Smart Sensing and Context-Cognitive Networking In and Beyond mmWave Band: Efficiency, Reliability, and Security

by

Xueyang Hu

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
May 4, 2024

Keywords: 5G, High Frequency Communication, Sensing, context-cognitive networking, security

Approved by

Tao Shu, Chair, Associate Professor of Computer Science and Software Engineering
Shiwen Mao, Samuel Ginn Endowed Professor of Electrical and Computer Engineering
Alvin Lim, Professor of Computer Science and Software Engineering
Richard Chapman, Associate Professor of Computer Science and Software Engineering
Xiaowen Gong, Associate Professor of Electrical and Computer Engineering

Abstract

The demand for mobile data rates has grown tremendously in recent years, but the current low-frequency spectrum is insufficient to support the rapidly growing demand for data rates. Therefore, the upcoming 5G and future 6G technologies move data transmissions into the unused higher frequency bands for more bandwidth. However, moving the network to a higher frequency band also brings additional difficulties, such as reduced network coverage and reliance on the line-of-sight (LoS) path. Fortunately, today's smart wireless systems have already transformed from purely communication networks into integrated systems that combine sensing, computing, and communication. This integration facilitates the aggregation of various data streams to form an intelligent and context-aware system that could overcome the limitations presented by high-frequency radio signals.

This dissertation focuses on improving the efficiency, reliability, and security of networks that work in the mmWave band and beyond, with a special perspective of using smart sensing to obtain environmental information and building cognitive networks to obtain context knowledge. In the first work, we explore the environment perception capability of the commercial off-the-shelf (COTS) mmWave device to build a reflector map of the environment. The map is used to find the backup Non-line-of-sight (NLoS) path directions to sustain the communication when the LoS path is blocked. In the second work, we optimize the Reconfigurable Intelligent Surface (RIS)-aided mmWave network topology by considering the directional communication property. A novel (k,α)-Coverage model is proposed to fully characterize the impact of the path difference on the availability of the path. Two deployment schemes are proposed to address the problem of using the least number of RIS to achieve (k,α)-coverage. In the third work, to ensure the data correctness of light detection and ranging (LiDAR) sensors in autonomous vehicles, we propose a Doppler frequency shift-based physical layer spoofing detection method. A statistical spoofing detection framework is also proposed to jointly consider the impact of short-term uncertainty in vehicle velocity.

Acknowledgments

First, I extend my profound gratitude to my doctoral advisor, Dr. Tao Shu, who has been a beacon of inspiration and wisdom throughout my academic journey. He instilled in me the understanding that pursuing a Ph.D. requires unwavering persistence and an indomitable spirit to never give up, regardless of the obstacles that lie ahead. From him, I learned that perseverance is not merely about enduring but about overcoming and thriving through the rigors of research.

I am also immensely thankful to my dissertation committee, including Dr. Shiwen Mao, Dr. Alvin Lim, Dr. Richard Chapeman, and Dr. Xiaowen Gong, for their invaluable feedback, which significantly enriched my work. I am also deeply grateful to Dr. Diep Nguyen from the University of Technology Sydney for the insightful comments I received.

I am honored to have had the opportunity to work and collaborate with the colleagues at the Wireless Networking and Security Lab (WINGS). My special thanks go to Drs. Tian Liu, Li Sun, Jing Hou, and Jian Chen, as well as Minarul Islam, Amit Das, Hairuo Xu and Guan Huang. In particular, I express my deepest gratitude to Dr. Tian Liu for her exceptional collaborative support and guidance.

Also, I want to acknowledge the financial support that has been received in my research. This includes contributions from the Department of Computer Science and Software Engineering and the National Science Foundation, under the awards CNS-2006998 and CNS-1837034.

My years at Auburn would have been different without all the friends. My special thanks goes to Jiaxiang Ren, Jingwei Liu, Yaoxuan Luan, Tong Li, Gaoxiang Li, Ruoyu Xu, among others.

Lastly, my heartfelt gratitude is extended to my parents, whose constant love and support have been instrumental in shaping me into the best version of myself. Thank you for your endless kindness and for everything you have provided me with, beyond what I could ever imagine.

Table of Contents

Abstract ii

Acknowledgments iv

1 Dissertation Introduction 1

1.1 Background and Motivation 1

1.2 Overview of Research Contributions 4

1.2.1 Environment Perception based Smart Beam Switching for Commercial Off-the-shelf (COTS) mmWave Product 4

1.2.2 Network Topology Optimization for Reconfigurable Intelligent Surface(RIS) Assisted mmWave Directional Communication Network. 5

1.2.3 Physical Layer Spoofing Attack Detection for LiDAR Sensors 6

1.3 Publication Contributions 6

1.4 Dissertation Overview 8

2 Fast and High-Resolution NLoS Beam Switching over Commercial Off-the-Shelf mmWave Devices 9

2.1 Introduction 9

2.2 System Design 14

2.2.1 Problem Statement and Solution Framework 14

2.2.2 Locating Dominant Reflectors 19

2.2.3 Model Driven RSSI Estimation 29

2.2.4 Overhead/Cost Analysis 31

2.2.5 Limitation of the Method and Extension to Larger Space 33

2.3 Testbed and Implementation 34

2.4 Evaluation 35

2.4.1 Test Setting 35

2.4.2 Experimental Results 38

2.5 Conclusions 47

3 (k,α)-Coverage for RIS-aided mmWave Directional Communication 49

3.1 Introduction 49

3.2 Related Work 53

3.2.1 Coverage in RIS 54

3.2.2 Coverage in mmWave Network 54

3.2.3 Coverage in Directional Sensor Network 55

3.3 (k,α)-Coverage Model and Verification 57

3.3.1 Notations 58

3.3.2 (k,α)-Coverage Model and Problem 59

3.3.3 (k,α)-Coverage Verification 61

3.4 (k,α)-Coverage under Deterministic Deployment Scheme 67

3.4.1 The Optimal k-sided Regular Polygon Pattern 68

3.4.2 Optimality Bound 79

3.5 (k,α)-Coverage Probability under Random Deployment Scheme 81

3.6 Numerical and Simulation Results 84

3.6.1 Simulation Setup 85

3.6.2 Simulation Results 85

3.7 Conclusion 94

4 Spoofing Detection for LiDAR in Autonomous Vehicles: A Physical-layer Approach 96

4.1 Introduction 96

4.2 Related Work 100

4.2.1 Attacks Against AV Sensors 100

4.2.2 Perception Model Level Defense Methods 100

4.2.3 Signal Level Defense Methods 101

4.3 LiDAR Working Principle and Vulnerability 102

4.3.1 LiDAR Working Principle 103

4.3.2 Motivating LiDAR Security via Real-world Observations 104

4.4 Doppler Frequency Shift in LiDAR Sensing 107

4.4.1 Doppler Frequency Shift Difference Between Legitimate Signal and Spoofing Signal 107

4.4.2 Feasibility Study of Extracting Doppler Frequency Shift 110

4.5 Doppler Shift based Spoofing Detection 114

4.5.1 Attack Model 1: Static Attacker and Moving LiDAR 115

4.5.2 Attack Model 2: Moving Attacker and Moving LiDAR 116

4.5.3 Attack Model 3: Moving Attacker That Controls Both Its Velocity and Signal Frequency 125

4.5.4 Limitations 128

4.6 Spoofing Detection with Joint Consideration of Velocity and Acceleration 129

4.6.1 Hypothesis Test Formulation 130

4.6.2 Joint Consideration of Velocity and Acceleration. 131

4.6.3 Formulation of $\mathscr{H}_a$ for Power Analysis 132

4.6.4 Settings for Power Analysis 135

4.6.5 Numerical Results of Power Analysis 137

4.6.6 Discussion on Implementation 139

4.7 Applicability Discussion, Future Works , and Conclusion 140

4.7.1 Applicability Discussion 140

4.7.2 Future Work 141

4.7.3 Conclusion 141

5 Conclusion and Future Work 142

5.1 Dissertation Conclusion 142

5.2 Future Work 143

5.2.1 Multi-modal Sensing with mmWave Radar 143

5.2.2 UAV-Assisted Reconfigurable Intelligent Surface Deployment 144

5.2.3 Autonomous Vehicle Cooperative Lidar Security 144

References 146

List of Figures

2.1 First order reflection model. 16

2.2 Single reflector scenario. 21

2.3 Multiple reflectors scenario. 22

2.4 Out-lobe resolving based on measured RSSI (scanning step $\omega = 5°$). 24

2.5 Normalized MSE v.s. N. 26

2.6 NLoS path matching between two different client locations. 28

2.7 System prototype. 36

2.8 Dominant reflector map. 38

2.9 RSSI color map. 39

2.10 Performance of different NLoS links. 40

2.11 RSSI distribution. 41

2.12 The CDF of RSSI estimation error. 42

2.13 Transport layer throughput trace in UDP. 43

2.14 Performance comparison with RSSI. 44

2.15 Transport layer throughput trace in TCP. 46

3.1 A circular list and a point (k,α)-coverage examples. 60

3.2 Feasible region. 63

3.3 Area (k,α)-coverage examples. 65

3.4 Boundary condition for region $\mathscr{A}$. 66

3.5 The $(k, \pi/k)$-coverage area for $k = 3, 4, 6$. 70

3.6 The maximum side-lengths for equilateral triangle, square and regular hexagon patterns w.r.t. the communication radius r. 70

3.7 The $(3, \alpha)$-coverage area in a triangle pattern. 73

3.8 A d_3^* example. 74

3.9 The $(4, \alpha)$-coverage area in square pattern 76

3.10 A d_4^* example. 77

3.11 The (k, α)-coverage probability example for $k = 3$. 83

3.12 The maximum feasible side length d_k^* and the maximum (k, α)-coverage area $\mathscr{S}^*_{(k,\alpha)}$ generated by a pattern. 86

3.13 Performance comparison results. 88

3.14 The simulation and estimation results for Lemma 3.5.1. 91

3.15 $f(k, \alpha, k')$ vs. k'. 92

3.16 The (k, α)-coverage probability result for uniform distribution (solid lines) and Poisson distribution (dashed lines). 93

3.17 $(3, \alpha)$-coverage probability ($r = 0.15$ and $r = 0.25$) for uniform distribution (solid lines) and Poisson distribution (dashed lines). 94

4.1 Normal LiDAR sensing. 103

4.2 Spoofing device. 105

4.3 Outdoor spoofing test environment and setup. 106

4.4 LiDAR point-cloud data. 106

4.5 Doppler frequency shift illustration. 109

4.6 Testbed design. 110

4.7 Doppler shift spectrum results. 113

4.8 Spoofing attack in Scenario 1. 115

4.9 The Cartesian coordinate system and signal tuple. 117

4.10 Cooperative LiDARs 123

4.11 Numerical examples for spoofing detection when attacker controls its velocity. . . 124

4.12 Illustration of a Statistical Test. 133

4.13 The number of samples needed for each road conditions under different spoofed signal proportion. 138

4.14 The number of samples needed for each road conditions under different SNR. . . 138

List of Tables

2.1 Multi-path channel resolving result. 27

2.2 Complexity of each step. 33

4.1 Comparison between related work (Def LV: defense method level, Def Strategies: defense strategies, Phy Inva: physical invariants used) 102

Chapter 1

Dissertation Introduction

1.1 Background and Motivation

As data traffic consumption grows exponentially in wireless networks, the contradiction between capacity requirements and spectrum shortage becomes increasingly prominent [77]. Currently, the widely used frequency bands are already congested with TV and radio signals, along with 4G LTE networks, which mainly span the frequency range between 800 to 3,000 MHz. This saturation challenges not only the efficiency of current wireless services, but also poses significant barriers to the deployment of emerging technologies that require higher data rates, such as virtual reality (VR), augmented reality (AR), 5G networks, and beyond [104, 8, 4].

To tackle the ever-increasing demand on higher transmission rate, one of the most effective solutions is to move the data transmissions into an unused higher section of the spectrum where enormous bandwidths are available. The next-generation 5G network, is planned for the first time to be deployed in a spectrum beyond 6 GHz, in the millimeter wave (mmWave) frequency bands [1, 99, 103]. Compared to existing wireless technologies, mmWave communication has several advantages. First, with a huge bandwidth from 30 to 300 GHz, it can easily support multi-gigabit communication services such as real-time high-definition (HD) streaming and ultra-high-definition video (UHDV). Second, due to the short wavelengths of the mmWave signal, mmWave devices can pack a large number of antenna arrays to support multiple-input and multiple-output (MIMO) [82, 64] transmission while only occupying a small physical space.

Third, with the large number of antennas being used, the mmWave device can utilize beamforming technology [95, 57] to focus the transmission energy to certain directions, which not only increases the signal quality by suppressing the multipath effect but also facilitate the development of other applications, such as detection radars. Furthermore, future 6G technology will operate in the Terahertz (THz) frequency band [29, 24] to harvest even more unused and unexplored spectrum to further improve network throughput.[97, 58, 136, 6].

Although moving the network to a higher frequency band can provide many appealing advantages, how can these physical layer advantages be fully capitalized by higher layers of the network, so that it is ultimately translated into an equally significant user-perceivable throughput/QoS performance gain, still remain as a critical challenge. First, due to the much higher carrier frequencies than those used in conventional wireless technologies, high-frequency signals, particularly those in the mmWave band, are severely attenuated by oxygen [42, 26], leading to increased free-space loss and reduced network coverage. Second, signals experience poor diffraction upon encountering blockages due to their short wavelengths. Specifically, mmWave communication heavily rely on the availability of the line of sight(LoS) path between users and a base station. When the LoS path is blocked by treetops or pedestrians, the mmWave signal cannot circumvent or penetrate through these obstacles, leading to a sudden loss of the signal.

Fortunately, with the expansion of communication bands, the landscape of wireless devices is also undergoing rapid changes, creating new opportunities that could overcome the limitations imposed by current design techniques. Unlike a traditional wireless device that is basically a radio for communication, today's smart wireless system is commonly equipped with powerful application processors and strong multi-modal, multi-functional sensors. Consequently, modern wireless networks are transforming from purely communication-focused networks to integrated systems that combine sensing, computing, and communication capabilities. This integration facilitates the aggregation of various data streams from various channels to form an intelligent, context-aware system that can make smart decisions. Generally, context is a collection of measured and inferred knowledge about the environment where the networking takes place, and context awareness refers to the ability of a system to acquire and reason about context information and adapt to the corresponding applications accordingly [67]. With environmental perception results and network context information, cross-layer optimization approaches can be adopted to achieve end-to-end performance improvement.

The overarching goal of this dissertation is to improve the efficiency, reliability, and security of networks and systems operating in the mmWave and beyond frequency bands, particularly from perspectives of harnessing environmental information and network context information to develop intelligent, cross-layer optimization approaches. This dissertation focuses on two key areas related to this topic.

- **Network Efficiency & Reliability Improvement:** By leveraging the perceptual capabilities and contextual information within the mmWave network, we can aggregate information from other channels with the special physical properties of the network to develope intelligent, context-aware algorithms/methods for efficiency and reliability improvement.

- **Sensing Security Protection:** Exploiting the open nature of the wireless medium, an adversary can easily launch attacks to compromise the security of a sensor and manipulate perception results. We focus on protecting the accuracy and trustworthiness of sensor data, which is a crucial and fundamental point in maintaining the normal operation of the network.

In the first work, we leverage the environmental perception capabilities of commercial off-the-shelf (COTS) mmWave devices to identify dominant reflectors in the environment, which are used to determine suitable NLoS path directions to sustain communication when the LoS path is being blocked. In the second work, our focus shifts from pure perception of the environment to proactive manipulation of the environment. Considering the unique context of directional communication networks, we investigate a network topology optimization problem for the Reconfigurable Intelligent Surface(RIS)-assisted mmWave directional communication network. A novel coverage model, the (k, α)-Coverage, is proposed to fully characterize the impact of the difference in path direction on the availability of the path. We propose two deployment schemes: random and deterministic deployment schemes, to address the problem of using the least number of RIS to achieve (k, α)-coverage. In the third work, we focus on addressing the sensing security problem associated with the modern light detection and ranging (LiDAR) sensor. As LiDARs are susceptible to malicious spoofing attacks, we propose a physical layer spoofing attack detection method that uses the Doppler frequency shift of the signal to verify the sender of the signal and identify potential spoofing attempts. In addition, a statistical spoofing detection framework is proposed to jointly consider the impact of short-term uncertainty in vehicle velocity, which can provide more accurate spoofing detection results in realistic environments.

1.2 Overview of Research Contributions

1.2.1 Environment Perception based Smart Beam Switching for Commercial Off-the-shelf (COTS) mmWave Product

The high directionality of mmWave communication makes its line-of-sight (LoS) path susceptible to blockage when the user is moving. Most existing solutions have very stringent requirements on the antennas of the transmitter and the receiver, which are hardly met by today's consumer-level commercial off-the-shelf (COTS) mmWave products. In reality, a COTS device uses low-resolution wide-beam antennas, and hence cannot support the aforementioned methods for NLoS beam switching in response to the LoS blockage. In this work, we develop a new method to support high-resolution mmWave multi-path channel resolving based on coarse-grained wide-beam phased array antennas. We design a novel real-time beam-switching algorithm that allows COTS devices to estimate the location and reflection coefficient of the dominant reflectors. Whenever the current LoS is blocked, our algorithm can compute in real-time the best alternative beam direction based on estimated reflectors to establish a strong

NLoS link. We implemented the proposed algorithm on a COTS mmWave device and evaluated the system's performance on the physical and transport layer. Our experiments demonstrate the effectiveness of our algorithm on estimating dominant reflectors and calculating strong alternative beam directions, and its efficacy in providing robust connections for COTS mmWave devices.

1.2.2 Network Topology Optimization for Reconfigurable Intelligent Surface(RIS) Assisted mmWave Directional Communication Network.

Reconfigurable Intelligent Surface (RIS) offers a new way to provide controllable non line-of-sight (NLoS) propagation paths for millimeter-wave (mmWave) directional communication to overcome the performance degradation caused by line-of-sight blockage. However, current coverage models do not consider the impact of path direction difference on path's availability, which is a crucial property of mmWave directional communication network. In the work, we propose a new coverage model called (k,α)-coverage. A receiver is (k,α)-covered if it is covered by at least k RISs to have k different NLoS path directions and the angular separation between any two adjacent path directions is at least α . In this case, when the current communication direction is blocked by an obstacle, other RIS created paths are still likely to be available for transmission, which increases the robustness of mmWave directional communication. To tackle the problem of using the least number of RISs to achieve the (k,α)-coverage, we formally define the (k,α)-coverage models and propose methods to verify if the target area is (k,α)-covered by the given set of RISs. Then, we solve the problem under both deterministic and random RIS deployment schemes. For the deterministic deployment scheme, we derive the optimal k-sided regular polygon deployment patterns and use it to achieve area (k,α)-coverage. An analytical performance bound on the number of RISs needed is also derived. For the random RIS deployment scheme, we derive the (k,α)-coverage probability under uniform and spatial-Poisson RIS distributions. Finally, extensive simulation results are provided to validate our analyses.

1.2.3 Physical Layer Spoofing Attack Detection for LiDAR Sensors

Recent years have witnessed the ever-growing interest and adoption of autonomous vehicles (AVs), thanks to the latest advancement in sensing and artificial intelligence (AI) technologies. The LiDAR sensor is adopted by most AV manufacturers for its high precision and high reliability. Unfortunately, LiDARs are susceptible to malicious spoofing attacks, which can lead to severe safety consequences for AVs. Most current work focuses on protecting LiDAR against spoofing attacks by using perception model-level defense methods, whose effectiveness unfortunately depends on the correctness of the LiDAR's sensing outcome. A spoofer thus can elude from these methods as long as it fabricates points that maintain the right contextual relationship held by the legitimate points. In this work, we propose to use the signal's Doppler frequency shift to verify the sender of the signal and detect potential spoofing attacks. To this end, we first thoroughly analyze the working principle of LiDAR and conduct real-world experiments to deeply understand and reveal the vulnerability of LiDAR sensors. We then prove that the Doppler frequency shifts of legitimate and spoofing signals present different characteristics, which can be used to fundamentally protect the LiDAR sensing outcome. For better demonstration purposes, we consider three attack models, including static attacker, moving attacker, and moving attacker with control of both velocity and signal frequency. For each of the models, we first show how the spoofing attack is performed and then present our countermeasures. We then propose a statistical spoofing detection framework to jointly consider the impact of short-term uncertainty in vehicle velocity, which can provide more accurate spoofing detection results in realistic environments. Extensive numerical results are provided in a wide range of settings and road conditions.

1.3 Publication Contributions

During my Ph.D. study, I have contributed to the following publications (listed chronologically)

1. **Hu, Xueyang**, Tian Liu, and Tao Shu. "Fast and High-Resolution Nlos Beam Switching over Commercial off-the-Shelf Mmwave Devices." *IEEE Transactions on Mobile Computing* 21, no. 11 (November 2022): 3956–70. https://doi.org/10.1109/tmc.2021.3064809.

2. **Hu, Xueyang**, Tian Liu, and Tao Shu. "(k, α)-Coverage for RIS-Aided mmwave Directional Communication." *IEEE Transactions on Mobile Computing* 22, no. 12 (December 2023): 7482–97. https://doi.org/10.1109/tmc.2022.3212902.

3. Liu, Tian, **Xueyang Hu**, Hairuo Xu, Tao Shu, and Diep N. Nguyen. "High-Accuracy Low-Cost Privacy-Preserving Federated Learning in IOT Systems via Adaptive Perturbation." *Journal of Information Security and Applications* 70 (November 2022): 103309. https://doi.org/10.1016/j.jisa.2022.103309.

4. Liu, Tian, **Xueyang Hu**, and Tao Shu. "Assisting Backdoor Federated Learning with Whole Population Knowledge Alignment in Mobile Edge Computing." *2022 19th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, September 20, 2022. https://doi.org/10.1109/secon55815.2022.9918550.

5. Liu, Tian, **Xueyang Hu**, and Tao Shu. "Facilitating Early-Stage Backdoor Attacks in Federated Learning with Whole Population Distribution Inference." *IEEE Internet of Things Journal* 10, no. 12 (June 15, 2023): 10385–99. https://doi.org/10.1109/jiot.2023.3237806.

6. **Hu, Xueyang**, Tian Liu, Tao Shu, and Diep Nguyen. "Spoofing Detection for Lidar in Autonomous Vehicles: A Physical-Layer Approach." *IEEE Internet of Things Journal*, 2024, https://doi.org/10.1109/jiot.2024.3371378.

7. Tian Liu, Yue Cui, **Xueyang Hu**, Yecheng Xu, and Bo Liu. On the Convergence of Gossip Learning in the Presence of Node Inaccessibility. *2024 IEEE International Conference on Communications(ICC)*, June 06, 2024.

1.4 Dissertation Overview

In the rest of this dissertation, three works are presented, each addressing a set of problems associated with the efficiency, reliability, and security of the network and system in and beyond the mmWave band. Each chapter focuses on presenting one work, along with comprehensive evaluations and comparisons between the solutions to the state-of-the-art methods. Specifically, the remainder of this dissertation is as follows:

In Chapter 2, we use the perception capabilities of COTS mmWave devices to identify dominant reflectors in the environment. We also introduce a novel method for high-resolution mmWave multipath channel resolving. The proposed algorithm is implemented and evaluated on a COTS mmWave device. Our experiments demonstrate the effectiveness of our algorithm in estimating dominant reflectors and its efficacy in providing robust connections for COTS mmWave devices.

In Chapter 3, we propose a novel coverage model called (k,α)-Coverage for RIS-aided mmWave networks to fully characterize the properties of directional communication. We address the challenge of achieving (k,α)-coverage with the minimum number of RIS units, using both deterministic and random deployment schemes. Extensive simulation results are provided to validate our analyses.

In Chapter 4, we propose a Doppler shift-based physical layer spoofing attack detection method to fundamentally protect the LiDAR result. We prove that Doppler frequency shifts can be used to verify the sender of the signal and detect possible spoofing attacks. A statistical spoofing detection framework is also proposed to provide more accurate spoofing detection results in realistic environments. Extensive numerical results are provided in a wide range of settings and road conditions.

Chapter 5 concludes the dissertation and discusses future research.

Chapter 2

Fast and High-Resolution NLoS Beam Switching over Commercial Off-the-Shelf mmWave Devices

2.1 Introduction

Millimeter-wave (mmWave) communication is considered as one of the most promising technologies for the next generation high-speed wireless networks [41, 91, 90]. In contrast to current WiFi and LTE-based 4G communications that operate at sub-6 GHz frequencies, mmWave network works at a much higher frequency band, therefore is able to provide much wider bandwidth for wireless applications [33, 77, 92]. For example, a mmWave link working at 60 GHz can support a data rate more than 7 Gbps [38]. Such a high-throughput transmission fits well with data-hungry real-time applications such as live high-definition video streaming and virtual reality (VR), which are envisioned to be the dominant killer applications in the era of 5G [14, 86].

Although mmWave provides many desirable features, a big challenge for its practical application is the susceptibility to line-of-sight (LoS) blockage [66, 65, 108]. In particular, the mmWave transmission relies on directional communication to overcome the high oxygen attenuation and the signal's propagation heavily relies on LoS [57, 89]. When the LoS is blocked by an obstacle, the mmWave signal can not penetrate through or circumvent around the obstacle, leading to a significant drop of the received signal strength. In this situation, one solution is to promptly steer the communication beam towards a strong non line-of-sight (NLoS) signal propagation path to maintain the communication [88].

Many methods have been proposed to find such a strong NLoS path in the literature, which can be divided into three categories. The first category uses beam-scanning to search over the space when a strong NLoS path is needed [80, 118, 143, 49]. The proposed methods include sequential scanning through the space and hierarchical scanning, which begins with a low-resolution scanning over the entire space, followed by iterative higher-resolution scannings over particular smaller (finer) ranges of directions that are selected based on the outcome of the previous round lower-resolution scanning. The average overhead of the scanning is usually around 100 ms to 200 ms [131, 25]. The second category assumes a nominal mmWave multi-path channel model, and then attempts to estimate the parameters of this model, including the amplitude, angle of departure (AoD), angle of arrival (AoA), and phase shift of each signal propagation path, by reverse engineering. For example, in [113, 152], based on a measured channel impulse response (CIR), reverse engineering is performed to find the optimal channel parameters that best match the nominal multi-path channel model with the measured CIR. The third category includes those well-studied array signal processing techniques, such as MUSIC [102] and ESPRIT [96, 125], that are pertinent to phased array antennas. These techniques conduct angular spectrum analysis over signals received at each antenna element to resolve the multi-path channel.

Despite their good performance, these existing methods all have very stringent/high requirements on the antennas of the transmitter and the receiver, which are hardly met by today's consumer-level commercial off-the-shelf (COTS) mmWave products. More specifically, since the methods in Category 1 require sequential scanning through N different beam patterns, each of which covers a non-overlapping $(360/N)^\circ$ slice of the space, the fundamental limitation of the methods is that the spatial resolution of the beam scanning is upper bounded by the minimum beam width of the phased-array antenna. When these methods are directly applied to the coarse-grained wide-beam antenna, they may fail to identify those paths that happen to fall within the same beam pattern. As a result, signal propagation paths whose AoDs (or AoAs) are separated less than the minimum beam-width of the antenna would not be distinguishable.

Therefore, to be able to accurately locate a strong NLoS path, these methods require the use of a narrow-beam horn antenna or a high-precision narrow-beam phased array antenna that has a large number of antenna elements, typically costs over \$10K. Clearly such a high price tag is unaffordable to a COTS device. As a matter of fact, current COTS mmWave products typically use a quasi-omni-directional antenna (e.g., 180° beam width) for reception, and a coarse-grained wide-beam (e.g., 60° beam width) phased-array antenna for transmission [15, 72]. Similar issue exists in the methods of Category 2. In particular, to measure the CIR, the receive antenna needs to be able to accurately separate, measure, and report both the amplitude and the phase of each lag of the CIR components (in a typical indoor environment where a COTS mmWave device operates, the length of a lag is in the order of nanoseconds). Such a high-time-resolution CIR time-sequence information is typically not provided by COTS devices. Similarly, the array-signal angular spectrum analysis techniques in Category 3 require the accurate amplitude and phase information of the signal received at each individual antenna element. While a COTS receiver indeed reports the amplitude information of the aggregate signal combined from all antenna elements, it typically does not provide detailed amplitude and phase information of the received signals at individual antenna elements.

Due to the above limitations, the aforementioned methods are not directly applicable to consumer-level COTS mmWave devices. Only recently, several new path resolving methods that are suitable for COTS devices are proposed. Among them, non-coherent compressive path tracking proves to be the most effective algorithm, e.g., see [94, 111, 93, 12, 22]. Based on reverse engineering, this algorithm aims to find the direction of the strongest NLoS path by using only the signal's amplitude information. Instead of relying on a measured CIR, the algorithm probes the channel by sending out a sequence of compressive beacons, each of which is separated apart in time by a dozen microseconds. By measuring the amplitude of each received beacon on the receiver side, the algorithm finds the optimal AoA and AoD that best match the sequence of amplitudes calculated according to the nominal channel model to that of the received beacons. While upon each channel probing (i.e., the transmission of a group of compressive beacons) this method is able to obtain the strongest NLoS path of that moment, it incurs high channel-probing overhead when one needs to keep tracing the change of the strongest NLoS path. For instance, this happens for a user playing an electronic VR game, whereby the direction of the strongest NLoS path keeps changing due to the user's movement. Due to this reason, compressive sensing-based approach is mostly suitable for dynamic application scenarios with frequently changing environment and moving users, under which repetitively probing the channel to catch up with the frequent movement of reflectors in the environment is necessary. However, for those application scenarios where the environment is mostly static but the user could be moving, which are typical in most of the household applications, such repetitive probing may not be an efficient solution, as the reflectors in the environment are hardly changed.

Keeping the limitations and weaknesses of existing methods in mind, in this paper we are interested in developing a new method to achieve high-resolution mmWave multi-path channel resolving result using coarse-grained wide-beam phased array antennas that are commonly equipped on today's COTS mmWave devices. Based on this new method, we further propose an efficient computation-based beam-switching algorithm that can directly predict a strong NLoS path (i.e., without the overhead of per-prediction probing) whenever the LoS blockage happens and a strong NLoS backup path is needed. With these efforts, it becomes feasible for a commercial available device to perform fast and high-resolution NLoS beam switching. Our proposed method is most suitable for static-environment application scenarios, thus fills into the regime where the non-coherent compressive path tracking method does not perform efficiently.

More specifically, to address the challenge of achieving high-resolution multi-path channel resolving based on a coarse-grained wide-beam antenna array, we perform fine-grained spatial scanning of the antenna array and exploit the high spatial resolution of the differential received signal strength (RSS) information measured when the antenna array is turned to point to different directions with small steps. One key insight here is that the wide beam-width of the antenna array does not prevent the array from turning to scan the space in a fine resolution (e.g., with a step of 1° increment in the direction of the antenna beam). The differential RSS information associated with the spatial scanning process, which naturally has a high spatial resolution (e.g., in a resolution of 1°), is then exploited by a novel two-step multi-path channel resolving algorithm. In particular, a low-resolution out-lobe resolving step is first performed to identify the clusters of paths that are separated more than the beam width of the antenna array. Then, for each cluster, a high-resolution in-lobe resolving step is performed, which utilizes reverse engineering to compute the optimal in-cluster fine-grained paths that offer the closest match with the measured RSS of that cluster.

Our reflector-based NLoS beam-switching mechanism is then built upon the above channel resolving process. In particular, our method consists of two phases: the offline site survey phase and the online operational phase. In the site survey phase, our model aims to construct a reflector map by estimating the locations and reflection coefficients of the dominant reflectors in the environment, through a sequence of coordinated differential RSS measurements at multiple locations. At each location, the above channel resolving process is called to compute the top-K strongest NLoS paths generated by the dominant reflectors. Exploiting the sparse nature of the mmWave channel, the NLoS paths computed at different locations are then used to estimate the location of the dominant reflectors. Furthermore, based on the Fresnel reflection model assumption, the reflection coefficient of each dominant reflector is calculated by a minimum mean square error (MMSE) estimator based on the RSS measurements. Note that the offline site survey phase is basically a one-time operation for static or quasi-static environments. The next offline site survey is not needed until there is a significant change on the layout of the environment (e.g., a new steel furniture is just added so the reflection layout is changed).

This reflector map is subsequently used in the online operational phase to calculate the supposedly strongest NLoS path at the current location of the user. Note that during this phase, our system can instantly calculate the NLoS path and does not require any additional probing effort. The main beam of the transmit antenna is then steered accordingly to maintain the ongoing connection when the LoS is blocked.

To verify the performance of the proposed method, we implement our algorithm on a COTS mmWave device MikroTik WAP 60G transceiver set [72].The system is tested in an indoor environment for both static and mobile applications. The results show that our system is able to accurately estimate the locations of the strong reflectors in the test environment. In case of LoS blockage, by steering the transmit antenna towards the directions indicated by the proposed algorithm, the system is able to achieve a 200% to 300% throughput gain over the case that the transmit antenna is always pointing to the LoS direction. Our received signal strength indicator (RSSI) measurement at the physical layer also shows that our algorithm can recover the link performance more quickly from blockage and achieves better stability than the device's built-in 802.11ad based method.

The remainder of the paper is organized as follows. We present our system design in Section 2. The test-bed implementation is described in Section 3. The test settings and test results are presented in Section 4. And we conclude our paper in Section 5.

2.2 System Design

2.2.1 Problem Statement and Solution Framework

We consider an indoor mmWave communication scenario where the room layout is static, and there is only one link consisting of one mmWave access point (AP) and one mmWave adapter (referred to as the client). Without loss of generality, we consider an uplink case: the AP is the receiver and the client is the transmitter. Just like those consumer-level off-the-shelf mmWave products, we assume the AP uses a quasi-omni-directional beam pattern for receiving, while the client is equipped with a low-end phased array antenna with coarse beamforming capability for directional transmission.

For indoor mmWave communication, a strong NLoS path is used to maintain the connection when LoS is being blocked. In a static communication scenario where both the transmitter and the receiver are located at fixed positions, the NLoS paths should also remain static therefore can be measured in advance. However, what we consider here is a more common but challenging scenario: the AP is static but the client is mobile, e.g., in a wireless VR game, making the NLoS paths change as the client moves . The pre-measurement method will fail under this scenario due to the infinite number of possible transmitter locations. So how to determine the real-time NLoS paths for a mobile transmitter is the problem we are trying to solve here.

Fortunately, the strong NLoS paths are not randomly distributed. Instead, they are heavily dependent on the locations of the transmitter, the receiver, and the mmWave signal reflectors in the environment. Specifically, the high frequency of mmWave and the usage of transmit antenna array make the communication quasi-optical: in a typical indoor environment such as an office or home, most of the strong NLoS paths are formed by the first-order specular reflection from reflectors in the environment. Note that we are mainly focused on indoor applications in a small-to-moderate-size room, where the likelihood for the existence of some reflective surfaces, such as concrete walls and book shelves, is high. Due to the wide availability of these reflective objects, the NLoS paths generated by these reflectors can cover most area of the room. Therefore the assumption that there exists a strong NLoS path should be reasonable for the indoor scenarios considered in this work.

Based on this fact, we take a generative method to solve our real-time NLoS paths resolution problem. In particular, we intend to create a model for dominant reflectors in the environment based on some site survey process. This model describes the location, orientation, and reflection coefficient of each dominant reflector. We then put the model into operation: at a given client location, the real-time NLoS paths are simply computed as the specular reflections generated by those dominant reflectors according to the model.

This idea is better illustrated in Figure 2.1, where the locations of the client and the AP are (x_t, y_t) and (x_r, y_r), respectively. And we assume there is only one reflector R (the solid black line) for simplicity of presentation. Clearly, given (x_t, y_t), (x_r, y_r), and the location and orientation of the reflector R, there is a unique path, highlighted in yellow in Figure 2.1, by which the

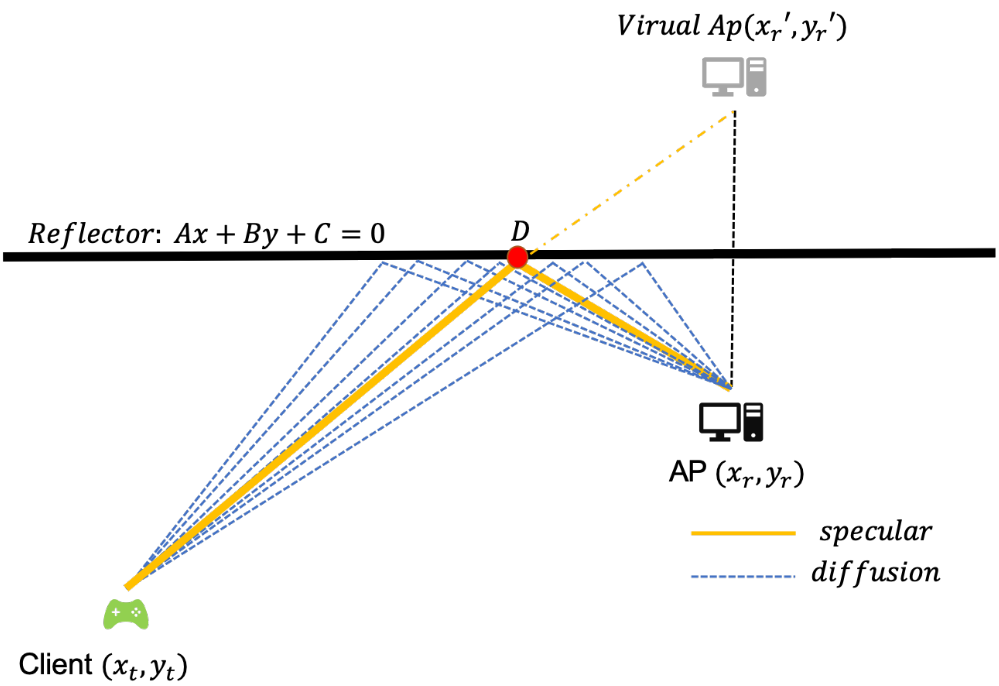


Figure 2.1: First order reflection model.

signal transmitted from the client can be specularly reflected by the reflector, and received by the AP. The uniqueness of this path is owing to the law of reflection, i.e., in the case of specular reflection, for each incident ray, the angle of incidence equals the angle of reflection. It is easy to verify that point D is the only position on the reflector through which the incident ray from the client can be specularly reflected and received by the AP, while other positions on the reflector contribute to the weak diffusive scattering, as denoted by the blue dotted lines. Here, point D is the intersection between reflector R and the line segment connecting the client and the mirrored image of the AP (referred to as virtual AP), defined w.r.t. the reflector. Consequently, the reflection-induced NLoS path cluster simply consists of a strong specular reflection path (the solid yellow path), surrounded by a set of weak diffusive reflection paths (the dotted blue paths).

