AUTHENTICATION AND MESSAGE INTEGRITY VERIFICATION FOR EMERGING WIRELESS NETWORKS

by

Ebuka Philip Oguchi

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfilment of Requirements

For the Degree of Doctor of Philosophy

Major: Computer Science

Under the Supervision of Professor Nirnimesh Ghose

Lincoln, Nebraska

August, 2025

AUTHENTICATION AND MESSAGE INTEGRITY VERIFICATION FOR EMERGING WIRELESS NETWORKS

Ebuka Philip Oguchi, Ph.D.

University of Nebraska, 2025

Adviser: Nirnimesh Ghose

This dissertation presents a comprehensive body of research on authentication and message integrity verification for emerging wireless networks, focusing on secret-free and physical layer security techniques across diverse, challenging, and unconventional environments. It comprises four first-author contributions that span underground wireless systems, over-the-air (OTA) channels, vehicular communications, and nanoscale molecular networks.

The first contribution, Soil-Assisted Trust Establishment for Underground Wireless Networks (STUN), introduces a physical-layer trust bootstrapping protocol that achieves authentication and message integrity without pre-shared secrets. Leveraging underground-to-air propagation laws and trusted relay nodes, STUN resists active signal injection attacks and demonstrates security comparable to the unbalanced oil and vinegar cryptographic scheme, with practical applicability to underground agricultural IoT deployments.

The second contribution, RF Fingerprint-Based Location Authentication for Over-the-Air and Underground Wireless Networks (LAOUWN), proposes a robust location authentication framework based on channel impulse response (CIR) features and deep learning. It employs convolutional neural networks (ResNet-18/34/50) enhanced with transfer learning and adversarial domain adaptation, achieving over 90% authentication accuracy across diverse testbeds. The system demonstrates complete

resistance to advanced adversaries including Friis-based and ray-tracing-enhanced attackers whose success rate is reduced to random guessing.

The third contribution, VET: Autonomous Vehicular Credential Verification using Trajectory and Motion Vectors, presents a lightweight, privacy-preserving authentication protocol for vehicular networks. VET verifies credential legitimacy using trajectory similarity and motion-based trust metrics (TMVs), achieving a 97% true positive rate under benign conditions and a 99.9% detection rate against remote signal-manipulating adversaries. It remains agnostic to wireless channel variability and scalable to multi-attacker scenarios.

Finally, a Systematization of Knowledge for Security in Molecular and Nano-Communications surveys current threats and defense mechanisms in nanoscale networks. It identifies critical gaps such as the lack of structured taxonomies, active threat mitigation, and cross-layer integration and proposes novel solutions, including bio-inspired cryptographic models and enhanced error correction strategies.

Together, these contributions advance the field of physical-layer security by delivering robust, practical, and hard-to-forge mechanisms for secure communication in next-generation emerging wireless networks, especially in unconventional and resource-constrained settings where traditional cryptographic approaches fall short.

Publications Resulting from This Dissertation

The research presented in this dissertation has led to the following publications and manuscripts:

Peer-Reviewed Conference Publications

- Ebuka Oguchi, Nirnimesh Ghose, and Mehmet Can Vuran, *"STUN: Secret-Free Trust Establishment For Underground Wireless Networks"*, IEEE INFOCOM Workshop on Wireless Security (Wireless-Sec), Virtual, pp. 1–6, May 2–5, 2022.
- Ebuka Oguchi and Nirnimesh Ghose, *"VET: Autonomous Vehicular Credential Verification using Trajectory and Motion Vectors"*, EAI SecureComm, Hong Kong SAR, pp. 1–23, Oct. 19–21, 2023.

Under Review / In Preparation

- Ebuka P. Oguchi, Nirnimesh Ghose, and Mehmet Can Vuran, *"Soil-Assisted Trust Establishment for Underground Internet-of-Things"*, Under review at IEEE TNSM, 2025.
- Malcolm I. Anderson, Truc T. Duong, Ebuka P. Oguchi, Anna Wisniewska, and Nirnimesh Ghose, *"Systematization of Knowledge for Security in Molecular and Nanocommunications"*, In preparation for IEEE TMBMC, 2025.
- Ebuka P. Oguchi, Hakim Lado, Nirnimesh Ghose, Boyang Wang, and Mehmet Can Vuran, *"RF Fingerprint-Based Location Authentication for Over-The-Air and Underground Wireless Networks"*, In preparation for NDSS, 2025.

DEDICATION

To my family and friends who gave me all the support they could to help me complete my Ph.D. program. Most importantly, to God almighty for his mercies and favors.

ACKNOWLEDGMENTS

I would like to express my profound gratitude to my advisor, Prof. Nirnimesh Ghose, for his unwavering guidance, support, and mentorship throughout my doctoral journey. His patience and insight have truly shaped my growth as a researcher.

I am sincerely thankful to my committee members: Prof. Mehmet C. Vuran, Prof. Yi Qian, and Prof. Massimiliano Pierobon. I appreciate their thoughtful feedback, encouragement, referrals, and valuable time. I am also especially grateful to Prof. Shuai Nie and Prof. Judy Goldsmith for their mentorship and for providing recommendations that have supported my academic and professional development.

I would like to thank my lab mates in the SWAN Lab and my colleagues from the CPN Lab. Their collaboration, technical insight, and support have played an important role in my research. I also appreciate the broader academic community and fellow graduate students whose encouragement and friendship made this journey meaningful. To the friends I have made at UNL and throughout Nebraska, thank you for the memories and companionship.

To my family, I thank you for your unconditional love, prayers, and belief in me. Your support, even from a distance, has sustained me through every challenge. I dedicate this dissertation to you.

Above all, I thank God for His grace, strength, and faithfulness in every step of this journey. All praise and glory belong to Him.

Finally, I would like to acknowledge the U.S. National Science Foundation (NSF) for the financial support provided to conduct the research needed for this dissertation. This research was supported in part by NSF awards CNS-2331191 and CNS-2225161. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

The rapid integration of wireless networks across agriculture, transportation, biological systems, and beyond has created vast opportunities for scalable, intelligent, and real-time applications. From buried sensors in smart farms [151,181], to swarms of autonomous vehicles [46, 77], and nanoscale molecular communication [39, 57], these systems are transforming how data is collected, processed, and acted upon. However, they are also increasingly exposed to adversarial threats that exploit credential leakage [25, 71], physical inaccessibility [52], and limitations of cryptographic solutions [48, 174]. In many of these environments, such as underground deployments, low-power ad-hoc networks, and bio-nano systems, traditional security primitives fall short due to their reliance on pre-shared secrets, computational overhead, or lack of location verification capability [105, 109, 176]. This dissertation addresses these limitations by developing authentication and message integrity mechanisms that rely on secret-free, hard-to-forge physical-layer properties across diverse and unconventional settings.

In agricultural Internet of Things (Ag-IoT) networks, buried sensors transmit vital data on soil moisture, precipitation, and temperature through underground wireless channels to aboveground gateways. Commercial systems like GroGuru [64] and SoilScout [172] exemplify the benefits of this approach by shielding sensors from en-

vironmental damage and minimizing operational costs [64, 172]. However, recent cyberattacks on Ag-IoT infrastructures, including WiFi jamming, spoofing, and signal injection, underscore the urgent need for scalable, lightweight, and secure communication protocols [30,187]. Traditional methods such as Over-the-Air Activation (OTAA) and Activation by Personalization (ABP) remain vulnerable to jamming and replay attacks and require secret management [13,156]. To overcome these shortcomings, this dissertation introduces Soil-Assisted Trust Establishment for Underground Wireless Networks (STUN), a scalable trust bootstrapping protocol that leverages the hard-to-replicate properties of underground wireless propagation [127]. STUN achieves authentication and message integrity without pre-shared secrets and proves resilient to signal injection and collusion-based attacks, offering security guarantees comparable to post-quantum cryptographic schemes such as Unbalanced Oil and Vinegar [83,127].

Beyond soil-based systems, verifying a device's physical location is essential in over-the-air and underground networks. RF fingerprinting based on Channel Impulse Response (CIR) emerges as a compelling solution, enabling location authentication without relying on prior shared keys or passwords. CIR features encode fine-grain spatial and temporal properties of the wireless channel that are difficult to replicate from different positions, even with identical hardware [86, 102]. This dissertation presents Location Authentication for Over-the-Air and Underground Wireless Networks (LAOUWN), a machine learning-based system that uses CNN models such as ResNet-18/34/50, enhanced with transfer learning and adversarial domain adaptation, to achieve over 90% accuracy in location verification. LAOUWN demonstrates resistance to both empirical Friis-based attackers and ray-tracing-enhanced adversaries, showing that spoofing attacks from physically separate locations degrade to random guessing [87,101]. Experiments conducted using BPSK signals and USRP devices confirm the robustness of this approach across varying environments, distances,

and device combinations.

In vehicular networks and autonomous systems, secure communication is vital for ensuring motion coordination, collision avoidance, and geofencing compliance. Yet, cryptographic credentials alone are insufficient to prevent ghost vehicles or spoofed location messages [25,71]. This dissertation introduces Veracity Evaluation using Trajectory and Motion Vectors (VET) [126], a location-and-motion-based authentication framework that evaluates the consistency between a node's claimed trajectory and its motion vectors estimated from physical-layer observations. Unlike prior schemes that require multiple verifiers or static channel conditions, VET operates with a single verifier and remains agnostic to wireless channel assumptions [15, 176]. It utilizes frequency-of-arrival measurements and random sampling to estimate position and velocity, offering immunity against signal manipulation attacks. Experimental results using software-defined radios show that VET can detect spoofing attempts with 99.9% accuracy, even under adversarial conditions.

Finally, the dissertation turns to molecular and nano-communication systems, which are gaining traction in medical, environmental, and biological domains. In these systems, nano-machines exchange information through molecular diffusion or biochemical signaling. Applications include targeted drug delivery, in-vivo health monitoring, and tissue engineering [3,37]. Unlike electromagnetic systems, molecular communication depends on chemical signals affected by temperature, fluid viscosity, and medium composition [80, 91]. These characteristics present new security challenges such as molecular jamming, eavesdropping, signal degradation, and bit flipping [32, 76]. This dissertation presents a systematization of knowledge that surveys existing security mechanisms for molecular networks and highlights critical gaps in confidentiality, authentication, and availability [14, 57]. It proposes novel directions for future research, including bio-inspired authentication models, DNA-based iden-

tifiers, enhanced error correction, and secure channel coding tailored to molecular dynamics [146, 160].

Taken together, the contributions in this dissertation present a coherent vision for advancing authentication and message integrity in unconventional wireless systems. By leveraging physical-layer characteristics such as CIR, frequency-of-arrival, and environmental coupling, this work provides robust, cryptography-independent security mechanisms that remain viable even in adversarial, resource-constrained, or physically inaccessible environments. These approaches not only reduce reliance on traditional key management but also offer spatial and temporal guarantees that raise the bar for adversaries. The proposed solutions represent foundational advancements in securing next-generation wireless networks.

1.1 Motivation

The proliferation of wireless technologies across diverse domains such as underground agricultural networks, over-the-air (OTA) communications, autonomous vehicular systems, and molecular communication has enabled real-time, distributed sensing and control. These advancements underpin critical infrastructure ranging from smart farming and intelligent transportation to bio-nano healthcare applications. However, with the rise of connectivity comes an increasing risk of adversarial attacks. In many of these settings, conventional cryptographic techniques face deployment barriers due to limited computational resources, lack of pre-shared secrets, or the physical inaccessibility of nodes. This motivates the exploration of physical-layer properties as an alternative security primitive for authentication and message integrity, particularly in unconventional and resource-constrained environments.

1.2 Problem Statement

Traditional authentication and integrity mechanisms rely heavily on pre-shared keys, digital certificates, or trusted third parties. These approaches are not always viable in emerging wireless scenarios where:

- Devices are deployed in inaccessible locations (e.g., buried sensors or embedded nano-machines).
- Communication occurs over dynamic or unknown environments (e.g., soil, air, or biological tissue).
- Network entities are mobile and operate in ad hoc, decentralized configurations (e.g., autonomous vehicles).
- Resource constraints preclude the use of heavyweight cryptographic protocols.
- Spoofing and signal manipulation attacks can bypass credential-based verification.

These challenges necessitate new methods that do not depend on secret exchange or computational hardness assumptions but instead exploit inherent physical-layer characteristics for robust, passive, and scalable authentication.

1.3 Research Objectives

This dissertation aims to address the above challenges by designing and evaluating a suite of authentication and message integrity protocols for unconventional wireless networks. The specific objectives are

1. Develop a physical-layer trust bootstrapping protocol for underground wireless networks that utilizes hard-to-forge soil propagation properties, resulting in a scalable and secret-free alternative to conventional key exchange (STUN).

2. Design a deep learning-based RF fingerprinting system using Channel Impulse Response (CIR) features to verify device locations across OTA and underground environments, and evaluate its robustness against realistic adversaries (LAOUWN).

3. Create a trajectory-based message authentication scheme for vehicular networks that leverages motion vectors and frequency-of-arrival features to prevent ghost vehicle attacks and verify physical claims even with only one verifier (VET).

4. Conduct a systematization of knowledge (SoK) for molecular and nano communication systems, identifying emerging threats and limitations of current security schemes, and proposing bio-inspired and context-aware solutions tailored to the chemical communication paradigm.

5. Demonstrate the feasibility and effectiveness of each proposed system through real-world experiments using software-defined radios (SDRs), simulated adversarial attacks, and rigorous statistical evaluations.

Together, these contributions aim to shift the security paradigm from secret-based mechanisms to secret-free authentication, enabling resilient and low-overhead trust establishment in emerging wireless systems using hard-to-forge physical layer properties.

CHAPTER 2

Related Work

Security in emerging wireless networks spanning underground sensor systems, OTA authentication, vehicular trust, and nanoscale molecular communication faces unique challenges. This section reviews related efforts across these domains, highlighting gaps our work addresses.

2.1 Underground Wireless Networks

Security for underground wireless sensor networks (WUSNs) remains underdeveloped. SPRIDE [189, 190] employs homomorphic encryption to protect location privacy, but lacks authentication and integrity mechanisms for underground-to-over-the-air communication. Traditional cryptographic key exchanges like Diffie-Hellman [48] are vulnerable to man-in-the-middle attacks over public channels.

Out-of-band (OOB) solutions use audio, visual, or tactile channels for secure pairing [110, 184], but require additional hardware and are impractical for buried deployments. In-band schemes like helper-aided pairing [62] and signal cancellation [130] are mostly OTA-specific and do not model the unique soil propagation behaviors [52, 149]. Our STUN protocol fills this gap by exploiting path loss asymmetry and soil decorrelation for secure, secret-free trust establishment in underground Ag-IoT deployments.

2.2 RF Fingerprinting and Location Authentication

RF fingerprinting leverages physical-layer features such as IQ imbalance, oscillator offset, or transient emissions to identify devices [29, 74]. While effective for device authentication [18, 43], these methods do not account for environmental context and are susceptible to spoofing.

Location-based authentication instead exploits spatially unique channel characteristics like channel state information (CSI), CIR, and received signal strength (RSS). CSI and CIR-based systems [26, 101] have shown promise in indoor/outdoor localization but are rarely extended to underground settings, where soil permittivity, moisture, and temperature significantly affect propagation [50, 150].

Our LAOUWN framework applies deep learning over CIR fingerprints to authenticate node locations in both OTA and underground environments, using ResNet CNNs, fine-tuning, and adversarial domain adaptation to achieve domain-robust performance.

2.3 Autonomous Vehicle Trust and Veracity

In vehicular ad hoc networks (VANETs), cryptographic credential verification alone is insufficient: attackers may possess valid credentials or compromise them [71]. Secure position and motion verification methods such as Doppler shift [153, 176] or time-of-arrival-based distance bounding [69] require trusted multi-verifier setups or rich multipath assumptions that do not hold in all environments.

Out-of-band solutions using radar, LiDAR, or camera-based verification [97, 109] suffer from hardware cost, privacy concerns, and susceptibility to spoofing. Our VET framework verifies trajectory and motion vectors (TMVs) using a single verifier, even

in line-of-sight (LoS) and non-line-of-sight (NLoS) conditions, enabling scalable, in-band physical authentication against spoofed vehicular behavior.

2.4 Security in Molecular and Nano Communication

Molecular communication (MC) systems rely on chemical signal propagation in biological media, introducing noise, delay, and non-determinism not present in traditional wireless networks [57]. Existing surveys focus on signal modulation [90], energy efficiency [1], or biomedical applications [39], but largely overlook robust security.

Bio-inspired security methods, including those modeled after immune system responses, are beginning to emerge [53, 105]. However, gaps remain in formal threat modeling, authentication techniques, and secure molecular coding. Our work provides a comprehensive systematization of MC security, proposing secret-free, context-aware solutions suited for the nanoscale domain, such as targeted drug delivery and in-vivo networks.

Across underground, vehicular, OTA, and molecular domains, existing authentication and message integrity frameworks remain fragmented, often tied to cryptographic assumptions [48,174], trusted infrastructure [71], or costly out-of-band modalities [97, 110, 184]. These limitations become especially pronounced in adversarial or resource-constrained settings such as buried Ag-IoT nodes [148,181], high-mobility vehicular networks [46, 176], or nanoscale biological systems [53, 105]. Prior approaches typically emphasize either device identity using RF fingerprints [29,43] or physical positioning [69,153], but fail to account for environmental variability, attacker proximity, or domain transferability. In contrast, this dissertation proposes a cohesive suite of secret-free, physical-layer-based mechanisms for location authentication and message integrity verification. Collectively, these frameworks demonstrate that robust, scal-

able, and passive authentication can be achieved through environment-coupled signal features, without reliance on pre-shared secrets or additional hardware.

CHAPTER 3

Soil Assisted Trust-Establishment For Underground Wireless Networks

To realize secure communication in underground agricultural environments, it is imperative to understand the unique physical and architectural characteristics of the wireless underground medium. This chapter presents STUN, a novel trust-establishment protocol that leverages hard-to-forge underground propagation characteristics to achieve in-band authentication and secure key establishment. Before detailing the protocol design, we introduce the foundational models that guide our approach specifically, the system model, adversarial capabilities, and characteristics of the underground-to-over-the-air wireless channel. These models set the stage for a principled security framework tailored to the constraints and opportunities of wireless underground networks.

3.1 Model and Preliminaries

In this section, we first define STUN's system and adversarial models, followed by the preliminaries of underground wireless channels. We begin by presenting Table 5.1, which summarizes the frequently used notations in this paper.

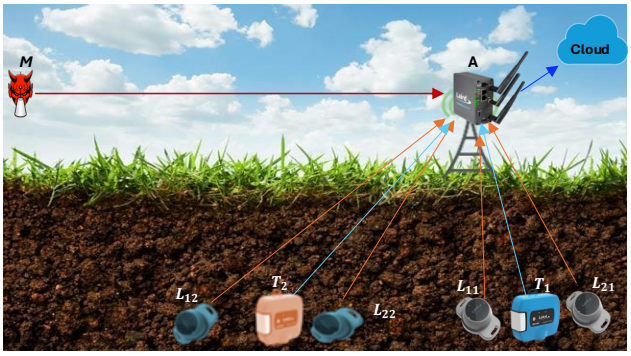


Figure 3.1: Several underground nodes $\mathbf{L}_i$ securely bootstrap with the gateway A assisted by the trusted node T_i in the presence of an adversary M.

3.1.1 System Model

As shown in Fig. 3.1, the system model consists of four types of nodes: (1) Gateway, A, (2) Trusted nodes, $\mathbf{T} = \{T_i : i \in \mathcal{I}\}$, where $\mathcal{I}$ is the index set of trusted nodes, and (3) Sets of legitimate nodes, $\mathbf{L}_i = \{L_{ij} : j \in \mathcal{J}_i\}$, where $\mathcal{J}_i$ is the index set of legitimate nodes for each trusted node T_i, and (4) an active advisory, M.

Gateway (A)**:** The aboveground gateway coordinates, captures, and authenticates the data transmitted by the deployed nodes. Functioning as a tower, it receives messages from the underground legitimate and trusted nodes. The gateway is responsible for authenticating the data received from these legitimate nodes. Moreover, the gateway is located within the farm under the user's control.

Trusted Nodes ($\mathbf{T}$)**:** The trusted nodes $\{T_1, T_2, \ldots, T_{|\mathcal{I}|}\}$ have high battery and computation power, enabling them to perform cryptographic functions efficiently. These nodes, deployed underground, collectively cover the entire farm in a sector-based deployment with a maximum sector size d^u_{max}, as shown in Fig. 3.2. A trusted channel between T_i and A is established using a shared secret K_{AT_i}, and transmissions are secured with authenticated encryption AE($\cdot$) [20]. This can be implemented as encrypt-then-MAC for symmetric cryptography or as a sign/encrypt/sign scheme for

Table 3.1: Table of Notations

Notation	Description
A	Gateway
$\mathbf{T}$	Set of trusted nodes, $\mathbf{T} = \{T_i : i \in \mathcal{I}\}$
$\mathcal{I}$	Index set of trusted nodes
$\mathbf{L}_i$	Set of legitimate nodes, $\mathbf{L}_i = \{L_{ij} : j \in \mathcal{J}_i\}$
$\mathcal{J}_i$	Index set of legitimate nodes for each trusted node T_i
M	Adversary
d^u_{xy}	Underground path length between x and y
d^u_x	Underground depth of node x
d^a_{xy}	Aboveground path length between x and y
α^u_x	Underground attenuation constant at node x
β^u_x	Underground phase shift constant at x
η	OTA attenuation constant
ϵ^a_x	OTA permittivity constant
ϵ^u_x	Underground permittivity constant
μ_x	Relative permeability
P^{tx}_x	Transmit power of node x
P^{rx}_{xy}	Received power at node x from y
G_x	Antenna gain of the node x
PL^u	Underground pathloss
PL^a	Over-the-air pathloss
PL^ρ	Refraction loss
m_x	Message transmitted by node x
τ	Threshold for detection of adversarial node among T_i and L_{ij}
τ^i_x	Threshold for identifying outlier received signal strength (RSS) for T_i
d^u_{max}	Maximum sector size

public key cryptography. The protocol, designed for a single T_i, can scale to multiple $\mathbf{T}$ nodes to ensure complete farm coverage.

Legitimate Nodes ($\mathbf{L}_i$): The nodes $\{L_{i1}, L_{i2}, \ldots, L_{i|\mathcal{J}_i|}\}$ are deployed underground within the user's control and are within the communication range of at least one trusted node. These legitimate nodes collect and transmit data directly to A through a wireless channel. The deployment of the legitimate nodes is randomized, subject to application requirements, covering the entire farm area as depicted in Fig.

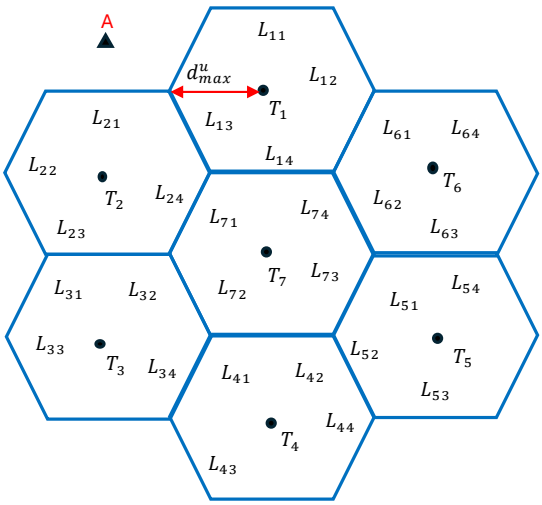


Figure 3.2: Several underground Trusted nodes $T_i \in \mathbf{T}$ with the over-the-air gateway A deployed in the field in a hexagonal sector pattern.

3.2.

3.1.2 Threat Model

An active adversary (M), controlling one or more colluding devices, operates outside the trusted farm's perimeter, such as on adjacent roads, without entering the fields[1]. The adversary aims to spoof messages and bootstrap at A by posing as a rogue node.

The adversary employs signal injection attacks, including overshadowing attacks, which require only 6dB stronger signal for LPWAN technologies [156]. While M knows the communication protocol, they lack physical access to the nodes and cannot perform jamming or physically block signals (e.g., via a Faraday cage).

There are two types of attackers with advanced capabilities considered in this scenario.

[1]Trespassing is assumed to be deterred by law enforcement.

Type 1 Adversary: This adversary attempts to inject its signal simultaneously at both A and T_i.

Type 2 Adversary: Type 2 adversary can deploy colluding aboveground and underground wireless nodes in addition to the capabilities of Type 1 M. By leveraging these additional nodes, the Type 2 adversary can achieve the required received signal strength (RSS) at both A and T_i.

3.1.3 Preliminaries: Underground-to-Air Wireless Channel Model

We present the underground-to-OTA channel model, essential for the security protocol due to the channel's inherent unpredictability [52]. This study focuses on underground nodes communicating with an aboveground gateway. In underground communication, electromagnetic wave propagation experiences significant attenuation influenced by soil properties such as soil composition and soil moisture, which lead to higher permittivity than in air [51]. Moreover, variations in soil moisture alter the resonance characteristics of underground antennas, leading to additional loss. Furthermore, the soil-air interface, through which the waves propagate, cause reflection and refraction, impacting the received signal.

Consider A, located aboveground, receiving wireless signal from an underground node L_{ij}, as illustrated in Fig. 3.3. The wireless signal travels a path length of $d_{AL_{ij}} = d^a_{AL_{ij}} + d^u_{L_{ij}}$, where $d^a_{AL_{ij}}$ represents the distance from the point where the signal crosses the soil-air border to node A, and $d^u_{L_{ij}}$ represents the vertical distance between the underground node and the soil-air interface. The underground path $d^u_{L_{ij}}$ is approximately equal to the depth of the underground node $d^u_{L_{ij}}$ since the dominant signal takes the shortest path to exit the underground environment [52]. The deployment depth of all the underground nodes and the trusted node remains consistent, ensuring uniform signal propagation characteristics. At the soil-air border,

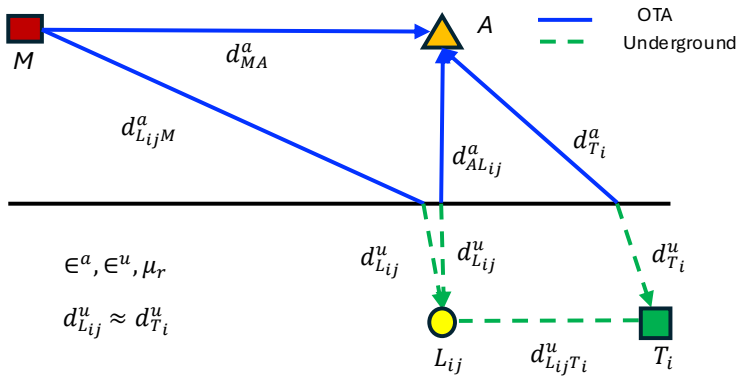


Figure 3.3: An adversary M attempts to hijack the bootstrap session between L_{ij} and A, while T_i performs simultaneous verification.

the signal undergoes refraction due to the changes in the propagation medium. The power received by L_{ij} in dB can be expressed as [52]:

$$P^{rx}_{L_{ij}A} = P^{tx}_A + G_A + G_{L_{ij}} - PL^u - PL^a - PL^p, \tag{3.1}$$

where P^{tx}_A is the transmit power of A, G_x is the antenna gain of the transceiver x, and PL^u, PL^a, PL^p are the losses due to underground and OTA paths, and refraction at the ground level, respectively in dB.

The path loss of the OTA wireless channel is given by [52]:

$$PL^a = C_a + 10\eta \log(d^a_{AL_{ij}}) + 20 \log(f), \tag{3.2}$$

where $d^a_{AL_{ij}}$ is the distance between the point where the wireless channel transitions from underground to air and node A, η is the attenuation factor, f is the center frequency, and $C_a = -147.6$.

The path loss attributed to the underground channel is given by [52]:

$$PL^u = C_u + 20\log(d^u_{L_{ij}}) + 20\log(\beta^u_{L_{ij}}) + \rho_u \alpha^u_{L_{ij}} d^u_{L_{ij}}, \tag{3.3}$$

where $d^u_{L_{ij}}$ represents the depth of the underground node L_{ij}, α^u_x and β^u_x denote the attenuation and phase-shifting constants of the soil, and $C_u = 6.4, \rho_u = 8.69$, respectively.

Lastly, the path loss caused by refraction can be calculated separately for air-to-underground as follows:

$$PL_{\rho,a-u} = 20\log\left(\frac{r+1}{4}\right),$$

where $r = \sqrt{(\sqrt{(\epsilon^a)^2 + (\epsilon^u)^2} + \epsilon^a)/2}$ represents the refractive index of the soil [52]. Alternatively, for the underground-to-air link, the signal propagates perpendicularly without refraction, resulting in $PL_{\rho,u-a} = 1$ [52].

3.2 Secret-Free Trust-Establishment for Underground Wireless Networks

We introduce STUN, a trust establishment protocol for an underground wireless networks that leverages the unique physical propagation properties of soil for secure in-band communication. STUN incorporates a unique PHY-layer trust verification primitive to authenticate the legitimacy of the underground nodes, $\mathbf{L}_i$, and ensure the integrity of their transmissions.

3.2.1 STUN: The Protocol

The central concept involves A coordinating the key establishment process with each $\mathbf{L}_i$ in their respective sectors using a time-division approach[2]. First, A creates

[2]In this work, we assume a time-division approach, whereas any medium access control approach could be adopted without loss of generality.

a transmission schedule for T_i and broadcasts this schedule to all trusted nodes $\mathbf{T}$. Subsequently, each T_i sends a message to coordinate all the legitimate nodes $\mathbf{L}_i$ within its communication range.

Once the legitimate nodes, $\mathbf{L}_i$, receive the coordination message from T_i, they transmit their key primitives in a time-division manner. These key primitives are then verified by the trusted node and the gateway to ensure their authenticity and integrity. The trust establishment protocol consists of the following steps:

1. **Initialization:** The protocol begins with A transmitting a synchronization message to $\mathbf{L}_i$ and $\mathbf{T}$. Next, A schedules all participating trusted nodes $ID_{T_1}, \ldots, ID_{T_{|\mathcal{I}|}}$. The gateway sends an initialization message to T_i as $\mathrm{AE}_{K_{AT_i}}(\mathrm{INIT}||\eta)$, where η is a nonce. $\mathrm{AE}_{K_{AT_i}}$ ensures the confidentiality and integrity of the message using a shared key K, allowing only T_i to decrypt and verify it. All entities agree on Diffie-Hellman (DH) public parameters: $\mathbb{G}, q, g$.

2. **Initialization of a sector:** The trusted node, T_i, broadcasts a Ready-to-Authenticate (RTA) message to all the underground nodes in the vicinity, $\mathbf{L}_i$.

3. **Primitive transmission from $\mathbf{L}_i$:** Each legitimate node, $L_{ij} \in \mathbf{L}_i$, picks a secret value, $X_{ij} \in_U \mathbb{Z}_q$, computes the public value, $z_{ij} \leftarrow g^{X_{ij}}$, and transmits its message, $m_{ij} \leftarrow \{ID_{ij}, z_{ij}\}$.

4. **Verification at T_i:** T_i synchronizes with the preamble and receives all the messages, $m'_{ij}\ \forall\ j = 1, \ldots, |\mathcal{J}_i|$, and records the corresponding received signal strength (RSS),

   $\mathbf{r}_j^{T_i} = \{r_j^{T_i}(1), r_j^{T_i}(2), \ldots, r_j^{T_i}(\ell)\}$. Finally, T_i performs the following verification if

$$\tau_{low}^{T_i} \leq \mathbf{r}_j^{T_i}(\kappa) \leq \tau_{high}^{T_i};\ \ \forall\ j = 1, \ldots |\mathcal{J}_i|;\ \kappa = 1 \ldots \ell.$$

After successful verification, the trusted node relays $m_i := \text{AE}_{K_{AT_i}}(m'_{i1}||ID_{i1}, \ldots, m'_{ij}||ID_{ij}, \ldots, m'_{i|\mathcal{J}_i|}||ID_{i|\mathcal{J}_i|})$ to A after $L_{i|\mathcal{J}_i|}$'s transmission in a time division fashion.

5. **Reception at A:** The gateway, A, records RSS samples, $\mathbf{r}_{ij} = \{r_{ij}(1), r_{ij}(2), \ldots, r_{ij}(\ell)\}$, while receiving m''_{ij}. Further, A records the RSS samples, $\mathbf{r}_{T_i} = \{r_i(1), r_i(2), \ldots, r_i(\ell')\}$, while receiving m'_i.

6. **Verification at A:** The gateway decrypts m'_i to obtain $m'_{ij}||ID_{ij}\ \forall\ j = 1, \ldots, |\mathcal{J}_i|$ and verifies its integrity using the corresponding verification function. A rejects all received messages if verification fails. Further, A verifies $m'_{ij} \stackrel{?}{=} m''_{ij}\ \forall\ j = 1, \ldots, |\mathcal{J}_i|$; and rejects if the verification fails. Finally, A computes:

$$\Gamma_{ij} = \{\gamma_{ij}(1), \gamma_{ij}(2), \ldots, \gamma_{ij}\ell\}, \quad \gamma_{ij}(\kappa) = \frac{r_i(\kappa)}{r_{ij}(\kappa)}$$

$\forall\ j = 1, \ldots, |\mathcal{J}_i|$. The gateway accepts m''_{ij} if $\tau_{low} \leq \gamma_{ij}(\kappa) \leq \tau_{high};\ \ \forall\ \kappa = 1 \ldots \ell$, $j = 1, \ldots, |\mathcal{J}_i|$.

7. **Primitive transmission from A:** Following successful verification, A picks a secret value, $X_A \in_U \mathbb{Z}_q$, computes the public value, $z_A \leftarrow g^{X_A}$, and transmits as $m_A \leftarrow \{ID_A, z_A\}$.

8. **Key establishment:** After reception of the message, $\mathbf{L}_j$ computes the pairwise keys as $K_{Aij} \leftarrow (z_A)^{X_{ij}}$ and A computes as $K_{Aij} \leftarrow (z_{ij})^{X_A}$. Immediately following the key agreement, L_{ij} and A engage in a key confirmation phase, initiated by L_{ij}. This can be done by executing a two-way challenge-response protocol [28].

9. **Repeat for all sectors:** All the above steps are repeated for $i = 1, \ldots, |\mathcal{I}|$. If

any of the steps fail for a legitimate node, L_{ij}, is detected at any of the entities, and the steps are repeated for the node. Otherwise, the gateway, A, notifies the user of the successful completion of the trust establishment.

STUN employs the Diffie-Hellman key exchange protocol [48] for trust establishment, enabling pairwise key generation in star topologies and group key creation for mesh networks to support secure group operations [27]. If trust establishment steps like key confirmation fail, it may indicate adversarial activity or communication issues. The affected node can notify a trusted node to retry the process, maintaining protocol integrity. Persistent failures may require human intervention to address potential security breaches.

3.2.1.1 Securing Downlink (A-to-$\mathbf{L}_i$) communication

In the protocol outlined in Section 3.2.1, underground nodes (L_{ij}) do not explicitly verify the authenticity of gateway messages (m_A) during the key confirmation process. An adversary (M) impersonating the gateway would disrupt the session with the legitimate gateway (A), triggering protocol failure and user notification. Trusted nodes (T_i) are employed to enable explicit message verification. After A validates messages from L_{ij} nodes, it sends an authenticated and encrypted version (m'_A) to T_i using a shared key (K_{AT_i}). Simultaneously, A broadcasts the plaintext message (m_A) to all L_{ij} nodes and their respective T_i nodes. Each T_i verifies the authenticity of m_A by comparing it with m'_A. Depending on the result, T_i broadcasts a success or failure message. While adversaries could exploit this to perform a DoS attack, such scenarios are outside the scope of this work.

3.2.1.2 Size of T_i's Sector

The size of the sector for T_i is determined by the maximum possible distance, denoted as d^u_{max}, between the trusted node, T_i, and the legitimate nodes, $\mathbf{L}_i$. We employ the underground path loss model given by (3.3) to calculate this distance. Considering the assumption that the legitimate nodes, $\mathbf{L}_i$, have lower capabilities and transmit at lower power compared to T_i, this determines the communication range of the sector. The size of T_i's sector can be calculated using (3.3), yielding the following expression:

$$d^u_{max} = \frac{20W\left(\alpha^u_x \rho_u e^X K\right)}{\alpha^u_x \rho_u \log(10)}. \tag{3.4}$$

where $K = \log(10)/20$ is a constant multiplier, and $X = [\log(10^{-C_u}) + PL^u \log(10) - 20\log(\beta^u_x)]/20$ is an intermediate variable that encapsulates the logarithmic and path loss dependencies.

3.2.2 Selection of Thresholds

Now, we first theoretically describe selecting the thresholds utilized in Steps 3 and 5 of STUN:

3.2.2.1 Thresholds for T_i

The detection thresholds at trusted nodes (T_i) for identifying outlier RSS values are determined using the Median Absolute Deviation (MAD) method, a robust measure less sensitive to outliers compared to standard deviation [96]. The thresholds are defined as:

$$\tau^{T_i}_{low} = \widetilde{\mathbf{r}^{T_i}_j} - \zeta * \nu, \quad \tau^{T_i}_{high} = \widetilde{\mathbf{r}^{T_i}_j} + \zeta * \nu, \tag{3.5}$$

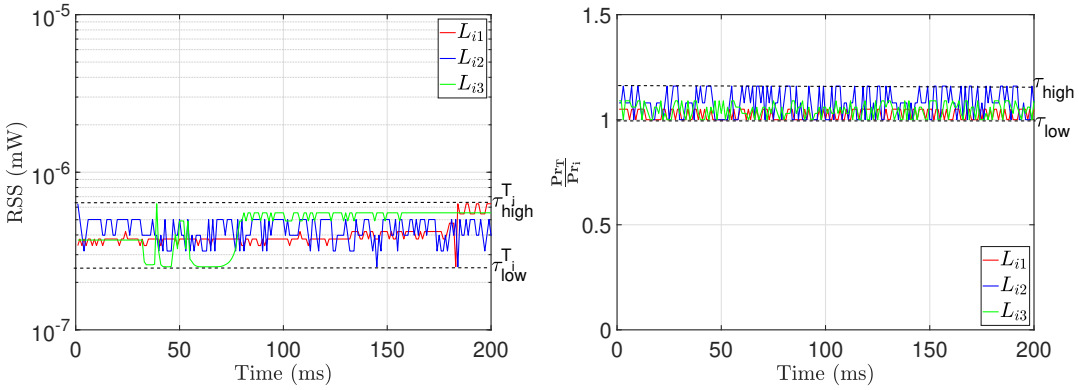


Figure 3.4: (a) Power received at the trusted node from three $\mathbf{L}_i$, and (b) RSS ratio at A of power received from T_i to power received from $\mathbf{L}_i$.

where $\widetilde{\mathbf{r}_j^{T_i}}$ is the median of all the RSS samples, ζ controls the strictness of the outlier rule, and ν is MAD

$$\nu = b \cdot |\widetilde{\mathbf{r}_j^{T_i}} - \widetilde{\mathbf{r}_j^{T_i}}|, \tag{3.6}$$

where $\tilde{\cdot}$ is the median of all the samples over all the nodes, and $b = 1/Q(0.75) = 1.4826$. Empirical results [202] in Fig. 3.4(a) illustrate that the RSS received at T_i from three underground nodes. We observe that the RSS is relatively stable over time and that the $\tau^{T_i}_{low} = 2.512 \times 10^{-7}$mW and $\tau^{T_i}_{high} = 6.309 \times 10^{-7}$mW, due to high path loss of the underground communication.

3.2.2.2 Thresholds for A

To calculate the threshold for the gateway, the power received from the trusted node, T_i, is divided by the power received from the legitimate node, $\mathbf{L}_i$. The threshold

can be expressed as:

$$\tau = 20\log(d^u_{T_i}) - 20\log(d^u_{L_{ij}}) + 20\log(\beta^u_{T_i}) - 20\log(\beta^u_{L_{ij}}) + \rho_u \alpha^u_{T_i} d^u_{T_i} - \rho_u \alpha^u_{L_{ij}} d^u_{L_{ij}} \ , \tag{3.7}$$

where $d^u_{T_i}$ is the underground depth of the trusted node, $d^u_{L_{ij}}$ is the underground depth of L_{ij}, and α^u_x and β^u_x are the parameters due to soil characteristics. The aboveground distance to the gateway A cancels out due to the proximity of the underground node. Since variations in the OTA path loss may affect the accuracy of this calculation, an adjustment factor δ_{low} and δ_{high} are introduced. Two thresholds are then defined as:

$$\tau_{low} = \tau - \delta_{low}, \quad \tau_{high} = \tau + \delta_{high}. \tag{3.8}$$

Empirical results in (Fig. 3.4(b)), illustrate that the RSS ratio between signals from T_i and $\mathbf{L}_i$ remains close to 1 over time. The calculated thresholds $\tau_{low} = 0.9$ and $\tau_{high} = 1.160$, indicate less than 10% error due to underground fading. This confirms the protocol's effectiveness in mitigating the impact of underground channel variability on OTA communication.

3.3 Security Analysis of STUN

We analyze the security of STUN against the adversary defined in Section 3.1.2. In an attempt to pair with the gateway A, the adversary can compute $z_M := g^{X_M} \mod p$ where X_M is uniformly chosen from the set $\mathbb{Z}_q$. Then compiling and injecting a message $m_M := \{ID_M, z_M\}$ to the gateway A. However, for A to accept the message, the adversary must pass the verification of Steps 4 and 6 at T_i and A, respectively, in STUN's verification. We will first analyze a Type 1 adversary followed by a Type

2 adversary.

3.3.1 Type 1 Adversary

A Type 1 M, is a remote aboveground entity outside the farm injecting m_M at power P_M^{tx} to A and T_i as shown in Fig. 3.3. M has to compute the power P_M^{tx} to defeat Step 4 and Step 6 simultaneously. In the next proposition, we evaluate the adversary's capability to compute the transmit power P_M^{tx}. It should be noted here that M does not have a visual channel to T_i and $\mathbf{L}_i$ to know their actual locations underground. We present the proof for single and multiple $\mathbf{T}$.

Proposition 1. A Type 1 aboveground adversary M simultaneously injecting messages at T_i, as well as A, can be detected with certainty if the distance between M and A does not satisfy

$$\begin{aligned} 10\eta \log(d^a_{MA}) = C_a + C_u + PL_{\rho,a-u} + \rho_u \alpha^u_{L_{ij}} d^u_{L_{ij}T_i} + \eta 10 \log(d^a_{AT_i}) \\ + \eta 10 \log(d^a_{MT_i}) + 20 \log(\beta^u_{T_i}) - 20 \log(\beta^u_{L_{ij}}) \\ + 40 \log(d^u_{T_i}) - 20 \log(d^u_{L_{ij}T_i}) + 20 \log(f), \end{aligned}$$

where d^a_{xy} is the aboveground distance between x and y, d^u_{xy} is the underground distance between x and y, d^u_x is the underground depth of x, α^u_x and β^u_x are the underground attenuation constant and phase shifting constants for transmission from x, respectively, η is the OTA attenuation constant, and $\delta = \delta^a \pm \delta^u$, where δ^a is the relaxation introduced by the outlier evaluation technique and δ^u is the relaxation introduced due to the threshold selection.

Proof. A Type 1 adversary M injecting its message from an aboveground location will need to pass the verification of Steps 4 and 6 for A to accept and bootstrap M.

We evaluate the strategy of M to compute the transmit power independently for steps 4 and 6, and then we evaluate the effect of one step on the other. Using (3.1), (3.2), (3.3), and Step 4, the transmit power required by M for passing the verification in Step 4. The power received at T_i from L_{ij} is given by for $C_u = 6.4$ and antenna gains G_X

$$\begin{aligned} P^{rx}_{T_iL_{ij}} &= P^{tx}_{L_{ij}} + G_{L_{ij}} + G_{T_i} - C_u \\ &\quad - 20\log(d^u_{L_{ij}T_i}) - 20\log(\beta^u_{L_{ij}}) - \rho_u \alpha^u_{L_{ij}} d^u_{L_{ij}T_i}. \end{aligned} \tag{3.9}$$

Moreover, the power received from M at T_i is given by for $C_a = -147.6$

$$\begin{aligned} P^{rx}_{T_iM} &= P^{tx}_{M} + G_{M} + G_{T_i} - C_a - C_u - 10\eta\log(d^a_{MT_i}) - 20\log(f) \\ &\quad - 20\log(d^u_{MT_i}) - 20\log(\beta^u_{T_i}) - \rho_u \alpha^u_{T_i} d^u_{MT_i} - PL_{\rho,a-u}. \end{aligned} \tag{3.10}$$

In Step 4, T_i accepts M's signal if (3.9) and (3.10) are satisfied with some relaxation, which gives M's transmit power as

$$\begin{aligned} P^{tx}_{M} &= P^{tx}_{L_{ij}} + G_{L_{ij}} - G_M + C_a - 20\log(d^u_{L_{ij}T_i}) + 20\log(d^u_{MT_i}) \\ &\quad - 20\log(\beta^u_{L_{ij}}) + 20\log(\beta^u_{T_i}) + \rho_u \alpha^u_{L_{ij}} d^u_{L_{ij}T_i} - \rho_u \alpha^u_{T_i} d^u_{MT_i} \\ &\quad + 10\eta\log(d^a_{MT_i}) + 20\log(f) + PL_{\rho,a-u} \pm \delta^a, \end{aligned} \tag{3.11}$$

where d^a_{xy} is the aboveground distance between x and y, d^u_{xy} is the underground distance between x and y, α^u_x and β^u_x the underground attenuation constants and phase shifting constants for x, respectively, η is the OTA attenuation constant, and δ^a is the relaxation for the outlier evaluation technique.

Now, for evaluating the capability of Type 1 M in defeating Step 6, we compute

the power received by A:

$$P^{rx}_{AM} = P^{tx}_{M} + G_M + G_A - C_a - 10\eta \log(d^a_{MA}) - 20\log(f). \tag{3.12}$$

Then the power received by the gateway from T_i is:

$$\begin{aligned} P^{rx}_{AT_i} = P^{tx}_{T_i} + G_A + G_i - C_a - C_u - 10\eta \log(d^a_{AT_i}) \\ - 20\log(f) - 20\log(d^u_{T_i}) - 20\log(\beta^u_{T_i}) - \rho_u \alpha^u_{T_i} d^u_{T_i}. \end{aligned} \tag{3.13}$$

Equating (3.13) and (3.12) and equating to the threshold (3.7), we compute the transmission power to pass Step 6 as:

$$\begin{aligned} P^{tx}_{M} = P^{tx}_{T_i} - G_M + G_{T_i} - C_u \\ - 10\eta \log(d^a_{AT_i}) + 10\eta \log(d^a_{MA}) \\ - 20\log(d^u_{T_i}) - 20\log(\beta^u_{T_i}) - \rho_u \alpha^u_{T_i} d^u_{T_i} \pm \delta^u, \end{aligned} \tag{3.14}$$

where δ^u is the relaxation introduced due to the threshold selection. Now, for the adversary to pass both the verification simultaneously, the transmit power in (3.11) should equate to the transmit power in (3.14). Equating (3.11) and (3.14) and approximating $d^u_{MT_i} \approx d^u_{AT_i} \approx d^u_{T_i}$ to depth of the node, we compute the distance between the adversary and the gateway as:

$$\begin{aligned} 10\eta \log(d^a_{MA}) = C_a + C_u + PL_{\rho,a-u} + \rho_u \alpha^u_{L_{ij}} d^u_{L_{ij}T_i} \\ + \eta 10\log(d^a_{AT_i}) + \eta 10\log(d^a_{MT_i}) \\ + 20\log(\beta^u_{T_i}) - 20\log(\beta^u_{L_{ij}}) \\ + 40\log(d^u_{T_i}) - 20\log(d^u_{L_{ij}T_i}) + 20\log(f), \end{aligned} \tag{3.15}$$

where $\delta = \delta^a \pm \delta^u$, and assuming $P^{tx}_{L_{ij}} + G_{L_{ij}} - P^{tx}_{T_i} - G_i = 0$, or the power transmitted by the legitimate and trusted node is the same, and the antenna gains are the same. Therefore, a Type 1 adversary at d^a_{MA} of the gateway can defeat STUN by transmitting from an aboveground location. □

When several trusted nodes are installed on the farm and transmit information to a single A, numerous legitimate nodes are also deployed in a single/multiple node(s)-to-single trusted node arrangement in each sector to cover the entire farm. Each trusted node is connected to at least one legitimate node, and all the nodes are transmitting their signal to A, where $L_{ij} \in \mathbf{L}_i$ and $T_i \in \mathbf{T}$. We provide empirical results to visualize this distance in Section 3.4.3.1. Our experimental results show that an adversary would require an infeasibly high transmission power of $\approx 10^6$W, and the adversary needs to be positioned at a very far distance reaching about $\approx$ 5Km range to execute this attack, making such an attack impractical.

3.3.2 Security Against a Rogue Gateway

An aboveground Type 1 M attempts to inject m_M at one or more legitimate nodes (L_{ij}), emulating a rogue A. The adversary, typically located outside the farm's perimeter, is assumed to be closer to the legitimate node than the trusted node, resulting in lower attenuation to L_{ij}. The adversary's message (m_M) is simultaneously received by both L_{ij} and the corresponding trusted node (T_i) in the same underground environment. For m_M to be reliably received at T_i, the SNR at T_i must equal the SNR at L_{ij}. To analyze this, the power received by the trusted node ($P^{rx}_{T_iM}$) and the legitimate node ($P^{rx}_{L_{ij}M}$) from the adversary is evaluated. The conditions under which $P^{rx}_{T_iM}$ is greater than or equal to $P^{rx}_{L_{ij}M}$ are derived, considering the adversary's transmission power (P^{tx}_M) and its relative proximity to L_{ij} compared to T_i.

Proposition 2. An aboveground Type 1 M attempting to pair with any underground node L_{ij} as a rogue gateway can be detected with certainty when located at a distance given by:

$$\begin{aligned} 10\eta \log(d^a_{MT_i}) - 20\log(d^u_{T_i}) - 20\log(\beta^u_{T_i}) - \rho_u \alpha^u_{T_i} d^u_{T_i} \\ \geq 10\eta \log(d^a_{L_{ij}M}) - 20\log(d^u_{L_{ij}}) - 20\log(\beta^u_{L_{ij}}) \\ - \rho_u \alpha^u_{L_{ij}} d^u_{L_{ij}}, \end{aligned}$$

where d^a_{xy} is the aboveground distance between x and y, and d^u_x is the underground depth of x. $\beta^u_{L_{ij}}$ and $\alpha^u_{L_{ij}}$ are the attenuation and phase-shifting constants, for each node L_{ij}, respectively.

Proof. In an attempt to force a legitimate node L_{ij} with itself, the adversary M injects their message m_M with transmit power P^{tx}_M to over overshadow A's message m_A. For evading detection, m_M should be received at L_{ij} and simultaneously not received at the corresponding T_i. We consider the best case for the adversary where the legitimate node is closer to the adversary than the trusted node, as shown in Fig. 3.3, such that the attenuation from the adversary to the legitimate node is lower than the trusted node.

We compute the power received by L_{ij} from M given by

$$\begin{aligned} P^{rx}_{L_{ij}M} = P^{tx}_M + G_M + G_{L_{ij}} - C_a - C_u - 10\eta \log(d^a_{L_{ij}M}) \\ - 20\log(f) - 20\log(d^u_{L_{ij}M}) - 20\log(\beta^u_{L_{ij}}) \\ - \rho_u \alpha^u_{L_{ij}} d^u_{L_{ij}M} - PL_{\rho,a-u}, \end{aligned} \tag{3.16}$$

where P^{tx}_M is the transmit power by M, G_x is the antenna gains of the transceivers,

$PL_{\rho,a-u}$ is the air-to-underground path loss, d^a_{xy} is the above ground distance between x and y, d^u_{xy} is the underground distance between x and y, f is the center frequency, α^u_x, β^u_x are the underground attenuation and phase shift constant from x.

Next, we compute the power simultaneously received by the corresponding trusted node T_i from M, given by

$$\begin{aligned} P^{rx}_{T_iM} = P^{tx}_M + G_M + G_{T_i} - C_a - C_u - 10\eta \log(d^a_{MT_i}) \\ - 20\log(f) - 20\log(d^u_{MT_i}) - 20\log(\beta^u_{T_i}) \\ - \rho_u \alpha^u_{T_i} d^u_{MT_i} - PL_{\rho,a-u}, \end{aligned} \tag{3.17}$$

where P^{tx}_M is the transmit power by M, G_x is the antenna gains of the transceivers, $PL_{\rho,a-u}$ is the air-to-underground path loss, d^a_{xy} is the above ground distance between x and y, d^u_{xy} is the underground distance between x and y, f is the center frequency, α^u_x, β^u_x are the underground attenuation and phase shift constant from x.

