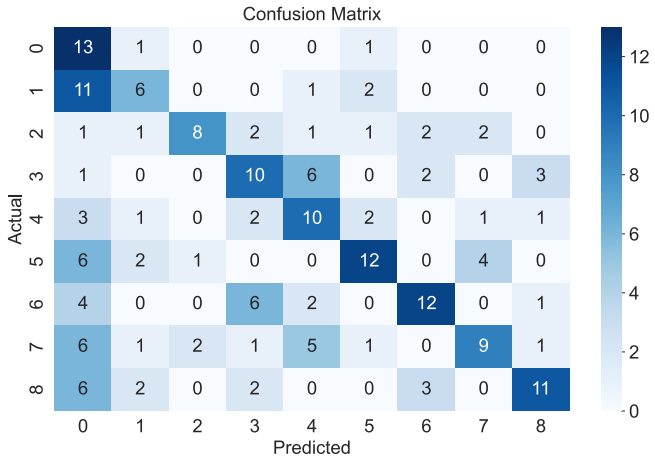


Figure 4.7: Confusion Matrix — Empty Floor (using RSSI).

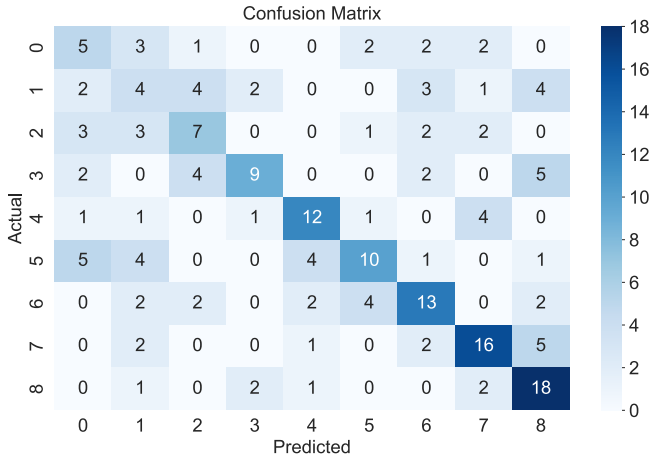


Figure 4.8: Confusion Matrix — Laboratory (using RSSI).

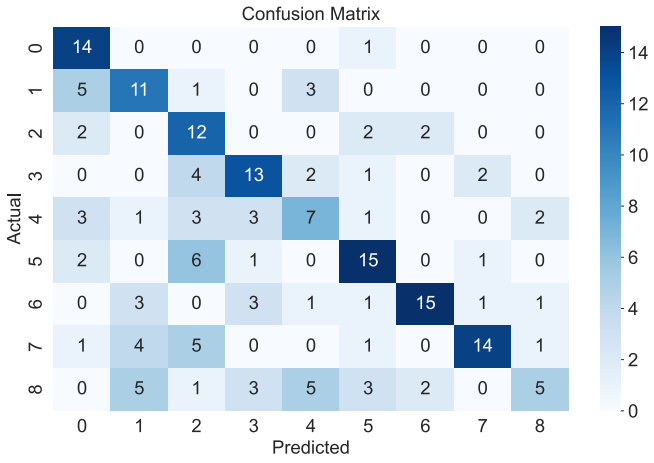


Figure 4.9: Confusion Matrix — Bedroom (using RSSI).

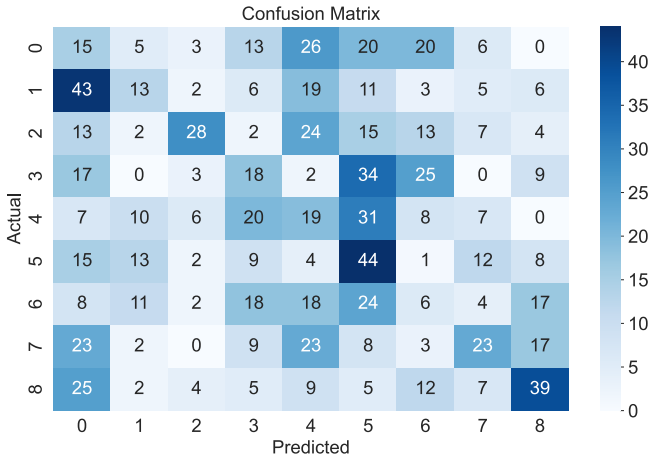


Figure 4.10: Confusion Matrix — Empty Floor Cross-Session (using RSSI).

4.2 Regression-based Approach

4.2.1 Using 9 coordinates

The performance of the regression model was assessed under both random train-test splits and cross-session evaluation settings. Two primary metrics were used: the MAE for each coordinate and the Coefficient of Determination (R^2 score), which provides insight into how well the model explains the variance in the target positions.

Table 4.1 presents the MAE and R^2 scores for the x and y coordinates across

different environments under a 60:40 train-test split. The empty room environment was evaluated using three independent data splits across two physical locations to observe performance consistency. The 4th year lab and bedroom setups were evaluated with a single split each.

Table 4.1: MAE and R^2 Score for train-test split evaluation.

Environment	**MAE (X)**	**MAE (Y)**	R^2 **(X)**	R^2 **(Y)**
Empty Room (Loc 1)	0.1062	0.1114	0.9393	0.9316
4th Year Lab	0.1474	0.1445	0.8997	0.9033
Bedroom	0.1927	0.1909	0.8378	0.8413

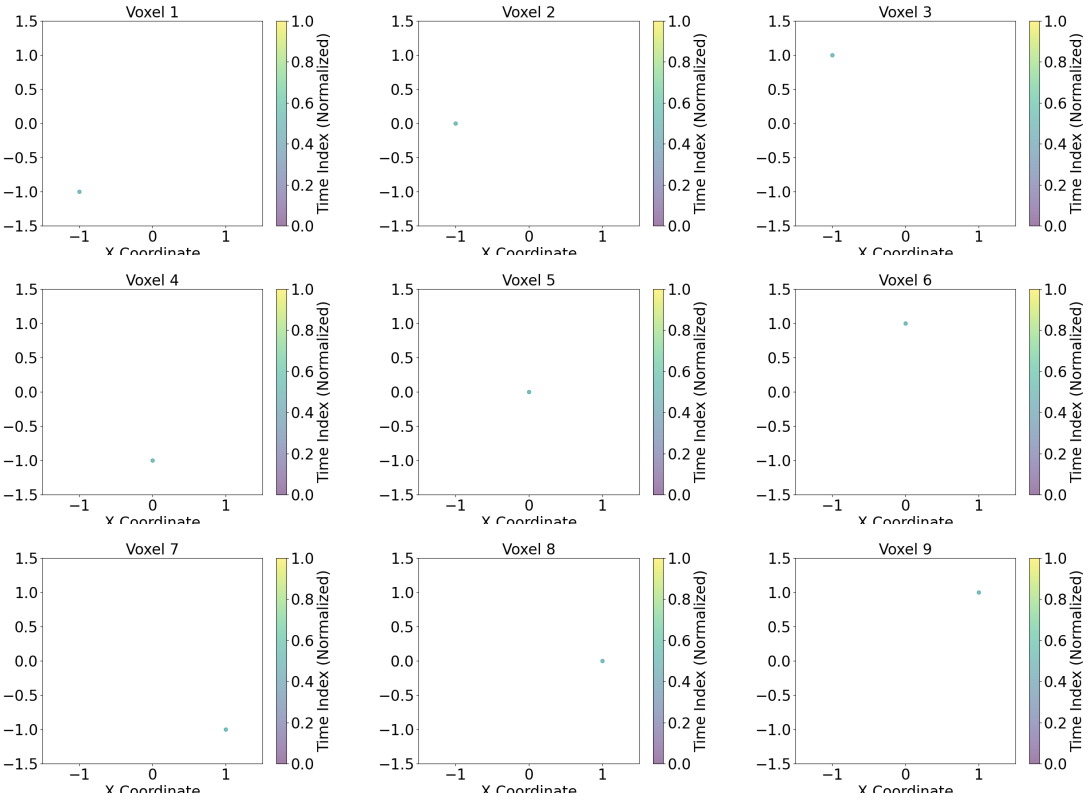


Figure 4.11: Scatterplots for voxels 1-9 with just the actual/expected location.