The insight is that point D can be uniquely decided by the AoD (say α) and the AoA (say β) of the specular reflection path: it is just the intersection between a line passing through the client with an orientation of α and a line passing through the AP with an orientation β. Therefore, given α and β are known, the location and orientation of the reflector R can be uniquely decided by two independent sets of AP and client locations. However, β is unknown in our problem setting, due to the omni-directional receive antenna of the AP. This condition poses challenges to the reflector modeling.

We propose the following approach to address the above challenge in the localization of R. The reflector in a 2-D space can be modeled as a line segment with math representation $Ax + By + C = 0$ $(A, B, C \in \mathbf{R})$, where $\mathbf{R}$ is the set of real numbers. In this assumption, we do not consider the actual length of the reflector, because in reality the size of a dominant reflector is usually big enough to cover most of the locations through reflection in a small-to-moderate-size room, which is the setting of interest considered in this work. Let the AP's position be (x_r, y_r), and the client's location be (x_t, y_t). Denote the location of the virtual AP by (x_r', y_r'). The location of the virtual AP satisfies the following condition:

$$\begin{cases} A\dfrac{x_r + x_r'}{2} + B\dfrac{y_r + y_r'}{2} + C = 0 \\ \dfrac{A(y_r' - y_r)}{B(x_r' - x_r)} = 1. \end{cases} \tag{2.1}$$

The virtual AP's location can be further calculated as:

$$\begin{cases} x_r' = \dfrac{(B^2 - A^2)x_r - 2ABy_r - 2AC}{(A^2 + B^2)} \\ y_r' = \dfrac{(A^2 - B^2)x_r - 2ABx_r - 2BC}{(A^2 + B^2)}. \end{cases} \tag{2.2}$$

Next we will find another constraint for the virtual AP's location. Let ϕ be the angle of departure (AoD) of the specular ray ($\phi \neq \frac{\pi}{2}$). Then the incidence part of the specular reflection path can be represented by:

$$tan(\phi)(x - x_t) = (y - y_t). \tag{2.3}$$

The incidence ray follows the law of reflection: the incident angle equals to the reflection angle. So the location of the virtual AP must satisfy Eq.(2.3). Substituting (x'_r, y'_r) into the equation, we get:

$$y'_r = tan(\phi)(x'_r - x_t) + y_t. \tag{2.4}$$

Substituting Eq.(2.4) into Eq.(2.2), we have:

$$\begin{aligned} tan(\phi)(\frac{(B^2 - A^2)x_r - 2ABy_r - 2AC}{(A^2 + B^2)} - x_t) \\ = (\frac{(A^2 - B^2)x_r - 2ABx_r - 2BC}{(A^2 + B^2)} - y_t). \end{aligned} \tag{2.5}$$

Eq.(2.5) provides an analytical condition that must be met by the locations of the AP and the client, the AoD, and the location and orientation of the reflector. Because the locations of the AP and the client can be measured, and (x_r, y_r) and (x_t, y_t) are considered known. Meanwhile, as will be clarified in Section 2.2.2, the AoD ϕ of the specular reflection path can be estimated through a sequence of RSS measurements accompanying the steering of the wide-beam transmit antenna. So ϕ is also considered as a known value. Therefore, Eq.(5) only depends on variables A, B, and C. As a result, we need at least three independent sets of $< \phi, (x_t, y_t) >$ to uniquely determine the location and orientation of the reflector.

To the best of our knowledge, this is the first framework in the literature that supports the computation of the reflector's location and orientation without knowing AoA, neither any phase information of the received signal. Our framework does require some knowledge on AoD, but the acquisition of this information does not rely on high-precision narrow-beam transmit antennas or any phase information of the signal, as will be described in Section 2.2.2. This is in sharp contrast to existing methods that rely on high precision phased array antennas and the phase information on both sides of the link to obtain accurate AoA and AoD in order to localize the reflectors [125].

2.2.2 Locating Dominant Reflectors

Design Philosophy

Dominant reflectors are reflectors that create strong NLoS paths for most of the indoor positions. Although the number of reflectors a mmWave radio can see vary with the device's location, there are only a few dominant reflectors in a realistic environment because of the sparsity of the mmWave channel, and most of dominant reflectors are static metallic surfaces that have low reflection losses.

We already show that a reflector can be uniquely determined by three independent sets of $<\phi, (x_t, y_t)>$, so deriving the AoD of the specular ray associated with the dominant reflectors (i.e., ϕ_i, for the i-th dominant reflector) is a crucial step in determining the dominant reflectors' location and orientation. We propose to determine ϕ_i's by measuring the RSS at the receiver as a function of the beam direction of the transmit antenna. In particular, let θ_{mid} denote the center angle of the main lobe of the transmit antenna. Given an omni-directional receive antenna, the RSS at the receiver is a function of θ_{mid}, denote as $P_r(\theta_{mid})$ and is given by

$$P_r(\theta_{mid}) = P_t \left\| \sum_{i=0}^{N} D(\theta_{mid}|\phi_i) g_i e^{j\delta_i} \right\|^2, \tag{2.6}$$

where P_t is the transmit power, and we have assumed that the mmWave channel has $N+1$ paths (so there are N dominant reflectors), and the i-th path has a path loss, AoD, and path phase shift of g_i, ϕ_i, and δ_i, respectively. e is the natural logarithm, and j is the imaginary unit. $D(\phi_i|\theta_{mid})$ is the transmit antenna gain at the AoD ϕ_i, given that the main beam of the antenna is pointing at θ_{mid}. Without loss of generality, we assume $D(\phi_i|\theta_{mid}) = 1$ if $|\phi_i - \theta_{mid}| \leq \Theta_{beam}$ and $D(\phi_i|\theta_{mid}) = 0$ otherwise, where $2\Theta_{beam}$ is the beam width of the main lobe of the transmit antenna.

Clearly, according to Eq.(2.6), ϕ_i's can be easily resolved by steering a narrow-beam transmit antenna with small Θ_{beam} to perform a 360°-scanning of the space. However, note that in our problem we have a low-end phased array antenna with a wide beam-width. The reduced angle resolution (i.e., large Θ_{beam}) makes it challenging to resolve the ϕ_i's.

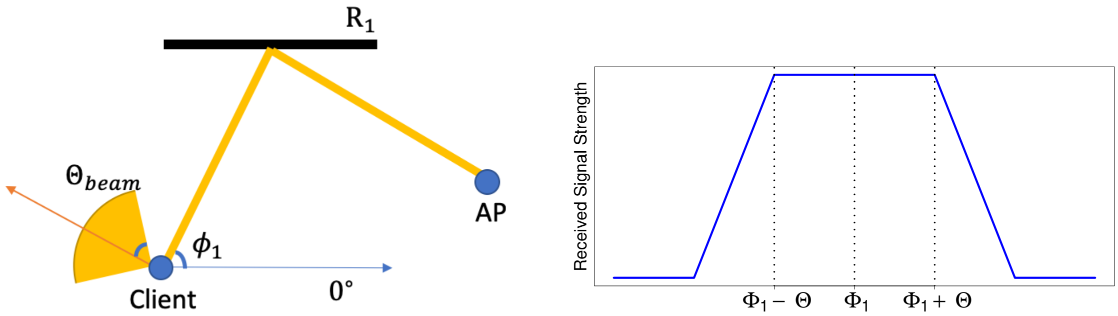
One way to resolve ϕ_i's is through reverse engineering, i.e., by considering (ϕ_i, g_i, δ_i)'s as variables, and then resolve them by solving a set of nonlinear equations defined by Eq.(2.6), where P_r is measured at multiple θ_{mid}'s. In this case, the optimal N^o and $(\phi_i^o, g_i^o, \delta_i^o)$, $0 \le i \le N^o$, can be simply calculated as the optimal solution to the following minimum mean square error (MMSE) problem:

$$\text{minimize}\frac{1}{2\pi}\int_0^{2\pi} |P_r(\theta_{mid}) - P_r^{\mathrm{m}}(\theta_{mid})|^2 d\theta_{mid}, \tag{2.7}$$

where $P_r(\theta_{mid})$ is defined in Eq.(2.6) and $P_r^{\mathrm{m}}(\theta_{mid})$ is the RSS measurements.

Although the above method fits in our problem and provides promising solutions, the computation complexity is high, which makes it unsuitable for COTS devices. In addition, it provides more than what the problem needs. Notice that, for the purpose of finding a NLoS backup path, what we are interested in is only the top few (say K, where K is a small integer) dominant reflectors that provide the strongest NLoS paths. Therefore, resolving the whole set of NLoS paths, as in the MMSE method, is unnecessary. Based on this observation, we propose to only consider the K particular θ_{mid}'s that correspond to the top-K RSS peaks in the function $P_r^{\mathrm{m}}(\theta_{mid})$, and use these K θ_{mid}'s to estimate the AoDs of the top-K dominant reflectors in the environment.

This idea can be better illustrated as follows. We begin with the simplest scenario: suppose we only have one reflector, say R_1, in the environment. The specular reflection direction to R_1 is denoted as ϕ_1, as shown in Figure 2.2(a). As the transmit antenna scans from $\theta_{mid} = 0$ to 360°, the RSS measured at the AP should present a trapezoid shape as shown in Figure 2.2(b), where the high RSS level in the range of $\phi_1 - \Theta_{beam} \le \theta_{mid} \le \phi_1 + \Theta_{beam}$ is due to the fact that the AoD of the specular ray ϕ_1 is within the main lobe when the transmit antenna is scanning in this range. So, for the single-reflector case, the peak of the RSS, defined as the center of the RSS high level, corresponds to the AoD of the specular ray ϕ_1, as shown in Figure 2.2(b).



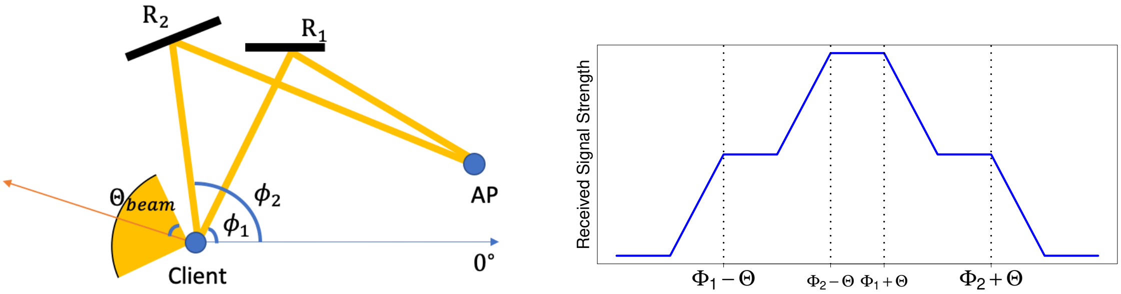
(a) Single reflector specular reflection case.

(b) RSS pattern of single reflector.

Figure 2.2: Single reflector scenario.

Now let us consider a more complicated scenario where there are two reflectors in the environment, as shown in Figure 2.3(a), where AoDs of the two specular rays are denoted by ϕ_1 and ϕ_2, respectively. Without loss of generality, we assume that $\phi_1 < \phi_2$. Note that if ϕ_1 and ϕ_2 are separated far apart such that $\phi_2 - \phi_1 > 2\Theta_{beam}$, then each of the two AoDs can be resolved separately as two independent single-reflector cases (i.e., the two peaks of the RSS curve correspond to ϕ_1 and ϕ_2 respectively). Now let us consider the case that ϕ_1 and ϕ_2 are close enough such that $\phi_2 - \phi_1 \leq 2\Theta_{beam}$. In this case, as the transmit antenna scans from $\theta_{mid} = 0$ to $360°$, the RSS measured at the AP should present the pattern shown in Figure 2.3(b), where the different levels of RSS are due to the fact that different combinations of the specular rays are in the main lobe as the transmit antenna scans. Clearly, in this case, the peak of the measured RSS corresponds to the center angle of ϕ_1 and ϕ_2, i.e., $\frac{1}{2}(\phi_1 + \phi_2)$. Physically, this means that because ϕ_1 and ϕ_2 are close to each other, the reflectors R_1 and R_2 are actually considered as a cluster of reflectors, and the peak of the measured RSS simply corresponds to the center AoD of this cluster.

To simplify the presentation, we have assumed that the beam pattern of COTS devices has a regular pie shape, but the insight here is general enough to accommodate any regular/irregular beam pattern function. This is because changing the beam pattern only changes the coefficients of antenna gain, but does not change the structure/nature of the problem.



(a) Multiple reflectors specular reflection case.

(b) RSS pattern of multiple reflectors.

Figure 2.3: Multiple reflectors scenario.

Note while Figures 2.2(b) and 2.3(b) are showing symmetric RSS peaks in concept, the measured RSS peaks in reality are rarely symmetric. As shown in Figure 2.4, the measured peaks are usually skewed, mainly due to the large number of scatters (i.e., small/minor reflectors) surrounding the major reflector and the heterogeneous reflection efficiency of the reflectors. Therefore, while a measured RSS peak provides a rough range of directions where one or several major reflectors could reside in, simply interpreting the mid-point of the RSS peak as the AoD of one major reflector is inaccurate and misleading, for the peak could be generated by multiple close-by major reflectors.

How to accurately resolve the major reflector(s) from the skewed RSS measurements constitutes a challenge.

Two-step Fine-Grained Multi-path Channel Resolving

We solve this challenge by a novel two-step multi-path channel resolving algorithm. The main idea is to first identify clusters of major reflectors by evaluating the peaks of measured RSSI, and then apply reverse engineering within each RSSI peak to resolve the optimal in-cluster reflector setting that offers the best match with the measured RSSI in that peak.

The detail of our algorithm is described below:

(1) Low-resolution out-lobe resolving: Since the goal of the algorithm is to resolve for k dominant reflectors, we first need to decide the ranges of θ_{mid} where these K dominant reflectors reside in. This is done by picking the K highest peaks in the curve $P_r^{\mathrm{m}}(\theta_{mid})$, as illustrated in Figure 2.4. Notably, the curve $P_r^{\mathrm{m}}(\theta_{mid})$ is obtained by using the fine-grained spatial scanning process. Specifically, the process steers the transmit antenna to point θ_{mid} to a sequence of N angles respectively, denoted by $\theta_1, \ldots, \theta_N$, which are evenly distributed between 0° and 360° with a step size $\omega = 360°/N$, i.e., $\theta_i = i\omega$ for $1 \le i \le N$. At each θ_i, let the measured RSSI at the receiver be $P_r^{\mathrm{m}}(\theta_i)$. Recall that we are only interested in NLoS paths, so those θ_i's that belong to the LoS should be excluded in our subsequent range selection. Let the LoS direction be θ_{LoS}. Since the half beam width of the transmit antenna is Θ_{beam}, a θ_i is considered belonging to the LoS if $\theta_{LoS} - \Theta_{beam} \le \theta_i \le \theta_{LoS} + \Theta_{beam}$. For example, in Figure 4, the LoS direction θ_{LoS} is 330°, and the half beam width of antenna is 30°. So RSS measurements between 300° to 360° are considered belonging to the LoS range and are ignored during subsequent NLoS path ranges selection procedure.

Our range decision is iterative: we decide one non-overlapping range for θ_{mid} in each iteration, and our decision concludes after K iterations, resulting in K non-overlapping ranges. In particular, in the k-th iteration we decide a range $\left[\theta_{low}^{(k)}, \theta_{high}^{(k)}\right]$ that includes the following angles:

$$\theta_k^o = \arg\max\left\{ P_r^{\mathrm{m}}(\theta_i) \middle| \theta_i \notin \bigcup_{j=1}^{k-1} \left[\theta_{low}^{(j)}, \theta_{high}^{(j)}\right] \right\}. \tag{2.8}$$

And all adjacent θ_i's that are within the half beam of θ_k^o but are not included in any of the ranges decided in previous iterations, i.e., $\left[\theta_{low}^{(k)}, \theta_{high}^{(k)}\right]$ includes the following θ_i's:

$$\{\theta_i | \theta_k^o - \Theta_{beam} \le \theta_i \le \theta_k^o + \Theta_{beam}\} \text{ and } \theta_i \notin \bigcup_{j=1}^{k-1} \left[\theta_{low}^{(j)}, \theta_{high}^{(j)}\right], \tag{2.9}$$

where $\theta_{low}^{(k)}$ and $\theta_{high}^{(k)}$ are the smallest and the largest elements in the above set, respectively. A θ_i is excluded from the subsequent iterations if it has been included in one of the ranges decided in previous iterations. Note that by picking the above K ranges, we do not mean that the AoDs of the top-K dominant reflectors should reside in each of these K ranges (i.e., one in each range). Instead, the AoDs of the top-K dominant reflectors should reside in the union of these K ranges.

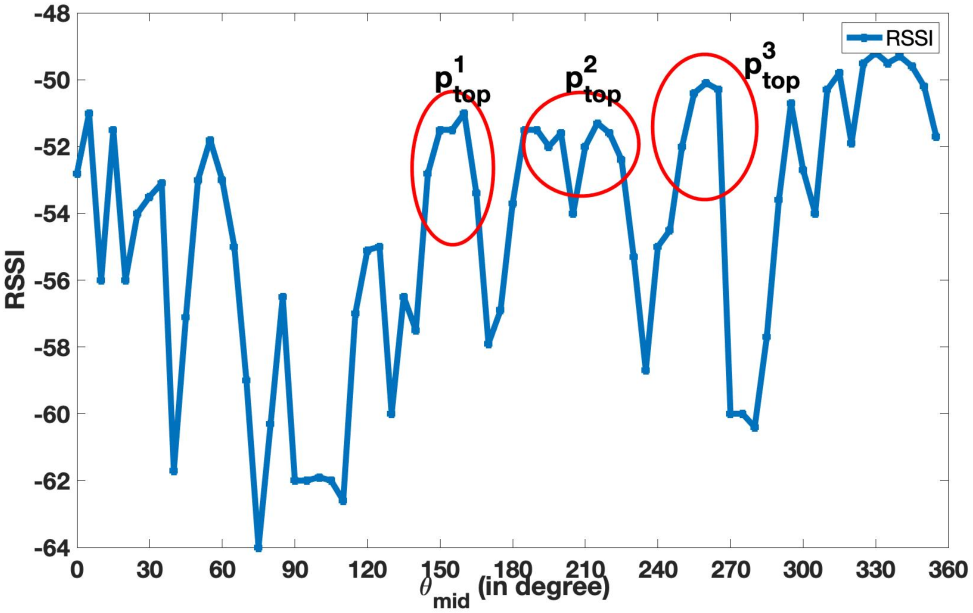


Figure 2.4: Out-lobe resolving based on measured RSSI (scanning step $\omega = 5°$).

(2) High-resolution in-lobe resolving: Now that we have the top-K ranges of θ_{mid} as $\left[\theta^{(k)}_{low}, \theta^{(k)}_{high}\right]$, $1 \le k \le K$, we need to resolve the reflectors whose AoDs are within these ranges. Without loss of generality, let us consider the k-th range. Suppose there are N reflectors in this range, and accordingly there are N NLoS paths (this is because each reflector generates exactly one NLoS path via its specular reflection) in the cluster defined by this range. And each path can be characterized by its propagation (and reflection) path loss, AoD, and path phase shift, denoted by g_i, ϕ_i, and δ_i respectively, for the i-th NLoS path, where $1 \le i \le N$. Considering the beam width of the transmit antenna, notice that for the k-th range, we have $\theta^{(k)}_{low} - \Theta_{beam} \le \phi_i \le \theta^{(k)}_{high} + \Theta_{beam}$ for all $1 \le i \le N$. Given that the main beam of the transmit antenna is pointing at θ_{mid}, the

transmit antenna gain for the i-th NLoS path is given by

$$D(\theta_{mid}|\phi_i) = \begin{cases} 1 & \phi_i - \Theta_{beam} \le \theta_{mid} \le \phi_i + \Theta_{beam} \\ 0 & \text{otherwise.} \end{cases} \tag{2.10}$$

Therefore, when the transmit antenna is scanning within the k-th range, the RSS at the receiver can be analytically described as

$$P_r(\theta_{mid}) = P_t \left\| \sum_{i=1}^{N} D(\theta_{mid}|\phi_i) g_i e^{j\delta_i} \right\|^2, \theta_{low}^{(k)} \le \theta_{mid} \le \theta_{high}^{(k)}. \tag{2.11}$$

To resolve for g_i's, ϕ_i's, and δ_i's, we use reverse engineering: we would like to decide the optimal g_i^o's, ϕ_i^o's, and δ_i^o's that would make $P_r(\theta_{mid})$ the closest match, in the mean square error (MSE) sense, with the measured RSS $P_r^{\mathrm{m}}(\theta_{mid})$ at the discrete angles $\theta_i \in \left[\theta_{low}^{(k)}, \theta_{high}^{(k)}\right]$, i.e.,

$$\underset{g_i^o, \phi_i^o, \delta_i^o, i \in N}{\text{minimize}} \sum_{\theta_i = \theta_{low}^{(k)}}^{\theta_{high}^{(k)}} \left| P_r(\theta_i) - P_r^{\mathrm{m}}(\theta_i) \right|^2. \tag{2.12}$$

In our evaluation part, the signal gain g is from -50 dB to -70 dB with step size 0.01 dB. The angle direction ϕ is from $[\theta_{low}^{(k)}, \theta_{high}^{(k)}]$ with a step size 0.1°. And phase δ is from 0° to 360° with a step size 0.1°

Compared with the global range (from 0 to 360°) reverse engineering in Eq.(2.7), the scale of the above local range optimization, in terms of the number of variables to be optimized, is much smaller therefore the optimization can be achieved much faster. This is for the ground truth number of reflectors in $\left[\theta_{low}^{(k)}, \theta_{high}^{(k)}\right]$ should be much smaller than that in [0°,360°], so a small N in Eq.(2.12) is usually sufficient to obtain small MSE in the objective function. To verify the point, we solve the optimization problem in Eq.(2.12) for the top-3 ranges highlighted in Figure 2.4 under various N's. As a representative outcome, Figure 2.5 plots the normalized MSE for the optimization in range 1 as a function of N, where the normalization is w.r.t. the square of the maximum measured RSS in range 1. It can be observed that the normalized MSE in this case goes down quickly as N increases, and remains almost flat after $N \geq 3$, implying that $N = 3$ is an acceptable estimation for the ground truth number of major reflectors in this range. Note that our resolved top-N dominant reflectors are in fact dominant reflector clusters, each of which represents an aggregation of multiple closed-by reflectors and do not have one-to-one correspondence with actual physical reflectors. Both specular and diffusive (or scattering) reflection effects have already been aggregated into these reflector clusters.

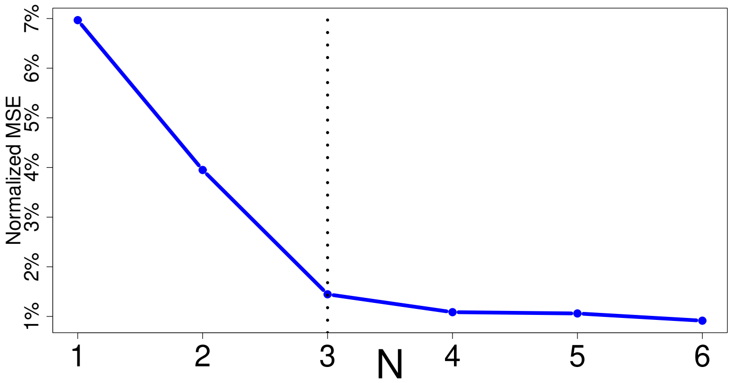


Figure 2.5: Normalized MSE v.s. N.

The AoDs of the top-K dominant reflectors are decided by pooling the reflectors resolved over all K ranges together, and picking the top-K among them with the smallest propagation losses (i.e., highest g_i's). For example, Table 2.1 lists the optimized g_i's and ϕ_i's for the top-3 ranges in Figure 2.4 for $N = 3$. The optimized δ_i's are not shown in the table due to space limit. The top-3 dominant reflectors are decided as: Reflector 1 in range 1 (path loss = -55.83 dB, AoD = 172.5°), Reflector 1 in range 3 (path loss = -55.96 dB, AoD = 275.8°), and Reflector 3 in range 3 (path loss = -56.19 dB, AoD = 235.2°). Among them, it can be observed that the AoDs of the last two dominant reflectors are separated less than the beam width of the transmit antenna. These results verify that the proposed two-step algorithm can achieve fine-grained multi-path channel resolving by only using coarse-grained wide-beam antennas. We then use a real testbed to evaluate the accuracy of the resolved paths and their effects in maintaining mmWave connections in Section 2.3.

Table 2.1: Multi-path channel resolving result.

Par	g_1	ϕ_1	g_2	ϕ_2	g_3	ϕ_3
$P_{top}^{(1)}$	-55.83	172.5°	-57.60	134.4°	-60.37	152.3°
$P_{top}^{(2)}$	-61.47	210.9°	-56.84	196.8°	-56.54	243.2°
$P_{top}^{(3)}$	-55.96	275.8°	-58.15	257.8°	-56.19	235.2°

Localization for the client and the AP

Aiming for real-time beam switching under the mobile scenario, the accurate location information of the client and the AP is essential in the proposed framework. However, the GPS localization is not suitable for the indoor usage scenario. Moreover, the resolution of the GPS system is low, typically in several meters, which does not meet the precision requirement of our problem. To obtain high precision localization information, we use the HTC VIVE VR system to track the real-time location of the client. Note that the VR system can be replaced by any indoor mmWave device localization method, such as the mmWave AP triangulation [15] or AP device localization mentioned in [16].

Matching the AoDs

To fully determine a dominant reflector using our framework, we need at least three independent sets of $<\phi,(x_t,y_t)>$. Therefore, in an environment with multiple dominant reflectors, how to identify those AoDs that are measured at different client locations but are associated with the same dominant reflector raises another challenge.

To address this issue, we exploit the sparsity of the mmWave channel, which dictates that the mmWave channels at two nearby locations are caused by the same set of dominant reflectors. So their spatial channel profiles (SCPs) are tightly correlated in the sense that their AoD realizations associated with the same reflector are also close-by to each other [152]. To utilize this property, we propose the following AoD measurement and matching process. We fix the AP's location, and measure the RSSI at the AP as a function of θ_{mid} when the client is positioned at several nearby locations, respectively. Denote this set of nearby locations as set S. We then identify the top-K AoDs at the first client location based on the RSSI-θ_{mid} measurement made at that location. Let $\phi_1^{(1)}, \phi_2^{(1)}, \ldots, \phi_K^{(1)}$ denote these top-K AoDs, associated with K strong dominant reflectors, say $R_1, R_2, \ldots, R_K$, respectively. Let $\phi_i^{(j)}$ denote the AoD realization associated with reflector R_i at a different client location $j \in S$. To decide $\phi_i^{(j)}$, one simply finds the peak RSSI measured at client location j that is nearest to $\phi_i^{(1)}$. The θ_{mid} corresponding to this peak RSSI is $\phi_i^{(j)}$.

The above process is illustrated in Figure 2.6, where the RSSI-θ_{mid} measurements have been made at two close-by locations, represented by the blue curve and the red curve, respectively. To decide the AoD realizations of three strong dominant reflectors at these two locations, we first pick the top-3 AoDs on the blue curve, and label them as NLoS1 through NLoS3 in blue. Then, the AoD realization of NLoS1 on the red curve is simply the red peak nearest to the blue peak of NLoS1. The AoD realization of NLoS2 and NLoS3 on the red curve can be decided in a similar way.

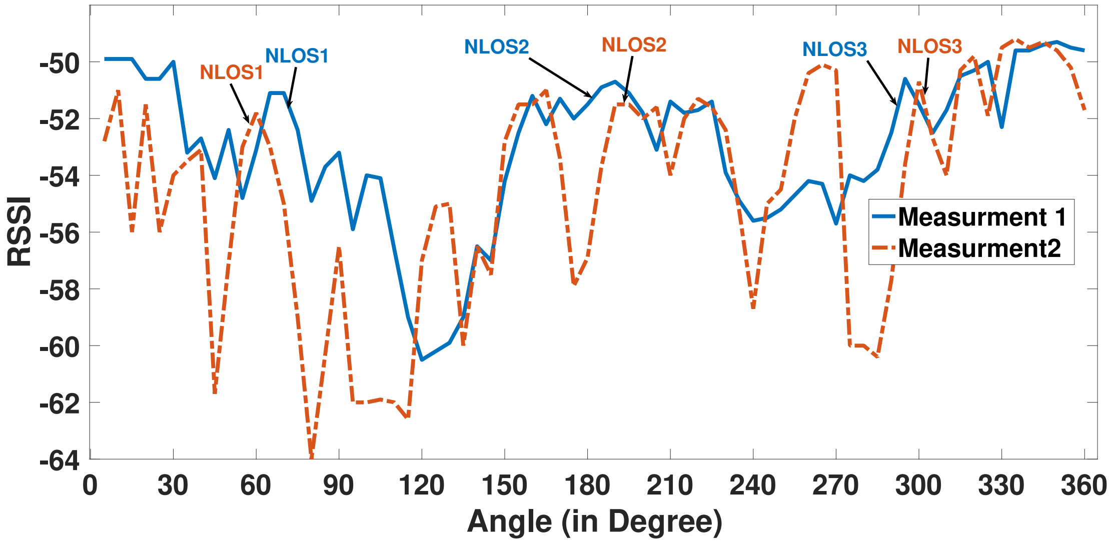


Figure 2.6: NLoS path matching between two different client locations.

Based on our methods, we are able to obtain a sufficient number of $< \phi_i, (x_t, y_t) >$'s for each of the K strongest dominant reflectors identified in the AoD matching process, based on which the proposed framework in Eq.(2.5) can be applied to calculate the location and orientation for each of these strong dominant reflectors.

2.2.3 Model Driven RSSI Estimation

In this section, we present a received signal strength estimation model to predict the link performance of the mmWave NLoS backup paths in real-time. When the LoS path is blocked, the aforementioned dominant reflector map provides K NLoS candidate paths for the transmit antenna to steer to. Selecting the best one among them will be critical to retain a comparable performance to that of the LoS path. However, the naive sequentially trial-based method will result in significant delays in beam switching, therefore undermining the stability of the connection. To achieve a better seamless beam switching performance, we use a model-driven approach on the transmitter side to predict the quality of each NLoS backup paths, so the transmit antenna can choose the best path directly.

Among the K NLoS candidate paths provided by the dominant reflector map, let us consider the one associated with the i-th dominant reflector. If the transmit antenna beam is switched to this path, then the received power P_r at the receiver is given by:

$$P_r = \frac{P_t G_t G_r}{L_f R_l}. \tag{2.13}$$

Here P_t is the total transmission power, G_t and G_r are the transmitter and receiver's antenna gain for the path; L_f and R_l are the free space loss and the reflection loss of the path, respectively. The total transmission power is a constant for a COTS device. The antenna gains are also fixed, because the receive antenna is omni-directional, and the path is in the main lobe of the transmit antenna (so $G_t = 1$). Next, we will explain how to calculate L_f and R_l.

1)Free space Loss: According to the Friis's law, the free space loss L_f is:

$$L_f(d) = \left(\frac{\lambda}{4\pi d}\right)^2, \tag{2.14}$$

where λ is the wavelength of the carrier frequency and d is the length of the NLoS path.

Our model considers the specular reflection, so the transmission path length d between the client and the AP is equal to the distance between the client and the virtual AP, which can be calculated using:

$$d = \sqrt{(x_t - x'_r)^2 + (y_t - y'_r)^2}, \tag{2.15}$$

where the virtual AP's location (x'_r, y'_r) is defined w.r.t. the i-th dominant reflector and is given by Eq.(2.2).

2)Reflection Loss: The reflection loss depends on the material of the reflector, and can be characterized using the Fresnel reflection coefficient(Γ) [135, 79]. There are two Fresnel equations for two different polarization cases. And we use a simplified version of the horizontally polarized model, under which the Fresnel coefficient is given by:

$$\Gamma_H = \frac{\sin\psi - \sqrt{\varepsilon_r - \cos^2\psi}}{\sin\psi + \sqrt{\varepsilon_r - \cos^2\psi}}, \tag{2.16}$$

where ε_r is the relative permittivity of the reflective material, ψ is the grazing angle. Notably, the ε_r remains as a constant and does not depend on the carrier frequency [135]. The grazing angle is the angle between the incident ray and the reflecting surface, since we have already modeled the reflector as a line segmentation $Ax + By + C = 0$, the grazing angle can be calculated by:

$$\psi = \arctan\left(\left|\frac{k - \tan(\phi)}{1 + k \cdot \tan(\phi)}\right|\right), \tag{2.17}$$

where $k = -A/B$ is the slope of the reflector surface, ϕ is the specular ray AoD, given by:

$$\tan(\phi) = \frac{(B^2 - A^2)x_r - 2ABy_r - 2AC - (A^2 + B^2)x_t}{(A^2 - B^2)x_r - 2ABx_r - 2BC - (A^2 + B^2)y_t}, \tag{2.18}$$

where (x_t, y_t) is the real-time location of the client, and (x_r, y_r) is the location of the AP. The reflection loss can then be represented as [135]:

$$R_l = \left|\frac{1}{\Gamma_H}\right|^2 = \left|\frac{\sin\psi + \sqrt{\varepsilon_r - \cos^2\psi}}{\sin\psi - \sqrt{\varepsilon_r - \cos^2\psi}}\right|^2. \tag{2.19}$$

Note that the above theoretical model only describes the general pattern followed by the RSS on this path. To make this model fit in our particular operation environment, we use a regression for model parameter fitting. In particular, we consider the following decibel form of the RSS for the NLoS path associated with the i-th dominant reflector:

$$P_r^{(i)}(d,\psi) = \kappa_i - 20\eta_i \log_{10}(d) - 20\gamma_i \log_{10}(\left|\frac{\sin\psi + \sqrt{\zeta_i - \cos^2\psi}}{\sin\psi - \sqrt{\zeta_i - \cos^2\psi}}\right|). \tag{2.20}$$

We estimate the parameters κ_i, η_i, γ_i and ζ_i of the model offline using regression when the transmit beam is switched to the i-th dominant reflector with known AP and client locations (so d and ψ can be calculated). The data is gathered using the empirical RSSI readings generated by the device's firmware (unit in decibel). The regression functions $P_r^{(i)}(d,\psi)$, $i = 1,\dots,K$, are then used online to model the received signal strength for the K NLoS paths offered by the dominant reflector map at new client locations. When the LoS is being blocked, our system can compute the RSS of different NLoS paths and directly switch the transmit beam to the best one among them.

2.2.4 Overhead/Cost Analysis

In this section, we provide a overhead/cost analysis for our proposed method. The total overhead of the system consists of the following two components:

1) Offline calibration phase: As we have discussed in Section 2.1, to fully determine the location and orientation of environmental reflectors, we need at least three independent sets of $< \phi, (x_t, y_t) >$ for algorithm processing. During the calibration phase, we fixed the AP location and move the client to three different locations to perform a 360° fine-grained spatial scanning. The scanning step is ω. For each step, we record the RSS from the AP side, so in total we collect $3 \cdot \frac{360^\circ}{\omega}$ RSS measurements. In our test, each RSS measurement is represented by a 4 byte float number, and ω is set to 5°, so the total information needed is 864 byte.

Low-resolution out-lobe resolving: In this part, for each ($360^\circ/\omega$) number of RSS measurements, our algorithm determines K RSS ranges for high-resolution in-lobe resolving. An iterative approach is used to determine these K ranges, so the time complexity is $\mathcal{O}(K)$. Since the number of dominant reflectors in an indoor environment is usually limited, a small K is sufficient to provide stable NLoS paths for robust communication. In our experiments, we set K to 3 and the actual time spent by low-resolution out-lobe resolving process is negligible.

High-resolution in-lobe resolving: For each RSS range extracted from the low-resolution out-lobe resolving phase, it contains ($2\Theta_{beam}/\omega$) RSS measurements, where Θ_{beam} denotes the half beam width of the main lobe. And we need to solve N sets of g_i^o's, ϕ_i^o's, and δ_i^o's for each selected range. In this step, we use an optimization tool to solve the proposed MMSE problem. In our test, we use the MatLab fmincon function with default interior-point method. We have validated that a small N (3 to 4) is usually sufficient to obtain a sufficient small error in the objective function. With an Apple iMac with 3.4 GHz Quad-Core Intel Core i5 CPU, the optimization can be done within 10 seconds for each range when N is set to 3.

RSSI Regression: For each of the K dominant reflectors, we estimate the parameters κ_i, η_i, γ_i and $\zeta_i, i \in K$ of the model offline using regression when the transmit beam is switched to the i-th dominant reflector with known AP and client locations. The data is gathered using the empirical RSSI readings generated by the device's firmware. In our evaluation part, 10 to 15 different client locations are enough for the regression model to reach a high accuracy. The regression is performed by using the MatLab curving fitting with the nonlinear least-squares fitting procedure. With an Apple iMac with 3.4 GHz Quad-Core Intel Core i5 CPU, the regression can be done within 1 seconds for each range when K is set to 3.

2) Online operation phase: In the online operation phase, our algorithm uses specular reflection model to calculate the NLoS paths for current locations. The time complexity to calculate K paths is $\mathcal{O}(K)$.

The complexity analyses are summarized in Table 2.2

Phase	Operation	Complexity
Offline sensing	Fine-grained 360° scanning	$\mathcal{O}(\frac{360°}{\omega})$ for 3 times
	Low-resolution out-lobe resolving	$\mathcal{O}(K)$ for each scanning
	High-resolution in-lobe resolving	Fmincon with interior-point method
	Path loss regression	Nonlinear least-squares fitting
Online operation	NLoS paths calculation and switching	$\mathcal{O}(K)$ to calculate K paths

Table 2.2: Complexity of each step.

2.2.5 Limitation of the Method and Extension to Larger Space

A dominant reflector is defined naturally in a local sense, because the strength of its reflected signal will go down with the distance between the reflector and the user increases. So a reflector being dominant when the user was close may not remain dominant when the user moves far away. Because of this, our proposed method can only be directly applied to a small-to-moderate room scenario, where at least one reflector defined in the top-K dominant reflector map remains to be dominant at any location of the room. In reality, this may correspond to practical application scenarios such as wireless VR/AR gaming, in which a player does not move too far but may frequently turn their body, or multiple players interact with each other in one game, so the LoS may be frequently blocked by the player's or the other player's body. In the multi-user scenario, other users can not only block the LoS but also the estimated NLoS path with each other. This issue can be trivially solved by simply turning to the second or next optimal estimated NLoS direction, as the proposed method is actually able to compute the top-K optimal NLoS directions based on the top-K dominant reflector map.