For the Type 1 adversary to be detected, the message m_M should be received simultaneously at T_i and L_{ij}, as the detection is performed at T_i. This happens with certainty if the received power received $P^{rx}_{T_iM} \geq P^{rx}_{L_{ij}M}$, or T_i receives higher power than L_{ij}. This is under the assumption both T_i and L_{ij} are in similar underground environment such as SNR at T_i is greater than or equal to that of L_{ij}. Now from (3.17) and (3.16); $P^{rx}_{MT_i} \geq P^{rx}_{L_{ij}M}$ gives:

$$\begin{aligned} & P^{tx}_M + G_M + G_{T_i} - C_a - C_u - 10\eta \log(d^a_{MT_i}) - 20\log(f) \\ & \quad - 20\log(d^u_{MT_i}) - 20\log(\beta^u_{T_i}) - \rho_u \alpha^u_{T_i} d^u_{MT_i} - PL_{\rho,a-u} \\ & \geq P^{tx}_M + G_M + G_{L_{ij}} - C_a - C_u - 10\eta \log(d^a_{L_{ij}M}) - 20\log(f) \\ & \quad - 20\log(d^u_{L_{ij}M}) - 20\log(\beta^u_{L_{ij}}) - \rho_u \alpha^u_{L_{ij}} d^u_{L_{ij}M} - PL_{\rho,a-u}. \end{aligned} \tag{3.18}$$

Now we can simplify the numerator if we assume $G_{T_i} = G_{L_{ij}}$, this is a valid assumption with the most underground sensors equipped with similar wireless modules. Under the assumption that T_i and L_{ij} are in similar underground environments, they will experience similar OTA-to-underground pathloss. Hence we can simplify $PL_{\rho,a-u}$ and approximating $d^u_{Mx} \approx d^u_{Ax} \approx \log(d^u_x)$ to depth of the node x which gives:

$$\begin{aligned} 10\eta \log(d^a_{MT_i}) - 20\log(d^u_{T_i}) - 20\log(\beta^u_{T_i}) - \rho_u \alpha^u_{T_i} d^u_{T_i} \\ \geq 10\eta \log(d^a_{L_{ij}M}) - 20\log(d^u_{L_{ij}}) - 20\log(\beta^u_{L_{ij}}) \\ - \rho_u \alpha^u_{L_{ij}} d^u_{L_{ij}}. \end{aligned} \tag{3.19}$$

where d^a_{xy} is the above ground distance between x and y, d^u_{xy} is the underground distance between x and y, $\beta^u_{L_{ij}}$ and $\alpha^u_{L_{ij}}$ are the attenuation and phase-shifting constants, for each node $L_{ij} \in \mathbf{L}_i$, respectively. □

The inequality in Proposition 2 can be simplified using numerical methods like the Newton-Raphson approximation [196]. In Fig. 3.5, we approximate the distances between the adversary (M), the legitimate node (L_{ij}), and the trusted node (T_i), assuming $d^u_{MT_i} = d^u_{L_{ij}M} + d^u_{L_{ij}T_i}$. Solving for these distances requires optimization techniques, where convergence depends on factors like the initial guess (x_0), tolerance (tol), and the number of iterations (N_{max}). Precision in setting tolerance and monitoring convergence is critical to balance computational cost and accuracy, as improper settings may prevent convergence or lead to suboptimal results.

Figures 3.5(a) and (b) illustrate that small changes in $d^u_{L_{ij}T_i}$ *significantly affect the underground distance between* M *and* L_{ij}, making it difficult for M, which lacks precise knowledge of T_i's location, to target legitimate nodes without being detected.

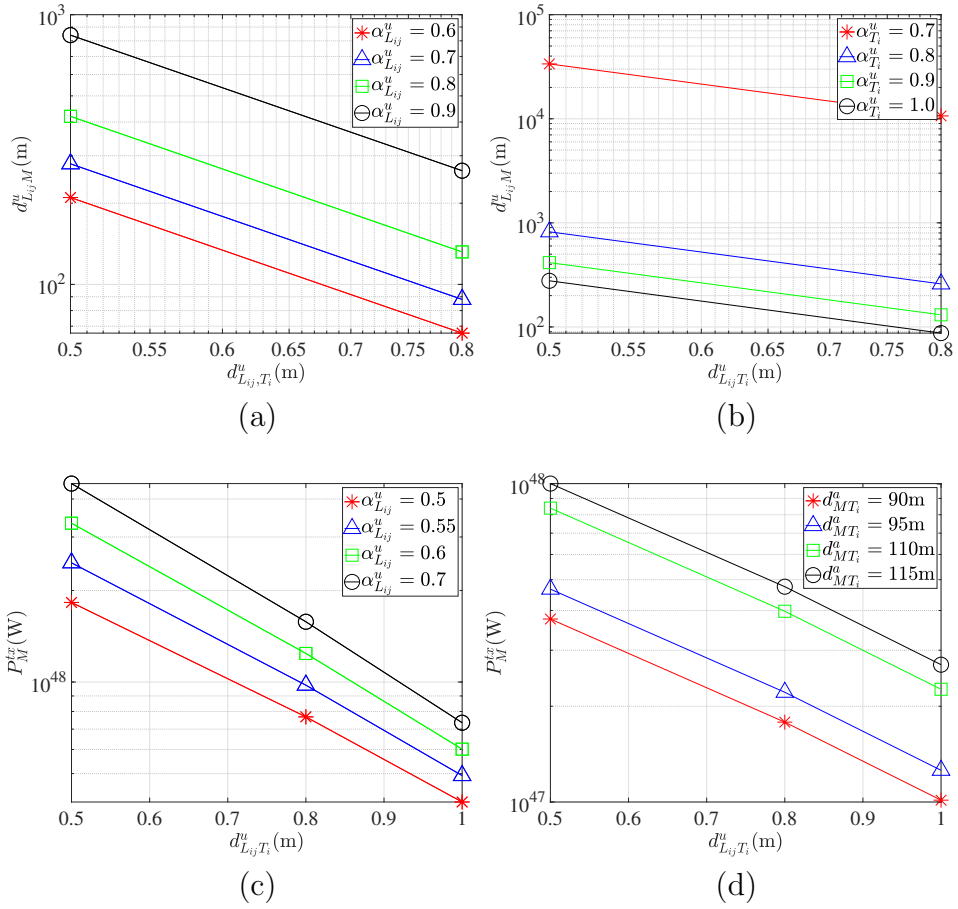


Figure 3.5: Plots of the distance of Type 1 M from the legitimate nodes against the distance of the legitimate nodes from the trusted nodes while varying (a) the attenuation of the legitimate nodes, and (b) the attenuation of the trusted node. Plots of the power transmitted from the Type 1 M to the legitimate node against the legitimate underground distances while varying (c) the attenuation $\alpha^u_{T_i}$, and (d) the aboveground M distance from T_i.

Similarly, Figures 3.5(c) and (d) show that even slight variations in legitimate underground distances require the adversary to transmit at very high power levels to compensate for attenuation. For instance, to avoid detection, M must be located approximately 1 km away and transmit at an infeasible power of about 10^{47} W. This supports the hypothesis that an adversary outside the farm's limits emulating a rogue gateway will be detected with certainty.

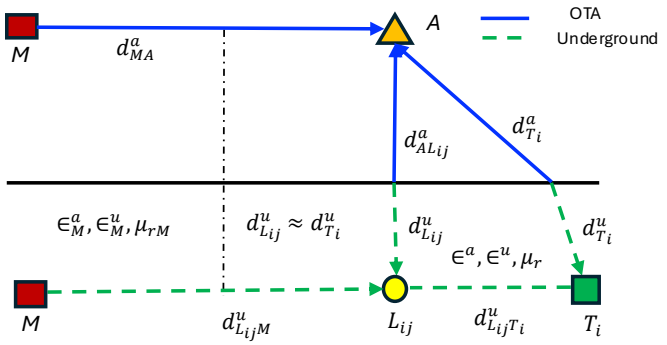


Figure 3.6: A Type 2 M attempting to inject m_M to highjack the bootstrap session of L_{ij} with A with T_i performing simultaneous STUN verification.

3.3.3 Type 2 Adversary

A Type 2 M is a remote aboveground adversary colluding with an underground adversarial node outside the farm injecting m_M transmitting at above-ground power P^{tx}_M, and underground P^{tx}_M to A and T_i, respectively as shown in Fig. 3.6. M with a visual channel to A may be able to predict the channels for computing the aboveground power P^{tx}_M to defeat Step 6. In the next proposition, we evaluate the adversary's capability to compute the transmit underground power P^{tx}_M for defeating Step 4. It should be noted here that M does not have a visual channel to T_i and $\mathbf{L}_i$. We present the proof for single then generalizing for multiple $\mathbf{T}$.

Proposition 3. STUN can detect a Type 2 aboveground and underground colluding adversary M injecting message at T_i and A with certainty if at least four legitimate nodes participate in STUN bootstrapping for every trusted node.

Proof. The colluding Type 2 adversary's underground node in an attempt to defeat Step 4 of STUN must overshadow all the transmissions from the legitimate nodes $\mathbf{L}_i$ in that sector. This is because if the adversary allows any one of the transmissions

from $\mathbf{L}_i$ to reach T_i, it will be the outlier and cause STUN to fail. Moreover, without a visual channel to the legitimate nodes and not knowing the location of the target legitimate node. M cannot just target one legitimate node. The adversary has to compute the transmit power to overshadow all the legitimate nodes $\mathbf{L}_i$ in a sector. Thus, according to MAD, it can pass the thresholds set at T_i. The underground node of the colluding Type 2 adversary has to transmit at least an underground power P_M^{tx} for each of $\mathbf{L}_i$ such that its power is within the acceptable range of the other legitimate underground nodes $\mathbf{L}_i$. Further, the adversary cannot just transmit a very high power as that will be detected by other trusted nodes . As the detection in Step 4 is performed by detecting the outlier, the transmit power has to satisfy (3.11).

Therefore, the adversary has to compute its transmit power according to the equation system S to overshadow all the underground nodes in a sector:

$$(S)\begin{cases} P_M^{tx} = P_{L_{i1}}^{tx} + G_{L_{i1}} - G_M + 20\log(d_{MT_i}^u) - 20\log(d_{L_{i1}T_i}^u) \\ \quad +20\log(\beta_M^u) - 20\log(\beta_{L_{i1}}^u) + \rho_u\alpha_M^u d_{MT_i}^u - \rho_u\alpha_{L_{i1}}^u d_{L_{i1}T_i}^u, \\ \vdots \\ P_M^{tx} = P_{L_{i|\mathcal{J}_i|}}^{tx} + G_{L_{i|\mathcal{J}_i|}} - G_M + 20\log(d_{MT_i}^u) - 20\log(d_{L_{i|\mathcal{J}_i|}T_i}^u) \\ \quad +20\log(\beta_M^u) - 20\log(\beta_{L_{i|\mathcal{J}_i|}}^u) + \rho_u\alpha_M^u d_{MT_i}^u - \rho_u\alpha_{L_{i|\mathcal{J}_i|}}^u d_{L_{i|\mathcal{J}_i|}T_i}^u, \end{cases} \tag{3.20}$$

The equation system S is an underdefined multivariate quadratic equation system. It is well known that the solution of such an equation system is NP-hard [83]. Moreover, even for small values of some equations (e), the best-known algorithms perform an exhaustive search [83]. Therefore, this type of system is known as an Unbalanced Oil and Vinegar (UOV) signature scheme. The security of UOV systems is proved for $3e \leq v \leq {}^{e(e+2)}/_2$ [83]. Where the system has e equations and v unknowns. Out of the total number of variables, e is known as the "oil" unknowns, and $(v-e)$ is called the "vinegar" unknowns. For our equation system S, we have seven variables in S which are known to underground transmit power of the legitimate node ($P_{L_{ij}}^{tx}$), gain of

antenna (G_x), operating frequency (f), and the relaxation introduced by the outlier evaluation technique (δ^a). Further, there are six variables in S which are unknown transmit power of the underground adversary (P_M^{tx}), distance from the adversary to the trusted node ($d_{MT_i}^a$), soil parameters (permittivity constants (ϵ^a, ϵ^u), relative permeability (μ_r)) as these control the variables ($\alpha_{L_{ij}}^u$, α_M^u, $\beta_{L_{ij}}^u$ and β_M^u), and distance between the legitimate node and trusted node $d_{L_{ij}T_i}^u$. Hence, S has $n+5$ number variables for n number of equations, where n is the number of legitimate node in our setup. Therefore, to satisfy the conditions for the UOV public cryptosystem, the minimum number of legitimate nodes required per trusted node is four. □

Hence, a Type 2 adversary cannot pass Step 4 with certainty. In addition, it is essential to note that in practice, the underground-to-underground wireless channel (e.g., M-T_i) has a limited communication range (i.e., less than 10m [152]). Hence, in most practical cases, an adversary outside the farm limits to pass Step 4 will be outside the communication range of T_i. Therefore, Type 2 attacks are practically rare and otherwise detectable, as per Proposition 3. This is even more challenging if the adversary wants to defeat multiple sectors because each sector contains at least four legitimate nodes per trusted node.

3.3.4 Security Against a Rogue Gateway

A colluding Type 2 adversary at a distance aboveground and underground outside the farm's perimeter injects its signal at the legitimate nodes $\mathbf{L}_i$ to emulate a rogue A's signal, as shown in Fig. 3.6. Without access to the secure channel between A-to-T_i, the adversary has to inject the signal such that it cannot be detected by corresponding T_i. Now we evaluate whether a Type 2 adversary can transmit underground power P_M^{tx} such that it is only received at $\mathbf{L}_i$, and not at the corresponding T_i.

Proposition 4. An aboveground and underground Type 2 colluding adversary M attempting to pair with underground nodes $\mathbf{L}_i$ as a rogue gateway can be detected with certainty when located at a distance.

$$20\log(d^u_{L_{ij}M}) + 20\log(\beta^u_{L_{ij}}) + \rho_u\alpha^u_{L_{ij}}d^u_{L_{ij}M} \geq 20\log(d^u_{L_{ij}T_i}) + 20\log(\beta^u_{T_i}) + \rho_u\alpha^u_{T_i}d^u_{L_{ij}T_i}. \tag{3.21}$$

where d^u_{xy} is the underground distance between entities x and y, and β^u_x and α^u_x are the attenuation and phase-shifting constants, respectively.

Proof. The underground node of the colluding Type 2 adversary has to transmit at an underground power P^{tx}_M to the legitimate nodes and must be at a distance that is less than the distance of the legitimate nodes L_{ij} and the trusted node T_i. The transmit power and the distance have to satisfy (3.25) and (3.21). In the best case, the adversary with the full knowledge of all the legitimate nodes can only be successful when transmitting in a location less than the distance between T_i and L_{ij}. Otherwise, the transmission will be overhead by T_i, and a failure message will be broadcast to other nodes. Note that the adversary does not know the exact location of T_i.

First, we compute the received power at L_{ij} from T_i, which is given by

$$P^{rx}_{L_{ij}T_i} = P^{tx}_{T_i} + G_{L_{ij}} + G_{T_i} - (C_u + 20\log(d^u_{L_{ij}T_i}) + 20\log(\beta^u_{T_i}) + \rho_u\alpha^u_{T_i}d^u_{L_{ij}T_i}), \tag{3.22}$$

Accordingly, we compute the received power at $L_{L_{ij}}$ from M, which is given by

$$P^{rx}_{L_{ij}M} = P^{tx}_M + G_M + G_{L_{ij}} - C_u - 20\log(d^u_{L_{ij}M}) - 20\log(\beta^u_{L_{ij}}) - \rho_u\alpha^u_{L_{ij}}d^u_{L_{ij}M}, \tag{3.23}$$

Since the underground power received $P^{rx}_{L_{ij}T_i} \geq P^{rx}_{L_{ij}M}$,

$$P^{tx}_{T_i} + G_{L_{ij}} + G_{T_i} - C_u - 20\log(d^u_{L_{ij}T_i}) - 20\log(\beta^u_{T_i}) - \rho_u \alpha^u_{T_i} d^u_{L_{ij}T_i} \geq$$
$$P^{tx}_{M} + G_{M} + G_{L_{ij}} - C_u - 20\log(d^u_{ML_{ij}}) - 20\log(\beta^u_{L_{ij}}) - \rho_u \alpha^u_{L_{ij}} d^u_{ML_{ij}}. \tag{3.24}$$

The adversary transmit power is provided by:

$$\begin{aligned} P^{tx}_{M} = P^{tx}_{T_i} + G_{T_i} - G_M + 20\log(d^u_{L_{ij}M}) + 20\log(\beta^u_{L_{ij}}) \\ - 20\log(d^u_{L_{ij}T_i}) - 20\log(\beta^u_{T_i}) + \rho_u \alpha^u_{L_{ij}} d^u_{L_{ij}M} - \rho_u \alpha^u_{T_i} d^u_{L_{ij}T_i}, \end{aligned} \tag{3.25}$$

From (3.24), with the assumption that the adversary has near perfect estimates of the transmits power and gains, which means that $P^{tx}_{T_i} + G_{T_i} - P^{tx}_{M} - G_M \approx \xi$, therefore it must be at a distance given by

$$20\log(d^u_{L_{ij}M}) + 20\log(\beta^u_{L_{ij}}) + \rho_u \alpha^u_{L_{ij}} d^u_{L_{ij}M} \geq 20\log(d^u_{L_{ij}T_i}) + 20\log(\beta^u_{T_i}) + \rho_u \alpha^u_{T_i} d^u_{L_{ij}T_i}. \tag{3.26}$$

where d^u_{xy} is the underground distance between x and y, β^u_x and α^u_x are the attenuation and phase-shifting constants, respectively.

□

The inequality in Proposition 4 can be simplified using the Newton-Raphson approximation [196], as illustrated in Fig. 3.7. Achieving convergence requires careful selection of parameters like tolerance (tol), maximum iterations (N_{max}), and the initial guess (x_0). For example, using $tol = 10^{-6}$, $N_{max} = 1000$, and $x_0 = 1$ fails to converge, even with known parameters like $\alpha^u_{L_{ij}}, \alpha^u_{T_i}, \beta^u_{L_{ij}}, \beta^u_{T_i}$, and $d^u_{L_{ij}T_i}$. While higher values can expedite convergence, they risk reducing accuracy. Thus, achieving convergence

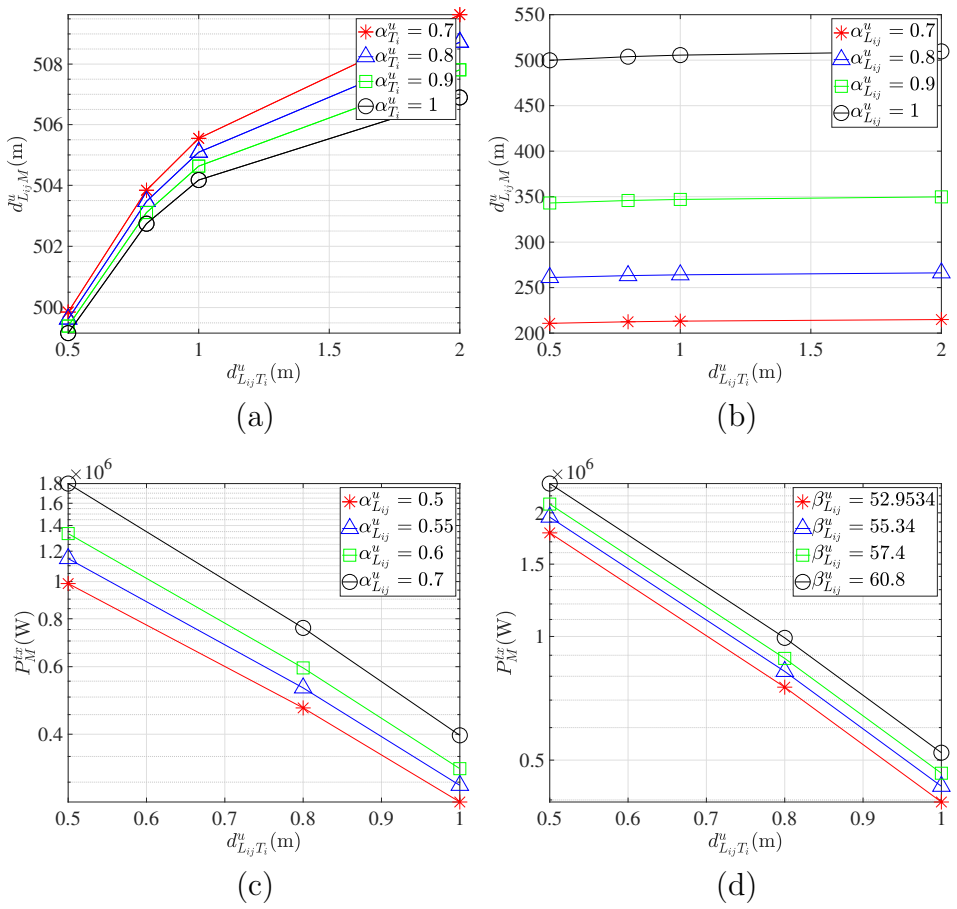


Figure 3.7: Plots of Type 2 M distance against the legitimate distance while varying (a) the legitimate node attenuation, and (b) the trusted node attenuation to determine the distance of the adversary from the legitimate nodes. Plots of the power transmitted by the underground adversary to the legitimate nodes against the legitimate underground distance while varying (c) the attenuation $\alpha^u_{L_{ij}}$, and (d) the phase shift constant to determine the power of M from $\mathbf{L}_i$.

often involves iterative optimization of these values. Therefore, an adversary must know the trusted node's location to perform this attack.

From Fig. 3.7(a)-(d), even with legitimate and trusted nodes separated by [0.5–2]m and varying attenuation and phase shift constants, the adversary must transmit enormous power from a farther distance. To evade detection, a Type 2 adversary must transmit approximately 10^6 W from ≈ 0.5 km away, which is impractical. These find-

ings confirm that a Type 2 adversary can only emulate a rogue gateway with precise knowledge of underground node locations.

3.4 Performance Evaluations

We evaluate STUN using experimental data from our wireless underground outdoor experiments in [52, 167, 168, 202]. First, we describe the experimental configuration and evaluate STUN's correctness and robustness.

3.4.1 Setup

Underground Testbed: We use data from outdoor experiments in [52, 167, 168, 202] under varying conditions of distances. In these outdoor experiments, a 433MHz dipole antenna and an underground wideband planar antenna [180] are utilized. Wideband planar antennas have been experimentally shown to be more suitable for underground communication, enhancing the communication range and increasing the channel link budget [180]. Additionally, a dipole antenna was employed for underground transmission when the aboveground node distance was fixed. We kept an underground depth of $d^u_{L_{ij}}$, and $d^u_{T_i}$ fixed at 0.2m, with a maximum distance of $d^a_{AT_i} = 115$m. The distance between the underground nodes and the trusted device was kept constant at $d^u_{L_{ij}T_i} = 2$m. The aboveground height was kept at 1.78m and the aboveground distances of the gateway $d^a_{AT_i}$ were varied at different intervals of [30, 60, 80, and 110m]. The outdoor testbed contains 13.09cm of silt clay loam soil with varying volumetric water content (VWC) ranging from 17% to 37% dry and wet. The VWC of the soil was kept at 37% for the most part. The transmit power ranges between [0.0032W - 0.2W] with a gain of 0.02W with a dipole antenna for the underground nodes. $d^a_{AT_i}$ was shown in [202] to cover a distance as far as 115m when a long-range device is

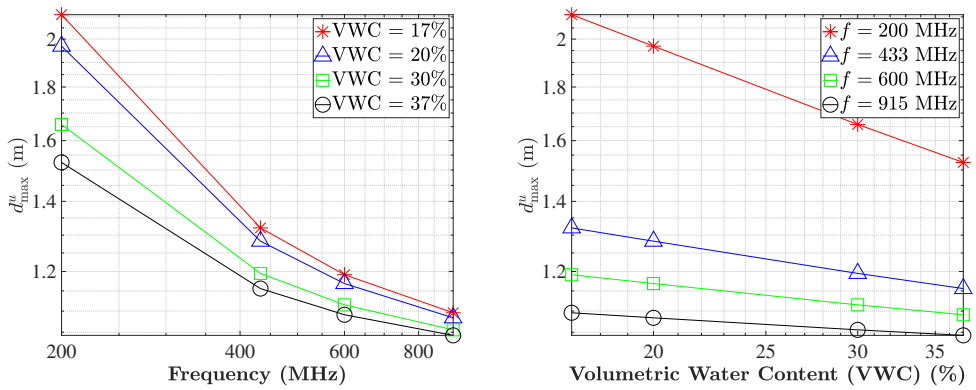


Figure 3.8: Plots of the maximum size of T_i's sector against (a) the frequency while varying volumetric water content of soil, and (b) the volumetric water content while varying the frequency to determine the placement of the legitimate nodes.

used and move up to 200m [182]. The $PL_{\rho,a-u}$ was observed to be 4.2063×10^{-9}W, equivalent to -53dBm.

3.4.2 Correctness Evaluations

First, we evaluate the maximum range of a trusted node, T_i, to assess the area of the sector for placement of legitimate nodes, $\mathbf{L}_i$. We use the information from the underground path loss to estimate the sector size of T_i in (3.4), giving us the maximum separation between the trusted node and the legitimate nodes underground within each sector. We show in Fig. 3.8(a) the maximum sector size of a trusted node against the frequency for various volumetric water content. We can deduce that the maximum separation in each sector is approximately 2.1m. Similarly, we show in Fig. 3.8(b) the maximum sector size against the soil moisture content for various center frequencies. We observe that the underground separation decreases as the water content in the soil increases. The maximum range of T_i depends on the operational frequency since lower frequency antennas can cover more distances. When placing nodes underground, the

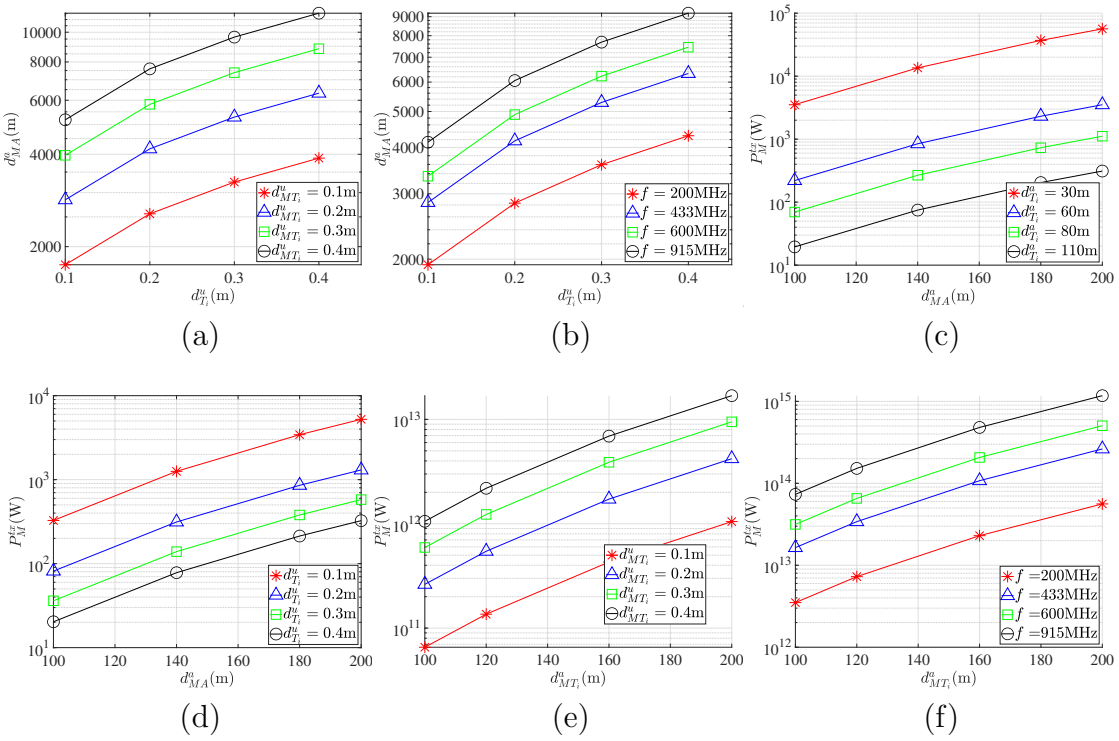


Figure 3.9: Plots of (a) the distance between a Type 1 M from A against the underground depth of T_i, (b) the distance between a Type 1 M from A against the underground depth of T_i, (c) the required power transmitted by the Type 1 M to defeat Step 6 of STUN, (d) the required power transmitted by the Type 1 M to defeat Step 6 of STUN, (e) the required power transmitted by the Type 1 M to defeat Step 4 of STUN, (f) the required power transmitted by the Type 1 M to defeat Step 4 of STUN.

separation of the nodes should be within 2.1m, and the soil moisture content influences the maximum range between the legitimate node and the trusted node.

3.4.3 Robustness Evaluation

In this section, we evaluate the robustness of STUN against Type 1 and Type 2 adversaries.

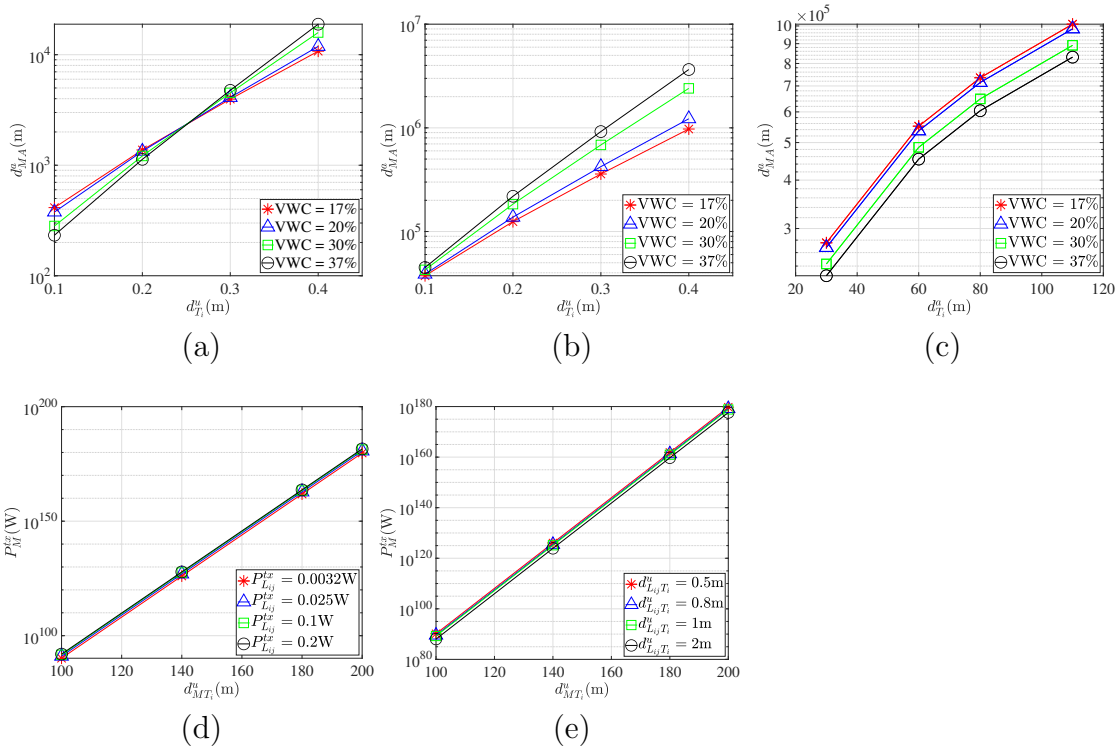


Figure 3.10: Plots of (a) the distance between a Type 1 M from A against the underground depth of T_i at different volumetric water content for $d_{MT_i}^{u} = 0.2$m, (b) the distance between a Type 1 M from A against the aboveground distance of A at different volumetric water content for $d_{MT_i}^{u} = 0.4$m, (c) the required power transmitted by the Type 1 M to defeat Step 4 of STUN at different volumetric water content, (d) the required power transmitted by the Type 2 M against the distance between M and T_i, and (e) the required power transmitted by the Type 2 M against the distance between M and T_i.

3.4.3.1 Type 1 Adversary

To evaluate the capabilities of an adversary, we emulate the adversarial data utilizing experimentally obtained wireless channel parameters, precisely measuring the effective soil permittivity and relative permeability of the soil from the experimental setup described in Section 5.5. The underground wireless channel is relatively stable and mostly deterministic compared to the OTA wireless channel [168].

In Fig. 3.9(a), we illustrate how the gateway-to-adversary distance changes with

the underground depth of the trusted node ($d^u_{T_i}$), considering various underground adversary distances ($d^u_{MT_i}$). Similarly, in Fig. 3.9(b), we vary the center frequencies and analyze the impact on the adversary distance. Our results show that M must position itself between 2000m and 10000m from the gateway to achieve an equivalent path loss in underground-to-OTA while located above ground.

In Fig. 3.9(c) and (d), we examine how successful an aboveground adversary can be in passing the verification at the gateway. We observed that the transmitter power is above the threshold at the gateway for various underground ($d^u_{T_i}$) and aboveground ($d^a_{AT_i}$) distances. Even when the adversary operates with the experimentally measured acceptable distance, our results reveal that the required power remains more than the threshold needed to pass Step 6. The adversary must transmit at a precise power level, maintaining an RSS ratio between 1 and 1.5. Adversary can compute the needed power for the gateway due to positional knowledge of the gateway aboveground but needs to do the same for the trusted node underground whose location is unknown.

Similarly, in Fig. 3.9(e), and (f), we show that the power transmitted from aboveground Type 1 adversary attempting to pass Step 4 verification, even when the adversary distance ($d^a_{MT_i}$) from the trusted node is within the communication range. Our results show that the adversary transmit power to the trusted node is very high while changing the underground depth, and center frequency. Our findings indicate that the adversary needs to transmit a precise power to maintain an RSS threshold between 2.512×10^{-7}mW and 6.309×10^{-7}mW for verification at the trusted node. The likelihood of a successful attack is low, given that the adversary typically lacks exact positional knowledge of the trusted nodes and does not possess physical access to the farm.

In Fig. 3.10(a), (b), we demonstrate how aboveground gateway-to-adversary distance changes with underground depth ($d^u_{T_i}$), while varying the volumetric water con-

tent (VWC). We observe that in Fig. 3.10(a) and (b), while we keep the $d^u_{MT_i}$ at 0.2m and 0.4m. We observe that water content alters the adversary's position, shifting the distance from the underground node to soil the surface. Additionally, Fig. 3.10(c) analyzes how the adversary's distance changes with the aboveground gateway distance ($d^a_{AT_i}$) while varying the VWC. Our results confirm that in Fig. 3.10(a), (b), (c), the adversary must operate at considerably large distances. M cannot reduce the transmit power rather than increase the distance, as this will cause the adversary to fail Step 4 due to high attenuation for OTA-to-underground transmissions.

3.4.3.2 Type 2 Adversary

In our evaluation, we assess the transmission power of an underground adversary to the trusted nodes as the distance between them increases. Our findings indicate that the adversary has to transmit an excessive amount of power. In contrast, the legitimate nodes maintain a consistent transmit power of [0.0032W, 0.025W, 0.1W, 0.2W], while the distance between the legitimate and trusted nodes remains at [0.5m, 0.8m, 1m, 2m]. The adversary succeeds if the received power of the trusted node is between $\tau^{T_i}_{low} = 2.512 \times 10^{-7}$mW and $\tau^{T_i}_{high} = 6.309 \times 10^{-7}$mW as obtained in Section 3.2.2. We emulate the adversary's path loss according to (3.3) with the soil parameters obtained from the testbed. We ran the experiment 10,000 times. Fig. 3.10(d) and (e) depict the underground power transmitted by the Type 2 adversary (P^{tx}_M) against the underground distance between the adversary and the trusted node ($d^u_{MT_i}$), with variations in the power transmitted by the legitimate nodes ($P^{tx}_{L_{ij}}$) and the underground distance between the trusted node and the legitimate node ($d^u_{L_{ij}T_i}$). Our observations reveal that the adversary's power, as illustrated in Fig. 3.10(d) and (e), is sufficient to cause damage to the sensors. Furthermore, as the adversary's underground distance from the farm's perimeter increases, it becomes increasingly challenging to counter

step 4 due to changes in soil conditions securely.

3.5 Chapter Summary

We address the problem of a secret-free secure bootstrapping for COTS underground nodes with an aboveground gateway. We propose STUN, which uses hard-to-forge underground wireless propagation laws to achieve node authentication and secret establishment in-band with the help of a trusted underground node. We demonstrate that STUN resists active signal injection attacks and scales well as the number of underground nodes increases. In addition, We theoretically prove that STUN has a security equivalent to the UOV scheme in public cryptography. We also validate our theoretical results with outdoor underground wireless testbed experiments. We evaluate the placements of the trusted nodes to cover an agricultural farm in a hexagonal sector format. To optimize the required number of trusted nodes, we investigated the distance of the trusted nodes from the farm boundary and the distance between each other. Finally, we developed security for the downlink communication and multiple trusted and legitimate nodes.

CHAPTER 4

RF Fingerprint-Based Location Authentication for Over-The-Air and Underground Wireless Networks

This chapter is based on joint work with Mr. Hakim Lado, Dr. Nirnimesh Ghose, Dr. Boyang Wang, and Dr. Mehmet Can Vuran. While the work reflects collaboration across team members, I led the conceptual design, data collection, modeling, and experimental evaluation of the RF fingerprinting system. My contribution to this work represents approximately 70% of the effort.

This chapter explores the use of wireless channel characteristics as unique fingerprints to verify the location of transmitting devices. Instead of relying on cryptographic secrets or trusted hardware, we utilize the channel impulse response (CIR) as a discriminative indicator of location. CIR encodes the multipath structure between a transmitter and receiver and reflects the surrounding physical environment, making it inherently difficult to forge or emulate from a different location. This enables robust authentication even in the presence of adversaries attempting to spoof legitimate transmissions. To ground our method in practical constraints, we first outline the system and adversarial models, followed by a description of how CIR is computed from pilot symbols. This foundation supports the development of a generalizable, authentication pipeline resilient to our attack settings.

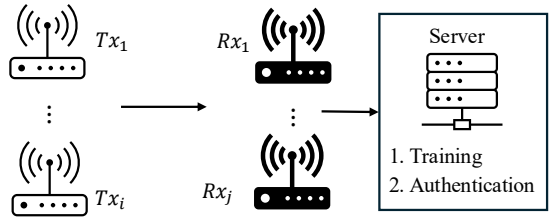


Figure 4.1: System model showing transmitters Tx_i sending wireless signals to the receivers Rx_j, which extract CIRs. Labeled CIRs from authorized location zones are used for training and authentication at the central server.

4.1 System Overview

4.1.1 Notations

We preprocess raw RF signals received at the receiver to extract the complex baseband I/Q samples. These samples are segmented into fixed-length traces, which serve as input to our location authentication system.

A single trace, representing a short sequence of complex I/Q samples, is denoted as $\mathbf{x} = [x_1, x_2, \ldots, x_L]$, where $x_k = x_k^{\mathrm{R}} + jx_k^{\mathrm{I}}$ and L is the trace length. We fix $L = 288$ in our experiments. The set of all traces is $\mathcal{X} = \{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}\}$, with each trace $\mathbf{x}^{(i)}$ associated with a ground truth label $\ell^{(i)} \in \mathfrak{L}$, where $\mathfrak{L}$ is the set of authorized location zones. The number of unique locations is $M = |\mathfrak{L}|$.

4.1.2 System Model

The system consists of a set of stationary or mobile transmitters $\mathbf{T}$ and stationary receivers $\mathbf{R}$. Let $\mathbf{T} = \{\mathrm{Tx}_1, \mathrm{Tx}_2, \ldots, \mathrm{Tx}_i\}$ denote the set of wireless transmitters, and let $\mathbf{R} = \{\mathrm{Rx}_1, \mathrm{Rx}_2, \ldots, \mathrm{Rx}_j\}$ denote the set of stationary receivers. Each transmitter $\mathrm{Tx}_i \in \mathbf{T}$ periodically sends known pilot signals, and each receiver $\mathrm{Rx}_j \in \mathbf{R}$ computes the channel impulse response (CIR) from the received baseband I/Q samples. The CIR captures the spatial and temporal characteristics of the wireless channel.

The system operates in two phases: training and authentication. During training,

Table 4.1: Summary of Notations

Symbol	Description
$\mathbf{x}^{(i)}$	i-th trace containing complex I/Q samples
$x_k^{\mathrm{R}}, x_k^{\mathrm{I}}$	Real and imaginary parts of the k-th complex sample
L	Trace length (set to 288)
w	Sliding window stride (set to 288)
M	Total number of I/Q samples in a sequence
$\mathcal{X}$	Set of all traces used for training/testing
$\ell^{(i)}$	Ground truth label (location) for trace $\mathbf{x}^{(i)}$
$\mathfrak{L}$	Set of all candidate location labels
$\mathcal{L}_{\text{loss}}$	Loss function used during training
$h_j[n]$	Time-domain CIR at location L_j
$\hat{h}_j[n]$	Estimated time-domain CIR at location L_j
$\hat{h}_j$	Estimated CIR vector via least squares
$H_j[f]$	Intermediate frequency-domain channel response at L_j
$\hat{H}_j[f]$	Estimated frequency-domain channel response at L_j
$\mathbf{Y}_{L_j}$	Received signal vector at L_j in matrix CIR model
$\mathbf{P}$	Toeplitz matrix constructed from pilot symbols
$\mathbf{W}$	Additive noise vector in matrix CIR model
$P[n]$	Known pilot symbols (time domain)
$P[f]$	Fourier transform of pilot symbols
$w[n]$	Additive noise in time domain
$Y_{L_j}[f]$	Received signal in frequency domain at location L_j
A_{filt}	Filtered amplitude component of CSI
Φ_{filt}	Filtered phase component of CSI
$P(\tau)$	Power Delay Profile (PDP)
$\bar{\tau}$	Mean excess delay
τ_{rms}	RMS delay spread
B_c	Coherence bandwidth, $B_c \approx 1/\tau_{\text{rms}}$
$\pi_{i,j}$	Predicted confidence for location ℓ_j on trace $\mathbf{x}^{(i)}$
$\mathcal{F}$	Classifier model
$\mathcal{D}_{\text{train}}, \mathcal{D}_{\text{test}}$	Training and test datasets
r_{avg}	Average rank of the true location across predictions
m'	Number of correctly classified test traces
$\theta_f, \theta_y, \theta_d$	ADA parameters (feature extractor, classifier, discriminator)
λ	ADA weighting factor
ϵ	Adversarial perturbation budget
α	Step size for adversarial attacks
$h_{\text{UG}}(t)$	Continuous-time underground CIR

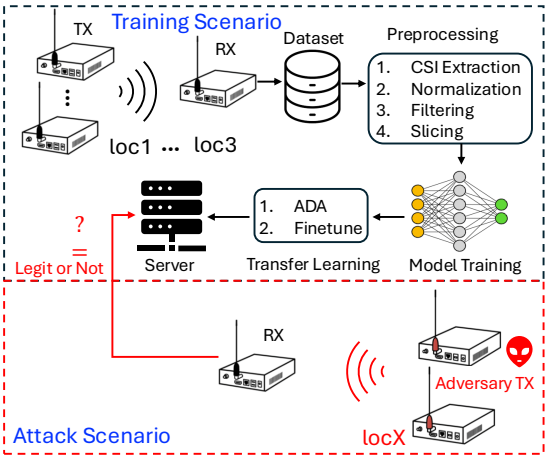


Figure 4.2: System overview showing the preprocessing, training, and transfer learning using fine-tuning and adversarial domain adaptation for authorized locations, followed by the adversarial attack from location X.

transmitters send pilot signals from known authorized location zones, and the receiver extracts and labels the corresponding CIR traces. These labeled traces are used to train a multi-class classifier, where each class represents a distinct location zone.

During authentication, the receiver extracts the CIR from new transmissions and queries the trained classifier. The system, as illustrated in fig 4.1, verifies whether the predicted location matches an authorized zone with high confidence. The architecture is agnostic to communication protocol, device type, and deployment environment, enabling flexible and scalable deployment.

4.1.3 Threat Model

We assume that the training phase occurs in a trustworthy environment with only legitimate transmitters operating from authorized zones. The adversary attempts to spoof location in the authentication phase by injecting CIR-like signals to mislead the classifier.

We define two attacker models:

4.1.3.1 Friis' Empirical Adversary:

Knows only device distances and estimates CIR using Friis' equation, ignoring multipath and noise effects.

4.1.3.2 Ray-Tracing Enhanced Adversary:

Uses advanced ray-tracing to simulate CIR with reflection, scattering, shadowing, and delay clusters. It attempts to closely match legitimate CIR patterns, representing a worst-case channel-aware attacker.

Our system demonstrates robustness against both attacker types across various environments. More details are provided in Section 4.5.

4.1.4 Motivation for Deep Learning

Traditional RF fingerprinting methods relying on RSS or CSI perform well under controlled settings but degrade in dynamic or underground environments due to temporal drift, noise, and multipath.

Even after preprocessing such as Butterworth filtering or denoising autoencoders, location-invariant distortions remain. Deep CNNs offer improved robustness by learning complex spatio-temporal features and effectively distinguishing intra- versus inter-location variation.

Each input trace $\mathbf{x}^{(i)} = [x_1^{(i)}, \ldots, x_L^{(i)}]$ is passed through a CNN to extract high-level features corresponding to coherence bandwidth, RMS delay, and PDP variations—features strongly tied to location. CNNs also adapt well to noise and domain shifts across OTA and underground conditions. The general overview our authentication system is shown in fig. 4.2

4.1.5 Impact of Hardware Imperfections

Even devices built from the same specification exhibit hardware-induced nonlinearities. These include phase noise, amplifier distortion, and oscillator drift, which can introduce either beneficial or harmful bias in the signal.

Such imperfections may increase separability between devices but reduce the fidelity of spatially induced channel effects. Hence, it is essential that the model learn to filter out hardware-specific noise while preserving channel-specific spatial features for reliable location authentication.

4.1.6 Impact of Channel Variability

The wireless channel is highly sensitive to environmental factors like multipath, soil properties, object motion, and weather, which directly impact CIR and its reliability as a location identifier.

To maintain authentication reliability, we combine spatial filtering, denoising, and domain adaptation techniques to reduce noise and emphasize persistent location-specific features. This ensures model resilience even when the channel conditions change.

4.1.7 Natural vs. Adversarial Channel Variations

4.1.7.1 Natural Variations:

OTA channels experience rich multipath and high temporal variation from motion and interference. Underground channels have higher spatial resolution but suffer from soil-induced drift. OTA scenarios require frequent recalibration, while underground setups demand robust handling of slow, non-stationary drift.

4.1.7.2 Adversarial Variations:

OTA adversaries may attempt replay or modulation spoofing, but fail to mimic CIR from a different location. Underground attackers cannot replicate soil conditions or air-soil boundary distortions, leading to detectable deviations.

Underground channels offer higher spoofing resistance due to environmental constraints, while OTA channels provide diverse propagation paths requiring stronger generalization from models.

4.2 Location Fingerprinting Architecture

Table 4.2: Complete Model Performance Analysis for Location Authentication.

Model	Best Performance	Stability	Reliability
ResNet-50	85–95%	High	Excellent across all scenarios
ResNet-34	80–92%	High	Very Reliable, best overall
ResNet-18	75–90%	High	Very Reliable
In-Lab Model	70–85%	Moderate to High	Reliable in controlled settings
GoogleNet	60–70%	Moderate	Reasonably Reliable in Filtered settings
VGG16	50–60%	Low	Inconsistent across device/distance
VGG19	∼33% (Random Baseline)	Very Low	Unreliable, fails to generalize

4.2.1 Convolutional Neural Network Architectures

We adopt multiple convolutional neural network architectures to extract spatio-temporal patterns in the filtered CIR traces for location-based authentication. These models are designed to provide robust authentication performance across over-the-air and underground scenarios in both indoor and outdoor environments by learning to distinguish between authorized location zones.

Our primary architectures include ResNet-18, ResNet-34, and ResNet-50 deep residual networks, which employ identity skip connections to mitigate vanishing gra-

dients. ResNet-18 offers a compact structure with fewer parameters, balancing computational efficiency and representational power, making it suitable for resource-constrained authentication systems. ResNet-34 and ResNet-50 extend this capability with deeper stacks of convolutional layers, incorporating bottleneck blocks to increase architectural depth while maintaining computational tractability. This facilitates deeper feature extraction for authentication decisions, enabling better discrimination between closely spaced authorized locations and capturing subtle variations in channel characteristics crucial for reliable authentication.

We selected ResNet architectures as our primary models based on comprehensive comparative analysis with alternative CNN architectures, including VGG16, VGG19, and GoogleNet, as detailed in Table 4.2. ResNet models demonstrate superior performance with 75–95% accuracy and high stability compared to alternatives that suffer from fundamental architectural limitations. VGG19 fails completely, achieving only random baseline performance of approximately 33% due to severe vanishing gradient problems in its deep architecture without skip connections. VGG16 exhibits poor stability with 50–60% accuracy and inconsistent performance across different transmission scenarios. GoogleNet provides moderate performance ranging from 60–70% but lacks the consistent reliability required for critical authentication applications. The critical advantage of ResNet architectures lies in their identity skip connections, which effectively mitigate gradient degradation issues that severely limit deeper networks without residual connections.

In addition, we designed a custom 5-layer CNN, referred to as the in-lab model, to provide baseline comparison against deeper architectures. This model comprises three convolutional layers with ReLU activations, a pooling layer, a fully connected layer, a softmax classifier, and two additional fully connected layers for location classification. Despite its simplicity, the in-lab model achieves 70–85% authentication

accuracy, demonstrating that effective location authentication can be achieved with properly designed simpler architectures. This validates that ResNet's superior performance stems from architectural advantages rather than mere parameter count, while providing a computationally efficient alternative for resource-constrained deployments.

These selected models offer a synergistic combination of depth, efficiency, and resilience, facilitating robust temporal and spatial authentication of location zones across diverse indoor and outdoor environments. The detailed performance evaluation of these models for location authentication is presented in Section 4.3.

4.2.2 Location Authentication Dataset

We consider a dataset that includes wireless transmissions in two environmental settings: (i) indoor and (ii) outdoor. The indoor environmental settings comprise wireless transmissions for OTA communication, demonstrating the feasibility of our authentication approach across diverse propagation environments. The outdoor environmental settings consist of CIR data samples with a sample rate of 6.4 MSps collected for OTA communication, enabling evaluation of authentication robustness under challenging environmental conditions.

Both communication scenarios utilize USRP SDR B200 and B205mini devices operating at a 2.4 GHz center frequency that act as transmitters and receivers. We utilize 1 receiver device and 4 transmitter devices positioned across 3, 4, 5, and 6 feet for the distance-based authentication evaluation. We use 6 transmitters and 6 receivers for the device-based authentication assessment while maintaining the same distance and changing each device to evaluate robustness against hardware variations.

Each transmitted data is collected for 2 minutes per transmission, and we move the devices across different distances to capture spatial variations in channel character-

istics. We have a variable-length vector of two dimensions representing the in-phase and quadrature transmissions. We repeat this procedure across three indoor and outdoor locations, creating authorized location zones for authentication evaluation. The experimental setup for data collection is shown in Figures 4.10 and 4.11 for outdoor and indoor environments, respectively. We refer to the data samples as transmissions for convenience. We save the IQ samples as binary files and perform the same experiment for different locations in both outdoor and indoor settings. Additional details regarding the dataset statistics are provided in Section 4.3.

4.2.3 Data Preprocessing

Once we capture the received signal, we perform various data preprocessing steps: extracting the channel state information, data normalization, and filtering. These are applied to both underground and OTA transmissions, indoors and outdoors. This preprocessing step removes noisy components that could compromise authentication reliability while retaining the essential location-specific information needed for accurate authentication decisions.

4.2.3.1 Data Normalization

We normalize our dataset to improve the numerical stability and convergence speed during training of our authentication models. Data normalization guarantees that the input data is uniformly scaled, eliminating anomalies that could adversely impact our machine learning models' capacity to learn temporal and spatial features crucial for reliable location authentication. Our procedure entails normalizing the input data to handle zeros and NaN values, which might affect the learning process and improve the robustness and precision of the authentication system in both indoor and outdoor environments.

4.2.3.2 Extracting The Channel Impulse Response

The Channel Impulse Response (CIR) serves as a critical characterization of location-specific wireless propagation signatures, essential for robust authentication mechanisms. This work extracts the CIR using pilot symbol-based channel estimation in the time domain, a method that effectively preserves the temporal multipath characteristics of the wireless channel.

For a transmitter communicating with receivers at locations L_j, the discrete-time received signal, $y_{L_j}[n]$, can be fundamentally modeled as the convolution of the transmitted signal $x_T[k]$, additive noise $w[n]$, and the location-dependent CIR, $h_j[n-k]$:

$$y_{L_j}[n] = \sum_{k=-\infty}^{\infty} x_T[k] \cdot h_j[n-k] + w[n] \tag{4.1}$$

When known pilot symbols $P[n]$ are transmitted, the received signal $y_{L_j}[n]$:

$$y_{L_j}[n] = \sum_{k=-\infty}^{\infty} P[k] \cdot h_j[n-k] + w[n] \tag{4.2}$$

To derive the channel frequency response, a Fourier transform is applied to both sides of the equation, yielding:

$$Y_{L_j} = P \cdot H_j + W \tag{4.3}$$

The channel frequency response, $\hat{H}_j$, is then obtained by solving for H_j:

$$\hat{H}_j = \frac{Y_{L_j}}{P} \tag{4.4}$$

Finally, the time-domain CIR, $\hat{h}_j[n]$, is recovered by applying the inverse Fourier transform:

$$\hat{h}_j[n] = \mathcal{F}^{-1}\left\{\frac{Y_{L_j}}{P}\right\} \tag{4.5}$$

From Equation (4.1), the CIR can also be written as

$$\hat{h}_j[n] = \mathcal{F}^{-1}\left\{\frac{Y_{L_j}}{X}\right\} \tag{4.6}$$

Where X is the transmitted signal envelope. This process effectively extracts location-specific signatures, encompassing multipath delays, amplitude variations, and delay spread patterns. The preservation of these temporal propagation characteristics within the time-domain CIR creates unique spatial fingerprints adversaries find difficult to replicate, thus enabling robust location-based authentication across diverse wireless environments.