To evaluate the generalization of the model across different data collection sessions, cross-session tests were conducted for the empty room environment. The results are summarized in Table 4.2.

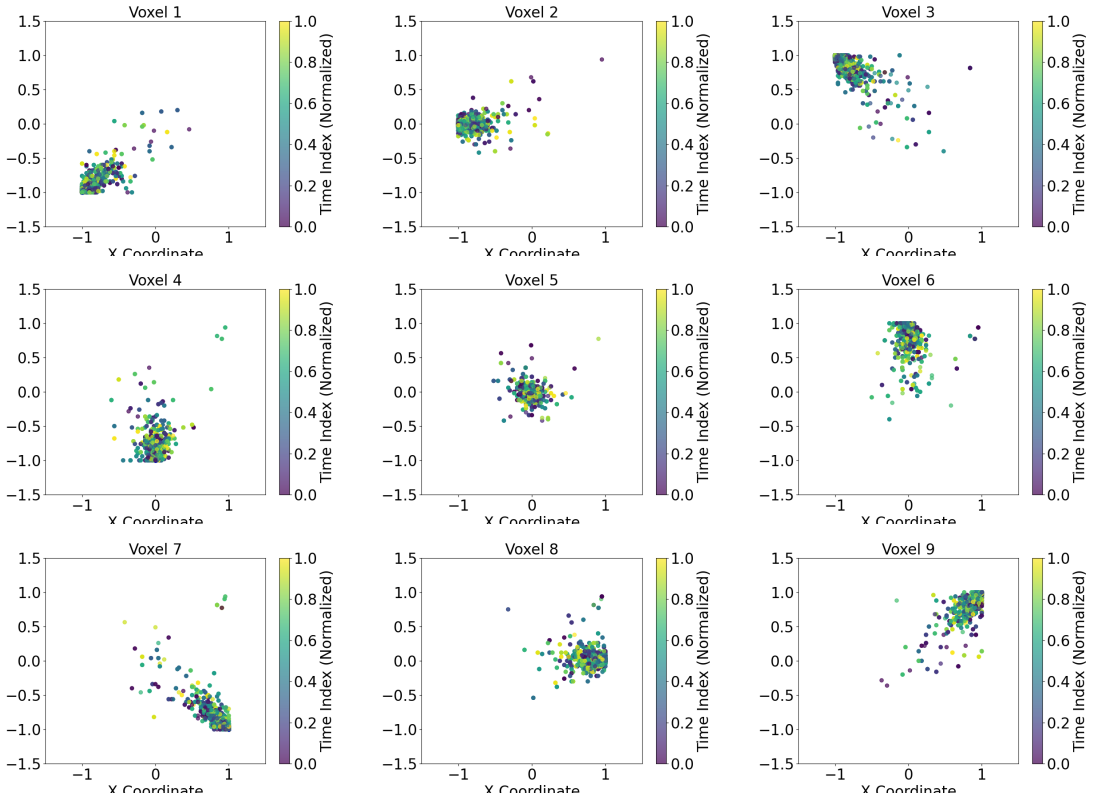


Figure 4.12: Scatterplots with time heatmap for voxels 1-9 in the empty floor environment (train-test split).

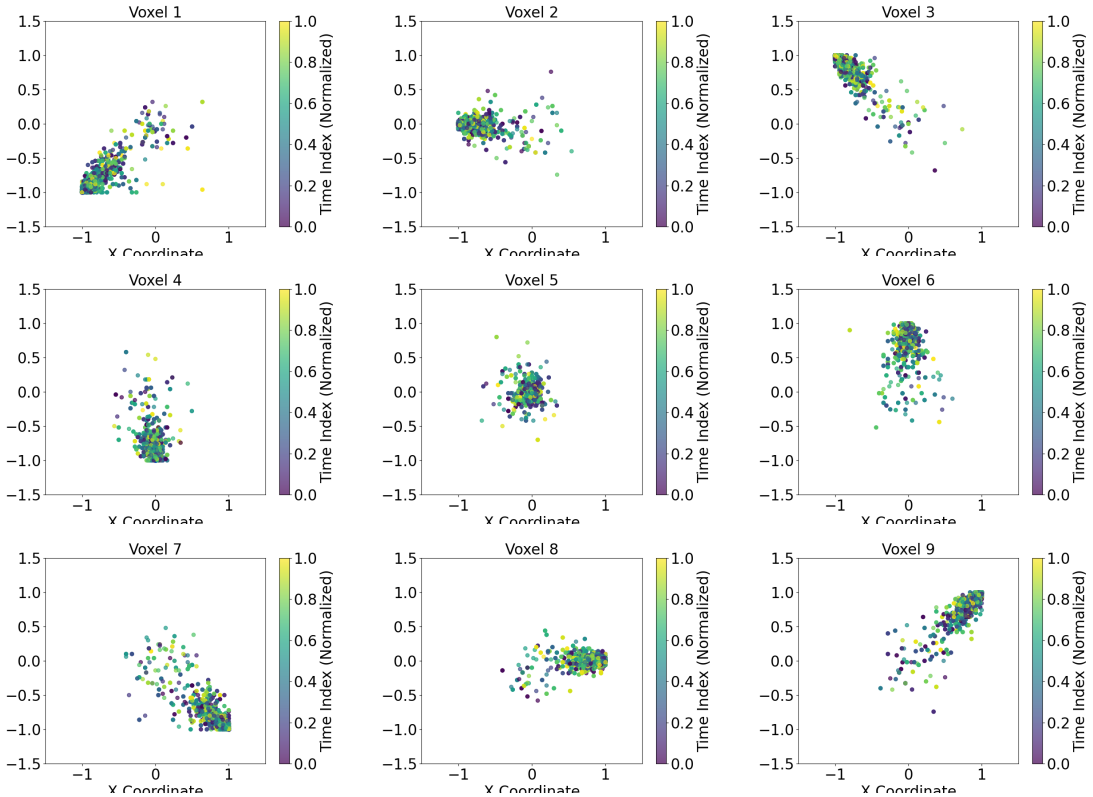


Figure 4.13: Scatterplots with time heatmap for voxels 1-9 in the laboratory environment (train-test split).