The proposed method can be trivially extended to a larger-space scenario (e.g., a ballroom or an auditorium) by partitioning the space of the room into smaller areas, and then applying the proposed method to each area to construct a individual dominant reflector map. The maps of individual areas are then aggregated and fused into a master map that describes the location and reflection efficiency of all dominant reflectors in the room. This master map is distributed to each user in the operational phase for their online strong NLoS path prediction and beam switching. Such an extension is out of the scope of this work and will be pursued in our future research.

Also, we want to point out that our proposed mechanism is suitable for most of the indoor household application scenarios, where the room layout (i.e., the environment) is static or quasi-static but the users could be moving. For instance, when a user is playing an electronic VR game or using wireless cell phone in indoor environment, the direction of the strongest NLoS path keeps changing due to the user's movement but the environment is static (i.e., the locations of the dominant reflectors do not change or remain static for a long period of time). This static environment assumption should be true in most of the indoor application scenarios, because those major reflectors are usually large-size furniture, walls, and windows of metallic surfaces, which are hardly mobile.

2.3 Testbed and Implementation

We implemented our prototype system based on COTS components. The system architecture and prototype are shown in Figure 2.7. Our system consists of four main parts: 1) two MikroTik WAP 60G mmWave radios [72] are used for mmWave communication; 2) a robotic arm is used for 360° mechanical steering of the transmit beam. This robotic arm is needed only during the one-time offline site survey (installation) phase to perform stepped scanning. In the online operational phase, this robotic arm is optional. If it is not available, the COTS device can simply steers the transmit beam to the particular beam mode (e.g., 64 beam modes are provided by MikroTik WAP 60G) that is the closest to and covers the desired NLoS direction, achieving an approximation to the original mechanism presented in Section 2.2 when the arm is available to steer the beam to the exact desired direction. 3) A VR system is used to provide accurate position information for the AP and the client. Note that this VR system is used only for convenience/ease of our implementation. It can be replaced by any state-of-the-art indoor WiFi-based localization algorithm that does not require any additional infrastructure [15]; 4) a PC host is used to control the beam switching procedure according to our proposed method.

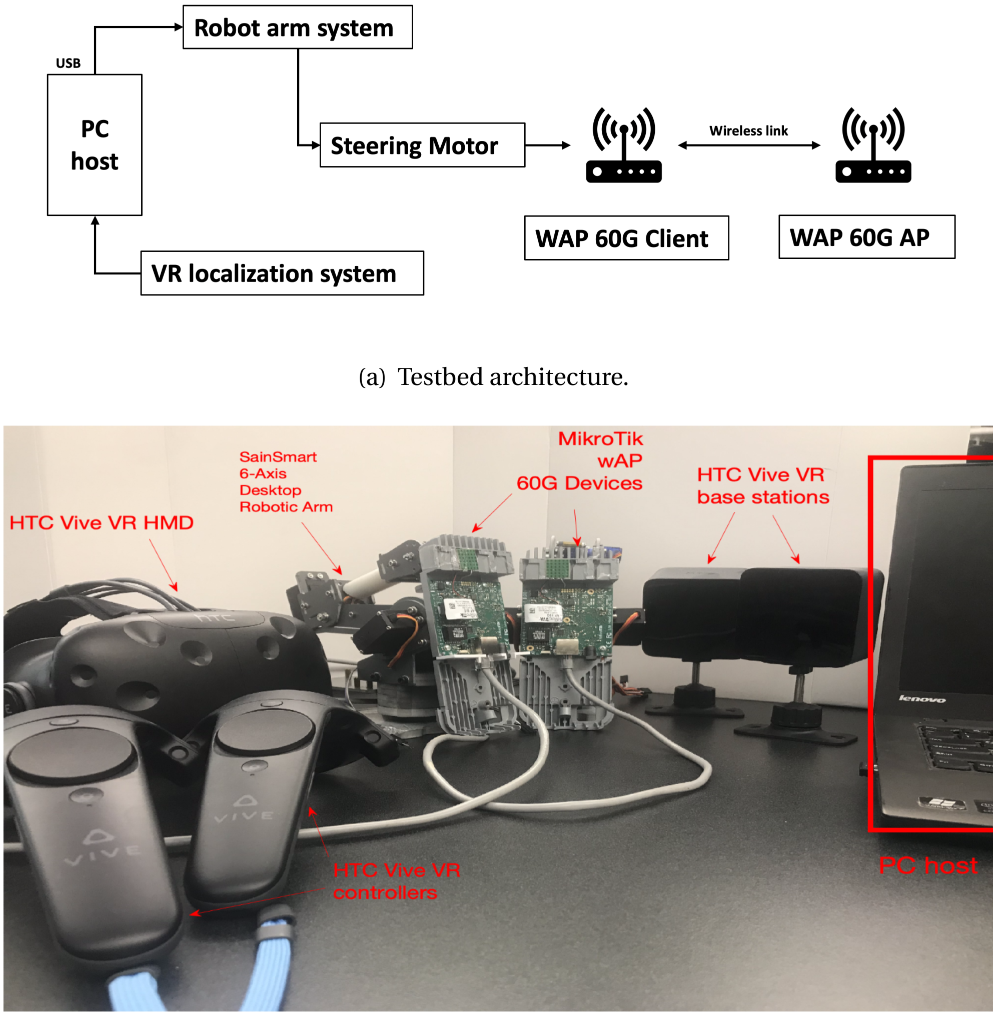
Our testbed is only intended to serve as a prototype to demonstrate the feasibility of the proposed method. The robotic arm is not an indispensable part to perform our algorithm. In particular, in the online NLoS prediction phase, instead of using the robotic arm for fine-grained mechanical steering of the antenna beam, the COTS firmware we are using allows a coarse-grained electronic steering of the beam by selecting an appropriate beam pattern that covers the desired direction to which the beam should be turned. In the offline measurement phase, using the robotic arm to do the automatic space scanning can significantly expedite the measurement process. However, in case that the robotic arm is not available, the above scanning can also be done manually.

When the LoS is blocked, in order to allow the user to turn its beam to the estimated NLoS direction, say α, we do need to know the direction/orientation of the beam right before the LoS blockage, i.e., the direction of the LoS path, denoted as β. So after the blockage the beam needs to turn $\alpha - \beta$ degrees from its current orientation in order to point to the estimated NLoS direction. Given the availability of the locations of the AP (denoted as (x_0, y_0)) and the user(denoted as (x_1, y_1)), the orientation of the user's beam before the LoS blockage can be calculated as $\beta = \arctan\frac{x_1 - x_0}{y_1 - y_0}$ (without loss of generality, here we are assuming that the direction of the Y-axis is the 0°).

2.4 Evaluation

2.4.1 Test Setting

Test environment: The performance test is conducted in an indoor lab with a 4.9 m × 4.8 m layout. We set up a pair of AP and client. The AP is placed at a fixed location with coordinate $(-1.44, 0.05)$, and its receive antenna is omni-directional. For the client, its location is random



(a) Testbed architecture.

(b) Testbed overview.

Figure 2.7: System prototype.

picked to cover the whole test area and a VR HMD is bounded with the client to track its location(VR base stations are mounted on wall for HMD position tracking). The client uses a beamforming mode that forms a 60° beam for transmission. We use the iPerf as the traffic source to drive the mmWave link, and the RSSI measurement is extracted from the integrated RouterOS operating system.

Reflectors reconstruction: To reconstruct dominant reflectors, we need multiple correlated tests to fully locate them. So we fix the AP to an anchor location and move the client to three different locations to perform a 360° scanning with a step angle 5°. The measured RSSI patterns are fed as input for the two-step AoD derivation method to extract the AoDs of dominant reflectors. Then we use these AoDs and location information as the input for specular reflection model to reconstruct the dominant reflectors' geometry.

The NLoS path directions are calculated using dominant reflectors' geometry and real-time client locations. To compare the performance of different NLoS paths, we conduct a link performance test under 30 different client locations and use the RSSI as the performance metric. In addition, a transport layer performance test under TCP/UDP, containing both static and mobile scenarios, is also conducted to show the performance of our system under different conditions. In the static test, both the AP and client's locations are fixed, whereas in mobile scenario, only the AP's location is fixed and the client moves across the room with 0.5 m/s velocity. The real time TCP/UDP throughput is used as the performance metric.

Performance benchmarks: For the purpose of performance comparison, we conduct two types of performance benchmarks. The first is the performance using LoS link for communication, which is the upper bound performance of the system. The second is the performance using auto beam steering method [71], which is the default beam steering method for the MikroTik WAP 60G devices. The method follows the IEEE 802.11ad standard and can automatically change the beam among the 64 predefined beam patterns to maximize the throughput.

2.4.2 Experimental Results

Dominant Reflector Reconstruction

In our test, we successfully reconstruct two dominant reflectors, namely the left side reflector 1 and bottom side reflector 2, as specified in Figure 2.8. The reflection coefficients obtained by our regression model is 1 and 0.224 for reflector 1 and reflector 2. The rectangle denotes the lab's layout. The blue and red lines represent the two reconstructed dominant reflectors.

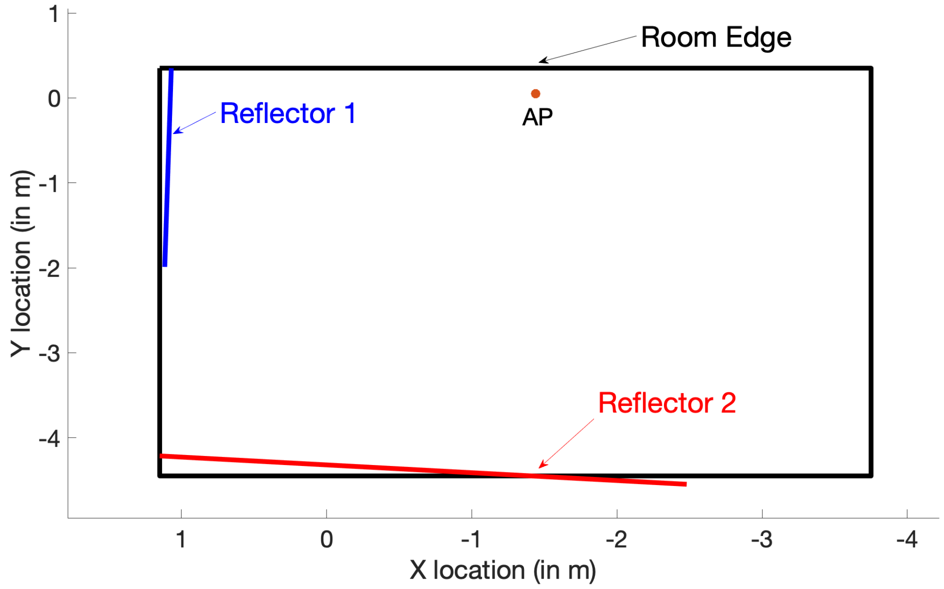


Figure 2.8: Dominant reflector map.

Link Performance Test

After locating the dominant reflectors in the environment, we set the AP to the anchor position as in Figure 2.8 and randomly select client test locations to conduct a comprehensive performance test. For each test location, we first measure the RSSI value of the LoS link under the blockage, then we calculate two NLoS directions corresponding to those two reflectors. A high RSSI value usually indicates a better channel status. For a LoS link without any blockage, the average RSSI value is -50 dB. Figure 2.9 plots the RSSI color maps for four different beam steering strategies under LoS blockage.

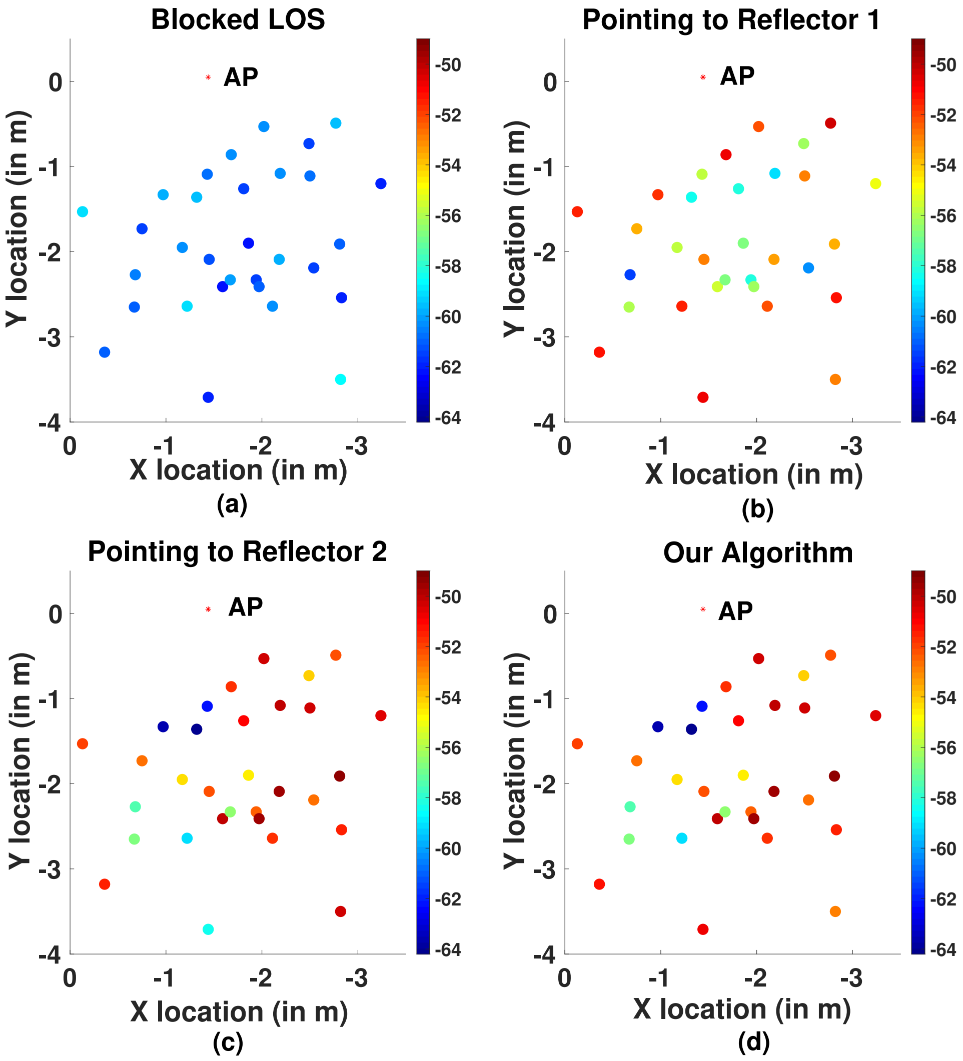


Figure 2.9: RSSI color map.

Figure 2.9a shows the performance under no beam steering strategy. We refer this as the baseline performance. The average RSSI drop is -12 dB. At some test points, the link even suffers from outage. Intuitively, we would think a broad width beam pattern to be beneficial to stabilize the linkage. However, in our test, a broad width beam pattern does not mitigate the performance drop when blockage happens. The finding indicates the LoS link is no longer available for stable mmWave communication under the blockage.

Figure 2.9b shows the performance when the beam direction changes to reflector 1's specular reflection direction. These paths are denoted as $NLoS_1$. In this case, the RSSI values are acceptable for most of the test locations. We observe that test points at the center of the environment usually have lower RSSI values comparing to other points. A possible reason is that although reflector 1 can always create available NLoS links, the performance for different NLoS links highly depend on the client locations.

Figure 2.9c shows the performance when the beam direction changes to reflector 2's specular reflection direction. These paths are denoted as $NLoS_2$. Compared to Figure 2.9b, the performance is better for most of the test locations. As mentioned before, the performance of the NLoS link highly depends on the reflector's physical properties, such as the material and area size. Reflector 2 contains a metal cabinet, which has a larger reflective area than that of reflector 1.

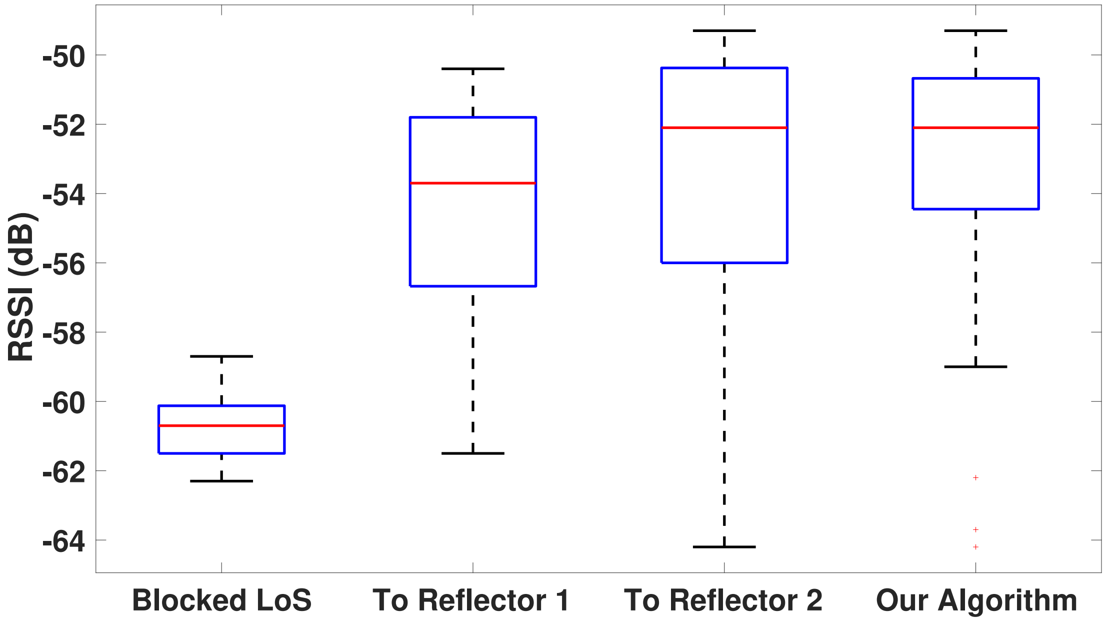


Figure 2.10: Performance of different NLoS links.

Figure 2.9d shows the result of our proposed beam steering algorithm, where the NLoS link is selected based on estimated RSSI values of different NLoS paths. The overall performance is better than using either $NLoS_1$ or $NLoS_2$.

Figure 2.10 shows the numerical results of Figure 2.9. The average RSSI of the LoS blockage case is -61 dB, which is far below that of $NLoS_1$ or $NLoS_2$. $NLoS_2$ has a higher average

RSSI value compared to that of $NLoS_1$(-52.1 dB over -53.5 dB). The RSSI values of $NLoS_1$ are bounded by a tighter range, indicating a greater stability. Our algorithm takes advantage of both reflectors. The average RSSI measured using our system is -52.1 dB, which is the same as the $NLoS_2$. The distribution of the RSSI values is less dispersed. In addition, most of the measurement locations achieve RSSI strength higher than -55 dB, which is a huge performance boost compared to the baseline.

Figure 2.11 shows the CDF of RSSI. From left to right, the lines represent the RSSI of ground truth, our system and LoS link under blockage, respectively. The ground truth is generated by comparing the NLoS link measurement result in each test location and choose the highest RSSI value, which is treated as the oracle value of the current location. The performance of our system is close to the oracle value, which indicates our system can successfully choose the best NLoS link by predicting the link performance using the RSSI estimation model.

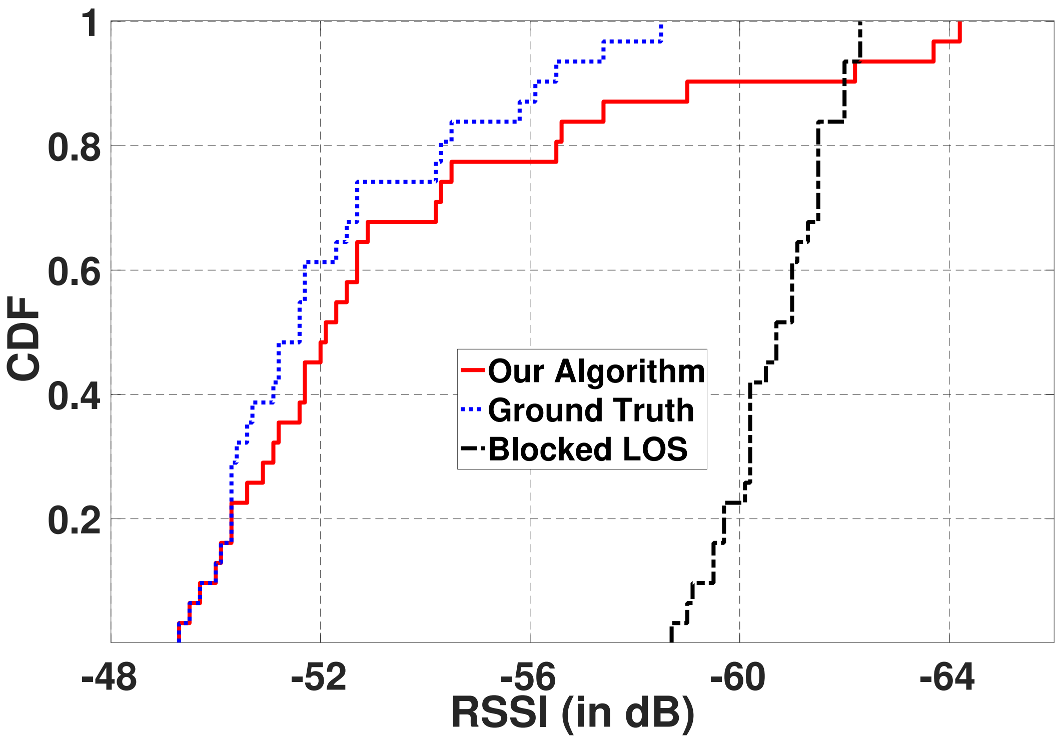
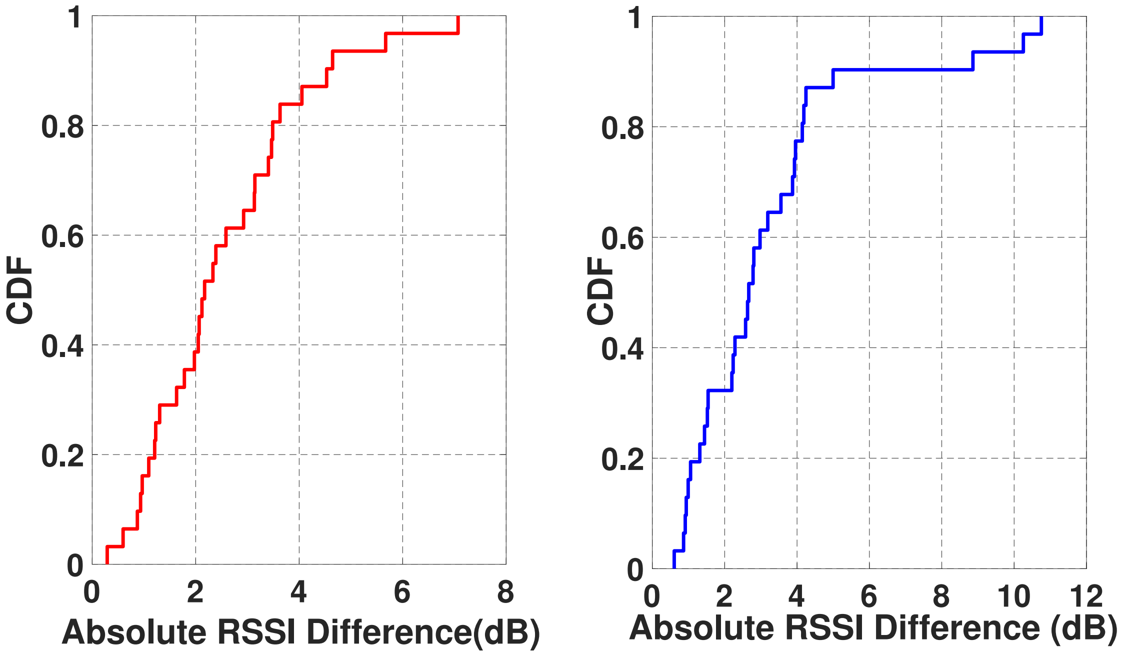


Figure 2.11: RSSI distribution.

Evaluation of Path RSSI Estimation Model

The RSSI estimation model is an essential part in deciding the best NLoS path direction, so we compare our prediction results with the real world measurements. Figure 2.12 illustrates the RSSI estimation model accuracy for the two dominant reflectors. The horizontal axis refers to the estimation error. For both reflectors, about 80% RSSI estimation errors are below 4 dB, and 40% are below 2 dB. Therefore, our RSSI estimation model accurately estimates the RSSI measurement, and thus efficiently assists our system in selecting the best NLoS.



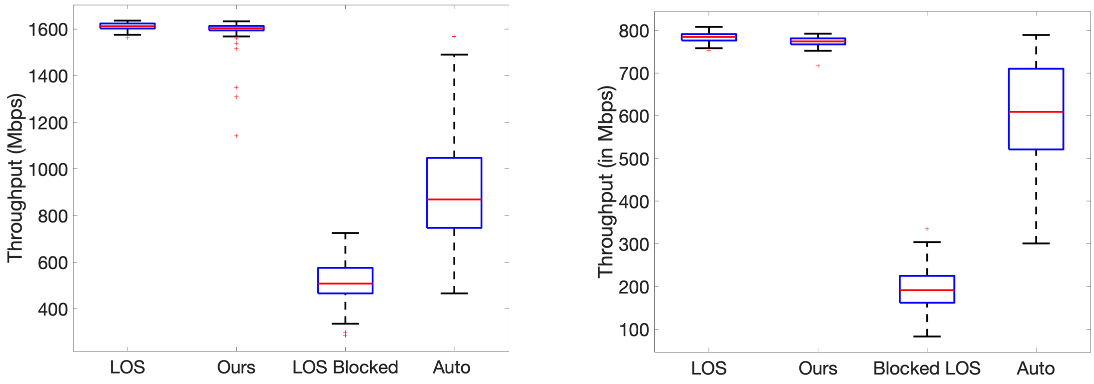
(a) Reflector 1

(b) Reflector 2

Figure 2.12: The CDF of RSSI estimation error.

System Performance under Static and Mobile Scenarios

Our system is designed to handle indoor mobile device communication, such as VR gaming. So we use a bandwidth testing software (integrated in the WAP 60G system) to conduct a transport layer throughput test. To provide more convincing results, we test our system on both UDP and TCP. For all test scenarios, the LoS link test is conducted without any blockage to serve as the upper bound performance, and the LoS direction is blocked by human bodies to test the performance under blockage.



(a) UDP test.

(b) TCP test.

Figure 2.13: Transport layer throughput trace in UDP.

Static scenario performance test: We first perform a static test where we fix the locations of client and AP. The purpose is to test the performance for different methods under a static scenario where the LoS is blocked and each method tries to recover the high performance.

Figure 2.13 shows the bandwidth test results of four methods. The “LoS”, “Ours”, “Blocked LoS” and “Auto” represent the throughput measured by the LoS link, our system, LoS being blocked and auto beam steering method, respectively. For each method, we conduct a 60-second bandwidth test and record the throughput sequence. The UDP result is shown in Figure 2.13(a), the average throughput of “LoS” is 1612 Mbps, and the throughput is stable, as the LoS link is the most reliable link. The average throughput of “Ours” is very close to that of “LoS”, which is 1603 Mbps. The “Blocked LoS” shows that as the RSSI values suffer from a drastic degradation when blockage, the throughput also drops dramatically. The average throughput is only 503 Mbps, which is about only 30% of the throughput of “Ours”. The “Auto” method, which automatically select a beam pattern among the 64 pre-defined patterns to maximize the throughput, has a higher average throughput (869 Mbps) than that of “Blocked LoS”, but only 50% of “Ours”. The distribution of the “Auto” spread widely, with a minimum value 466 Mbps and maximum value of 1569 Mbps.

Figure 2.13(b) shows the TCP test results. Compared to UDP, TCP is more reliable and can tolerate severer signal strength drop. In compensation, the maximum TCP throughput is lower than that of UDP. This is reflected in our result that the average throughput for the "LoS" in TCP (791 Mbps) is about half value of that in UDP. Similar to UDP case, the average throughput of "Ours" in TCP (774 Mbps) is very close to the "LoS" in TCP (791 Mbps). Compared to UDP, the throughput of "Auto" in TCP is closer to "LoS".

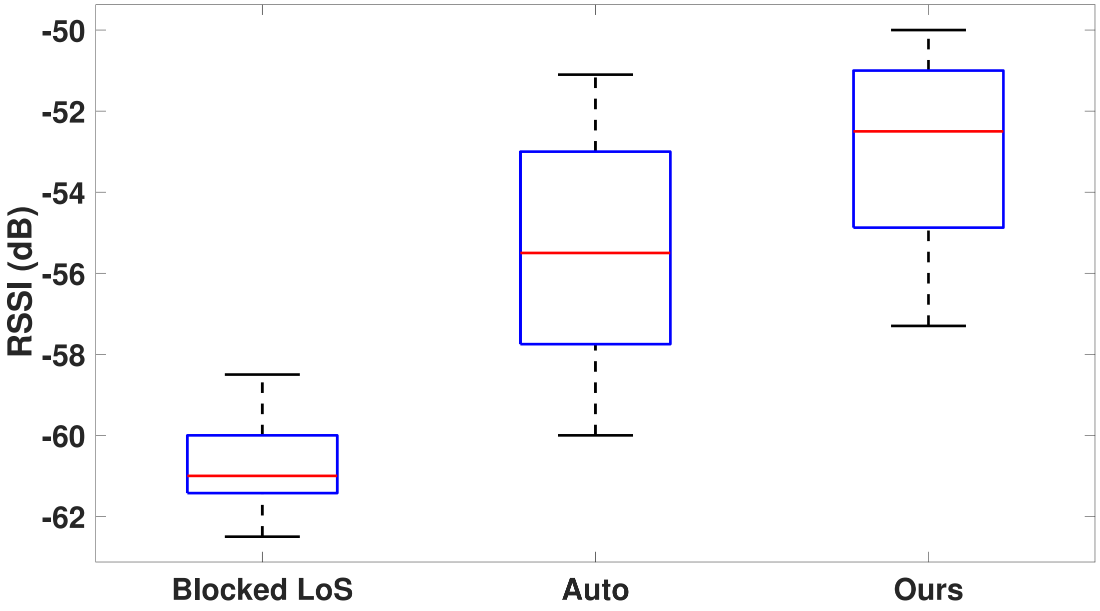


Figure 2.14: Performance comparison with RSSI.

The "Ours" method has a higher TCP/UDP throughput than the "Auto" method since "Ours" can achieve higher RSSI values than the "Auto" method. To justify this point, we test the RSSI of "Ours" and "Auto" under a static scenario and plot the measured RSSIs in Figure 2.14. More specifically, our test is performed under a static scenario where we block the LoS at 15 random locations and collect the RSSIs under "Ours" and "Auto" methods respectively at each location. The RSSI of the blocked LoS is also collected to provide a baseline for the comparison. When the LoS is blocked, the average RSSI of the LoS link drops to -60.6 dB. While both the "Auto" and "Ours" methods bring in some RSSI gains over the blocked LoS, the "Ours" method can achieve a higher RSSI than the "Auto" method. In particular, as shown in Figure 2.14, the average RSSI of

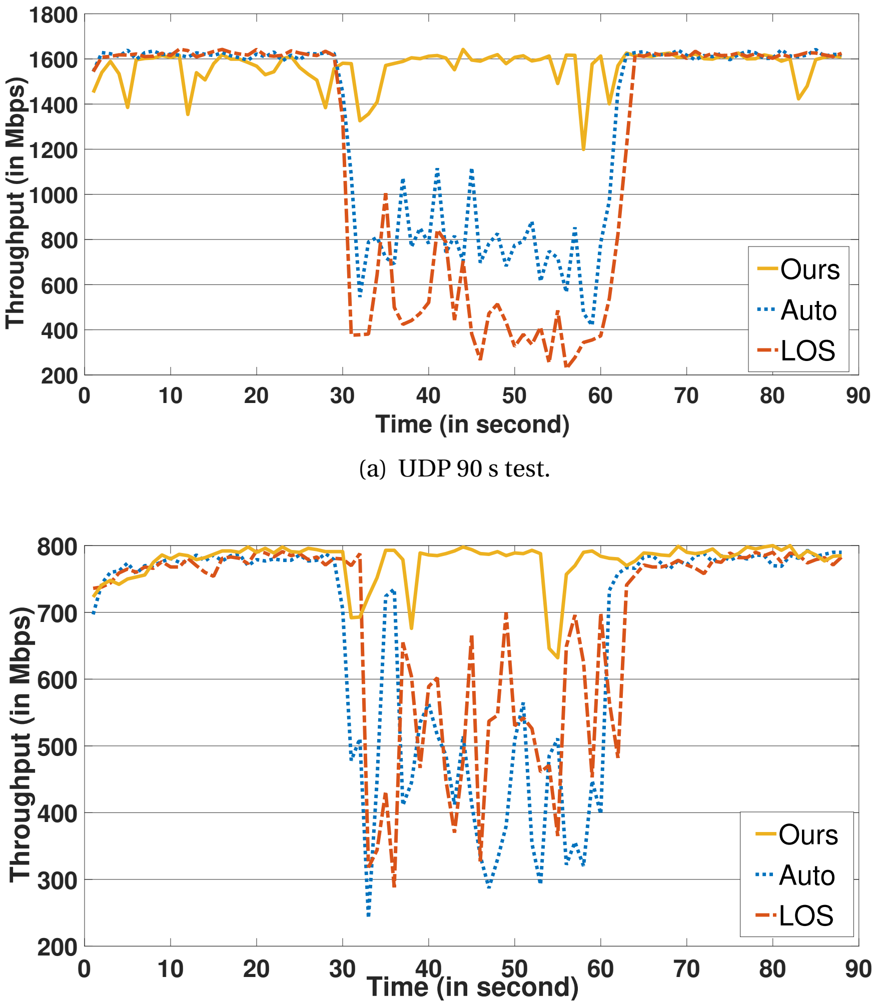
“Auto” is -55.4 dB, while the average RSSI of “Ours” is -52.8 dB, so a 2.6 dB gain over the “Auto” method. In addition, it can be observed that the variance of RSSI under “Ours” is smaller than that of the “Auto”, which indicates that the performance of “Our” is more consistent at different locations than that of the “Auto” method.

The throughput result of “Ours” should be considered as the upper bound performance of our proposed method, which can be achieved when fine-grained steering of the beam (either electronically or mechanically) is available. Note that such an upper bound cannot be achieved by the “Auto” search method, even if fine grained beam steering is available (e.g., by having more higher-resolution patterns in the codebook). This is because a finer grained beam scanning will require the “Auto” method to scan through a larger number of beam patterns, and thus increases the delay for the method to select the optimal pattern, undermining the overall average throughput (where the increased delay should be accounted for as overhead) that can be achieved by the method. On the other hand, when fine grained beam steering is not available to our proposed method, our method will directly pick the pattern in the codebook that covers the estimated optimal NLoS direction. In this case, our method still outperforms the “Auto” method due to its much shorter beam switching delay and faster response time.

In summary, our system outperforms other methods from stability and throughput perspectives, in UDP, TCP and RSSI.

Mobile scenario performance test: We conduct a system level usage test, using TCP and UDP throughput as evaluation metrics. We fix the AP position and move the client across the lab. We conduct a 90-second system test: no blockage to the LoS direction in the first 30 s (0-30 s); human body blockage to the LoS direction continuously following the client movement in the second 30 s (30 s-60 s); blockage moved away from LoS direction in the third 30 s (60 s-90 s). We also test the auto beam steering method for performance comparison.

Figure 2.15(a) shows the test result based on UDP. During the first 30 s, due to no blockage, all three methods select the LoS link for communication. Hence all methods reach a throughput of 1600 Mbps. Then the LoS link is blocked during the second 30 s. When a blockage happens,



(a) UDP 90 s test.

(b) TCP 90 s test.

Figure 2.15: Transport layer throughput trace in TCP.

both “Auto” and “LoS” suffer from instantaneous performance drop. Due to the fact that the “LoS” only uses the direct LoS direction whereas the “Auto” selects among different beam patterns to maximize the throughput, the performance for “Auto” is better than “LoS” when blockage happens. However, both “Auto” and “LoS” throughput drop below 800 Mbps, which is only 50% of the maximum speed of the system.

As for our system, the throughput slightly drops when the blockage occurs. Then the throughput swiftly restores close to the maximum throughput, which is around 1600 Mbps.

The overall throughput is stable during entire blockage period, but there still exist some unstable points. This is because the NLoS path signal strength is not uniformly distributed in the environment. Therefore, the throughput fluctuates as the client goes through strong and weak NLoS signal strength areas. This is consistent with our previous finding in Figure 2.9: the NLoS link performance highly depends on the client's location. In the meantime, the client changes the beam direction accordingly to maintain the high performance. We remove the blockage at the 60th second, bringing the LoS link available again. So all three methods can use the LoS direction and the throughput restores to the maximum level.

Figure 2.15(b) shows the TCP test results under the same test setting. Due to the error handling mechanism in TCP, the performance drop of "Auto" and "LoS" during the 30 s to 60 s is less compared to that of UDP, but still catastrophic for the mmWave communication.

Similar to UDP case, our system can maintain a throughput to an "almost LoS link" (800 Mbps) throughput level with a little fluctuation. Our system shows a superiority in stability and performance aspects.

In summary, our system successfully handles the LoS blockage under mobile scenarios for COTS mmWave devices and provides robust link for mmWave communication.

2.5 Conclusions

In this paper, we develop a NLoS beam switching algorithm for off-the-shelf mmWave devices to maintain a stable connection when its LoS communication path is blocked. The main idea of our method is to leverage the sparsity of the mmWave channel and the spatial correlation of the close-by mmWave channels to resolve for the location and orientation of the dominant reflectors in the environment. Strong NLoS backup paths are then computed based on these resolved dominant reflectors. We also propose a model-driven RSSI estimation algorithm, which allows the transmitter to predict the quality of each backup NLoS path and pick the best one among them.

In contrast to existing methods, our model does not rely on high precision phased array antennas, nor does our model require accurate phase information of the received signals, and therefore is applicable to a wide line of COTS mmWave products. We validate the feasibility and effectiveness of our system on a mmWave off-the-shelf testbed and demonstrate that it supports efficient and stable mmWave communication under human blockage.

Our system can serve as a prototype for off-the-shelf mmWave devices to handle the LoS blockage. The simplicity and low cost of our system can benefit a wide range of low-end commercial mmWave devices.