Matrix Representation for Computational Modeling For computational modeling and algorithmic implementation, the convolution operation described above can be efficiently represented in matrix form. Considering finite-length signals, the convolution of the pilot symbols $P[n]$ and the CIR $h_j[n]$ to produce the received signal $y_{L_j}[n]$ can be expressed as a matrix-vector product.

Let P be a vector of N_p pilot symbols, and h_j be a vector of N_h CIR coefficients. The received signal $\mathbf{Y}_{L_j}$ (a vector of length $N_p + N_h - 1$) can be expressed as:

$$\mathbf{Y}_{L_j} = \mathbf{P}h_j + \mathbf{W} \tag{4.7}$$

Here, $\mathbf{Y}_{L_j} \in \mathbb{C}^{(N_p+N_h-1)\times 1}$ represents the received signal vector, $\mathbf{h}_j \in \mathbb{C}^{N_h\times 1}$ is the CIR vector, $\mathbf{W}$ is the additive noise vector, and $\mathbf{P}$ is a Toeplitz matrix constructed from the pilot symbols $P[n]$. This Toeplitz matrix is formed by shifting

the pilot symbol vector in successive rows, effectively implementing the convolution operation as a linear transformation. For instance, if $P = [P_0, P_1, \ldots, P_{N_p-1}]^T$ and $h_j = [h_{j,0}, h_{j,1}, \ldots, h_{j,N_h-1}]^T$, the Toeplitz matrix $\mathbf{P}$ would have dimensions $(N_p + N_h - 1) \times N_h$, where each column is a shifted version of the pilot signal vector.

The matrix $\mathbf{P}$ can be visualized as:

$$\mathbf{P} = \begin{pmatrix} P_0 & 0 & 0 & \cdots & 0 \\ P_1 & P_0 & 0 & \cdots & 0 \\ P_2 & P_1 & P_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{N_p-1} & P_{N_p-2} & P_{N_p-3} & \cdots & P_{N_p-N_h} \\ 0 & P_{N_p-1} & P_{N_p-2} & \cdots & P_{N_p-N_h+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & P_{N_p-1} \end{pmatrix}$$

In practical channel estimation, particularly when pilot symbols are transmitted, the CIR h_j can be estimated using techniques such as least squares:

$$\hat{h}_j = (\mathbf{P}^H \mathbf{P})^{-1} \mathbf{P}^H \mathbf{Y}_{L_j} \tag{4.8}$$

where $\mathbf{P}^H$ denotes the conjugate transpose of $\mathbf{P}$. Note that Toeplitz systems can be solved efficiently using Fast Fourier Transform (FFT) algorithms, reducing computational complexity from $O(N^3)$ to $O(N \log N)$ for large pilot sequences.

This matrix formulation not only simplifies CIR estimation computationally, but also provides a structured input format ideal for deep learning models trained on

location-specific channel characteristics.

4.2.4 Filtering Process for Authentication Enhancement

We evaluate multiple filtering approaches to extract stable location-specific signatures while removing noise and hardware imperfections that could compromise authentication reliability. Our comparative analysis considers several filtering techniques including Butterworth low-pass filtering [33,199], elliptic filtering [128], moving average filtering [171], and Denoising Autoencoder (DAE) approaches [179, 198].

After extensive evaluation, we select Butterworth filtering and DAE techniques based on their superior performance in preserving authentication-relevant channel characteristics while providing effective noise reduction. Butterworth filters were chosen for their maximally flat frequency response in the passband and excellent phase linearity, which preserves the temporal structure crucial for CIR-based authentication. Alternative filters such as elliptic and moving average filters were considered but rejected due to their inferior performance in maintaining location-specific signatures.

Elliptic filtering was evaluated but showed suboptimal results due to its ripple characteristics in both passband and stopband, which could distort the fine-grained location-specific features essential for authentication. Moving average filtering, while computationally simple, demonstrated poor frequency selectivity and phase response, making it unsuitable for preserving the complex spectral characteristics of CIR data required for reliable location authentication.

The quantitative results presented in Table 4.3 demonstrate the clear superiority of Butterworth filtering over the rejected alternatives. Moving average filtering achieved only moderate performance (50-60%) with very poor stability, while elliptic filtering, despite reaching up to 70% performance in some cases, exhibited extremely poor stability that rendered it completely unreliable for authentication applications.

Table 4.3: Comparative Performance Analysis of Filtering Methods for Location Authentication.

Filtering Method	Best Performance	Stability	Reliability
Butterworth	80–90%+	High	Excellent
Moving Average	∼50–60%	Very Poor	Unreliable
Elliptic	∼60–70%	Extremely Poor	Completely Unreliable

4.2.4.1 Butterworth Filtering of CIR

We apply Butterworth low-pass filtering separately to the amplitude and phase components of the estimated time-domain CIR, $\hat{h}_j[n]$, to enhance location-specific features and suppress noise. This two-part filtering process improves fingerprinting accuracy by reducing small-scale fading and hardware-induced fluctuations while preserving large-scale propagation effects.

Amplitude Filtering. The amplitude-based filter smooths rapid magnitude fluctuations in the CIR while preserving features critical for location authentication. This is particularly effective in controlled indoor environments where signal variations are more stable. We apply the Butterworth transfer function to the amplitude component of the CIR:

$$\hat{h}_{\text{filt}}[n] = G(z) \cdot \hat{h}_j[n] \tag{4.9}$$

where $G(z)$ is the low-pass Butterworth filter defined as:

$$G(z) = \frac{B(z)}{A(z)} = \frac{b_1 + b_2 z^{-1} + \cdots + b_{n+1} z^{-n}}{1 + a_2 z^{-1} + \cdots + a_{n+1} z^{-n}} \tag{4.10}$$

with **b** and **a** as filter coefficient vectors. We apply zero-phase filtering (forward and reverse directions) to avoid introducing phase distortion. The filtered amplitude is denoted as $A_{\text{filt}}[n] = |\hat{h}_{\text{filt}}[n]|$.

Phase Filtering. To reduce phase noise and hardware-induced artifacts, we extract the phase of the CIR, $\angle\hat{h}_j[n]$, and apply the same Butterworth filter. This isolates the large-scale phase behavior from small-scale fluctuations. The filtered phase component is denoted as $\Phi_{\text{filt}}[n] = \angle\hat{h}_{\text{filt}}[n]$. Zero-phase filtering ensures phase symmetry is preserved and avoids time distortion.

Recombination. After applying both filters, we reconstruct the filtered CIR by recombining the filtered amplitude and phase:

$$\tilde{h}_j[n] = A_{\text{filt}}[n] \cdot e^{j\Phi_{\text{filt}}[n]} \tag{4.11}$$

This final filtered CIR, $\tilde{h}_j[n]$, enhances the stability and discriminative power of location-specific features for fingerprinting under various propagation conditions.

4.2.4.2 Denoising Autoencoder (DAE)

In addition to traditional signal processing filters, we leverage DAEs for adaptive noise reduction, particularly in complex and noisy environments where their learned representation offers superior performance. A DAE is a type of artificial neural network designed to learn a robust representation of input data by attempting to reconstruct a clean output from a corrupted, or noisy, input [179]. In principle, a DAE consists of two main parts: an Encoder which maps the noisy input $\mathbf{x}_{\text{noisy}}$ to a lower-dimensional latent representation $\mathbf{z}$, such that $\mathbf{z} = f(\mathbf{x}_{\text{noisy}})$; and a Decoder which reconstructs the original, clean input $\mathbf{x}_{\text{clean}}$ from the latent representation, such that $\mathbf{x}_{\text{reconstructed}} = g(\mathbf{z})$. During training, the DAE is fed noisy versions of the CIR data and is optimized to minimize the difference between its reconstructed output and the original, uncorrupted CIR signal. This forces the network to learn to identify and

remove noise patterns while preserving the essential underlying signal structure. For CIR data, this means the DAE learns to distinguish between true location-specific channel characteristics and transient environmental noise or hardware imperfections. The network effectively denoises the CIR by projecting it into a latent space where noise is suppressed, and then reconstructing the signal from this cleaner representation. This enables the DAE to selectively preserve spatial and temporal features crucial for location fingerprinting while removing environment-specific noise patterns through its learned representation. Our evaluation in section 4.4 reveal that DAEs excel in high-noise outdoor environments, where their adaptive noise reduction capabilities are critical for reliable authentication decisions, in contrast to Butterworth filtering, which performs better in controlled indoor settings.

Algorithm 1: Sliding Window-Based I/Q Trace Extraction and Preprocessing.

Input: Sequence of complex I/Q samples derived from filtered CIR: $\{x_1, x_2, \ldots, x_M\}$

Output: Preprocessed trace set $\mathcal{X} = \{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}\}$

Set trace length L and stride w
Initialize trace index $i \leftarrow 1$
for $k = 1$ **to** $M - L + 1$ w **do**
    Extract trace: $\mathbf{x}^{(i)} \leftarrow [x_k, x_{k+1}, \ldots, x_{k+L-1}]$
    Normalize $\mathrm{Re}(\mathbf{x}^{(i)})$ and $\mathrm{Im}(\mathbf{x}^{(i)})$ via min-max scaling
    Optionally apply STFT to $\mathbf{x}^{(i)}$ for spectrogram generation
    Store $\mathbf{x}^{(i)}$ in $\mathcal{X}$
    $i \leftarrow i + 1$
return $\mathcal{X}$

4.2.5 Extracting I/Q Samples Using Sliding Window

We preprocess the extracted CIR data using a sliding window approach, segmenting complex I/Q representations into fixed-length traces. Each trace serves as an

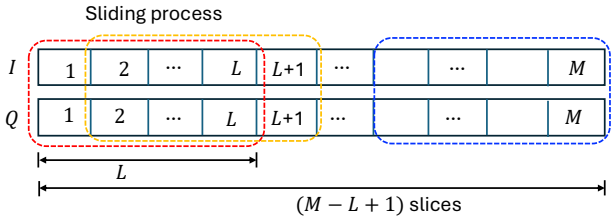


Figure 4.3: Sliding window segmentation of complex I/Q samples with trace length L, producing $(M - L + 1)$ overlapping traces from a sequence of M samples.

input instance for location authentication, providing the temporal context necessary for reliable location verification.

The choice of stride length significantly impacts both computational efficiency and authentication performance. We select a stride $w = 288$ equal to the window length $L = 288$ to create non-overlapping windows for several reasons: computational efficiency, by eliminating redundant processing of the same samples; statistical independence, by ensuring independence between training samples and preventing overfitting to temporal correlations within overlapping segments; and authentication robustness, as our empirical evaluation shows that overlapping windows ($w < L$) lead to memorization of local temporal patterns rather than learning generalizable location-specific features.

Comparative analysis reveals that overlapping windows initially appear to improve training accuracy due to increased sample size, but this improvement stems from data leakage rather than genuine learning of location signatures. During testing with temporally separated data, non-overlapping window training demonstrates superior generalization performance, achieving 3-5% higher authentication accuracy compared to overlapping approaches.

We apply a fixed-length window $L = 288$ and stride $w = 288$ over a sequence of M complex CIR-derived I/Q samples where $L \ll M$. Each trace is defined as $\mathbf{x}^{(i)} = [x_1^{(i)}, x_2^{(i)} \ldots, x_L^{(i)}]$, where $x_k^{(i)} = x_k^{\mathrm{R},(i)} + j x_k^{\mathrm{I},(i)}$ represents the k-th complex I/Q

sample within the i-th trace, derived from the filtered CIR. The window slides across the sequence with a stride w, producing one trace at each step. This continues until we obtain a dataset $\mathcal{X} = \{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}\}$.

For $w = L = 288$, we obtain $(M-L+1)$ total traces from a sequence of M samples, where consecutive traces capture independent temporal segments of the CIR evolution, providing robust training data for location authentication while maintaining computational tractability.

We normalize each trace using min-max normalization to improve numerical stability and model convergence. The I/Q data is then optionally converted to the frequency domain using Short-Time Fourier Transform (STFT), extracting spectrogram features by computing the power spectrum over local windows. This spectrogram emphasizes transient frequency content and enhances the model's robustness to temporal variation. STFT helps the model focus on discriminative channel signatures rather than raw signal variations by ensuring numerical stability, reducing hardware bandwidth to focus on spatial or channel variations instead of absolute signal levels, and capturing time-frequency patterns, which is useful since CIR exhibits temporal and spectral changes due to multipath, soil conditions, and motion. STFT also enhances generalization since spectrograms emphasize features robust to noise, making the model less sensitive to transient fluctuations, minor hardware, or environmental shifts. It is important to note that the training pipeline has no STFT or spectrogram computation. Each input sample is formatted as a real-valued tensor of shape $(\text{batch size}, 2, L)$, where the two channels correspond to the in-phase (I) and quadrature (Q) components, and L is the trace length.

Each trace captures CIR's short-term spectral and temporal properties that are essential for location-based authentication. This slicing method applies equally to the time and frequency-domain representations of the CIR. Each trace is treated as

a two-channel sequence, with one channel for the in-phase (I) and the other for the quadrature (Q) component. This ensures a uniform trace structure and length across all samples, facilitating consistent input for the location authentication model.

4.2.6 Channel Impulse Response in Underground Wireless Communication

Multipath effects from lateral, direct, and reflected propagation components significantly alter the temporal structure of the wireless channel in underground environments, creating unique location-specific signatures that are particularly valuable for authentication in critical infrastructure applications. The CIR captures these effects, which characterize the channel's output in response to an ideal impulse input. The CIR for underground communication [150] can be approximated as a sum of delayed and weighted delta functions given as

$$h_{\mathrm{UG}}(t) = \sum_{l=0}^{L-1} \alpha_l \delta(t - \tau_l) + \sum_{d=0}^{D-1} \alpha_d \delta(t - \tau_d) + \sum_{r=0}^{R-1} \alpha_r \delta(t - \tau_r) \tag{4.12}$$

Here, L, D, R are the numbers of lateral, direct, and reflected wave components. $\alpha_l, \alpha_d, \alpha_r$ are the complex gains for each wave component. These complex gains represent the amplitude and phase changes each wave component undergoes as it propagates through the underground channel. τ_l, τ_d, τ_r are the delays associated with the lateral, direct, and reflected waves, respectively. The CIR inherently captures the environmental characteristics of underground propagation, including the lossy and dispersive nature of soil, and serves as a basis for extracting fingerprintable features such as power delay profile (PDP), RMS delay spread, and coherence bandwidth.

We obtain the PDP by squaring the magnitude of the CIR given by

$$P(\tau) = |h(\tau)|^2, \tag{4.13}$$

The mean excess delay uses the power from the PDP to compute the average delays, the time domain characterized by the channel based on the CIR's multipath structure. This is given by

$$\bar{\tau} = \frac{\int_{\tau_{\min}}^{\tau_{\max}} \tau P(\tau)\, d\tau}{\int_{\tau_{\min}}^{\tau_{\max}} P(\tau)\, d\tau}, \tag{4.14}$$

where $P(\tau)$ is the power at a delay τ, $\bar{\tau}$ is the mean excess delay, $\tau_{\min}$ and $\tau_{\max}$ are the range of delays. The numerator is the weighted sum of delays, where the weights are the power values from the CIR, and the denominator is the total received power integral of the PDP. Equation (4.14) can also be represented in discrete form as

$$\bar{\tau} = \frac{\sum_k P_k \tau_k}{\sum_k P_k}, \tag{4.15}$$

where $P_k = |\alpha_k|^2$ is the power of the k-th multipath component from the impulse response, τ_k is the delay of the k-th multipath component, and α_k are all the multipath amplitudes.

The root mean square (RMS) delay spread quantifies the temporal dispersion of multipath components around the mean excess delay and is a key indicator for channel time dispersion. RMS delay spread evaluates the temporal dispersion, which directly affects communication performance. RMS delay spread is defined as the square root of the second central moment of the power delay profile, which is given as

$$\tau_{\text{rms}} = \sqrt{(\tau^2) - (\bar{\tau})^2}, \tag{4.16}$$

which can be expanded in discrete form as

$$\tau_{\text{rms}} = \sqrt{\frac{\sum_k P_k \tau_k^2}{\sum_k P_k} - \left(\frac{\sum_k P_k \tau_k}{\sum_k P_k}\right)^2}, \tag{4.17}$$

where $\bar{\tau} = \frac{\sum_k P_k \tau_k}{\sum_k P_k}$ is the mean excess delay, τ_k is the k-th multipath component.

The RMS delay spread indicates the extent of dispersion in the delays. A larger RMS delay spread signifies greater temporal dispersion of the signal, potentially resulting in overlapping symbols (ISI) in high-data-rate communication. RMS Delay spread is inversely related to the coherence bandwidth.

$$B_c \approx \frac{1}{\tau_{\text{rms}}}. \tag{4.18}$$

The inverse relationship between the RMS delay spread and coherence bandwidth reflects the effects of multipath dispersion on frequency selectivity. Coherence bandwidth characterizes the frequency band over which the channel response remains approximately constant and is typically estimated from the channel transfer function. Channels with large RMS delay spreads exhibit smaller coherence bandwidths, indicating higher frequency selectivity and reduced diversity.

Underground channels typically exhibit large RMS delay spread and low coherence bandwidth due to significant multipath propagation within the lossy and dispersive underground medium, creating robust location-specific signatures that are difficult for adversaries to replicate. This limits frequency diversity but provides strong authentication features and favors operation in the low-frequency zone where attenuation and signal stability create consistent authentication signatures.

4.2.7 Large-scale vs Small-scale Fading for Authentication

Large-scale fading refers to variations in signal strength over relatively large distances or time durations, providing stable location-specific signatures that are valuable for authentication. These variations are primarily caused by path loss and shadowing due to macroscopic terrain features. Under large-scale fading, the CIR records deterministic effects that create consistent location fingerprints, including path attenuation and blockage patterns that are difficult for adversaries to replicate.

Small-scale fading results from reflection, diffraction, and scattering effects that cause rapid signal variations. While these fluctuations can degrade authentication consistency in dynamic situations, they also provide fine-grained location-specific features that enhance security against sophisticated spoofing attempts.

We utilize amplitude and phase-based filtering approaches to separate large-scale fading from small-scale fading, preserving stable authentication features while removing rapidly varying noise that could compromise location verification reliability. We assess the filtered signals using multiple statistical metrics to ensure that location-discriminative features essential for authentication are preserved. This filtered CIR is critical for reliable authentication decisions in machine learning-based location verification systems.

We also investigate temporal and spatial correlation in the channel data for authentication system design. While temporal correlation shows the stability of authentication features over time, spatial correlation captures the uniqueness of channel behavior across different positions. Together, these metrics provide crucial insights for wireless channel-based authentication in dynamic environments, enabling robust location verification that can adapt to changing conditions while maintaining security guarantees.

4.2.8 Spatio-Temporal Channel Correlation

4.2.8.1 Temporal Correlation

Temporal correlation computes the statistical relationship between wireless channel properties at different times, directly impacting authentication system stability. It captures how channel characteristics change and whether location-specific features remain consistent over time.

Temporal correlation guides location-based authentication by determining channel predictability for verification decisions. High temporal correlation indicates stable location signatures, allowing longer authentication validity periods. Low temporal correlation signifies rapid channel variation, requiring more frequent re-authentication or adaptive thresholds to maintain authentication accuracy.

4.2.8.2 Spatial Correlation

Spatial correlation quantifies the statistical relationship between wireless channel properties at different spatial locations, providing the foundation for location-based authentication security. It captures the uniqueness of channel behavior across different positions, enabling discrimination between authorized and unauthorized locations.

Understanding spatial variations in channel characteristics is essential for robust authentication system design. Low spatial correlation across locations ensures that each authorized zone has distinct signatures that adversaries cannot easily replicate from different positions. This information guides the development of robust location authentication techniques capable of precisely distinguishing between legitimate and spoofed transmissions across space.

4.2.9 Adversarial Domain Adaptation for Location Authentication

We applied Adversarial Domain Adaptation (ADA) to improve the generalization of location authentication models across indoor and outdoor settings for both OTA and underground scenarios. Environmental variations such as soil moisture, multipath fading, and signal attenuation cause significant variation in wireless channel properties that could compromise authentication effectiveness when models trained in one environment are deployed in another.

ADA tackles this problem by aligning feature distributions across domains using adversarial learning, ensuring robust authentication performance regardless of environmental conditions. This is particularly crucial for authentication systems that must maintain security guarantees across diverse deployment scenarios.

The propagation of wireless signals varies with the environment. While outdoor environments experience fading and diffraction due to large-scale obstructions, indoor environments induce dense multipath from walls and objects. Because of the considerable attenuation in the soil and the refractive effects at the soil-air interface, underground to above-ground transmission adds even more complexity. These domain-dependent effects restrict the generalizing capability of conventional machine learning systems. ADA helps to reduce this restriction by ensuring that the feature learned from one domain, for instance, indoor, remains effective when applied to another, such as outdoor settings.

The ADA architecture consists of three components: a feature extractor θ_f, a source classifier θ_y, and a domain discriminator θ_d, as illustrated in Figure 4.9. The feature extractors θ_f learn the channel attributes like the power delay profile, RMS delay spread, and coherence bandwidth. The classifier θ_y predicts the location utilizing the retrieved attributes, whereas the domain discriminator θ_d attempts to distinguish

between various domains, including indoor and outdoor settings. A gradient reversal layer (GRL) nullifies the gradients from the discriminator, forcing the feature extractor to acquire domain-invariant features for robust classification over diverse settings.

Let $G_f(.;\theta_f)$ denote the neural network-based feature extractor with parameters θ_f. Let $G_y(.;\theta_y)$ represent the label prediction module with parameter θ_y, and $G_d(.;\theta_d)$ denote the domain discriminator with parameters θ_d. We define the total ADA objectives as

$$\mathcal{L}(\theta_f, \theta_y, \theta_d) = \mathcal{L}_{\text{clas}}(\theta_f, \theta_y) - \lambda \cdot \mathcal{L}_{\text{dom}}(\theta_f, \theta_d) \tag{4.19}$$

Where;

$$\mathcal{L}_{\text{clas}}(\theta_f, \theta_y) = \mathcal{L}_y\left(G_y(G_f(x_i;\theta_f);\theta_y), y_i\right) \tag{4.20}$$

$$\mathcal{L}_{\text{dom}}(\theta_f, \theta_d) = \mathcal{L}_d\left(G_d(G_f(x_i;\theta_f);\theta_d), d_i\right) \tag{4.21}$$

Where G_y, G_f, and G_d are the label classifier, feature extractor, and domain discriminator. y_i, and d_i are the true class and domain labels. $\mathcal{L}_{\text{clas}}$ is the supervised cross-entropy loss for the source classification on source domain data, $\mathcal{L}_{\text{dom}}$ is the binary cross-entropy loss function for the domain discrimination (source vs. target), and λ is the weighting factor.

During training, we used the source and target training data. After training, only θ_f and θ_y are retained for inference. This enables location prediction on unseen domains without requiring domain labels at test time.

ADA enhances authentication robustness by enabling models trained in one environment to provide reliable location verification in different environmental contexts.

The learned channel features are robust to environmental variations, resulting in enhanced authentication reliability in real-world applications. By strengthening feature learning, ADA additionally improves the security of location authentication systems by reducing susceptibility to adversarial attacks that attempt to exploit environmental domain shifts.

4.2.10 Fine-tuning for Location Authentication Models

We fine-tuned pre-trained CNNs to enhance their adaptability to our wireless CIR data's spatial and temporal characteristics for improved authentication performance. Our process involves optimizations including hyperparameter tuning, learning rate adjustment, and regularization techniques to prevent overfitting while maintaining authentication accuracy.

Fine-tuning addresses the challenge of limited labeled authentication data by adapting existing CNN models to emphasize domain-specific channel features crucial for location verification. These include coherence bandwidth, RMS delay spread, and power delay profile characteristics that provide location-specific authentication signatures.

The fine-tuning technique begins with initializing a pre-trained CNN and substituting its final classification layers with new layers tailored for location classification. The network undergoes additional training with a smaller learning rate to prevent significant weight updates while enabling it to learn information from the underground and the aboveground channel data. This technique guarantees that the model can extract generalizable features while enhancing its understanding of the distinct characteristics of wireless propagation settings.

Optimizing the fine-tuned model necessitates the selection of suitable hyperparameters, such as reducing learning rate schedules and batch sizes of 32, and using the

Adam optimizer and regularization methods to mitigate overfitting. If memory errors occur, the system dynamically reduces the batch size to 8 and retries training.

Transfer learning is essential for authentication systems as it minimizes the requirement for extensive training data while enhancing location verification accuracy. This approach ensures that authentication models can extract generalizable security-relevant features while adapting to the specific characteristics of the deployment environment.

During evaluation, we compute authentication accuracy and obtain rank statistics using predicted probability distributions, enabling assessment of authentication confidence and reliability across different threat scenarios.

4.2.11 Evaluating CIR Processing Techniques for Authentication

Using statistical correlation metrics and error analyses, we assess the efficacy of different CIR processing approaches by comparing original, filtered, and denoised CIR data samples. These measures evaluate the effectiveness of Butterworth filtering and denoising autoencoder (DAE) techniques in maintaining location-discriminative features essential for reliable authentication while enhancing signal clarity.

4.2.11.1 The Spearman Rank Correlation Coefficient (ρ)

This gauges the monotonic relationship between the CIR dataset's ranked values. Spearman's Rank correlation coefficient captures both linear and non-linear patterns. Low p-value and high Spearman coefficient suggest a consistent rank order is maintained after processing [173].

4.2.11.2 The Pearson Correlation Coefficient

The Pearson correlation coefficient [131] evaluates the linear correlation between the original and processed CIR datasets. Unlike Spearman's rank-based approach, Pearson's is sensitive to the actual magnitude and distribution of the signal values, making it particularly useful for assessing amplitude preservation in CIR processing. A high Pearson coefficient and a low p-value indicate that the filtering process maintains the underlying signal structure while effectively removing high-frequency artifacts, ensuring that location-discriminative features remain intact for reliable fingerprinting. Pearson's quantifies the strength and direction of linear correlation, ranging from -1 to $+1$. A high Pearson coefficient (approaching 1) with a low p-value indicates that the processed CIR data maintains a strong linear relationship with the original signal, suggesting that the essential amplitude and phase characteristics are preserved during filtering while removing unwanted noise components [59].

4.2.11.3 The R-squared coefficient of determination (R^2)

R^2 quantifies the extent to which the processed CIR data preserves the variability of the original CIR data. A R^2 value approaching 1 indicates that the processed data accurately represents the structure and variability of the original data, preserving essential spatial and temporal properties [41, 72].

4.2.11.4 The MSE, RMSE, and MAE Error metrics

We also compute normalized Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) to measure variations introduced by processing. While MAE gives a robust average error that is less sensitive to outliers, MSE emphasizes larger errors, and RMSE provides an interpretable error

magnitude in the original signal units.

Low error measures of MSE, RMSE, MAE, and high correlation coefficient of Spearman and the coefficient of Pearson demonstrate that filtering preserves important relations. A high R^2 further supports that the processed signal preserves the core spatial and temporal variability. These measures guarantee that the processed CIR data maintains the spatial and temporal attributes crucial for precise location fingerprinting in indoor and outdoor settings.

4.2.12 Comprehensive CIR Processing Performance Analysis for Authentication

We contrast the original CIR, filtered CIR, and denoised CIR data for both indoor and outdoor settings in same-RX different-TX experiments, evaluating their effectiveness for location authentication applications. Our results are presented in Figures 4.4, 4.5, 4.7, 4.6, and 4.8, along with detailed statistics in Table 4.4, demonstrating environment-specific effectiveness of different processing techniques.

4.2.12.1 Original CIR vs. Butterworth Filtered CIR

We evaluate Butterworth filtering performance by comparing original and filtered CIR data across indoor and outdoor environments, assessing preservation of authentication-relevant features.

For the indoor environment, Butterworth filtering demonstrates moderate preservation of signal characteristics with a Spearman rank correlation coefficient of $\rho = 0.6032$, $R^2 = 0.224$, and a Pearson correlation coefficient of 0.6243. The normalized error metrics show MAE $= 1.04 \times 10^{-1}$, RMSE $= 1.31 \times 10^{-1}$, and MSE $= 1.71 \times 10^{-2}$, with a PSNR of 17.67 dB, indicating reasonable signal preservation with controlled noise reduction. These results are illustrated in Figure 4.4(c,f) and summarized in

Table 4.4.

In the outdoor environment, filtering exhibited similar correlation performance with Spearman $\rho = 0.6024$, $R^2 = 0.191$, and a Pearson correlation coefficient of 0.6340. The normalized error metrics were MAE = 9.12×10^{-2}, RMSE = 1.15×10^{-1}, and MSE = 1.32×10^{-2}, with a higher PSNR of 18.80 dB. The slightly better error performance in outdoor settings suggests that Butterworth filtering is more effective in environments with higher baseline noise levels, as shown in Figure 4.5(c,f).

4.2.12.2 Original CIR vs. Denoised CIR

The denoising autoencoder (DAE) exhibits environment-specific performance, revealing varying effectiveness across settings as illustrated in Figure 4.4.

In the indoor environment, DAE denoising shows poor correlation with the original signal, achieving only a Spearman $\rho = 0.3276$, a Pearson coefficient of 0.3592, and a negative $R^2 = -1.215$, indicating that the denoised signal explains less variance than a simple mean model. The normalized error metrics are substantially higher: MAE = 1.64×10^{-1}, RMSE = 2.22×10^{-1}, and MSE = 4.91×10^{-2}, with a PSNR of 13.09 dB. These results, shown in Figure 4.4(a,b), suggest that the DAE over-processes the relatively clean indoor signals.

Conversely, in the outdoor environment, DAE demonstrates excellent performance, with Spearman $\rho = 0.8766$, Pearson coefficient of 0.8872, and $R^2 = 0.7537$, indicating strong preservation of signal structure. The normalized error metrics are significantly lower: MAE = 4.70×10^{-2}, RMSE = 6.06×10^{-2}, and MSE = 3.68×10^{-3}, achieving the highest PSNR of 24.01 dB across all comparisons. This superior performance, illustrated in Figure 4.5(a,b), suggests that DAE excels in high-noise outdoor environments where its learned noise patterns effectively separate signal from environmental interference.

4.2.12.3 Denoised CIR vs. Butterworth Filtered CIR

Direct comparison between the two processing approaches reveals their relative strengths and trade-offs.

In the indoor environment, the comparison shows poor correlation with a Spearman $\rho = 0.1864$, a Pearson coefficient of 0.2159, and $R^2 = -1.411$, indicating that the two methods produce substantially different outputs. The normalized errors are MAE = 1.82×10^{-1}, RMSE = 2.31×10^{-1}, and MSE = 5.33×10^{-2}, with a PSNR of 6.054 dB, suggesting fundamental differences in their processing approaches for clean indoor signals. These results are evident in the correlation analysis shown in Figure 4.7(a,c).

For the outdoor environment, the comparison shows moderate correlation with a Spearman $\rho = 0.5199$, Pearson coefficient of 0.5491, and $R^2 = 0.157$, with normalized errors of MAE = 8.23×10^{-2}, RMSE = 1.04×10^{-1}, and MSE = 1.09×10^{-2}, and a PSNR of 19.65 dB. This indicates that both methods provide similar processing outcomes in noisy outdoor environments, as demonstrated in Figure 4.7(d,e).

4.2.12.4 Comparative Analysis and Implications

Our comprehensive evaluation reveals environment-specific effectiveness.

Butterworth filtering consistently outperforms DAE denoising in indoor settings, providing moderate but stable signal preservation ($R^2 = 0.2237$ vs. -1.215). The controlled indoor environment contains minimal noise for DAE to suppress, which causes over-processing and signal degradation.

In contrast, in outdoor environments, DAE denoising significantly outperforms Butterworth filtering ($R^2 = 0.7537$ vs. 0.1912), demonstrating superior capability in high-noise conditions. DAE learns and removes complex environmental noise patterns while preserving essential channel characteristics.

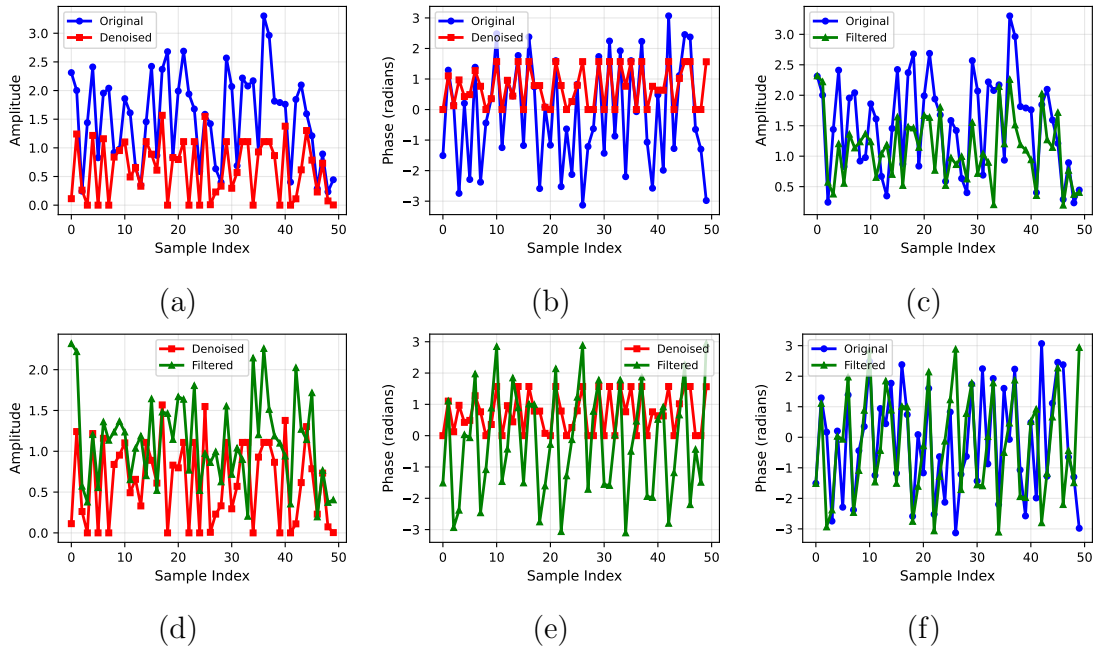


Figure 4.4: CIR signal processing comparison for outdoor environments showing amplitude and phase characteristics across the first 50 samples in the same-RX, different-TX setup: (a) Amplitude comparison between original and denoised signals, (b) Phase comparison between original and denoised signals, (c) Amplitude comparison between original and filtered signals, (d) Amplitude comparison between denoised and filtered signals, (e) Phase comparison between denoised and filtered signals, (f) Phase comparison between original and filtered signals.

While Butterworth filtering offers predictable, frequency-domain noise reduction, it uniformly attenuates high-frequency components that often contain discriminative channel information. In comparison, DAE selectively preserves spatial and temporal features critical for location fingerprinting by leveraging learned representations of environment-specific noise.

These findings indicate that optimal CIR processing requires environment-aware technique selection: Butterworth filtering for controlled indoor settings and DAE denoising for complex outdoor scenarios. This adaptive strategy ensures maximum preservation of discriminative channel properties necessary for accurate location authentication across diverse deployment environments.

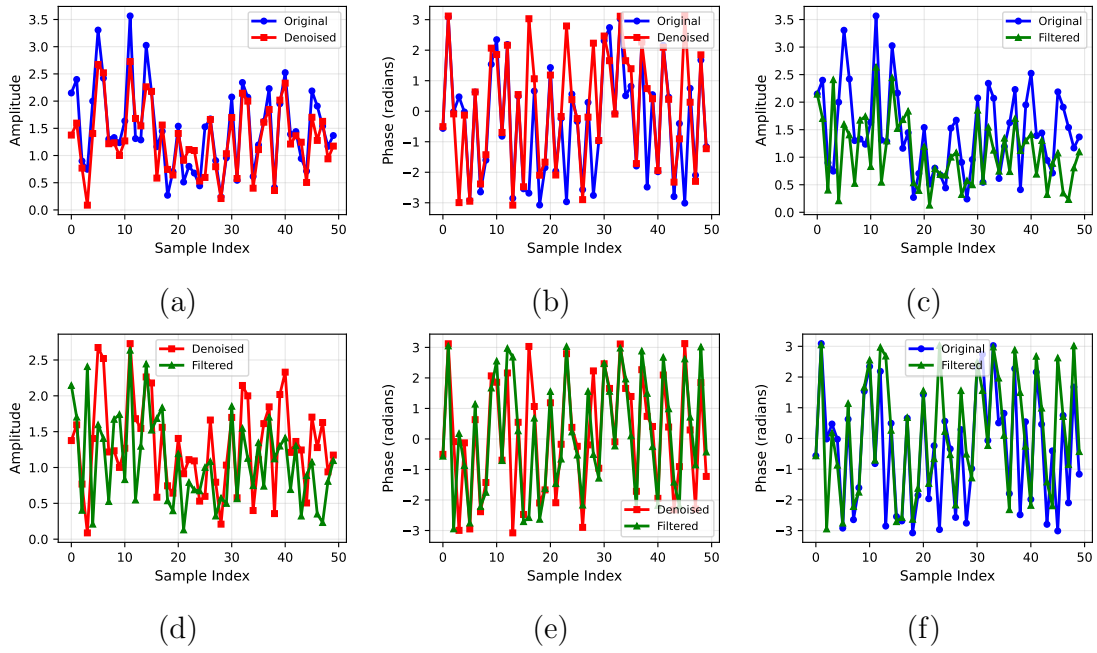


Figure 4.5: CIR signal processing comparison for outdoor environments showing amplitude and phase characteristics across the first 50 samples in the same-RX, different-TX setup. (a) Amplitude comparison between original and denoised signals. (b) Phase comparison between original and denoised signals. (c) Amplitude comparison between original and filtered signals. (d) Amplitude comparison between denoised and filtered signals. (e) Phase comparison between denoised and filtered signals. (f) Phase comparison between original and filtered signals. These comparisons highlight the signal preservation and transformation behaviors of denoising and filtering methods.

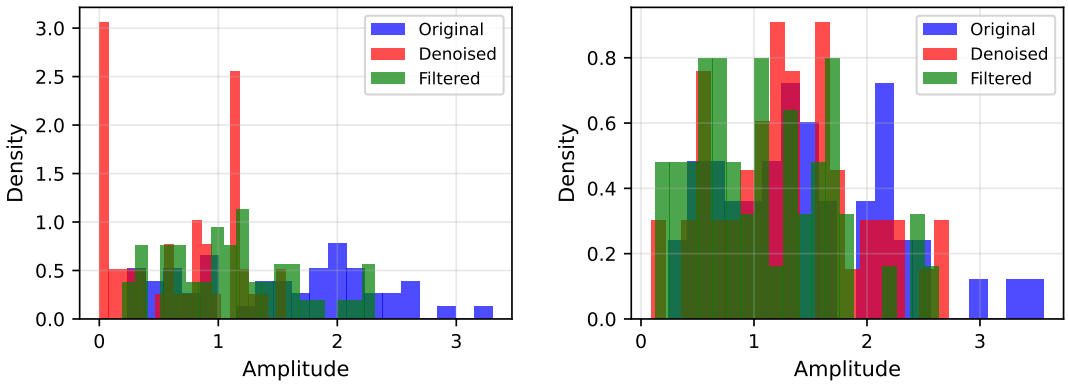
Table 4.4: Quantitative comparison of different CIR processing approaches for indoor and outdoor settings.

Setting	**Comparison**	**Pearson**	**P-Value**	**Spearman ρ**	**MAE**	**RMSE**	**MSE**	R^2	**PSNR (dB)**
Indoor	Orig. vs Denoised	0.3592	0.0	0.3276	1.64×10^{-1}	2.22×10^{-1}	4.91×10^{-2}	-1.215	13.09
	Orig. vs Filtered	0.6243	0.0	0.6032	1.04×10^{-1}	1.31×10^{-1}	1.71×10^{-2}	0.224	17.67
	Denoised vs Filtered	0.2159	0.0	0.1864	1.82×10^{-1}	2.31×10^{-1}	5.33×10^{-2}	-1.411	6.054
Outdoor	Orig. vs Denoised	0.8872	0.0	0.8766	4.70×10^{-2}	6.06×10^{-2}	3.68×10^{-3}	0.754	24.01
	Orig. vs Filtered	0.6340	0.0	0.6024	9.12×10^{-2}	1.15×10^{-1}	1.32×10^{-2}	0.191	18.80
	Denoised vs Filtered	0.5491	0.0	0.5199	8.23×10^{-2}	1.04×10^{-1}	1.09×10^{-2}	0.157	19.65

4.3 Experimental Evaluation

4.3.1 OTA Experimental Setup

We collect over-the-air (OTA) BPSK samples at three indoor locations using GNU Radio and software-defined radio (SDR) USRP devices, specifically the B200 and



(a) Indoor amplitude distribution (b) Outdoor amplitude distribution

Figure 4.6: Amplitude histogram distribution plots of CIR data for (a) indoor and (b) outdoor environments comparing original, filtered, and denoised signals.

B205mini, to evaluate our location authentication system. BPSK is selected for its simplicity, robustness to noise, and low bit error rate (BER) under low signal-to-noise ratio (SNR) conditions. It enables efficient channel extraction with minimal processing, preserving authentication-relevant features. Furthermore, BPSK operates in the time domain, offering an accurate representation of the wireless channel—critical for reliable location-based authentication in resource-constrained environments.

4.3.2 OTA Experimental Setup

We collect OTA BPSK samples at three indoor locations using GNU Radio and software-defined radio (SDR) USRP devices, specifically the B200 and B205mini, to evaluate our location authentication system. We select the BPSK modulation technique and direct signal-based CIR extraction over pilot-based channel estimation due to several critical advantages for authentication performance and data efficiency.

BPSK Selection Rationale: BPSK offers several advantages for location authentication applications. Its binary nature provides simplicity and robustness to

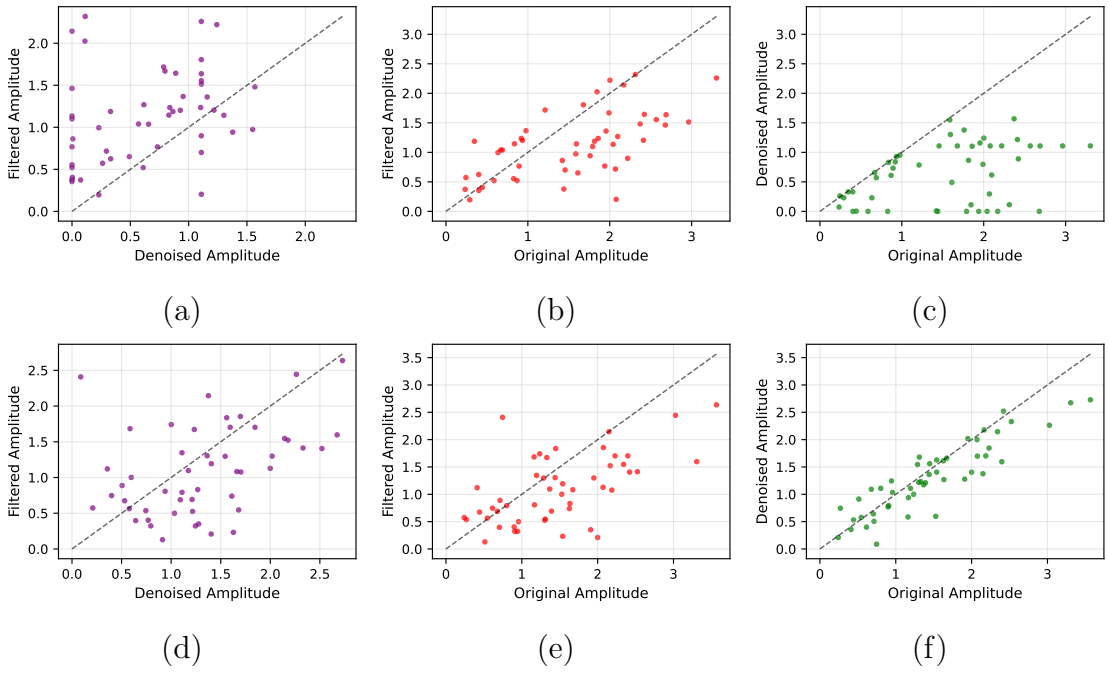


Figure 4.7: Pairwise CIR correlation scatter plots for indoor (top row) and outdoor (bottom row) environments: (a) Indoor denoised amplitude against filtered amplitude, (b) Indoor original amplitude against filtered amplitude, (c) Indoor original amplitude against denoised amplitude, (d) Outdoor denoised amplitude against filtered amplitude, (e) Outdoor original amplitude against filtered amplitude, (f) Outdoor original amplitude against denoised amplitude.

noise while maintaining a low bit error rate (BER) under low signal-to-noise ratio (SNR) conditions. It requires minimal processing for channel extraction and preserves authentication-relevant channel characteristics, making it computationally efficient and well-suited for resource-constrained authentication systems. The constant envelope property of BPSK ensures consistent power transmission, which is essential for reliable channel characterization across different spatial locations.

Direct Signal-Based CIR vs. Pilot-Based Estimation: We extract CIR directly from the entire BPSK signal rather than relying on sparse pilot symbols for several authentication-specific reasons. Pilot-based channel estimation typically uses only $10 - 20\%$ of the transmission frame, providing limited temporal samples for

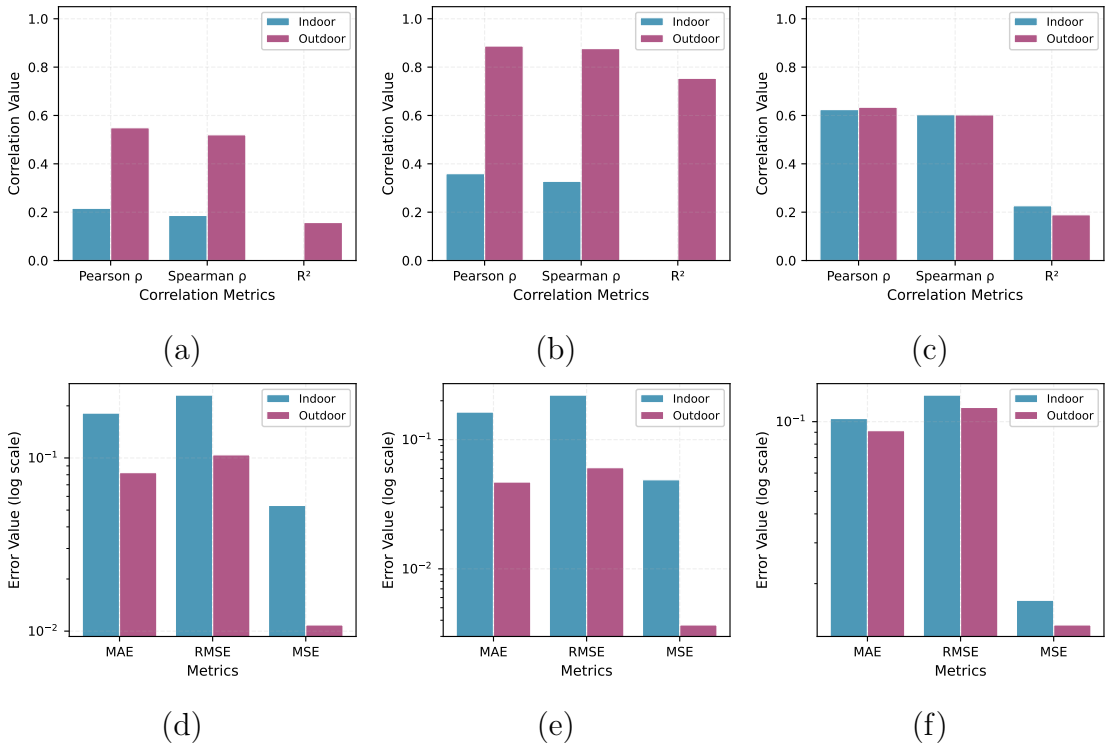


Figure 4.8: Quantitative performance comparison of CIR signal processing methods in indoor vs. outdoor environments. The top row shows correlation metrics: (a) denoised vs. filtered signals correlation (Pearson ρ, Spearman ρ, R^2), (b) original vs. denoised signals correlation, and (c) original vs. filtered signals correlation. The bottom row presents error metrics on a logarithmic scale: (d) denoised vs. filtered error comparison (MAE, RMSE, MSE), (e) original vs. denoised error comparison, and (f) original vs. filtered error comparison. Indoor environments (blue bars) consistently demonstrate higher correlation values and lower error metrics than outdoor environments (magenta bars), indicating more stable channel conditions. All processing methods show strong correlation preservation (¿0.8), with filtering approaches achieving superior noise reduction performance across both environments.

location fingerprinting. In contrast, our approach utilizes the complete BPSK signal for CIR extraction, generating significantly larger datasets essential for robust CNN-based authentication models.

The authentication task requires extensive training data to learn subtle location-specific channel variations that distinguish authorized from unauthorized positions. Pilot-based approaches yield insufficient data volume for effective deep learning model

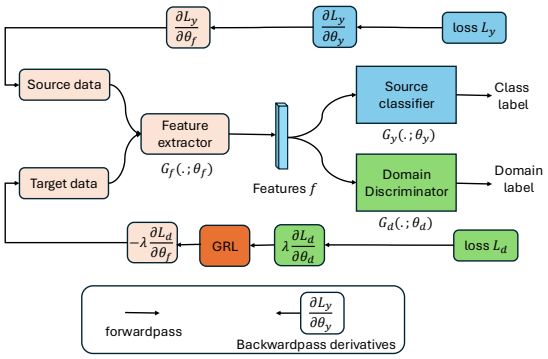


Figure 4.9: Adversarial Domain Adaptation architecture with a shared feature extractor, class label predictor for source domain classification, and a domain discriminator optimized via a gradient reversal layer.

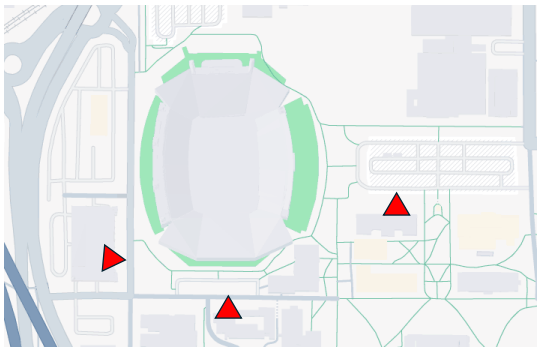


Figure 4.10: Schematic map of the locations used for outdoor wireless data collection.

training, particularly for capturing the complex spatial relationships necessary for reliable location verification. Our BPSK-based approach generates the substantial datasets shown in Table 4.5, with total volumes exceeding 50 GB per experimental configuration critical for training robust authentication models.

Furthermore, continuous signal-based CIR extraction captures finer temporal variations in channel characteristics compared to sparse pilot-based estimation. These detailed temporal signatures provide richer location-specific features that enhance authentication accuracy and reliability across diverse environmental conditions. The approach enables comprehensive characterization of multipath, fading, and propagation effects that serve as unique spatial fingerprints for location-based authentication.

BPSK operation in the time domain enables accurate channel representation while

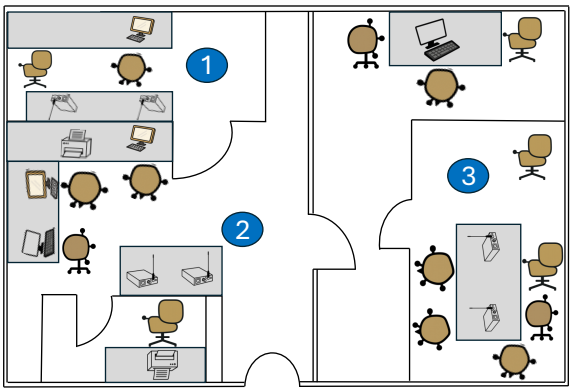


Figure 4.11: Schematic layout of the indoor wireless data collection locations.

ensuring the data volume necessary for effective machine learning model development.

4.3.2.1 Indoor Experimental Settings

Our indoor testbed setup comprises aboveground transmitter and receiver pairs implemented using USRP B200 and B205mini software-defined radios (SDRs). We select these USRPs for their portability and plug-and-play support, enabling flexible authentication system deployment. A Lenovo ThinkPad T14 runs GNU Radio BPSK modulation code with a 6.4 MSps sampling rate, 4 samples per symbol, and a 2.45 GHz center frequency. We choose 2.45 GHz as the operating frequency to prioritize bandwidth over range for controlled indoor authentication evaluation.

In the distance-based authentication experiment, we first move the transmitter USRP while keeping the receiver USRP fixed, evaluating authentication performance across spatial separations. We transmit signals at 3 ft, 4 ft, 5 ft, and 6 ft, distances selected to thoroughly evaluate system performance within typical operational ranges and demonstrate feasibility in near-field indoor scenarios for authentication analysis. In the second setup, we vary the transmitter USRP at each distance while keeping the receiver fixed, enabling evaluation of authentication robustness against device variations at specific locations.