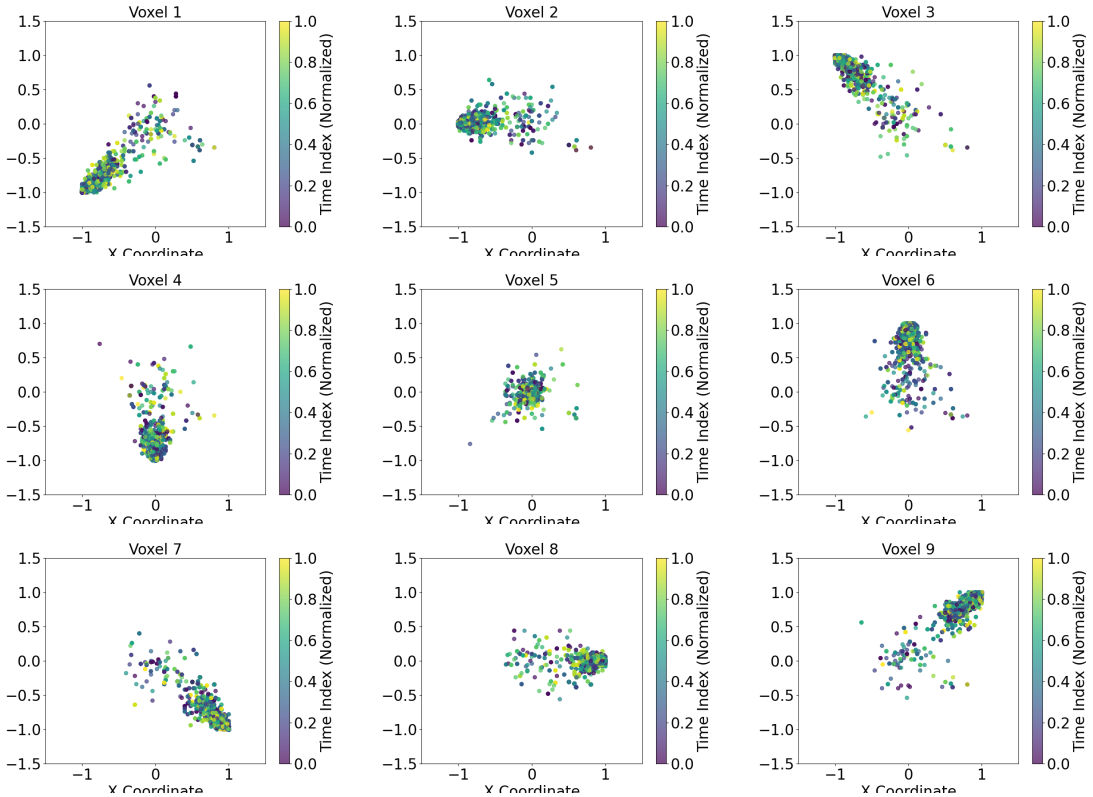


Figure 4.14: Scatterplots with time heatmap for voxels 1-9 in the bedroom environment (train-test split).

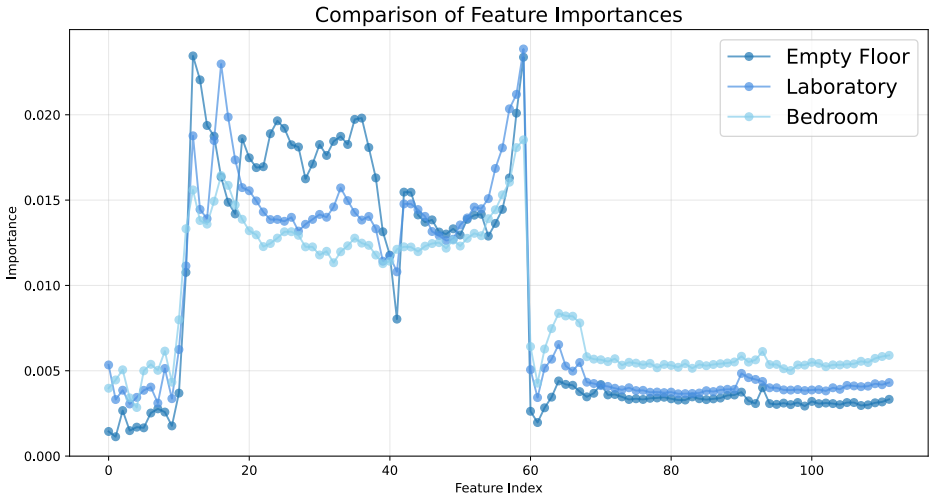


Figure 4.15: Overall feature importance within the same session across all environments.

Table 4.2: MAE and R^2 Score for cross-session evaluation.

Session	**MAE (X)**	**MAE (Y)**	R^2 **(X)**	R^2 **(Y)**
Session 1 (Loc 1)	0.5725	0.5877	0.2495	0.2101
Session 2 (Loc 2)	0.6443	0.6459	0.0867	0.0998
Session 3 (Loc 2)	0.6586	0.6376	0.0437	0.1091

The model exhibits strong predictive performance under controlled train-test splits, particularly in the empty room environment, where both MAE and R^2 scores are optimal. Performance is stable across locations within the empty room and remains reasonably high in the lab setting. The bedroom, likely due to a more complex physical layout and potential for multipath interference, shows a slight drop in accuracy. The scatter plots in Figure 4.12, Figure 4.13, and Figure 4.14 illustrate the predicted and actual coordinate positions for each voxel under the train-test split configuration The predicted coordinates closely align with the actual positions, forming compact clusters with minimal dispersion.

In contrast, cross-session evaluations reveal a significant reduction in predictive accuracy, with increased MAE values and low R^2 scores. These findings under-

score the model's sensitivity to environmental and temporal variability, highlighting challenges for deployment in dynamic or previously unseen conditions. The scatter plots, as shown in Figure 4.16, Figure 4.17, and Figure 4.18 shows predicted positions that are more widely scattered across the coordinate space. This increased spread indicates a reduced spatial agreement between predicted and actual values in the cross-session setting.

The feature importance profiles observed across all three environments, as illustrated in Figure 4.15, exhibit patterns consistent with those identified in the classification task shown in Figure 4.4. This similarity may be attributed to the shared physical relevance of dominant features such as amplitude and phase-related statistics, which are sensitive to spatial positioning and multipath effects, thereby influencing both classification and regression tasks in a comparable manner.