Chapter 3

(k, α)-Coverage for RIS-aided mmWave Directional Communication

3.1 Introduction

Millimeter-wave (mmWave) communication is an essential technology in the next generation high-speed 5G networks. However, a fundamental problem of mmWave communication is the susceptibility to line-of-sight (LoS) blockage [91, 46]. In particular, the mmWave transmission relies on directional communication to overcome the high free-space attenuation and the signal's propagation heavily relies on LoS paths [66, 65]. When the LoS is blocked by an obstacle, the mmWave signal cannot penetrate through or circumvent around the obstacle, resulting in a significant drop of the received signal strength. In this situation, a possible solution is to steer the communication beam towards a strong non line-of-sight (NLoS) propagation path to maintain the communication quality. However, the NLoS paths generated by a natural environment are inherently unstable and uncontrollable, whose quality may not be good enough to be used for mmWave communication. To prevent the sudden performance drop caused by random blockage to the LoS path and increase the communication robustness, a novel concept called reconfigurable intelligent surface (RIS) [139, 129, 28, 40] is proposed, which can provide controllable high quality NLoS paths for mmWave directional communication.

RIS, also called intelligent reflecting surface or smart reflect-array, is a planar surface formed by a large group of passive reflecting meta-elements, each of which is able to control the phase and amplitude of incident signal independently by dynamically tuning an on-board positive-intrinsic-negative (PIN) diode [130]. The joint effect generated by massive number of reflecting meta-elements can change the propagation property of the reflected signal, and achieve numerous functionalities such as anomalous reflection, beam refocus [47] and beam split [7]. By smartly deploying multiple RISs in the environment and coordinating their reflection properties, the wireless channels in the environment can be artificially reconfigured, which provides new means to fundamentally increase the quality of NLoS channels.

However, how to deploy RISs to provide effective NLoS paths for mmWave directional communication, which is often modeled as a coverage problem, still remains a challenge, as directions of NLoS paths have decisive influence on the paths' availability. In particular, if the available NLoS paths are all from the same or nearby directions, when an obstacle blocks one of the paths, it is likely that other paths are also blocked, leading to failure of multiple NLoS paths and degradation of the communication quality. Therefore, a desirable case in mmWave directional communication is that the receiver is not only within the transmission ranges of (i.e., being covered by) multiple RISs at the same time, but also the signal propagation paths from these RISs to the receiver are separated far enough in their directions. In this case, when the current communication direction is blocked by an obstacle, other paths are still likely to be available.

In the literature, although there have been extensive results on the coverage problem in the context of wireless sensor networks, few of them assure the minimal angular separation between different signal propagation/sensing paths when designing the coverage strategies. In traditional isotropic sensor networks, most sensor deployment strategies aim to guarantee that every point in the area is within the sensing ranges of a certain number of sensors. For example, in the wildly studied k-coverage problem [35, 36], a receiver is required to be within the sensing ranges of at least k sensors. However, the k-coverage problem only considers the distance between receiver and sensors as the coverage criteria and there is no requirement on the angular separations between the paths from each sensor to the receiver (a.k.a. angle of arrivals or AoAs). In the directional sensor networks, i.e., sensors with a limited angle of view, e.g., cameras, existing literature on coverage mainly study the following three types of problems: (1) single-perspective coverage [61, 18], which ensures that any point in a target area is always within at least one sensor's angle of view; (2) multi-perspective coverage [76], which provides at least k views at the same time for the same object from k different directions (as seen by k directional sensors); (3) full-perspective (360°) angular coverage [150, 133], which guarantees that the union of the angle of views of the sensors covers every point of the object, i.e., there is no blind point on the object that cannot be viewed/sensed by any of the sensors. Despite their different goals, these works mainly aim to combine the limited angle of views of several directional sensors to partially or fully cover an object. Angular separation between AoAs is simply not a consideration/objective in these problems.

Motivated by the drawbacks of existing coverage models, in this paper we propose a novel min-number (k,α) area coverage (MNkaAC) problem. In particular, we define that a receiver is (k,α)-covered if it is within the transmission range of at least k RISs so that not only it has k different NLoS communication paths, but also, from the receiver's perspective, the angular separation between any two adjacent NLoS paths must be not smaller than a predefined angle α. The angular separation parameter α may vary for different application scenarios to provide different levels of anti-path-blockage robustness. Our MNkaAC problem aims to minimize the number of RISs deployed in a target area such that every point in the area is (k,α)-covered. The proposed MNkaAC problem not only ensures that the number of available NLoS paths is at least k, but also guarantees that the directions of these NLoS paths are sufficiently separated to avoid simultaneous paths failure, at the minimum cost of RIS deployment. With the consideration of angular separation, the model can enhance the robustness of RIS-aided mmWave communication network, especially in the scenarios where random blockages are common.

Solving the MNkaAC problem faces the following unique challenges. Firstly, to satisfy the angular separation requirement α, the deployment of an RIS has to be dependent on the deployment of neighboring RISs (this is because the angular separation of two NLoS paths is defined by two RISs and the receiver). This is in sharp contrast to the traditional k-coverage problem, in which the deployment of multiple RISs is essentially a range-coverage problem and thus can be considered independently. The dependent RISs deployment makes the (k,α)-coverage problem much more challenging. Secondly, as will be shown shortly, in contrast to the traditional coverage model that assumes a regular disk or pie-shaped coverage area for each RIS, the coverage area of an RIS in (k,α)-coverage model is irregular and has a complex relationship with the parameter α. How to characterize such a complicated and irregular coverage area is challenging.

This paper takes the first step to systemically study the MNkaAC problem under the aforementioned challenges. Our contributions are summarized as follows:

- We formulated and solved the MNkaAC problem for RIS-aided mmWave directional communication networks respectively under two different RIS deployment conditions, namely, deterministic RIS deployment and randomized RIS deployment (and hence a deterministic version and a random version of MNkaAC problem). In the former case one has full control over the exact position where each RIS should be placed, while in the latter case there is no such control and instead RISs are deployed randomly so that their positions after the deployment follow certain spatial distribution. The deterministic deployment model is usually applicable for installing a mild number of RISs to cover a limited target area (e.g., a shopping mall or an apartment building), while the random deployment model applies to the scenarios of large-scale RIS deployment.

- To develop solutions to the problem of both versions, we first give formal definitions for the concepts of point-(k,α)-coverage and area-(k,α)-coverage, and then propose and prove a necessary and sufficient condition for an area to be (k,α)-covered. Based on this condition, two efficient feasibility-check methods are proposed to decide if a target area is (k,α)-covered by a given set of RISs.

- We derive a quasi-optimal solution to the deterministic version of the MNkaAC problem. Specifically, we first find an optimal regular k-sided polygon deployment pattern and use it to obtain an approximate solution (i.e., a good feasible solution) to the deterministic

MNkaAC problem. We then derive a performance bound on this solution in terms of its optimality gap, i.e., an upper bound that curbs the ratio of the number of RISs required by this approximate solution to that required by the optimal solution.

- Under the random RIS deployment condition, for a given RIS deployment density, we show that the target area can only be (k,α)-covered with a probability (defined as the probability that an arbitrary point in the area is (k,α)-covered). So the random version of MNkaAC problem is formulated as an optimization to minimize the required RIS density under the constraint of a desired coverage probability for the target area. We solve this problem under two RIS distributions, uniformly random distribution and Spatial Poisson distribution. Under each distribution, we derive the functional relationship between the RIS density and its (k,α)-coverage probability, based on which the minimum RIS density to reach the desired coverage probability can be readily decided.

The rest of the paper is organized as follows. Section 3.2 reviewed the related work. Section 3.3 defines the (k,α)-coverage problem and proposes two (k,α)-coverage verfication methods. Section 3.4 proposes a quasi-optimal solution to the MNkaAC problem and derives analytical performance bound of the solution under the deterministic deployment scenario. Section 3.5 derives the (k,α)-coverage probability under the random deployment scenario. Section 3.6 provides numerical results for our analyse and we conclude our work in Section 3.7.

3.2 Related Work

Coverage problem in wireless sensor network (WSN) has been extensively studied in the past few years. There are many existing works in area coverage in traditional scalar wireless sensor networks (WSN) [85, 31, 48, 5, 154]. In this section, we review the most relevant works.

3.2.1 Coverage in RIS

As the concept of RIS becomes more and more popular, coverage problem in RIS has attracted extensive interests [13, 32, 100, 115]. Although most of the works are under directional communication scenarios, none of their works considers the angular separation between two communication paths. Authors in [78] determined the optimal RIS placement w.r.t. the transmitter and receiver antenna positioning. Their results indicated that the RIS should be optimally placed closer to the receiver than the transmitter so that the signal to noise ratio is maximized. In [144], the authors focused on the joint effect of RIS orientation and placement. In particular, their work maximized the cell coverage by adjusting the RIS orientation and the horizontal distance between the RIS and the base station. However, most of their research focus on increasing the overall RIS network performance by changing the distance between transmitter and RIS, and none of their works pays attention to the angular separation among different RIS-created paths.

3.2.2 Coverage in mmWave Network

The coverage problem in the mmWave network has drawn much attention [74, 10]. Since the mmWave network is highly dependent on LoS path for signal transmission, most studies consider the coverage and placement of the mmWave nodes with the aim of maximizing the LoS coverage area. For example, in [114, 81], the authors automated the process of placing mmWave access points (APs) by computational geometry to maximize the LoS coverage area in dense cities. In [87], the authors considered the impact of the height of the AP placement on the probability of LoS coverage. The deployments of high-rise and low-rise APs were jointly optimized to maximize the probability of LoS coverage of the target field. In addition to the LoS coverage area, the authors in [9] proposed a geometry-and-blockage-aided coverage model for mmWave communication with RISs, which also considered the area covered by the first-order reflection of RISs. Then, the problem of deploying APs and RISs is formulated as a maximization of the total coverage area under several practical constraints.

To adapt to the mobile user scenario and increase the reliability of mmWave communication, some research works consider the angular separation between different paths in the coverage model. In [137], the authors formulated the problem of placing mmWave APs to maximize both the number of multipath channels and the angular spread between these channels. And in [30], the authors developed an AP deployment scheme that maximizes the average minimum angular separation between paths to increase the robustness of communication. However, their scheme is based on mixed-integer linear programming and can be applied only to point coverage problems. In contrast, in our paper, we build a geometry model for path separation and develop a pattern-based deployment scheme, in which the requirement of angular separation is always guaranteed for both the point and area coverage problems.

3.2.3 Coverage in Directional Sensor Network

In directional sensor network, the sensor can only detect objects within a limited angle of view. Any target that lies out of the angle of view cannot be detected, regardless of the distance to the sensor. Also, the viewing direction of the sensor have important influence on the coverage quality. So based on the coverage goal, we divide the work in direction sensor network into three categories (1) single-perspective coverage, which ensures any point in a target area is within at least one sensor's angle of view. For instance, in [18], the authors investigated the problem of using the least number of directional sensors to cover 3-D space while maintaining the connectivity and link quality. And a novel distributed parallel multi-objective evolutionary algorithm is also proposed to solve the problem. And in [83], the authors formulated a coverage and energy consumption optimization problem and solve it by the improved adaptive particle swarm optimization. (2) Multi-perspective coverage. In the multi-perspective coverage model, the goal is to provide distinct views for an object from k different directions, so an object must be covered by at least k sensor at a time. As in [76], the authors proposed multi-perspective coverage, in which the cameras are used to gather disparate views of events from different perspectives. A novel metric is proposed to measure the multi-perspective coverage for a region from a given number of perspectives. Further, in [140], an exact algorithm based on binary-level mixed integer programming was proposed to provide the multi-perspective coverage with the least number of camera sensors. Although the authors pointed out that for the optimal multi-perspective coverage, the angular separation between different camera's viewing directions must be maximized, they do not constrain the minimal angular separation. (3) Full-perspective coverage, or full angular coverage, which is to select a set of cameras circle around an object and the union of their angle of view arcs fully covers the object (i.e., creates a 360° circle). For instance, the work in [150] considered the problem of determining optimal camera placement to achieve angular coverage continuously over a given region. A bi-level algorithm is proposed to find the minimum-cost camera placement. And in [23], the angle coverage problem is combined with an additional image resolution requirement, the author transform it into a set cover problem and propose a greedy algorithm to solve it. However, all the aforementioned fails guarantee the angular separation among different viewing directions.

Full-view coverage is another important coverage model in camera sensor networks [138, 52, 128, 45, 132]. Wang *et al.* [138] proposed a novel coverage model called full-view coverage. The full-view coverage model requires the object's face direction to be captured by at least one camera sensor. He *et al.* [45] further showed that full-view area coverage problem can be transferred to full-view point coverage problem, and several set cover based algorithms were therefore proposed to solve the full-view point coverage problem. Wu *et al.* [132] investigated the necessary and sufficient conditions to achieve full view coverage under different random deployment schemes. Their work provides a performance bound for randomly deployed cameras sensor to guarantee area full-view coverage. Note that to achieve the full-view coverage, an object is required to be monitored by a set of cameras and the angular separation between two adjacent cameras must be less than a predefined effective angle. While in the proposed (k,α)-coverage model, an object is required to be covered by k different RISs and the angular separation between two adjacent RISs must be greater than the parameter α, which is different from the full-view coverage model. In addition, in this work, we investigate the (k,α)-coverage problem under both random and deterministic deployment scenarios and study the problem of selecting the minimal number of RISs to ensure the (k, α) area coverage.

Among all the existing literature, [119, 134] are the most relevant to our work. Especially in [119], Tseng *et al.* proposed the k-angle object coverage model for camera sensor network. The k-angle object coverage model requires an object to be simultaneously monitored by at least k sensors from multiple angles satisfying certain angle constraints ω. Although the definition is similar to our (k, α)-coverage model, their work only defines the coverage for discrete points and investigates the problem of using the minimum number of sensors to k-angle cover the maximum number of discrete objects. In particular, they defined a coverage contribution function of each sensor and propose a greedy algorithm to pick the sensor with the highest contribution to the total coverage. In contrast, in our work, we not only define the (k, α)-coverage for discrete points, but also investigate the geometry relationship of (k, α)-coverage and provide a necessary and sufficient condition for a continuous region to be (k, α)-covered. Meanwhile, we derive the coverage probability under random deployment scheme and use deployment patterns to achieve (k, α)-coverage under the deterministic scheme, which provide more insightful results than their works.

3.3 (k, α)-Coverage Model and Verification

In this section, we provide basic definitions for the (k, α)-coverage and propose two methods to verify if a target area is (k, α)-covered by a given set of RISs. We first define the (k, α)-coverage and the MNkaAC problem. Then, we derive a necessary and sufficient condition for an area to be (k, α)-covered (Lemma 3.3.2). Finally, we show that (k, α)-covering an area is equivalent to (k, α)-covering the boundary of the area (Theorem 3.3.1). Based on this equivalence, we propose two methods to verify whether a region is (k, α)-covered. Our analysis is inspired by the works in [138, 45].

3.3.1 Notations

We consider that multiple mmWave-based RISs are deployed in a bounded area $\mathscr{A}$ (target region) to provide controllable high-quality NLoS paths for mmWave communication. And the channel between the RIS and the receiver is the LoS channel.The motivation to consider the RIS-aided mmWave communication, rather than other mmWave technologies such as the conventional cooperative relay, is that RIS is a more relevant and suitable application of the (k,α)-coverage problem studied in this paper. The proposed (k,α)-coverage problem is most valuable when the problem setting involves a large-scale and dense deployment of nodes to cover a large target area. Since RISs are small planar surfaces made up of passive reflecting elements, which are low-cost, nearly energy-free, and easy to deploy, they can be flexibly and pervasively attached to almost any surface such as billboards and buildings, to provide strong and pervasive reflection for mmWave signals

Denote the set of RISs deployed in the region by $\mathscr{N}$ and $|\mathscr{N}|$ represents the cardinality of $\mathscr{N}$. We use R_i to represent the i-th RIS ($1 \leq i \leq |\mathscr{N}|$). Without ambiguity, we also use R_i to denote the RIS position, and all RIS has the same transmission radius r. Let $d(P_1, P_2)$ denote the Euclidean distance between points P_1 and P_2. The communication coverage model for RIS is a disk coverage model: an RIS R_i can cover point P if P is within the transmission range of R_i, i.e., $d(P, R_i) \leq r$. And let $CS(P)$ denote the coverage set of point P, which is the set of all RISs that cover P.

Note that the novelty of the (k,α)-coverage model lies in the special requirement for the angular separation between two adjacent NLoS paths. To quantify the path angular difference, we use the vector $\overrightarrow{PR_i}$ to represent the NLoS path direction from point P to RIS R_i, and let $\mathrm{ANG}(\overrightarrow{PR_i},\overrightarrow{PR_j})$ denote the path angular difference between two path directions $\overrightarrow{PR_i}$ and $\overrightarrow{PR_j}$, which ranges from 0 to π (0 to 180°).

3.3.2 (k, α)-Coverage Model and Problem

We first give basic definitions for the (k, α)-coverage model and problem. Since the angular separation α is required between two adjacent paths, we use the concept of coverage circular list to accurately define which paths are adjacent.

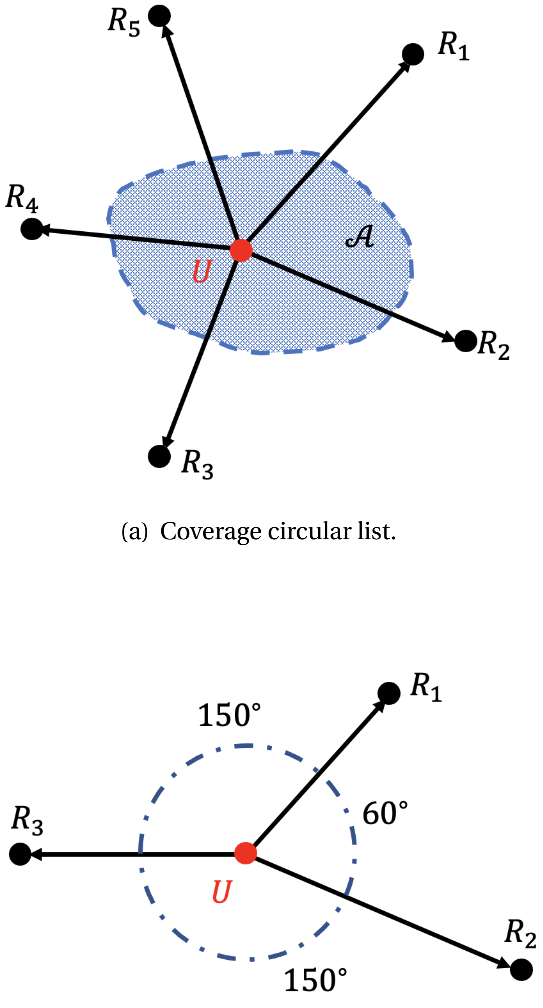
Definition 3.3.1 (**Coverage circular list**)**.** The coverage circular list for a point P, denote as $CCL(P)$, is a sequence of RISs in $CS(P)$ in counterclockwise (or clockwise) direction. And the coverage circular list for region $\mathscr{A}$, denoted as $CCL(\mathscr{A})$, is a sequence of RISs that preserves the same counterclockwise (or clockwise) direction for every point in $\mathscr{A}$.

For example, in Figure 3.1(a), point U is an interior point of $\mathscr{A}$ and $\mathscr{A}$ is covered by R_1, R_2, R_3, R_4 and R_5, i.e., $CS(\mathscr{A}) = CS(U) = \{R_1, R_2, R_3, R_4, R_5\}$. The coverage circular list for point U and region $\mathscr{A}$ is $CCL(\mathscr{A}) = CCL(U) = \{R_1, R_2, R_3, R_4, R_5\}$. As the list is circular, it can be also represented as $CCL(\mathscr{A}) = CCL(U) = \{R_3, R_4, R_5, R_1, R_2\}$. In the rest of the paper, the adjacent paths are the paths created by two consecutive RISs in the coverage circular list. e.g., In Figure 3.1(a), two adjacent paths can be $\overrightarrow{UR_1}$ and $\overrightarrow{UR_2}$ or $\overrightarrow{UR_5}$ and $\overrightarrow{UR_1}$. Based on the coverage circular list, we can further define the (k, α)-coverage.

Definition 3.3.2 (**Point** (k, α)**-coverage**)**.** A point P is (k, α)-covered if there exists a $CCL(P)$ with k elements ($k \geq 2$), denoted as $CCL(P) = \{R_1, R_2, \dots, R_k\}$, in which the angle separation between any two adjacent path vectors $\overrightarrow{PR_i}$ and $\overrightarrow{PR_{i'}}$ is greater than or equal to α, i.e., $\text{ANG}(\overrightarrow{PR_i}, \overrightarrow{PR_{i'}}) \geq \alpha$, for any $i \in [k]$ and $i' = \mod(i+1, k)$, where $\mod$ is the modulo operator and $0 \leq \alpha \leq \frac{2\pi}{k}$.

Figure 3.1(b) shows an example for a point U to be $(3, 50°)$-covered. In this case, the coverage circular list for point U is $CCL(U) = \{R_1, R_2, R_3\}$. The angular separation between any two adjacent path vectors is greater than or equal to 50°, i.e., $\text{ANG}(\overrightarrow{UR_1}, \overrightarrow{UR_2}) = 60° \geq 50°$, $\text{ANG}(\overrightarrow{UR_2}, \overrightarrow{UR_3}) = 150° \geq 50°$ and $\text{ANG}(\overrightarrow{UR_3}, \overrightarrow{UR_1}) = 150° \geq 50°$.

Definition 3.3.3 (**Area** (k, α)**-coverage**)**.** An area is (k, α)-covered if and only if every point in the area is (k, α)-covered. In the rest of the paper, the (k, α)-coverage refers to area (k, α)-coverage unless otherwise specified.



(a) Coverage circular list.

(b) A (3,50°)-coverage case.

Figure 3.1: A circular list and a point (k, α)-coverage examples.

Definition 3.3.4 (**Minimum-number** (k,α) **area coverage (MNkaAC) problem**)**.** For a target region and a collection of RISs $\mathscr{N}$ in the region, the minimum-number (k,α) area coverage (MNkaAC) problem is to find a (k,α) coverage set C of the target region, such that the cardinality of C, i.e., $|C|$, is minimized.

3.3.3 (k,α)-Coverage Verification

To tackle the MNkaAC problem, we need to first verify whether an area is (k,α)-covered by a given set of RISs. However, it is practically infeasible to verify if an area is (k,α)-covered by Definition 3.3.3, because there are infinite number of points in the area. Therefore, in this subsection, we propose two methods to detect if a target region is (k,α)-covered by a set of deployed RISs.

Given a set of deployed RISs $\mathscr{N}$, the target area $\mathscr{A}$ can be partitioned into sub-areas, and each sub-area $\mathscr{S}\mathscr{A}$ is defined as a set of points with the same coverage circular list. This can be done by first dividing the area by the communication border of all deployed RISs and the border of $\mathscr{A}$, such that each sub-area is covered by the same set of RISs. Then, for each sub-area with the same coverage set, we further divide it by the method provided in [132] to obtain the sub-areas with the same coverage circular list. In the following discussion, the sub-area refers to the sub-area with the same coverage circular list.

It is clear that area $\mathscr{A}$ is (k,α)-covered if every sub-area $\mathscr{S}\mathscr{A}$ is (k,α)-covered. However, verifying if $\mathscr{S}\mathscr{A}$ is (k,α)-covered is still time-consuming as it is a 2-D region. To tackle this problem, we first provide and prove a necessary and sufficient condition for the (k,α)-coverage, which can be used to verify if a sub-area is (k,α)-covered. Then, we further reduce the problem's dimension by showing that $\mathscr{S}\mathscr{A}$ is (k,α)-covered if and only if its boundary, denoted as $\partial\mathscr{S}\mathscr{A}$, is (k,α)-covered. As a result, the area (k,α)-coverage verification problem is reduced to a boundary (k,α)-coverage verification problem. And two methods are proposed to verify if the boundary is (k,α)-covered.

Necessary and Sufficient Condition for (k,α)-Coverage

We begin by introducing the concept of feasible region, which is defined as the $(2,\alpha)$-coverage area generated by two RISs. The concept of feasible region serves as a building block for our area (k,α)-coverage verification.

Definition 3.3.5 (**Feasible region and infeasible region**)**.** For any two RISs R_i and R_j, the feasible region is defined as the region in which any point P satisfies $\text{ANG}(\overrightarrow{PR_i}, \overrightarrow{PR_j}) \geq \alpha$ for a given α ($\alpha \leq \pi$); and the infeasible region is defined *vice versa*, in which any point P satisfies $\text{ANG}(\overrightarrow{PR_i}, \overrightarrow{PR_j}) < \alpha$ for a given α ($\alpha \leq \pi$).

Next, we show how to find the feasible and infeasible region.

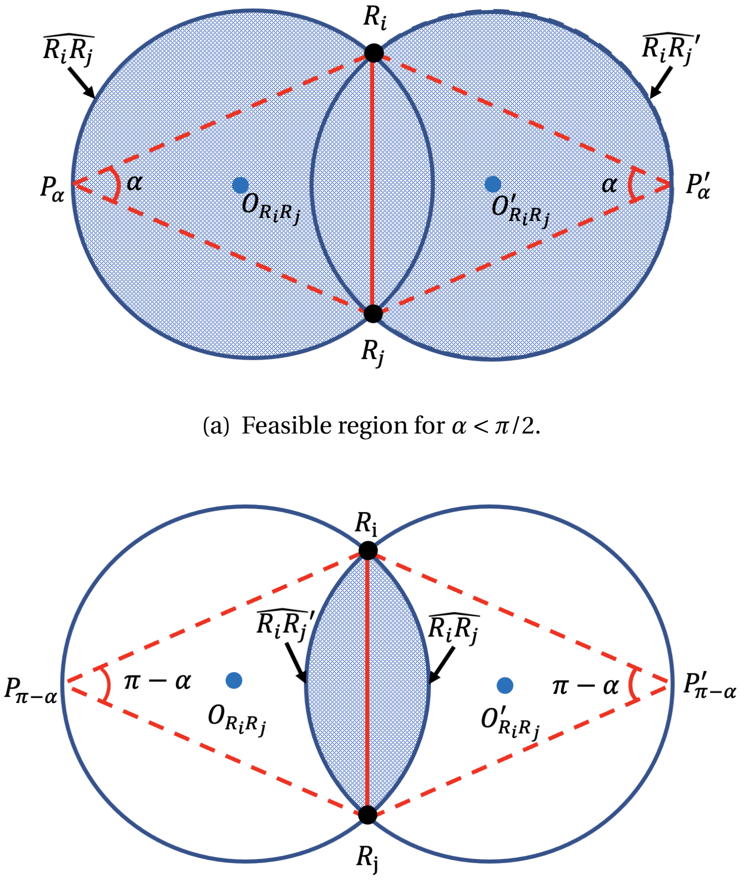
Lemma 3.3.1. Given R_i and R_j, there are two arcs $\widehat{R_iR_j}$ and $\widehat{R_iR_j}'$, which connect R_i and R_j and are symmetrical with respect to line $\overline{R_iR_j}$. The feasible region is the enclosed region bounded by the two arcs and the infeasible region is the complementary region of the feasible region.

Proof. We proof the lemma by finding the two arcs. The cases for $0 \leq \alpha \leq \pi/2$ and $\pi/2 < \alpha \leq \pi$ are discussed separately.

(1) For $0 \leq \alpha \leq \pi/2$ (illustrated in Figure 3.2(a)), we first find a point P_α on the mid-perpendicular line of $\overline{R_iR_j}$ such that $\angle R_iP_\alpha R_j = \alpha$. Based on basic geometry, there are two points that satisfy this condition. Denote the other point as P'_α. We obtain two triangles $\triangle R_iR_jP_\alpha$ and $\triangle R_iR_jP'_\alpha$. Next, we draw the circumcircles for triangles $\triangle R_iR_jP_\alpha$ and $\triangle R_iR_jP'_\alpha$ respectively, and denote the centers of the two circumcircles as $O_{R_iR_j}$ and $O'_{R_iR_j}$, respectively. Based on geometric relationships, the feasible region is the union of the two circumcircles $\odot O_{R_iR_j}$ and $\odot O'_{R_iR_j}$. Then the arc $\widehat{R_iR_j}$ is a portion of the perimeter of $\odot O_{R_iR_j}$ on the left and arc $\widehat{R_iR_j}'$ is a portion of the perimeter of $\odot O'_{R_iR_j}$ on the right. As shown in Figure 3.2(a), the shaded area is the feasible region for $\alpha \leq \pi/2$;

(2) For the $\pi/2 < \alpha \leq \pi$ case (illustrated in Figure 3.2(b)), the steps are similar but the result is different. First, we find two points on the mid-perpendicular line of $\overline{R_iR_j}$ such that

$\angle R_i P_\alpha R_j = \pi - \alpha$. Denote these two points as $P_{\pi-\alpha}$ and $P'_{\pi-\alpha}$, respectively. Then, we find the circumcircles for $\triangle R_i R_j P_{\pi-\alpha}$ and $\triangle R_i R_j P'_{\pi-\alpha}$ as $\odot O_{R_i R_j}$ and $\odot O'_{R_i R_j}$, respectively. In this case, the feasible region is the intersection of the two circumcircles. And the arc $\widehat{R_i R_j}$ is the right portion of the perimeter of $\odot O_{R_i R_j}$ and arc $\widehat{R_i R_j}'$ is the left portion of the perimeter of $\odot O'_{R_i R_j}$. As shown in Figure 3.2(b), the shaded area is the feasible region for $\pi/2 < \alpha < \pi$. □



(a) Feasible region for $\alpha < \pi/2$.

(b) Feasible region for $\alpha > \pi/2$.

Figure 3.2: Feasible region.

Based on the definitions of the feasible and infeasible region, we can provide a necessary and sufficient condition for area (k, α)-coverage.

Lemma 3.3.2 (**Necessary and sufficient condition**). For a sub-area $\mathscr{SA}$ with a coverage circular list with cardinality k, i.e., $CCL(\mathscr{SA}) = \{R_1, R_2, \ldots, R_k\}$, $\mathscr{SA}$ is (k, α)-covered if and only if

$\mathscr{S}\mathscr{A}$ is contained in the intersection of feasible regions of all pairs of R_i and R'_i, where $i \in [k]$, $i' = \mod(i+1, k)$ and $0 \leq \alpha \leq 2\pi/k$.

Proof. This is a result from Lemma 3.3.1 and (k, α)-coverage model. □

By using Lemma 3.3.2, we can easily verify if a region is (k, α)-covered. Figure 3.3(a) shows two area (k, α)-coverage examples. In Figure 3.3(a), the dotted circles represent the feasible regions for the four pairs of RISs, R_1R_2, R_2R_3, R_3R_4 and R_4R_1. The region of interests, $\mathscr{A}$, is $(4, \alpha)$-covered since it is fully contained in the intersection of all feasible regions (the shaded area). As for Figure 3.3(b), the feasible regions for RIS pairs, R_1R_2, R_2R_3, R_3R_4, R_4R_5 and R_5R_1, do not have a common overlapping area, therefore $\mathscr{A}$ is not $(5, \alpha)$-covered.

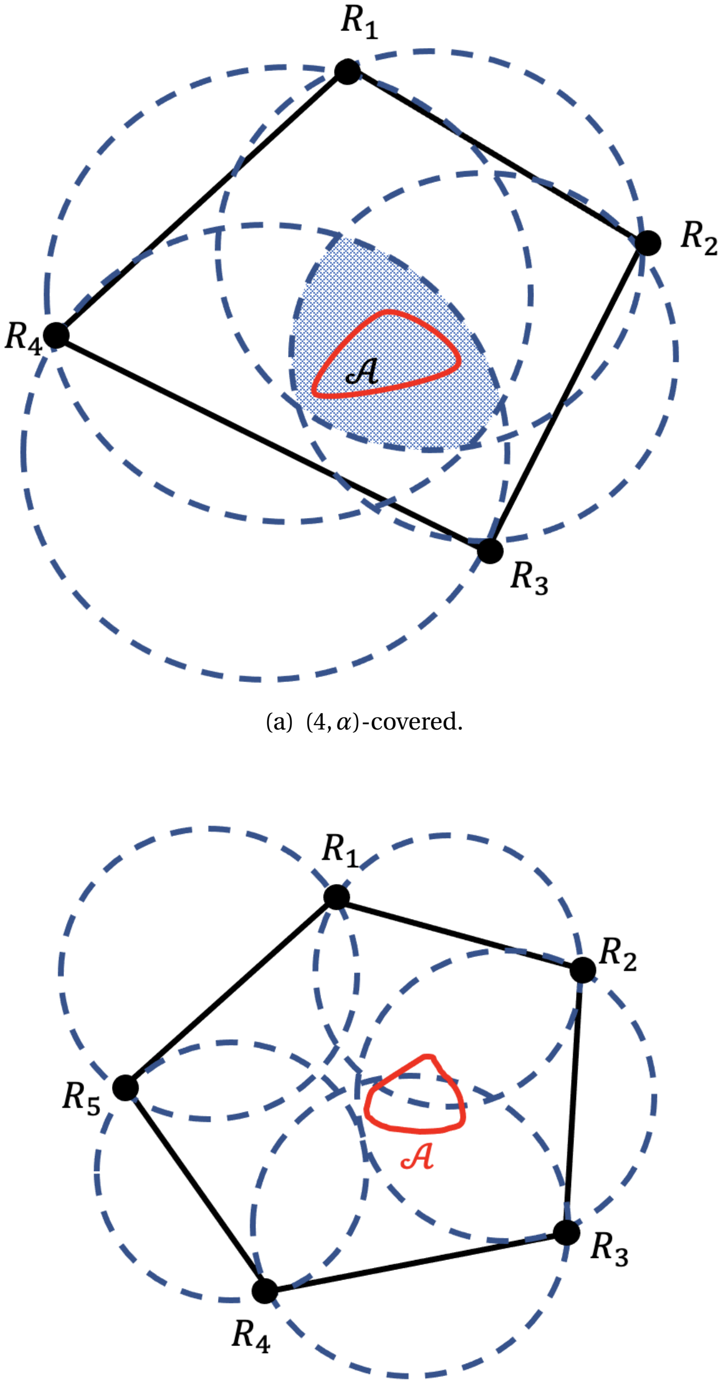
The Boundary Condition and (k, α)-coverage Verification Methods

Next, we shows that to (k, α)-cover an area is equivalent to (k, α)-cover the boundary of the area. And propose two (k, α)-coverage verification methods.

Lemma 3.3.3 (**Boundary condition**). Given an area $\mathscr{A}$ is k-covered by k different RIS. The area is (k, α)-covered by the k RISs if and only if its boundary $\partial\mathscr{A}$, is (k, α)-covered by the given set of RISs

Proof. When the region is (k, α)-covered, because the boundary of $\mathscr{A}$ is a part of the $\mathscr{A}$, it is obvious that $\partial\mathscr{A}$ is also (k, α)-covered by the given set of RISs. This proves the "only if" part.

Now we prove the "if" part: if the boundary of the region is (k, α)-covered by the given set of RISs, then the whole region is (k, α)-covered. We prove it by contradiction. Suppose all points on $\partial\mathscr{A}$ are (k, α)-covered, and there exists an interior point V in $\mathscr{A}$ that is not (k, α)-covered by the set of RISs. By definition, there exist at least two RISs, R_i and R_j, being adjacent in $CCL(\mathscr{A})$, such that $d(V, R_i) \leq r, d(V, R_j) \leq r$ and $\text{ANG}(VR_i, VR_j) < \alpha$. Now consider a point Q on $\partial\mathscr{A}$, and points Q, V and R_j is on the same line. As shown in Figure 3.4, Q is a boundary point and (k, α)-covered, i.e., $\text{ANG}(\overrightarrow{QR_i}, \overrightarrow{QR_{i'}}) \geq \alpha$ for $i \in [k]$ and $i' = \mod(i+1, k)$, indicating that $\alpha \leq \angle R_iQR_j \leq \angle R_iVR_j$. However, based on geometric relationships, we have $\angle R_iQR_j \leq$



(a) $(4, \alpha)$-covered.

(b) Not $(5, \alpha)$-covered.

Figure 3.3: Area (k, α)-coverage examples.

$\angle R_i V R_j < \alpha$, which is a contradiction. Hence we have proved that any interior point in $\mathscr{A}$ is (k, α)-covered if the boundary of $\mathscr{A}$ is (k, α)-covered. □

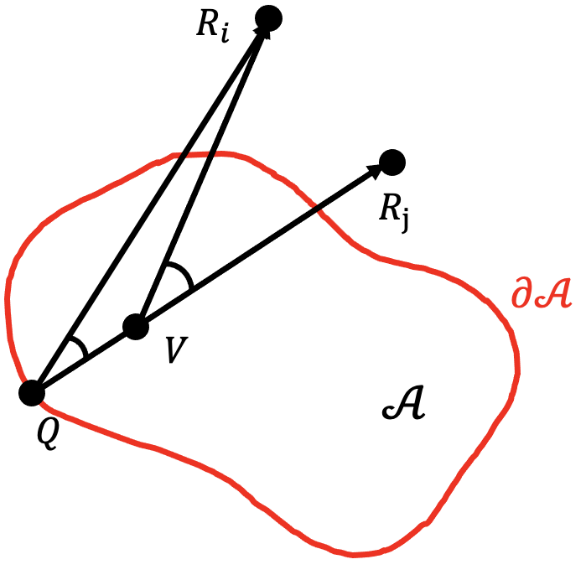


Figure 3.4: Boundary condition for region $\mathscr{A}$.

Theorem 3.3.1 (Boundary (k, α)-coverage). Given $\mathscr{S}\mathscr{A}$ with $CCL(\mathscr{S}\mathscr{A}) = \{R_1, R_2, \ldots, R_k\}$, $\mathscr{S}\mathscr{A}$ is (k, α)-covered if and only if its boundary $\partial\mathscr{S}\mathscr{A}$ is contained in the intersection of feasible regions of all pairs of R_i and R'_i, where $i \in [k]$, $i' = \mod (i+1, k)$ and $0 \leq \alpha \leq 2\pi/k$.

Proof. This is a result from Lemma 3.3.1, 3.3.2 and 3.3.3. □

By Theorem 3.3.1, we have proved that the area (k, α)-coverage is equivalent to the boundary (k, α)-coverage. Here, we further provide two methods to check whether a boundary is (k, α)-covered or not. A simple method is to select a collection of points on the boundary as control points such that the distance between any two adjacent control points is less than a constant δ. When δ is sufficiently small and all control points are (k, α)-covered, the area is also (k, α)-covered. Another method is to directly check if every segment of the sub-area boundary is (k, α)-covered. As the total number of boundaries generated is $\mathscr{O}(|\mathscr{N}|^4)$, where $|\mathscr{N}|$ is the total number of RISs deployed. We can check if every boundary segment is (k, α)-covered for the area (k, α)-coverage verification.