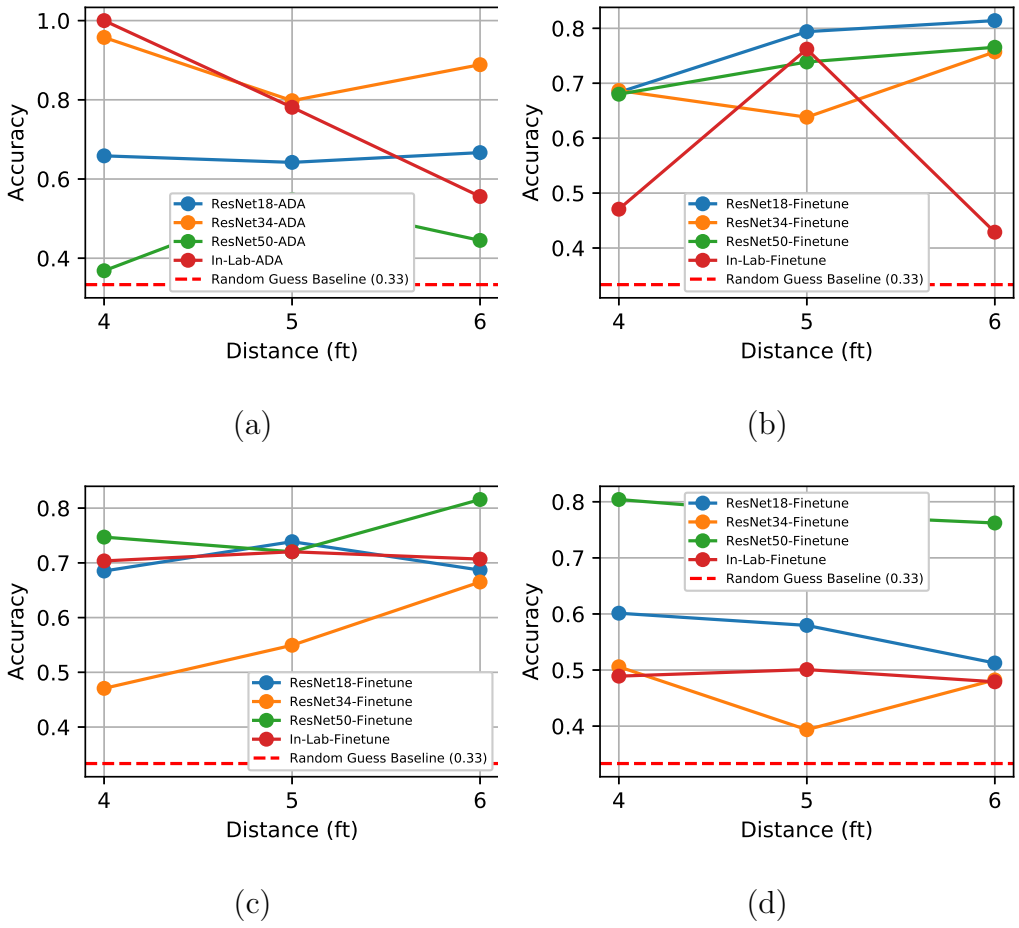


Figure 4.12: Outdoor experiments using different CNN-based models for authentication under varying transmitter-receiver pairs. Plots show accuracy versus distance for: (a) ADA models with ReLU activation and Butterworth filtering in same-receiver, different-transmitter setting, (b) fine-tuning with ReLU and Butterworth filtering in the same setting, (c) fine-tuning with Linear activation and no filtering, and (d) fine-tuning with ReLU activation and no filtering, all for same-receiver, different-transmitter settings.

In the device-based authentication experiment, we fix the transmitter-receiver USRP separation at 5 ft to evaluate authentication performance across different hardware configurations. We maintain the same receiver and vary the transmitter among six USRP devices, collecting CIR data for each authentication scenario. This setup evaluates the system's ability to authenticate different transmitters from the same

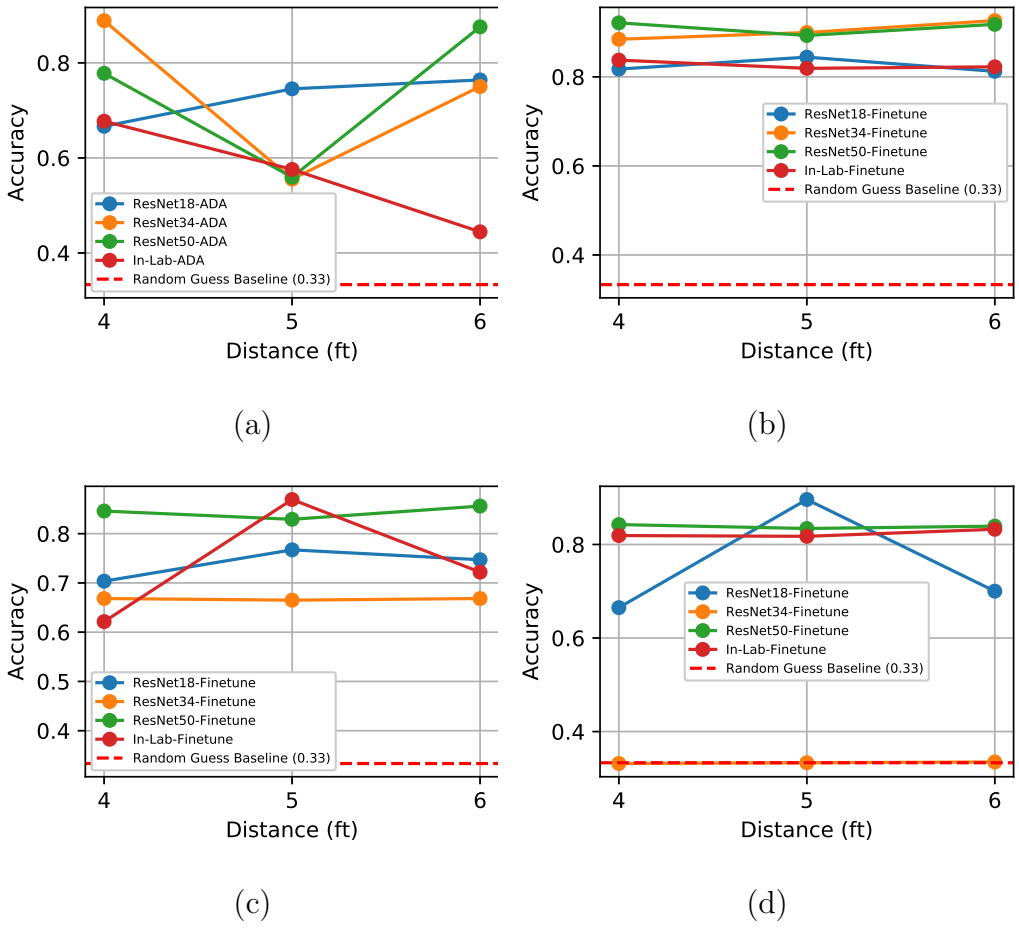


Figure 4.13: Outdoor experiments using different CNN-based models for authentication under varying transmitter-receiver pairs. Plots show accuracy versus distance for: (a) ADA with ReLU activation and Butterworth filtering for same receiver and same transmitter setting, (b) fine-tuning with ReLU activation and Butterworth filtering for same receiver and same transmitter setting, (c) fine-tuning with Linear activation and no filtering, and (d) fine-tuning with ReLU activation and no filtering, all under the same-receiver, same-transmitter setting.

location while maintaining receiver consistency.

We store distinct transmitter and receiver signals in binary files for each authentication experiment. These files serve as input for the preprocessing phase, where we compute CIR data essential for location-based authentication decisions. This experiment validates OTA authentication performance across three indoor locations with

varying environmental characteristics.

4.3.2.2 Outdoor Experimental Settings

The outdoor environment introduces additional noise and multipath effects, providing a challenging testbed for evaluating authentication robustness. In the outdoor authentication experiments, we replicate the setup used in indoor settings to enable comparative authentication performance analysis. We conduct data collection across three distinct outdoor locations, each representing different environmental conditions that affect authentication reliability.

At each location, we move the transmitter USRP while keeping the receiver USRP stationary, transmitting authentication signals at 3 ft, 4 ft, 5 ft, and 6 ft. These distances are selected to assess system performance within representative short-range outdoor operational environments. We record the received signals in separate binary files corresponding to each distance, enabling comprehensive authentication performance analysis across spatial variations. This setup evaluates the authentication system's ability to maintain location verification accuracy despite environmental interference.

In the second phase of the experiment, we keep the receiver USRP device stationary and vary the transmitter USRP at each distance. This configuration allows us to evaluate authentication robustness against both spatial and hardware variations, simulating real-world deployment scenarios where different devices operate from authorized locations.

For the device-based authentication evaluation, we fix the transmitter-receiver separation at 5 ft and vary both the transmitter and receiver across six distinct USRP units. This comprehensive evaluation ensures that the authentication system reliably verifies location-based trust regardless of specific hardware configurations, demon-

Table 4.5: Dataset Sizes Before and After CIR Processing

Setting	Condition	Before CIR (GB)	After CIR (GB)
Indoor Devices	Same RX, Diff TX	27.9	13.9
	Diff RX, Diff TX	28.1	14.0
Indoor Distances	Same RX, Same TX	18.7	9.3
	Same RX, Diff TX	18.7	9.3
Outdoor Devices	Same RX, Diff TX	28.2	14.1
	Diff RX, Diff TX	28.0	14.0
Outdoor Distances	Same RX, Same TX	18.7	9.3
	Same RX, Diff TX	18.7	9.3

strating practical deployment feasibility.

4.3.3 Dataset Size Summary

We comprehensively summarize the dataset sizes collected across different environmental settings and experimental conditions:

4.3.3.1 Indoor Device-Based Dataset

The total data volume is approximately 56GB, comprising same receiver–different transmitter (27.9GB before CIR processing; 13.9GB after) and different receiver–different transmitter (28.1GB before CIR processing; 14GB after).

4.3.3.2 Indoor Distance-Based Dataset

We collected 37.4GB of raw data, consisting of same receiver–different transmitter (18.7GB before CIR processing; 9.3GB after) and same receiver–same transmitter (18.7GB before CIR processing; 9.3GB after).

4.3.3.3 Outdoor Device-Based Dataset

The dataset amounts to approximately 56.2GB, including different receiver–different transmitter (28GB before CIR processing; 14GB after) and same receiver–different

Condition	Diff. Rx – Diff. Tx				Same Rx – Diff. Tx			
	ResNet-18	ResNet-34	ResNet-50	InLab	ResNet-18	ResNet-34	ResNet-50	InLab
ADA-Butter-ReLU	77.4 ± 8.8	92.4 ± 9.3	47.4 ± 12.0	86.6 ± 4.8	71.9 ± 8.3	76.6 ± 13.3	45.1 ± 11.7	84.5 ± 9.7
Fine-ReLU	53.9 ± 4.3	73.9 ± 1.6	74.8 ± 2.6	79.1 ± 3.6	64.2 ± 2.9	33.3 ± 0.1	74.3 ± 2.3	75.7 ± 1.5
Fine-Linear	71.1 ± 5.6	70.0 ± 4.8	76.1 ± 1.9	55.5 ± 2.2	67.4 ± 1.1	66.5 ± 0.5	70.9 ± 1.8	87.8 ± 3.0
Fine-Butter-ReLU	74.3 ± 6.0	74.0 ± 13.1	86.3 ± 1.3	58.9 ± 7.1	66.4 ± 0.4	68.2 ± 2.8	74.6 ± 2.9	67.7 ± 1.4

Table 4.6: Mean ± standard deviation of classification accuracy (%) for each CNN model and condition in both "Different Rx – Different Tx" and "Same Rx – Different Tx" settings in the devices' outdoor experiments.

Condition	Same Rx – Same Tx				Same Rx – Diff. Tx			
	ResNet-18	ResNet-34	ResNet-50	InLab	ResNet-18	ResNet-34	ResNet-50	InLab
ADA-Butter-ReLU	72.5 ± 4.2	73.1 ± 13.7	73.8± 13.2	56.6 ± 9.5	65.6 ± 1.0	88.1 ± 6.5	45.4 ± 7.4	77.9 ± 18.1
Fine-ReLU	75.4 ± 10.2	33.3 ± 0.1	83.9 ± 0.3	82.3 ± 0.7	56.4 ± 3.8	46.1 ± 4.8	78.1 ± 1.7	49.0 ± 0.9
Fine-Linear	73.9 ± 2.7	66.7 ± 0.2	84.4 ± 1.1	73.8 ± 10.2	70.4 ± 2.5	56.2 ± 8.0	76.1 ± 4.0	71.0 ± 0.7
Fine-Butter-ReLU	82.5 ± 1.4	90.3 ± 1.7	91.1 ± 1.3	82.6 ± 0.8	76.4 ± 5.7	69.4 ± 4.9	72.8 ± 3.6	55.4 ± 14.8

Table 4.7: Mean ± standard deviation of classification accuracy (%) for each CNN model and condition in both "Same Rx – Same Tx" and "Same Rx – Different Tx" outdoor distances' experiments.

transmitter (28.2GB before CIR processing; 14.1GB after).

4.3.3.4 Outdoor Distance-Based Dataset

We collected 37.4GB of raw data, consisting of same receiver–different transmitter (18.7GB before CIR processing; 9.3GB after) and same receiver–same transmitter (18.7GB before CIR processing; 9.3GB after).

4.3.4 Model Training

We train several CNN models, including ResNet-18, ResNet-34, ResNet-50, and In-lab architectures, to provide location-based authentication decisions utilizing CIR data. The training focuses on reducing loss and improving location authentication accuracy, enabling the models to efficiently learn spatial and temporal features relevant to reliable location verification and trust establishment.

4.3.4.1 Model Loss and Accuracy Over Time

During training, we monitor loss and authentication accuracy curves to assess the model's convergence and stability for location verification tasks. While authentication accuracy describes the model's ability to correctly verify location-based trust across spatial areas, the loss function measures the difference between predicted and true location labels. Over several epochs, we assess these measures to find the early stopping point that balances generalization with training efficiency for robust authentication performance.

4.3.4.2 Hardware Specifications, and Computational Requirements

We employ a high-performance computing system with an NVIDIA RTX A6000 GPU, 48GB GPU memory, and third-generation tensor cores TF32, significantly accelerating the authentication model training process and providing up to 5X training throughput. The system runs on CUDA version 12.4 and executes authentication experiments using the TensorFlow framework with the Keras API. We train each authentication model for a fixed number of epochs, with the duration depending on the model architecture and dataset size. This hardware configuration was particularly critical for deeper CNN authentication models, significantly reducing training time and increasing convergence rates for reliable location verification systems.

All authentication experiments are conducted on a Dell Precision 7920 Tower Desktop, configured with 384GB RAM, 4.1 TB disk capacity, and an Intel Xeon Gold 625 CPU @ 3.90GHz (32 cores), running Ubuntu 22.04. This configuration ensures reproducible authentication evaluation results and supports the computational demands of comprehensive location verification analysis.

4.3.4.3 Hyperparameter Settings for Authentication Models

We configure the CNN authentication models with the following hyperparameters to ensure optimal convergence and generalization for location verification tasks. We use a batch size of 64, a learning rate of 1×10^{-4} with a cosine decay schedule, and train the models for 100 epochs. We employ the Adam optimizer with default $\beta_1 = 0.9$ and $\beta_2 = 0.999$ and use categorical cross-entropy as the loss function for location authentication decisions.

Weight initialization follows the He-normal distribution, and ReLU activations are applied after each convolutional layer to enable robust feature extraction for authentication. We set trace length L and stride w to 288, and we train on 80% and validate on 20% of our authentication dataset, ensuring robust model evaluation for location verification performance.

4.3.5 Model Evaluation

We evaluate the optimized authentication models using various performance indicators, including authentication accuracy and classification reports across different threat scenarios. The best-performing models demonstrate high authentication reliability across diverse transmitter and receiver configurations at different locations, confirming the effectiveness of our Butterworth filtering and data preprocessing methods for location-based trust establishment.

By analyzing training patterns and tuning hyperparameters, we ensure that the CNN models effectively capture the spatial and temporal attributes of the CIR data that are essential for accurate and reliable location-based authentication. This comprehensive evaluation demonstrates the practical feasibility of our authentication approach across a wide range of deployment scenarios and environmental conditions.

Condition	Diff. Rx – Diff. Tx				Same Rx – Diff. Tx			
	ResNet-18	ResNet-34	ResNet-50	InLab	ResNet-18	ResNet-34	ResNet-50	InLab
ADA-Butter-ReLU	63.9 ± 9.0	84.5 ± 12.9	56.1 ± 14.2	86.7 ± 11.5	83.3 ± 13.4	80.2 ± 2.0	57.0 ± 5.2	84.9 ± 4.7
Fine-ReLU	66.6 ± 0.2	66.6 ± 0.2	75.5 ± 10.9	66.6 ± 0.2	67.3 ± 9.4	69.3 ± 7.1	75.5 ± 2.0	76.3 ± 2.4
Fine-Linear	66.8 ± 0.1	66.7 ± 0.2	73.4 ± 13.3	66.7 ± 0.2	73.7 ± 4.6	71.5 ± 4.0	78.1 ± 4.2	75.1 ± 1.4
Fine-Butter-ReLU	66.7 ± 0.2	66.5 ± 0.3	80.0 ± 16.3	66.7 ± 0.2	89.3 ± 6.5	88.8 ± 7.2	91.6 ± 4.0	94.1 ± 1.1

Table 4.8: Mean ± standard deviation of classification accuracy (%) for each CNN model and condition in both "Different Rx – Different Tx" and "Same Rx – Different Tx" scenarios in the devices' indoor experiments.

4.4 Performance Evaluation and Results

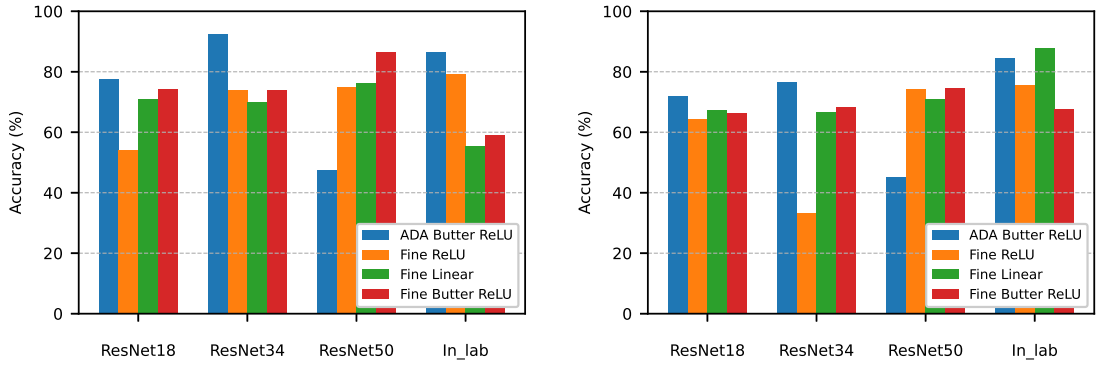


Figure 4.14: Bar chart showing classification accuracy (%) for each CNN model and condition in both "Different Rx – Different Tx" and "Same Rx – Different Tx" settings from the outdoor device experiments.

Condition	Same Rx – Same Tx				Same Rx – Diff. Tx			
	ResNet-18	ResNet-34	ResNet-50	InLab	ResNet-18	ResNet-34	ResNet-50	InLab
ADA-Butter-ReLU	59.8 ± 21.1	80.0 ± 8.1	55.4 ± 11.9	77.2 ± 12.5	93.1 ± 3.8	60.0 ± 11.0	66.2 ± 16.1	60.1 ± 4.8
Fine-ReLU	65.4 ± 3.7	53.7 ± 8.8	72.3 ± 0.6	58.1 ± 4.9	61.6 ± 4.6	50.5 ± 5.8	77.1 ± 0.9	49.6 ± 5.3
Fine-Linear	94.6 ± 0.6	33.2 ± 0.1	66.3 ± 4.5	55.6 ± 0.5	43.9 ± 3.4	54.9 ± 2.5	64.1 ± 6.4	52.8 ± 2.6
Fine-Butter-ReLU	81.5 ± 2.0	57.0 ± 7.8	93.1 ± 1.6	38.0 ± 6.6	92.9 ± 0.4	58.7 ± 8.4	43.7 ± 6.7	63.0 ± 2.7

Table 4.9: Mean ± standard deviation of classification accuracy (%) for each CNN model and condition in both "Same Rx – Same Tx" and "Same Rx – Different Tx" settings in the indoor distance experiments.

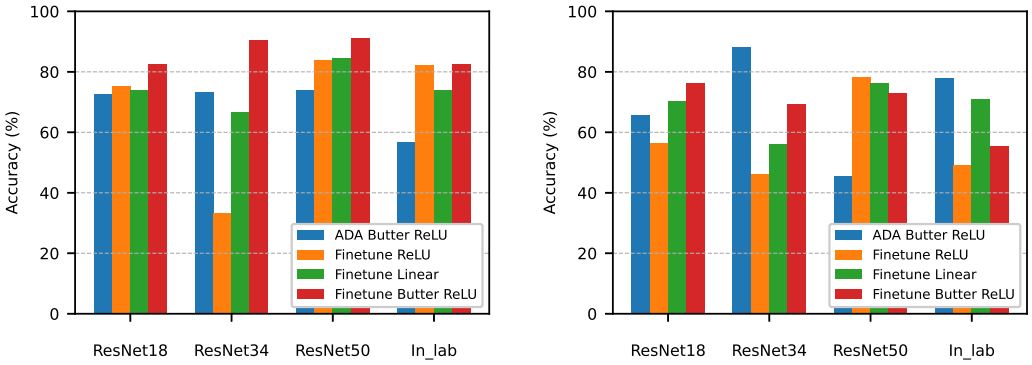


Figure 4.15: Bar chart showing classification accuracy (%) for each CNN model and condition in both "Same Rx – Same Tx" and "Same Rx – Different Tx" settings from the outdoor distance-based experiments.

4.4.1 Evaluation Metrics

4.4.1.1 Location Ranking for Authentication Performance Evaluation

We present location ranking as an additional metric to accuracy to assess the efficiency of location-based authentication. Authentication accuracy measures the percentage of correctly verified CIR traces; it does not show how well the authentication system ranks the correct location among all authorized candidates. On the other hand, location rank reflects the true location in the arranged list of prediction scores over several CIR traces, providing insight into authentication confidence levels.

For a collection of CIR traces associated with a transmitter at a fixed authorized location, the authentication system generates score vectors for each trace: $(s_{i,1}, ..., s_{i,|\mathfrak{L}|})$, where $|\mathfrak{L}|$ is the total number of authorized locations and $s_{i,j}$ is the predicted confidence for location ℓ_j on the i-th trace. These scores are aggregated across n' traces

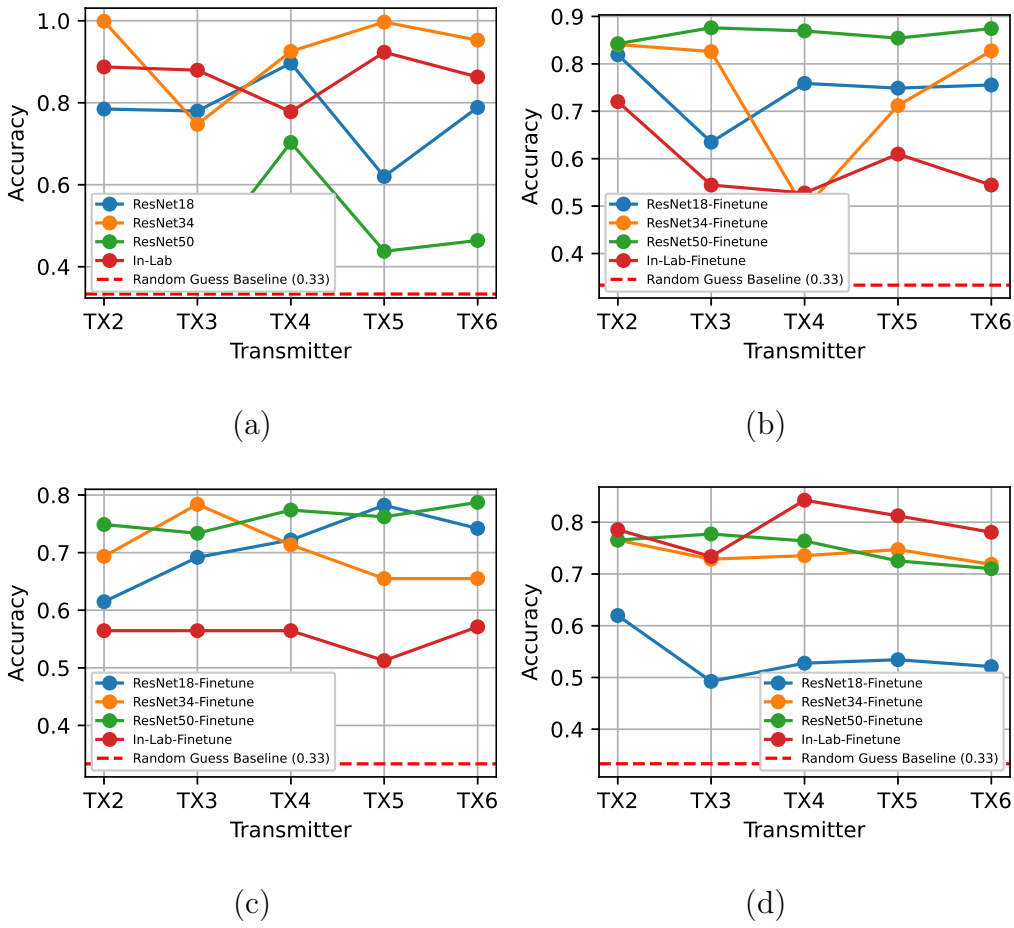


Figure 4.16: Outdoor experiments using different CNN-based models for authentication under varying transmitter–receiver pairs. Plots show accuracy versus transmitters for: (a) ADA models with ReLU activation and Butterworth filtering in different-receiver, different-transmitter settings, (b) fine-tuning with ReLU activation and Butterworth filtering, (c) fine-tuning with Linear activation and no filtering, and (d) fine-tuning with ReLU activation and no filtering, all in different-receiver, different-transmitter settings.

given as

$$s_k = \sum_{i=1}^{n'} s_{i,k}, \quad \text{for } 1 \leq k \leq |\mathfrak{L}|, \tag{4.22}$$

The aggregated score vector is then arranged in decreasing order. The true location's

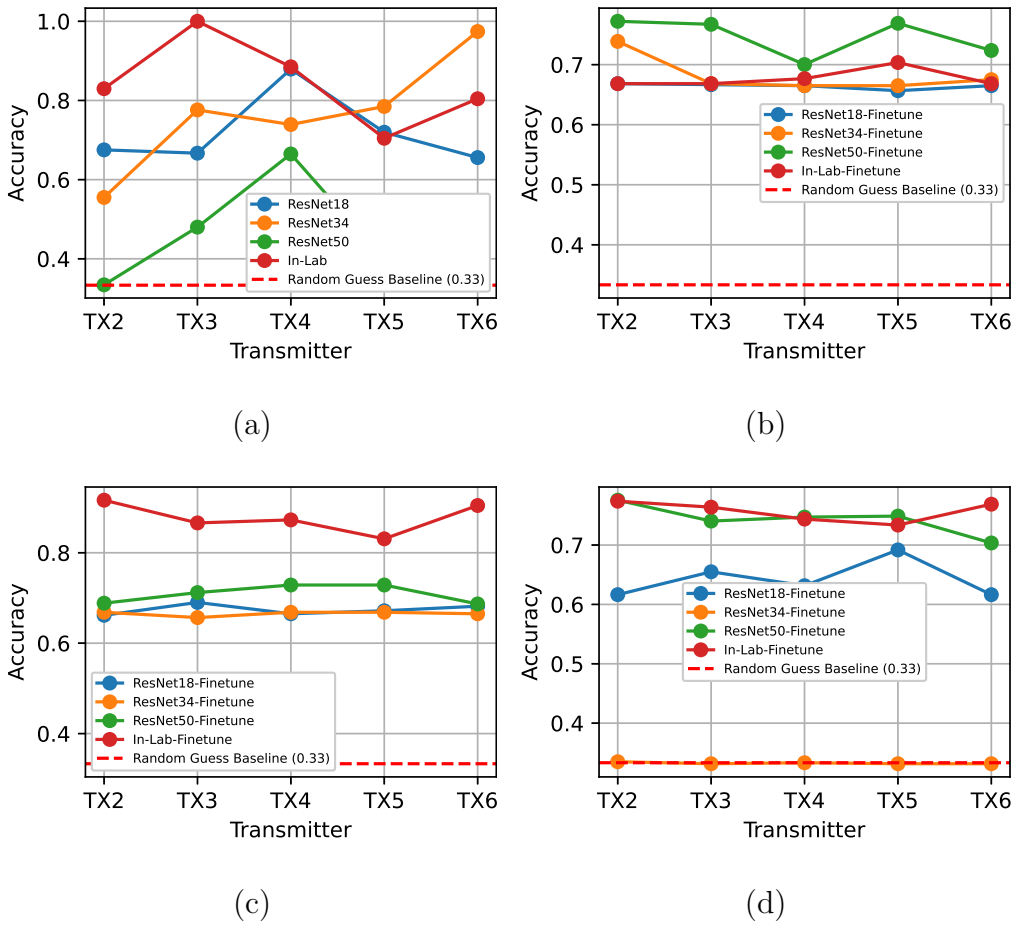


Figure 4.17: Outdoor experiments using different CNN-based models for authentication under varying transmitter–receiver pairs. Plots show accuracy versus transmitters for: (a) ADA models with ReLU activation and Butterworth filtering for the same receiver and different transmitter setting, (b) fine-tuning with ReLU activation and Butterworth filtering, (c) fine-tuning with Linear activation and no filtering, and (d) fine-tuning with ReLU activation and no filtering, all under the same receiver and different transmitter setting.

authentication rank is defined as its index in the sorted score list. A rank of 1 indicates that the authentication system identifies the correct authorized location as the top match based on cumulative evidence from multiple traces.

We find the average location rank to evaluate authentication performance over all

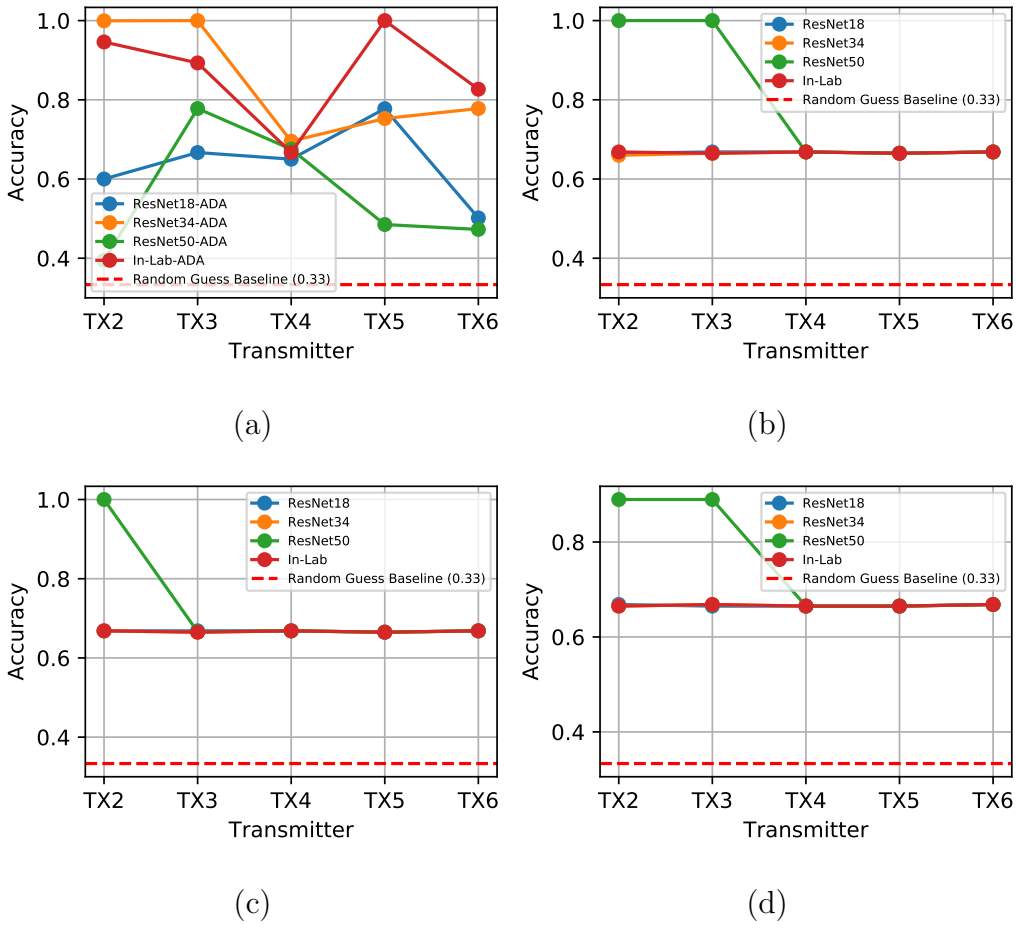


Figure 4.18: Indoor experiments using different CNN-based models for authentication under varying transmitter–receiver pairs. Plots show accuracy against transmitters for: (a) ADA with ReLU activation and Butterworth filtering, (b) Fine-tuning with ReLU activation and Butterworth filtering, (c) Fine-tuning with Linear activation and no filtering, (d) Fine-tuning with ReLU activation and no filtering — all for different receiver and different transmitter settings.

authorized sites, which is given as

$$r_{\text{avg}} = \frac{1}{|\mathfrak{L}|} \sum_{j=1}^{|\mathfrak{L}|} r_j, \tag{4.23}$$

Location rank is especially useful in low-SNR or adversarial scenarios when individual authentication decisions may be uncertain. While single trace authentication

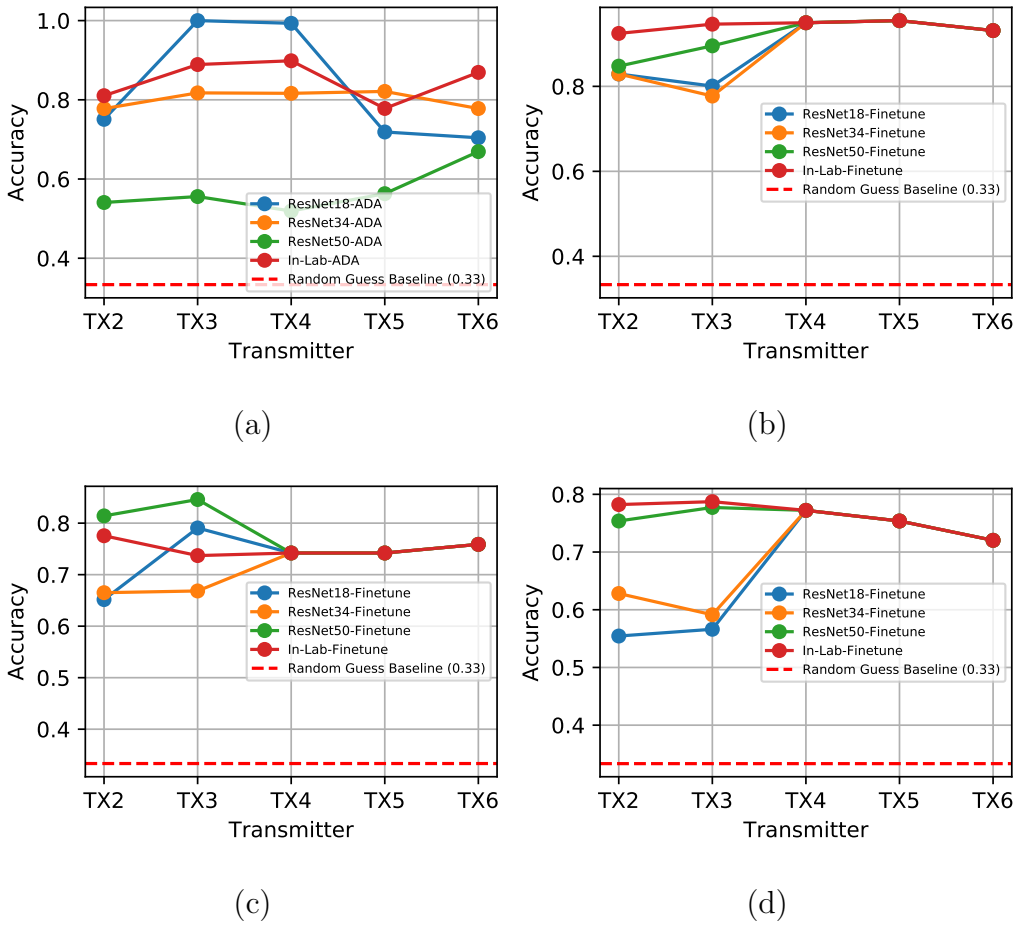


Figure 4.19: Indoor experiments using different CNN-based models for authentication under varying transmitter–receiver pairs. Plots show accuracy versus transmitters for: (a) ADA models with ReLU activation and Butterworth filtering for the same receiver and different transmitter setting, (b) fine-tuning with ReLU activation and Butterworth filtering, (c) fine-tuning with Linear activation and no filtering, and (d) fine-tuning with ReLU activation and no filtering, all under the same receiver and different transmitter setting.

accuracy may be limited, a low average location rank (near 1) suggests that the authentication system frequently ranks the correct authorized location highly. This allows for accurate location verification using a short sequence of CIR traces.

In our experiments, we observe that location rank converges rapidly. For example, even with suboptimal per-trace authentication accuracy in the OTA setting, the true

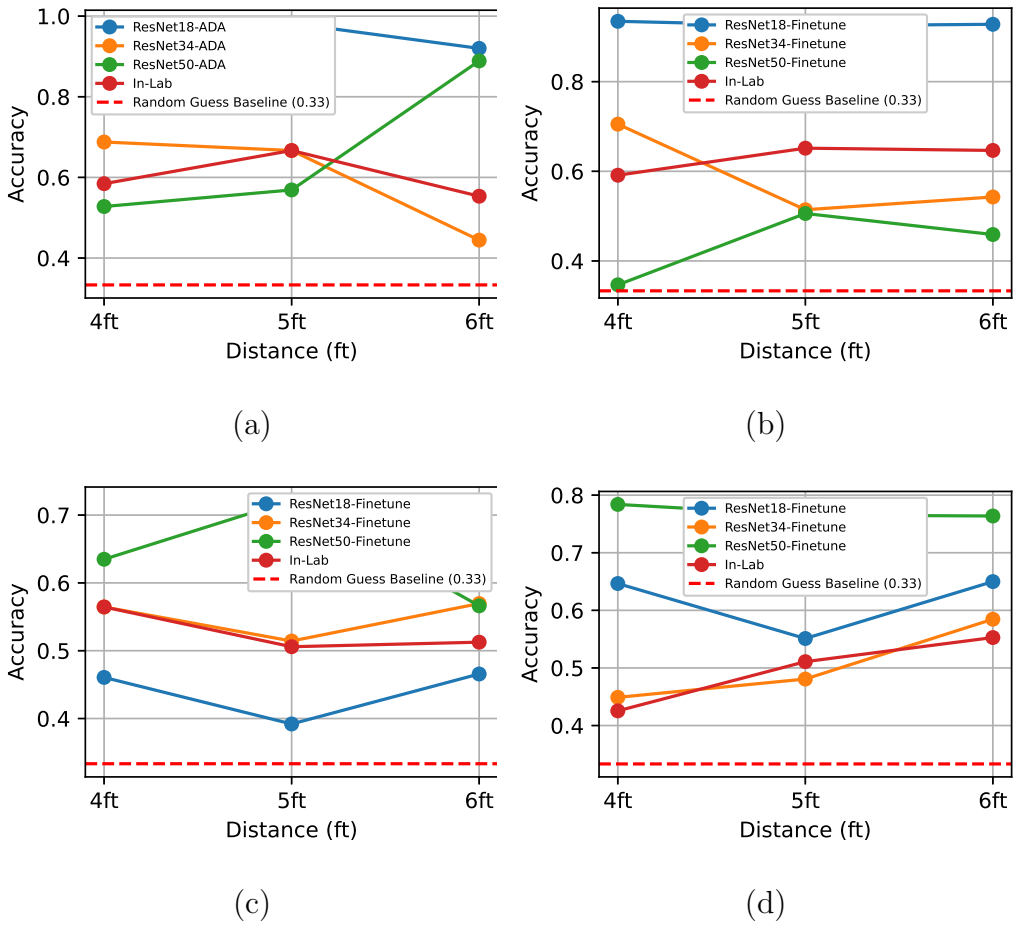


Figure 4.20: Indoor experiments using different CNN-based models for authentication under varying transmitter–receiver pairs. Plots show the accuracy against distances for: (a) ADA with ReLU activation and Butterworth filtering, (b) Fine-tuning with ReLU activation and Butterworth filtering, (c) Fine-tuning with Linear activation and no filtering, (d) Fine-tuning with ReLU activation and no filtering — all for the same receiver and different transmitter setting.

authorized location is often ranked in the top 3 after only a few CIR traces, demonstrating the utility of location rank as a robust and confidence-aware performance metric for location authentication systems.

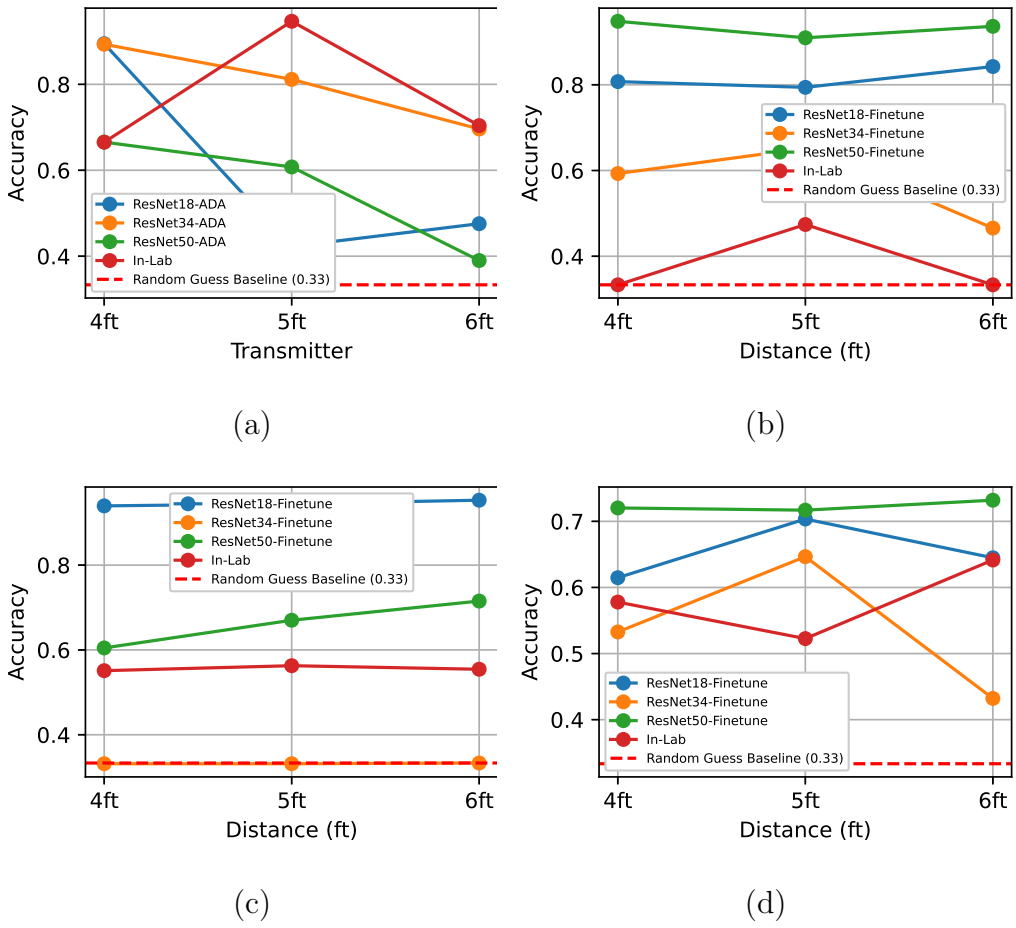
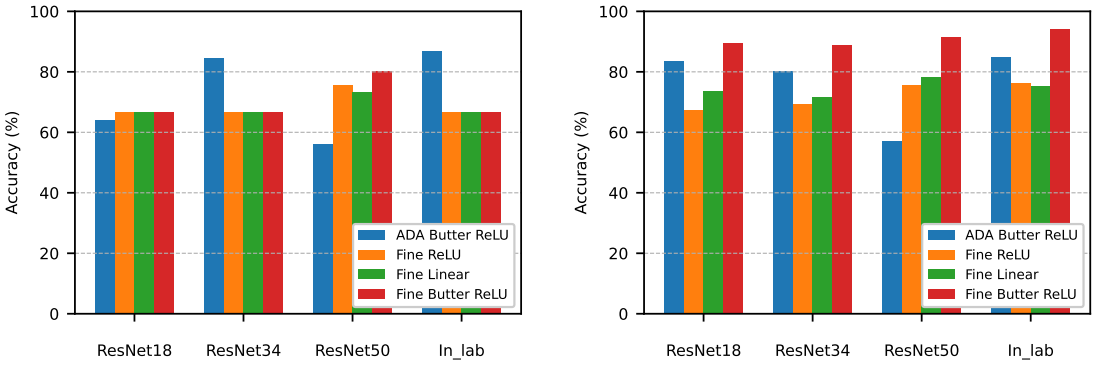


Figure 4.21: Indoor experiments using different CNN-based models for authentication under varying transmitter–receiver pairs. Plots show accuracy versus distances for: (a) ADA with ReLU activation and Butterworth filtering for same receiver and same transmitter setting, (b) fine-tuning with ReLU activation and Butterworth filtering, (c) fine-tuning with Linear activation and no filtering, and (d) fine-tuning with ReLU activation and no filtering, all in same receiver and same transmitter settings.

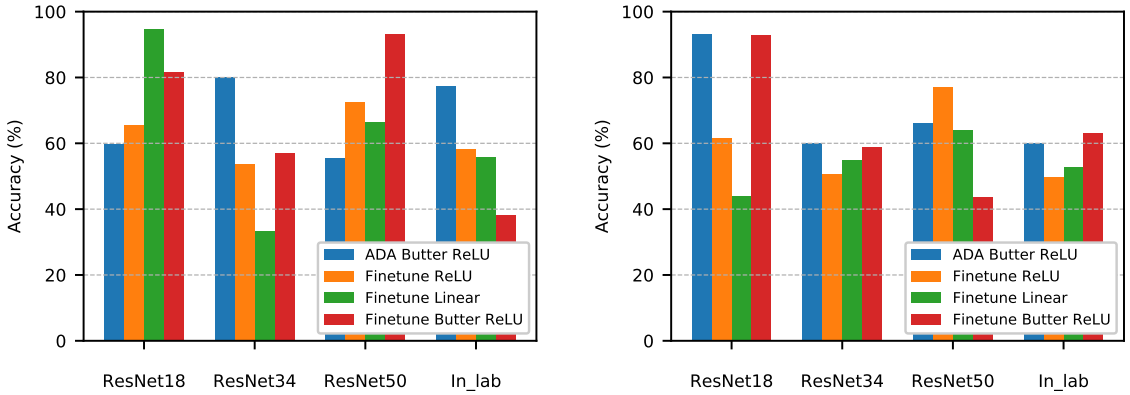
4.4.1.2 Accuracy

Given a training dataset $\mathcal{D}_{\text{train}} = \{(\mathbf{x}^{(1)}, \ell^{(1)}), \ldots, (\mathbf{x}^{(N)}, \ell^{(N)})\}$, where each input trace $\mathbf{x}^{(i)} \in \mathcal{X}$ is labeled with a ground truth authorized location $\ell^{(i)} \in \mathfrak{L}$, we train an authentication model $\mathcal{F}$ to map CIR traces to their corresponding authorized location zones.



(a) (b)

Figure 4.22: Bar chart showing classification accuracy (%) for each CNN model and condition in both "Different Rx – Different Tx" and "Same Rx – Different Tx" settings in the devices' indoor experiments.



(a) (b)

Figure 4.23: Bar chart showing classification accuracy (%) for each CNN model and condition in both "Same Rx – Same Tx" and "Same Rx – Different Tx" settings in the indoor distance experiments.

During the authentication evaluation phase, our system is tested on a separate dataset $\mathcal{D}_{\text{test}} = \{(\mathbf{x}'^{(1)}, \ell'^{(1)}), \ldots, (\mathbf{x}'^{(N')}, \ell'^{(N')})\}$. For each test trace $\mathbf{x}'^{(i)}$, our authentication system produces a confidence vector $\boldsymbol{\pi}_i = (\pi_{i,1}, \ldots, \pi_{i,|\mathfrak{L}|})$, where $\pi_{i,j}$ is the predicted confidence score for authorized location ℓ_j.

Our authentication decision is correct if the highest-scoring label matches the true

authorized location. For instance, $\ell'^{(i)} = \ell_j$ where $\pi_{i,j} = \max_k \pi_{i,k}$. Let m' denote the number of correctly authenticated test traces among the N' total traces. Then, the authentication accuracy is computed as

$$\text{Authentication Accuracy} = \frac{m'}{N'}$$

As a baseline, we set the random guess probability to $\frac{1}{M} = \frac{1}{3} \approx 0.33$, since we have $M = 3$ distinct authorized location zones. This benchmark provides a lower bound for authentication system performance under uniform random prediction.

4.4.2 Outdoor Experimental Results

We evaluate the authentication performance of the CNN models introduced in Section 4.1 using datasets collected in outdoor settings. We conduct two main types of authentication evaluations: device-based and distance-based experiments. In device-based authentication experiments, we evaluate scenarios involving transmitter and receiver USRP device variations. Specifically, we investigate (1) authentication scenarios where both transmitter and receiver devices differ and (2) authentication scenarios with a constant receiver and varying transmitters while maintaining a fixed separation of 5ft. For distance-based authentication experiments, we evaluate scenarios involving non-changing transmitter and receiver devices and scenarios where only transmitters vary across different distances.

4.4.2.1 Device-Based Evaluation

In a different receiver-different transmitter authentication scenario (Fig. 4.14(a), Fig. 4.16(a), and Table 4.6), ADA significantly improves authentication accuracy beyond random guess levels, especially for ResNet-18, ResNet-34, and the in-lab model

using Butterworth filtering and ReLU activation. ResNet-34 particularly obtains the best authentication performance ($92.4\% \pm 9.3\%$), highlighting ADA's power in controlling significant domain shifts by promoting domain-invariant learning for robust location authentication. Deeper models like ResNet-50, however, show modest authentication gains from ADA with performance drops up to $47.4\% \pm 12.0\%$, presumably due to memorization of domain-specific characteristics and limited dataset size, suggesting that increasing the training data may enhance authentication performance through improved generalization across spatial variation of the CIR data.

Fine-tuned networks using Butterworth filtering and ReLU activation as shown in Fig. 4.16(b) and Fig. 4.14(b) consistently outperform other variants for location authentication, affirming the joint significance of applying filtering and nonlinear activations. These fine-tuned models show better spatial and temporal stability than ADA models for authentication tasks. Linear activation models without filtering in Fig. 4.14(b) and Fig. 4.16(c,d) also perform reliably for authentication, underscoring fine-tuning's capability to identify and leverage dominant, generalizable signal patterns essential for robust location verification.

In the same receiver-different transmitter authentication scenario as shown in Fig. 4.14(a) and Fig. 4.17(a), lighter models such as ResNet-18 and ResNet-34 outperform deeper models like ResNet-50 under ADA, with ResNet-34 achieving $76.6\% \pm 13.3\%$ authentication accuracy, benefiting from regularization effects. However, fine-tuned models generally surpass ADA models for authentication tasks, with notable accuracy improvements. Particularly, fine-tuning combined with Butterworth filtering remains the most effective for location authentication, consistently delivering stable accuracy and resilience against minimal domain shifts as shown in Fig. 4.17(b,c,d). Fine-tune linear excels in the in-lab authentication scenarios with accuracy up to $87.8\% \pm 3.0\%$, underscoring the scenario-dependent effectiveness of linear activation

for location verification, while fine-tune ReLU performs poorly for ResNet-34 with authentication accuracy of $33.3\% \pm 0.1\%$, illustrating potential model overfitting or inability to generalize well for authentication under that setting.

4.4.2.2 Distance-Based Evaluation

The results of the distance-based authentication evaluation are summarized in Fig. 4.15, Fig. 4.12, Fig. 4.13, and Table 4.7. In authentication scenarios involving the same receiver with different transmitters, shown in Fig. 4.15, ADA models exhibit robust and consistent authentication performance across spatial separations, particularly with shallower networks like ResNet-18 and ResNet-34, reaffirming ADA's domain invariant learning strength for location authentication. However, authentication performance significantly varied with ADA models in same receiver-different transmitter settings; for instance, the In-lab model achieves authentication accuracy of $77.9\% \pm 18.1\%$, reflecting sensitivity to location variations in authentication scenarios.

Fine-tuned models with Butterworth filtering and ReLU activation, shown in Fig. 4.15, achieve good authentication accuracy and consistency with ResNet-18 reaching up to $76.4\% \pm 5.7\%$, confirming the substantial benefits of frequency-domain smoothing and nonlinear feature extraction for robust location authentication.