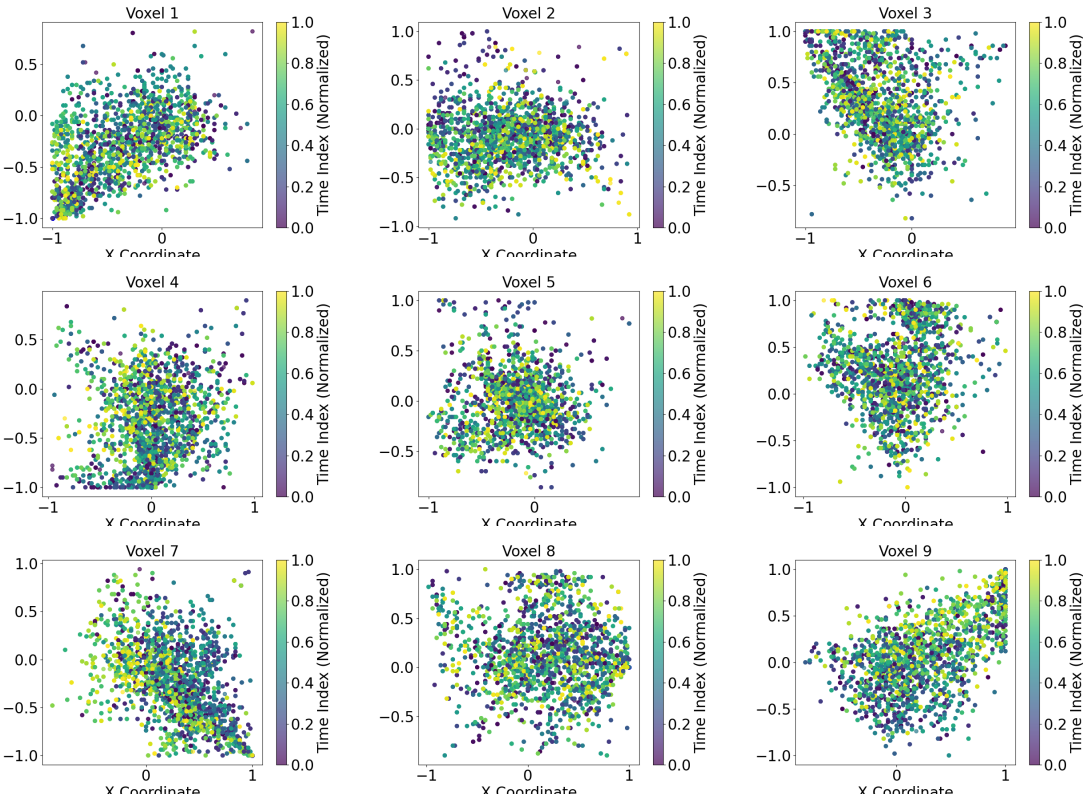


Figure 4.16: Scatterplots with time heatmap for voxels 1-9 in the empty floor environment (Loc 1) (cross-session).

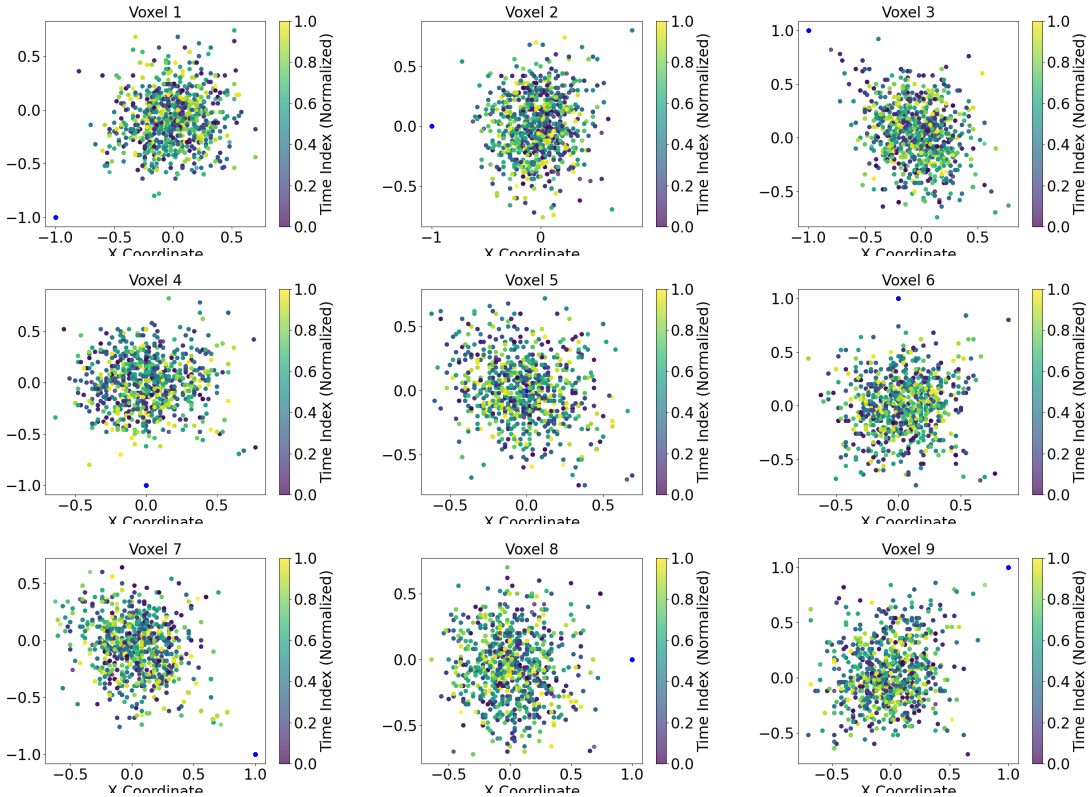


Figure 4.17: Scatterplots with time heatmap for voxels 1-9 in the empty floor environment (Loc 2) (cross-session).

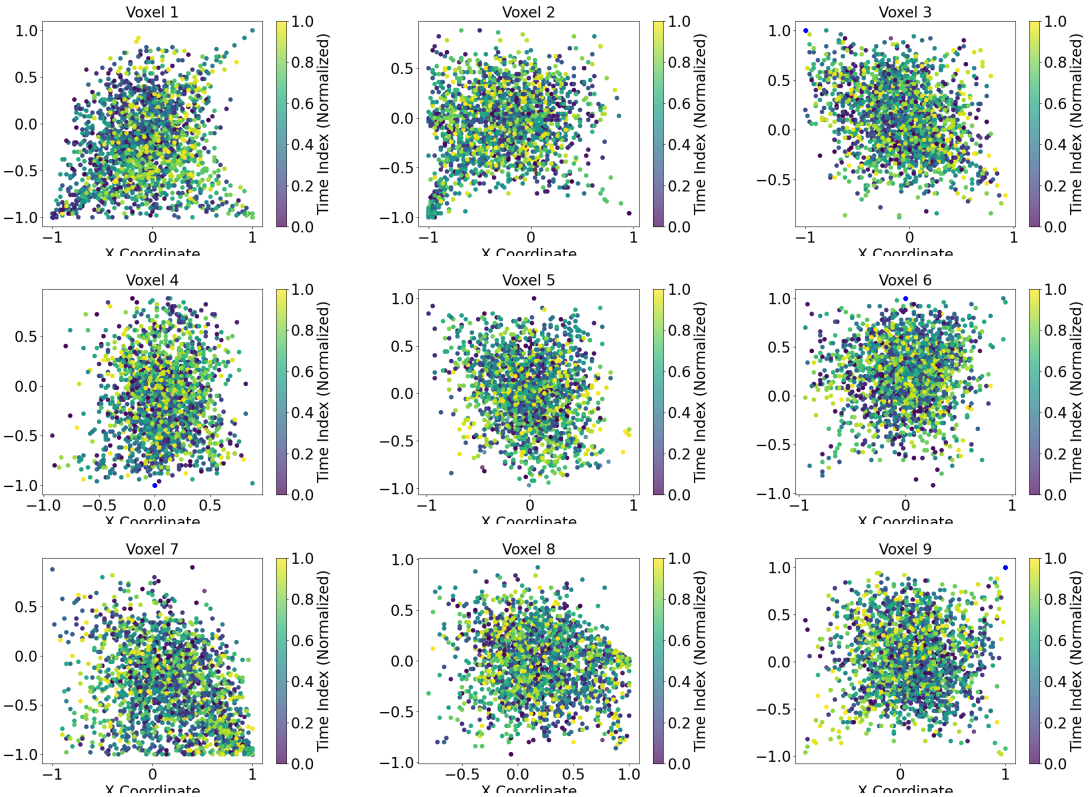


Figure 4.18: Scatterplots with time heatmap for voxels 1-9 in the empty floor environment (Loc 2) taken at a different time (cross-session).