To tackle the MNkaAC problem under deterministic deployment scheme, we first approximate the continuous target region by discrete target points. According to Theorem 3.3.1, a region is (k, α)-covered if and only if its boundary is (k, α)-covered. Therefore, we can select a collection of points on the boundary as target points such that the distance between any two adjacent target points is less than δ. And if all target points are (k, α)-covered, the target region is also (k, α) covered. Approximating the continuous region by discrete points is a commonly used method to tackle the coverage problem in a 2-D region. And the MNkaAC is thereby transferred to the MNkaPC problem, which is defined as follows.

Definition 3.4.1 (**MNkaPC**)**.** Given a RISs network $\mathcal{N}$ and a set of target points, which is (k, α) covered by $\mathcal{N}$. The minimum-number (k, α) point coverage (MNkaPC) problem is to find a set C of RISs with the minimal cardinality that ensures (k, α) coverage for every target point.

Since the MNkaPC problem is NP-complete (proofed in Lemma 3.4.1), the computational complexity to solve the MNkaPC problem grows exponentially with the number of RISs. We will derive an approximation solution to the MNkaPC problem.

Proposition 3.4.1. The MNkaPC problem is NP-complete.

Proof. We prove its NP-completeness by showing a special case of MNkaPC problem to be NP-complete. When $\alpha = 0$, the MNkaPC problem can be reduced to a minimum-number k point coverage problem, which is to identify the minimal number of RISs such that every target point is within the transmission range of k different RISs. And this problem is NP-complete[56, 39]. □

Specifically, in this section, we use regular k-sided polygon (regular k-gon) deployment patterns to solve the MNkaPC problem approximately. A regular k-gon deployment pattern is formed by deploying k RISs at k vertices of a regular k-sided polygon, for example, equilateral triangle and square. In the following discussion, the deployment pattern refers to the regular

k-gon deployment pattern. The basic idea of using deployment patterns for the (k,α) area coverage is similar as tiling tiles. Specifically, each deployment pattern generates an (k,α)-coverage area, and by repetitively applying the deployment pattern to the target region, we can achieve the full area (k,α)-coverage. Among all the regular k-gon deployment patterns, we are more interested in the pattern that generates the maximum (k,α)-coverage area, which is referred as the optimal deployment pattern, because it reduces the number of RIS needed to the utmost extent.

In the next few subsections, we focus on the optimal deployment pattern and our analyses are twofold: **(1) The optimal pattern derivation:** We resolve the optimal deployment pattern by finding its side length and calculate the (k,α)-coverage area size generated by the optimal deployment pattern. Specifically, we first derive a general formula to calculate the (k,α)-coverage area for any regular k-gon deployment pattern. Then, we find the maximum side length for the deployment pattern under the constraint of the communication radius and use it to calculate the (k,α)-coverage area for the optimal pattern. **(2)The optimality bound**: We point out that applying the optimal pattern to (k,α)-cover the target region yields a good feasible solution to the MNkaPC problem. And we derive the optimality bound for this feasible solution in terms of the approximation ratio to quantify how good the solution is.

3.4.1 The Optimal k-sided Regular Polygon Pattern

The optimal pattern refers to the k-sided regular polygon pattern with the maximum (k,α)-coverage area. To optimize the pattern, we need to find the maximum side length of the k-sided regular polygon pattern, because a larger side length of the deployment pattern leads to larger (k,α)-coverage area. But the side length cannot be too large since the (k,α)-coverage area must be fully enclosed in the k-coverage area of the pattern (i.e., every point in the (k,α)-coverage area must be within the communication radius of all k RISs in the pattern). Therefore, the maximum side length also has complex relationship with the (k,α)-coverage area. In this section, we will derive the maximum side length of the k-sided regular polygon pattern considering the communication radius r, and use it to calculate the maximum (k,α)-coverage area of the pattern.

We begin by introducing the notations. Let $\mathscr{S}(k,\alpha,d_k)$ be the size of the (k,α)-coverage area generated by a regular k-sided polygon deployment pattern with side length d_k, which is a function of k,α and d_k. Denote the maximum value of d_k as d_k^*, which is the maximum feasible side length for the k-sided regular polygon deployment pattern. Then $\mathscr{S}(k,\alpha,d_k^*)$ is the correspondingly maximum (k,α)-coverage area. For the ease of the presentation, we define $\mathscr{S}^*_{(k,\alpha)} \overset{\text{def}}{=} \mathscr{S}(k,\alpha,d_k^*)$. Next, we discuss $\alpha < \pi/k$ and $\pi/k \le \alpha < 2\pi/k$ separately in finding d_k^* and $\mathscr{S}^*_{(k,\alpha)}$.

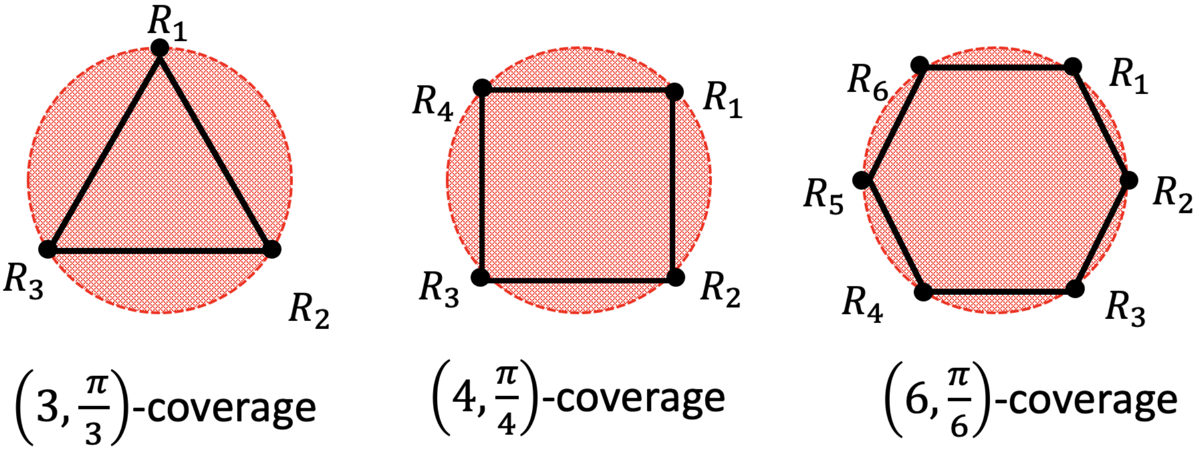
$\alpha < \pi/k.$

As proved in Theorem 3.4.1, any point in the circle that circumscribes a regular k-sided polygon is (k,α)-covered with $\alpha \le \pi/k$. Therefore, $\mathscr{S}^*_{(k,\alpha)}$ is the area size of the circle that circumscribes the regular k-sided polygon pattern. In this case, we can directly derive d_k^* and $\mathscr{S}^*_{(k,\alpha)}$. Since the diameter of the circle that circumscribed the regular k-sided polygon must be no greater than r to ensure all (k,α)-coverage area is k-covered, the maximum side length for the regular k-sided polygon deployment patterns $d_k^* = r\sin(\pi/k)$ and $\mathscr{S}^*_{(k,\alpha)} = 0.25\pi r^2$.

Theorem 3.4.1. For a regular k-sided polygon deployment pattern, the interior area of its circumcircle is $(k,\pi/k)$-covered.

Proof. For a regular polygon with k vertices v_1, v_2, ..., v_k (in clockwise), its circumcircle passes through all vertices of the polygon. For any two adjacent vertices v_i and $v_{i'}$ ($i' = \mod(i+1,k)$) and a point P on the circumcircle, the inscribed angle of chord $v_iv_{i'}$, i.e., $\angle v_iPv_{i'} = \pi/k$ or $\angle v_iPv_{i'} = \pi - \pi/k$. And for any point P' within $\odot O_k$, we have $\angle v_iP'v_{i'} \ge \pi/k$ for $i = [k]$. □

Figure 3.5 shows the $(k,\pi/k)$-coverage area for equilateral triangle, square and regular hexagon patterns, respectively. The interior area of the circumcircle satisfy $(3,\pi/3)$, $(4,\pi/4)$, $(6,\pi/6)$-coverage, respectively. Figure 3.6, from left to right, shows the maximum side-lengths for equilateral triangle, square and regular hexagon patterns, where r is the communication radius of RIS. For each of the three cases, the small red circle represents the (k,α)-coverage area and the big dotted circle is the communication circle of a RIS. The d_k^* is obtained by letting the radius of the big circle, i.e., r, equal to the diameter of the red circle.



R_1
R_2
R_3
$\left(3,\frac{\pi}{3}\right)$-coverage
R_1
R_2
R_3
R_4
$\left(4,\frac{\pi}{4}\right)$-coverage
R_1
R_2
R_3
R_4
R_5
R_6
$\left(6,\frac{\pi}{6}\right)$-coverage

Figure 3.5: The $(k,\pi/k)$-coverage area for $k = 3,4,6$.

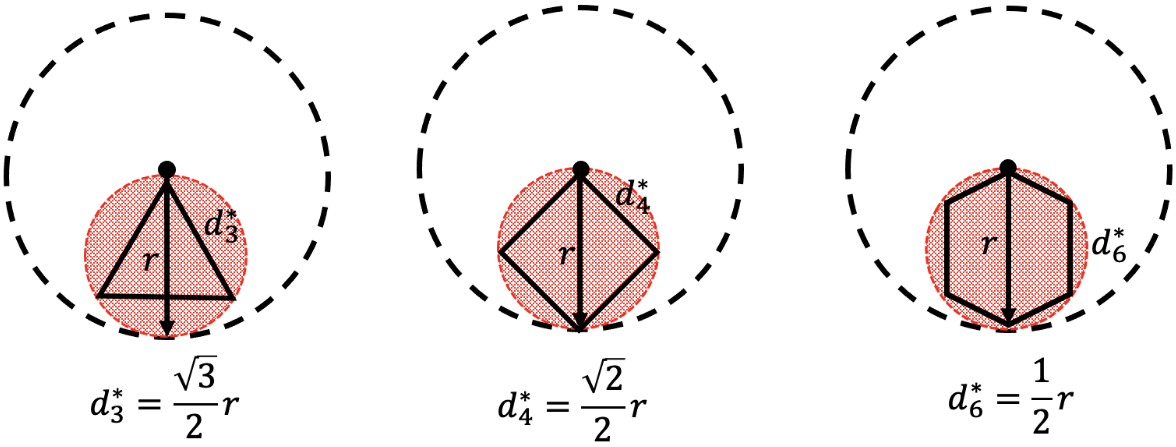


Figure 3.6: The maximum side-lengths for equilateral triangle, square and regular hexagon patterns w.r.t. the communication radius r.

$\pi/k \leq \alpha < 2\pi/k$.

For $\pi/k \leq \alpha < 2\pi/k$, we aim to derive an explicit expression for $\mathscr{S}(k,\alpha,d_k)$ to find the d_k^*, and then calculate $S^*_{(k,\alpha)}$. However, the derivation of $\mathscr{S}(k,\alpha,d_k)$ is difficult since finding the intersection area of all feasible regions is complicated. Recall that in Section 3.3, we have pointed out that the (k,α)-coverage area is the intersection area of k feasible regions. So, for a regular k-sided polygon pattern, we can also draw k feasible regions for every two adjacent RISs, and $\mathscr{S}(k,\alpha,d_k)$ is the intersection area of these k feasible regions, and each feasible region is either the union or overlap of two circumcircles. The $\mathscr{S}(k,\alpha,d_k)$ calculation thus involves finding the common intersection area of all $2k$ circles, which is a NP-complete problem.

Fortunately, we can decrease the problem difficulty by reducing the number of circles involved. According to Property 3.4.1, instead of calculating the intersection area of k feasible regions ($2k$ circles), we can calculate the intersection area of k feasible circles. Based on this observation, we use the feasible circle to calculate $\mathscr{S}(k,\alpha,d_k)$.

Property 3.4.1. For a regular polygon pattern, the intersection area of all feasible regions is only enclosed in one circumcircle of every feasible region, and we refer to this circle as the feasible circle.

In addition, in a regular k-sided polygon pattern, the feasible circles have the same radius and their centers also form a regular k-sided polygon. In this case, the intersection area of these k feasible circles, i.e., $\mathscr{S}(k,\alpha,d_k)$, is a curvilinear polygon with k identical curved sides. And the area size of this curvilinear polygon can be analytically derived.

In the following, we will use equilateral triangle deployment patterns and square patterns to illustrate the detailed calculation steps. For each of the patterns, we begin with calculating the radius of feasible circles and the side length of the polygon formed by centers of feasible circles. Then we calculate the area of the curvilinear polygon as $\mathscr{S}(k,\alpha,d_k)$.

Equilateral triangle pattern for $(3,\alpha)$**-coverage:** For an equilateral triangle deployment pattern $\triangle R_1R_2R_3$, the side length $d(R_1,R_2) = d(R_2,R_3) = d(R_3,R_1) = d_3$ and the internal angle is

$\pi/3$. As shown in the black triangle in Figure 3.7(a). To find the $(3,\alpha)$-coverage area, we draw the feasible circles for all adjacent RISs with centers as $O_{R_1R_2}$, $O_{R_2R_3}$ and $O_{R_3R_1}$, and the radii are $d_3/(2\sin\alpha)$. The $(3,\alpha)$-coverage area $\mathscr{S}(3,\alpha,d_3)$ is the overlapping area of three feasible circles $\odot O_{R_1R_2}$, $\odot O_{R_2R_3}$ and $\odot O_{R_3R_1}$, which is shown as the red dashed area in Figure 3.7(a).

By connecting $O_{R_1R_2}$, $O_{R_2R_3}$ and $O_{R_3R_1}$, the $\triangle O_{R_1R_2}O_{R_2R_3}O_{R_3R_1}$ is also an equilateral triangle. Denote its side length as d_3', as shown in Figure 3.7(b). We have

$$d_3' = \frac{d_3}{2\sin\alpha}\sqrt{2+2\cos(2\alpha+\pi/3)}. \tag{3.1}$$

Then, $S(3,\alpha,d_3)$ is the overlapping area of which the three feasible circles $\odot O_{R_1R_2}$, $\odot O_{R_2R_3}$ and $\odot O_{R_3R_1}$, each has the radius of $d_3/(2\sin\alpha)$, and the distance between any two centers of feasible circles is d_3'.

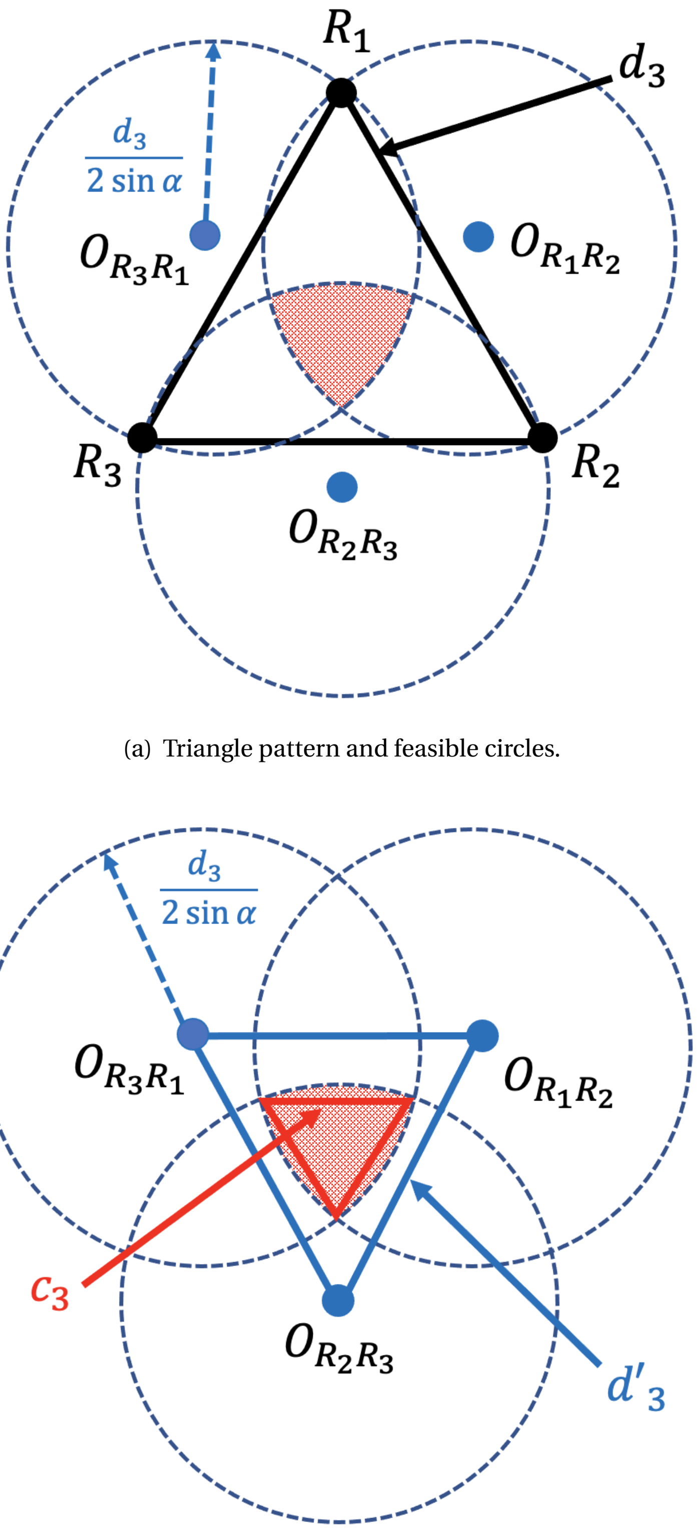
The $\mathscr{S}(3,\alpha,d_3)$ is a circular triangle, which can be decomposed into three circular segments and an equilateral triangle, and the chord length of the circular segment is equal to the side length of the equilateral triangle, as shown in Figure 3.7(b). Denote the chord length of the circular segment as c_3, we have

$$c_3 = \sqrt{3}\sqrt{\left(\frac{d_3}{2\sin\alpha}\right)^2 - \left(\frac{d_3'}{2}\right)^2} - \frac{d_3'}{2}. \tag{3.2}$$

And $S(3,\alpha,d_3)$ can thereby be calculated by summing the area of three circular segments with chord length c_3 and an equilateral triangle with side length c_3, which gives

$$\begin{aligned}&\mathscr{S}(3,\alpha,d_3) =\\ &\frac{\sqrt{3}}{4}c_3^2 + 3\Big((\frac{d_3}{2\sin\alpha})^2\arcsin\frac{c_3\sin\alpha}{d_3} - \frac{c_3}{4}\sqrt{(\frac{d_3}{\sin\alpha})^2 - c_3^2}\Big).\end{aligned} \tag{3.3}$$

After obtaining the shape and size of $\mathscr{S}(3,\alpha,d_3)$, we derive the maximum feasible side length d_3^* to obtain $\mathscr{S}^*_{(3,\alpha)}$. To ensure that the $(3,\alpha)$-coverage area is within the communication radii of R_1,R_2 and R_3, the largest distance between a RIS and a point in the $(3,\alpha)$-coverage



(a) Triangle pattern and feasible circles.

(b) Overlap area calculation.

Figure 3.7: The $(3, \alpha)$-coverage area in a triangle pattern.

area must be smaller than r, which gives

$$\frac{c_3}{2\sin(\pi/3)} + \frac{d_3}{2\sin(\pi/3)} \le r. \tag{3.4}$$

The maximum feasible side length for the equilateral triangle pattern, i.e., d_3^*, can be reached by taking the equality in Eq. (3.4), which is

$$d_3^* = \frac{4\sqrt{3}r\sin\alpha}{(6-6\cos(2\alpha+\pi/3))^{\frac{1}{2}} - (2+2\cos(2\alpha+\pi/3))^{\frac{1}{2}} + 4\sin\alpha}. \tag{3.5}$$

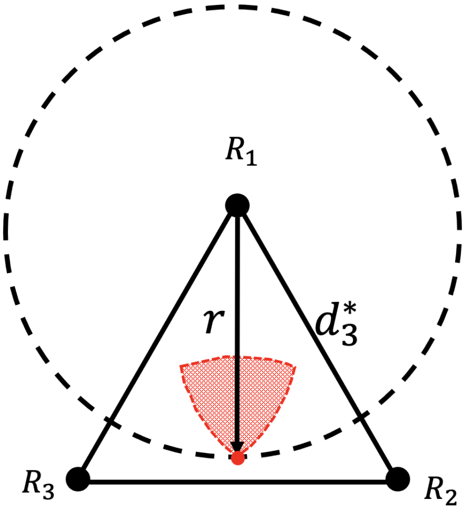


Figure 3.8: A d_3^* example.

Figure 3.8 shows the case where d_3^* is the side length of the equilateral triangle pattern $\triangle R_1R_2R_3$. In this case, the largest distance between R_1 and a point in the $(3,\alpha)$-coverage area (red dashed area) is r.

Square pattern for $(4,\alpha)$**-coverage:** For the square pattern $\square R_1R_2R_3R_4$, the analysis follows the similar procedure. Let d_4 be the side length of $\square R_1R_2R_3R_4$, we have $d(R_1,R_2) = d(R_2,R_3) = d(R_3,R_4) = d(R_4,R_1) = d_4$ and the internal angle of $\square R_1R_2R_3R_4$ as $\pi/2$. The feasible circles of adjacent RISs are $\odot O_{R_1R_2}$, $\odot O_{R_2R_3}$, $\odot O_{R_3R_4}$ and $\odot O_{R_3R_4}$, and the $(4,\alpha)$-coverage area is the overlapping area of the four feasible circles, as shown in the red dashed area in Figure 3.9(a).

The quadrilateral $O_{R_1R_2}O_{R_2R_3}O_{R_3R_4}O_{R_3R_4}$ is a square. Denote its side length as d_4', and we have

$$d_4' = \frac{d_4}{2\sin\alpha}\sqrt{2-2\sin(2\alpha)}. \tag{3.6}$$

And $\mathscr{S}(4,\alpha,d_4)$ is the overlapping area of four feasible circles $\odot O_{R_1R_2}$, $\odot O_{R_2R_3}$, $\odot O_{R_3R_4}$ and $\odot O_{R_3R_4}$, each with the radius of $d_4/(2\sin\alpha)$, and the distance between any two adjacent RISs is d_4', as shown in in Figure 3.9(b).

Similar to the equilateral triangle pattern analysis, $\mathscr{S}(4,\alpha,d_4)$ is the summation area of four circular segments and a square, and the chord length of each circular segment is equal to the side length of the square. Denote the chord length of the circular segment as c_4, as shown by the red dashed area in Figure 3.9(b), we have

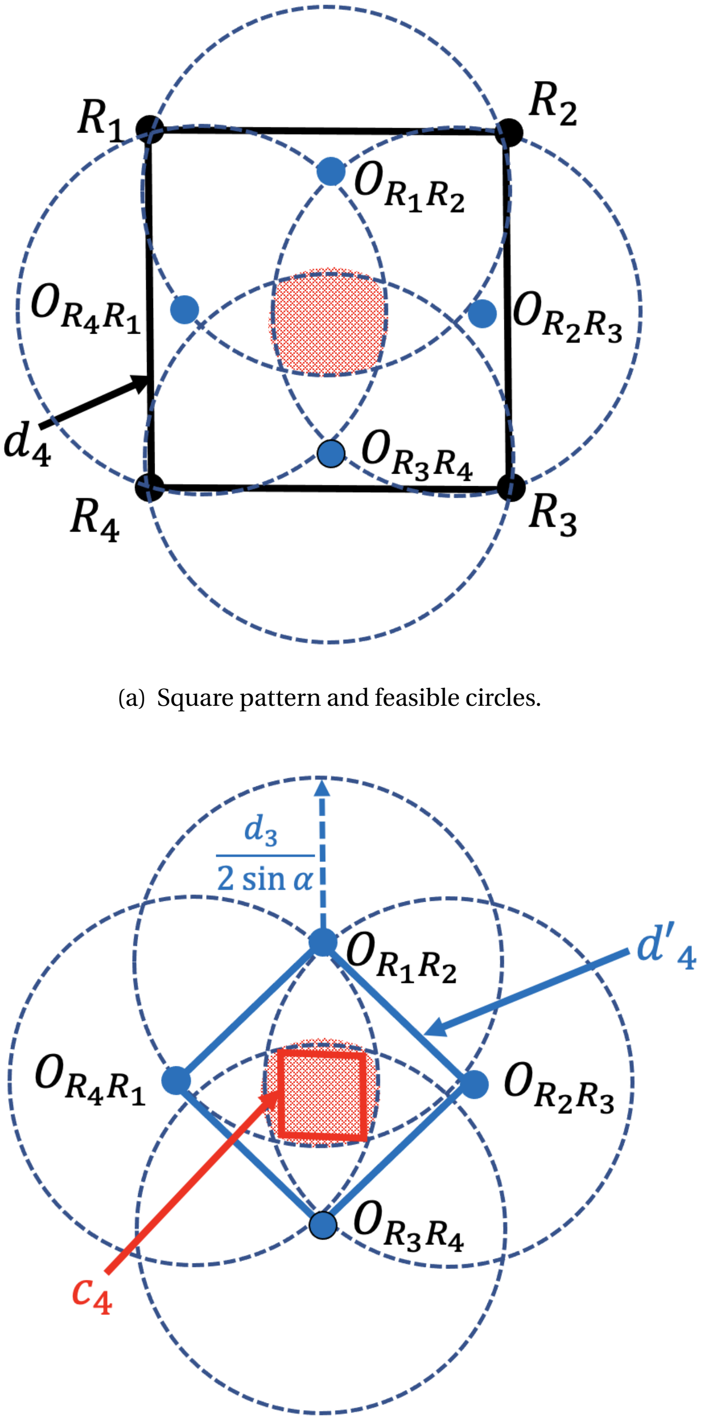
$$c_4 = \sqrt{2}\cdot\sqrt{(\frac{d_4}{2\sin\alpha})^2-(\frac{d_4'}{2})^2}-\frac{\sqrt{2}}{2}d_4'. \tag{3.7}$$

And $\mathscr{S}(4,\alpha,d_4)$ can be calculated by summing four circular segments with chord length c_4 and one square with side length c_4, which is

$$\begin{aligned}&\mathscr{S}(4,\alpha,d_4) = \\ &c_4^2+4\Big((\frac{d_4}{2\sin\alpha})^2\arcsin\frac{c_4\sin\alpha}{d_4}-\frac{c_4}{4}\sqrt{(\frac{d_4}{\sin\alpha})^2-c_4^2}\Big).\end{aligned} \tag{3.8}$$

And to calculate the maximum feasible side length for square deployment pattern, i.e., d_4^*, we have to ensure any point in the $(4,\alpha)$-coverage area is within the transmission radii of all four RISs. As shown in Figure 3.10, the red dashed area is the $(4,\alpha)$-coverage area of square pattern $\square R_1R_2R_3R_4$. To insure that any point in the $(4,\alpha)$-coverage area is within distance r to R_1, the largest distance between R_1 and a point in the red dashed area must be smaller than or equal to r, therefore

$$\frac{c_4}{2\sin(\pi/4)}+\frac{d_4}{2\sin(\pi/4)} \le r. \tag{3.9}$$



(a) Square pattern and feasible circles.

(b) Overlap area calculation.

Figure 3.9: The $(4, \alpha)$-coverage area in square pattern

By taking the equality, the maximum side length d_4^* for the square pattern $\square R_1 R_2 R_3 R_4$ is

$$d_4^* = \frac{2\sqrt{2} r \sin\alpha}{(1+\sin 2\alpha)^{\frac{1}{2}} - (1-\sin 2\alpha)^{\frac{1}{2}} + 2\sin\alpha}. \tag{3.10}$$

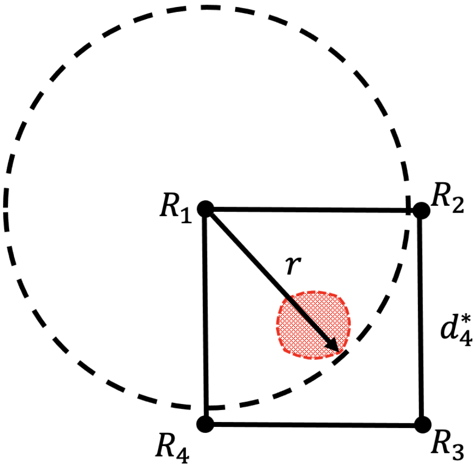


Figure 3.10: A d_4^* example.

$\mathscr{S}^*_{(4,\alpha)}$ is calculated by substituting Eq.(3.10) into Eq.(3.8).

General Formulations for d_k^* and $\mathscr{S}^*_{(k,\alpha)}$

Based on the above analyses, we can extend the results to any k-sided regular polygon deployment pattern for $\mathscr{S}^*_{(k,\alpha)}$ and d_k^* calculation.

For $0 \geq \alpha \geq \pi/k$, as proved in Theorem 3.4.1, $\mathscr{S}^*_{(k,\alpha)}$ is the area size of the circle that circumscribes the regular k-sided polygon pattern. Therefore, $d_k^* = r\sin(\pi/k)$ and $\mathscr{S}^*_{(k,\alpha)} = \pi r^2/4$.

As for $\pi/k \geq \alpha \geq 2\pi/k$. Let d'_k be the side length of k-sided regular polygon created by the centers of the k feasible circles, we have

$$d'_k = \frac{d_k}{2\sin\alpha}\sqrt{2+2\cos(2\alpha+(\pi-2\pi/k))}. \tag{3.11}$$

Then, $\mathscr{S}(k,\alpha,d_k)$ is composed of a regular k-sided polygon with side length c_k and k circular segments with chord length c_k, we have

$$c_k = 2\sin(\frac{\pi}{k})\Big(\sqrt{(\frac{d_k}{2\sin\alpha})^2 - (\frac{d'_k}{2})^2} - \frac{d'_k \cot\frac{\pi}{k}}{2}\Big), \tag{3.12}$$

And $\mathscr{S}(k,\alpha,d_k)$ is

$$\begin{aligned} \mathscr{S}(k,\alpha,d_k) &= \frac{k\sin\frac{2\pi}{k}}{4(1-\cos\frac{2\pi}{k})}c_k^2 \\ &+ k\Big((\frac{d_k}{2\sin\alpha})^2 \arcsin\frac{c_k\sin\alpha}{d_k} - \frac{c_k}{4}\sqrt{(\frac{d_k}{\sin\alpha})^2 - c_k^2}\Big). \end{aligned} \tag{3.13}$$

As d_k must be chosen to ensure the (k,α)-coverage area is also k-covered, we have

$$\frac{c_k}{2\sin(\pi/k)} + \frac{d_k}{2\sin(\pi/k)} \le r. \tag{3.14}$$

The maximum feasible side length of the k-sided regular polygon pattern, i.e., d_k^*, is

$$\begin{aligned} d_k^* &= \\ &\frac{2\sqrt{2}r\sin(\alpha)}{\big(1-\cos(\Phi(k,\alpha))\big)^{\frac{1}{2}} - \cot(\frac{\pi}{k})\big(1+\cos(\Phi(k,\alpha))\big)^{\frac{1}{2}} + \frac{\sqrt{2}\sin\alpha}{\sin(\pi/k)}}, \end{aligned} \tag{3.15}$$

where

$$\Phi(k,\alpha) = 2\alpha + \pi - 2\pi/k.$$

In summary, the general formulations for d_k^* and $\mathscr{S}^*_{(k,\alpha)}$ are

$$d_k^* = \begin{cases} r\sin(\frac{\pi}{k}) & 0 \le \alpha \le \pi/k \\ \text{Eq.(3.15)} & \pi/k < \alpha \le 2\pi/k \end{cases}, \tag{3.16}$$

and

$$\mathscr{S}^*_{(k,\alpha)} = \begin{cases} \frac{1}{4}\pi r^2 & 0 \le \alpha \le \pi/k \\ \text{Eq.(3.15) into Eq.(3.13)} & \pi/k < \alpha \le 2\pi/k \end{cases}. \tag{3.17}$$

3.4.2 Optimality Bound

After obtaining the optimal deployment pattern, we can apply it to (k,α)-cover the target region and derive an approximation solution to the MNkaPC problem. The solution can be obtained by using a **greedy algorithm**: *At each iteration, select the pattern that* (k,α)*-cover the most number of uncovered target points until there is no point left as un-*(k,α)*-covered.* Since the optimal deployment pattern has the largest (k,α)-coverage area, the approximation solution obtained by using the optimal pattern is a good feasible solution to the MNkaPC problem. Next, we will analyze the performance guarantee of the approximation solution in terms of the approximation ratio.

Let $C^{(k,\alpha)}$ be the set of RISs selected by applying the optimal deployment pattern to the MNkaPC problem. And let $OPT^{(k,\alpha)}$ be the optimal solution to the MNkaPC problem. The approximation ratio ρ is defined as

$$\rho = \frac{|C^{(k,\alpha)}|}{|OPT_{(k,\alpha)}|}, \tag{3.18}$$

which is the ratio between the number of RIS selected by the proposed method to the optimal solution, and it can be used to quantify how good the approximation solution is. A smaller ρ represents a better approximation.

To calculate ρ, we must know $|C^{(k,\alpha)}|$ and $|OPT_{(k,\alpha)}|$. Although we can apply the optimal pattern to the target region and numerically derive $|C^{(k,\alpha)}|$, $|OPT^{(k,\alpha)}|$ is hard to derive. Therefore, instead of calculating the exact value of ρ, we derive an upper-bound for it. We have the following theorem about ρ.

Theorem 3.4.2. Denote $C^{(k)}$ be the solution given in [153], which is an approximation solution to the problem of using the minimal of number RIS to k cover the target points, we have

$$\rho \leq \frac{2r\log(k|\mathcal{N}|)|C^{(k,\alpha)}|}{|C^{(k)}|}. \tag{3.19}$$

where r is the transmission radius of RIS and $|\mathcal{N}|$ is the size of the RIS network.

Proof. Consider the minimal-number k point coverage problem, which is to select the minimal number of RIS such that the every target point is within the communication radius of at least k different RISs. The minimal-number k point coverage problem is a special case of the MNkaPC problem where the angular constrain $\alpha = 0$ (no angular constrain). Therefore, the optimal solution to the minimal-number k point coverage problem, denoted as $OPT^{(k)}$, is a subset of the optimal solution to the MNkaPC problem $OPT^{(k,\alpha)}$, i.e., $OPT^{(k)} \subset OPT^{(k,\alpha)}$. And we have

$$\frac{|C^{(k,\alpha)}|}{|OPT^{(k,\alpha)}|} \leq \frac{|C^{(k,\alpha)}|}{|OPT^{(k)}|} \tag{3.20}$$

Since the minimal-number k point coverage problem is NP-complete, only approximation solution is derived for it. According to [153], the authors gave a solution to achieve the k-coverage with size at most $2r\log(k|\mathcal{N}|)|OPT^{(k)}|$, where r is transmission radius of a RIS and $|\mathcal{N}|$ is the size of the RIS network. Denote the approximate solution given in [153] as $C^{(k)}$, we have $|C^{(k)}| \leq 2r\log(k|\mathcal{N}|)|OPT^{(k)}|$. Therefore,

$$\rho \leq \frac{|C^{(k,\alpha)}|}{|OPT^{(k)}|} \leq \frac{2r\log(k|\mathcal{N}|)}{|C^{(k,\alpha)}|}|C^{(k)}| \tag{3.21}$$

□

Given a target region, both $|C^{(k,\alpha)}|$ and $|C^{(k)}|$ can be derived. And we can use them to calculate the optimality bound in Theorem 3.4.2 for a worst-case performance guarantee.

Remark: In addition, we can further derive another optimality bound for ρ since we have derived the (k,α)-coverage area size of each optimal pattern, i.e., $S^*_{(k,\alpha)}$. Denote the area size of the target region as S_0, and assume it is far greater than the (k,α)-coverage area generated by an optimal pattern, i.e., $S_0 \gg S^*_{(k,\alpha)}$, which is a common scenario in most real-world cases. we have

$$|C^{(k,\alpha)}| \le \frac{kS_0}{S^*_{(k,\alpha)}}. \tag{3.22}$$

And the optimality bound in Theorem 3.4.2 can be further derived as

$$\rho \le \frac{2krS_0 \log(k|\mathcal{N}|)}{S^*_{(k,\alpha)}|C^{(k)}|}. \tag{3.23}$$

3.5 (k,α)-Coverage Probability under Random Deployment Scheme

In this section, we calculate the probability that an arbitrary point in the target region is (k,α)-covered given a number of RISs are randomly deployed in the region. Due to its convenience, random deployment scheme is widely used to deploy RISs in a large complex environment. The calculation we derived can be used to estimate the least number of RISs that are needed to achieve (k,α)-coverage with a required coverage probability.

Without loss of generality, we assume the target region is a 1×1 squared region. And the coverage area of an RIS, i.e., πr^2, is less than 1. We first consider the case where there are k uniformly deployed RISs to find the probability that an arbitrary point is (k,α)-covered by these k RISs. Then we consider the number of RISs to be $N(N > k)$ to derive the (k,α)-coverage probability. Finally, we consider the case where the RISs are distributed following a homogeneous spatial Poisson distribution with density λ for more general applications.

When RISs are deployed in a bounded region, boundary effect takes place in which the number of RISs in the area that is close to the boundary is likely to be less than that in the interior area. As a result, the area close to the boundary is less likely to satisfy the coverage condition than the interior area. However, the boundary effect can be trivially avoided by deploying the RISs in a larger deployment region, which can be achieved by slightly increasing the side length of the original target region. When the target region is large enough, the difference is negligible. Therefore, the boundary effect is ignored in our analysis.

Lemma 3.5.1. Given k uniformly distributed RISs in area $\mathscr{A}$, the probability that an arbitrary point P is (k,α) (α in rad and $\alpha \le 2\pi/k$) covered by the set of RISs is

$$P_{(k,\alpha)}^{(k)} = (\pi r^2)^k (\frac{2\pi - k\alpha}{2\pi})^{k-1}. \tag{3.24}$$

Proof. The RISs are uniformly distributed, therefore the probability that all k RISs are within distance r to the point P is $p = (\pi r^2)^k$.