Conversely, models trained with ReLU activation and no filtering show moderate authentication accuracy, indicating their limited ability to capture meaningful spatial features for location verification compared to filtered approaches, reinforcing that filtering contributes more significantly to spatial discrimination for authentication than activation choice alone. Our baseline in-lab model, despite performing very well sometimes, can be unstable for authentication tasks, which emphasizes the benefit of ResNet residual learning capability with skip connections to mitigate the vanishing

gradient effect and enable better gradient flow and stable feature extraction for reliable authentication, even in deeper signal hierarchies.

In the same receiver-same transmitter authentication scenario, ADA preserves competitive authentication accuracy, though the gains are smaller due to reduced domain shifts. Fine-tuned models retain superior authentication performance across all configurations, with ReLU activation and Butterworth filtering producing the most consistent authentication accuracy, underscoring the importance of filtering and fine-tuning for reliable location verification.

Butterworth filtering greatly increases authentication model stability and robustness over different spatial configurations from the outdoor experiments for both distance and device scenarios. Fine-tuning paired with Butterworth filter and ReLU activation usually produces the most robust and reliable authentication performance across experiments. Larger ResNet models, such as ResNet-50, tend to underperform in some authentication scenarios, likely due to insufficient training data. ADA models show authentication advantages mostly in settings with notable domain shifts, but their effectiveness diminishes in stable authentication conditions.

4.4.2.3 Outdoor Denoised Experimental Results

Distance-Based Evaluation The denoised outdoor distance-based results significantly improve model stability and performance consistency. As shown in Table 4.10, Figure 4.26, Figure 4.24, and Figure 4.25, the denoised preprocessing effectively enhances authentication reliability across spatial variations.

In the same receiver-same transmitter configuration, finetune ReLU achieves exceptional performance with ResNet-50 reaching 89.8% $\pm$ 2.6% accuracy, representing a substantial improvement over ADA and fine-tuned linear results. The denoising

process particularly benefits deeper architectures, with ResNet-50 consistently outperforming shallower networks. ADA-ReLU shows robust performance with ResNet-18, achieving 79.8% ± 11.9% accuracy, demonstrating the effectiveness of domain adaptation when combined with denoising techniques.

The denoised results show more balanced performance across architectures for the same receiver—different transmitter settings. Fine-ReLU maintains strong authentication accuracy with the in-lab model reaching 81.5% ± 4.4%, while Fine-Linear demonstrates consistent performance across ResNet architectures with ResNet-50 achieving 75.8% ± 1.9% accuracy. The reduced variance in denoised results indicates improved model stability and enhanced generalization capabilities for location authentication.

Device-Based Evaluation The outdoor device-based denoised results, presented in Table 4.11, Figure 4.29, Figure 4.27, and Figure 4.28, show significant performance improvements across transmitters. In different receiver-different transmitter scenarios, Fine-Linear with ResNet-50 achieves 84.5% ± 1.2% accuracy, demonstrating the robust combination of denoising and linear activation for authentication tasks. ADA-ReLU shows competitive performance with the in-lab model reaching 84.9% ± 11.1%, highlighting the effectiveness of domain adaptation in challenging device variation scenarios.

For the same receiver—different transmitter configurations, Fine-ReLU demonstrates exceptional performance with the in-lab model achieving 84.6% ± 6.2% accuracy and ResNet-50 reaching 81.5% ± 2.1%. The denoising process significantly reduces the impact of hardware-induced variations, enabling more reliable location-based authentication across different transmitter devices while maintaining receiver consistency.

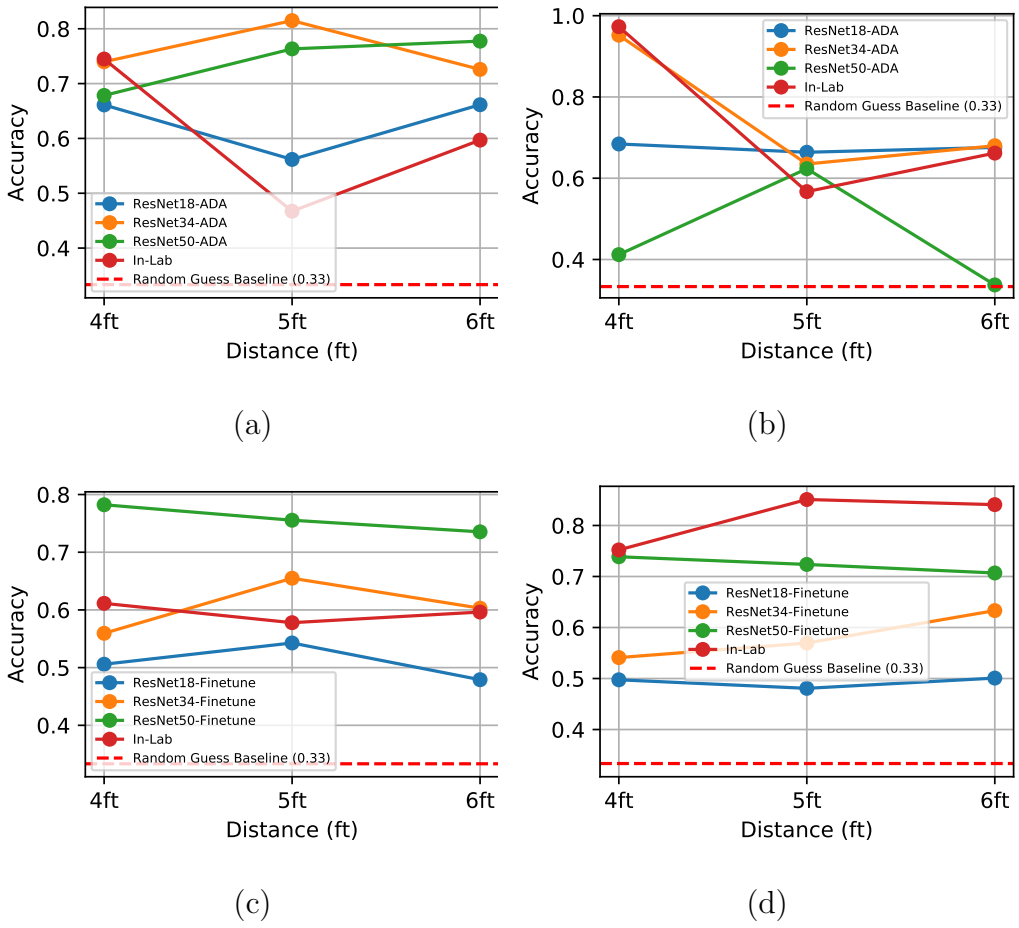


Figure 4.24: Outdoor distance-based denoised results for CNN-based authentication models showing accuracy versus distance. Plots compare different model configurations: (a) ADA Vanilla with Linear activation, same RX–different TX, (b) ADA Vanilla with ReLU activation, same RX–different TX, (c) Fine-tuned Vanilla with Linear activation, same RX–different TX, (d) Fine-tuned Vanilla with ReLU activation, same RX–different TX.

4.4.3 Indoor Experimental Results

4.4.3.1 Device-Based Evaluation

We summarized the indoor device-based authentication results in Fig. 4.22, Fig. 4.18, Fig. 4.19, and Table 4.8. In the Different RX-Different TX authentication scenario, ADA models benefit most from domain adaptation when retaining Butterworth fil-

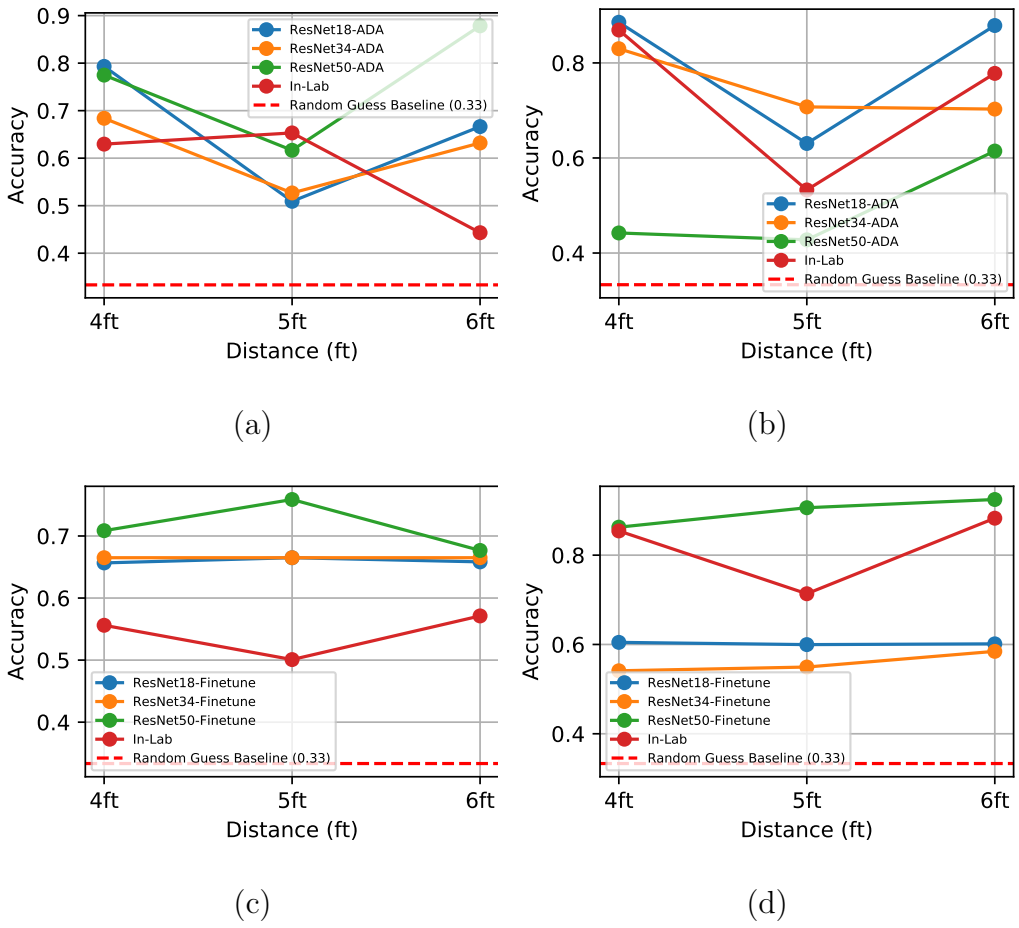


Figure 4.25: Outdoor distance-based denoised results for CNN-based authentication models showing accuracy versus distance. Plots compare different model configurations: (a) ADA Vanilla with Linear activation, same RX–same TX, (b) ADA Vanilla with ReLU activation, same RX–same TX, (c) Fine-tuned Vanilla with Linear activation, same RX–same TX, (d) Fine-tuned Vanilla with ReLU activation, same RX–same TX. All results reflect denoised performance in outdoor experimental conditions with varying transmitter–receiver pair configurations.

tering and ReLU activation, with ResNet-34 reaching $84.5\% \pm 12.9\%$ authentication accuracy and the lightweight in-lab architecture peaking at $86.7\% \pm 11.5\%$. Fine-tune-ReLU and Fine-tune-Linear without filtering hover around the mid 60% range for authentication accuracy on ResNet-18/34 and improve only when depth increases in ResNet-50, which ranges from 73–76% authentication performance. Adding Butter-

Condition	Same Rx – Same Tx				Same Rx – Diff. Tx			
	ResNet-18	ResNet-34	ResNet-50	InLab	ResNet-18	ResNet-34	ResNet-50	InLab
ADA-Linear	65.6 ± 11.6	61.4 ± 6.5	75.7 ± 10.8	57.5 ± 9.4	62.8 ± 4.7	76.0 ± 3.9	74.0 ± 4.4	60.3 ± 11.3
ADA-ReLU	79.8 ± 11.9	74.7 ± 5.9	49.5 ± 8.5	72.7 ± 14.2	67.5 ± 0.8	75.6 ± 14.0	45.8 ± 12.1	73.4 ± 17.3
Fine-Linear	66.0 ± 0.4	66.5 ± 0.0	71.5 ± 3.4	54.3 ± 3.0	50.9 ± 2.6	60.6 ± 3.9	75.8 ± 1.9	59.5 ± 1.4
Fine-ReLU	60.2 ± 0.2	55.8 ± 1.9	89.8 ± 2.6	81.7 ± 7.4	49.3 ± 0.9	58.1 ± 3.9	72.3 ± 1.3	81.5 ± 4.4

Table 4.10: Mean ± standard deviation of classification accuracy (%) for each CNN model and condition in both "Same Rx – Same Tx" and "Same Rx – Different Tx" scenarios in the outdoor distance experiments on denoised CIR.

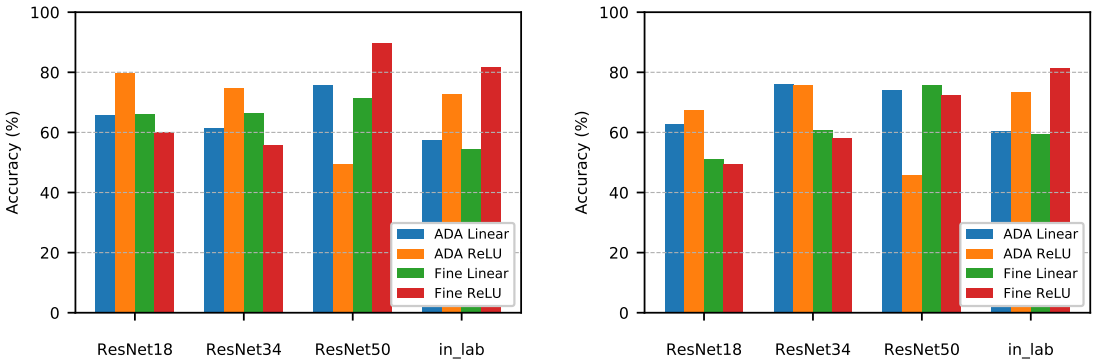


Figure 4.26: Bar chart showing classification accuracy (%) for each CNN model and condition in both "Same Rx – Same Tx" and "Same Rx – Different Tx" settings in the distances' outdoor experiments using denoised CIR data.

Condition	Diff. Rx – Diff. Tx				Same Rx – Diff. Tx			
	ResNet-18	ResNet-34	ResNet-50	InLab	ResNet-18	ResNet-34	ResNet-50	InLab
ADA-Linear	78.2 ± 7.5	54.6 ± 9.2	76.2 ± 9.4	50.8 ± 8.0	75.4 ± 14.1	45.3 ± 4.6	64.6 ± 7.8	44.7 ± 4.7
ADA-ReLU	78.4 ± 7.1	81.5 ± 14.5	50.9 ± 8.8	84.9 ± 11.1	72.1 ± 18.8	77.2 ± 10.8	59.1 ± 8.4	68.6 ± 13.3
Fine-Linear	64.5 ± 4.1	69.9 ± 6.7	84.5 ± 1.2	61.9 ± 3.5	77.8 ± 5.5	66.7 ± 0.2	83.3 ± 1.4	79.4 ± 12.4
Fine-ReLU	74.4 ± 10.2	38.0 ± 9.2	68.9 ± 2.2	48.0 ± 2.5	67.0 ± 0.9	62.6 ± 1.8	81.5 ± 2.1	84.6 ± 6.2

Table 4.11: Mean ± standard deviation of classification accuracy (%) for each CNN model and condition in both "Different Rx – Different Tx" and "Same Rx – Different Tx" scenarios in the outdoor device experiments on denoised CIR.

worth filtering at fine-tune time narrows the authentication performance gap, letting ResNet-50 break 80% authentication accuracy (80.0% ± 16.3%), but depth-shallow variants remain capped near the low-70s for authentication tasks. Overall, adaptation plus filtering is essential for robust authentication when both transmitter and

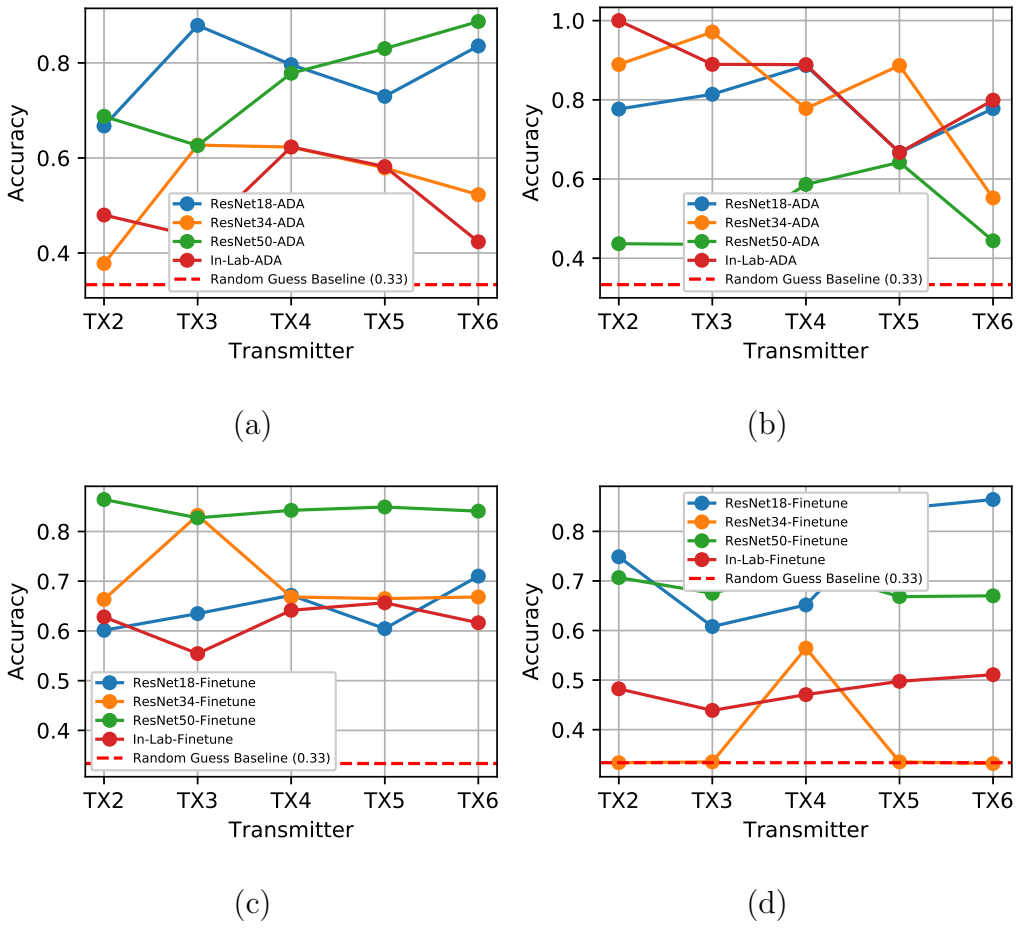


Figure 4.27: Outdoor device-based denoised results for CNN-based authentication models. Plots compare different model configurations: (a) ADA Vanilla with Linear activation, different RX–different TX, (b) ADA Vanilla with ReLU activation, different RX–different TX, (c) Fine-tuned Vanilla with Linear activation, different RX–different TX, (d) Fine-tuned Vanilla with ReLU activation, different RX–different TX.

receiver positions vary; depth alone cannot compensate for channel-transfer mismatch in authentication scenarios.

In the same RX-Different TX authentication setting, fine-tuned networks combining Butterworth filtering and ReLU activation dominate authentication performance. ResNet-50 achieves $91.6\% \pm 4.0\%$ authentication accuracy, ResNet-34 reaches $88.8\% \pm 7.2\%$, and the in-lab model attains the overall best authentication perfor-

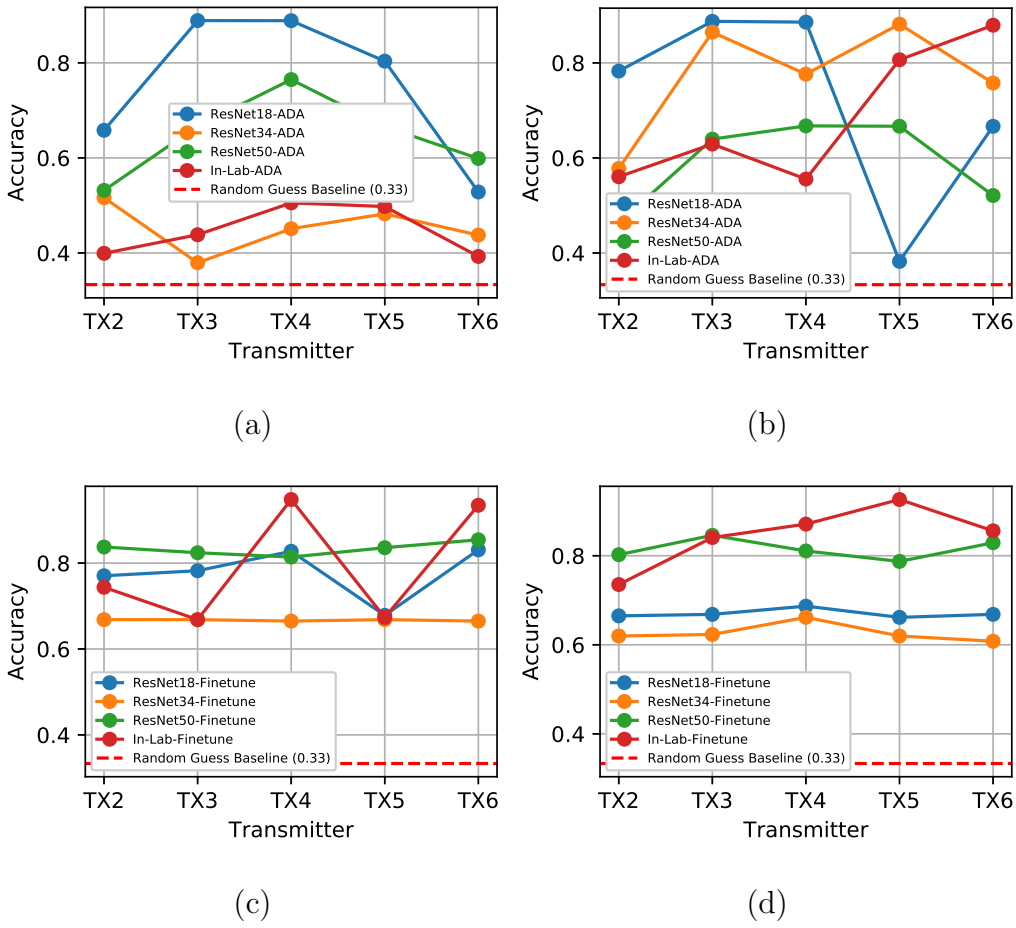


Figure 4.28: Outdoor device-based denoised results for CNN-based authentication models. Plots compare different model configurations: (a) ADA Vanilla with Linear activation, same RX–different TX, (b) ADA Vanilla with ReLU activation, same RX–different TX, (c) Fine-tuned Vanilla with Linear activation, same RX–different TX, (d) Fine-tuned Vanilla with ReLU activation, same RX–different TX. All results show denoised performance for outdoor experimental conditions with varying transmitter–receiver device pair configurations.

mance of $94.1\% \pm 1.1\%$. ADA-Butter-ReLU still performs solidly for authentication ($\approx 80-85\%$) but no longer leads, suggesting that once the receiver is fixed, the heavy domain adaptation step adds less value than lighter fine-tuning for authentication tasks. Removing Butterworth filtering (Fine-ReLU) again suppresses authentication accuracy by $10-20$ percentage points, underscoring that frequency domain smoothing

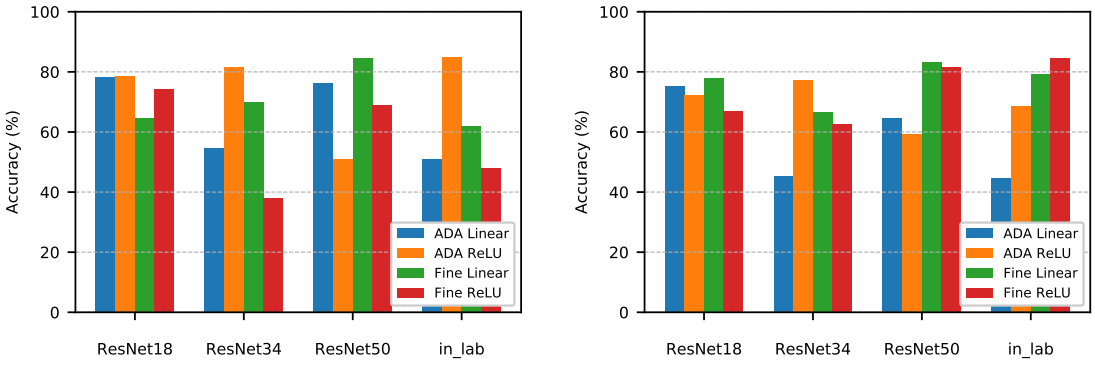


Figure 4.29: Bar chart showing classification accuracy (%) for each CNN model and condition in both "Different Rx – Different Tx" and "Same Rx – Different Tx" settings in the devices' outdoor experiments.

is the single most reliable ingredient for robust authentication across architectures. These results reaffirm that the Butterworth filter is the key determinant of authentication robustness, while the choice between ADA and fine-tuning depends on whether the receiver location variability affects authentication performance.

4.4.3.2 Distance-Based Evaluation

We summarize the indoor distance-based authentication results in Fig. 4.23, Fig. 4.20, Fig. 4.21, and Table 4.9. In the same receiver-different transmitter authentication scenario, ADA models with ReLU activation and Butterworth filtering achieve strong authentication accuracy, with ResNet-18 reaching $93.1\% \pm 3.8\%$, indicating robust adaptation to spatial changes in authentication performance. However, the same model shows reduced stability in the same receiver-same transmitter authentication settings, implying limited benefit from domain adaptation when spatial variance is minimized for authentication tasks.

Fine-tuned models incorporating Butterworth filtering and ReLU activation perform competitively across both authentication conditions. For instance, ResNet-50

attains $93.1\% \pm 1.6\%$ authentication accuracy in the same receiver-same transmitter setting, validating the effectiveness of frequency smoothing and non-linear activation for reliable location authentication. Notably, the fine-tuned linear model with ResNet-18 achieves the highest authentication accuracy of $94.6\% \pm 0.6\%$, though this peak authentication performance does not generalize consistently across other architectures.

Models without Butterworth filtering but using ReLU, such as fine-tune ReLU, usually underperform for authentication tasks, emphasizing that filtering has more impact on authentication robustness than activation choice. Although the in-lab models sometimes show great authentication accuracy up to $77.2\% \pm 12.5\%$, its variability across authentication settings confirms that deeper architectures with residual learning provide better feature consistency for reliable location authentication. These results demonstrate that Butterworth filtering generates stable and accurate models for indoor location authentication when combined with fine-tuning and suitable activation functions.

4.5 Security Analysis

In this section, we first discuss our adversary's ability to model the CIR using Friis' equation and the ray-tracing channel reconstruction method. This is followed by the robustness analysis of LAOUWN against the adversary presented in Section 4.1, focusing on the security guarantees for location-based authentication.

4.5.1 Friis' Empirical Adversary

Our Friis' empirical adversary is the baseline for evaluating our system's resilience against a low-resource adversary attempting to bypass location authentication with

minimal or no knowledge of multipath presence in a communication setting. It sets the lower bound of adversarial effectiveness against our authentication system, illustrating the optimistic scenario that neglects realistic environmental complexities. This adversary attempts to forge channel characteristics to spoof authorized locations based solely on distance estimation.

Let us denote the transmitter complex baseband signal as X and the received complex baseband signal as Y. In a legitimate authentication scenario, the relationship is:

$$Y = h_{\text{RX-TX}} \cdot X + w \tag{4.24}$$

Where $h_{\text{RX-TX}}$ represents the actual channel between the legitimate authorized transmitter (TX) and the receiver (RX), and w represents the additive noise.

The Friis-based authentication adversary attempts to impersonate the legitimate authorized transmitter by estimating the legitimate channel using only the Friis transmission equation. When the adversary transmits to bypass authentication, the system observes:

$$Y = h_{\text{RX-Adv}} \cdot X_{\text{Adv}} + w \tag{4.25}$$

Where $h_{\text{RX-Adv}}$ denotes the actual channel between the adversary and the receiver, and X_{Adv} is the adversary's transmitted signal attempting to spoof authentication.

Unlike the more sophisticated ray-tracing enhanced adversary, the Friis-based authentication adversary does not attempt to compensate for its own channel. Instead, it simply applies a Friis-based channel estimate for its transmission, hoping that the received signal will approximate the legitimate transmission and bypass authentication:

$$h_{\text{RX-Adv}} \cdot X_{\text{Adv}} \approx h_{\text{RX-TX}} \cdot X \tag{4.26}$$

To achieve authentication bypass, the adversary sets:

$$X_{\text{Adv}} = h_{\text{Friis}}(d_{\text{TX-RX}}) \cdot X \tag{4.27}$$

Where X represents a legitimate message, and $h_{\text{Friis}}(d_{\text{TX-RX}})$ is the Friis-based estimate of the legitimate channel based solely on distance.

The Friis-based authentication adversary uses the free-space path loss model to approximate the legitimate channel given as:

$$h_{\text{Friis}}(d) = \sqrt{G_t G_r} \cdot \frac{\lambda}{4\pi d} \cdot e^{-j\frac{2\pi d}{\lambda}} \tag{4.28}$$

Where G_t and G_r are the transmitter and receiver gains, λ is the wavelength, and d is the distance. The first part of equation $\sqrt{G_t G_r} \cdot \frac{\lambda}{4\pi d}$ models the attenuation of the amplitude, while the exponential term $e^{-j\frac{2\pi d}{\lambda}}$ accounts for the phase change due to the delay of propagation.

In practice, the authentication adversary faces significant challenges in accurately measuring distance $d_{\text{TX-RX}}$. Although technologies such as GPS can provide location estimates, they typically have error margins of several meters, which introduces substantial phase errors at GHz frequencies that compromise authentication spoofing attempts. Underground settings further complicate authentication attacks with variable signal propagation speeds through different soil compositions.

Beyond the channel mismatch issue, the Friis model itself represents a simplistic single-tap channel response given by

$$h_{\text{Friis}}(t) = \alpha \cdot \delta(t - \tau_0) \tag{4.29}$$

Where $\alpha = \sqrt{G_t G_r} \cdot \frac{\lambda}{4\pi d}$ is the complex amplitude derived from the Friis equation,

$\delta(t)$ is the Dirac delta function, and $\tau_0 = \frac{d}{c}$ is the propagation delay with c being the speed of light.

The Friis-based adversary's minimal knowledge approach is ineffective against our LAOUWN authentication framework. Some of these authentication security limitations are discussed below.

4.5.1.1 Ignores Adversary's Channel in Authentication Bypass

By failing to account for $h_{\text{RX-Adv}}$, the adversary introduces a systematic error in the characteristics of the received signal that compromises authentication spoofing attempts. The channel mismatch happens because the minimal knowledge Friis-based adversary ignores its own channel $h_{\text{RX-Adv}}$. The actual received signal will be

$$Y = h_{\text{RX-Adv}} \cdot h_{\text{Friis}}(d_{\text{TX-RX}}) \cdot X + w \tag{4.30}$$

Equation (4.30) is obtained by substituting (4.27) into (4.25). For successful authentication bypass, we need $h_{\text{RX-Adv}} \cdot h_{\text{Friis}}(d_{\text{TX-RX}}) \approx h_{\text{RX-TX}}$, which is highly unlikely given the complex nature of wireless channels, especially in underground environments. This mismatch provides strong authentication security against Friis-based spoofing attempts.

4.5.1.2 Single-tap Approximation Limits Authentication Spoofing

The Friis model produces a simplistic channel representation with only a direct path component, failing to capture the rich multipath structure of actual underground-to-air channels that our authentication system relies upon. While the legitimate

authorized channel consists of multiple components given as

$$h_{\mathrm{UG}}(t) = \sum_{l=0}^{L-1} \alpha_l \delta(t - \tau_l) + \sum_{d=0}^{D-1} \alpha_d \delta(t - \tau_d) + \sum_{r=0}^{R-1} \alpha_r \delta(t - \tau_r) \tag{4.31}$$

Equation (4.31) represents the lateral, direct, and reflected wave components as described in Section 4.2; the Friis model reduces this to a single component, making authentication spoofing ineffective.

4.5.1.3 Missing Soil Effects Enhance Authentication Security

The model ignores the crucial impact of soil composition, moisture, and temperature, which dramatically alter signal propagation in an underground environment and provide unique authentication signatures. These factors affect amplitude attenuation and phase velocity, which are not captured in the Friis model but are essential for our authentication security.

4.5.1.4 No Frequency-selective Fading in Authentication Attacks

Underground channels exhibit significant frequency-selective characteristics that the Friis model cannot reproduce, limiting authentication spoofing effectiveness. The model assumes flat fading across all frequencies, which is unrealistic in complex propagation environments and fails to replicate the signatures our authentication system uses.

4.5.1.5 Inability to Model Narrow Beam Patterns for Authentication Bypass

Unlike the ray-tracing approach, which models signal propagation as a large number of very narrow beams reflecting off surfaces and penetrating materials with differ-

ent properties, the Friis model treats propagation as a signal direct path without any scattering or reflection, making authentication spoofing attempts easily detectable.

4.5.1.6 Channel Metrics Comparison

The legitimate authorized power delay profile (PDP) is given by

$$\mathrm{PDP}_{\mathrm{legitimate}}(\tau) = |h_{\mathrm{UG}}(\tau)|^2 = \sum_i |\alpha_i|^2 \cdot \delta(\tau - \tau_i) \tag{4.32}$$

And the Friis-based authentication spoofing estimation is given as

$$\mathrm{PDP}_{\mathrm{Friis}}(\tau) = |h_{\mathrm{Friis}}(\tau)|^2 = |\alpha|^2 \cdot \delta(\tau - \tau_0) \tag{4.33}$$

Equation (4.32) shows the sum squared of the magnitude of the CIR for authorized locations, while Equation (4.33) shows a single tap squared of the magnitude of the CIR for spoofing attempts. This highlights that the Friis model produces a simplistic single-spike PDP that cannot bypass our authentication system. It validates why the Friis-based adversary cannot successfully impersonate a legitimate authorized location; it simply cannot reproduce the rich multipath structure that characterizes authentic location signatures.

This fundamental difference leads to dramatic disparities in key channel metrics such as RMS delay spread given by

$$\tau_{\mathrm{rms,legitimate}} = \sqrt{\frac{\sum_i P_i \tau_i^2}{\sum_i P_i} - \left(\frac{\sum_i P_i \tau_i}{\sum_i P_i}\right)^2} \gg \tau_{\mathrm{rms,Friis}} = 0 \tag{4.34}$$

Where $P_i = |\alpha_i|^2$ represents the power of each multipath component, $\tau_{\mathrm{rms,legitimate}}$ is the RMS delay of the legitimate authorized location and $\tau_{\mathrm{rms,Friis}}$ is the Friis RMS delay spread. The RMS delay spread of the Friis model is effectively zero due to

its single-tap nature, while underground channels exhibit significant delay spread, which provides strong authentication security guarantees and justifies why a Friis-based adversary will be unsuccessful in spoofing the location of legitimate authorized nodes.

These limitations directly translate into robust authentication security guarantees in our system. Even with perfect knowledge of the distance between the legitimate transmitter and receiver, this minimal-knowledge approach fails to capture the essential features LAOUWN leverages for location authentication. Therefore, a Friis-based adversary cannot successfully replicate the complex channel characteristics essential to bypass our authentication method in either an OTA or an underground scenario. Unlike the ray-tracing approach discussed in the next section, which models signal propagation as many narrow beams reflecting off surfaces and penetrating materials with different properties, the Friis model treats propagation as a single direct path without scattering or reflection. Next, we analyze the more sophisticated ray-tracing-enhanced adversary that attempts to overcome these limitations.

4.5.2 Ray-Tracing Enhanced Adversary

We consider the adversary's ability to replicate legitimate authorized channel characteristics to bypass authentication. Contrary to Friis' adversary, in the enhanced ray-trace adversary, we leverage advanced computational modeling methods via a hybrid 3D geometric ray-tracer. These techniques numerically reconstruct the channel by modeling interactions between transmitted signals and surrounding environments to create sophisticated authentication bypass attempts. We consider key parameters when simulating our adversary CIR dataset for authentication attacks, which includes log-distance path loss, multipath reflections, shadow fading, Ricean fading, delay spread, and diffuse scattering.

We implement our adversary framework, which is a hybrid 3D ray-tracer that synthesizes CIR data using a combination of deterministic specular reflections, diffuse scattering, shadow fading, Ricean K-factor balancing, and delay clusters. Our output CIR binary files represent a realistic adversarial condition. We simulate CIR data for LoS and NLoS paths. We include two scenarios for NLoS settings, including occlusion, like a wall, and shifting the adversary's position 3 meters from the LoS.

We simulate CIR data for clean, occluded attack settings. For instance, clean LoS attacks are characterized by a clear, unobstructed, direct path between the simulated transmitter and receiver; the clean NLoS attack scenario is when the direct path is obstructed, and the signals primarily propagate via reflections and scattering, simulated by shifting the adversary's receiver target position $3m$ away from the direct LoS path, and the occluded NLoS attack settings are where we explicitly introduce an additional obstacle, such as a wall, into the simulation to block dominant paths further and test authentication security. We save unique CIR data for various distances, including 10m, 20m, 30m, and 40m across indoor and outdoor settings, generating distinct files for clean LoS, clean NLoS, and occluded NLoS attack conditions.

The enhanced ray-tracing adversary embodies a sophisticated attacker capable of accurately replicating legitimate transmitter signatures for authentication bypass, representing the most challenging threat to our location authentication system. Ray tracing offers accurate environmental modeling, calculation of numerous multipath components, and diffuse scattering at the expense of substantial computational resource demand. This adversary has comprehensive knowledge of the environmental layout and propagation characteristics.

4.5.3 Adversary Setup Analysis

We conducted experiments comparing the impact of both adversarial types on our authentication method. We rigorously evaluate our methodology to assess the robustness of our proposed authentication system against sophisticated attacks. It encompasses the datasets utilized, the specific model under scrutiny, and the comprehensive software framework we developed for this authentication security analysis. The empirical Friis attacker provides a baseline attack model reflecting minimal adversarial capability. In contrast, the ray-tracing adversary represents a worst-case attack scenario.

4.5.3.1 Friis' Empirical Adversary

Our Friis' empirical adversary serves as the baseline for evaluating our authentication system's resilience against a low-resource attacker with minimal or no knowledge of multipath presence in a communication setting. This adversary attempts to forge channel characteristics based solely on distance estimation using Friis' transmission equation to bypass location authentication. It sets the lower bound of adversarial effectiveness against our authentication system, representing an optimistic scenario that neglects realistic environmental complexities. The model assumes a simplistic single tap channel response, ignoring crucial impacts of soil composition, moisture, temperature in underground environments, and frequency-selective fading that provide authentication security. These limitations mean the Friis-based adversary cannot successfully replicate the complex channel characteristics essential to bypass our authentication method in either an OTA or an underground scenario.

4.5.3.2 Ray-tracing Enhanced Adversary

For our authentication security evaluation, we leverage both real-world legitimate data and synthetically generated adversarial CIR data to test our authentication system under diverse attack conditions. The foundation of our evaluation rests on two primary datasets. The legitimate authorized dataset consists of CIR data we collected from real-world wireless transmissions across indoor, outdoor, and underground propagation environments. Our process for capturing and pre-processing authentic channel conditions using software-defined radios was discussed in Section 4.2. This dataset serves as the ground truth for training our authentication model and as the baseline for evaluating its performance under non-adversarial conditions. We organized it in a designated directory, categorized by environment and specific measurement configurations, such as varying transmitter-receiver distances of 4 ft, 5 ft, and 6 ft for authorized locations.

To stringently test our authentication framework's resilience, we use our synthetically generated adversarial CIR dataset, primarily via sophisticated hybrid 3D ray-tracing simulation designed to bypass authentication. We ensured our adversarially generated CIRs are saved in the same binary format and scaled to match the amplitude range of the real-world legitimate data, ensuring consistency for authentication attack testing. We simulated adversary transmitters at positions offset from legitimate authorized sites, targeting these locations from various distances from 10m to 40m to test authentication bypass attempts. We stored this synthetic adversarial data in a separate, structured directory compatible with our authentication evaluation framework.

We engineered a comprehensive Python-based evaluation framework to automate and standardize the assessment of the authentication model's performance under var-

ious attack conditions. This framework manages data loading, model inference, calculation of authentication performance metrics, generation of additional adversarial attacks against the authentication system, simulation of channel noise and hardware impairments, and visualization of authentication security results.

Our authentication security framework establishes a structured output directory to store all generated results, including diagnostics, detailed analyses, authentication attack outcomes, and noise study data. We employ initial diagnostic utilities to inspect the raw binary format of legitimate and adversarial CIR files in the sample, aiding in data integrity verification for authentication testing. Subsequently, we load the fine-tuned ResNet18 model for authentication, with provisions for custom objects such as accuracy metrics for authentication performance. We perform preliminary validation using purely synthetic data to confirm the operational integrity of the basic authentication prediction pipeline before proceeding to more complex authentication security evaluations.

4.6 Adversarial Robustness Evaluation Results

4.6.1 Complete Ray-Tracing Authentication Bypass Failure

The most significant findings from our authentication security evaluation are the complete failure of the sophisticated ray-tracing enhanced adversary to successfully bypass location authentication by spoofing legitimate authorized nodes' locations. Across all environmental conditions, including indoor, outdoor, and occluded scenarios, the adversarial authentication accuracy remains approximately 0.33, corresponding to random guessing performance for our three-location authentication problem, irrespective of the distance.

Figure 4.30 (a), (b), (c), and (d) demonstrate the authentication adversarial ro-

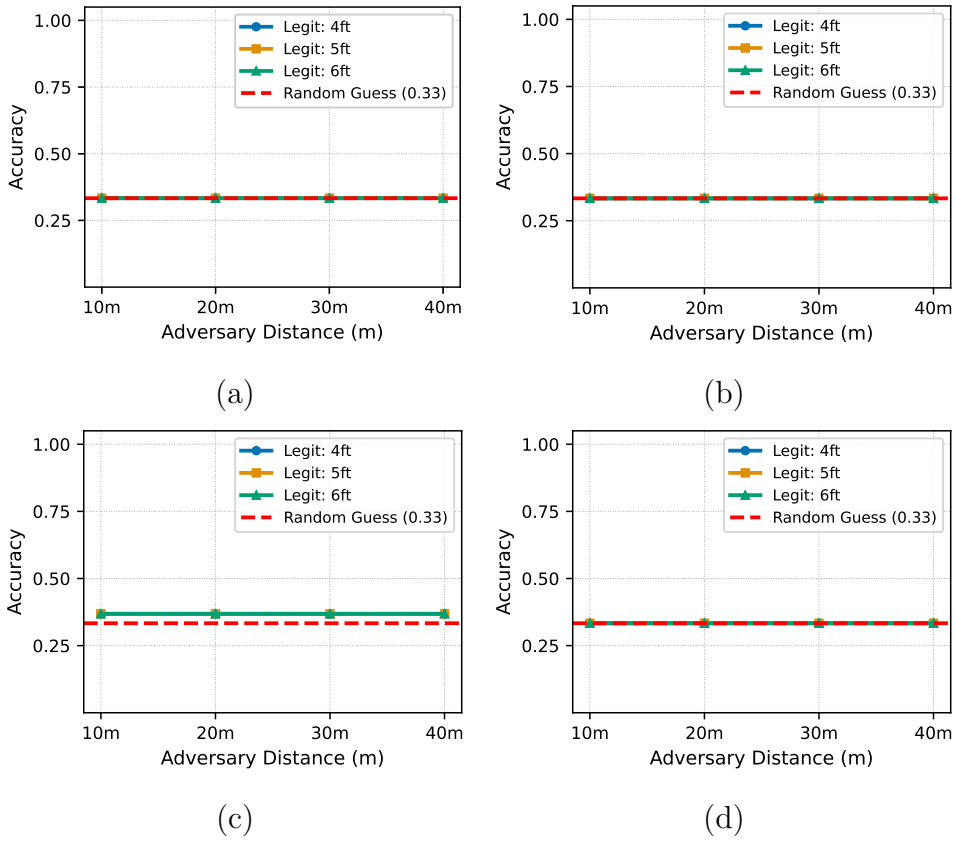


Figure 4.30: Plots of adversarial impact on authentication accuracy across different environments and conditions. Subfigures show the authentication accuracy of our location fingerprinting system against enhanced raytrace adversaries at varying distances and occlusion conditions: (a) indoor clean environment, (b) indoor occluded environment, (c) outdoor clean environment, and (d) outdoor occluded environment.

bustness results across different environmental conditions, showing remarkable consistency in authentication defense performance. The ray-tracing adversary achieves no better than random authentication accuracy across all tested scenarios. This indicates that the CNN authentication classifier treats sophisticated adversarial samples as illegitimate authentication attempts rather than misclassifying them to the wrong authorized locations.

4.6.2 Distance-Independent Security Guarantees

The adversarial robustness exhibits exceptional consistency across spatial separations from 10m to 40m. Even at the closest adversary distance of 10m, where the attacker theoretically possesses the most accurate channel knowledge and strongest signal coupling, the authentication bypass success remains at random guessing levels. These distance-independent authentication security characteristics indicate that unique spatial signatures captured by underground-to-air and OTA propagation cannot be effectively replicated through sophisticated channel modeling for authentication spoofing, even with detailed environmental knowledge.

The flat response curves across all adversary distances indicate systematic authentication security properties rather than distance-dependent vulnerabilities. This provides strong authentication security guarantees for practical deployments where adversary positioning cannot be controlled entirely, establishing a minimum authentication security perimeter of 10m with complete protection against sophisticated authentication spoofing attempts.

4.6.3 Environmental Security Consistency

The identical authentication defensive performance across indoor, outdoor, and occluded scenarios demonstrates that LAOUWN's authentication security properties are inherent to underground-to-air channel characteristics rather than dependent on specific environmental conditions. This consistency provides strong authentication security guarantees for diverse agricultural deployment scenarios while maintaining operational flexibility in sensor positioning across different environments for location authentication.

4.6.4 Advanced Threat Model Evaluation

Our comprehensive adversarial evaluation assesses the security properties of LAOUWN against sophisticated attack scenarios that represent a realistic threat to location-based authentication systems in critical agricultural IoT deployments. We evaluate three primary categories: mobile adversaries with dynamic positioning capabilities, coordinated multi-adversary attacks, and hardware trojans representing insider threats with physical access to system components.

4.6.4.1 Mobile Adversary Attack Analysis

Mobile adversaries represent a significant threat to location authentication systems due to their ability to dynamically position themselves and adapt their attack strategies based on spatial and temporal factors. Our evaluation implements mathematical simulations of mobile adversaries that can move along predetermined trajectories while attempting to spoof legitimate location signatures.

Temporal Consistency Against Mobile Attacks Figure 4.31(a) demonstrates the temporal consistency of our defense against mobile adversary attacks over a 15-second evaluation period. The attack accuracy remains consistently at approximately 0.33 throughout the entire duration, representing only marginal improvement over the random guess baseline of 0.33 for our three-location authentication system. This temporal stability indicates that mobile adversaries gain no significant advantage through dynamic positioning or timing based attack strategies.

Spatial Attack Success Patterns The spatial attack success patterns, illustrated in Figure 4.31(c), reveal uniform resistance across different adversary positions. The attack success rate varies minimally between 0.30 and 0.36 across the evaluated spatial

grid, indicating that our authentication system exhibits no exploitable spatial vulnerabilities. The consistent performance across diverse positions demonstrates that location-specific channel signatures cannot be effectively replicated through mathematical position models.

The heatmap visualization shows that regardless of where mobile adversaries position themselves within the evaluated area, their attack success rates remain clustered around the random guess baseline. This spatial uniformity is particularly significant because it indicates that proximity to legitimate transmitter locations provides no meaningful advantage for authentication bypass attempts.

Distance-Independent Security Properties Distance-based robustness evaluation, shown in Figure 4.31(b), demonstrates exceptional security properties that are independent of adversary proximity. The attack success rate remains constant at approximately 0.33 across distances ranging from 0 to 50 meters. This distance-independent security characteristic in particularly significant because it indicates that even adversaries with close proximity and potentially superior signal coupling cannot achieve meaningful improvements in authentication bypass success rates.

The consistent attack failure across all evaluated distances establishes a fundamental security property: LAOUWN's location authentication does not depend on maintaining physical separation from potential adversaries. This is crucial for agricultural IoT deployments where adversaries may have physical access to deployment areas but cannot compromise authentication integrity through positioning strategies alone.

4.6.4.2 Coordinated Attack Resistance

Coordinated attacks represent sophisticated threat scenarios where multiple adversaries collaborate to overcome authentication defenses through synchronized operations, temporal coordination, or signal amplification strategies. Our evaluation implements three distinct coordination strategies to assess the collective threat potential of multiple adversaries operating in concert.

Complete Failure of Coordination Strategies As demonstrated in Figure 4.31(f), all coordination strategies achieve an identical attack success rate of 0.333, indicating complete failure to improve upon individual adversary performance. This represents exactly the random guess baseline for our three-location authentication system, compared to the legitimate baseline accuracy of 0.92.

Synchronized Attacks: These attacks involve multiple adversaries transmitting simultaneously to create signal superposition effects, attempting to overwhelm or confuse the authentication system through coordinated signal interference. Despite the theoretical potential for signal amplification through constructive interference, synchronized attacks provide no advantage over single-adversary scenarios, achieving only random guess performance.

Time-Shifted Average Accuracy Attacks: These sophisticated attacks are designed to exploit potential temporal vulnerabilities through coordinated transmission scheduling, where multiple adversaries coordinate their transmissions at different time intervals to probe for temporal weaknesses in the authentication system. The identical 0.333 success rate demonstrates that temporal coordination strategies similarly fail to enhance attack effectiveness against LAOUWN's robust authentication framework.

Security Architecture Validation The coordinated attack resistance validates LAOUWN's defense-in-depth security architecture:

Multi-Adversary Resilience: The system maintains security against external threats through spatial correlation properties that remain robust even when multiple adversaries collaborate with perfect coordination and unlimited communication capabilities.

Scalable Security Guarantees: The uniform failure across coordination strategies indicates that increasing the number of coordinated adversaries would not improve attack success rates, providing scalable security guarantees regardless of adversary resources.

Physical Layer Foundation: The results reinforce that LAOUWN's security is fundamentally grounded in physical layer properties that cannot be circumvented through coordination strategies, establishing a theoretical foundation for the inherent security of underground-to-air location authentication systems.

4.6.4.3 Hardware Trojan Impact Assessment

Hardware trojans represent the most severe threat category in our evaluation, simulating insider attacks where adversaries have physical access to system components and can implement malicious modifications at the hardware level. Our evaluation implements four distinct trojan types that target fundamental signal characteristics used in RF fingerprinting authentication.

Complete Hardware Trojan Immunity Figure 4.31(d), presents the model accuracy analysis across all hardware Trojan types, demonstrating complete immunity to hardware-level attacks. All Trojan implementations achieve exactly 0.333 model accuracy, which corresponds to random guessing performance for our three-location

authentication system, compared to the baseline accuracy of 0.92. This represents a remarkable security property where hardware-level modifications cannot improve attack success beyond random chance.

The uniform failure across all trojan types indicates that:

Phase Shifter Trojans: These trojans introduce systematic phase corruptions to exploit timing-based authentication features but fail to compromise system security. Despite targeting the phase characteristics that are fundamental to CIR-based location authentication, these modifications cannot bypass the robust feature learning of our CNN architecture.

Amplitude Manipulator Trojans: Designed to alter signal magnitude characteristics, these trojans similarly provide no improvement in authentication bypass capability. The consistent 0.333 accuracy demonstrates that amplitude modifications at the hardware level cannot replicate the complex spatial signatures required for location spoofing.

Timing Jammer Trojans: These implement circular shifts to corrupt temporal signal structure but achieve no impact on system integrity. Even systematic timing manipulations cannot overcome the sophisticated spatio-temporal feature extraction of the ResNet-based authentication system.

Noise Injector Trojans: Adding targeted interference based on signal variance characteristics, these trojans also fail to compromise authentication effectiveness. The robustness against noise injection demonstrates that the learned location signatures transcend simple statistical signal properties.

Performance Degradation Analysis Figure 4.31(e), reveals the performance degradation analysis across all hardware trojan types. All hardware trojan types show exactly 63.9% performance degradation, which represents the difference be-

tween the baseline accuracy (0.92) and the random guess performance (0.333). This indicates that hardware trojans, rather than improving attack success, actually reduce the system to random guessing performance across all attack vectors.

This uniform degradation pattern demonstrates:

Algorithmic Tamper Resistance:

The CNN-based feature extraction has learned authentication signatures that are inherently robust to hardware-level signal corruptions. The learned representations capture fundamental propagation characteristics that cannot be effectively spoofed through hardware modifications.

Multi-layered Security Architecture: The residual learning structure of our ResNet implementation provides multiple pathways for authentication decisions, ensuring that localized hardware corruptions cannot compromise overall system security.

4.6.5 Security Analysis and Implications

4.6.5.1 Algorithmic Tamper Resistance

The comprehensive failure of all attack categories demonstrates that LAOUWN, achieves algorithmic tamper resistance, where the authentication algorithm maintains security properties even with corrupted or maliciously modified inputs. The consistent attack success rates across mobile, coordinated, and hardware Trojan scenarios indicate that our deep learning approach has identified authentication features that are inherently robust to our attack strategies.