4.2.2 Using 12 coordinates

Similarly, the configuration utilizing 12 coordinates was assessed using both random train-test splits and cross-session evaluation settings. The results, presented in Table 4.3, indicate that no significant improvement was observed. Although the overall amount of collected data has increased, the amount of data per voxel remains unchanged. Additionally, the reduction in relative distances between voxels may make it more difficult for the model to distinguish between the corners of voxel 5 and its surrounding coordinates.

Table 4.3: MAE and R^2 Score for 12 coordinates.

Session	**MAE (X)**	**MAE (Y)**	R^2 **(X)**	R^2 **(Y)**
Same session	0.1290	0.1277	0.9112	0.9127
Cross-session	0.5958	0.5687	0.0418	0.1224

Previous research involving CSI data used Wi-Fi packets with unchanging payloads to perform RTI. In the custom protocol used in this research, the payload contains the latest measurements of CSI data from other devices. It is not known and no research has been done to study the effect of changing payloads to the localization accuracy using the CSI data collected.

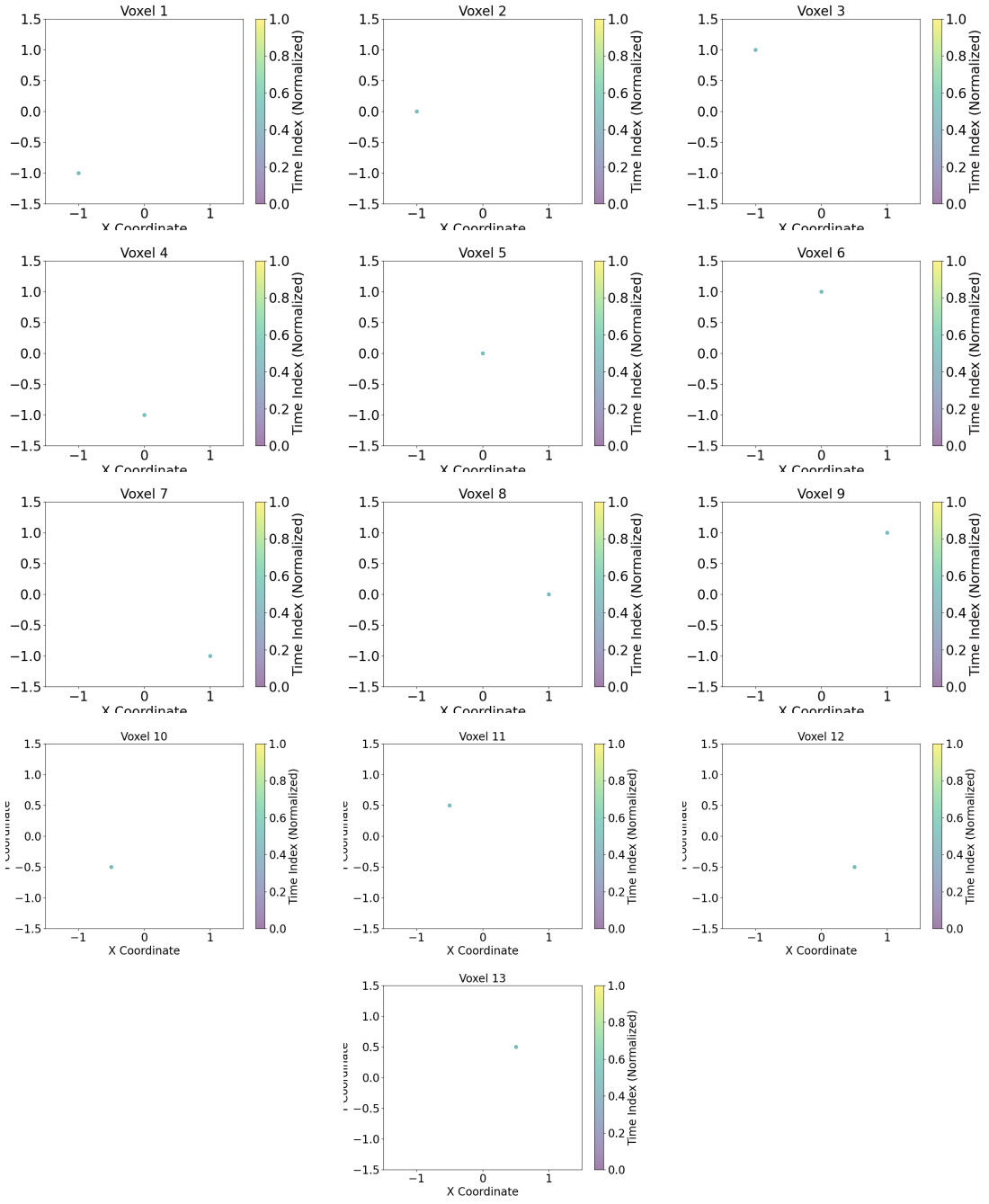


Figure 4.19: Scatterplots for the 12 coordinates with just the actual/expected location.