Next, we consider the probability that k RISs can (k,α) cover P. Consider a disk area with radius r and centered at P. Since RISs are uniformly distributed in $\mathscr{A}$, they are also uniformly distributed in the disk area. For each RIS $R_i, i \in [k]$ within the disk area, consider its projection point R_i' on the perimeter of the disk, which is the intersection point of vector $\overrightarrow{PR_i}$ and the perimeter of the disk. R_i' is also uniformly distributed on the perimeter of the circle, and all R_i's divide the perimeter of the circle into k non-overlapping arc segments. Therefore, the probability that P is (k,α)-covered is equal to the probability that the length of any k arc segments is no less than $r\alpha$. This probability can be calculated through geometric probability[109]. We use $k=3$ as an example, and the results of other k values can be derived using similar approach. Suppose $r=1$, and the length of the j-th arc segment is denoted as x_j. So, we have $x_j > 0$ for any $j \in [1,2,3]$ and $\sum_{j=1}^{3} x_j = 2\pi$. Draw the constraints on a Cartesian coordinate system with the x-axis as x_2 and y-axis as x_1, as shown in Figure 3.11. The area formed by these constrains, which is denoted as S_0, is $\frac{(2\pi)^2}{2}$, as shown in the black dashed area in Figure 3.11. And if the lengths of all 3 arc segments are no less than α, we have $x_j > \alpha$ for $j \in [1,2,3]$ and $\sum_{j=1}^{3} x_j = 2\pi$. The area, which is denoted as $S_{(3,\alpha)}$, is $\frac{(2\pi-3\alpha)^2}{2}$ (The red dashed area in Figure 3.11). Therefore, the probability Pr(all 3 arc segments are no less than α) $= \frac{S_{(3,\alpha)}}{S_0} = (1-\frac{3\alpha}{2\pi})^2$.

Following this idea, the probability that the k RISs can (k,α)-cover P can thus be derived, which is $(1-\frac{k\alpha}{2\pi})^{k-1}$. The lemma is proved. □

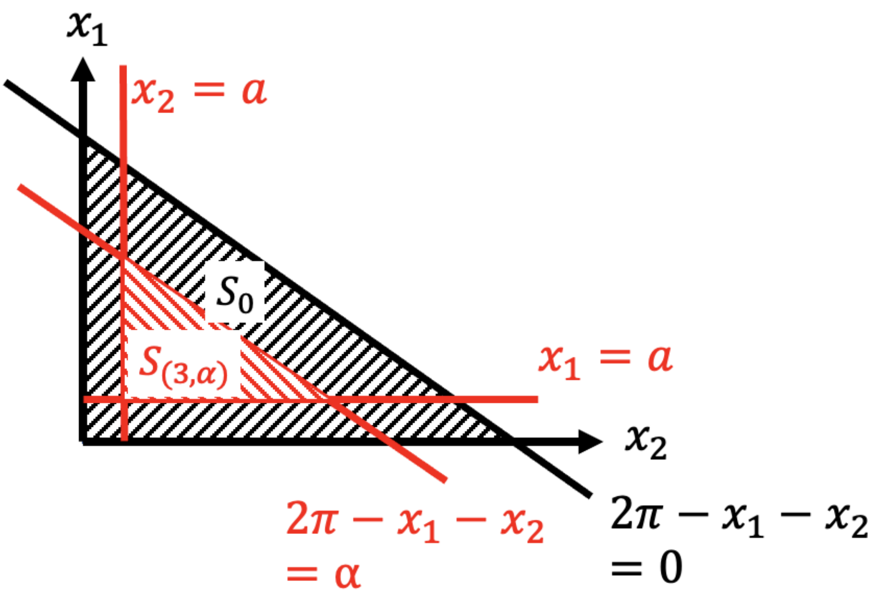


Figure 3.11: The (k, α)-coverage probability example for $k = 3$.

Theorem 3.5.1. Given N $(N > k)$ RISs uniformly distributed in area $\mathscr{A}$, the probability that an arbitrary point P is (k, α) (α in rad and $\alpha \le 2\pi/k$) covered by the set of RISs is

$$P^{(N)}_{(k,\alpha)} = \sum_{k'=k}^{N} \binom{N}{k'} (\pi r^2)^{k'} (1 - \pi r^2)^{N-k'} f(k, \alpha, k'), \tag{3.25}$$

where

$$f(k, \alpha, k') = \left(1 - \left(1 - (\frac{2\pi - k\alpha}{2\pi})^{k-1}\right)^{\binom{k'}{k}}\right). \tag{3.26}$$

Proof. Given N RISs in $\mathscr{A}$, the probability that point P is covered by exactly $k'(k' \le N)$ RISs is

$$p = \binom{N}{k'} (\pi r^2)^{k'} (1 - \pi r^2)^{N-k'}. \tag{3.27}$$

Given point P covered by k' $(k' \ge k)$ RISs, we need to calculate the probability that point P is (k, α)-covered. Denote the probability as $f(k, \alpha, k')$, and let $\tilde{f}(k, \alpha, k')$ be the probability that P is not (k, α)-covered by k' RISs, we have

$$f(k, \alpha, k') = 1 - \tilde{f}(k, \alpha, k'). \tag{3.28}$$

By Definition 3.3.2, $f(k,\alpha,k')$ is the probability that given point P covered by k' RISs, there exists at least one $CCL(P)$ with k elements that satisfies the α angular constraint. Therefore, $\tilde{f}(k,\alpha,k')$ is the probability that given point P covered by k' RISs, all $CCL(P)$ of P with k elements do not satisfy the α angular constraint. Given k' RISs, the probability that P is not (k,α)-covered is

$$\tilde{f}(k,\alpha,k') = (1-(\frac{2\pi-k\alpha}{2\pi})^{k-1})^{\binom{k'}{k}}. \tag{3.29}$$

(k,α)-coverage requires at least k RISs, so k' ranges from k to N. By summing k' from k to N, the lemma is proved. □

Theorem 3.5.2. Given that RISs are deployed following a homogeneous spatial Poisson distribution with density λ in region $\mathscr{A}$, the probability that an arbitrary point P is (k,α) (α in rad and $\alpha \le 2\pi/k$) covered by the set of RISs is

$$P_{(k,\alpha)} = \sum_{k'=k}^{\infty} \frac{(\lambda\pi r^2)^{k'} e^{-\lambda\pi r^2}}{k'!} f(k,\alpha,k'), \tag{3.30}$$

where $f(k,\alpha,k')$ is given in Theorem 3.5.1.

Proof. This is proved by putting together the homogeneous Poisson probability and the probability in Theorem 3.5.1. □

3.6 Numerical and Simulation Results

In this section, we provide simulation results for both random and deterministic schemes for (k,α)-coverage. In deterministic deployment scheme, we first show the d_k^* and $\mathscr{S}^*_{(k,\alpha)}$ to illustrate the performance of using k-sided regular polygon patterns to achieve the (k,α)-coverage. Then, we compare the proposed deterministic schemes with a modified k-coverage algorithm to highlight the performance improvement of our method. In the random deployment scheme, we first validate the probability derivation in Section 3.5 and get a pictorial view of the relationship between the RIS density and the (k,α)-coverage probability under different settings.

3.6.1 Simulation Setup

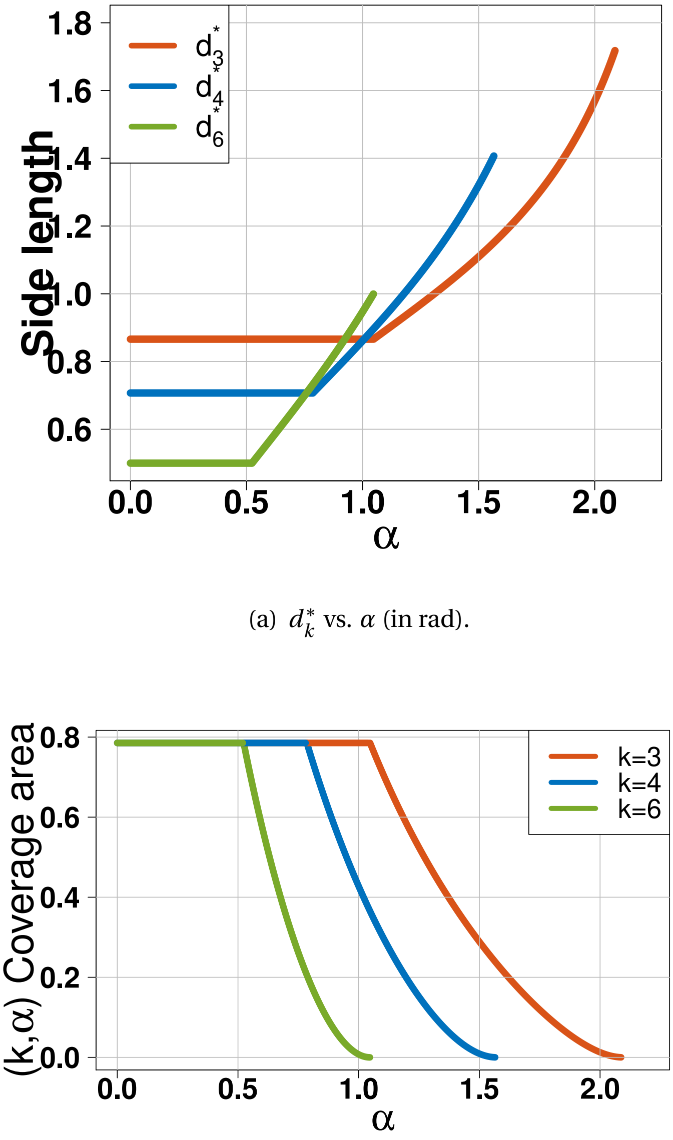
To normalize the results, the area size of the target region $\mathscr{A}$ is set to 1 (unit area). The experiments are done by MatLab for both the random and deterministic deployment simulations. The detailed simulation setup is described along with the figure explanations.

3.6.2 Simulation Results

Deterministic Deployment Scheme.

Theoretical results: In the deterministic deployment scheme, we first show the theoretical results derived from section 3.4. Specifically, we show the maximum (k,α)-coverage area $\mathscr{S}^*_{(k,\alpha)}$ and the maximum feasible side length d^*_k under different α. The communication radius r is set to 1 (unit length) in the simulation.

Figure 3.12(a) shows the maximum feasible side length d^*_k for the k-sided regular polygon deployment pattern vs. α. We show d^*_k for k = 3, 4, 6. For $\alpha \leq \pi/k$, the interior area of the k-sided regular polygon's circumcircle is $(k,\pi/k)$-covered. In this case, d^*_k does not vary with the α, and remains as $r\sin\pi/k$. For $\pi/k \geq \alpha \geq 2\pi/k$, d^*_k increases as α increases. The reason is that as α increases, the k-sided regular polygon pattern must also increase to ensure that every (k,α)-covered point is k-covered. Figure 3.12(b) shows the maximum (k,α)-coverage area $\mathscr{S}^*_{(k,\alpha)}$ for the k-sided regular polygon deployment pattern with maximum feasible side length d^*_k for different α. Here, we show the results for k = 3, 4, 6. For $\alpha \leq \pi/k$, the interior area of the circumcircle is $(k,\pi/k)$-covered, therefore, $\mathscr{S}^*_{(k,\alpha)} = (0.5r)^2\pi = 0.25\pi$. For $\pi/k \geq \alpha \geq 2\pi/k$, the $\mathscr{S}^*_{(k,\alpha)}$ is calculated through Eq.(3.13). Although d^*_k increases as α increases, $\mathscr{S}^*_{(k,\alpha)}$ decreases as α increases. The reason is that the angular constrain α is the most strict requirement in the (k,α)-coverage model, as α increase, the (k,α)-coverage area generated by the deployment drops fast. When $\alpha = 2\pi/k$, $\mathscr{S}^*_{(k,\alpha)} = 0$, because only one point satisfy the $(k,2\pi/k)$-coverage requirement, and the area size of a point is 0.



(a) d_k^* vs. α (in rad).

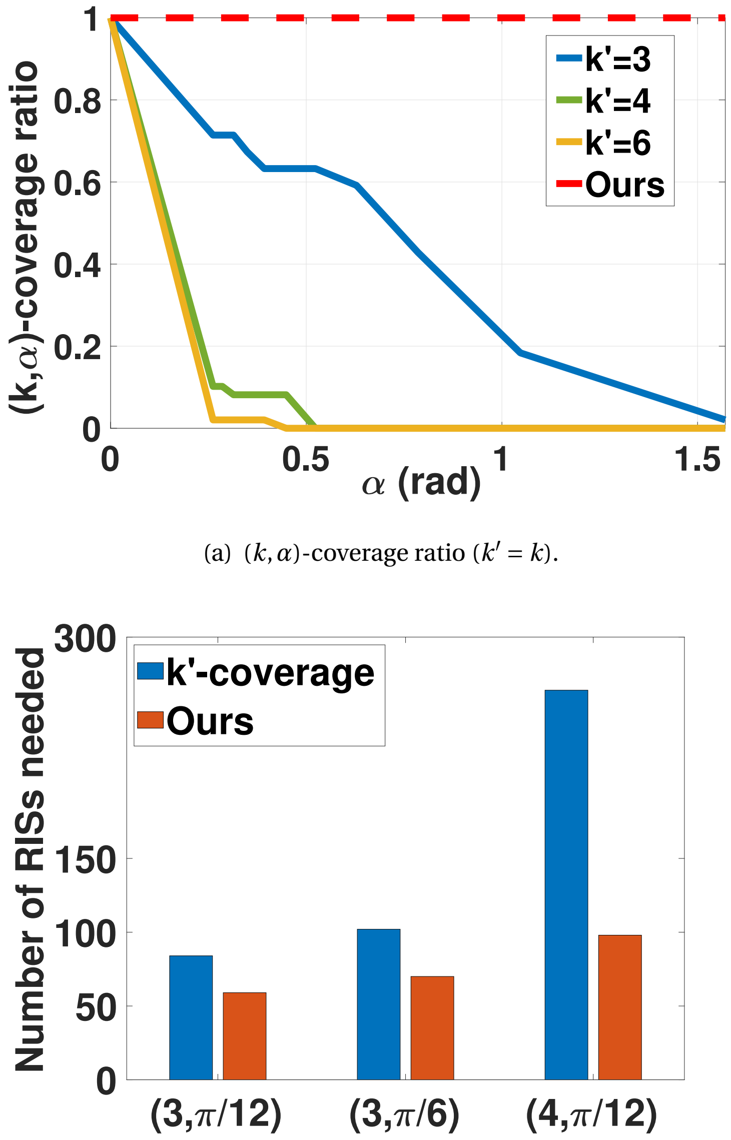
(b) $\mathscr{S}^*_{(k,\alpha)}$ vs. α (in rad).

Figure 3.12: The maximum feasible side length d_k^* and the maximum (k,α)-coverage area $\mathscr{S}^*_{(k,\alpha)}$ generated by a pattern.

Performance comparison. We then compare our deterministic deployment pattern-based scheme with a modified k-coverage algorithm to highlight our contribution. We must emphasize that, to the best of our knowledge, our paper is the first work in the literature that studies the (k,α)-coverage problem. In the absence of any existing scheme that considers similar constraints to those in the (k,α)-coverage problem, the best we can do in order to obtain a meaningful comparison is to modify an existing algorithm that was originally developed to solve a different coverage problem (e.g., the k-coverage problem in our simulations) in such a way that the modified algorithm can give a feasible solution to the (k,α)-coverage problem.

More specifically, we consider the minimum-number (k,α) point coverage problem (MNkaPC), which finds the minimum number of RISs that provides (k,α)-coverage for all points in a target area. To demonstrate the performance gain of our proposed method, we compare the solution offered by our proposed deterministic deployment scheme with that by a well-known greedy algorithm that solves the classical k-coverage problem [153]. Note that the greedy algorithm by itself only yields a quasi-optimal solution to the k-coverage problem, so we have to modify it as follows to obtain a feasible solution to the MNkaPC problem. To avoid any ambiguity in the following presentation, we replace the letter k in the k-coverage by k'. The modified greedy algorithm works as follows: Given a desired (k,α)-coverage requirement, we start by setting $k' = k$, and test whether the solution offered by the greedy algorithm constitutes a feasible solution to the (k,α)-coverage problem for every point in the target area. If it is not a feasible solution, we increase k' by 1 and repeat the above process. The iteration will continue until the smallest k' for which the greedy algorithm offers a feasible solution to the (k,α)-coverage problem. This feasible solution can then be compared with the solution offered by our proposed algorithm.

Our simulation is performed as follows. We consider a target area of 300×300 m^2, which is fully confined to an RIS deployment area of 500×500 m^2. We partition the target area and the RIS deployment area into grids of 50×50 m^2 and 1×1 m^2, respectively. Each cross-point of grids is denoted as a target point or a RIS candidate position. There are 49 target points and 251,001 RIS candidate positions in total. The communication radius of an RIS is set to 75 m. We compare our method with the k'-coverage method by two metrics: (1) the (k,α)-coverage ratio, which is the percentage of target points that are (k,α)-covered under a given RIS deployment scheme; (2) the number of RISs needed to achieve (k,α)-coverage for all target points.



(a) (k,α)-coverage ratio ($k' = k$).

(b) Number of RIS needed.

Figure 3.13: Performance comparison results.

Figure 3.13(a) plots the (k,α)-coverage ratio against the angular separation parameter α under various RIS deployment schemes. Our deterministic deployment scheme (marked by the red dotted line) maintains a ratio of 1, as it guarantees that all target points are (k,α)-covered. For the k'-coverage, we consider $k' = k$, where $k = 3,4,6$, respectively. As α or k' increases, the (k,α)-coverage ratio achieved by the k'-coverage method decreases rapidly. An extreme case is when $k' = 6$, the 6-coverage method hardly provides $(6,0.5)$-coverage for any target points. Another observation is that the k'-coverage method can only achieve an acceptable (k,α)-coverage ratio when α is small. For example, for the special case of $\alpha = 0$, the coverage ratio achieved by the modified greedy algorithm conforms with our proposed method (i.e., 100%). In case of blockage, multiple RIS generated paths with small angular separation are more likely to be blocked at the same time. In contrast, our method can maintain a high (k,α)-coverage ratio for large α, which is more robust than k'-coverage when multiple path blockages occur.

Figure 3.13(b) shows the number of RISs needed for different (k,α)-coverage scenarios by our method and the modified k'-coverage method. Specifically, we consider $(3,\pi/12)$-coverage, $(3,\pi/6)$-coverage, and $(4,\pi/12)$-coverage. As mentioned above, k' could be greater than k to achieve the (k,α)-coverage. For example, we find that the smallest k' to achieve $(3,\pi/12)$-coverage and $(4,\pi/12)$-coverage by the k'-coverage method is 10 and 30, respectively. This explains why the number of RIS needed by the k'-coverage method is significantly higher than that of our method. For example, it requires 264 RISs for k'-coverage to achieve $(4,\pi/12)$-coverage, while it only needs 98 RISs by our method, a reduction of 62.8%.

In conclusion, compared to k'-coverage, our method is more robust against multiple path blockages and requires significantly fewer RISs to be deployed.

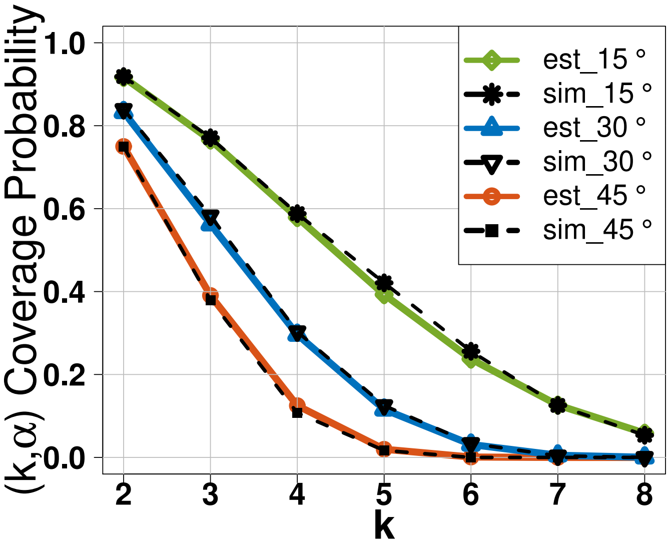
Random Deployment Scheme.

In the random deployment case, we first use Monte Carlo simulation to verify the correctness of our probability derivation in Section 3.5. Then we present the theoretical results for the probability function to better illustrate the impact of different parameters in the (k,α)-coverage model.

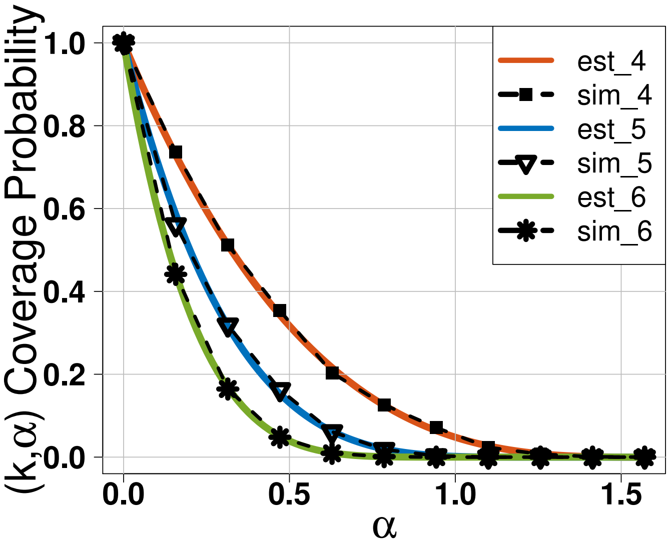
Result verification. In this simulation, each experiment is run 10000 times and the results are averaged. For comparison, we also give the theoretical estimation for each configuration. To decrease the complexity of the simulation and highlight the (k,α)-coverage probability, we omit the k-coverage requirement, i.e., the probability πr^2 is set to 1. Since the k-coverage problem is well studied and the probability of k-coverage can be simply calculated by using Bernoulli distribution function.

We first validate the probability result derived from Lemma 3.5.1. Figure 3.14 shows the the simulation and estimation result for (k,α)-coverage probability. In the simulation, we uniformly random place k number of RISs in the target region 10000 times and calculate the ratio of points that satisfy the required (k,α)-coverage, and the results are drawn by dotted black lines. The estimation results are calculated through Eq. 3.24 and are marked as colored solid lines. In Figure 3.14(a), the x-axis represents the number of RISs k in the (k,α)-coverage model, and the y-axis is the (k,α)-coverage probability. We vary α from $15°, 30°, 45°$ to test different scenarios. In Figure 3.14(b), the x-axis represents the angular separation parameter α in the (k,α)-coverage model and the y-axis is the (k,α)-coverage probability. The k is chosen from $4,5,6$. From both figures, the estimation based on our derivation accord with the Monte Carlo simulation results, which validate Lemma 3.5.1.

Next, we validate the probability result derived from Theorem 3.5.1 Eq. 3.26. Note that if the derivation in Eq. 3.26 is correct, it will also validate the correctness of Theorem 3.5.1. The result in Eq. 3.26 stands for the probability that a point can be (k,α) covered when given $k'(k' \geq k)$ RISs available for transmission. Figure 3.15 shows the simulation and estimation results. For the simulation, we uniform randomly place k' number of RISs in the target region 10000 times and calculate the ratio of points that satisfy the required (k,α)-coverage, and the results are drawn by dotted black lines. The estimation results are calculated through Eq. 3.26 in Theorem 3.5.1 and are marked as colored solid lines. To cover different scenarios, we simulate the $(3,90°)$-coverage (red line), $(4,60°)$-coverage (blue line) and $(5,50°)$-coverage (green line) cases. The estimation results accord with the simulation results, which validate Theorem 3.5.1.



(a) (k, α) Probability vs. k.



(b) (k, α) Probability vs. α (in rad).

Figure 3.14: The simulation and estimation results for Lemma 3.5.1.

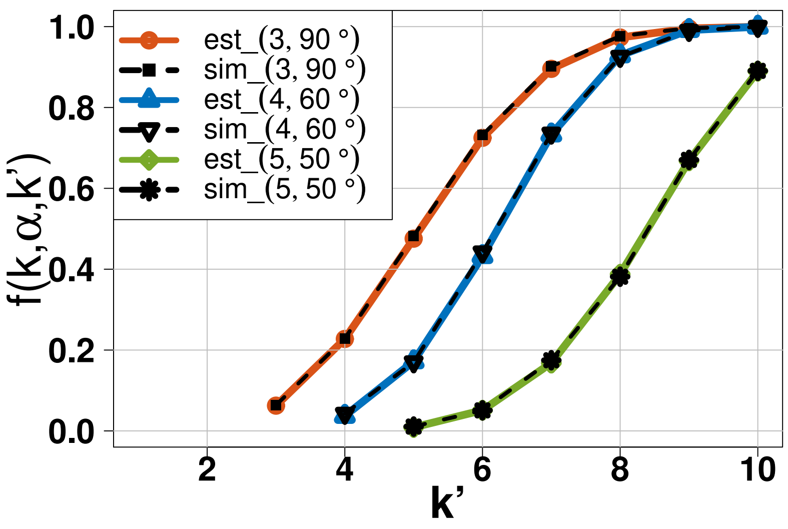
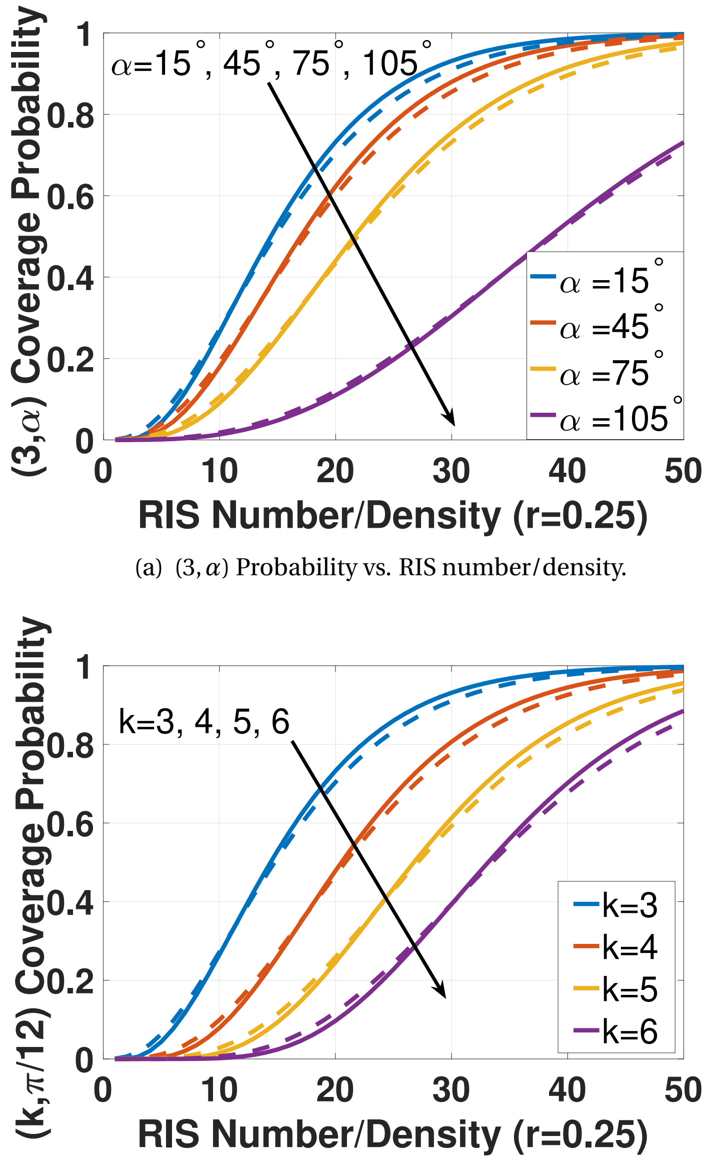


Figure 3.15: $f(k,\alpha,k')$ vs. k'.

Theoretical result. To better understand the impact of k and α on (k,α)-coverage probability under random deployment scheme, we present the theoretical results for uniform random deployment scheme derived from Theorem 3.5.1 (the solid lines) and Poisson distribution based deploy scheme derived from Theorem 3.5.2 (the dashed lines) under different (k,α) settings.

Figure 3.16 shows the results of the (k,α)-coverage probability vs. RIS number/density under different k and α. The y-axis represents the (k,α)-coverage probability and x-axis represents both the RIS number and the density (λ) of Poisson point distribution. Note that the area size is 1 in our simulation, the density of Poisson point distribution (λ) and the number of RISs deployed are the same in quantity. The transmission radius is set to 0.25. In Figure 3.16(a), k is fixed to 3 and $\alpha = \pi/12, \pi/4, 5\pi/12, 7\pi/12$ (or $15^\circ, 45^\circ, 75^\circ, 105^\circ$ in degree). Both exact number deployment scheme and Poisson deployment scheme show the same trend as the number/density increases. The (k,α)-coverage is hard to satisfy in the random deployment scheme and when α becomes larger, the RIS number/density increases rapidly. In Figure 3.16(b), α is fixed as $\pi/12$ and $k = 3,4,5,6$. As k increases, the RISs needed for (k,α)-coverage also increases.

However, compared to Figure 3.16(a), the density increment caused by k is much smaller, as the angular separation requirement is harder to satisfy in random deployment scheme.



(a) $(3, \alpha)$ Probability vs. RIS number/density.

(b) $(k, \pi/12)$ Probability vs. RIS number/density.

Figure 3.16: The (k, α)-coverage probability result for uniform distribution (solid lines) and Poisson distribution (dashed lines).

Figure 3.17 shows the results for the $(3, \alpha)$-coverage probability vs. RIS number/density under different r and α. We set the transmission radius of a RIS r as 0.25 and 0.15, and we vary α from $15°, 45°, 75°$ to highlight the difference. As r decreases, the coverage area of a RIS decreases, leading to the RIS density needed to achieve the 100% coverage probability increases dramatically. Also, although the number of RISs needed to achieve 100% coverage probability is very high, the density needed to achieve high coverage probability is much lower. For example, when the RIS density is 23 and $r = 0.25$, $(3, 15°)$-coverage probability reaches about 80%, while it takes another 25 RISs density to increase the $(3, 15°)$-coverage probability from 80% to 100%.

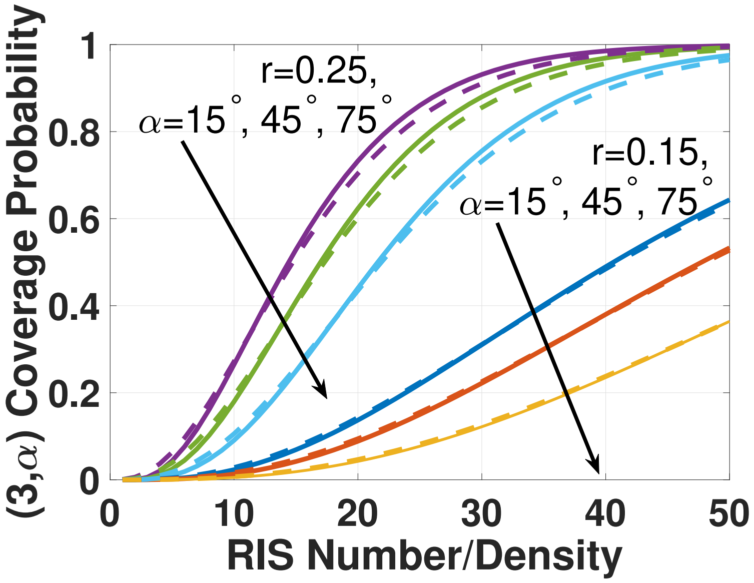


Figure 3.17: $(3, \alpha)$-coverage probability ($r = 0.15$ and $r = 0.25$) for uniform distribution (solid lines) and Poisson distribution (dashed lines).

3.7 Conclusion

In this paper, we have studied the minimal-number (k, α) area coverage problem for RIS-aided mmWave directional communication network. To overcome the limitations in the traditional coverage models, we introduced a novel coverage model called (k, α)-coverage, which requires a receiver to be covered by k different RISs to provide k different NLoS path directions while the angular separation difference between adjacent path directions is no less than α. Two methods were proposed to detect if the target area is (k, α)-covered by the given set of RISs. For the deterministic deployment scheme, we derived a quasi-optimal solution for the (k, α)-coverage problem using deployment patterns. An analytical optimality bound was also obtained for this solution. For the random RIS deployment scheme, we derive the (k, α)-coverage probability under uniform and spatial-Poisson RIS distributions. In the simulation, we first compare our deterministic deployment pattern based scheme with a modified k-coverage algorithm to highlight our contribution. And the results show that our method is more robust against multiple path blockages and requires significantly fewer RISs than the k-coverage method. Then, we conducted extensive simulations to verify the correctness of our theoretical derivation.

Chapter 4

Spoofing Detection for LiDAR in Autonomous Vehicles: A Physical-layer Approach

4.1 Introduction

In recent years, the development of autonomous vehicles (AVs), i.e., vehicles that can drive by themselves without the real-time intervention of human drivers, is rapidly progressing with the advancement of sensing and artificial intelligence (AI) technologies [27, 121]. Some AVs are already operating on public roads, e.g., Google's Waymo One self-driving taxis [124]. For all these AVs, *driving safety* is always the No.1 requirement. To this end, all existing AVs are equipped with certain types of environment-perception sensors, such as cameras, mmWave radar, ultrasonic sensors, and light detection and ranging sensor (LiDAR). With the rise of the Internet of Things (IoT), AVs can connect to other devices and systems, such as traffic lights and road sensors, to collect real-time data and make more informed decisions. This can enhance the accuracy of the AVs' sensing and decision-making abilities, leading to improved safety and efficiency. Additionally, IoT-enabled AVs can communicate with each other, allowing them to coordinate their movements and further improve safety on the road.

Among the various environment-perception sensors used by AVs, LiDAR sensor is adopted by almost all AV manufacturers due to its high precision and high reliability [59, 17, 75]. The LiDAR sensor employs highly directional laser pulses to probe the surrounding environment. An accurate depth image of the surrounding objects is then collected by the time of flight (ToF) of the received pulse, on which a high-resolution 3D point cloud map of the environment can be built. In addition, the usage of an infrared laser signal not only makes LiDAR less affected by ambient light in the environment, but also enables LiDAR to remain functional even under poor light conditions.

Ensuring correct and truthful sensing outcome from all environment-perception sensors is essential to ensure reliable safety-critical decision making in autonomous driving. Unfortunately, recent studies have found that LiDARs are susceptible to malicious spoofing attacks that aim to alter LiDAR's sensing outcome by adding fake objects to and removing real objects from the LiDAR's sensed point cloud map, and hence leading to severe safety consequences. For example, the feasibility of injecting fake points into the LiDAR's sensed point cloud was first demonstrated in [84]. They showed that LiDAR sensing results can be easily manipulated by a black-box attack using low-cost commodity hardware (less than 60 US dollars). Subsequent work in [19] launched LiDAR spoofing attacks that successfully fooled a real-world AV perception system, Baidu Apollo 2.5, to detect (faked) objects that do not actually exist in reality. The work in [112] further demonstrated that by spoofing only a small number of points (up to 100), the LiDAR object detection system can be fooled to detect non-existing objects. Their work shows the severity of the threats posed by spoofing attacks on AV LiDARs, which urgently calls for promising countermeasures that can better guarantee the safety of autonomous driving, so as to offer a peace of mind to users when they are using the technology.

In the last couple of years, many works have been focused on mitigating the *effect* of LiDAR spoofing attack using perception model-level defense methods [37, 62, 117, 147]. For example, the work in [112] proposed CARLO, which harnesses occlusion patterns between objects in the LiDAR point cloud for spoofed vehicle detection. The intuition is that, if there are many LiDAR points appearing to pass through a detected object, the object is likely to be a fake object Another anomaly detection system, Shadow-Catcher [44], identifies spoofed ghost objects by checking the contextual consistency between the object and its shadow. Treating the LiDAR's sensed point cloud as a depth image, these methods essentially follow the image-recognition research ideas in AI, which mainly consider the high-level contextual relationship, i.e., the perception, between the points to decide the presence of a spoofer. A critical weakness of these post-sensing methods is that their effectiveness fully depends on the correctness/truthfulness of their input, i.e., the LiDAR's sensing outcome (the point cloud). Therefore, a spoofer will be able to elude from these methods as long as it fabricates/fakes points that maintain the right contextual relationship among them.

Keeping the weakness of the above model-level methods in mind, another category of work is dedicated to fundamentally protect LiDAR from spoofing attacks based on physical-layer authentication (PLA). These methods work on the signal level, and try to authenticate LiDAR's signal based on some physical properties of the light so as to ensure the correctness of LiDAR's sensing outcome. For example, the work in [69] uses amplitude modulation (AM) to directly encrypt LiDAR signals with side channel information leaked from a cryptographic device. Since side channel information cannot be recreated without the knowledge of the secret key, attackers cannot inject spoofing signals while remaining undetected. In [70], the authors used the signal-to-noise ratio (SNR) of the received signal as an authentication metric and developed a probabilistic approach based on the Neyman-Pearson criterion to select the best SNR threshold for spoofing attack detection. However, a major limitation of their methods is that they use the intensity of the received signal for spoofing detection, which is not a robust metric for LiDARs. In LiDAR sensing, the intensity of the reflected signal faces complicated distortions that are related to the material, size, and roughness of the reflector. Therefore, the sensing signals encrypted by the method in [69] may become unrecognizable after reflections. Furthermore, the SNR of the sensing signal used in [70] has a large variance due to the dynamics of the environment (e.g., reflectors are moving), making it difficult to accurately identify the spoofing signal.

In this paper, we find that the intrinsic vulnerability of LiDAR is caused by the fact that current LiDAR sensors blindly accept incoming signals without verifying the sender of the signal. Therefore, we propose to use the signal's Doppler frequency shift to verify the sender of the signal and detect potential spoofing attacks. The fundamental difference between a spoofing signal and a legitimate signal is that the spoofing signal is generated by the attacker and directly sent to the LiDAR receiver, while the legitimate signal is originally sent by the LiDAR transmitter and then echoed/reflected by some objects. Based on this observation and through experiments on real-world testbed, we find that the propagation differences between legitimate and spoofing signals can be characterized by the Doppler shift of the received signal, which can then be used for spoofing attack detection. Specifically, the major contributions of our work are four-folds:

- To have a deep understanding on the vulnerability of today's LiDAR sensors, we thoroughly analyze the working principle of LiDAR and conduct real-world experiments to demonstrate how easily a spoofing attack can be launched against LiDAR, so as to show such attacks are realistic to current LiDAR technology, and hence the urgency of a promising countermeasure.

- We prove that the Doppler frequency shifts of legitimate and spoofing signals present different characteristics, and this signal-level difference can be used to fundamentally protect the sensing outcome of LiDAR. We then build a testbed to verify the feasibility of extracting Doppler shift from LiDAR signals with only minor modifications to the LiDAR system. Compared to amplitude and AM-based authentication methods [69, 70], the signal's Doppler frequency shift is a more robust and reliable decision statistic for spoofing detection, because it is decided by the motion between the LiDAR and sensed object and is less affected by the RF environment.