The CNN architecture's ability to maintain authentication integrity despite systematic signal corruptions suggests that the learned feature representations capture fundamental propagation characteristics that cannot be effectively spoofed through current attack methodologies. The residual learning structure of our ResNet im-

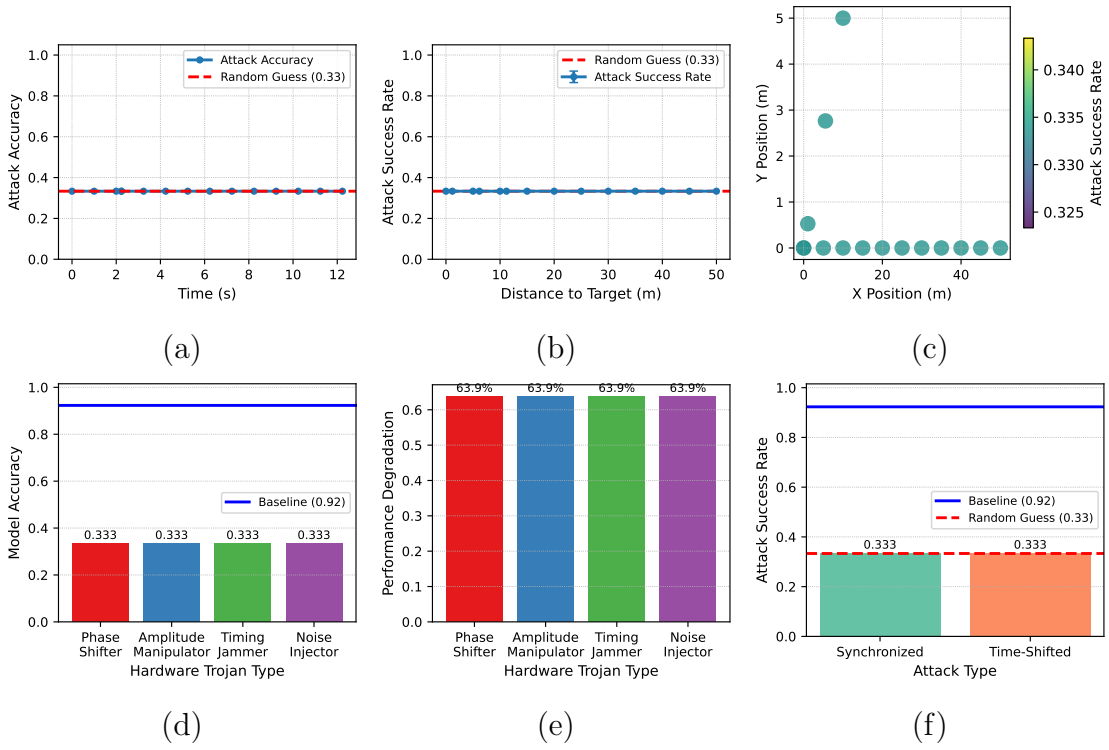


Figure 4.31: Advanced adversary attack analysis across multiple scenarios and metrics. (a) Attack accuracy evolution over time showing performance degradation patterns, (b) Attack success rate versus distance demonstrating range-dependent effectiveness, (c) Attack success heatmap showing spatial distribution of effectiveness, (d) Hardware trojan effectiveness across different implementations, (e) Hardware trojan performance degradation over time, and (f) Coordinated attack comparison showing performance and success rates for synchronized versus independent attack vectors.

plementation provides multiple pathways for authentication decisions, ensuring that localized signal corruptions cannot compromise overall system security.

4.6.5.2 Spatial Decorrelation Security Properties

Our evaluation confirms that the spatial correlation properties of underground-to-air propagation provide inherent security advantages for location authentication. The uniform attack failure across different spatial positions and distances demonstrates that location-specific signatures cannot be replicated through positioning strategies alone. The complex interaction between soil composition, moisture content, and

electromagnetic propagation creates authentication features that are fundamentally tied to specific geographic locations.

The distance-independent security characteristics observed in our evaluation indicate that proximity-based attacks, which are effective against many wireless authentication systems, provide no advantage against underground-to-air channel fingerprinting. This property is particularly valuable for agricultural IoT deployments where adversaries may have physical access to deployment areas but cannot compromise authentication through positioning strategies.

4.6.5.3 Defense-in-Depth Security Architecture

The comprehensive resistance across all threat categories demonstrates that LAOUWN implements an effective defense-in-depth security architecture. The system maintains security against external threats through spatial correlation properties, resists coordinated attacks through robust feature learning, and provides insider threat protection through algorithmic tamper resistance. This multi-layered security approach ensures authentication integrity across the complete threat spectrum relevant to critical agricultural IoT applications.

4.6.6 Practical Security Guarentees

4.6.6.1 Deployment Security Assurance

Our evaluation results provide strong evidence for the practical security of LAOUWN in real-world agricultural IoT. The consistent attack failure across diverse threat scenarios and minimal improvement over random guessing performance indicate that adversaries cannot achieve reliable authentication bypass even with sophisticated attack capabilities.

The hardware Trojan resistance provides particular assurance for insider threat protection. Agricultural IoT systems often involve distributed sensor deployments with limited physical security, making hardware modification attacks a realistic concern. Our demonstration of complete immunity to hardware modification attacks is a realistic concern. Our demonstration of complete immunity to hardware-level attacks ensures that authentication integrity remains intact even if individual system components are compromised.

4.6.6.2 Minimum Security Perimeter

The distance-independent attacks success rates establish that LAOUWN maintains consistent security properties across practical deployment ranges. The 10-meter minimum evaluation distance protects against proximity-based attacks while allowing flexible sensor positioning for agricultural monitoring applications. The consistent security performance across extended ranges supports scalable deployment architectures without degradation.

4.6.6.3 Environmental Robustness

The uniform attack resistance across different spatial positions and coordination strategies indicates that LAOUWN maintains security properties despite environmental variations that could affect wireless propagation. The robust authentication performance under diverse attack scenarios suggests that real-world environmental factors, such as seasonal changes, weather conditions, and agricultural activities, are unlikely to create exploitable security vulnerabilities.

Our comprehensive adversarial evaluation demonstrates that LAOUWN provides exceptional security properties suitable for critical agricultural IoT applications where

location authentication integrity is paramount for system operation and data trustworthiness.

4.7 Chapter Summary

This paper presents LAOUWN, a novel secret-free authentication framework that leverages unique channel impulse response (CIR) signatures from underground-to-air wireless channels to establish robust location-based authentication without cryptographic key management. We developed a comprehensive system combining advanced signal processing techniques with deep learning architectures (ResNet-18/34/50) enhanced through transfer learning and adversarial domain adaptation, achieving over 90% authentication accuracy across diverse indoor and outdoor scenarios. Our comprehensive security analysis demonstrates complete resistance to sophisticated adversaries, with both Friis empirical and ray-tracing enhanced attackers achieving only random guessing performance (33%), proving that even adversaries with perfect electromagnetic knowledge cannot compromise authentication integrity. We establish distance-independent security guarantees across 10-60 meter ranges and environmental robustness across indoor, outdoor, and underground scenarios. This makes LAOUWN particularly valuable for precision agriculture applications where unauthorized sensor access could lead to crop damage and financial losses. The demonstrated spatial decorrelation properties and resistance to advanced electromagnetic modeling establish a theoretical foundation for inherent security in underground wireless sensor networks, positioning LAOUWN as a foundational technology for securing critical agricultural IoT infrastructure.

CHAPTER 5

VET: Autonomous Vehicular Credential Verification using Trajectory and Motion Vectors

This chapter introduces VET (Vehicular credential Verification using Trajectory and motion vectors), a novel in-band authentication framework designed for autonomous vehicular networks. Instead of relying solely on traditional cryptographic credentials, VET verifies the physical veracity of the sender by cross-validating claimed trajectory and motion vectors (TMVs) with those independently estimated from wireless signal characteristics—specifically, frequency-of-arrival (FoA) measurements. The framework addresses a critical security gap by detecting adversaries who may have valid credentials but attempt to spoof their physical location or movement. VET supports single-verifier deployment, does not require line-of-sight conditions, and is robust even in the presence of remote adversaries with signal manipulation capabilities. The chapter presents the system and threat models, outlines the TMV-based verification protocol, and provides both theoretical analysis and experimental validation using USRP testbeds. Results demonstrate that VET achieves high true positive rates and resilience to advanced spoofing attacks, offering a practical and lightweight enhancement to vehicular authentication. We first describe the system and threat model, then we describe our developed verification protocol.

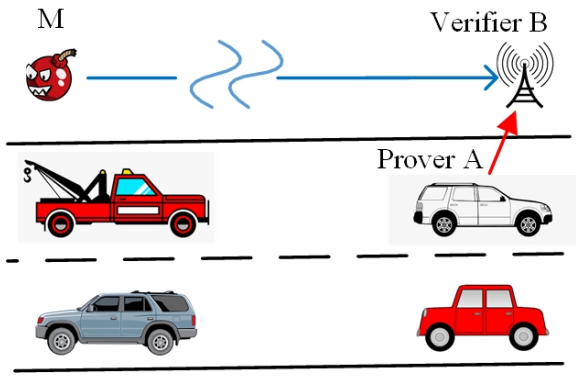


Figure 5.1: The verifier B performs verification of a prover A's credentials based on motion state vectors in the presence of an attacker M capable of spoofing trajectories.

5.1 Models

In this section, we first present the system model followed by the threat model for VET. We present Table 5.1, which summarizes the frequently used notations in this paper.

5.1.1 System Model

The Legitimate Prover (A)**:** The prover A has legitimate credentials, which can be either PKI credentials (pK, sK), or symmetric key credentials K. We assume that A uses an omnidirectional antenna to transmit wireless signals.

The Verifier (B)**:** The signal transmitted by A is received by the verifier B, when the prover is within the communication range. The trust is established by performing source authentication of the prover. We assume that there are one or more truthful verifiers X within the communication range. This is a valid assumption for CAVs, and there are other vehicles or roadside units or UAVs with more than one trusted controller and UAVs. However, these *verifiers do not require mutual trust. Each verifier performs VET independently and can broadcast a failure message in case of a failure.* It should be noted that an adversary can exploit to launch a denial-of-

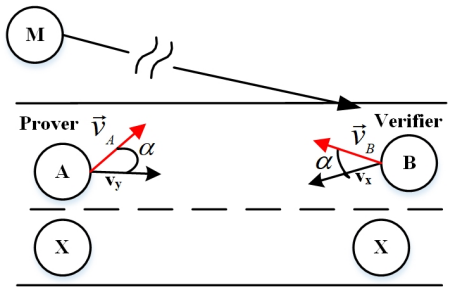


Figure 5.2: The prover A attempts to authenticate with the verifier B in the presence of an adversary M and other entities X within the communication range.

service; such an adversary can be manually removed. Also, this is orthogonal to VET presented in this paper.

5.1.2 Threat Model

We consider a Dolev-Yao attacker [49]. The adversary M has a valid credential, which can be either PKI credential (pk_M, sK_M) or symmetric credential K_M, and injecting its messages to disrupt the acceptable functionalities of a vehicular ad-hoc network. The attacker knows the locations of all the verifiers but does not physically control the verifier. M transmit a message with an intention for the verifier (B) to accept as the legitimate prover. M also knows all the channels between legitimate entities. The adversary can either be within the communication range of B or can compromise static wireless nodes connected to the internet to realize the attack. Finally, we do not put any restrictions on the motion of the adversary. Hence, the adversary can be either static or moving. The attack scenarios for this work are:

Remote Attacker: We consider an attacker located within the verifiers' communication range and attempting to inject his messages without intentional modification of PHY-layer data.

Remote Advanced Attacker: In addition to the capability of a remote attacker, the advanced attacker can intentionally modify the transmitted PHY-layer level wireless signal.

Notation	Description
A	Prover
B	Verifier
X	One or more truthful verifiers within the communication range
M	Adversary
$\mathcal{L}$	Claimed trajectory for k time-ordered locations for initial TMVs verification
$\mathcal{V}$	Claimed motion for k time-ordered locations for initial TMVs verification
$\mathcal{L}'$	Estimated trajectory for k time-ordered locations for initial TMVs verification
$\mathcal{V}'$	Estimated motion for k time-ordered locations for initial TMVs verification
$\tilde{\mathcal{L}}'$	Estimated trajectory for non A-to-B communication for final TMVs verification
$\tilde{\mathcal{V}}'$	Estimated motion for non A-to-B communication for final TMVs verification
$\mathcal{M}$	Set of transmitted k messages $(\{(m(1), t(1)), \ldots, (m(k), t(k)\})$
$\widehat{\mathcal{M}}$	Captured k messages
$\mathcal{F}$	Frequency of Arrival
$\tilde{\mathcal{F}}$	Frequency of Arrival for non A-to-B communication, where $t(i) \neq t'(i)$
ϵ	Acceptable error for location
μ	Acceptable error for velocity
Π_A	Prover oracle
Π_B	Verifier oracle
Π_X	Entity X oracle
Π_M	Adversary oracle
tx	Message transmissions
p^k	Probability of success of Adversary with No -Matching
d_{MB}	Distance between M and B
h_{MB}	Wireless channel between M and B
h_{MX}	Wireless channel between M and X
k	Number of trajectory data points
$RMSE(\cdot)$	Normalized root mean square error function for TMV Verification.

Table 5.1: Table of Notations

5.2 Primitives used in VET

Before diving into the details of VET, we present its building blocks. First, we present the method utilized to estimate the velocity. Followed by the method to compute the position and combine both to compute the trajectory and motion vectors (TMVs).

Frequency of Arrival (FoA) for Velocity Estimation: The FoA captures the effect of the velocity on the center frequency. In other words, it is the Doppler effect experienced by the moving verifier B with respect to the moving prover A at speed v. From Fig. 5.2, the prover vehicle A is within the communication range of the verifier B. The frequency of arrival when the verifier and the prover are moving towards each other, so the Doppler effect experienced by the verifier increases, is given by

$$\mathcal{F} = f_0 \times \frac{c + \overrightarrow{v}_B \; cos \; \alpha}{c - \overrightarrow{v}_A \; sin \; \alpha}, \tag{5.1}$$

where $\mathcal{F}$ is the Doppler shift on verifier B, c is the propagation speed, f_0 is the prover's center frequency.

From (5.1), the velocity of the prover at the ith sample is given by

$$\overrightarrow{v}_B(i) = \left[c - \frac{\mathcal{F}(i)(\overrightarrow{v}_A(i) \; sin \; \alpha)}{f_0}\right] cos^{-1} \; \alpha. \tag{5.2}$$

The Doppler effect of the signal measured by B is dependent on the radial velocity and the center frequency. The relative velocity observed $\overrightarrow{v}(i) = \overrightarrow{v}_B(i) \; cos \; \alpha - \overrightarrow{v}_A(i) \; sin \; \alpha$.

5.2.0.1 Direct Location Estimation

For estimation of the location, it is important to note that the verifier B has more than one antenna. This assumption is valid for vehicular networks as the roadside units are MIMO enabled, and in the case of UAV swarms, multiple single antenna UAVs can collude as the verifier. We used maximum likelihood estimation to directly estimate the position, which maximizes the likelihood for the prover [9] when the prover is broadcasting within the expected verification range. This is a one-step process that does a 2-D or 3-D grid search of the prover's position.

The location $\ell(i)$ of the prover B is the position that maximizes the log-likelihood function. This position is expressed as

$$\ell(i) = \arg\max_{\ell}\{L_i\}. \tag{5.3}$$

Here, the log-likelihood function is written as

$$L_i = \sum_{j=1}^{J} \lambda_{max}(Q_j). \tag{5.4}$$

The log-likelihood function is the summation of all the maximum eigenvalues of the Hermitian matrix Q_j. The matrix contains received signals multiplied by the frequency difference of arrival (FDoA) at the different antennas.

Effect of NLoS on the FoA: Typically, when vehicular wireless signals propagate in the real world, it does that in multipath [176]. The motion claim of the prover reaches the verifier in two or more paths. For simplicity, in our research, we use two path components. The NLoS path exists due to signal reflection before getting to the verifiers. This means there exists a reflection point that is always changing due to the dynamicity of the environment and this change affects the position by some factor δ.

Therefore, the prover's position for verifiers is $c - \overrightarrow{v_B} cos\ \alpha + \delta$. This means that for a moving verifier and a moving prover due to NLoS, the Doppler effect will be given by

$$\mathcal{F} = f_0 \times \frac{c + \overrightarrow{v}_B\ cos\ \alpha + \delta}{c - \overrightarrow{v}_A\ sin\ \alpha} + \epsilon, \tag{5.5}$$

This NLoS component is embedded in the signal path attenuation which maximizes the likelihood [9], at each verifier and the function that contains the unknown provers position and velocity.

5.3 VET: Credential Verification using Trajectory and Motion Vectors

We present a secured, in-band vehicular access control method to verify the authenticity and integrity of a set of messages transmitted from a legitimate vehicle A at the verifier, as shown in fig. 5.1. B implements a location based strategy for verification, where B does not trust A in the start of the communication. For the verification, B generates a set of trajectory and motion vectors from the carrier frequencies.

5.3.1 Vehicular Motion State Verifier

The basic idea is for verifier B to authenticate the claimed trajectory observed for a prover A via a location-based authentication strategy. We exploit the characteristics of the direct position and velocity estimation via the arrival frequency to verify the prover. The protocol is presented as without generality between a prover A and a verifier B. It should be noted that simultaneous runs of the protocol can be initiated between the same prover and different verifiers as the prover A can simultaneously communicate with various entities. First, we describe TMVs utilized to develop VET:

Trajectory and Motion Vectors: The trajectory $\mathcal{L}$ of a moving vehicle is defined as k time-ordered locations

$$\mathcal{L} = \{(\ell(1), t(1)), (\ell(2), t(2)), \ldots, (\ell(k), t(k))\}, \tag{5.6}$$

where each location $\ell(i) = (x(i), y(i))$ is the geospatial location coordinate at time $t(i)$. Where $1 \leq i \leq k$ for $t(i)$ and $t(i) < t(j)$ for $i < j$. Further, the motion $\mathcal{V}$ is defined as k time-ordered locations in the same epoch

$$\mathcal{V} = \{(\overrightarrow{v}(1), t(1)), (\overrightarrow{v}(2), t(2)), \ldots, (\overrightarrow{v}(k), t(k))\}, \tag{5.7}$$

where $\overrightarrow{v}(i)$ is the velocity at time $t(i)$. These locations and velocities are obtained using the method described in Section 5.2

The protocol is initiated when the prover A is within the communication range of B. The prover A sends *Request to Authenticate* message with authenticated encryption using issued credentials. An authenticated encryption function AE$(\cdot)$ utilizing the shared secret K [20]. This will guarantee the source's authenticity, message integrity, and confidentiality. When verifiers share a common secret, AE$(\cdot)$ can be implemented as an encrypt-then-MAC operation. Whereas for the public key cryptographic scenario, AE$(\cdot)$ can be implemented as a sign/encrypt/sign (or encrypt/sign/encrypt). Here, the credential can either be the actual one issued by a trusted authority or a pseudonym credential for preserving privacy. The verifier B provides the prover limited connectivity if the credentials are verified.

During the limited connectivity, the verifier B captures the message transmitted to it and extracts the claimed k TMVs $\mathcal{L}$ and $\mathcal{V}$, and estimated TMVs $\mathcal{L}'$ and $\mathcal{V}'$. First, B verifies the claimed and estimated using a root mean square error (RMSE)

function. If successful, in the same time epoch, the verifier B captures the frequency of arrival (FoA) $\tilde{\mathcal{F}}$ for the messages transmitted by A but not intended for the B. From these FoA, the verifier B estimates TMVs $\tilde{\mathcal{L}}'$ and $\tilde{\mathcal{V}}'$, which are shifted in time as compared to claimed. Further, B maps the claimed TMVs to the same time as estimated TMVs using kinematic equations. The estimated and claimed TMVs are compared; if these are within the accepted errors, A's messages are accepted and granted full access. Formally, the vehicular motion state verification steps are:

1. **Initial Request :** Once the prover A is within the communication range of the verifier B. A transmits a request to authenticate $\mathrm{AE}_K(RTA)$ to the verifier B to join.

2. **Limited Access Connection:** After verifying the authenticity of A's credential K, B grants it *limited access*. During the limited access B captures k messages transmitted by A as $\mathcal{M} = \{(m(1), t(1)), \ldots, (m(k), t(k)\}$, containing claimed TMVs: velocity vectors $\mathcal{V} = \{(\overrightarrow{v}(1), t(1)), \ldots, (\overrightarrow{v}(k), t(k)\}$ and $\mathcal{L} = \{(\ell(1), t(1)), \ldots, (\ell(k), t(k))\}$. B also records the FoA $\mathcal{F} = \{(f(1), t(1)), \ldots, (f(k), t(k)\}$, and computes TMVs: velocity vectors $\mathcal{V}' = \{(\overrightarrow{v}(1)', t(1)), \ldots, (\overrightarrow{v}(k)', t(k)\}$ and $\mathcal{L}' = \{(\ell(1)', t(1)), \ldots, (\ell(k)', t(k))\}$.
   It should be emphasized that the verifier has not yet acknowledged any of the critical directives. It is only used for the verifier to extract the relevant trajectory information of the incoming prover for verification.

3. **Initial TMVs Verification:** B computes the Root Mean Square Error (RMSE) of location as:

$$RMSE(\ell(i), \ell(i)') = \sqrt{\frac{\sum_{i=1}^{k} \left(\frac{\ell(i) - \ell(i)'}{\ell(i)'}\right)^2}{k}}.$$

The RMSE of velocity is as follows;

$$RMSE(\overrightarrow{v}(i), \overrightarrow{v}(i)') = \sqrt{\frac{\sum_{i=1}^{k}\left(\frac{\overrightarrow{v}(i)-\overrightarrow{v}(i)'}{\overrightarrow{v}(i)'}\right)^2}{k}}.$$

B then performs verification:

$$RMSE(\ell(i), \ell(i)') \overset{?}{\leq} \epsilon \quad \forall\ 1 \leq i \leq k,$$

$$RMSE(\overrightarrow{v}(i), \overrightarrow{v}(i)') \overset{?}{\leq} \mu \quad \forall\ 1 \leq i \leq k,$$

where RMSE$(\cdot)$ is a normalized root mean square error function, and ϵ and μ are the acceptable error. If B passes the check, A grants B partial access and accepts $\mathcal{M}$ as valid. Else, B disregards $\mathcal{M}$ and terminates the connection of A. Also, B broadcasts a signal notifying FAILED authentication of A.

4. **Estimating TMVs for non A-to-B communication:** During the same time epoch, verifier records FoA $\tilde{\mathcal{F}} = \{(\tilde{f}(1),\ t'(1)),\ \ldots, (\tilde{f}(k), t'(k)\}$, where $t(i) \neq t'(i)$, from the transmissions (tx) from A not intended for B. Next the verifier B computes corresponding velocity vectors $\tilde{\mathcal{V}}' = \{(\overrightarrow{\tilde{v}'}(1), t'(1)),\ \ldots, (\overrightarrow{\tilde{v}'}(k), t'(k)\}$, and trajectory vectors $\tilde{\mathcal{L}}' = \{(\tilde{\ell}'(1), t'(1)),\ \ldots, (\tilde{\ell}'(k), t'(k))\}$.

5. **Interpolating Claimed TMVs:** The estimated TMVs $\overrightarrow{\tilde{v}'}(i)$, $t'(i)$ and $(\tilde{\ell}'(i), t'(i))$ are interpolated to synchronize with claimed TMVs $(\overrightarrow{v}(i), t(i))$ and $(\ell(i), t(i))$ using cubic spline interpolation methods for a timeseries data [113].

6. **Final TMVs Verification:** The RMSE of location is calculated as:

$$RMSE(\ell(i), \tilde{\ell}'(i)) = \sqrt{\frac{\sum_{i=1}^{k}\left(\frac{\ell(i)-\tilde{\ell}'(i)}{\tilde{\ell}'(i)}\right)^2}{k}}.$$

The RMSE of velocity is computed as:

$$RMSE(\overrightarrow{v}(i), \overrightarrow{\tilde{v}'}(i)) = \sqrt{\frac{\sum_{i=1}^{k}\left(\frac{\overrightarrow{v}(i)-\overrightarrow{\tilde{v}'}(i)}{\overrightarrow{\tilde{v}'}(i)}\right)^2}{k}}.$$

Finally, B performs verification:

$$RMSE(\ell(i), \tilde{\ell}'(i)) \overset{?}{\leq} \epsilon \quad \forall\ 1 \leq i \leq k,$$

$$RMSE(\overrightarrow{v}(i), \overrightarrow{\tilde{v}'}(i)) \overset{?}{\leq} \mu \quad \forall\ 1 \leq i \leq k,$$

where RMSE$(\cdot)$ is a normalized root mean square error function, and ϵ and μ are the acceptable error. If B passes the check, A grants B full access and accepts $\mathcal{M}$ as valid. Else, B disregards $\mathcal{M}$ and terminates the connection of A. Also, B broadcasts a signal notifying FAILED authentication of A.

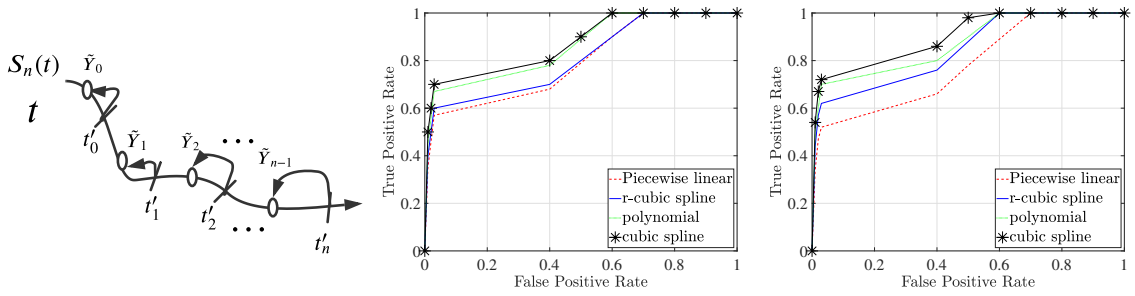
Figure 5.3 formally presents the steps of VET. A remote adversary M who cannot modify the physical characteristics of the transmitted signal is detected in Step 3, as the claimed TMVs are for the emulated trajectory while the estimated TMVs are the actual trajectory of M. Further, an advanced adversary M who with the knowledge of channel to the verifier B can craft the FoA to match the emulated and the claimed TMVs. Such as the adversary is detected by Step 6, as in Step 4 the verifier B captures the FoAs when the adversary will be communicating with any other entity present in the vicinity. Such communication can be detected by noting the sender and receiver in the header [21].

We will present a more detailed discussion on the robustness of the protocol in the next section. Here, it is assumed that the advanced adversary is attempting to emulate different trajectories at different verifiers. This is an acceptable assumption

	A		B
Initialization:		$\xrightarrow{AE_K(RTA)}$	Grants limited Access
Captures Claimed Messages:	$\mathcal{M} = \{(m(1), t(1)), \ldots, (m(k), t(k)\}$	$\xrightarrow{AE_K(\mathcal{M})}$	$\widehat{\mathcal{M}}$, also capture $\mathcal{F} = \{(f(1), t(1)), \ldots, (f(k), t(k)\}$, extract $\mathcal{V} = \{(\overrightarrow{v}(1), t(1)), \ldots, (\overrightarrow{v}(k), t(k)\}$ $\mathcal{L} = \{(\ell(1), t(1)), \ldots, (\ell(k), t(k))\}$. compute $\mathcal{V}' = \{(\overrightarrow{v}(1)', t(1)), \ldots, (\overrightarrow{v}(k)', t(k)\}$ $\mathcal{L}' = \{(\ell(1)', t(1)), \ldots, (\ell(k)', t(k))\}$.
Initial TMVs Verification:	Partial access.		$RMSE(\overrightarrow{v}(i), \overrightarrow{v}(i)') = \sqrt{\frac{\sum_{i=1}^{k}\left(\frac{\overrightarrow{v}(i)-\overrightarrow{v}(i)'}{\overrightarrow{v}(i)'}\right)^2}{k}}$, $RMSE(\ell(i), \ell(i)') = \sqrt{\frac{\sum_{i=1}^{k}\left(\frac{\ell(i)-\ell(i)'}{\ell(i)'}\right)^2}{k}}$. $RMSE(\overrightarrow{v}(i), \overrightarrow{v}(i)') \stackrel{?}{\leq} \epsilon$, $RMSE(\ell(i), \ell(i)') \stackrel{?}{\leq} \mu$,
Computes TMVs's :	tx (Transmission not to B)	$\xrightarrow{Sniff\ Packets}$	$\tilde{\mathcal{F}} = \{(\tilde{f}(1), t'(1)), \ldots, (\tilde{f(k)}, t'(k)\}$, $t(i) \neq t'(i)$, compute $\tilde{\mathcal{V}}' = \{(\overrightarrow{\tilde{v}}(1), t'(1)), \ldots, (\overrightarrow{\tilde{v}}(k), t'(k)\}$, $\tilde{\mathcal{L}}' = \{(\tilde{\ell}'(1), t'(1)), \ldots, (\tilde{\ell}'(k), t'(k))\}$.
TMVs Interpolation:			From $t(i)'$ to $t(i)$ using cubic spline interpolation: $\mathcal{V} = \{(\overrightarrow{v}(1), t(1)), \ldots, (\overrightarrow{v}(k), t(k)\}$, $\mathcal{L} = \{(\ell(1), t(1)), \ldots, (\ell(k), t(k))\}$.
Final TMVs Verification:			$RMSE(\overrightarrow{v}(i), \overrightarrow{\tilde{v}}(i)) = \sqrt{\frac{\sum_{i=1}^{k}\left(\frac{\overrightarrow{v}(i)-\overrightarrow{v'(i)}}{\overrightarrow{v'(i)}}\right)^2}{k}}$, $RMSE(\ell(i), \tilde{\ell}'(i)) = \sqrt{\frac{\sum_{i=1}^{k}\left(\frac{\ell(i)-\tilde{\ell}'(i)}{\tilde{\ell}'(i)}\right)^2}{k}}$. $RMSE(\overrightarrow{v}(i), \overrightarrow{\tilde{v}}(i)) \stackrel{?}{\leq} \epsilon$, $RMSE(\ell(i), \tilde{\ell}'(i)) \stackrel{?}{\leq} \mu$,
Decision:	Full access.	$\xleftarrow{Broadcast\ Decision}$	Decision ¡FAIL or SUCCESS¿, Accept $\mathcal{M}$, grant B full access.

Figure 5.3: Vehicular Motion Vectors Verifier Protocol.

as all the verifiers will have different physical locations. Hence, emulating the same physical trajectory will force M to emulate different perceived trajectories at different verifiers.



(a) (b) (c)

Figure 5.4: (a) A timeline for the interpolation, (b) ROC curve for location data using various interpolation techniques, and (c) ROC curve for velocity data using various interpolation techniques.

5.3.2 Interpolating TMVs

In Step 5, of VET the estimated TMVs need to be interpolated for synchronizing with claimed TMVs, as shown in Fig. 5.4(a) such that the comparison can be made between the estimated and claimed trajectories. We perform the interpolation utilizing cubic spline interpolation [113, 155] because of its high accuracy, smoothness, flexibility, robustness, and less noisy interpolation when modeling trajectory motion profiles. Compared to other interpolation techniques like piecewise linear interpolation [95], r-cubic spline [143], and polynomial interpolation [61], cubic spline produces a smoother curve. Piecewise linear interpolation [95] has a high granularity of the TMV data but only does well when the vehicle is moving on a straight line at constant velocity. It is not as robust as cubic-spline for interpolating TMV in the real world. Linear interpolation works well in ideal scenarios, but in our experiments, the vehicles move at changing speeds at different times. Piecewise polynomial interpolation, like quadratic spline, is not the best interpolation technique compared to cubic spline and r-cubic spline regarding accuracy and smoothness, especially for complex scenarios. r-cubic spline [143] is simpler and faster but less accurate because its interpolation is based on simple recurrence equations, unlike the cubic spline which requires solving

tri-diagonal matrix-vector equations. In Fig. 5.4 (b) and (c), we show that cubic spline has the best performing ROC curve as compared to other techniques for location and velocity, respectively, when we account for a trajectory with 90° turn. We used the data we collected for evaluations; please refer to Section 5.5.1 for more details. Although we do note that cubic spline is more computationally expensive, it is acceptable for our model as a not computationally limited verifier performs all the computations. Moreover, we need accuracy and smoothness of the curve, especially for irregular data points, rather than speed for VET. The cubic spline interpolation of both the location and velocity is performed using the following equation:

$$\begin{cases} S_0(t) = \tilde{Y}_0 + b_0(t - t'_0) + c_0(t - t'_0)^2 + d_0(t - t'_0)^n \ \forall \ t \in [t'_0, t'_1], \\ \vdots \\ S_n(t) = \tilde{Y}_{n-1} + b_{n-1}(t - t'_n) + c_{n-1}(t - t'_n)^2 + d_{n-1}(t - t'_n)^n \ \forall \ t \in [t'_{n-1}, t'_n]. \end{cases} \tag{5.8}$$

where $\tilde{Y}$ can be either velocity $\overrightarrow{v}$ or location ℓ of the vehicle at time t', computing the parameters for b, c, d is obtained from solving a system of linear equations and substitution. The result will be a TMV curve that is smooth and more continuous than other forms of interpolations.

5.4 Security Analysis

In this section, first, we analyze the correctness of VET followed by robustness analysis against the adversary presented in Section 5.1.2. For the formal analysis of the protocol, we will utilize the idea of matching conversation [21]. The main idea is that two entities can mutually authenticate each other in the presence of an adversary if and only if they have the same chronology of exchanged messages.

5.4.1 Correctness Analysis

We discuss the correct implementation of VET when there is no adversary present. We consider the prover A and verifier B to be modeled by an Oracle model. We define the protocol transcript at A and B as Π_A and Π_B, respectively as observed by the oracle Π. In the transcripts, the received messages are denoted by a hat notation. The transcripts of the messages exchanged between A and B are:

$$\Pi_A = \{AE_K(RTA); m(1); tx(1); \ldots; m(k); tx(k)\}, \tag{5.9}$$

$$\Pi_B = \{\widehat{AE_K(RTA)}; \widehat{m(1)}; \widehat{tx(1)}; \ldots; \widehat{m(k)}; \widehat{tx(k)}\}, \tag{5.10}$$

for ease of depiction, we have skipped the timestamps for the messages. Several communicating oracles are also possible in a distributed way, but each oracle is unique.

The matching conversation is a way of authenticating an entity, which is the prover A. Both A and B will get the same long-lived key K, which would be unknown to anyone else. Once the communication is correct, the verifier B confirms or denies the prover A. That is, at the end of the conversation, the decision (η), from the verifier B is to confirm ($\mathcal{C}$) or reject ($\mathcal{R}$) the prover A ($\eta, \mathcal{C}, \mathcal{R}$). Although rejection can occur before the end of the conversation, confirmation only happens at the end.

The prover oracle (Π_A) sends a message $AE_K(RTA)$, which contains the request to authenticate. The verifier oracle (Π_B) receives the message $\widehat{AE_K(RTA)}$. It decrypts the message to check if the correct key K was used and grants the prover partial access. Next the prover oracle (Π_A) transmits the message $m(1)$ where Π_B extracts trajectory and motion vectors (TMVs) $\ell(1)$ and $v(1)$. This is followed by Π_A transmits a message $tx_A(1)$ not intended for Π_B. From $\widehat{tx_A(1)}$ transmission Π_B records the

Frequency of Arrival (FoA) $\tilde{f}(1)$. Now the verifier estimates the velocity $\overrightarrow{\tilde{v}'}(1)$ and location $\tilde{\ell}'(1)$. Finally, compare the estimated and claimed velocities and locations based on the RMSE in Step 6 after interpolating the estimated to synchronize with the claimed in Step 5. It is straightforward to show if the message $m(1)$ and the transmission $tx(1)$ are from the same prover oracle Π_A. The estimated and claimed will be within the acceptable error ϵ for location and μ for velocity. This is repeated for all k transmissions.

5.4.2 Robustness Analysis

Next, we will analyze the robustness of VET against the threat model we defined in Section 5.1.2. First, we will analyze VET against a remote attacker who injects messages. This is followed by the remote advanced attacker, who can modify its physical layer envelope in an attempt to force the verifier B to accept the messages.

Remote Attacker: The remote adversary (M) is inside the communication range of the verifier B. Here, in the oracle model, we have an adversary oracle Π_M and the verifier oracle Π_B. The transcripts of the messages exchanged between M and B for VET execution is:

$$\Pi_M = \{AE_{K_M}(RTA); m_M(1); tx_M(1); \ldots; m_M(k); tx_M(k)\}, \tag{5.11}$$

$$\Pi_B = \{\widehat{AE_{K_M}(RTA)}; \widehat{m_M(1)}; \widehat{tx_M(1)}; \ldots; \widehat{m_M(k)}; \widehat{tx_M(k)}\}, \tag{5.12}$$

For ease of depiction, we have skipped the timestamps for the messages. For the case of a legitimate prover (A), there can be two different scenarios: (1) A is not present, and M initiates VET, and (2) A is present and M hijacks the execution of VET. In

the first case, the prover oracle's transcript is:

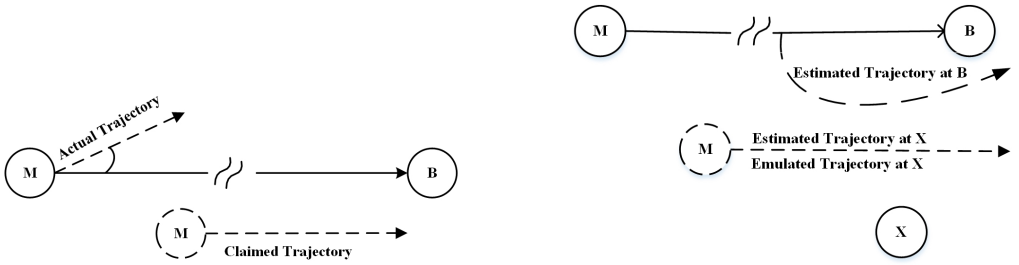
$$\Pi_A = \{\emptyset\}. \tag{5.13}$$

In the second case, the transcript is:

$$\Pi_A = \{AE_K(RTA); m(1); tx(1); \ldots; m(k); tx(k)\}. \tag{5.14}$$

To prove the robustness of VET against a remote attacker. We aim to prove that the acceptance or authentication at the verifier (B) with non-matching conversations at prover and verifier oracles Π_A and Π_B, respectively, is negligible. Let us dive into the individual messages exchanged between the adversary oracle Π_M and the verifier oracle Π_B, (5.11) and (5.12). The first message that is exchanged between them is $AE_{K_M}(RTA)$, this message is accepted by Π_B even with mismatch with Π_A, (5.13) and (5.14). This is because the credential used by Π_M is either issued by a valid Trusted Authority (TA) or compromised from the legitimate prover A. Thus Π_M is able to initiate the session. Now, let us focus on the k messages exchanged for the verification. There are two sets of messages $m_M(i)$ are the messages intended for the verifier oracle Π_B and Π_B estimates the TMVs. And the transmissions $tx_M(i)$ from adversary oracle Π_M intended for other oracles present such as Π_X. Which can be some other verifier in the same vicinity. Such that the verifier oracle Π_B can be a roadside unit and other oracles Π_X can be another vehicle in the vicinity.

For this type of adversary, the claimed and the estimated TMVs because the adversary is present at a remote location, as shown in Fig. 5.5(a). This will force Π_B to estimate the remote TMVs detected by Step 3 of VET. It should be noted here this applies to both static and moving adversaries. Hence, this type of adversary



(a) (b)

Figure 5.5: (a) A remote adversary M attempting to authenticate with a spoofed trajectory inside the communication range of the verifier B, and (b) a remote advanced attacker M attempting to authenticate an emulated trajectory to verifier B with other verifiers X in the vicinity.

will be detected and removed from the system. As well as the verifier will notify the presence of an adversary to other entities in the communication range.

Remote Advanced Attacker: Similar to the analysis against a remote attacker, to prove the robustness of VET against a remote advanced attacker. We aim to prove that the acceptance or authentication at the verifier (B) with non-matching conversations at prover and verifier oracles Π_A and Π_B, respectively, is negligible. The individual messages exchanged between the adversary oracle Π_M and the verifier oracle Π_B, (5.11) and (5.12). The first message that is exchanged between them is $AE_{K_M}(RTA)$, this message is accepted by Π_B even with mismatch with Π_A, (5.13) and (5.14). Here, in addition to injecting the set of claimed TMVs $\mathcal{M}$, the advanced adversary can change the envelope and FoA to emulate a trajectory estimated from sniffed packets indented for Π_X. Hence, an advanced adversary oracle Π_M is capable of emulating a trajectory to Π_B, as shown in Fig. 5.5(b).

The emulated trajectory of Π_M is accepted at Π_B without matching conversation with Π_A if the RMSE of all the TMVs in Step 6 is within the acceptable range. This cannot happen with certainty as Π_M because even if emulating the same trajectory to

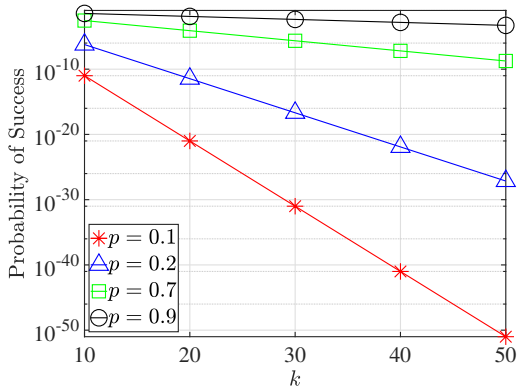


Figure 5.6: The plot shows the success probability of M for different numbers of TMVs in the trajectory k.

Π_B and Π_X. The estimated trajectories will be different; this is because the estimated trajectory $(\mathcal{V}', \mathcal{L}')$ in Step 2 is emulated for Π_B. Whereas the estimated trajectory $(\tilde{\mathcal{V}}', \tilde{\mathcal{L}}')$ in Step 4 is emulated for Π_X. The adversary does this to pass Step 3, where the claimed and estimated trajectories must match at respective verifiers. Note here that Step 4 for Π_B captures the messages for Step 2 of Π_X. Also, verifiers inform all other entities about the failure of the authentication of any entity. It should be noted here this can be utilized to launch a denial-of-service (DoS) where a legitimate entity is forcibly disconnected. This is orthogonal to the application of VET. This can be trivially tackled by cryptographic verification of the failure broadcast. Next, we show that both types of adversaries have negligible success probability in defeating VET.

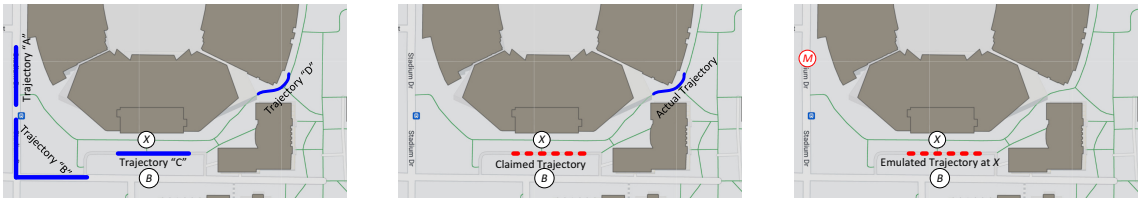
Formal Proof: For both the adversary models, we can model the success of the adversary oracle Π_M for claimed TMVs to match the estimated TMVs. Let the probability for Π_M for $\overrightarrow{v}(i)$ and $\ell(i)$ match with $\overrightarrow{\tilde{v}'}(i)$ and $\tilde{\ell}'(i)$, respectively at Π_B be p. This probability depends on the distance of the adversary M from the emulated trajectory. As the wireless channel outdoors decorrelates [188]. We evaluate this probability in the evaluation section. Thus, for k TMVs, the probability of an adversary succeeding with no matching is

$$\Pr[B \text{ accept} \wedge \text{No-matching}] = p^k, \tag{5.15}$$

which is a negligible probability [135], as shown in Fig. 5.6. Even for a high probability $p_1 = 0.9$, for 50, 40, 30, 20, and 10 TMVs, the success probability is 5×10^{-3}, 1.4×10^{-2}, 4.2×10^{-2}, 1.2×10^{-1}, and 3.5×10^{-1}. Please note here for a single execution of VET, the attacker has only one chance to inject all the TMVs online. Hence, a higher probability of success is acceptable here relative to traditional crypto-algorithm (similar values are acceptable for other online protocols with short authentication strings [123]).

5.4.3 Discussion on Shortcomings

One of the areas we need to recognize is in the absence of at least two truthful verifiers, a remote advanced adversary can be successful. But it should be noted here that a novice remote adversary who cannot craft the physical layer envelope can be detected with only one verifier. Thus, only detecting an advanced adversary can craft the physical layer envelope with the knowledge of all the channels within the entities. We need more than one truthful verifier, which is not a reasonable requirement for detecting the strongest possible adversary.



(a) (b) (c)

Figure 5.7: (a) Verifying different trajectories of a legitimate prover A, (b) a remote adversary injecting a claimed trajectory, and (c) a remote advanced adversary manipulating the signal's physical properties to emulate trajectory.

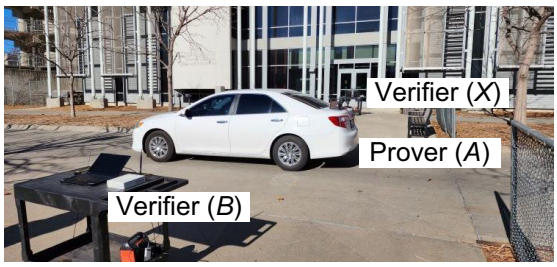


Figure 5.8: Experimental setup with Prover car (A) and Verifiers (B) and (X).

5.5 Experimental Evaluation

In this Section, we evaluate the correctness, robustness, and protocol parameters utilizing a USRP platform with well-defined experiments. First, we describe the experimental setup followed by correctness and robustness analysis.

5.5.1 Experimental Setup

Our experimental setup includes a prover (A) vehicle and stationary verifier (B), as shown in Fig. 5.7(a) and Fig. 5.8. We have a secondary verifier (X) present in the system for demonstration purposes only; we do not use the data collected at (X) for evaluations. The prover vehicle contains the signal transmitter USRP 2922 inside a car, which continuously broadcasts the BPSK signal at 915MHz using an omnidirectional antenna (VERT-900). The transmitter USRP is connected to a Lenovo ThinkPad T14 running the GNU Radio transmitter code. We choose 915MHz center frequency with a bandwidth, f_0, instead of 2.45GHz, which is in the Wi-Fi band because it is less congested and has a longer range. The verifiers are two stationary USRP 2922s connected to two individual computers placed on the opposite side of the road, which acts as receivers. The center frequency is also set to 915MHz, with a target sampling rate of 32000Sps and an actual sample rate of 195312Sps. The

receivers also run GNU Radio code to capture the transmitted data packets from the moving prover. A GPS-enabled phone collects the ground truth of location and velocity data as the prover vehicle drives around the verifiers. We synchronize all three computers and the phone to use the United States Internet Time Server (ITS) of The Network Time Protocol server [185]. The verifier collects timestamped data as the prover drives around at a constant speed.

5.5.2 Correctness Analysis

First, we focus on evaluating the correct performance of VET. For this, we evaluate the location estimation and velocity estimation individually. The performance of VET is the worst of either of the estimations. We captured the physical layer envelope and frequency of arrival of the signal received from A. We implemented the methods mentioned in Appendix A to estimate velocity and position. Then, we compute the key performance indicator (Receiver operating characteristic) by comparing the estimated velocity and position to the ground truth recorded on the phone kept inside A.

Receiver operating characteristic (ROC) curve: We compute two separate Receiver operating characteristic (ROC) curves for velocity and location data. We use the ROC curves to evaluate three parameters. First, the acceptable errors to set the thresholds (ϵ, μ) for RMSE of location and velocity, respectively, in Steps 3 and 6. Second, k is the number of trajectory points required to complete the verification with an acceptable true positive rate. Finally, we evaluate the acceptable errors for the straight or turning trajectory of the vehicle.

Figure 5.9(a) shows the plot between true positive rate (TPR) and false positive rate (FPR) for various ϵ RMSE errors and $k = 3$ for the location data. From the figure, we observe that for $\epsilon = 0.2$, we observe a 0.92 true positive rate for 0.03 false

positive rate. In Fig. 5.9(c), we show the location data ROC curve for various k number of trajectory points for $\epsilon = 0.2$. We observe that for $k = 3$ VET can achieve TPR = 0.96 for FPR or 0.03. Further, in Fig. 5.9(b) and (d), we plot the velocity ROC curve for various RMSE threshold (μ) and k, respectively. We observe that for velocity, VET achieves a TPR of 0.9 for $\mu = 0.2$ and a TPR of 0.94 for $k = 3$. We also observe that each of the curves are acceptable ROC curve as the TPR goes close to 1 before the FPR reaches 0.05. For the rest of the experimental analysis, we set the values of the thresholds from the ROC curves. Specifically, we fix the $\epsilon = 0.2$ and $\mu = 0.2$ for location RMSE and velocity RMSE, respectively. We selected these values as they achieve optimum TPR for acceptable FPR. Finally, we compute the ROC curves for various trajectories, as shown in Fig. 5.7(a). The trajectory "A" and "C" are straight line while "B" and "D" involves turns. From the curves in Fig. 5.9(e) and (f), we observe better performance for straight-line trajectories as compared to ones involving turns. However, all of the TPR and FPR values are acceptable, with TPR reaching 1 for the trajectories involving turns before FPR reaches 0.08.

5.5.3 Robustness Analysis

Next, we evaluate the robustness of VET against both adversaries defined in Section 5.1.2. First, we evaluate the remote attacker who injects a spoofed trajectory as a message. Next, we evaluate the performance of an advanced remote attacker M who can change the physical parameters of the signal to emulate a target trajectory. We compute the success probability using the $RMSE(\cdot)$ function.

Remote Attacker: We utilized the data collected to emulate the remote attacker. Here, the attacker's actual trajectory differed from the claimed trajectory, as shown in Fig. 5.7(b). Using the data, we plot two graphs for location and velocity. In Fig. 5.10(a) and Fig. 5.10(b), we plot the probability of success (p^k) from (5.15) for

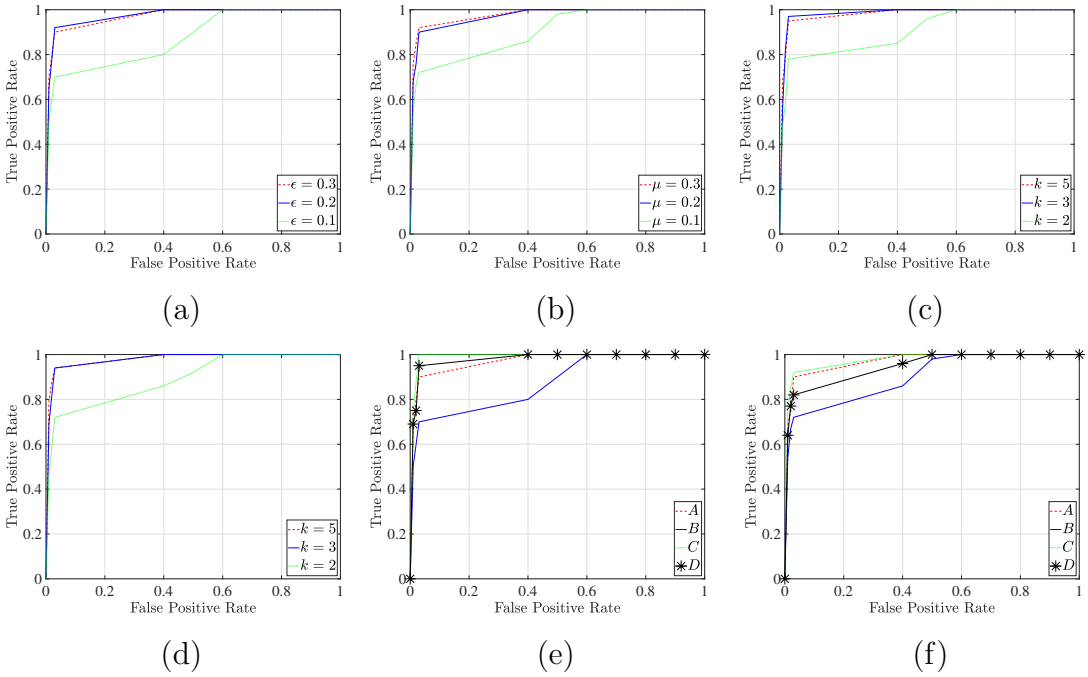


Figure 5.9: (a) ROC curve for location data for various RMSE thresholds ϵ, (b) ROC curve for velocity data for various RMSE thresholds μ, (c) ROC curve for location data for varying k (number of trajectory points), (d) ROC curve for velocity data for varying k, (e) ROC curve for location data across different trajectories as shown in Fig. 5.7(a), (f) ROC curve for velocity data across the same trajectories.