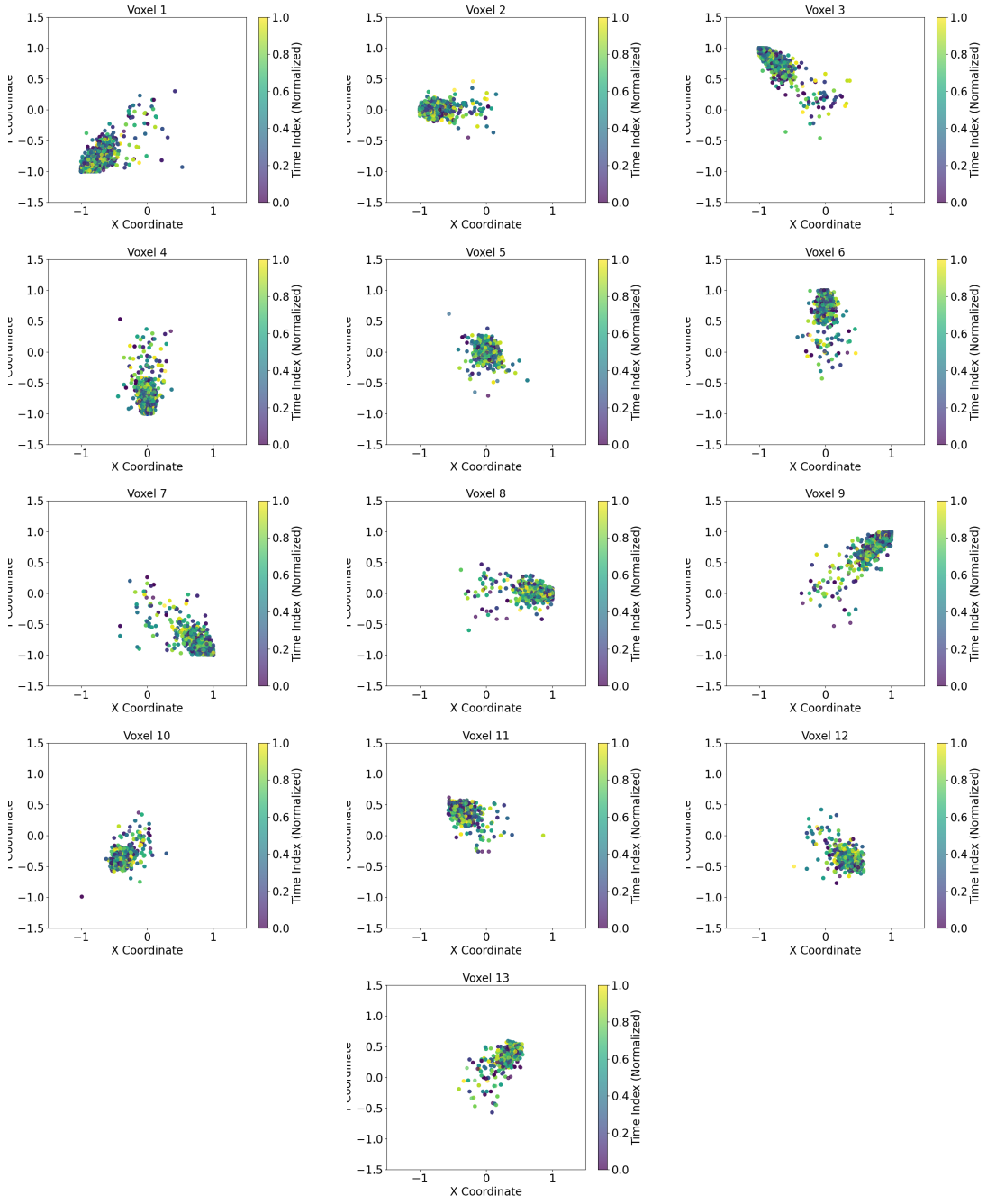


Figure 4.20: Scatterplots with time heatmap for the 12 coordinates in the empty floor environment (Loc 2) (train-test split).

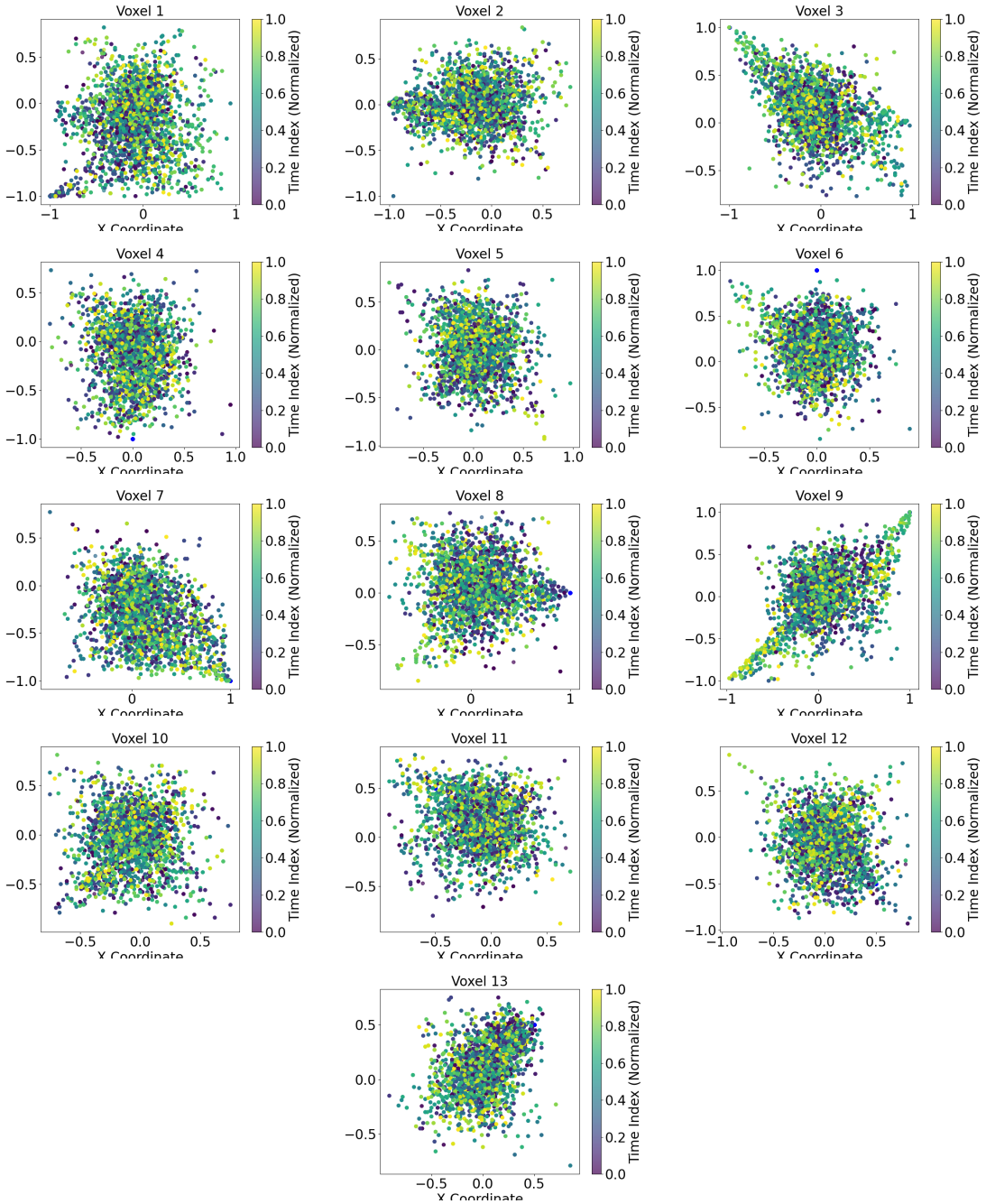


Figure 4.21: Scatterplots with time heatmap for the 12 coordinates in the empty floor environment (Loc 2) (cross-session).

Chapter 5

Conclusion

5.1 Conclusion on the Research Questions

5.1.1 How can we reduce the number of nodes required for effective RTI be reduced?

The results show that the number of nodes required for RTI can be significantly reduced by using CSI with the custom round-robin protocol. A minimal setup of four ESP32 nodes arranged in a square layout gives us $P(4, 2) = 4 \times 3 = 12$ links which would normally require 6 or 12 pairs of devices depending on whether the data collection is bi-directional or uni-directional. Data collection is centralized making it easier to perform real-time processing without requiring complex data aggregation.

5.1.2 Does CSI outperform RSSI for RTI tasks in sparse network deployments?

The results confirm that CSI consistently outperforms RSSI across all environments tested, especially when using a small number of nodes.

In both classification and regression tasks:

1. Models using CSI features outperformed those using RSSI in terms of accuracy and robustness.

2. The advantage of CSI was more significant in complex environments where multipath effects are prominent.

3. Feature importance analysis showed that CSI, particularly amplitude data, provided more stable and informative signals for localization.

5.2 Contributions

This work makes the following contributions to the field of device-free localization:

1. A data collection protocol using ESP-NOW was designed to enable round-robin transmission of CSI data without requiring full network synchronization.

2. A full data processing pipeline was developed, including CSI preprocessing, feature extraction, and both classification and regression modeling, all of which are lightweight and suitable for real-time inference.