- To show how the Doppler shift can be used to detect spoofing attacks under different scenarios of attacker capabilities, we thoroughly consider three attack models, including static attacker, moving attacker, and moving attacker with control of both velocity and signal frequency. In each of these models, we first show how spoofing attacks can be performed and then present our countermeasures for spoofing detection.

- We make the proposed detection mechanisms more accurate and practical by further accounting for the short-term variance/uncertainty in the vehicle's velocity, caused by the vehicle's acceleration and random perturbation on its movement by the road condition. A statistical spoofing detection framework is proposed to jointly consider the impact of velocity and acceleration on the Doppler shift, which can provide more accurate spoofing detection in realistic application environments. Extensive numerical results are provided in a wide range of settings and road conditions.

The remainder of the paper is as follows. We begin by briefly reviewing related work in Section 4.2. Then, we analyze the working principle and vulnerability of LiDAR in Section 4.3. We analyze the difference in Doppler frequency shift between legitimate and spoofing signals in Section 4.4. We consider three attack models and present the spoofing detection method in Section 4.5. The statistic-based spoofing detection framework is presented in Section 4.6. And finally, we conclude our paper in Section 4.7

4.2 Related Work

4.2.1 Attacks Against AV Sensors

Attacks against AV sensors can be classified into three categories according to the physical channel used by the attacker [106, 70], namely, the regular, side, and transmission channel attacks. Regular channel attacks use the same working channel as the sensor (e.g., laser for LiDAR) to directly alter the sensing results. Side channel attacks use a physical channel other than the sensor's working channel to attack the LiDAR [60, 98]. Lastly, transmission channel attacks focus on the transmission channel that connects the sensor and other parts of the system [20, 3, 50].

4.2.2 Perception Model Level Defense Methods

Since the point cloud data generated by LiDAR is used by the AI-based perception model for 3D object detection, many research works focus on mitigating the effect of spoofing attack by the perception model level defense methods. For example, in [44], the authors proposed Shadow-Catcher, which validates object identities by examining the shadow of the object in the LiDAR point cloud. The idea is that, for the genuine object representations in the LiDAR point cloud, they are closely followed by regions void of measurements (shadow region). For the injected spoofed object, it is either does not have shadow regions or its shadow regions are inconsistent with the object's size or shape. In [141, 147], the authors leveraged the spatio-temporal consistency of the genuine object for spoofing attack detection. The authors utilized a motion prediction framework to analyze the spatio-temporal consistency of objects across consecutive frames in a driving scene. The spoofed object is detected if it violates the law of temporal consistency. However, the major limitation of the above model-level defense methods is that they rely on the geometric formation of points in the LiDAR point cloud and its evolution over time (i.e., the contextual relationship between points) to detect spoofing. These mechanisms first aggregate multiple points in the point-cloud to establish an object representation, and then check whether the object representation remains contextually consistent over a certain time period. Therefore, if an attacker can maintain the correct contextual relationship among the fabricated points, it can evade from being detected by these spoofing detection methods. In contrast, our proposed method works in the signal space and evaluates each point in the LiDAR point-clould individually, by testing whether the Doppler shift of the received signal matches with the expected Doppler shift caused by the velocity of the LiDAR. A spoofing LiDAR signal (i.e., a point in the point cloud) causes mismatch between the received Doppler shift and the expected Doppler shift, and hence will be detected by the proposed method, irrespective of its geometric relationship with the other points in the point-clould.

4.2.3 Signal Level Defense Methods

The signal-level defense method mainly uses physical-layer authentication (PLA) for spoofing detection. Unlike perception model-level defense methods, PLA protects LiDAR sensors against spoofing attacks by identifying the malicious signal in the analog domain [11, 120]. The most widely used PLA method is to endow the probes used by active sensors with a special designed feature and use the feature to authenticate the responses. For example, in [107], Shoukry *et al.* proposed PyCRA, which identifies spoofing attacks for magnetic sensors and radio-frequency identification (RFID) tags. PyCRA turns off the probe signal at random instants to verify the existence of any spoofers. If there is no spoofer, it will receive nothing; otherwise, the spoofing attack is identified. However, PyCRA does not meet the high availability requirement in safety-critical systems, such as an autonomous driving system. When using PyCRA, an AV LiDAR should be turned off at random times for attack detection. As a result, the LiDAR sensor becomes unavailable for environmental sensing during that period, which may cause safety problems for AVs.

Def LV	**Ref**	**Attack Model**	**Phy Inva**	**Def Strategies**
Model	[141]	Add/remove Points in Point-Cloud	N/A	Spatio-temporal Consistency
	[112]			Occlusion Pattern Verification
	[44]			Shadow Pattern Verification
	[147]			Disparity Errors Verification
Transmission	[11]	Data Tamping in Transmission		Dynamic Watermark
	[20]			QIM-based Watermark
Signal	[69]	Change Signal ToF	Amplitude	Signal Amplitude Encryption
	[107]		Time	Challenge-Response Authentication
	[70]		Amplitude	SNR Distribution Analysis
	Ours		**Frequency**	**Doppler Shift Verification**

Table 4.1: Comparison between related work (Def LV: defense method level, Def Strategies: defense strategies, Phy Inva: physical invariants used)

For better readability, we summarize the related work in Table 4.1, to highlight the difference between our work and other works.

4.3 LiDAR Working Principle and Vulnerability

To defend LiDAR against spoofing attack, we first need to understand the working principle and vulnerability of LiDAR. In this section, we first analyze the working principle and vulnerability of LiDAR. Then, we conduct real-world experiments to demonstrate the practicability and easiness of conducting spoofing attacks against LiDAR.

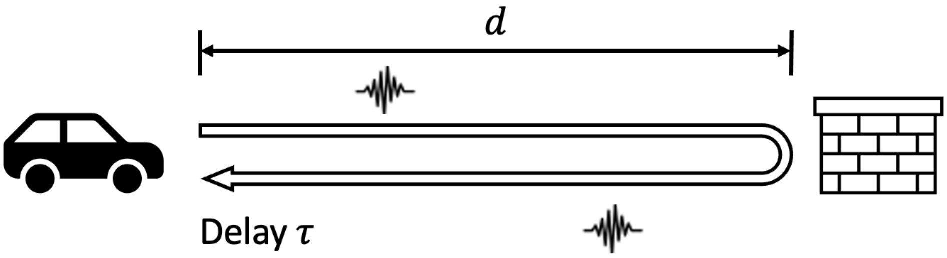


Figure 4.1: Normal LiDAR sensing.

4.3.1 LiDAR Working Principle

LiDARs detect and localize objects by actively probing objects with pulses of infrared laser signals between 750 nm to 1.5 μm. Figure 4.1 shows a typical LiDAR sensing scenario. A LiDAR sensor consists of two parts: a laser diode as transmitter and a photodetector as receiver. During LiDAR sensing, the transmitter periodically emits laser pulses to the environment. After the pulses reach the objects in the environment, they are reflected back and received by the LiDAR photodetector. The reflected signals are called echo signals. The time difference between the emitting and arriving time of the signal, i.e. Time of Flight (ToF), is used to calculate the distance d between the LiDAR and the object. Let t_s denote the time of the laser pulse being sent by the transmitter, and t_a denote the time of the echo signal being received by the photodetector. The ToF of the received signal τ is $t_a - t_s$, and the distance d between the LiDAR and the detected object is

$$d = \frac{c}{2n}\tau, \tag{4.1}$$

where n is the refractive index of the propagation medium ($n = 1$ for air) and c is the speed of light. By mechanically or electronically steering the laser pulses towards different directions and calculating the ToF distances of the echo signals, LiDAR is able to generate a point cloud, which is a high-resolution depth image of the environment.

4.3.2 Motivating LiDAR Security via Real-world Observations

Existing LiDAR only accepts the first arrival signal and uses the signal's arrival time for ToF distance calculation without verifying whether that signal was sent out by the LiDAR's laser diode (i.e., the legitimate transmitter). This leaves a sufficient loophole for many possible forms of spoofing attacks. In the following, we first present a simple toy example implemented in [19] to illustrate a basic type of spoofing attack that fakes a point in the LiDAR point cloud through ToF manipulation. Such a basic attack can be used as building blocks by the attacker to create more sophisticated spoofing attacks, e.g., those that fake an object. We then present our real-world experiments that are built upon two commercial YD X2L LiDARs to demonstrate how spoofing attacks can actually take place in real-world applications. The main purpose of this section is two-fold: (1) To better motivate the LiDAR spoofing attack problem studied in the paper. In particular, by demonstrating a real LiDAR spoofing attack over a commercially available LiDAR system, we wish to show that such an attack is very realistic for LiDAR systems available in today's market. Note that even though such attacks have been demonstrated in the past, most of them were based on experimental testbeds in a lab rather than directly over a commercially available LiDAR product. We believe that showing the spoofing attack on a commercial LiDAR product will make the attack more convincing, especially for readers not familiar with LiDAR and its vulnerabilities. (2) To better show the compelling nature of the problem: By showing how easy it is to launch a spoofing attack against current LiDARs, we highlight the urgent need for solutions to this compelling security problem.

A Toy Example for Spoofing Attack

A basic point-faking attack was proposed and implemented in [19], as illustrated in Figure 4.2. This system features a photodiode, a time delay component, and a laser diode. The total cost of the system is less than 50 US dollars. The goal of this spoofing system is to deceive the LiDAR by sending signals with false ToF that simulate a fake object. The spoofing signal can be injected to LiDAR via an attacker-controlled laser diode, whose working wavelength is the same as the victim's LiDAR. By properly controlling the timing of the spoofing signal, the attacker can alter the ToF measurements of the victim LiDAR, which in turn results in a counterfeit point at the distance that the attacker desires. More specifically, suppose that the attacker aims to mislead the LiDAR in detecting a counterfeit point at distance d_{spoof}, while the actual physical distance between the LiDAR and the attacker is d. To achieve the attack goal, the attacker first needs to synchronize with the victim LiDAR to obtain the sending time of the laser pulses. The attacker then sends a spoofing signal to LiDAR and ensures that the arrival time of the spoofing signal t_a^{spoof} is:

$$t_a^{spoof} = t_s + \tau_{spoof}. \tag{4.2}$$

where $\tau_{spoof} = 2d_{spoof}/c$. In this case, when the spoofing signal is received by LiDAR, the calculated ToF distance between the LiDAR and the attacker is now manipulated to be d_{spoof} (instead of being d), resulting in a faked point in the LiDAR's point cloud.

To launch a real-world spoofing attack, the photodiode in Figure 4.2 serves as a synchronization device to trigger the delay component whenever it captures laser signals from the victim LiDAR. And the delay component activates the laser diode to send a spoofing signal towards the victim LiDAR after a specified time delay τ_{spoof}.

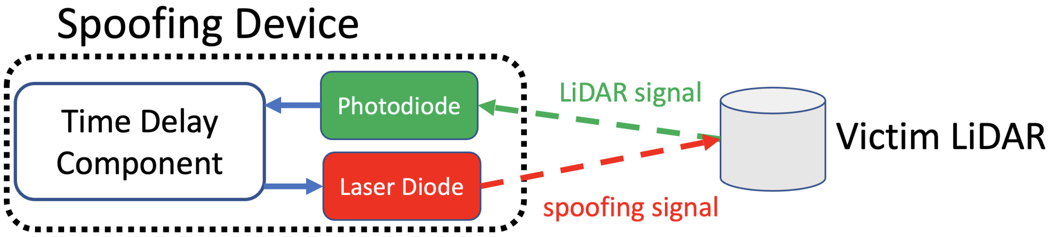


Figure 4.2: Spoofing device.

Our Real-world Experiments

To show how easily a spoofing attack can be launched against current LiDAR systems, we conduct the following real-world experiment. We use two YD X2L LiDARs [116]. One of the LiDARs acts as the victim LiDAR to generate point cloud data for the test environment. The other LiDAR is configured as a spoofing attacker that periodically generates spoofing signals with random ToF to attack the victim LiDAR. We conduct our spoofing attack experiments in an outdoor parking lot, and the test environment and test location are shown in Figure 4.3. In the test, the victim LiDAR is running normal operation to sense the environment and generate point cloud data, and the spoofing attacker is located 1.5 meters away from the victim LiDAR and shooting signals with random time delays.

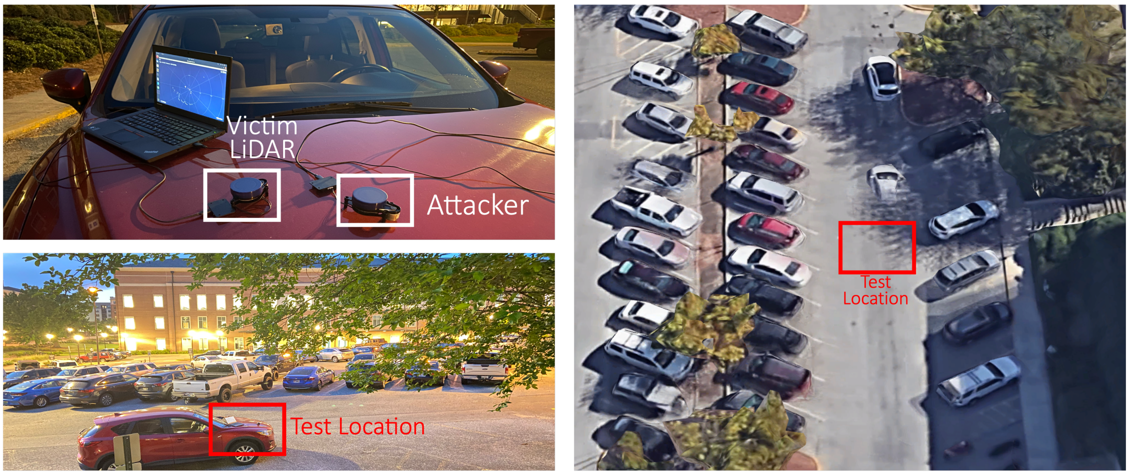
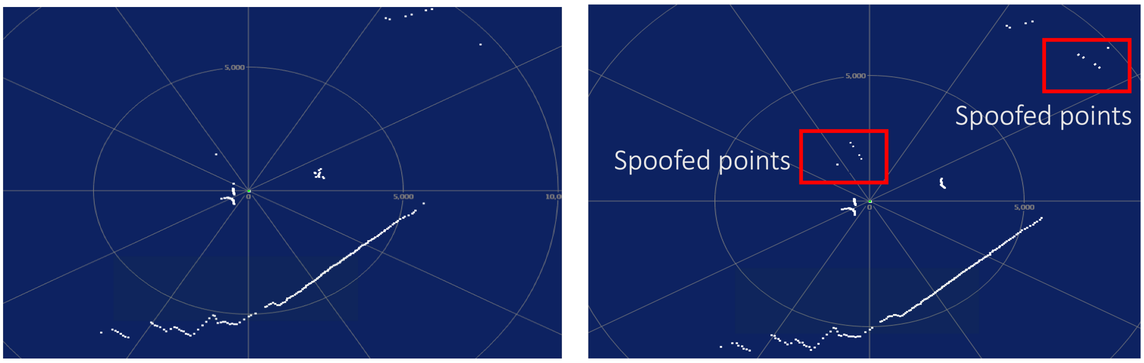


Figure 4.3: Outdoor spoofing test environment and setup.



(a) Normal point cloud.

(b) Point cloud under attack.

Figure 4.4: LiDAR point-cloud data.

Figure 4.4 shows a comparison of LiDAR point cloud with and without spoofing attack. In Figure 4.4(a), the point cloud with no spoofing attack clearly captures the shape and distance of the surrounding objects. In contrast, in Figure 4.4(b), there are multiple spoofed points shown on the point cloud map (marked by red squares). It can be seen that under the random attack, two small clusters of spoofed points are generated in the LiDAR point cloud. These clusters of spoofed points may deceive the LiDAR system to mis-interpret them as two small objects in front of the LiDAR: one in the 12 o'clock direction and the other in the 2 o'clock direction, which actually do not exist at all in reality. This indeed poses a serious safety threat for the AVs. Note that the experiment in Figure 4.4 is just a simple example. In reality, instead of a random attack, the attacker can enhance their attack effects (i.e., generate a bigger cluster of spoofed points in the point cloud) by launching more sophisticated attacks.

4.4 Doppler Frequency Shift in LiDAR Sensing

To fundamentally protect LiDAR against spoofing attacks in the analog domain, it is crucial to distinguish legitimate sensing signals and spoofing signals based on signal-level features. However, choosing an appropriate physical feature that can correctly represent the difference between legitimate and spoofing signals is a challenge.

In this section, we first prove that the Doppler frequency shift of the received signal can properly characterize the propagation difference between the legitimate and spoofing signal and distinguish the spoofing signal. Then, we build a real-world testbed to show the practicability of extracting Doppler shift from LiDAR's laser signal.

4.4.1 Doppler Frequency Shift Difference Between Legitimate Signal and Spoofing Signal

The Doppler effect, or Doppler frequency shift, is the change in frequency of a signal in relation to the relative movement between the signal's transmitter and receiver. In LiDAR sensing, due to the relative motion between the LiDAR and the detected object, the echoed signal presents a frequency shift caused by the Doppler effect.

In LiDAR sensing, the legitimate sensing signal is sent by LiDAR's transmitter, reflected by an object in the environment and then received by the LiDAR's receiver, which travels through a round trip. Let us consider a two-dimensional case to derive the Doppler frequency shift of the legitimate sensing signal. Let the velocity vector of LiDAR be $\vec{v}_L$ and the velocity vector of the object detected be $\vec{v}_{(ob)}$. The Doppler shift of the legitimate sensing signal is determined by the relative radial speed between the LiDAR and the object, which is defined as the rate of change of the distance between them. The relative radial speed Δv between the LiDAR and the object is calculated as $\Delta v = (\vec{v}_L - \vec{v}_{(ob)}) \cdot \vec{l}$, where $\cdot$ is the dot product and $\vec{l}$ is the direction of arrival (DoA) vector of the signal (i.e., the direction along the line connecting the LiDAR and the detected object). Due to the large magnitude of the speed of light, the DoA of the received signal is considered to be the same as the sending direction of the signal, which is considered as known.

Recall that the legitimate sensing signal travels through a round trip. We introduce an intermediate signal frequency f_r', which is the frequency of the signal that reaches the object. Let f_0 be the frequency of the signal transmitted by LiDAR and f_r be the frequency of the received signal, as shown in Figure 4.5(a). In the forward trip of the round trip, the signal with frequency f_0 is sent to the object, and f_r' is

$$f_r' = (\frac{c}{c - \Delta v}) f_0. \tag{4.3}$$

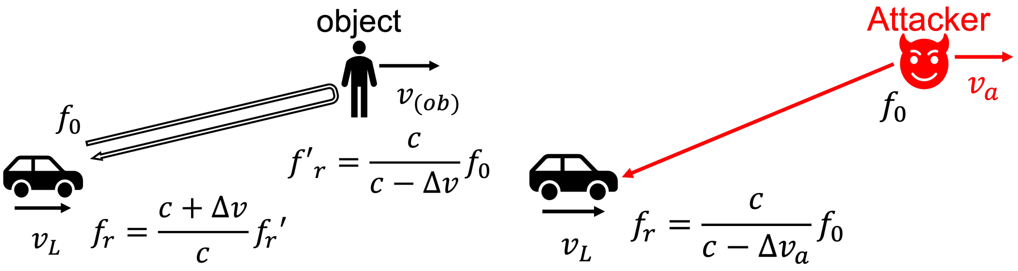
Then, the signal with frequency f_r' is reflected back to LiDAR on the same route, and the received signal frequency f_r is

$$f_r = (\frac{c + \Delta v}{c}) f_r'. \tag{4.4}$$

The Doppler frequency shift Δf of the received signal is calculated as the frequency difference between the transmitted and received signals, which is:

$$\Delta f = f_r - f_0 = (\frac{c + \Delta v}{c - \Delta v} - 1) f_0 \approx \frac{2 f_0}{c} \Delta v. \tag{4.5}$$

The approximation is valid since Δv is much smaller than the speed of light c ($3 \times 10^8 m/s$).


object
$v_{(ob)}$
f_0
$f'_r = \frac{c}{c - \Delta v} f_0$
v_L
$f_r = \frac{c + \Delta v}{c} f_r'$
Attacker
v_a
f_0
v_L
$f_r = \frac{c}{c - \Delta v_a} f_0$

(a) Normal Sensing (round trip).

(b) Spoofing attack (single-way).

Figure 4.5: Doppler frequency shift illustration.

In contrast, the spoofing signal is sent directly to LiDAR by the attacker, which only travels one-way. Let $\vec{v}_a$ be the velocity vector of the attacker. The relative radial speed between the LiDAR and the attacker is $\Delta v_a = (\vec{v}_L - \vec{v}_a) \cdot \vec{l}$, as shown in Figure 4.5(b). Since the frequency of the transmitted spoofing signal is also f_0, the Doppler frequency shift of the spoofing signal Δf_a is

$$\Delta f_a = (\frac{c}{c - \Delta v_a}) f_0 - f_0 = \frac{\Delta v_a}{c - \Delta v_a} f_0 \approx \frac{f_0}{c} \Delta v_a. \tag{4.6}$$

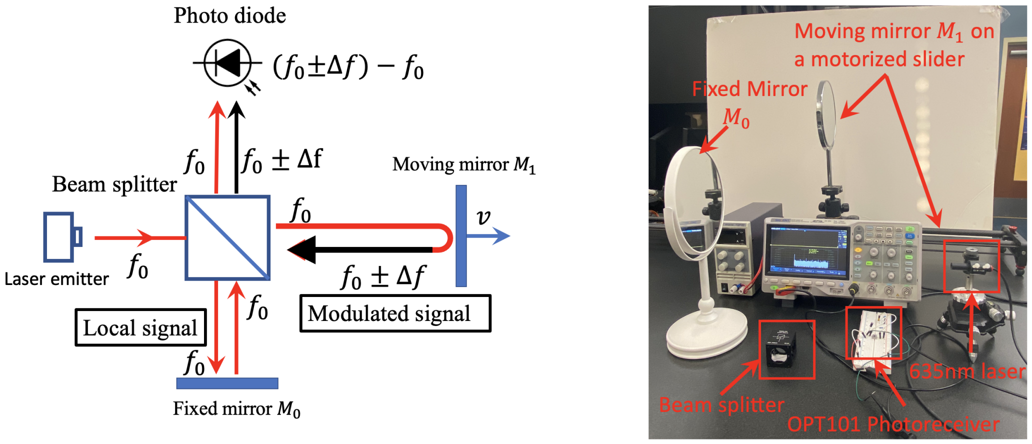
Based on the above analysis, it is clear that the Doppler shifts of the legitimate and spoofing signals are different due to their different propagation paths: The Doppler frequency shift of the legitimate sensing signal is twice as much as that of the spoofing signal under the same radial speed due to its round trip propagation. As will be elaborated shortly in Section 4.5, the above margin (a factor of 2) between the Doppler frequency shifts of legitimate and spoofing signals can be utilized to construct reliable and accurate spoofing detection mechanisms under various attack conditions.

Next, we present a proof-of-concept testbed to demonstrate the feasibility of extracting Doppler frequency shift from the laser signal of a moving LiDAR.

4.4.2 Feasibility Study of Extracting Doppler Frequency Shift

Proof-of-concept Testbed Design

We design and build a proof-of-concept testbed to test the feasibility of extracting the velocity-based Doppler shift from the laser similar in nature to those used in LiDAR systems. The schematic diagram of the testbed is shown in Figure 4.6(a). The Doppler shift is extracted by using the self-mixing effect of the signal. Specifically, the laser signal with frequency f_0 is first split into two orthogonal beams by a 3 dB beam splitter in the middle. Then, one beam of signal, which is called the local signal, is reflected back by the fixed mirror M_0. And the other beam of signal, termed the modulated signal, is reflected back by a moving mirror M_2 whose velocity is v. The local signal and the modulated signal are mixed together and received by the photodiode to extract the Doppler frequency shift of the modulated signal.



(a) Schematic diagram.

(b) Testbed layout.

Figure 4.6: Testbed design.

The Doppler shift of the modulated signal is extracted by the homodyne detection method. The local signal Y_{LO} and the modulated signal Y_M can be expressed as

$$Y_{LO} = A_{LO} \cdot e^{-j(2\pi f_0 t + \phi_{LO})},$$

$$Y_M = A_M \cdot e^{-j(2\pi (f_0 \pm \Delta f) t + \phi_M)},$$

where $A_{LO}, \phi_{LO}, A_M, \phi_M$ denote the amplitude and phase shift of the local and modulated signal, respectively. j is the imaginary unit and $\Delta f = \frac{2v}{c} f_0$ is the Doppler frequency shift caused by the movement of the mirror M_1.

The output of the photodetector is the combined signal power of Y_{LO} and Y_M. Due to the low pass filtering effect of the photodetector, the high frequency components of $Y_{LO} + Y_M$ are filtered out, and the output power is

$$P_{out} = \frac{A_M^2}{2} + \frac{A_{LO}^2}{2} + A_M A_{LO} \cos(2\pi \Delta f t + \phi_M - \phi_{LO}),$$

which is a beat signal with frequency Δf. By filtering out the direct current (DC) signal, Δf can be extracted by fast Fourier transform (FFT).

Testbed Implementation

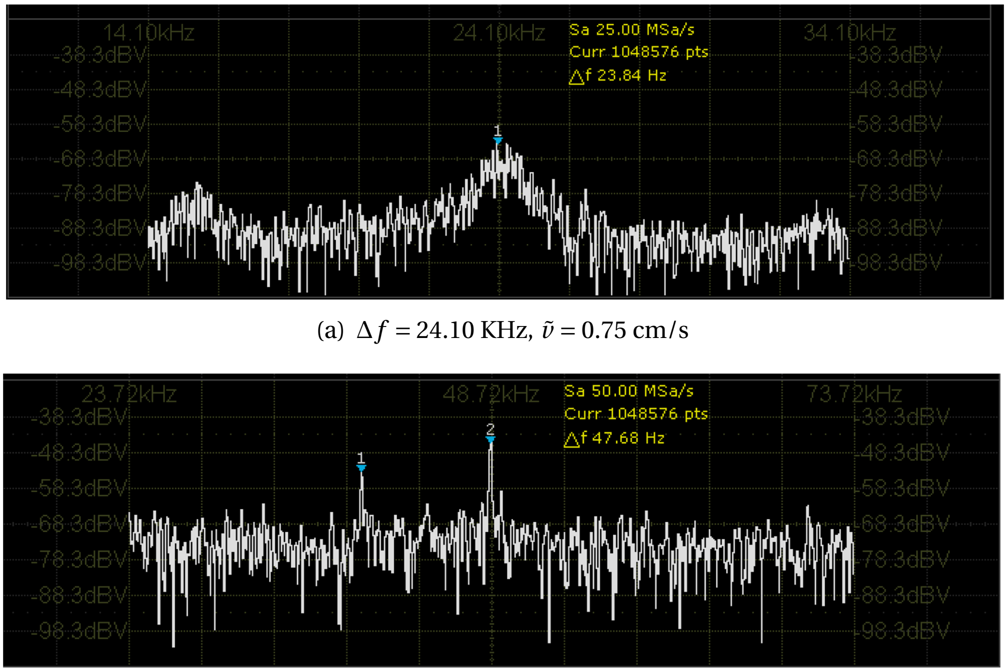
Regarding the implementation of the testbed, we use a 635 nm ThorLabs PL202 laser diode to send laser signals. A cubic beam splitter, ThorLabs CCM1-BS013, is used to split the beam. The photoreceiver is OPT101 from Texas Instruments. The moving mirror is attached to a motorized camera slider for stable and continuous movement. An oscilloscope is connected to the photoreceiver for data collection and visualization. The layout of the testbed is shown in Figure 4.6(b).

Note that for the demonstration purpose, we use mirrors instead of real obstacles. In real-world scenarios, the surface roughness and color of the obstacle can affect the received signal's SNR. Rough surfaces can scatter laser light, and darker colors absorb more light, resulting in weaker reflections (i.e., smaller reflection coefficient of the obstacle) and hence lower SNR.

However, in practice, the reduced reflection coefficient can be well compensated by using a higher power laser emitter and filter lenses, which are commonly adopted by vehicle LiDARs. Therefore, in real-world use cases, the LiDAR's SNR should be sufficient for reliable and accurate Doppler shift extraction.

Test Results

We then use the above testbed to extract the Doppler frequency shift of the received signal and estimate M_1's velocity v. Figure 4.7 shows the Fourier spectrum of the signals for different velocities of M_1. In the experiment, the moving speed of M_1 is set to 0.75 cm/s and 1.50 cm/s, respectively. The estimated velocity $\tilde{v}$ of M_1, which is calculated from the Doppler shift Δf, is $\tilde{v} = \frac{\Delta f}{2f_0}c$. In Figure 4.7(a), the Doppler shift of the signal is 24.10 KHz, which corresponds to $\tilde{v}$= 0.75 cm/s and is match with the M_1's ground truth speed v = 0.75 cm/s. In Figure 4.7(b), there are two peaks found, and the peak with the highest value is chosen as the Doppler shift that corresponds to the real signal. In this case, the Doppler frequency of the signal is 48.72 KHz, which corresponds to $\tilde{v}$= 1.55 cm/s. Under real-world conditions, the Doppler spectrum of the received signal may contain multiple peaks due to random noise and subtle movement of the object. A general principle of identifying the real signal is to choose the frequency component with the highest energy, as this is caused by the dominant movement of the object. Compared to the ground truth speed of v = 1.50 cm/s, the small variance between v and $\tilde{v}$ is caused by noise in the photodiode. This small variance does not affect the accuracy of our proposed spoofing attack detection method. As will be shown in later sections, the velocity detection error (about 3% as shown in Figure 4.7(b)) caused by random noise is much smaller than the separation between the detected velocity of a real object and the detected velocity of a spoofed object (the former is twice as much as the latter). Furthermore, the impact of random noise can be reduced by our statistical spoofing detection framework presented in Section 4.6.



(a) $\Delta f = 24.10$ KHz, $\tilde{v} = 0.75$ cm/s

(b) $\Delta f = 48.72$ KHz, $\tilde{v} = 1.55$ cm/s

Figure 4.7: Doppler shift spectrum results.

In summary, this experiment establishes the feasibility of extracting the Doppler shift of high-frequency LiDAR signals over a testbed that is open for redevelopment. The same structure of the testbed can be integrated into real-world LiDAR systems to extract the Doppler shift and detect spoofing attacks. Specifically, the testbed is based on an interferometer structure and can be integrated into LiDARs. The potential challenges of incorporating our method into the LiDAR system include: (1) Cost Issue: Implementing the structure shown in Section IV.B requires an additional frequency mixer and A/D converters, which increases the manufacturing cost of LiDAR sensors. (2) Standardization Issues: The lack of industry-wide standards for LiDAR systems can cause compatibility and interoperability issues between different AV models and brands. At this point, all commercial LiDAR products available on the market are proprietary and are not open for redevelopment.

We understand that there have been numerous existing commercial products on the market that are capable of extracting Doppler shift from laser signals. However, these products are often proprietary, and hence are not friendly to redevelopment. The spoofing detection measures developed in the subsequent sections can be implemented on the testbed presented in this section.

4.5 Doppler Shift based Spoofing Detection

In the previous section, we demonstrated that the Doppler shift of the laser signal can be used to distinguish between a spoofing signal and a legitimate sensing signal. In this section, we present the detailed designs that utilize the Doppler frequency shift for LiDAR spoofing attack detection under various attack models. Specifically, we first study the uniform-motion scenario, where the velocities of the attacker, the LiDAR, and genuine objects in the environment are assumed to be constants during the window of detection (we will relax this assumption and consider accelerations in the next section). We consider three different spoofing attack models, respectively: a static attacker, a mobile attacker, and a mobile attacker that controls both its velocity and signal frequency. Each of these models can be considered as a generalization of the model before it. We start off our detection design with the simplest attack model – the static attacker, and gradually make the design more general by considering more realistic conditions in the attack. For each attack model, we first show how the spoofing attacks are performed. Then, we illustrate the countermeasure that uses the signal Doppler shift to identify the spoofing attack.

We need to point out that the spoofing attack models adopted by the related works are essentially based on the same assumption of the attacker's most basic attack capability considered in this work. In particular, no matter it is the fake object injection attack or the target object removal attack, they are all built upon the attacker's foundational capability of being able to manipulate the time-of-flight of the LiDAR signal, so that the attacker can either inject a fake point into or remove a real point from the LiDAR's point cloud. Our work considers exactly the same foundational capability of the attacker, as shown in Figure 2 and Section III.B.(1). In this regard, the comparison between our work and those related works is fair. In addition, our work not only considers the same foundational capability of the attacker, but also studies how such a foundational capability can be achieved by an attacker and how such capability can be countered under various realistic scenarios, e.g., when the attacker is static, or when the attacker is mobile, or when the attacker can control its movement and the frequency of the LiDAR signal, etc. Because the detection methods proposed in our work essentially target detecting the manipulation of the time-of-flight of the LiDAR signals, they are also able to detect those fake object injection attacks and the target object removal attacks which are based on the above manipulations.

4.5.1 Attack Model 1: Static Attacker and Moving LiDAR

Spoofing Attack in Model 1

We first consider the case where only LiDAR is moving with constant velocity $\vec{v}_L$, and any other objects and the attacker remain static. This is a common scenario for LiDAR spoofing attacks. For example, the attacker can place the spoofing device on the roadside to shoot malicious laser pulses to AVs passing by. We also assume that the LiDAR system already knows that all genuine objects are static. In this scenario, similar to the example illustrated in Section 4.3.2, the attacker aims to mislead the LiDAR in detecting a counterfeit point at distance d_{spoof} while the real distance between the LiDAR and the attacker is d. This is achieved by sending spoofing signal with time delay τ_{spoof} to the victim LiDAR, as shown in Figure 4.8.

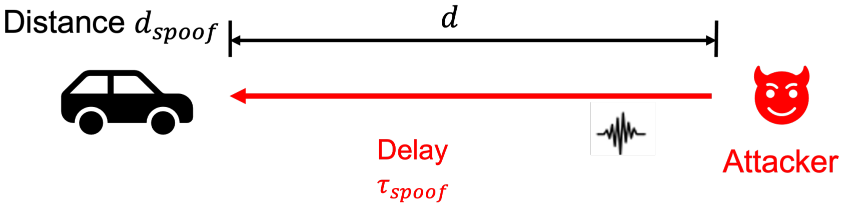


Figure 4.8: Spoofing attack in Scenario 1.

In this attack scheme (and also the subsequent two attack models), it is assumed that the attacker is aware of the working frequency of the victim LiDAR, and the transmitted spoofing signal has the same frequency as the victim LiDAR's working frequency. This assumption is practical because the working frequency of a vehicle's LiDAR can be easily obtained through the product specification. We also assume that the attacker is aware of its distance to the victim LiDAR, so that it can decide the timing of emitting the spoofing signal that misleads the victim LiDAR to calculate d_{spoof}. This assumption is reasonable because the attacker can simply use its own LiDAR to monitor its distance to the victim in real time.

Spoofing Detection in Attack Model 1

In Attack Model 1, a spoofing signal can be identified by testing whether the Doppler shift of the received signal matches the expected Doppler shift caused by the velocity of the LiDAR. Specifically, for the legitimate sensing signal sent to direction $\vec{l}$, since only LiDAR is moving with velocity $\vec{v}_L$, the expected Doppler frequency shift of the reflected signal is $\frac{2f_0}{c}\vec{v}_L \cdot \vec{l}$. Here, due to the small field of view of LiDAR receiver (less than 1°), the transmission direction of the signal is the same as the receive direction. Let the Doppler shift of the received signal be Δf_r (Δf_r can be measured as illustrated in Section 4.4.2). To detect a spoofing signal, the following should be tested:

$$\Delta f_r \stackrel{?}{=} \frac{2f_0}{c}(\vec{v}_L \cdot \vec{l}). \tag{4.7}$$

For the spoofing signal sent by the attacker from direction $\vec{l}$, since the attacker is static and the spoofing signal travels one way, its Doppler shift is only $\frac{f_0}{c}\vec{v}_L \cdot \vec{l}$, – a margin of a factor of 2. Therefore, the spoofing signal can be detected.

4.5.2 Attack Model 2: Moving Attacker and Moving LiDAR

Next, we consider a more general attack model, where the LiDAR, the attacker, and the object in the environment are moving. This scenario is more common than attack model 1. For example, the attacker can drive a vehicle in close proximity to the victim AV, e.g., in the same lane or adjacent lanes, to shoot the laser pulses to the victim AV's LiDAR. To better present the spoofing attack and the proposed spoofing detection in this model, we first introduce some basic notation and definitions.

Let us consider a 2-D Cartesian coordinate system shown in Figure 4.9. Let the LiDAR's velocity be $\vec{v}_L$. Without loss of generality, we assume that the direction of $\vec{v}_L$ is the same as the y-axis, and the x-axis is perpendicular to $\vec{v}_L$. With the movement of the LiDAR and the object, the LiDAR receives a series of signals emitted by the LiDAR and then reflected by the object at different locations. In particular, at times $t_1, t_2, ..., t_K$, let the locations of the LiDAR and the object be LiDAR$_{t_1}$, Object$_{t_1}$, LiDAR$_{t_2}$, Object$_{t_2}$, ..., and LiDAR$_{t_K}$, Object$_{t_K}$, respectively. Denote the signal that is emitted from the LiDAR, reflected by the object, and then received by the LiDAR at time t_k by S_{t_k}, where $k = 1, 2, \dots, K$. The signal S_{t_k} can be presented as a tuple $S_{t_k} = [\Delta f_{t_k}, d_{t_k}, \theta_{t_k}]$, where Δf_{t_k}, d_{t_k}, and θ_{t_k} represent the signal's Doppler frequency shift, ToF distance, and angle of arrival (AoA), respectively, at time t_k, as shown in Figure 4.9. Let $\mathbb{S}$ denote the set of signals reflected by the object and received by the LiDAR from time t_1 to time t_K.