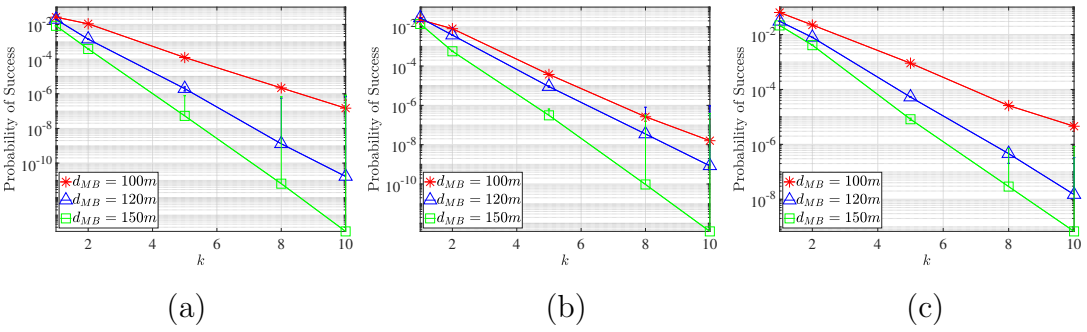


Figure 5.10: (a) Probability of success for remote M in defeating velocity verification, (b) probability of success for remote M in defeating location verification, and (c) probability of success for remote advanced M in defeating velocity verification.

the adversary for B to accept velocity and location, respectively against k the number of messages for d_{MB} distances between B and M, varied between 100m and 150m. From the plot, we observe that for $k = 3$ data points of the trajectory, the probability

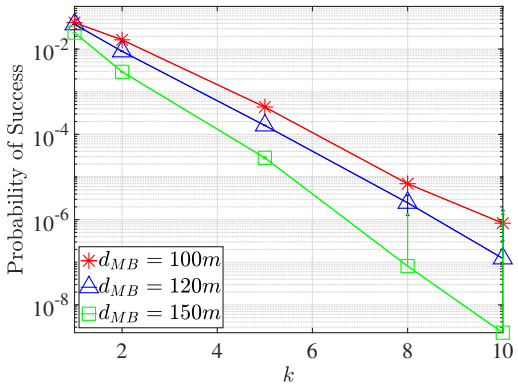


Figure 5.11: Probability of success for remote advanced M in defeating location verification.

of success for the adversary goes down to the level of 10^{-5} for both the velocity and location. Further, we observe that an adversary farther than 120m from the verifier B has a significantly low success probability in defeating VET. Thus, VET can detect an adversary who might be using compromised infrastructure to inject data. Moreover, VET can detect a remote-moving adversary attempting to inject rogue messages. *This attests to our theoretical finding that the probability of success for the remote adversary is a negligible probability.*

Remote Advanced Attacker: Finally, we performed emulation to evaluate the remote advanced attacker using Matlab. We first computed the wireless channels h_{MB} and h_{MX} between the vehicle and the verifier B, and the vehicle and the second verifier X, respectively. The adversary utilized the knowledge of the channel h_{MX} to emulate the trajectory at X, as shown in Fig. 5.7(c). This signal is received by B on the h_{MB} emulated by a ray tracing model. B computed the estimated trajectory using the emulated trajectory to compute the probability of success for the adversary. Figure 5.10(c) and Fig 5.11 show the plot between the probability of success (p^k) from (5.15) for the adversary against k, for the velocity and location, respectively. We observed that an advanced adversary M could defeat VET with a success probability

of 10^{-6} for $k = 5$ trajectory data points. Also, here an adversary further than the distance than $d_{MB} \geq 120m$ is detected with probability $(1 - 10^{-6})$, for both velocity and location. *This attests to our theoretical finding that the probability of success for the remote advanced adversary is a negligible probability.* Hence, the advanced adversary M has to be close to B for defeating VET. Even when M is close to B, the adversary can be detected with certainty when more number k of trajectory points are collected for authentication.

5.6 Chapter Summary

We propose VET: a framework that verifies the veracity of the crypto-credentials by authenticating them against physical trajectory and motion vectors (TMVs). The verifier implements a location and motion-based authentication strategy and verifies the crypto-credentials based on the acceptability of claimed TMVs against randomly estimated TMVs. This detects any adversary from remotely injecting spoofed messages when it is not physically present where it claims to be. We formally analyze the correctness and robustness of VET using matching conversations. Finally, we attest to the findings of theoretical analysis with experimental analysis of VET on the USRP platform. Our experiments show that VET has 97% true positives when operating without an adversary. We fix the threshold values for evaluation based on the ROC curve plotted for the location and velocity data. We evaluate both novice and advanced adversarial behaviors. In the experiments, we show that VET can detect an advanced remote adversary with 99.9% who is capable of manipulating signals with absolute channel knowledge. *In the future*, we plan to expand the experimental evaluations on a UAV platform with both moving provers and verifiers.

CHAPTER 6

Systematization of Knowledge for Security in Molecular and Nano-communications

This chapter is based on joint work with Mr. Malcolm I. Anderson, Mr. Truc T. Duong, Dr. Nirnimesh Ghose, and Dr. Anna Wisniewska. All student authors contributed equally to the research design, literature review, and systematization of knowledge presented herein. My primary contribution focused on developing the security taxonomy and leading the structuring of the bio-inspired security discussion, specifically regarding cybersecurity solutions for molecular and nanocommunications and related ideas.

This chapter presents a comprehensive systematization of knowledge on the security challenges and defense mechanisms in molecular and nano-communication (MC) networks. Molecular and nano-based systems face threats like eavesdropping, denial of service, spoofing, and jamming attacks, as shown in Fig. 6.1. Unlike conventional wireless communication, MC leverages chemical and biological signal propagation, introducing unique vulnerabilities in confidentiality, authentication, integrity, and availability. The chapter surveys prior work that summarizes physical-layer characteristics, energy efficiency, modulation techniques, and error correction in MC systems. It highlights that security remains an underexplored area, with limited systematization of threat models, attack surfaces, and countermeasures tailored to the molecular

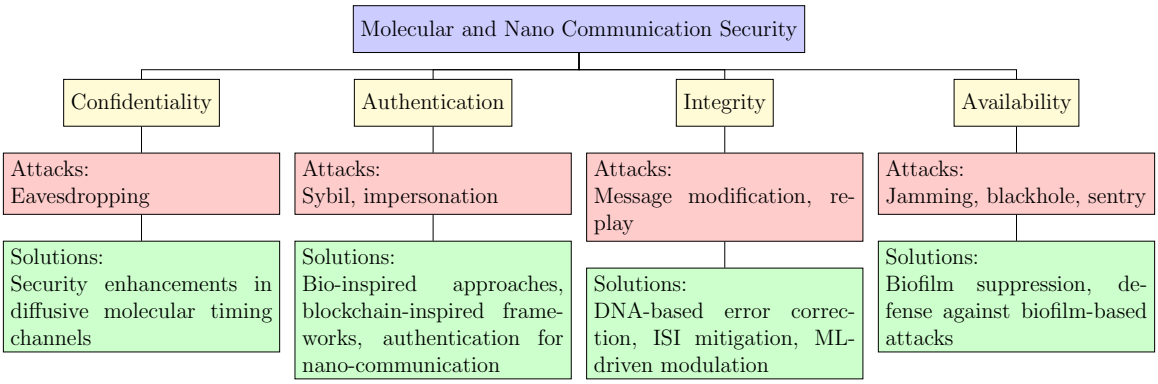


Figure 6.1: Overview of molecular and nano communication security.

scale. By compiling existing surveys and research trends, this chapter identifies critical gaps, such as the lack of end-to-end security frameworks, insufficient analysis of bio-inspired cryptography, and the need for adaptive, noise-resilient mechanisms for sensitive applications like in-body healthcare and targeted drug delivery.

6.1 Prior Surveys and Gaps

The security challenges in molecular and nano-network communication have been a growing concern. Various researchers have surveyed these challenges, focusing on the limitations of traditional methods and exploring bio-inspired cryptography as a potential solution—the need to address vulnerabilities and improve the robustness of molecular and nano-communication systems. For example, a Time-Division Diffusion (TDD) network is increasingly critical due to the devastating consequences of security attacks. Despite these risks, security considerations within molecular communication (MC) systems have historically been underexplored. The unique nature of molecular communication, which relies on transmitting and receiving information through molecules, introduces distinct security challenges that differ significantly from conventional communication systems. This section reviews existing survey efforts in this

domain, emphasizing their contributions to understanding security in molecular and nano-communication networks. Moreover, it identifies critical gaps, such as the lack of comprehensive threat models and insufficient exploration of system-specific vulnerabilities, underscoring the need for targeted research to effectively address security concerns in MC systems.

6.1.1 Prior Systematization of Knowledge

Numerous surveys have summarized the physical layer of molecular communication (MC) systems, providing foundational insights into their unique characteristics and challenges. Abbas *et al.* summarized energy efficiency, emphasizing optimizing power utilization for nano-devices to enable sustainable communication in resource-constrained environments [1]. Hasan *et al.* [70] and Darchini *et al.* [44] surveyed transmission paradigms inspired by social behaviors and biological interactions, aiming to enhance the reliability and effectiveness of molecular networks. Similarly, Kuran *et al.* compiled modulation techniques tailored to molecular communication, presenting strategies to improve signal accuracy and reliability [90].

Farsad *et al.* summarized error correction mechanisms critical for data integrity in the noisy environments characteristic of MC channels [57]. Additionally, Qiu *et al.* reviewed advanced signal processing techniques to enhance system robustness under dynamic and unpredictable conditions [138]. Applications such as in-body healthcare networks were explored by Chude-Okonkwo *et al.*, highlighting the practical implications of molecular communication technologies in real-world scenarios [39].

Despite these advancements, these surveys often provide only cursory treatment of security challenges in MC systems, recognizing them as an essential yet underexplored research area [90, 105, 138]. Loscri *et al.* highlighted the unique vulnerabilities of MC networks, such as the random mobility of molecules and the shared medium,

which make conventional cryptographic approaches ineffective [105]. This gap highlights the need for dedicated security solutions tailored to the molecular propagation environment.

Researchers advocate for bio-inspired security techniques that leverage biological processes to address these challenges and develop innovative cryptographic protocols. These adaptive mechanisms respond to environmental changes, and robust error correction systems are designed for the intrinsic noise and interference of MC systems [105, 139]. These solutions are vital for ensuring the security, reliability, and scalability of MC systems as they expand into critical applications, including healthcare and environmental monitoring.

The prior surveys have significantly systemized the knowledge of molecular and nano-network communication challenges. Dressler *et al.* explored the limitations of conventional cryptographic methods in the context of molecular and nano-communication systems and proposed bio-inspired cryptography as a promising alternative [53]. They highlighted the necessity of developing robust security frameworks tailored to the unique constraints of these systems. Despite the critical consequences of security breaches in molecular and nano-based systems, such as targeted drug delivery (TDD) networks and in-vivo communication systems, the field has historically paid limited attention to addressing security comprehensively [53, 76, 137].

In addition to Dressler's work, several researchers have surveyed and sought to address unresolved security challenges within molecular communication (MC) networks. Loscri *et al.* [105], Andreasson *et al.* [12], and Nakano *et al.* [120] surveyed the development of defense mechanisms tailored to the distinct nature of MC systems. Efforts include leveraging biological systems, such as the immune system, to create novel, bio-inspired techniques for enhancing security [183]. These works highlight the importance of interdisciplinary collaboration and innovative methodologies to ensure

the security and privacy of MC systems while addressing the challenges posed by their unconventional communication paradigms.

6.1.2 Gaps in Systematization of Knowledge

Substantial obstacles persist in the organization of security within molecular and nano-communication systems, although considerable advancements in the examination of essential aspects such as message integrity, confidentiality, authentication and availability which are vital elements for secure communication in MC contexts. Despite progress, current research inadequately addresses some essential facets of security in a systematic manner. The deficiencies in the organization and synthesis of knowledge impede the establishment of a coherent framework for secure molecular communication systems and diminish the practical applicability of existing results. Principal deficiencies emcompass:

6.1.2.1 Lack of Cross-Layer Integration

Existing surveys frequently limit security talks to specific layers of communication protocol stacks, such as the physical or network layer, without investigating how vulnerabilities and solutions spread across layers. For example, Kuran *et al.* outline modulation strategies at the physical layer [90], whereas Farsad *et al.* concentrate on error correction techniques [57]. However, most studies do not explore how remedies at one layer may interact with or impact vulnerabilities at other layers. This divided approach inhibits a comprehensive understanding of end-to-end security, which is required for creating strong molecular communication systems.

6.1.2.2 Limited Focus on Active Threats

There aren't many surveys that methodically cover current dangers unique to molecular communication systems, like denial-of-service (DoS) attacks, spoofing, and malicious interference. Although Dressler *et al.* highlight the drawbacks of conventional cryptography techniques and offer bio-inspired substitutes, they do not go into detail about particular threat models or countermeasures [53] . Similar to this, Loscri *et al.* discuss weaknesses brought on by molecular mobility and shared media, but they don't go into great depth on how active attackers may take use of these characteristics [105]. Furthermore, reviews such as those by Qiu *et al.* underscore secrecy rate and physical layer security however fails to summarize the nature of attacks in developing bio-nano settings [137]. Significant knowledge gaps exist about the effects of active threats on molecular communication systems due to a lack of attention.

6.1.2.3 Inadequate Application-Specific Context

Surveys often generalize security concerns without customizing their study for unique applications, such as targeted drug delivery (TDD), in vivo health monitoring, or environmental monitoring. Chude-Okonkwo *et al.* highlights applications such as TDD but inadequately consider the distinct security needs in these areas [39]. While both Chude-Okonkwo *et al.* and Wang *et al.* emphasizes on biomedical and environmental uses, the generalization restricts the applicability of security insights to certain use situations, hindering the advancement of context-sensitive security systems [39,183]. Abbas *et al.* explored energy efficiency in nano-devices but neglect the distinct security ramifications of healthcare applications, where data confidentiality and integrity are paramount [1]. Current studies rarely examine how the specific demands of these applications affect the design of security protocols, hence constraining

their practical applicability.

6.1.2.4 Absence of Structured Taxonomies

Numerous surveys lack systematic categorization of risks, vulnerabilities, and security measures. Nakano *et al.* offers a comprehensive overview of molecular communication, although they fail to deliver a systematic taxonomy of security concerns and responses [121] . A structured taxonomy would facilitate comparison assessments as noted in Huang *et al.* [76]. The lack of such taxonomies hinders the methodical evaluation of previous work, the comparison of various techniques, and the effective identification of neglected regions.

6.1.2.5 Insufficient Discourse on Standardization and Protocols

Existing surveys seldom address the necessity for standardized security standards in molecular communication systems. Unlike regular communication methods, molecular communication systems do not possess specified security criteria. Surveys conducted by Farsad *et al.* [57] and Kuran *et al.* [90] neglect to highlight this crucial element, which is vital for guaranteeing interoperability and promoting extensive adoption.

6.1.2.6 Inadequate Interdisciplinary Perspectives

Molecular communication entails interdisciplinary collaboration, integrating biology, chemistry, physics, and engineering. Nevertheless, current surveys frequently exhibit a limited scope and neglect to incorporate ideas from burgeoning disciplines such as quantum computing, machine learning, and bioinformatics. Dressler *et al.* [53] and Loscri *et al.* [105] address interdisciplinary approaches but fail to examine how these domains could enhance novel security solutions.

6.2 Cybersecurity Overview for Molecular and Nano-Communication

Molecular communication (MC), nano-communication, and the Internet of Bio-Nano Things (IoBNT) are emerging paradigms enabling communication between nano-scale devices or within biological systems [5, 58, 120]. In MC, information is transmitted by encoding it into the properties or distribution of molecules, allowing communication among biological cells, organisms, or nano-scale devices. Molecules are released into mediums such as water, air, or biological fluids, with information encoded in the molecular concentration [92,98], type [10], or release timing [108,193]. These encoded molecules propagate through the medium to deliver information to a receiver.

MC has been extensively applied in nanotechnology, synthetic biology, and biomedicine. Notable applications include targeted drug delivery [7, 8, 125], environmental monitoring [147], and hybrid communication between synthetic and biological systems [7]. These advancements demonstrate the significant potential of MC in addressing real-world challenges across various domains.

Nano-communication enables nanoscale devices to exchange information using electromagnetic (EM), acoustic, and molecular communication mechanisms. These systems find applications in nanotechnology, nanorobotics, nanomedicine, and nanoscale sensor networks, allowing nano-machines to coordinate and cooperate on tasks such as drug delivery, sensing, and nanoscale assembly activities.

The Internet of Bio-Nano Things (IoBNT) extends the Internet of Things (IoT) paradigm by integrating nanoscale devices and biological components into the IoT ecosystem. IoBNT facilitates seamless communication, data exchange, and control between bio-nano devices, nano-machines, biological systems, and conventional

IoT devices through nano-communication techniques and traditional wireless communication protocols. Researchers leverage IoBNT for diverse applications, including implantable medical devices in healthcare, smart sensors in agriculture, and environmental monitoring systems, where the fusion of biological and nano-scale systems provides innovative and efficient solutions.

Although MC, nano-communication, and IoBNT are primarily designed for nanoscale operations, they often function at the micro-scale. In MC, information-carrying molecules interact with micro-scale biological structures such as cells, tissues, or microfluidic channels. Nano-communication facilitates interactions between nanoscale devices and more extensive systems via EM waves or molecular signaling within microfluidic environments. Similarly, IoBNT applications may involve communication at the micro-scale, enabling interactions between embedded micro-scale sensors, actuators, and devices in biological systems, bridging the gap between nano- and macroscale systems.

Cybersecurity in molecular and nano-communication presents unique challenges due to these technologies' novel and evolving nature. Unlike traditional communication systems, molecular and nano-communication rely on fundamentally different principles, such as the stochastic diffusion of molecules or nanoscale EM wave propagation, which limit the direct applicability of classical encryption methods and security models [76]. The unique constraints of these systems—such as limited computational capabilities, energy constraints, and the physical properties of the communication medium—further exacerbate these challenges.

Research in this field has primarily focused on the physical layer of molecular communication (MC) networks, investigating signal propagation, noise analysis, and channel capacity [2, 66, 67, 103, 132, 133]. While this work has advanced the reliability of MC, it has not fully addressed the security implications. Current methods lack

tailored approaches to ensure confidentiality, integrity, and availability—critical components for securing communication involving sensitive data, such as in healthcare, targeted drug delivery, and environmental monitoring. Next let us dive into various cyber attacks and their effects on MC or nano-communications.

6.2.1 Eavesdropping

Vulnerabilities in molecular communication networks can result in harmful attacks, particularly in scenarios involving malicious devices within the propagation range. These devices can intercept information molecules (IMs) as they diffuse through the medium toward their intended recipient. If the molecular communication (MC) signals propagate in the direction of a malicious device, the attacker can eavesdrop on the transmitted message [118]. This issue is especially critical in diffusion-based and flow-based MC systems, where IMs are freely broadcast throughout the medium to maximize the likelihood of message delivery. These methods inadvertently increase the risk of interception by malicious entities [76, 78, 162, 170].

In healthcare applications, such vulnerabilities can expose private patient data to adversaries, leading to significant privacy breaches. Beyond passive eavesdropping, attackers can execute more damaging active attacks [12]. For instance, intercepted data can be altered to introduce false information, disrupting network operations. Such actions can compromise critical mechanisms, such as the precise timing of drug delivery in targeted drug delivery systems, potentially causing adverse effects on patient health or sabotaging medical treatments. These risks highlight the need for secure communication protocols specifically designed for the unique characteristics of molecular communication systems.

6.2.2 Message Modification Attack

In molecular and nano-communication networks, the limited communication range of nano-devices often necessitates multi-hop transmission, where messages pass through multiple intermediary nodes to reach their intended destination. This multi-hop architecture introduces a vulnerability to message modification attacks, as shown in Figure 6.2a. Malicious nodes between the sender and the receiver can intercept the transmitted message, alter its content, and relay the modified version [140]. Such interference can lead to unintended and potentially harmful outcomes, such as issuing false commands to trigger actions like the premature release of medicine in healthcare applications [40]. Addressing this threat requires robust integrity verification mechanisms and secure routing protocols designed for the unique constraints of molecular and nano-communication networks.

6.2.3 Replay Attack

Malicious nano-devices can exploit previously intercepted messages through eavesdropping to perform replay attacks, as shown in Figure 6.2b. By capturing legitimate commands from authorized devices, attackers can resend these messages repeatedly to disrupt system functionality [122]. This can result in harmful outcomes, such as triggering the unintended and repetitive release of drugs in healthcare applications, potentially leading to overdoses or other severe consequences. To mitigate such threats, molecular and nano-communication networks must implement robust countermeasures, including message authentication codes (MACs), timestamping, and nonce-based mechanisms to ensure the uniqueness and validity of each transmitted command. These solutions are essential to safeguard against unauthorized retransmissions in these sensitive networks [124, 136].

6.2.4 Wormhole Attack

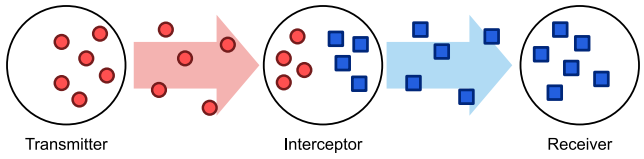
An attacker can create a network wormholes by relaying messages between distant nodes, bypassing the established communication pathways, as shown in Figure 6.3a. This unauthorized link can be exploited to intercept messages, modify their content, or disrupt the network's logical flow, causing confusion and potential system failures. Such attacks compromise the integrity, confidentiality, and reliability of the communication network, making the development of robust routing protocols and secure authentication mechanisms critical to mitigating these vulnerabilities [169].

6.2.5 Sybil Attack

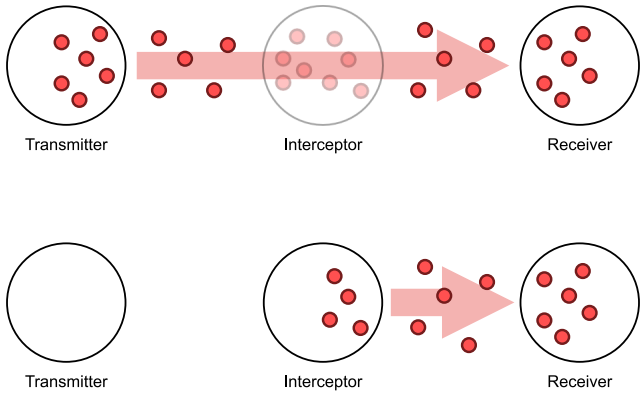
An attacker can deploy multiple fake identities within the network to manipulate communication or gain unauthorized access to the system, as shown in Figure 6.3b. These counterfeit identities can disrupt network operations by propagating misleading data, causing network congestion, or undermining trust-based protocols [105, 191]. Such attacks compromise the integrity and reliability of the system, highlighting the need for robust identity verification and trust mechanisms to safeguard molecular and nano-communication networks.

6.2.6 Impersonation Attack

Unauthorized nano-machines can mimic legitimate devices within the network, gaining access to critical data or disrupting operations, as shown in Figure 6.4. These impersonation attacks pose significant security risks, particularly in systems where accurate device identification is vital [197]. For instance, in healthcare networks, an attacker could impersonate a valid medical nano-machine, allowing access to sensitive patient information or issuing harmful commands [88]. Such actions could lead to de-



(a) Message modification attack



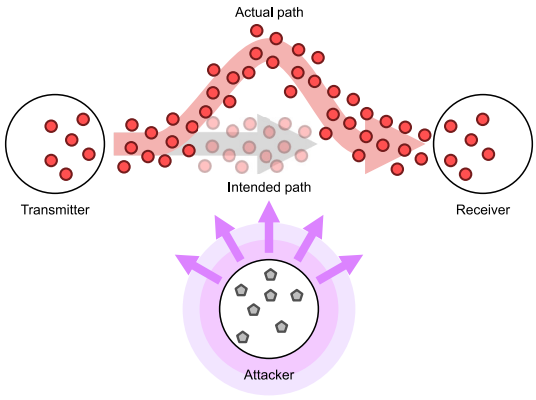
(b) Replay attack

Figure 6.2: Attacks that compromise confidentiality and integrity.

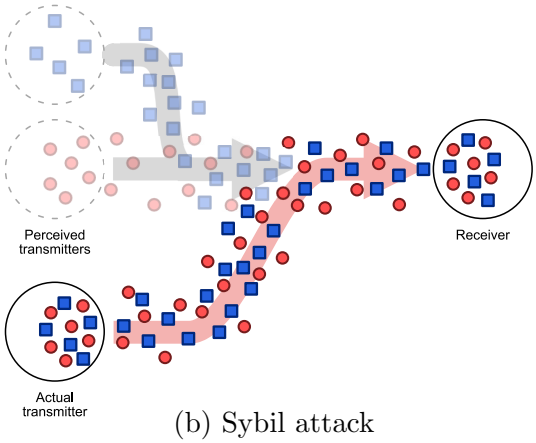
vice malfunctions, erratic behavior, or even failure of critical therapeutic functions, emphasizing the need for stringent authentication protocols and secure communication mechanisms in molecular and nano-communication systems.

6.2.7 Jamming Attack

Attackers can introduce noise or interfering molecules into the communication channel, effectively disrupting signal transmission and reception, as shown in Figure 6.6a. These jamming attacks degrade the signal-to-noise ratio (SNR), making it challenging for receivers to interpret the intended messages accurately [158, 166]. Such attacks are brutal in life-critical systems, where precise molecular communication is essential to trigger medical responses or manage autonomous nano-machines. Disrupted communication in these contexts could result in delayed or incorrect thera-



(a) Wormhole attack



(b) Sybil attack

Figure 6.3: Attacks that compromise confidentiality and integrity.

peutic actions, leading to severe consequences for patient safety or system reliability. Robust noise mitigation and interference detection techniques are critical to countering such threats effectively [22, 65, 112].

6.2.8 Blackhole Attack

Malicious nano-machines can deploy chemo-attractants to lure legitimate nano-machines toward them, as described in prior studies [55,164], as shown in Figure 6.6b. This tactic disrupts the communication flow by diverting nano-machines from their intended trajectories, which is critical due to the limited communication range of these devices. The resulting diversion prevents legitimate nano-machines from perform-

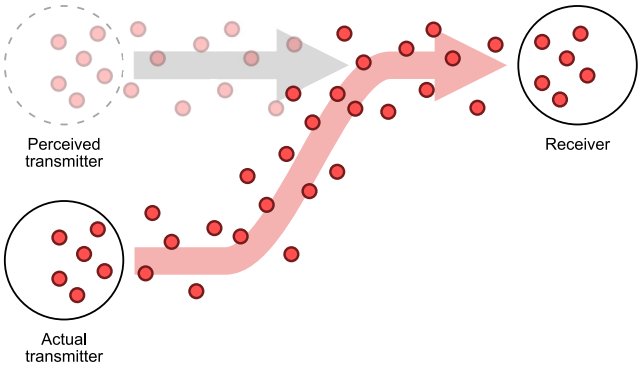


Figure 6.4: Impersonation attack that compromises confidentiality and integrity.

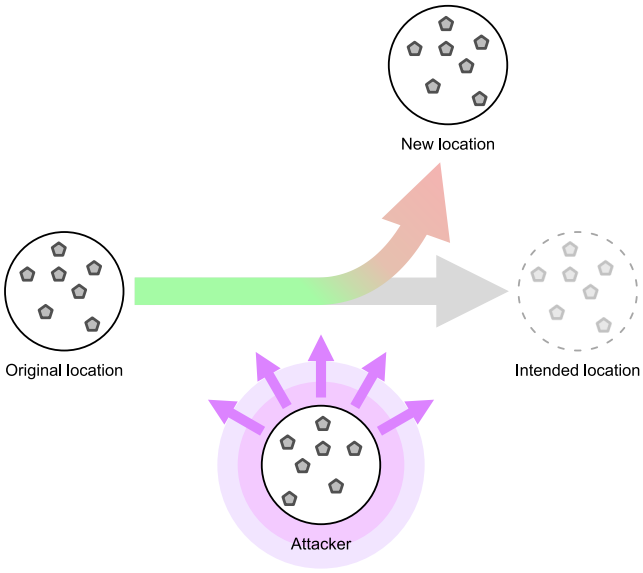


Figure 6.5: Sentry attack that compromises availability.

ing essential tasks, such as drug delivery, or enables other malicious nano-machines to avoid detection and inflict further damage. Additionally, chemo-attractants may interfere with other systems within the medium. For instance, in biological environments, such as the human body, these attractants could lure immune cells like lymphocytes away from infection sites, weakening the immune system's ability to suppress harmful pathogens. This exploitation of chemo-attractant signaling demonstrates a dual threat, disrupting nano-machine networks and compromising natural biological defense mechanisms [63].

6.2.9 Sentry Attack

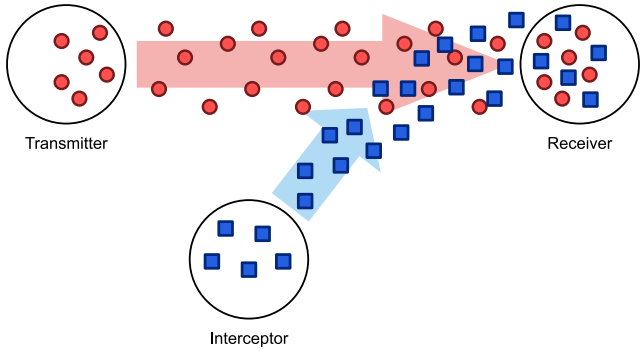
Malicious nano-devices can deploy chemo-repellents to obstruct legitimate nano-devices or nano-machines from reaching their intended destinations, as noted in previous studies [55, 164], as shown in Figure 6.5. This tactic hinders the ability of legitimate devices to carry out their tasks, such as drug delivery or detecting abnormalities caused by malicious nano-devices. By delaying or preventing legitimate nano-machines from sensing and neutralizing abnormal behavior, chemo-repellent attacks amplify the damage inflicted by other malicious entities within the network. Furthermore, this attack can have far-reaching consequences beyond the nano-communication system itself. For instance, chemo-repellents could block the interaction of specific organisms or environmental chemicals, impeding processes such as pollutant breakdown. In such cases, chemical treatments designed to mitigate environmental damage might be ineffective, leading to long-term ecological harm [63].

6.3 Cybersecurity Solutions for Molecular and Nano Communications

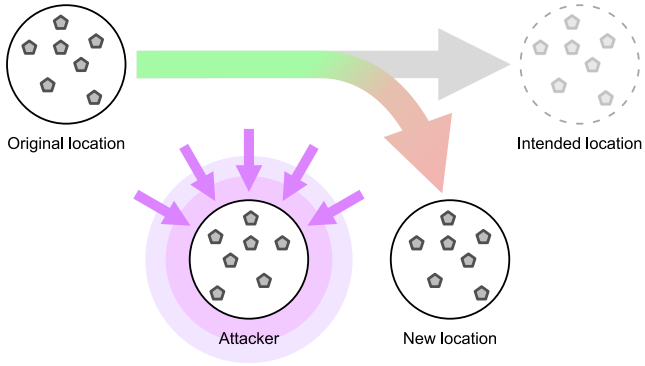
In this Section, we survey the existing solutions to the attacks we discussed in the previous section.

6.3.1 Confidentiality and Privacy

Research on molecular and nano-communication security has primarily focused on ensuring Confidentiality and Privacy [68, 78, 118, 145, 160, 161, 163]. Confidentiality protects information molecules (IMs) from being accessed by unintended entities. In contrast, privacy ensures that the identities and actions of nano-machines remain



(a) Jamming attack.



(b) Blackhole attack.

Figure 6.6: Attacks that compromise availability.

undisclosed. However, maintaining secrecy in these systems is challenging due to the shared propagation medium, which allows eavesdroppers to intercept and decode molecular signals. Such attacks may even involve benign nano-machines unintentionally capturing IMs due to their proximity to the communication range. Detecting eavesdroppers in molecular communication (MC) is particularly difficult because of the nanoscale environment and the stochastic nature of molecular diffusion. Furthermore, nano-machines limited computational power and energy constraints hinder the deployment of classical cryptographic methods for Confidentiality and privacy protection. Confidentiality and privacy techniques such as diffusive molecular timing (DMT) [160] augment secrecy against eavesdropping and replay attacks by incorpo-

rating randomization and obfuscation for adversaries.

6.3.1.1 Secrecy Enhancements in Diffusive Molecular Timing Channels

Sharma *et al.* [160] proposed methods to improve secrecy in Diffusive Molecular Timing (DMT) channels. Their work analyzed the effects of an interferer and an eavesdropper (Eve) on secrecy performance. Introducing an interfering node in the Alice-Eve communication link demonstrated how confusion could be created for Eve, enhancing secrecy in noise and interference-limited environments. In another work, Sharma *et al.* [161] studied the mutual relationship between Bob and Eve by examining their distances from Alice. They found that increasing the variance of Eve's distance negatively affected the system's secrecy. However, their studies primarily focused on point-source scenarios, neglecting the complexities introduced by dynamic channel conditions and more generalized receiver designs.

6.3.1.2 Information Theoretic bounds

Understanding molecular and nano-communication systems requires the application of information theory to establish theoretical bounds on performance metrics, such as channel capacity, mutual information, and secrecy capacity [159]. These bounds are critical for designing efficient and secure communication protocols, especially in complex environments like biological systems. Researchers leverage these methods to optimize molecular communication systems for diverse applications, ensuring reliable and secure data transmission despite the inherent challenges of diffusion-based and nanoscale communication [118, 163, 170].

Channel capacity defines the maximum information transfer rate a communication channel can reliably support. For molecular communication systems, this depends on factors like diffusion rates, molecular degradation, distance, and envi-

ronmental conditions [78]. For instance, in diffusion-based systems, the capacity is constrained by stochastic molecular propagation and inter-symbol interference (ISI). Theoretical expressions for channel capacity in Alice-to-Bob and Alice-to-Eve communication have been modeled and explored extensively [118, 134]. These formulations help characterize and improve communication efficiency in molecular systems.

Mutual information quantifies the information shared between a transmitted signal and the received signal, accounting for the stochastic nature of diffusion and environmental noise [78]. Despite unpredictable diffusion dynamics, this metric evaluates how well information molecules transmit meaningful data. Researchers have analyzed mutual information to optimize system designs for reliable communication in noisy environments [134].

Secrecy capacity represents the maximum secure data transmission rate between two parties (e.g., Alice and Bob) while ensuring that an eavesdropper (Eve) cannot decode the intercepted data. Researchers employ various techniques, such as chemical camouflage or receptor-specific targeting, to enhance secrecy in molecular communication. Probabilistic models for molecule absorption and ISI provide insights into secrecy vulnerabilities and solutions [78]. Secrecy capacity can be expressed mathematically as:

$$C_s = \max_{P_{X|Y}} I(X;Y)_{\mathrm{AB}} - \max_{P_{X|Z}} I(X;Z)_{\mathrm{AE}}. \tag{6.1}$$

where $I(X;Y)_{\mathrm{AB}}$ is the mutual information between the transmitted signal X and the received signal at Bob Y, $I(X;Z)_{\mathrm{AE}}$ is the mutual information between the transmitted information molecule signal X and the received signal at Eve Z. $P_{X|Y}$ and $P_{X|Z}$ are the conditional probability distributions of X given Y and Z respectively. Secrecy capacity can also be represented in terms of their channel capacity [78] as:

$$C_s = \max\{C_{AB} - C_{AE},\ 0\}. \tag{6.2}$$

Interference-Based Security: Sharma *et al.* [160] introduced interference nodes to disrupt eavesdropper communication in diffusive molecular timing (DMT) channels, enhancing secrecy by confusing Eve.

Dynamic Environmental Analysis: Studies have explored how varying distances and diffusion conditions affect secrecy performance [118, 161], providing insights into designing adaptive protocols.

Multi-Node Communication: While most studies focus on single transmitter-receiver scenarios, research by Sabu *et al.* [145] and Huang *et al.* [75] expanded this to multi-receiver setups. These analyses consider factors like molecule hitting probabilities and receiver distance but have yet to incorporate eavesdropping scenarios comprehensively.

Most studies on MC security [117,118,160,161] have focused on single-transmitter, single-receiver setups with a single eavesdropper. However, real-world molecular communication systems often involve multiple transmitters and receivers. Sabu *et al.* [145] extended this by exploring communication between a single transmitter and multiple fully absorbing receivers in a 3D medium. Their analysis provided an analytical expression for the hitting probability of IMs and examined the influence of receiver distance on detection performance. Similarly, Huang *et al.* [75] characterized communication between a single transmitter and two absorbing receivers in a 1D medium. However, neither study incorporated the presence of eavesdroppers, leaving a gap in understanding security implications in multi-node scenarios.

6.3.2 Authentication

Due to the unique nature of molecular communication (MC) and the constraints in developing lightweight security mechanisms, traditional authentication schemes are not directly applicable to molecular and nano-communication systems. Designing authentication methods for nano- and micro-scale biological environments requires significant adaptation to overcome challenges such as limited computational capacity, memory, energy constraints, and environmental variability. Nano-machines often lack the processing power and memory to perform complex encryption or decryption tasks, which makes implementing conventional security mechanisms impractical. Environmental factors, such as molecule degradation in biological fluids or variable diffusion rates, create fluctuating communication channels that complicate consistent authentication. Communication in MC systems is prone to noise, inter-symbol interference (ISI), and unintended interactions with other molecules, affecting the reliability of authentication signals. DNA-based and bio-inspired systems like Bioblock [115] guarantee origin integrity mitigating impersonation, sentry, and Sybil attacks.

6.3.2.1 Bio-Inspired Authentication Approaches

Recent work has explored deoxyribonucleic acid (DNA)-based molecular security devices that adapt traditional access control methods to molecular environments. Zhang *et al.* [200] introduced a robust two-factor authentication system combining DNA strand displacement and enzyme cleavage mechanisms. Their approach utilized molecular locks activated by the correct sequence of nicking enzymes, offering security comparable to electronic password systems. Bu *et al.* [31] leveraged peptides' recognition properties for sensing, encoding, and encryption. They demonstrated secure molecular-level information processing and protection using Pb^{2+}-binding pep-

tides. Liu *et al.* developed cryptographic methods using artificial molecules like p-nitrophenol (PNP) to encode and encrypt data. They incorporated fluorescent molecules to expand detection capabilities and introduced security functions that allowed hidden and encrypted information to be revealed under specific conditions.

6.3.2.2 Blockchain-Inspired Frameworks

Misra *et al.* [115] proposed Bioblock, a blockchain-inspired framework for MC that uses protein-based cryptography. The system uses plasmids as carriers of genetic signatures to create blocks that serve as decentralized records. While Bioblock enhances security and eliminates the need for centralized controllers, its consensus mechanism increases delays and energy consumption, posing a trade-off between performance and robustness.

6.3.2.3 Authentication for Nano-communication

Nano-devices often interact with macro-scale networks, forming a nano-IoT ecosystem. Rana *et al.* [141] proposed an ECC-based framework for secure communication between nano-devices, nano routers, and macro devices. Their system incorporated electromagnetic and molecular communication modes, maintaining a lightweight yet secure approach. Zafar *et al.* [197] introduced an artificial neural network (ANN)-based technique for profiling normal and abnormal behaviors at nano-device interface points. Their particle swarm optimization algorithm enhanced accuracy while maintaining low computational costs. Galal *et al.* [60] explored machine learning models for traffic analysis in electromagnetic nanonetworks. They demonstrated improved packet filtering and anomaly detection using techniques like support vector machines (SVM) and random forests.

6.3.3 Integrity Verification

naIntegrity verification in molecular and nano communication is critical to ensuring the correctness and reliability of transmitted data. However, due nto unique challenges—such as interference, biological noise, environmental variability, and the inherent constraints of nano-devices—innovative, interdisciplinary approaches are required. This section discusses various strategies for achieving data integrity, categorized by error correction, lightweight coding methods, and machine learning applications. Error correction techniques, such as DNA coding and lightweight channel protocols, protect against modification and deterioration during transmission [84, 201].

6.3.3.1 Error Detection and Correction Techniques

Error correction and detection techniques ensure data integrity in molecular and nano communication systems. These methods enable nano-machines to identify and correct errors introduced by noisy channels or malicious interference.

DNA-Based Error Correction: Numerous researchers have employed DNA as a medium for error correction due to its stability and high information density. Erlich *et al.* [54] introduced DNA Fountain, a coding-theoretic approach capable of storing complex datasets, such as an operating system and a movie, within DNA oligonucleotides. This approach enables accurate data retrieval using minimal sequencing coverage. Similarly, Zhang *et al.* [201] proposed a hybrid encoding scheme that combines minimal Hamming distance and multiple sequence alignment to address sequence defects. Kloosterman *et al.* [84] further developed a correction technique for DNA microscopy data, achieving up to 20% error correction, with potential applications for adjacency-based datasets such as Hi-C data [178].

6.3.3.2 Lightweight Channel Coding

Lightweight coding techniques are essential in nano-communication systems, where devices have constrained energy, memory, and processing capacity. These techniques mitigate issues like inter-symbol interference (ISI) while remaining computationally efficient.

ISI Mitigation: Jing *et al.* [79] proposed a lightweight channel coding scheme to mitigate ISI in biosensor networks. By ensuring that codewords lacked consecutive or trailing 1-bits, the scheme avoided molecular accumulation within the channel, improving the bit error rate (BER). This method offers simplicity, reduced time complexity, and suitability for resource-constrained nano-devices.

Probabilistic Constellation Shaping: Tang *et al.* [177] combined Reed-Solomon error correction codes with probabilistic constellation shaping (PCS) to optimize communication in diffusion-based molecular systems. Their approach uses molecular shell mapping to reduce energy consumption while addressing ISI. By integrating error correction with PCS, their scheme improved BER and robustness, even under resource limitations and channel variability.

6.3.3.3 Machine Learning Applications

Machine learning (ML) techniques have been increasingly explored to improve the reliability of molecular communication systems. By leveraging ML models at nano-receivers, it is possible to compensate for channel imperfections and reduce ISI.

ML-Driven Modulation Schemes: Kim *et al.* [82] proposed a modulation scheme called concentration-shifted key position (CSKP) for molecular communication. They integrated a classification-based machine learning model at the nano-receiver, enabling it to outperform conventional maximum likelihood receivers. The

ML model effectively learned and compensated for channel bias, enhancing communication reliability. However, despite these advancements, the application of machine learning in molecular and nano-communication security still needs to be explored.

6.3.4 Availability

The availability and reliability of molecular and nano communication channels are essential for continuous and dependable operation, particularly in critical applications like early disease detection, food safety, and environmental monitoring. Ensuring robust communication requires addressing environmental variability, biological interference, molecular degradation, and resource constraints. However, malicious nano-machines can disrupt communication pathways, rendering legitimate systems ineffective. This section explores the challenges of jamming, biofilm suppression, and denial-of-service (DoS) attacks and discusses proposed strategies to mitigate these disruptions. Resilient availability protocols can counteract wormhole, jamming, and blackhole attacks with adaptive interference detection and quorum-sensing biofilm suppression techniques.

6.3.4.1 Disruptive Attacks in Molecular and Nano Communication

Jamming attacks are intentional disruptions in molecular communication, wherein rogue devices emit signals or reactive molecules to interfere with legitimate transmissions. Such attacks distort communication pathways, degrade signal quality, and increase error rates, potentially causing false alarms or undetected errors.

Luo *et al.* [107] developed a proof-of-concept for nano-plate interfacial jamming, demonstrating its application in controlled molecular diffusion across liquid-liquid interfaces. Using disk-shaped kaolinite nano-particles, they created smart jamming platelets capable of dynamically blocking molecular transport. While disruptive in

malicious contexts, this concept has also inspired targeted applications in drug delivery and chemical separations.

Natural systems offer analogous examples of malicious jamming. For instance, parasitic wasps emit chemical compounds mimicking host pheromones to infiltrate colonies and lay eggs [47]. These biological strategies emphasize the importance of robust security protocols to prevent similar disruptions in molecular systems.

Inter-symbol interference (ISI) and jamming share overlapping effects on molecular communication but differ fundamentally in their origins. ISI occurs unintentionally due to overlapping molecular signals while jamming is a deliberate attack using external interference. Both, however, degrade signal quality and communication reliability. Liu *et al.* [99] explored mitigation strategies for ISI and jamming attacks. While ISI can be managed with error correction methods, jamming necessitates additional measures, such as encoding schemes to obscure molecular signals. Shahbaz *et al.* [158] proposed a jamming-resistant encoding scheme based on hopped repetition codes. The scheme effectively countered memory-less and adaptive jammers by dynamically altering the pre-shared transmission pattern, ensuring secure and robust communication.

6.3.4.2 Biofilm Suppression in Communication Systems

Biofilms can disrupt molecular communication by forming protective barriers that prevent signal propagation. Addressing biofilm formation is critical for maintaining communication efficiency in nano-communication systems.

Strategies for Biofilm Suppression: Martins *et al.* [111] demonstrated how bacteria could cooperate to block nutrients from reaching biofilms, effectively starving and dismantling the biofilm structure. Similarly, Gulec *et al.* [65] leveraged quorum sensing mimickers to disrupt biofilm communication pathways. Their state-based

chemical reaction model effectively simulated biofilm breakdown. Martins *et al.* [112] further utilized proteomic data to identify proteins critical for biofilm formation. By deploying synthetically engineered bacteria to emit jamming signals, they suppressed protein production essential for biofilm stability. The study revealed how the distance and delay of jamming signals influenced biofilm disruption, providing insights for designing effective suppression techniques.

Defense Against Biofilm-Based Attacks: Bernal *et al.* [22] explored strategies to mitigate cyber-bio-attacks, such as distributed denial-of-service (DDoS) attacks on bacteria-based sensors. By employing quorum quenching mechanisms, they demonstrated the ability to reduce biofilm resilience and mitigate DDoS impacts. Quorum amplification is another approach adapted to varying attack intensities, showcasing flexibility in defensive strategies.

6.3.4.3 Impact of Microtubule Jamming

Microtubules, vital components of the cellular cytoskeleton, are responsible for intracellular transport. Jamming of these transport pathways has been linked to neurological disorders, as motor proteins like kinesin and dynein fail to deliver essential cellular cargo. Chahibi *et al.* [35] proposed end-to-end mechanisms to model artificial molecular motor networks and study cargo movements in neurofilaments. Their findings highlight the role of microtubule jamming in brain illnesses and offer insights into creating artificial systems to counteract these effects. Hirokawa *et al.* [73] examined the roles of molecular motors in neural development, plasticity, and disease. Their comprehensive review underscores the importance of efficient transport networks for maintaining cellular health and mitigating the effects of jamming.

6.4 Gaps in Existing Security Solutions

Current security methodologies for molecular and nano communication (MC) systems overwhelmingly emphasize confidentiality. However, they must address the intricate challenges of the shared medium and biological randomness inherent in these systems. Due to the nature of MC, chemical signals are prone to interception by surrounding nano-devices, making it challenging to ensure only the intended recipient accesses the transmitted information.

Despite the extensive focus on confidentiality in recent research [78, 117, 164, 165, 170], practical strategies to secure MC channels remain inadequate. For instance, many models assume simplified point-source transmitters and static interferers [160], limiting their applicability in real-world molecular communication systems. A broader range of release models and receiver designs should be explored to generalize secrecy metrics like diffusive molecular timing channel secrecy.

Furthermore, most evaluations of security mechanisms rely on simulations conducted in controlled, bounded environments [163]. These fail to capture the complexities of real-world scenarios, where variable molecular concentrations, unpredictable movements, and ambient noise significantly impact system security and reliability.

6.4.1 Environmental Influences on Security

Key factors—temperature, distance, diffusion coefficient, and biological diversity—directly affect the performance and security of MC systems. These influences are insufficiently studied despite their profound impact:

Temperature: Temperature variations alter molecular kinetic energy, impacting diffusion rates, reaction kinetics, and signal stability [89]. For example, even minor fluctuations can increase noise or introduce vulnerabilities in signal propagation.

Distance and Diffusion Coefficients: Signal attenuation and delay over long distances amplify interference and reduce effective communication, increasing the likelihood of eavesdropping or denial-of-service (DoS) attacks [56, 117, 160].

Biological Variability: Environmental randomness—such as fluid dynamics and cellular barriers—compounds the challenges of designing secure MC systems. Failure to account for such variability may result in unreliable communication and higher susceptibility to interception or signal disruption [4].

6.4.2 Vulnerabilities to Eavesdropping

Most existing works [53, 76, 137] need to address the risks of eavesdropping in dynamic environments. In dense molecular networks, attackers can intercept chemical signals without specialized equipment, making it crucial to develop robust countermeasures. Research often overlooks real-world complexities, such as fluctuating molecular concentrations and unanticipated noise [117, 118]. Comprehensive security frameworks are needed to safeguard MC systems against unauthorized access and alteration of signals.

6.4.3 Integrity Challenges in MC Systems

Data integrity is essential but increasingly challenging due to noise and biological interference. Noisy molecular channels distort transmitted data, potentially altering information molecules (IMs). For example, reactive oxygen species may oxidize signaling molecules, leading to corrupted communication. Existing integrity mechanisms [32, 93] remain inadequate for dealing with biological noise and environmental changes, necessitating the development of advanced error detection and correction techniques.

6.4.4 Resource Constraints and Lightweight Security Mechanisms

Nano-machines face severe resource limitations, restricting the implementation of traditional security protocols such as AES or RSA encryption [6, 53]. Lightweight alternatives inspired by biological systems, such as immune system modeling and cellular communication pathways, have shown promise [114, 119]. However, research into biologically inspired security solutions remains in its infancy.

6.4.5 Secrecy Capacity in Molecular Channels

The concept of secrecy capacity—the maximum secure data transmission rate—has limited applicability in MC systems due to their stochastic behavior and susceptibility to diffusion and degradation. Studies exploring cooperative jamming, relay nodes, and adaptive modulation [36,81] highlight potential strategies but often require high levels of coordination and computational resources, which are infeasible for nano-devices.

6.5 Possible Solution Ideas to Fill the Gaps

Despite the various gaps in security solutions, we propose some possible ideas for developing robust and accurate security mechanisms tailored to molecular and nano communication systems.

6.5.1 Advanced Cryptographic Methods

The unique constraints and capabilities of molecular and nanoscale communication systems necessitate the development of specialized cryptographic methods. These approaches must be lightweight, resource-efficient, and capable of leveraging the inherent properties of molecular signals and nanoscale materials. Secure and efficient commu-

nication mechanisms can be achieved by integrating advanced encryption techniques with the distinctive characteristics of nano-materials and biochemical processes.

Biochemical and DNA-Based Cryptography: One promising avenue is using biochemical cryptographic techniques, which exploit molecular interactions and natural processes to encode information securely. For example, DNA-based cryptography utilizes the complexity and diversity inherent in DNA sequences to encode and decode data [94]. DNA strands can store vast amounts of information compactly, making this approach particularly effective for systems with strict resource limitations. Unlike conventional digital encryption, which may be computationally infeasible for nano-machines, DNA encryption leverages biological properties to achieve both security and efficiency.

Molecular Steganography: Molecular steganography represents another innovative direction. In this approach, information is concealed within organic carriers, such as the chemical signals that cells use to communicate. By embedding data within these biological pathways, the method ensures discretion and reduces the risk of interception. For instance, chemical signaling molecules can carry encoded information through established biological channels, thereby hiding messages in plain sight [85]. This approach can complement traditional cryptographic techniques by adding a layer of obfuscation.

Enzyme-Based and Bio-Inspired Cryptographic Algorithms: Molecular cryptographic algorithms can further leverage nanoscale biochemical reactions. Enzyme-based cryptographic systems, for example, use highly selective biological reactions to encode and decode data [19, 100, 175]. These processes are energy-efficient and secure, making them ideal for constrained environments like nano-machines.

Another bio-inspired approach involves drawing from natural biological processes, such as protein folding or cellular signaling pathways. These systems are inherently

complex and difficult to predict or replicate, offering high levels of security. Protein folding patterns, for instance, could serve as cryptographic keys, providing immense variability due to the vast number of possible folding configurations [142, 154]. This natural complexity makes such systems highly resistant to brute-force attacks.

Nano-Particle Cryptography: Nano-particles also present significant potential for cryptographic applications. These particles can be engineered to store and transmit cryptographic keys securely. Due to their minuscule size, nano-particles are challenging to detect and intercept, enhancing communication security. Moreover, their unique ability to interact with biological systems offers additional opportunities for secure, context-specific applications [16]. Nano-particles can encode information using physical or chemical properties, such as surface modifications or fluorescence patterns, enabling innovative methods for secure communication.

6.5.2 Enhanced Error Correction Protocols

Ensuring data integrity in molecular communication (MC) systems prone to high noise levels and interference is a critical challenge. To address this, error detection and correction protocols must be tailored to the unique characteristics of MC channels, providing robust and efficient solutions for reliable communication.

DNA-Based Error Correction: A promising approach to error correction in MC involves DNA-based error correction mechanisms. DNA molecules, with their inherent redundancy and capacity to store vast amounts of information, can be optimized for encoding and correcting data errors. When adapted to DNA sequences, techniques like Reed-Solomon (RS) codes introduce redundancy to transmitted data, allowing the receiver to detect and correct errors efficiently [195]. RS codes are particularly advantageous in environments with constrained computational and energy resources, making them well-suited for molecular and nano-communication systems.

These lightweight methods provide high reliability while maintaining low computational complexity, addressing key challenges in MC systems.

Parity and LDPC Codes: Parity codes and low-density parity-check (LDPC) codes are other error correction strategies applicable to MC channels. Parity-based methods introduce additional bits to the transmitted data, enabling error detection and correction at the receiver. LDPC codes, in particular, are highly efficient and effective in noisy environments due to their ability to handle high error rates [144]. The structured design of LDPC codes allows for scalable implementation in MC systems, where resource limitations and noise are significant constraints.