3. A direct empirical comparison between CSI and RSSI under identical experimental conditions was presented, offering clear evidence of the benefits of CSI in sparse deployments.

4. Feature importance analysis provided insights into which signal characteristics contribute most to localization performance, aiding in model interpretation and future improvements.

5.3 Limitations

Despite promising results, this study faces several limitations:

1. Cross-session performance is significantly lower than same-session performance, suggesting that CSI features are sensitive to environmental changes and hardware drift.

2. The ESP-NOW protocol has strict payload size limits, which restricts the number of CSI samples that can be transmitted in each packet, which in turn limits the total number of devices that can be used with this protocol.

3. The firmware used limits the maximum CSI array length, particularly in high-bandwidth Wi-Fi modes, potentially excluding useful data.

4. The system lacks time synchronization between nodes, which may introduce inconsistencies in data alignment and reduce localization accuracy.

5. While the chosen models are efficient and interpretable, they may not capture more complex temporal or spatial patterns needed for better generalization.

5.4 Future Work

To address these limitations and build upon current findings, the following directions are proposed for future research:

1. Explore phase calibration methods or session-invariant representations to improve robustness to environmental changes across different sessions.

2. Investigate domain adaptation and transfer learning techniques to enhance generalization across time and space.

3. Improve the round-robin protocol to support larger payloads or incorporate lightweight time synchronization.

4. Experiment with more complex models, such as convolutional or recurrent neural networks, to better capture spatial and temporal relationships in CSI data.

5. Expand the testbed to larger and more complex environments to further evaluate the scalability and effectiveness of the system.

6. Investigate the effect of changing payloads in the Wi-Fi packet to localization.

Bibliography

Al-qaness, M. A. A., Elaziz, M. A., Kim, S., Ewees, A. A., Abbasi, A. A., Alhaj, Y. A. & Hawbani, A. (2019), 'Channel state information from pure communication to sense and track human motion: A survey', *Sensors* **19**, 3329.
URL: *https://www.mdpi.com/1424-8220/19/15/3329*

Bahl, P. & Padmanabhan, V. (2000), 'Radar: an in-building rf-based user location and tracking system', *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)* **2**, 775–784.
URL: *http://ieeexplore.ieee.org/document/832252/*

Brinke, J. K., Chiumento, A. & Havinga, P. (2023), Channel state information for human activity recognition with low sampling rates, *in* '2023 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)', IEEE, pp. 646–652.

Chapre, Y., Ignjatovic, A., Seneviratne, A. & Jha, S. (2015), 'Csi-mimo: An efficient wi-fi fingerprinting using channel state information with mimo', *Pervasive and Mobile Computing* **23**, 89–103.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S1574119215001406*

Chen, H., Zhang, Y., Li, W., Tao, X. & Zhang, P. (2017), 'Confi: Convolutional neural networks based indoor wi-fi localization using channel state information', *IEEE Access* **5**, 18066–18074.

Choi, H., Fujimoto, M., Matsui, T., Misaki, S. & Yasumoto, K. (2022), 'Wi-

CaL: WiFi sensing and machine learning based device-free crowd counting and localization', *IEEE Access* **10**, 24395–24410.

Dang, X., Tang, X., Hao, Z. & Liu, Y. (2019), 'A device-free indoor localization method using csi with wi-fi signals.', *Sensors (Basel, Switzerland)* **19**.
URL: *http://www.ncbi.nlm.nih.gov/pubmed/31340502 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6679537*

Denis, S., Kaya, A., Berkvens, R. & Weyn, M. (2020), 'Device-free localization and identification using sub-ghz passive radio mapping', *Applied Sciences (Switzerland)* **10**.

Espressif Systems (2024), 'Wi-fi driver – esp32 — esp-idf programming guide v5.4.1 documentation', `https://docs.espressif.com/projects/esp-idf/en/stable/esp32/api-guides/wifi.html#wi-fi-channel-state-information`. [Accessed 20-Apr-2025].

Farahani, S. (2008), *Location Estimation Methods*, Elsevier, pp. 225–246.
URL: *https://linkinghub.elsevier.com/retrieve/pii/B9780750683937000078*

Gholamhosseini, L., Sadoughi, F. & Safaei, A. (2019), 'Hospital real-time location system (a practical approach in healthcare): A narrative review article.', *Iranian journal of public health* **48**, 593–602.
URL: *http://www.ncbi.nlm.nih.gov/pubmed/31110969*

GPS.gov: GPS Accuracy (2022).
URL: *https://www.gps.gov/systems/gps/performance/accuracy/*

Halperin, D., Hu, W., Sheth, A. & Wetherall, D. (2011), 'Tool release', *ACM SIGCOMM Computer Communication Review* **41**, 53–53.
URL: *https://dl.acm.org/doi/10.1145/1925861.1925870*

Hernandez, S. M. & Bulut, E. (2020), Lightweight and Standalone IoT Based WiFi Sensing for Active Repositioning and Mobility, *in* '21st International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM) (WoWMoM 2020)', Cork, Ireland.

Hoang, M. T., Yuen, B., Dong, X., Lu, T., Westendorp, R. & Reddy, K. (2019), 'Recurrent neural networks for accurate rssi indoor localization', *IEEE Internet of Things Journal* **6**, 10639–10651.
URL: *https://ieeexplore.ieee.org/document/8830368/*

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

Ikegami, T. & Kim, M. (2022), 'An indoor device-free localization technique using multipath wireless channel characteristics', *2022 IEEE International Conference on Consumer Electronics-Asia, ICCE-Asia 2022* .

Kak, A., Slaney, M., in Medicine, I. E. & Society, B. (1988), *Principles of Computerized Tomographic Imaging*, IEEE Press.
URL: *https://books.google.lk/books?id=RKhrAAAAMAAJ*

Kaltiokallio, O., Bocca, M. & Patwari, N. (2012*a*), 'Enhancing the accuracy of radio tomographic imaging using channel diversity', *2012 IEEE 9th International Conference on Mobile Ad-Hoc and Sensor Systems (MASS 2012)* pp. 254–262.
URL: *http://ieeexplore.ieee.org/document/6502524/*

Kaltiokallio, O., Bocca, M. & Patwari, N. (2012*b*), 'Follow @grandma: Long-term device-free localization for residential monitoring', *37th Annual IEEE Conference on Local Computer Networks – Workshops* pp. 991–998.
URL: *http://ieeexplore.ieee.org/document/6424092/*

Kim, M., Miyake, Y., Ikegami, T., Tsukada, H. & Kumakura, K. (2022), 'Device-free localization using millimeter-wave double-directional channel sounding measurements', *2022 16th European Conference on Antennas and Propagation (EuCAP)* pp. 1–4.
URL: *https://ieeexplore.ieee.org/document/9769204/*

Kim, M., Tasaki, T. & Yamakawa, S. (2019), Millimeter-wave radio tomographic

imaging technique using multipath components for indoor localization, *in* '2019 International Symposium on Antennas and Propagation (ISAP)', pp. 1–3.

Kumar, P., Reddy, L. & Varma, S. (2009), 'Distance measurement and error estimation scheme for rssi based localization in wireless sensor networks', *2009 Fifth International Conference on Wireless Communication and Sensor Networks (WCSN)* pp. 1–4.
URL: *http://ieeexplore.ieee.org/document/5434802/*

Liu, Z., Dai, B., Wan, X. & Li, X. (2019), 'Hybrid wireless fingerprint indoor localization method based on a convolutional neural network', *Sensors (Switzerland)* **19**.

Lu, X., Li, Y., Cui, W. & Wang, H. (2023), 'CeHAR: CSI-based channel-exchanging human activity recognition', *IEEE Internet Things J.* **10**(7), 5953–5961.

Lu, X., Zou, H., Zhou, H., Xie, L. & Huang, G. B. (2016), 'Robust extreme learning machine with its application to indoor positioning', *IEEE Transactions on Cybernetics* **46**, 194–205.

Mager, B., Lundrigan, P. & Patwari, N. (2015), 'Fingerprint-based device-free localization performance in changing environments', *IEEE Journal on Selected Areas in Communications* **33**, 2429–2438.
URL: *http://ieeexplore.ieee.org/document/7102680/*

Mahfouz, S., Mourad-Chehade, F., Honeine, P., Farah, J. & Snoussi, H. (2014), 'Target tracking using machine learning and kalman filter in wireless sensor networks', *IEEE Sensors Journal* **14**, 3715–3725.
URL: *https://hal.science/hal-01965571*

Patwari, N. & Wilson, J. (2010), 'Rf sensor networks for device-free localization: Measurements, models, and algorithms', *Proceedings of the IEEE* **98**, 1961–1973.
URL: *http://ieeexplore.ieee.org/document/5523907/*

Poulose, A. & Han, D. S. (2021), 'Hybrid deep learning model based indoor positioning using wi-fi rssi heat maps for autonomous applications', *Electronics (Switzerland)* **10**, 1–15.

Rao, X., Li, Z., Yang, Y. & Wang, S. (2020), 'DFPhaseFL: a robust device-free passive fingerprinting wireless localization system using CSI phase information', *Neural Comput. Appl.* **32**(18), 14909–14927.

Rodrigo, M. M. N. D. & Sayakkara, A. P. (2024), Indoor localization with radio tomographic imaging using channel state information. Manuscript, University of Colombo School of Computing. Not peer-reviewed.

Sanam, T. F. & Godrich, H. (2018), 'An improved csi based device free indoor localization using machine learning based classification approach', *2018 26th European Signal Processing Conference (EUSIPCO)* pp. 2390–2394.
URL: *https://ieeexplore.ieee.org/document/8553394/*

Schäfer, A. (2022), 'Self-localization of IoT devices - development and implementation of a system for self-localization of embedded devices based on Wi-Fi information'.

Seifeldin, M. & Youssef, M. (2009), 'Nuzzer: A large-scale device-free passive localization system for wireless environments', *IEEE Transactions on Mobile Computing* **12**, 1321–1334.
URL: *http://arxiv.org/abs/0908.0893*

Su, Y. H., Ren, J., Qian, Z., Fouhey, D. & Sample, A. (2023), 'Tomoid: A scalable approach to device free indoor localization via rfid tomography', *Proceedings - IEEE INFOCOM* **2023-May**.

Subhan, F., Saleem, S., Bari, H., Khan, W. Z., Hakak, S., Ahmad, S. & El-Sherbeeny, A. M. (2020), 'Linear discriminant analysis-based dynamic indoor localization using bluetooth low energy (ble)', *Sustainability (Switzerland)* **12**, 1–12.

Suroso, D. J., Cherntanomwong, P. & Sooraksa, P. (2021), 'Indoor device-free localization using received signal strength indicator and illuminance sensor for random-forest-based fingerprint technique', *Sensors and Materials* **33**(12), 4331–4345.

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

Wang, Y., Liu, J., Chen, Y., Gruteser, M., Yang, J. & Liu, H. (2014), E-eyes, *in* 'Proceedings of the 20th annual international conference on Mobile computing and networking', ACM, New York, NY, USA.

Wilson, J. & Patwari, N. (2010), 'Radio tomographic imaging with wireless networks', *IEEE Transactions on Mobile Computing* **9**, 621–632.

Wilson, J. & Patwari, N. (2011), 'See-through walls: Motion tracking using variance-based radio tomography networks', *IEEE Transactions on Mobile Computing* **10**, 612–621.

Wilson, J. & Patwari, N. (2012), 'A fade-level skew-laplace signal strength model for device-free localization with wireless networks', *IEEE Transactions on Mobile Computing* **11**, 947–958.

Wu, K., Xiao, J., Yi, Y., Gao, M. & Ni, L. M. (2012), 'Fila: Fine-grained indoor localization', *2012 Proceedings IEEE INFOCOM* pp. 2210–2218.
URL: *http://ieeexplore.ieee.org/document/6195606/*

Xiao, J., Wu, K., Yi, Y. & Ni, L. M. (2012), 'Fifs: Fine-grained indoor fingerprinting system', *2012 21st International Conference on Computer Communications and Networks (ICCCN)* pp. 1–7.
URL: *http://ieeexplore.ieee.org/document/6289200/*

Xu, C., Firner, B., Moore, R. S., Zhang, Y., Trappe, W., Howard, R., Zhang, F. & An, N. (2013), 'Scpl', *Proceedings of the 12th international conference on*

Information processing in sensor networks pp. 79–90.
URL: *https://dl.acm.org/doi/10.1145/2461381.2461394*

Xu, C., Firner, B., Zhang, Y. & Howard, R. E. (2016), 'The case for efficient and robust rf-based device-free localization', *IEEE Transactions on Mobile Computing* **15**, 2362–2375.

Yang, Z., Zhou, Z. & Liu, Y. (2013), 'From rssi to csi: Indoor localization via channel response', *ACM Computing Surveys* **46**.

Yin, G., Zhang, J., Shen, G. & Chen, Y. (2024), 'FewSense, towards a scalable and cross-domain WI-fi sensing system using few-shot learning', *IEEE Trans. Mob. Comput.* **23**(1), 453–468.

Youssef, M. & Agrawala, A. (2005), 'The horus wlan location determination system', *Proceedings of the 3rd international conference on Mobile systems, applications, and services* pp. 205–218.
URL: *https://dl.acm.org/doi/10.1145/1067170.1067193*

Zafari, F., Gkelias, A. & Leung, K. (2017), 'A survey of indoor localization systems and technologies', *IEEE Communications Surveys & Tutorials* **21**, 2568–2599.
URL: *http://arxiv.org/abs/1709.01015*

Zhang, D., Ma, J., Chen, Q. & Ni, L. M. (2007), 'An rf-based system for tracking transceiver-free objects', *Fifth Annual IEEE International Conference on Pervasive Computing and Communications (PerCom'07)* pp. 135–144.
URL: *http://ieeexplore.ieee.org/document/4144758/*

Zhang, Z., Ishida, S., Tagashira, S. & Fukuda, A. (2019), 'Danger-pose detection system using commodity WI-Fi for bathroom monitoring', *Sensors (Basel)* **19**(4), 884.

Zhao, Y., Patwari, N., Phillips, J. M. & Venkatasubramanian, S. (2013), 'Radio tomographic imaging and tracking of stationary and moving people via kernel

distance', *Proceedings of the 12th international conference on Information processing in sensor networks* pp. 229–240.
URL: *https://dl.acm.org/doi/10.1145/2461381.2461410*