Adaptive error correction methods can dynamically adjust the scheme in response to real-time noise and interference conditions. By applying more robust correction protocols during high-noise periods and reducing redundancy during low-noise intervals, adaptive methods enhance communication efficiency while ensuring reliability [194].

Enhancing Security in Error Correction: While standard error correction protocols address data integrity, they often lack security measures to protect against tampering or unauthorized interception. Enhanced secure error correction protocols, tailored for molecular environments, can address this gap by integrating lightweight cryptographic algorithms with error correction methods:

XOR-Based Masking: Researchers can ensure confidentiality by combining data with a secret key or mask before applying error correction. For instance, XORing the data with a shared secret key renders the information inaccessible to unauthorized parties. This operation adds a layer of protection with minimal computational overhead, making it ideal for resource-limited MC systems.

Lightweight Cryptographic Algorithms: Algorithms such as the Tiny Encryption Algorithm (TEA) and its variation, the Extended Tiny Encryption Algorithm (XTEA),

provide secure and computationally efficient encryption options for molecular communication. These algorithms rely on simple operations like addition, XOR, and bit shifts, ensuring compatibility with the constraints of molecular systems [11,186]. Additionally, lightweight stream ciphers, such as Grain and Trivium, can encrypt data bit by bit, offering efficient encryption for MC systems with modest transmission rates.

Authentication and Integrity Verification: Integrating keyed hash functions, such as HMAC (Hash-based Message Authentication Code), with error correction ensures that only recipients with the correct key can verify message authenticity and integrity [116]. This method helps secure transmitted data against unauthorized modifications while maintaining low computational complexity.

Pseudo-Random Permutations: Using pseudo-random permutations to scramble data with a shared secret key before applying error correction provides an additional layer of obfuscation. This technique ensures only legitimate parties can reverse the permutation and access the transmitted message.

6.5.3 Bio-inspired Security Mechanisms

Designing security mechanisms for molecular and nano-communication (MC) systems using principles inspired by biological systems can provide robust protection against unauthorized access. Leveraging molecular camouflage, selective receptor activation, and other adaptive biological strategies ensures that only designated recipients can decode communication while increasing system resilience against interception.

Molecular Camouflage: Molecular camouflage adapts a biological concept where organisms blend into their environment to evade predators or threats. In the context of MC, this approach involves altering the chemical properties of signal

molecules to mix with environmental background noise, making the communication signal indistinguishable from common metabolites or environmentally benign chemicals. Such a strategy reduces the likelihood of signal detection by unauthorized entities [47, 104]. For example, signal molecules could be engineered to mimic naturally occurring compounds within the medium, effectively disguising the communication process while maintaining functionality for authorized recipients.

Selective Receptor Activation: Selective receptor activation mirrors receptor-ligand interactions in biological systems, where only specific receptors can bind to corresponding ligands. Applying this concept to MC, messages are encoded to ensure that only devices or recipients equipped with matching receptors can decode and interpret the signal [17, 106, 129]. This creates a secure communication channel where intercepted signals remain inaccessible to unauthorized entities. For instance, ligands in the transmitted signal could be designed to bind exclusively to receptors with unique structural or chemical properties, ensuring that only legitimate recipients respond.

Dynamic Security Mechanisms: Biological systems inherently adapt to changing environmental conditions, a principle that can enhance security in MC. By designing communication protocols capable of dynamic security adjustments, systems can respond to potential threats or interference in real-time. For example:

- Signal encoding schemes can modify their structure or encryption keys in response to detected environmental changes or the presence of eavesdroppers [38].
- Dynamic alterations in the properties of signal molecules or receptors could further obscure communication processes, making interception increasingly tricky.

Immune-Inspired Biomimetic Security: Drawing inspiration from the immune system, which dynamically recognizes and neutralizes foreign entities, MC sys-

tems can implement biomimetic security mechanisms to identify and mitigate unauthorized access. Similar to how immune responses adapt to evolving threats, communication networks could:

- Continuously monitor for anomalies in the communication environment.
- Implement countermeasures such as altering molecular signals, scrambling encoded messages, or deploying decoy signals to mislead unauthorized entities.

6.5.4 Hybrid Communication Networks

Hybrid communication networks, which combine electromagnetic (EM) and molecular communication (MC), offer a versatile and robust solution to address the inherent weaknesses of each individual communication mode. By dynamically switching between these modalities based on environmental factors and security needs, these systems improve reliability, adaptability, and security in diverse operational contexts.

Addressing Electromagnetic Interference with Molecular Communication: Electromagnetic interference (EMI) presents a significant challenge to traditional communication systems, particularly in environments with high EM noise, such as industrial settings or conflict zones. Hybrid networks effectively mitigate this limitation by integrating molecular communication, which operates discreetly and independently of electromagnetic channels. Due to its short range and low detectability, MC ensures secure transmission when EM channels are compromised or disrupted. For instance, during signal jamming or eavesdropping, the system can switch to MC to maintain message integrity and confidentiality.

Enhancing Security Through Redundancy: Hybrid networks leverage multichannel redundancy to enhance security and reliability. These systems ensure robust data delivery even during disruptions in one channel by transmitting critical messages

through both MC and EM modalities. This redundancy strengthens the network's resilience to failures caused by signal deterioration, jamming, or other security vulnerabilities [120]. For example, the simultaneous use of EM and MC can safeguard communication by providing alternative pathways for message transmission, effectively reducing the likelihood of data loss or interception.

Advantages of Molecular Communication in Secure Transmissions: MC offers unique advantages in secure communication due to its covert and biologically-inspired nature. Unlike EM signals, which are relatively more straightforward to intercept and decode, molecular signals require specific receptors for detection, making them inherently resistant to eavesdropping [57]. This property makes MC particularly valuable in sensitive applications, such as military operations or industrial settings, where maintaining secrecy is paramount. The ability to encode messages into biochemical signals further enhances security by masking communication within biological processes.

Dual-Layered Security with Modern Encryption Techniques: Hybrid networks enhance security by combining the stealth capabilities of MC with advanced encryption techniques applied to EM communication. This dual-layered approach mitigates security threats, including data interception and unauthorized access attempts. For example, encryption algorithms like AES or lightweight cryptographic methods can secure EM transmissions, while the intrinsic properties of MC ensure low detectability and restricted access to molecular signals. These measures create a comprehensive security framework that protects data integrity and confidentiality.

Application Scenarios for Hybrid Networks:

- **Biomedical Applications:** The Internet of Bio-Nano Things (IoBNT) hybrid networks play a critical role in biomedical applications, such as drug delivery,

where nano-machines interact with the human body to release medication at precise locations. Secure biochemical communication between nano-machines ensures only intended recipients activate, reducing errors and adverse effects. This capability significantly improves the precision and safety of medical interventions, making hybrid networks an essential component of IoBNT systems [157].

- **Environmental Monitoring:** In high-EMI environments, such as industrial zones or disaster areas, hybrid networks enhance the accuracy and security of environmental data collection. Molecular communication ensures reliable transmission of ecological data despite electromagnetic noise, preserving data integrity and improving monitoring precision [192]. This ensures consistent and secure data flow, even in challenging operational settings.

Challenges in Implementing Hybrid Communication Networks: While hybrid communication networks provide significant benefits, several challenges must be addressed to realize their full potential:

- **Mode Switching Protocols:** Designing efficient protocols for seamless transitions between EM and MC modes without data loss or delay is critical.
- **Interference Management:** Ensuring that molecular signals and electromagnetic waves do not interfere with one another during concurrent operation.
- **Energy Efficiency:** Nano-machines involved in molecular communication typically have limited energy resources. Optimizing energy use is essential to prolong operational lifespans.

- **Data Integrity and Synchronization:** Coordinating data transmission across two distinct communication modalities while maintaining synchronization is complex.

6.5.5 Decentralized Authentication Systems

Adaptable and decentralized authentication systems can be designed for nanodevices by leveraging their intrinsic properties and interactions, mimicking biological processes such as quorum sensing, the immune system, and swarm intelligence. These approaches establish trust and authenticate devices without reliance on centralized control, ensuring resilience and scalability.

Quorum Sensing for Decentralized Authentication: Quorum sensing, a biological mechanism bacteria use to coordinate behavior based on population density, offers an effective model for decentralized authentication. Nanodevices can utilize quorum sensing principles to collectively determine device authenticity by exchanging signaling molecules. When a sufficient concentration of signals is detected, a coordinated action—such as granting access or verifying authenticity—can be triggered. This mechanism ensures distributed decision-making and improves security by reducing reliance on a central authority [34].

Immune System-Inspired Key Distribution: The immune system is another biological analogy for decentralized key distribution. As the immune system disseminates antibodies to detect and neutralize pathogens, nanodevice networks can dynamically distribute cryptographic keys. These keys are generated and shared based on interaction patterns within the network, ensuring no single point of failure. This dynamic, distributed approach enhances resilience to attacks, as keys are adapted and redistributed in response to changing conditions or threats [45].

Adaptive Authentication Protocols: Adaptive authentication protocols can

dynamically adjust according to the network's operational requirements and threat levels. These protocols may update trust thresholds, quorum signal requirements, or cryptographic key rotation intervals in response to environmental changes or detected risks. For example, during high-threat periods, the network may increase the number of authentication signals needed or implement stricter validation criteria. This dynamic adaptability enhances the system's robustness, enabling it to respond effectively to evolving threats.

Swarm Intelligence for Self-Organizing Authentication: Nanodevice networks can employ swarm intelligence algorithms inspired by natural systems like ant colonies and bird flocks. In swarm intelligence, individual entities follow simple, localized rules, resulting in emergent, organized behavior. Nanodevices can use similar algorithms to self-organize and manage authentication processes without centralized oversight. Such systems can adapt seamlessly to the addition or removal of devices and continuously optimize authentication protocols in response to network changes. These methods increase scalability and resilience while minimizing vulnerabilities associated with centralized systems [23, 24].

Autonomous Verification Systems: Nanomachines can independently incorporate autonomous verification systems to validate other devices' authenticity. These systems can employ lightweight cryptographic methods, such as public-key cryptography, digital signatures, or hash-based algorithms, to verify device identity and communication integrity. Autonomous verification reduces dependence on a central authority and enhances network resilience against attacks targeting centralized components. Distributed ledger technologies, such as blockchain, could further bolster these systems by ensuring tamper-proof transaction records and facilitating trustless verification [42, 203].

6.6 Chapter Summary

The advancements in molecular and nanoscale communication (MC) systems can potentially revolutionize fields such as healthcare, environmental monitoring, and smart technology. However, these cutting-edge systems introduce unique security vulnerabilities, including susceptibility to eavesdropping, data integrity breaches, and malicious interference. Addressing these challenges necessitates the development of novel security mechanisms tailored to the molecular communication paradigm, such as bio-inspired encryption, efficient error correction methods, and adaptive, resource-efficient protocols.

Future research should focus on integrating security measures that align seamlessly with the biological environments in which these systems operate. Leveraging innovative solutions like DNA-based cryptographic techniques, multi-layered security frameworks, and biomimetic strategies offers a promising path forward. The confidentiality, integrity, and availability of transmitted data must remain at the forefront of MC system design to ensure these networks are robust, reliable, and suitable for real-world applications.

This study provides a comprehensive overview of the current state of security in MC systems, identifies existing gaps, and highlights potential research directions. By addressing these challenges, researchers can drive the development of secure molecular and nanoscale communication systems, facilitating their adoption in critical applications and unlocking their full transformative potential.

Bibliography

[1] A. M. Abbas. Molecular nano communication networks: Architectures, protocols and technologies. In *Proc. of 5th International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT)*, pages 1–5, 2022.

[2] A. Ahmadzadeh, A. Noel, and R. Schober. Analysis and design of multi-hop diffusion-based molecular communication networks. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 1(2):144–157, 2015.

[3] O. B. Akan, H. Ramezani, T. Khan, N. A. Abbasi, and M. Kuscu. Fundamentals of molecular information and communication science. *Proceedings of the IEEE*, 105(2):306–318, 2016.

[4] D. Aktas, B. E. Ortlek, M. Civas, E. Baradari, A. B. Kilic, F. E. Bilgen, A. S. Okcu, M. Whitfield, O. Cetinkaya, and O. B. Akan. Odor-based molecular communications: State-of-the-art, vision, challenges, and frontier directions. *IEEE Communications Surveys & Tutorials*, 2024.

[5] I. F. Akyildiz, F. Fekri, R. Sivakumar, C. R. Forest, and B. K. Hammer. Monaco: fundamentals of molecular nano-communication networks. *IEEE Wireless Communications*, 19(5):12–18, 2012.

[6] I. F. Akyildiz and J. M. Jornet. Electromagnetic wireless nanosensor networks. *Nano Communication Networks*, 1(1):3–19, 2010.

[7] I. F. Akyildiz and J. M. Jornet. The internet of nano-things. *IEEE Wireless Communications*, 17(6):58–63, 2010.

[8] I. F. Akyildiz, M. Pierobon, S. Balasubramaniam, and Y. Koucheryavy. The internet of bio-nano things. *IEEE Communications Magazine*, 53(3):32–40, 2015.

[9] A. Amar and A. J. Weiss. Localization of narrowband radio emitters based on doppler frequency shifts. *IEEE Transactions on Signal Processing*, 56(11):5500–5508, 2008.

[10] G. Aminian, M. Mirmohseni, M. N. Kenari, and F. Fekri. On the capacity of level and type modulations in molecular communication with ligand receptors. In *Proc. of IEEE International Symposium on Information Theory (ISIT)*, pages 1951–1955. IEEE, 2015.

[11] V. R. Andem. *A cryptanalysis of the tiny encryption algorithm*. PhD thesis, University of Alabama Alabama, 2003.

[12] J. Andréasson and U. Pischel. Molecules for security measures: from keypad locks to advanced communication protocols. *Chemical Society Reviews*, 47(7):2266–2279, 2018.

[13] E. Aras, G. S. Ramachandran, P. Lawrence, and D. Hughes. Exploring the security vulnerabilities of LoRa. In *Proc. of IEEE International Conference on Cybernetics (CYBCONF)*, pages 1–6. IEEE, 2017.

[14] H. F. Atlam, R. J. Walters, and G. B. Wills. Internet of nano things: Security issues and applications. In *Proc. of the 2018 2nd international conference on cloud and big data computing*, pages 71–77, 2018.

[15] R. Baker and I. Martinovic. Secure location verification with a mobile receiver. In *Proc. of ACM Workshop on Cyber-Physical Systems Security and Privacy*, pages 35–46. ACM, 2016.

[16] A. C. Balazs, T. Emrick, and T. P. Russell. Nanoparticle polymer composites: where two small worlds meet. *Science*, 314(5802):1107–1110, 2006.

[17] G. Bao and S. Suresh. Cell and molecular mechanics of biological materials. *Nature materials*, 2(11):715–725, 2003.

[18] M. Barbeau, J. Hall, and E. Kranakis. Detection of rogue devices in bluetooth networks using radio frequency fingerprinting. In *proceedings of the 3rd IASTED International Conference on Communications and Computer Networks, CCN*, pages 4–6, 2006.

[19] J. Bath and A. J. Turberfield. DNA nanomachines. *Nature nanotechnology*, 2(5):275–284, 2007.

[20] M. Bellare and C. Namprempre. Authenticated encryption: Relations among notions and analysis of the generic composition paradigm. In *Proc. of International Conference on the Theory and Application of Cryptology and Information Security*, pages 531–545. Springer, 2000.

[21] M. Bellare and P. Rogaway. Entity authentication and key distribution. In *Proc. of Annual international cryptology conference*, pages 232–249. Springer, 1993.

[22] S. L. Bernal, D. P. Martins, and A. H. Celdrán. Towards the mitigation of distributed denial-of-service cyberbioattacks in bacteria-based biosensing systems. *Digital Signal Processing*, 118:103241, 2021.

[23] C. Blum and D. Merkle. *Swarm intelligence: introduction and applications*. Springer Science & Business Media, 2008.

[24] E. Bonabeau, M. Dorigo, and G. Theraulaz. *Swarm intelligence: from natural to artificial systems*. Oxford university press, 1999.

[25] R. Borgaonkar and M. G. Jaatun. 5G as an enabler for secure iot in the smart grid. In *Proc. of First International Conference on Societal Automation (SA)*, pages 1–7. IEEE, 2019.

[26] D. Boshoff, M. M. Zhou, R. E. Nkrow, B. Silva, and G. P. Hancke. Uwb physical layer key sharing using the frequency domain cir magnitude. *IEEE Transactions on Industrial Informatics*, 2025.

[27] C. Boyd, A. Mathuria, and D. Stebila. *Protocols for authentication and key establishment*, volume 1. Springer, 2003.

[28] V. Boyko, P. MacKenzie, and S. Patel. Provably secure password-authenticated key exchange using diffie-hellman. In *Proc. of International conference on the theory and applications of cryptographic techniques*, pages 156–171. Springer, 2000.

[29] V. Brik, S. Banerjee, M. Gruteser, and S. Oh. Wireless device identification with radiometric signatures. In *Proceedings of the 14th ACM international conference on Mobile computing and networking*, pages 116–127, 2008.

[30] Brownfield Ag News. Cyberattacks a growing concern for ag industry. https://www.brownfieldagnews.com/news/cyberattacks-a-growing-concern-for-ag-industry/, 2024. Accessed: Feb. 28, 2025.

[31] Z. Q. Bu, Q. F. Yao, Q. Y. Liu, M. X. Quan, J. Y. Lu, and W. T. Huang. Peptide-based sensing, logic computing, and information security on the antimonene platform. *ACS Applied Materials & Interfaces*, 14(6):8311–8321, 2022.

[32] P. K. Bulasara and S. R. Sahoo. A robust and secure drug delivery with single transmitter and dual symmetrical receivers in an internet of bio-nano things. *IEEE Internet of Things Journal*, 2024.

[33] S. Butterworth. On the theory of filter amplifiers. *Wireless Engineer*, 7(6):536–541, 1930.

[34] A. Camilli and B. L. Bassler. Bacterial small-molecule signaling pathways. *Science*, 311(5764):1113–1116, 2006.

[35] Y. Chahibi, I. F. Akyildiz, and I. Balasingham. Propagation modeling and analysis of molecular motors in molecular communication. *IEEE transactions on nanobioscience*, 15(8):917–927, 2016.

[36] G. Chang, L. Lin, and H. Yan. Adaptive detection and ISI mitigation for mobile molecular communication. *IEEE Transactions on nanobioscience*, 17(1):21–35, 2017.

[37] L. Chouhan and M.-S. Alouini. Interfacing of molecular communication system with various communication systems over internet of every nano things. *IEEE Internet of Things Journal*, 10(16):14552–14568, 2023.

[38] S. Chowdhury, S. Castro, C. Coker, T. E. Hinchliffe, N. Arpaia, and T. Danino. Programmable bacteria induce durable tumor regression and systemic antitumor immunity. *Nature medicine*, 25(7):1057–1063, 2019.

[39] U. A. K. Chude-Okonkwo, R. Malekian, B. T. Maharaj, and A. V. Vasilakos. Molecular communication and nanonetwork for targeted drug delivery: A survey. *IEEE Communications Surveys & Tutorials*, 19(4):3046–3096, 2017.

[40] U. A. K. Chude-Okonkwo, R. Malekian, B. T. Maharaj, and A. V. Vasilakos. Molecular Communication and Nanonetwork for Targeted Drug Delivery: A Survey. *IEEE Communications Surveys & Tutorials*, 19(4):3046–3096, 2017.

[41] J. Cohen. Statistical power analysis for the behavioral sciences. *Hillsdale, NJ: Lawrence Erlbaum Associates*, 1989.

[42] M. Conoscenti, A. Vetro, and J. C. De Martin. Blockchain for the internet of things: A systematic literature review. In *Proc. of IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*, pages 1–6. IEEE, 2016.

[43] B. Danev and S. Capkun. Transient-based identification of wireless sensor nodes. In *2009 International Conference on Information Processing in Sensor Networks*, pages 25–36. IEEE, 2009.

[44] K. Darchini and A. S. Alfa. Molecular communication via microtubules and physical contact in nanonetworks: A survey. *Nano Communication Networks*, 4(2):73–85, 2013.

[45] D. Dasgupta and N. Attoh-Okine. Immunity-based systems: A survey. In *Proc. of IEEE International Conference on Systems, Man, and Cybernetics.*

Computational Cybernetics and Simulation, volume 1, pages 369–374. IEEE, 1997.

[46] B. DeBruhl, S. Weerakkody, B. Sinopoli, and P. Tague. Is your commute driving you crazy? a study of misbehavior in vehicular platoons. In *Proc. of ACM Conference on Security & Privacy in Wireless and Mobile Networks*, pages 1–11, 2015.

[47] K. Dettner and C. Liepert. Chemical mimicry and camouflage. *Annual review of entomology*, 39(1):129–154, 1994.

[48] W. Diffie and M. Hellman. New directions in cryptography. *IEEE transactions on Information Theory*, 22(6):644–654, 1976.

[49] D. Dolev and A. Yao. On the security of public key protocols. *IEEE Transactions on information theory*, 29(2):198–208, 1983.

[50] X. Dong and M. C. Vuran. Spatio-temporal soil moisture measurement with wireless underground sensor networks. In *Proc. of IFIP Med-Hoc-Net*, pages 1–8. IEEE, 2010.

[51] X. Dong and M. C. Vuran. A channel model for wireless underground sensor networks using lateral waves. In *Proc. of IEEE Global Telecommunications Conference-GLOBECOM*, pages 1–6. IEEE, 2011.

[52] X. Dong, M. C. Vuran, and S. Irmak. Autonomous precision agriculture through integration of wireless underground sensor networks with center pivot irrigation systems. *Ad Hoc Networks*, 11(7):1975–1987, 2013.

[53] F. Dressler and F. Kargl. Towards security in nano-communication: Challenges and opportunities. *Nano Communication Networks*, 3(3):151–160, 2012.

[54] Y. Erlich and D. Zielinski. DNA fountain enables a robust and efficient storage architecture. *science*, 355(6328):950–954, 2017.

[55] A. Etemadi, M. Farahnak-Ghazani, H. Arjmandi, M. Mirmohseni, and M. Nasiri-Kenari. Abnormality Detection and Localization Schemes Using Molecular Communication Systems: A Survey. *IEEE Access*, 11:1761–1792, 2023.

[56] N. Farsad, N.-R. Kim, A. W. Eckford, and C.-B. Chae. Channel and noise models for nonlinear molecular communication systems. *IEEE Journal on Selected Areas in Communications*, 32(12):2392–2401, 2014.

[57] N. Farsad, H. B. Yilmaz, A. Eckford, C.-B. Chae, and W. Guo. A comprehensive survey of recent advancements in molecular communication. *IEEE Communications Surveys & Tutorials*, 18(3):1887–1919, 2016.

[58] L. Felicetti, M. Femminella, G. Reali, and P. Liò. Applications of molecular communications to medicine: A survey. *Nano Communication Networks*, 7:27–45, 2016.

[59] D. Freedman, R. Pisani, and R. Purves. *Statistics*. W. W. Norton & Company, 4th edition, 2007.

[60] A. Galal and X. Hesselbach. Machine Learning Models for Traffic Classification in Electromagnetic Nano-Networks. *IEEE Access*, 10:38089–38103, 2022. Conference Name: IEEE Access.

[61] M. Gasca and T. Sauer. Polynomial interpolation in several variables. *Advances in Computational Mathematics*, 12:377–410, 2000.

[62] N. Ghose, L. Lazos, and M. Li. In-band Secret-Free Pairing for COTS Wireless Devices. *IEEE Transactions on Mobile Computing*, 2020.

[63] A. Giaretta, S. Balasubramaniam, and M. Conti. Security Vulnerabilities and Countermeasures for Target Localization in Bio-NanoThings Communication Networks. *IEEE Transactions on Information Forensics and Security*, 11(4):665–676, Apr. 2016.

[64] GroGuru. Groguru wireless underground system (wugs). `https://www.groguru.com/products/`, 2025. Accessed: Feb. 28, 2025.

[65] F. Gulec and A. W. Eckford. A stochastic biofilm disruption model based on quorum sensing mimickers. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 2023.

[66] W. Guo, M. Abbaszadeh, L. Lin, J. Charmet, P. Thomas, Z. Wei, B. Li, and C. Zhao. Molecular physical layer for 6G in wave-denied environments. *IEEE Communications Magazine*, 59(5):33–39, 2021.

[67] W. Guo, T. Asyhari, N. Farsad, H. B. Yilmaz, B. Li, A. Eckford, and C.-B. Chae. Molecular communications: Channel model and physical layer techniques. *IEEE Wireless Communications*, 23(4):120–127, 2016.

[68] W. Guo, Y. Deng, B. Li, C. Zhao, and A. Nallanathan. Eavesdropper localization in random walk channels. *IEEE Communications Letters*, 20(9):1776–1779, 2016.

[69] G. P. Hancke and M. G. Kuhn. An RFID distance bounding protocol. In *Proc. of SECURECOMM*, pages 67–73. IEEE, 2005.

[70] M. Hasan, E. Hossain, S. Balasubramaniam, and Y. Koucheryavy. Social behavior in bacterial nanonetworks: challenges and opportunities. *IEEE Network*, 29(1):26–34, 2015.

[71] H. Hasrouny, A. E. Samhat, C. Bassil, and A. Laouiti. VANET security challenges and solutions: A survey. *Vehicular Communications*, 7:7–20, 2017.

[72] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2nd edition, 2009.

[73] N. Hirokawa, S. Niwa, and Y. Tanaka. Molecular motors in neurons: transport mechanisms and roles in brain function, development, and disease. *Neuron*, 68(4):610–638, 2010.

[74] N. Hu and Y.-D. Yao. Identification of legacy radios in a cognitive radio network using a radio frequency fingerprinting based method. In *2012 IEEE International Conference on Communications (ICC)*, pages 1597–1602. IEEE, 2012.

[75] X. Huang, Y. Fang, A. Noel, and N. Yang. Channel characterization for 1-D molecular communication with two absorbing receivers. *IEEE Communications Letters*, 24(6):1150–1154, 2020.

[76] Y. Huang, M. Wen, L. Lin, B. Li, Z. Wei, D. Tang, J. Li, W. Duan, and W. Guo. Physical-layer counterattack strategies for the internet of bio-nano things with molecular communication. *IEEE Internet of Things Magazine*, 6(2):82–87, 2023.

[77] S. James, R. Raheb, and A. Hudak. UAV swarm path planning. In *Proc. of Integrated Communications Navigation and Surveillance Conference*, pages 2G3–1. IEEE, 2020.

[78] Z. Jia, L. Ma, S. Shen, and X. Jiang. On secrecy performance in D-MoSK-based 3-D diffusive molecular communication system. *IEEE Transactions on NanoBioscience*, 2023.

[79] D. Jing and A. W. Eckford. Lightweight Channel Codes for ISI Mitigation in Molecular Communication Between Bionanosensors. *IEEE Sensors Journal*, 23(13):13859–13867, July 2023. Conference Name: IEEE Sensors Journal.

[80] D. Jing, L. Lin, and A. W. Eckford. Energy allocation for multi-user cooperative molecular communication systems in the internet of bio-nano things. *IEEE Internet of Things Journal*, 2024.

[81] D. Kilinc and O. B. Akan. An information theoretical analysis of nanoscale molecular gap junction communication channel between cardiomyocytes. *IEEE Transactions on Nanotechnology*, 12(2):129–136, 2012.

[82] S.-J. Kim, P. Singh, and S.-Y. Jung. A machine learning-based concentration-encoded molecular communication system. *Nano Communication Networks*, 35:100433, 2023.

[83] A. Kipnis, J. Patarin, and L. Goubin. Unbalanced oil and vinegar signature schemes. In *Proc. of EUROCRYPT*, pages 206–222. Springer, 1999.

[84] A. Kloosterman, I. Baars, and B. Högberg. An error correction strategy for image reconstruction by DNA sequencing microscopy. *Nature Computational Science*, pages 1–9, 2024.

[85] A. Koch, N. Kumar, L. Weber, H. Keller, J. Imani, and K.-H. Kogel. Host-induced gene silencing of cytochrome P450 lanosterol C14α-demethylase–

encoding genes confers strong resistance to Fusarium species. *Proceedings of the National Academy of Sciences*, 110(48):19324–19329, 2013.

[86] R. Kong and H. Chen. Csi-rff: Leveraging micro-signals on csi for rf fingerprinting of commodity wifi. *IEEE Transactions on Information Forensics and Security*, 2024.

[87] R. Kong and H. Chen. Deepcrf: Deep learning-enhanced csi-based rf fingerprinting for channel-resilient wifi device identification. *IEEE Transactions on Information Forensics and Security*, 2024.

[88] N. Kumar and R. Ali. A smart contract-based 6g-enabled authentication scheme for securing internet of nano medical things network. *Ad Hoc Networks*, 163:103606, 2024.

[89] S. Kumari, S. Singh, R. K. Singh, V. Pandey, D. K. Singh, S. P. Singh, and M. Lakshmanan. Performance investigation of molecular nano communication over channels under dynamic scenarios. *Wireless Personal Communications*, 131(1):471–488, 2023.

[90] M. Ş. Kuran, H. B. Yilmaz, I. Demirkol, N. Farsad, and A. Goldsmith. A survey on modulation techniques in molecular communication via diffusion. *IEEE Communications Surveys & Tutorials*, 23(1):7–28, 2021.

[91] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and I. F. Akyildiz. Modulation techniques for communication via diffusion in nanonetworks. In *Proc. of IEEE international conference on communications (ICC)*, pages 1–5. IEEE, 2011.

[92] M. Ş. Kuran, H. B. Yilmaz, T. Tugcu, and B. Özerman. Energy model for communication via diffusion in nanonetworks. *Nano Communication Networks*, 1(2):86–95, 2010.

[93] M. S. Leeson and M. D. Higgins. Forward error correction for molecular communications. *Nano Communication Networks*, 3(3):161–167, 2012.

[94] A. Leier, C. Richter, W. Banzhaf, and H. Rauhe. Cryptography with DNA binary strands. *Biosystems*, 57(1):13–22, 2000.

[95] M. Lepot, J.-B. Aubin, and F. H. Clemens. Interpolation in time series: An introductive overview of existing methods, their performance criteria and uncertainty assessment. *Water*, 9(10):796, 2017.

[96] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of experimental social psychology*, 49(4):764–766, 2013.

[97] B. Li, T. Zhang, and T. Xia. Vehicle detection from 3d lidar using fully convolutional network. *arXiv preprint arXiv:1608.07916*, 2016.

[98] W.-A. Lin, Y.-C. Lee, P.-C. Yeh, and C.-h. Lee. Signal detection and ISI cancellation for quantity-based amplitude modulation in diffusion-based molecular communications. In *Proc. of IEEE Global Communications Conference (GLOBECOM)*, pages 4362–4367. IEEE, 2012.

[99] Q. Liu, P. He, K. Yang, and S. Leng. Inter-symbol interference analysis of synaptic channel in molecular communications. In *Proc. of IEEE International Conference on Communications (ICC)*, pages 4424–4429. IEEE, 2014.

[100] Q. Y. Liu, Z. Q. Bu, M. X. Quan, Y. Wu, X. Ding, L. Q. Xia, J. Y. Lu, and W. T. Huang. A molecular paradigm: "plug-and-play" chemical sensing and crypto-steganography based on molecular recognition and selective response. *Biosensors and Bioelectronics*, 209:114260, 2022.

[101] X. Liu, R. Wu, H. Zhang, Z. Chen, Y. Liu, and T. Qiu. Graph temporal convolution network-based wifi indoor localization using fine-grained csi fingerprint. *IEEE Sensors Journal*, 2025.

[102] Y. Liu, H. Si, G. O. Boateng, X. Guo, Y. Cao, B. Qian, and N. Ansari. Hard sample meta-learning for cir nlos identification in uwb positioning. *IEEE Internet of Things Journal*, 2025.

[103] I. Llatser, I. Pascual, N. Garralda, A. Cabellos-Aparicio, M. Pierobon, E. Alarcón, and J. Solé-Pareta. Exploring the physical channel of diffusion-based molecular communication by simulation. In *Proc. of IEEE Global Telecommunications Conference-GLOBECOM 2011*, pages 1–5. IEEE, 2011.

[104] K. Loeffler-Henry, C. Kang, and T. N. Sherratt. Evolutionary transitions from camouflage to aposematism: Hidden signals play a pivotal role. *Science*, 379(6637):1136–1140, 2023.

[105] V. Loscrí, C. Marchal, N. Mitton, G. Fortino, and A. V. Vasilakos. Security and privacy in molecular communication and networking: Opportunities and challenges. *IEEE Transactions on NanoBioscience*, 13(3):198–207, 2014.

[106] Y. Lu, A. A. Aimetti, R. Langer, and Z. Gu. Bioresponsive materials. *Nature Reviews Materials*, 2(1):1–17, 2016.

[107] J. Luo, M. Zeng, B. Peng, Y. Tang, L. Zhang, P. Wang, L. He, D. Huang, L. Wang, X. Wang, et al. Electrostatic-driven dynamic jamming of 2D nanoparticles at interfaces for controlled molecular diffusion. *Angewandte Chemie International Edition*, 57(36):11752–11757, 2018.

[108] M. U. Mahfuz, D. Makrakis, and H. Mouftah. Spatiotemporal distribution and modulation schemes for concentration-encoded medium-to-long range molecular communication. In *Proc. of 25th Biennial Symposium on Communications*, pages 100–105. IEEE, 2010.

[109] Y. Man, M. Li, and R. Gerdes. {GhostImage}: Remote perception attacks against camera-based image classification systems. In *Proc. of International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2020)*, pages 317–332, 2020.

[110] Y. Man, M. Li, and R. Gerdes. Remote perception attacks against camera-based object recognition systems and countermeasures. *ACM Transactions on Cyber-Physical Systems*, 8(2):1–27, 2024.

[111] D. P. Martins, M. T. Barros, and S. Balasubramaniam. Using competing bacterial communication to disassemble biofilms. In *Proc. of the 3rd ACM International Conference on Nanoscale Computing and Communication*, pages 1–6, 2016.

[112] D. P. Martins, K. Leetanasaksakul, M. T. Barros, A. Thamchaipenet, W. Donnelly, and S. Balasubramaniam. Molecular communications pulse-based jamming model for bacterial biofilm suppression. *IEEE transactions on nanobioscience*, 17(4):533–542, 2018.

[113] S. McKinley and M. Levine. Cubic spline interpolation. *College of the Redwoods*, 45(1):1049–1060, 1998.

[114] B. T. McLelland, B. Lin, A. Mathur, R. B. Aramant, B. B. Thomas, G. Nistor, H. S. Keirstead, and M. J. Seiler. Transplanted hESC-derived retina organoid sheets differentiate, integrate, and improve visual function in retinal degenerate rats. *Investigative ophthalmology & visual science*, 59(6):2586–2603, 2018.

[115] S. Misra, S. Pal, and A. Mukherjee. BioBlock: A Blockchain Analogous Mechanism for Integrity in IoBNT-Based Drug Delivery Systems. *IEEE Systems Journal*, 17(1):1000–1007, Mar. 2023. Conference Name: IEEE Systems Journal.

[116] D. M'Raihi, M. Bellare, F. Hoornaert, D. Naccache, and O. Ranen. RFC 4226: HOTP: An HMAC-Based One-Time Password Algorithm. Technical report, RFC Editor, USA, 2005.

[117] L. Mucchi, A. Martinelli, S. Caputo, S. Jayousi, and M. Pierobon. Secrecy capacity of diffusion-based molecular communication systems. In *Proc. of 13th EAI International Conference on Body Area Networks 13*, pages 103–114. Springer, 2020.

[118] L. Mucchi, A. Martinelli, S. Jayousi, S. Caputo, and M. Pierobon. Secrecy capacity and secure distance for diffusion-based molecular communication systems. *IEEE Access*, 7:110687–110697, 2019.

[119] T. Nakano, S. Ando, N. Takata, M. Kawada, K. Muguruma, K. Sekiguchi, K. Saito, S. Yonemura, M. Eiraku, and Y. Sasai. Self-formation of optic cups and storable stratified neural retina from human escs. *Cell stem cell*, 10(6):771–785, 2012.

[120] T. Nakano, M. J. Moore, F. Wei, A. V. Vasilakos, and J. Shuai. Molecular communication and networking: Opportunities and challenges. *IEEE transactions on nanobioscience*, 11(2):135–148, 2012.

[121] T. Nakano, T. Suda, Y. Okaie, M. J. Moore, and A. V. Vasilakos. Molecular communication among biological nanomachines: A layered architecture and research issues. *IEEE transactions on nanobioscience*, 13(3):169–197, 2014.

[122] V. Nellore, S. Xi, and C. Dwyer. Self-assembled resonance energy transfer keys for secure communication over classical channels. *ACS nano*, 9(12):11840–11848, 2015.

[123] L. H. Nguyen and A. W. Roscoe. Authentication protocols based on low-bandwidth unspoofable channels: a comparative survey. *Journal of Computer Security*, 19(1):139–201, 2011.

[124] V.-L. Nguyen, P.-C. Lin, B.-C. Cheng, R.-H. Hwang, and Y.-D. Lin. Security and privacy for 6G: A survey on prospective technologies and challenges. *IEEE Communications Surveys & Tutorials*, 23(4):2384–2428, 2021.

[125] N. Ntetsikas, S. Kyriakoudi, A. Kirmizis, B. D. Unluturk, A. Pitsillides, I. F. Akyildiz, and M. Lestas. Engineering yeast cells to facilitate information exchange. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 2024.

[126] E. Oguchi and N. Ghose. Vet: Autonomous vehicular credential verification using trajectory and motion vectors. In *International Conference on Security and Privacy in Communication Systems*, pages 140–164. Springer, 2023.

[127] E. Oguchi, N. Ghose, and M. C. Vuran. STUN: Secret-Free Trust-Establishment For Underground Wireless Networks. In *Proc. of IEEE INFOCOM 2022-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 1–6. IEEE, 2022.

[128] A. V. Oppenheim, R. W. Schafer, and J. R. Buck. *Discrete-time signal processing*. Prentice Hall, 1999.

[129] C. O'Connor, E. Brady, Y. Zheng, E. Moore, and K. R. Stevens. Engineering the multiscale complexity of vascular networks. *Nature Reviews Materials*, 7(9):702–716, 2022.

[130] Y. Pan, Z. Xu, M. Li, and L. Lazos. Man-in-the-middle attack resistant secret key generation via channel randomization. In *Proc. of International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, pages 231–240, 2021.

[131] K. Pearson. Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58:240–242, 1895.

[132] M. Pierobon. A systems-theoretic model of a biological circuit for molecular communication in nanonetworks. *Nano Communication Networks*, 5(1-2):25–34, 2014.

[133] M. Pierobon and I. F. Akyildiz. A physical end-to-end model for molecular communication in nanonetworks. *IEEE Journal on Selected Areas in Communications*, 28(4):602–611, 2010.

[134] M. Pierobon and I. F. Akyildiz. Capacity of a diffusion-based molecular communication system with channel memory and molecular noise. *IEEE Transactions on Information Theory*, 59(2):942–954, 2012.

[135] D. Pointcheval and J. Stern. Security proofs for signature schemes. In *Proc. of International conference on the theory and applications of cryptographic techniques*, pages 387–398. Springer, 1996.

[136] S. Prajapat, A. Rana, P. Kumar, and A. K. Das. Quantum safe lightweight encryption scheme for secure data sharing in internet of nano things. *Computers and Electrical Engineering*, 117:109253, 2024.

[137] S. Qiu, Z. Wei, Y. Huang, M. Abbaszadeh, J. Charmet, B. Li, and W. Guo. Review of physical layer security in molecular internet of nano-things. *IEEE Transactions on NanoBioscience*, 2023.

[138] S. Qiu, Z. Wei, Y. Huang, M. Abbaszadeh, J. Charmet, B. Li, and W. Guo. Review of physical layer security in molecular internet of nano-things. *IEEE Transactions on NanoBioscience*, 23(1):91–100, 2024.

[139] S. Qiu, Z. Wei, Y. Huang, M. Abbaszadeh, J. Charmet, B. Li, and W. Guo. Review of Physical Layer Security in Molecular Internet of Nano-Things. *IEEE Transactions on NanoBioscience*, 23(1):91–100, Jan. 2024.

[140] A. Rana, S. Prajapat, P. Kumar, D. Gautam, and C.-M. Chen. Designing a security framework based on hybrid communication in the internet of nano things. *IEEE Internet of Things Journal*, 2023.

[141] A. Rana, S. Prajapat, P. Kumar, D. Gautam, and C.-M. Chen. Designing a Security Framework Based on Hybrid Communication in Internet of Nano

Things. *IEEE Internet of Things Journal*, 11(4):7265–7284, Feb. 2024. Conference Name: IEEE Internet of Things Journal.

[142] F. Rao and A. Caflisch. The protein folding network. *Journal of molecular biology*, 342(1):299–306, 2004.

[143] P. Z. Revesz. A recurrence equation-based solution for the cubic spline interpolation problem. *International Journal of Mathematical Models and Methods in Applied Sciences*, 9(1):446–452, 2015.

[144] T. J. Richardson and R. L. Urbanke. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Transactions on information theory*, 47(2):599–618, 2001.

[145] N. V. Sabu, N. Varshney, and A. K. Gupta. 3-D diffusive molecular communication with two fully-absorbing receivers: Hitting probability and performance analysis. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 6(3):244–249, 2020.

[146] M. Şahin, B. E. Ortlek, and O. B. Akan. Molecular arithmetic coding (MAC) for internet of bio-nano things (IoBNT). *arXiv preprint arXiv:2403.04672*, 2024.

[147] Z. Sakkaff, A. Freiburger, N. Gupta, M. Pierobon, and C. S. Henry. Information- and communication-centric approach in cell metabolism for analyzing behavior of microbial communities. *bioRxiv*, pages 2023–08, 2023.

[148] A. Salam and M. C. Vuran. Wireless underground channel diversity reception with multiple antennas for internet of underground things. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–7. IEEE, 2017.

[149] A. Salam and M. C. Vuran. EM-based wireless underground sensor networks. In *Proc. of Underground sensing*, pages 247–285. Elsevier, 2018.

[150] A. Salam, M. C. Vuran, and S. Irmak. Pulses in the sand: Impulse response analysis of wireless underground channel. In *IEEE INFOCOM 2016-The 35th annual IEEE international conference on computer communications*, pages 1–9. IEEE, 2016.

[151] A. Salam, M. C. Vuran, and S. Irmak. Towards internet of underground things in smart lighting: A statistical model of wireless underground channel. In *Proc. of IEEE 14th International Conference on Networking, Sensing and Control (ICNSC)*, pages 574–579. IEEE, 2017.

[152] A. Salam, M. C. Vuran, and S. Irmak. A statistical impulse response model based on empirical characterization of wireless underground channels. *IEEE Transactions on Wireless Communications*, 19(9):5966–5981, 2020.

[153] M. Schäfer, P. Leu, V. Lenders, and J. Schmitt. Secure motion verification using the doppler effect. In *Proc. of ACM Conference on Security & Privacy in Wireless and Mobile Networks*, pages 135–145, 2016.

[154] H. A. Scheraga, M. Khalili, and A. Liwo. Protein-folding dynamics: overview of molecular simulation techniques. *Annu. Rev. Phys. Chem.*, 58(1):57–83, 2007.

[155] I. J. Schoenberg. Contributions to the problem of approximation of equidistant data by analytic functions: Part a.—on the problem of smoothing or graduation. a first class of analytic approximation formulae. *IJ Schoenberg Selected Papers*, pages 3–57, 1988.

[156] SEMTECH. What are LoRa® and LoRaWAN®? *URL https://lora-developers.semtech.com/library/tech-papers-and-guides/lora-and-lorawan/*, 2020.

[157] A. Setia, R. K. Sahu, S. Ray, R. Widyowati, W. Ekasari, and S. Saraf. Advances in hybrid vesicular-based drug delivery systems: improved biocompatibility, targeting, therapeutic efficacy and pharmacokinetics of anticancer drugs. *Current drug metabolism*, 23(9):757–780, 2022.

[158] S. Shahbaz, M. Mirmohseni, and M. Nasiri-Kenari. A jamming resistant molecular communication scheme. *IEEE Transactions on Molecular, Biological, and Multi-Scale Communications*, 2024.

[159] C. E. Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[160] G. Sharma, R. K. Mallik, N. Pandey, and A. Singh. Effect of interfering transmitter on the secrecy of diffusive molecular timing channels. *IEEE Transactions on Communications*, 2024.

[161] G. Sharma, N. Pandey, A. Singh, and R. K. Mallik. Impact of mutual influence between bob and eve on the secrecy of diffusion-based molecular timing channels. *IEEE Wireless Communications Letters*, 11(11):2255–2259, 2022.

[162] G. Sharma, N. Pandey, A. Singh, and R. K. Mallik. Security in diffusive molecular timing channels: An amount of confusion level perspective. *IEEE Transactions on Molecular, Biological and Multi-scale communications*, 8(3):190–201, 2022.

[163] G. Sharma and A. Singh. Secrecy performance of diffusion based molecular timing channels. *IET Communications*, 15(2):289–304, 2021.

[164] G. Sharma and A. Singh. Secrecy Loss in Diffusive Molecular Timing Channels. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 8(4):297–304, Dec. 2022.

[165] G. Sharma and A. Singh. Secrecy loss in diffusive molecular timing channels. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 8(4):297–304, 2022.

[166] R. Shrestha, H. Guerboukha, Z. Fang, E. Knightly, and D. M. Mittleman. Jamming a terahertz wireless link. *Nature Communications*, 13(1):3045, 2022.

[167] A. R. Silva and M. C. Vuran. Empirical evaluation of wireless underground-to-underground communication in wireless underground sensor networks. In *Proc. of International Conference on Distributed Computing in Sensor Systems*, pages 231–244. Springer, 2009.

[168] A. R. Silva and M. C. Vuran. Development of a testbed for wireless underground sensor networks. *EURASIP Journal on Wireless Communications and Networking*, 2010:1–14, 2010.

[169] H. Singh, M. Bala, S. S. Bamber, and M. Angurala. Detection of wormhole attack via bio-inspired ant colony optimization based trust model in WSN assisted IoT network. *Wireless Personal Communications*, pages 1–22, 2024.

[170] S. P. Singh, S. Yadav, R. K. Singh, V. Kansal, and G. Singh. Secrecy capacity of diffusive molecular communication under different deployments. *IEEE Access*, 10:21670–21683, 2022.

[171] S. W. Smith. *The Scientist and Engineer's Guide to Digital Signal Processing*. California Technical Publishing, 1997.

[172] Soil Scout. Wireless soil moisture sensor for precision agriculture. `https://soilscout.com/solution/wireless-soil-moisture-sensor`, 2025. Accessed: Feb. 28, 2025.

[173] C. Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101, 1904.

[174] D. R. Stinson. *Cryptography: theory and practice*. CRC press, 2005.

[175] M. N. Stojanovic and D. Stefanovic. A deoxyribozyme-based molecular automaton. *Nature biotechnology*, 21(9):1069–1074, 2003.

[176] M. Sun, Y. Man, M. Li, and R. Gerdes. SVM: secure vehicle motion verification with a single wireless receiver. In *Proc. of ACM Conference on Security and Privacy in Wireless and Mobile Networks*, pages 65–76, 2020.

[177] Y. Tang, F. Ji, Q. Wang, M. Wen, C.-B. Chae, and L.-L. Yang. Reed-Solomon Coded Probabilistic Constellation Shaping for Molecular Communications. *IEEE Communications Letters*, 28(2):258–262, Feb. 2024. Conference Name: IEEE Communications Letters.

[178] N. L. Van Berkum, E. Lieberman-Aiden, L. Williams, M. Imakaev, A. Gnirke, L. A. Mirny, J. Dekker, and E. S. Lander. Hi-C: a method to study the three-dimensional architecture of genomes. JoVE, Issue 39, e1869, 2010. DOI: 10.3791/1869.

[179] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning*, pages 1096–1103, 2008.

[180] M. C. Vuran, X. Dong, and D. Anthony. Antenna for wireless underground communication. U.S. Patent No. 9,532,118, Dec. 2016. Issued December 27, 2016.

[181] M. C. Vuran, A. Salam, R. Wong, and S. Irmak. Internet of underground things in precision agriculture: Architecture and technology aspects. *Ad Hoc Networks*, 81:160–173, 2018.

[182] X.-f. Wan, Y. Yang, J. Cui, and M. S. Sardar. Lora propagation testing in soil for wireless underground sensor networks. In *Proc. of Sixth Asia-Pacific Conference on Antennas and Propagation (APCAP)*, pages 1–3. IEEE, 2017.

[183] J. Wang, B. Yin, and M. Peng. Diffusion based molecular communication: principle, key technologies, and challenges. *China Communications*, 14(2):1–18, 2017.

[184] Y. Wang, T. Gu, Y. Zhang, M. Lyu, and H. Li. Exploring a secure device pairing using human body as a conductor. *IEEE Transactions on Mobile Computing*, 2024.

[185] N. P. webmaster. NTP: The network time protocol. `0.us.pool.ntp.org`, 2022.

[186] D. J. Wheeler and R. M. Needham. TEA, a tiny encryption algorithm. In *Fast Software Encryption: Second International Workshop Leuven, Belgium, December 14–16, 1994 Proceedings 2*, pages 363–366. Springer, 1995.

[187] Wisdiam. Recent cyber attacks on the food and agriculture sector. https://wisdiam.com/publications/recent-cyber-attacks-food-agriculture-sector/, 2024. Accessed: Feb. 28, 2025.

[188] Z. Xu, J. Li, Y. Pan, L. Lazos, M. Li, and N. Ghose. PoF: Proof-of-following for vehicle platoons. In *Proc. of The Network and Distributed System Security Symposium (NDSS 2022), San Diego, CA*, pages 1–18. Internet Society, 2022.

[189] Q. Yan, J. Lou, M. C. Vuran, and S. Irmak. Scalable privacy-preserving geo-distance evaluation for precision agriculture iot systems. *ACM Transactions on Sensor Networks (TOSN)*, 17(4):1–30, 2021.

[190] Q. Yan, H. Yang, M. C. Vuran, and S. Irmak. SPRIDE: Scalable and private continual geo-distance evaluation for precision agriculture. In *Proc. of IEEE Conference on Communications and Network Security (CNS)*, pages 1–9. IEEE, 2017.

[191] K. Yang, D. Bi, Y. Deng, R. Zhang, M. M. U. Rahman, N. A. Ali, M. A. Imran, J. M. Jornet, Q. H. Abbasi, and A. Alomainy. A Comprehensive Survey on Hybrid Communication in Context of Molecular Communication and Terahertz Communication for Body-Centric Nanonetworks. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 6(2):107–133, Nov. 2020.

[192] K. Yang, D. Bi, Y. Deng, R. Zhang, M. M. U. Rahman, N. A. Ali, M. A. Imran, J. M. Jornet, Q. H. Abbasi, and A. Alomainy. A comprehensive survey on hybrid communication in context of molecular communication and terahertz

communication for body-centric nanonetworks. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 6(2):107–133, 2020.

[193] H. B. Yilmaz and C.-B. Chae. Simulation study of molecular communication systems with an absorbing receiver: Modulation and ISI mitigation techniques. *Simulation Modelling Practice and Theory*, 49:136–150, 2014.

[194] H. B. Yilmaz, A. C. Heren, T. Tugcu, and C.-B. Chae. Three-dimensional channel characteristics for molecular communications with an absorbing receiver. *IEEE Communications Letters*, 18(6):929–932, 2014.

[195] P. Yin, H. M. Choi, C. R. Calvert, and N. A. Pierce. Programming biomolecular self-assembly pathways. *Nature*, 451(7176):318–322, 2008.

[196] T. J. Ypma. Historical development of the Newton–Raphson method. *SIAM review*, 37(4):531–551, 1995.

[197] S. Zafar, M. Nazir, A. Sabah, and A. D. Jurcut. Securing Bio-Cyber Interface for the Internet of Bio-Nano Things using Particle Swarm Optimization and Artificial Neural Networks based parameter profiling. *Computers in Biology and Medicine*, 136:104707, Sept. 2021.

[198] H. Zhang, X. Wu, and Y. Liu. Channel estimation using denoising autoencoder in ofdm systems. *IEEE Access*, 7:110160–110169, 2019.

[199] X. Zhang and et al. Robust csi-based authentication using deep neural networks and signal pre-processing. *IEEE Transactions on Information Forensics and Security*, 16:1355–1370, 2021.

[200] X. Zhang, Y. Liu, B. Wang, S. Zhou, P. Shi, B. Cao, Y. Zheng, Q. Zhang, and N. Kirilov Kasabov. Biomolecule-driven two-factor authentication strategy for access control of molecular devices. *ACS nano*, 17(18):18178–18189, 2023.

[201] X. Zhang and F. Zhou. An encoding table corresponding to ASCII codes for DNA data storage and a new error correction method HMSA. *IEEE Transactions on NanoBioscience*, 2024.

[202] B. Zhou, V. S. S. L. Karanam, and M. C. Vuran. Impacts of soil and antenna characteristics on LoRa in Internet of Underground Things. In *Proc. of IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE, 2021.

[203] G. Zyskind, O. Nathan, et al. Decentralizing privacy: Using blockchain to protect personal data. In *Proc. of IEEE security and privacy workshops*, pages 180–184. IEEE, 2015.