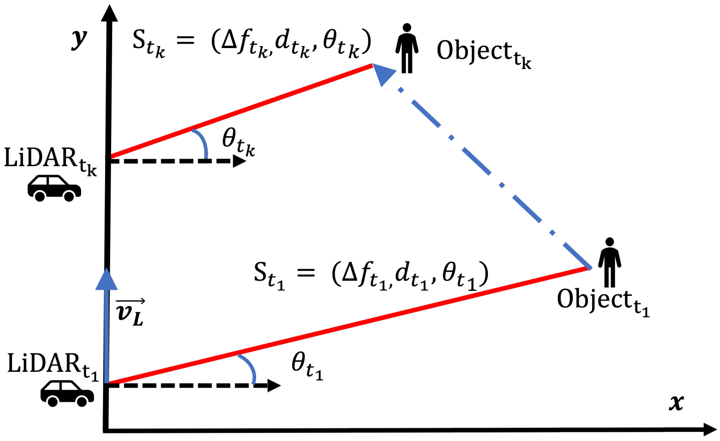

y
$S_{t_k} = (\Delta f_{t_k}, d_{t_k}, \theta_{t_k})$
Object$_{t_k}$
θ_{t_k}
LiDAR$_{t_k}$
$S_{t_1} = (\Delta f_{t_1}, d_{t_1}, \theta_{t_1})$
Object$_{t_1}$
$\vec{v}_L$
LiDAR$_{t_1}$
θ_{t_1}
x

Figure 4.9: The Cartesian coordinate system and signal tuple.

Using $\mathbb{S}$, we can determine the velocity of the object in one of two ways: (1) by the signal's Doppler shift or (2) by the ToF distance. We refer to the velocity determined from the ToF distance as the object's ToF velocity, and the velocity determined by the Doppler frequency shifts as the Doppler velocity. More specially, these velocities can be calculated as follows.

ToF Velocity: The ToF velocity of the object, denoted as $\vec{v}_{ToF}$, can be determined based on the ToF distances of the signals. In particular, the velocity vector can be represented as $\vec{v}_{ToF} = |\vec{v}_{ToF}|(\cos\phi_{ToF}, \sin\phi_{ToF})$, where $|\vec{v}_{ToF}|$ and ϕ_{ToF} denote the magnitude and direction angle of $\vec{v}_{ToF}$. Given any two signals received at time t_m and t_n ($t_m < t_n$), i.e., $S_{t_m} = [\Delta f_{t_m}, d_{t_m}, \theta_{t_m}]$ and $S_{t_n} = [\Delta f_{t_n}, d_{t_n}, \theta_{t_n}]$, $\vec{v}_{ToF}$ can be calculated as

$$\begin{aligned} |\vec{v}_{ToF}| = &[(d_{t_n}\sin\theta_{t_n} - d_{t_m}\sin\theta_{t_m})^2 + \\ &(d_{t_n}\cos\theta_{t_n} + |\vec{v}_L|\Delta t - d_{t_m}\cos\theta_{t_m})^2]^{\frac{1}{2}}, \end{aligned} \tag{4.8}$$

and

$$\phi_{ToF} = \arctan\frac{d_{t_n}\cos\theta_{t_n} + |\vec{v}_L|\Delta t - d_{t_m}\cos\theta_{t_m}}{d_{t_n}\sin\theta_{t_n} - d_{t_m}\sin\theta_{t_m}}, \tag{4.9}$$

where $\Delta t = |t_n - t_m|$.

Doppler Velocity: The object's Doppler velocity $\vec{v}_{Dop}$, can be represented as $\vec{v}_{Dop} = |\vec{v}_{Dop}|(\cos\phi_{Dop}, \mathrm{si}$ where $|\vec{v}_{Dop}|$ and ϕ_{Dop} are the magnitude and direction angle of the velocity. Given two signals $S_{t_m} = [\Delta f_{t_m}, d_{t_m}, \theta_{t_m}], S_{t_n} = [\Delta f_{t_n}, d_{t_n}, \theta_{t_n}] \in \mathbb{S}$, $|\vec{v}_{Dop}|$ and ϕ_{Dop} can be calculated by solving the following set of nonlinear equations:

$$\begin{cases} |\vec{v}_L|\sin(\theta_{t_m}) - |\vec{v}_{Dop}|\cos(\theta_{t_m} - \phi_{Dop}) = \frac{c}{2f_0}\Delta f_{t_m} \\ |\vec{v}_L|\sin(\theta_{t_n}) - |\vec{v}_{Dop}|\cos(\theta_{t_n} - \phi_{Dop}) = \frac{c}{2f_0}\Delta f_{t_n} \end{cases} \tag{4.10}$$

Depending on whether the attacker controls its velocity to facilitate the spoofing, the attacker's spoofing attack schemes can be divided into the following two cases.

Spoofing Attack When Attacker Does Not Control Its Velocity

We first consider a simple spoofing attack in which the attacker only manipulates the ToF distance of the probing signal, but does not control its velocity to facilitate the attack. Specifically, to launch a spoofing attack, the attacker injects spoofing signals into the victim LiDAR so that the legitimate signal set $\mathbb{S}$ that corresponds to a genuine object is replaced by the spoofing signal set $\mathbb{S}^{(spf)}$, where $S_{t_k}^{(spf)} = [\Delta f_{t_k}^{(spf)}, d_{t_k}^{(spf)}, \theta_{t_k}] \in \mathbb{S}^{(spf)}$. Note that due to the small field of view of the LiDAR receiver, the spoofing signal can only be injected when the LiDAR is transmitting to and receiving from the attacker's direction, and the AoAs of the spoofing signal can not be changed by the attacker. The goal of the attacker is to mislead the LiDAR's calculation of its distance to the faked object by manipulating $d_{t_k}^{(spf)}$, similar to that in Section 4.5.1. The ToF of the spoofing signal is determined by the attacker according to its attack goal, i.e., how far does it want the faked object to be from the LiDAR, based on Eq.(4.2).

Spoofing Detection When Attacker Does Not Control Its Velocity

The key insight in the above attack model is that $\Delta f_{t_k}^{(spf)}$ and $d_{t_k}^{(spf)}$ are not independent between each other. This is because both quantities are related to the velocity of the attacker/faked object, and both can be used to calculated that velocity according to Eqs.(4.8), (4.9), and (4.10). Since the attacker does not adjust its velocity according to the ToF distance it claims to be, there exists a mismatch between the Doppler velocity $\vec{v}_{Dop}$ and the ToF velocity $\vec{v}_{ToF}$. This allows us to detect spoofing by testing the following:

$$\vec{v}_{ToF} \stackrel{?}{=} \vec{v}_{Dop}. \tag{4.11}$$

For legitimate signals reflected by genuine objects, its ToF distance is authentic (i.e., not manipulated), and therefore $\vec{v}_{Dop} = \vec{v}_{ToF}$. Otherwise, a mismatch indicates the presence of a spoofing attack.

Spoofing Attack When Attacker Controls Its Velocity

An attacker can tailor its velocity to its claimed ToF distance to ensure that the calculated Doppler velocity $\vec{v}_{Dop}$ matches the ToF velocity $\vec{v}_{ToF}$. In particular, this can be achieved according to the following:

Proposition 4.5.1. Given the attacker's velocity $\vec{v}_a = |\vec{v}_a|(\cos\phi_a, \sin\phi_a)$, where ϕ_a is the direction angle of $\vec{v}_a$, to maintain consistency between the Doppler velocity and ToF velocity of the spoofing signals, for any two spoofing signals $S_{t_m}^{(spf)}, S_{t_n}^{(spf)} \in \mathbb{S}^{(spf)}$, where $S_{t_m}^{(spf)} = [\Delta f_{t_m}^{(spf)}, d_{t_m}^{(spf)}, \theta_{t_m}]$ and $S_{t_n}^{(spf)} = [\Delta f_{t_n}^{(spf)}, d_{t_n}^{(spf)}, \theta_{t_n}]$, $\theta_{t_m} \neq \theta_{t_n}$, $d_{t_m}^{(spf)}$ and $d_{t_n}^{(spf)}$ must satisfy the following equation set:

$$d_{t_m}^{(spf)} = \frac{\cos(\theta_{t_n})|\vec{v}_L|\Delta t + \frac{f(\theta_{t_m})\sin(\theta_{t_n}-\Phi)}{2\cos(\theta_{t_m}-\phi_a)}\Delta t}{\sin(\theta_{t_n}-\theta_{t_m})}, \tag{4.12}$$

$$d_{t_n}^{(spf)} = \frac{\cos(\theta_{t_m})|\vec{v}_L|\Delta t + \frac{f(\theta_{t_n})\sin(\theta_{t_m}-\Phi)}{2\cos(\theta_{t_n}-\phi_a)}\Delta t}{\sin(\theta_{t_n}-\theta_{t_m})}. \tag{4.13}$$

where

$$f(\theta) = |\vec{v}_L|\sin(\theta) + |\vec{v}_a|\cos(\theta - \phi_a),$$

$$\Phi = \arctan\frac{f(\theta_{t_n}) * \cos\theta_{t_m} - f(\theta_{t_m}) * \cos\theta_{t_n}}{f(\theta_{t_m}) * \sin\theta_{t_n} - f(\theta_{t_n}) * \sin\theta_{t_m}}.$$

and $\Delta t = |t_n - t_m|$.

Proof. The ToF velocity derived from the spoofing signals $\vec{v}_{ToF}$ must be equal to the Doppler velocity of the spoofing signals $\vec{v}_{Dop}$. Denote $\vec{v}_{Dop} = |\vec{v}_{Dop}|(\cos\Phi, \sin\Phi)$. Since the attacker is directly sending the spoofing signals to the victim LiDAR, the Doppler shifts of the spoofing signals are determined by the relative radial velocity between the LiDAR and the attacker. We also have the Doppler shifts of the spoofing signals as

$$\Delta f_{t_m}^{(spf)} = \frac{f_0}{c}(|\vec{v}_L|\sin(\theta_{t_m}) - |\vec{v}_a|\cos(\theta_{t_m} - \phi_\alpha))$$
$$\Delta f_{t_n}^{(spf)} = \frac{f_0}{c}(|\vec{v}_L|\sin(\theta_{t_n}) - |\vec{v}_a|\cos(\theta_{t_n} - \phi_\alpha))$$

And $\vec{v}_{Dop}$ can be obtained by substituting $\Delta f_{t_m}^{(spf)}$ and $\Delta f_{t_n}^{(spf)}$ into equation (4.10), which gives

$$|\vec{v}_{Dop}| = \frac{f(\theta_{t_m})}{2\cos(\theta_{t_m} - \Phi)} \tag{4.14}$$

$$\Phi = \arctan\frac{f(\theta_{t_n}) * \cos\theta_{t_m} - f(\theta_{t_m}) * \cos\theta_{t_n}}{f(\theta_{t_m}) * \sin\theta_{t_n} - f(\theta_{t_n}) * \sin\theta_{t_m}}.$$

where $f(\theta)$ is the function value of θ, and we have

$$f(\theta) = |\vec{v}_L|\sin(\theta) + |\vec{v}_a|\cos(\theta - \phi_a),$$

The ToF velocity can be obtained by equations (4.8) and (4.9). Letting $\vec{v}_{ToF} = \vec{v}_{Dop}$, we can obtain $d_{t_m}^{(spf)}$ and $d_{t_n}^{(spf)}$ as specified in the proposition. □

According to Proposition 4.5.1, given a pair of desired spoofing ToF distances $d_{t_m}^{(spf)}$ and $d_{t_n}^{(spf)}$ at time t_m and t_n, the attacker can calculate the required velocity that ensures a match between $\vec{v}_{ToF}$ and $\vec{v}_{Dop}$ by solving equations (4.12) and (4.13), so as to elude from being detected by the aforementioned detection mechanisms.

Spoofing Detection When Attacker Controls Its Velocity

A key insight of Proposition 4.5.1 is that the attacker's velocity must be coordinated with the ToF of the spoofing signals for a successful spoofing attack. Specifically, according to Eqs. (4.12) and (4.13), given a pair of desired fake ToF distances and the victim LiDAR's velocity, the attacker's velocity $\vec{v}_a$ is fully determined. Therefore, when there exist two LiDARs of different velocities, both are scanning the attacker at the same time, then there is no way for the attacker to adjust its velocity to satisfy the requirements from both LiDARs – one key cannot open two locks. In this case, there will be at least one LiDAR, whose calculated ToF velocity is inconsistent with the Doppler velocity. Based on the above insight, we propose a cooperatiave LiDAR sensing scheme [21, 68, 149, 55] for our spoofing detection. A basic cooperative LiDAR system is shown in Figure 4.10, which consists of two LiDARs: a Coop-LiDAR and an Ego-LiDAR. In cooperative LiDAR sensing, each LiDAR independently senses the environment and generates the data, and the generated sensing data are shared between them [149]. Note that in this scenario, it is essential to ensure the trustworthiness of the Cooperative LiDAR system, which can be guaranteed by using secured vehicle-to-vehicle (V2V) communication [63, 127, 54].

To detect the spoofing attack, we require LiDARs in the cooperative LiDAR system to move at different velocities. Each LiDAR computes its ToF velocity and Doppler velocity based on its received signals. The spoofing detection is conducted by checking whether the computed ToF velocity is consistent with the Doppler velocity at every LiDAR. To be more specific, suppose that we have N LiDARs in the cooperative LiDAR system with velocities $\vec{v}_L^{(1)},...,\vec{v}_L^{(N)}$, respectively. There exists at least a pair of LiDARs, say LiDAR i and LiDAR j, where $1 \le i, j \le N$, whose velocities are not equal, i.e., $|\vec{v}_L^{(i)}| \neq |\vec{v}_L^{(j)}|$. Each LiDAR calculates the Doppler velocity and the ToF velocity based on its received signals, which gives $\vec{v}_{Dop}^{(1)}$ and $\vec{v}_{ToF}^{(1)}$,...,$\vec{v}_{Dop}^{(N)}$ and $\vec{v}_{ToF}^{(N)}$, respectively.

For legitimate signals, the ToF and Doppler velocities computed by each LiDAR are consistent, i.e., $\vec{v}_{Dop}^{(1)} = \vec{v}_{ToF}^{(1)} = \ldots = \vec{v}_{Dop}^{(N)} = \vec{v}_{ToF}^{(N)}$, because they all correspond to the velocity of the same object. However, when a spoofing attacker is in place, it faces the following dilemma: On one hand, given the velocity of LiDAR i and the desired ToF distances to LiDAR i at time t_n and t_{n+1}, the attacker must set its velocity to, say $\vec{v}_a^{(i)}$, where $\vec{v}_a^{(i)}$ is decided based on Proposition 4.5.1, in order to elude from the detection of LiDAR i. On the other hand, given the velocity of LiDAR j and the desired ToF distances to LiDAR j at time t'_n and t'_{n+1}, where t'_n is close to t_n, and t'_{n+1} is close to t_{n+1}, the attacker must set its velocity to, say $\vec{v}_a^{(j)}$, where $\vec{v}_a^{(j)}$ is decided based on Proposition 4.5.1, in order to elude from the detection of LiDAR j. Because $|\vec{v}_L^{(i)}| \neq |\vec{v}_L^{(j)}|$, we can expect that in general $\vec{v}_a^{(i)} \neq \vec{v}_a^{(j)}$. Therefore, no matter which velocity the attacker chooses, at least one of LiDAR i and LiDAR j will be able to detect the attacker by testing the inconsistency between its calculated ToF velocity and Doppler velocity.

Alternatively, the attacker may just choose to move at velocity $\vec{v}_a^{(i)}$, and instead customize the spoofing ToF distances to LiDAR j at time t'_n and t'_{n+1} according to Eqs.(4.12) and (4.13). In this way, the ToF velocity is consistent with the Doppler velocity at each of the LiDARs i and j, i.e., $\vec{v}^{(i)}_{Dop} = \vec{v}^{(i)}_{ToF}$ and $\vec{v}^{(j)}_{Dop} = \vec{v}^{(j)}_{ToF}$, however, it must be true that $\vec{v}^{(i)}_{Dop} \neq \vec{v}^{(j)}_{Dop}$. Therefore, by sharing their ToF velocities and Doppler velocities with each other, LiDARs i and j can also detect the spoofing attack based on the inconsistency between their respective Doppler velocities.

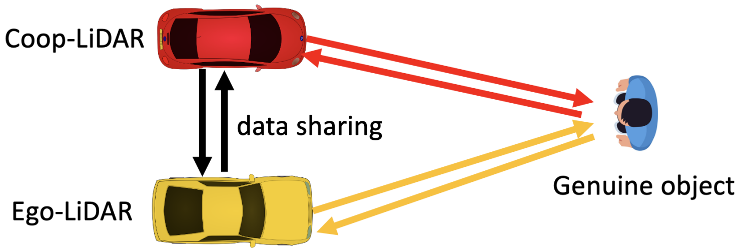
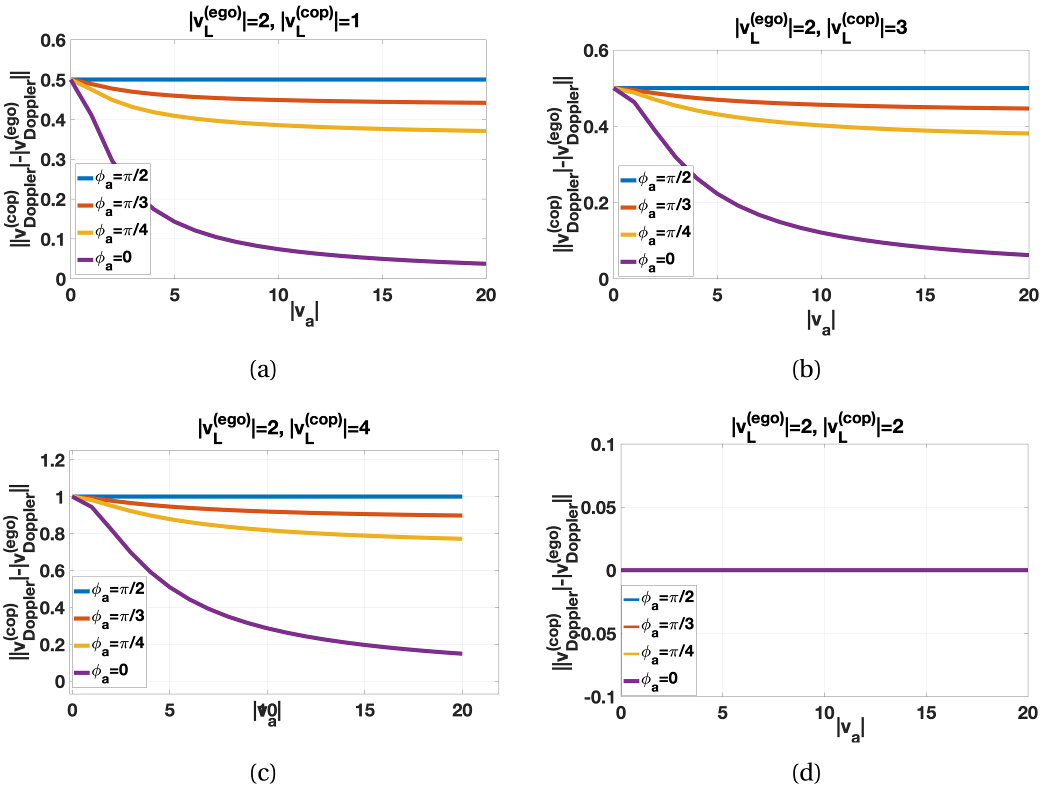


Figure 4.10: Cooperative LiDARs

The proposed spoofing detection can be better illustrated by the following numerical examples. Without loss of generality, we use the 2-LiDAR cooperative LiDAR system shown in Figure 4.10 as an example. The cooperative LiDAR system has one ego-LiDAR and one coop-LiDAR, and their velocities are denoted as $\vec{v}^{(cop)}_{L}$ and $\vec{v}^{(ego)}_{L}$, respectively. The Doppler velocities and ToF velocities computed by the two LiDARs for the same object are denoted as $\vec{v}^{(cop)}_{Dop}$, $\vec{v}^{(cop)}_{ToF}$ and $\vec{v}^{(ego)}_{Dop}$, $\vec{v}^{(ego)}_{ToF}$. For the attacker, denote its velocity by $\vec{v}_a = |\vec{v}_a|(\cos\phi_a, \sin\phi_a)$. The attacker sends spoofing signals $\mathbb{S}^{(spf)}_{ego}$ and $\mathbb{S}^{(spf)}_{cop}$ to ego-LiDAR and coop-LiDAR, respectively. And the ToF distances of $\mathbb{S}^{(spf)}_{ego}$ and $\mathbb{S}^{(spf)}_{cop}$ are designed according to Proposition 4.5.1 to maintain that for each LiDAR, the calculated Doppler velocity is consistent with the ToF velocity, i.e. $\vec{v}^{(cop)}_{Dop} = \vec{v}^{(cop)}_{ToF}$ and $\vec{v}^{(ego)}_{Dop} = \vec{v}^{(ego)}_{ToF}$.

The numerical results are shown in Figure 4.11. In each subfigure, the x axis denotes $|\vec{v}_a|$, which varies from 0 to 20 m/s. The y axis represents the difference between the magnitudes of the two Doppler velocities, i.e., $|\vec{v}^{(cop)}_{Dop}| - |\vec{v}^{(ego)}_{Dop}|$. Recall that $\vec{v}^{(cop)}_{Dop}$ and $\vec{v}^{(ego)}_{Dop}$ are 2-D vectors, therefore if $|\vec{v}^{(cop)}_{Dop}| - |\vec{v}^{(ego)}_{Dop}| \neq 0$, then we must have $\vec{v}^{(cop)}_{Dop} \neq \vec{v}^{(ego)}_{Dop}$. We plot $|\vec{v}^{(cop)}_{Dop}| - |\vec{v}^{(ego)}_{Dop}|$ as functions of $|\vec{v}_a|$ in different combinations of $|\vec{v}^{(ego)}_{L}|$ and $|\vec{v}^{(cop)}_{L}|$ in the four subfigures: (a)$|\vec{v}^{(ego)}_{L}| = 2\ m/s, |\vec{v}^{(cop)}_{L}| = 1\ m/s$. (b)$|\vec{v}^{(ego)}_{L}| = 2\ m/s, |\vec{v}^{(cop)}_{L}| = 3\ m/s$. (c)$|\vec{v}^{(ego)}_{L}| = 2\ m/s, |\vec{v}^{(cop)}_{L}| = 4\ m/s$. (d)$|\vec{v}^{(ego)}_{L}| = 2\ m/s, |\vec{v}^{(cop)}_{L}| = 2\ m/s$. In each subfigure, we also vary the angle of the attacker's velocity, i.e., ϕ_a, by setting $\phi_a = \frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{4}, 0$, respectively.



(a)

(b)

(c)

(d)

Figure 4.11: Numerical examples for spoofing detection when attacker controls its velocity.

In Figure (a), (b), and (c), the two LiDARs in the cooperative LiDAR system have different velocity magnitudes, i.e., $|\vec{v}_L^{(ego)}| \neq |\vec{v}_L^{(cop)}|$. Although the spoofing attack maintains that the ToF velocity is consistent with the Doppler velocity at each of LiDARs ($\vec{v}_{Dop}^{(cop)} = \vec{v}_{ToF}^{(cop)}$ and $\vec{v}_{Dop}^{(ego)} = \vec{v}_{ToF}^{(ego)}$), when the Doppler velocities are shared in the cooperative LiDAR system, ego-LiDAR and coop-LiDAR can detect spoofing attacks because $\vec{v}_{Dop}^{(cop)} \neq \vec{v}_{Dop}^{(ego)}$ ($|\vec{v}_{Dop}^{(cop)}| - |\vec{v}_{Dop}^{(ego)}| \neq 0$). A special case is shown in Figure (d), when the two LiDARs have the same velocity magnitude, that is, $|\vec{v}_L^{(cop)}| = |\vec{v}_L^{(ego)}|$, we have $|\vec{v}_L^{(cop)}| - |\vec{v}_L^{(ego)}| = 0$ even when the spoofing attack is in place. In this case, the cooperative LiDAR system cannot detect spoofing attacks based on the inconsistency between their respective Doppler velocities. Therefore, our spoofing detection scheme requires that the LiDARs in the cooperative LiDAR system have different velocities to successfully detect the spoofing attack.

4.5.3 Attack Model 3: Moving Attacker That Controls Both Its Velocity and Signal Frequency

A basic assumption in Attack Models 1 and 2 is that the attacker transmits spoofing signals of the same frequency as that of the victim LiDAR and it does not manipulate the frequency of the spoofing signal during the attack. Although this assumption is valid for many spoofing attack scenarios and has been adopted by many existing studies, e.g., [84, 19, 105], an attacker may use frequency modulation or a tunable laser source to dynamically change the frequency of the spoofing signal, so as to create a faked Doppler frequency shift to mislead those spoofing detection mechanisms proposed in the previous sections. This is elaborated as follows.

Spoofing Attack in Attack Model 3

When an attacker can dynamically adjust the frequency of the spoofing signal, besides sending spoofing ToF signals to the victim LiDAR, the attacker also compensates for the frequency offset caused by the Doppler effect by changing the frequency of the transmitted spoofing signal, making the frequency offset of the spoofing signal received by the victim LiDAR identical to the Doppler frequency shift of the legitimate signal.

Specifically, let us consider a typical spoofing attack scenario, where at the current moment the distance between the (victim) LiDAR and the attacker is d. The relative radial velocity between the victim LiDAR and the attacker is $\Delta v_a = (\vec{v}_L - \vec{v}_a) \cdot \vec{l}$, where $\vec{l}$ is the unit vector along the direction from the LiDAR to the attacker. The goal of the attacker is to create a fake object that is d' away from the LiDAR, in the same direction of $\vec{l}$ (so the LiDAR, the attacker, and the fake object are collinear) and of a relative radial velocity of Δv_{spoof}, where $\Delta v_{spoof} = (\vec{v}_L - \vec{v}_{spoof}) \cdot \vec{l}$, and $\vec{v}_{spoof}$ denotes the velocity of the fake object. With time continues, the trajectory of the faked object (i.e., d's) should be consistent with $\vec{v}_{spoof}$.

To achieve the attack goal, in the time domain, the attacker sends spoofing signals with faked ToF distance of d'. In the frequency domain, the attacker adjusts the frequency of the transmitted spoofing signal to mimic the Doppler shift experienced by a legitimate signal. Specifically, if a genuine object of velocity $\vec{v}_{spoof}$ is at the location of the fake object, then the Doppler shift experienced by a legitimate signal (this is the signal sent out by the LiDAR, reflected by the object, and then received by the LiDAR) is given by $\Delta f_r = \frac{2f_0}{c} \Delta v_{spoof}$, where f_0 is the frequency of the transmitted (legitimate) signal. Therefore, the frequency of the received legitimate signal is given by $f_0 + \Delta f_r$. To mimic the legitimate signal, the attacker chooses a frequency f_a for the transmitted spoofing signal, such that when the spoofing signal is received by the victim LiDAR, the frequency of the received spoofing signal is identical to that of the received legitimate signal. Since the spoofing signal is sent directly to the LiDAR, its Doppler shift is given by $\Delta f_a = \frac{f_a}{c} \Delta v_a$. So the frequency of the received spoofing signal is $f_a + \Delta f_a$. Therefore, the f_a that satisfies the aforementioned requirement is given by

$$f_a = \frac{c + 2\Delta v_{spoof}}{c + \Delta v_a} f_0. \tag{4.15}$$

In this way, the Doppler shift measured by the victim LiDAR happens to be Δf_r. As a result, the calculated Doppler velocity is consistent with the ToF velocity (both are equivalent to $\vec{v}_{spoof}$), and hence the fake object will be accepted by the LiDAR as a genuine one.

Spoofing Detection When Attacker Controls Signal Frequency

The cooperative LiDAR system can also be used for spoofing detection when the attacker controls its signal frequency. Specifically, according to Equation (4.15), the attacker must adjust the frequency of the transmitted signal each time when sending a spoofing signal to a LiDAR. When there exist multiple LiDARs with different velocities (so they have different Δv_a's and Δv_{spoof}'s), the attacker must choose different transmission frequencies when sending to different LiDARs to spoof each of them.

Based on this observation, we can use the cooperative LiDAR system and require all LiDARs in the system be synchronized to send probing laser pulses that will hit the object at the same time (and hence will be reflected by the object at the same time too), so that an attacker is not able to simultaneously change the frequency of spoofing signals for all LiDARs at once. The key point in achieving full synchronization among a group of cooperative LiDARs, i.e., making them point to the same object at the same time, is to realize that the first LiDAR that detects the object actually can compute and then communicate the location of that object to all other collalborating LiDARs, and hence allow all LiDARs in the group to compute their respective angles of departure for their laser beams in order for them to point to the same object.

The basic idea of using multiple LiDARs for spoofing detection is that an attacker can only send out a spoofing signal with a certain frequency at one time. Given that our Coop-LiDAR system synchronizes multiple LiDARs to monitor the same object at the same time, it is hard for an attacker to send a single spoofing signal that can simultaneously satisfy the frequency requirements from all LiDARs. In the case where the attacker has k coordinated dynamic-frequency laser transmitters, at least $k+1$ synchronized LiDARs are needed, so that at least one LiDAR is able to detect the spoofing by testing the inconsistency between its calculated Doppler velocity and ToF velocity. Note that here, the goal of the spoofing detection mechanism is to serve as a filter (a gate-keeper) that identifies and rejects spoofed LiDAR sensing outcomes. Therefore, a collective decision-making process is adopted among all $(k+1)$ LiDARs: a sensed point in the point cloud will be accepted only if none of the $k+1$ synchronized LiDARs has a negative detection outcome.

4.5.4 Limitations

Although in previous subsections we have demonstrated that the Doppler-shift-based method is effective for detecting spoofing attacks across various real-world attack scenarios, there still remain some scenarios where our method may be less effective or not suitable, as elaborated below:

1. Static or low relative velocities scenarios: Doppler shift is the change of signal frequency due to the movement of the transmitter in relative to the receiver. In the LiDAR case, if the relative velocity between the LiDAR and the sensed object is 0 or close to 0, then the Doppler shift will be negligible. In these scenarios, our method is not applicable.

2. Large velocity variation during small time interval scenarios: A basic assumption in our attack models 2 and 3 is that the relative velocity between the LiDAR and the object remains constant between the moments of two consecutive LiDAR measurements (usually this is over the span of a fraction of a second), so that our proposed algorithm is able to resolve the Doppler velocity and the ToF velocity of the object. While this assumption is valid in most cases, in reality there are special situations where the relative velocity between the LiDAR and the object changes significantly during the aforementioned small interval. Such changes in velocity could be caused by, e.g., a bumpy road condition, or a complicated traffic condition that requires frequent maneuvers (e.g., sudden acceleration, deceleration, or braking) of the car. In these special situations, the accuracy of the proposed method will be reduced. To deal with this issue, in Section 4.6, we have proposed a statistical spoofing detection scheme, which accounts for the short-term variation/perturbation in the vehicle's velocity. However, the proposed statistical detection scheme still faces limitations as it is based on certain assumed statistical models (i.e., the distribution) for the velocity variation. In the real-world scenario, if the actual velocity variation deviates significantly from the assumed distribution, then the accuracy of this statistical scheme will be reduced. In this case, a combination of our method with existing model-level defense methods would be a good solution. As model-level defense methods utilize high-level contextual relationships between multiple data points for spoofing detection, they well compensate for the limitations of the Doppler shift-based method that works only at the physical layer.

We want to clarify that our proposed spoofing detection method is not a panacea - a "solution to all" that intends to replace existing methods. Instead, it serves as the "first line of defense" that operates in the signal space and is designed to complement existing model-level defense methods. Our method uses the physical property of an individual data point within the point-cloud for spoofing attack detection, which is a validation in the signal space to check whether the signatures (Doppler shift) of the signal follow physical principles. Because of its physical feature, our proposed method can fundamentally ensure that the LiDAR sensing results that are fed to the subsequent high-level processing are authentic. In contrast, current perception models-level defense methods work at a higher level: They first aggregate multiple data points to establish a geometric representation for the sensed object, and then examine whether this geometric representation presents a reasonable contextual consistency over time. It is clear that our method works in an orthogonal space compared to these model-level defense methods. In practice, both methods can be applied at the same time to improve the overall detection accuracy against LiDAR spoofing attacks.

4.6 Spoofing Detection with Joint Consideration of Velocity and Acceleration

In the previous section, we assumed a uniform motion model, so that the relative velocity between the LiDAR and the object can be seen as constant. And we propose to verify the consistency between the ToF velocity and the Doppler velocity for spoofing attack detection. Although, due to the high scanning rate of LiDAR, the motion of an object with acceleration can be seen as a uniform motion, the presence of acceleration introduces additional variance in velocity estimation, which makes spoofing detection based only on velocity unreliable. In this section, we present a hypothesis-test-based spoofing detection framework that jointly considers velocity and acceleration. We first formulate the hypotheses for the attack and non-attack cases on the basis of our previous findings. Then, we demonstrate the necessity to jointly consider acceleration and velocity for spoofing detection and provide the test statistic designing strategies. Finally, we perform power analysis under various conditions and numerically determine the smallest test sample size required to achieve an expected performance level.

4.6.1 Hypothesis Test Formulation

According to our discussion in the previous section, the velocity of an object can be obtained based on the Doppler shift or ToF of the received signal, namely $\vec{v}_{Dop}$ and $\vec{v}_{ToF}$. The inconsistency between the two velocities, $\vec{v}_{Dop}$ and $\vec{v}_{ToF}$, can only be caused by spoofing attacks or noise. Consider a sequence of n Doppler and ToF velocity samples $\{\vec{v}_{Dop}\}_n$ and $\{\vec{v}_{ToF}\}_n$, respectively. For convenience, let v_{Dop} and v_{ToF} denote the magnitudes of $\vec{v}_{Dop}$ and $\vec{v}_{ToF}$, respectively. And their population means are denoted by μ_{Dop} and μ_{ToF}, respectively. The spoofing detection can be formulated as a hypothesis test, which essentially tests whether the two means are equal or not, that is, $\mu_{Dop} \overset{?}{=} \mu_{ToF}$. The null and alternative hypotheses can be formulated as follows:

$$\begin{aligned} &\mathscr{H}_0 : \text{no spoofing attack.}(\mu_{Dop} = \mu_{ToF}) \\ &\mathscr{H}_a : \text{the presence of a spoofing attack.}(\mu_{Dop} \neq \mu_{ToF}). \end{aligned} \tag{4.16}$$

When only velocity is taken into account for spoofing detection, the two-sample t-test is used. The test statistic is calculated as

$$t = \frac{|\mu_{Dop} - \mu_{ToF}|}{S_{pooled}\sqrt{2/n}}, \tag{4.17}$$

where $S_{pooled} = \frac{s_1^2+s_2^2}{2}$, and s_1^2 and s_2^2 are the sample variances of $v_{Dop}$and v_{ToF}, respectively.

Then t is compared with the critical value with the degree of freedom of $n-1$ and the significance level α, $t_{n-1}(\alpha/2)$. Hypothesis $\mathscr{H}_0$ is rejected if $t > t_{n-1}(\alpha/2)$, which indicates a spoofing attack.

4.6.2 Joint Consideration of Velocity and Acceleration.

In real driving scenarios, the AV's motion not only has velocity but also has acceleration. Such an acceleration could lead to a broadening spectrum in the Doppler frequency, which increases the variance in velocity estimations derived from the Doppler shift spectrum. This variance becomes more significant for the small velocity and large acceleration cases. For example, suppose that we have $v = 0.5\, m/s$ and $a = 0.5\, m/s^2$, the Doppler spectrum of the received signals is likely to display two dominant peaks at velocities of $0.5\, m/s$ and $1\, m/s$. This phenomenon can lead to ambiguity in velocity estimation, with potential values ranging between $0.5\, m/s$ or $1\, m/s$, thus introducing a maximal error of $0.5\, m/s$. Hence, when acceleration exists, it increases the risk of misidentifying a legitimate signal as a spoofing attack, resulting in an increased false alarm rate in spoofing attack detection. Realizing the limitation of considering velocity alone in spoofing attack detection, we introduce an advanced detection mechanism that jointly incorporates the effect of both velocity and acceleration, which can provide more robust and accurate results in identifying spoofing attacks in realistic driving scenarios.

Let a denote the acceleration and let $\boldsymbol{x} = [v, a]$ denote the multivariate variable that consists of both the velocity v and the acceleration a, which is used for the hypothesis test. We first use maximum likelihood estimation (MLE) to estimate v_{ToF} and a_{ToF}. Let $\boldsymbol{\mu}_{Dop}, \boldsymbol{\Sigma}_{Dop}$ and $\boldsymbol{\mu}_{ToF}, \boldsymbol{\Sigma}_{ToF}$ denote the mean and variance of the population for $\bar{\boldsymbol{x}}_{\boldsymbol{Dop}}$ and $\bar{\boldsymbol{x}}_{ToF}$, respectively. We assume that $\{\boldsymbol{x}_{Dop}\}_n$ is a random sample of size n from the normal distribution $\mathcal{N}(\boldsymbol{\mu}_{Dop}, \boldsymbol{\Sigma}_{Dop})$ and $\{\boldsymbol{x}_{ToF}\}_n$ is a random sample of size n from normal distribution $\mathcal{N}(\boldsymbol{\mu}_{ToF}, \boldsymbol{\Sigma}_{ToF})$. Note that $\bar{\boldsymbol{x}}_{Dop} - \bar{\boldsymbol{x}}_{ToF}$ follows the normal distribution $\mathcal{N}(\boldsymbol{\mu}_{Dop} - \boldsymbol{\mu}_{ToF}, \frac{1}{n}(\boldsymbol{\Sigma}_{Dop} + \boldsymbol{\Sigma}_{ToF}))$. Therefore, the hypothesis test is simplified accordingly to test if $\boldsymbol{\mu}_{Dop} = \boldsymbol{\mu}_{ToF}$ or not, and *Hotelling's* T^2 test is used, whose test statistic is

$$
\begin{aligned}
T_0^2 =& [\bar{\boldsymbol{x}}_{Dop} - \bar{\boldsymbol{x}}_{ToF} - (\boldsymbol{\mu}_{Dop} - \boldsymbol{\mu}_{ToF})]'[\frac{2}{n} S_{pooled}]^{-1} \\
& [\bar{\boldsymbol{x}}_{Dop} - \bar{\boldsymbol{x}}_{ToF} - (\boldsymbol{\mu}_{Dop} - \boldsymbol{\mu}_{ToF})],
\end{aligned}
\tag{4.18}
$$

where $S_{pooled} = \frac{s_1^2+s_2^2}{2}$, and $\bar{\boldsymbol{x}}_{Dop}$ and $\bar{\boldsymbol{x}}_{ToF}$, and s_1^2 and s_2^2 are the sample mean and sample variance of $\boldsymbol{x}_{Dop}$ and $\boldsymbol{x}_{ToF}$, respectively. T_0^2 follows a non-central F distribution $\frac{4n-4}{2n-3}F_{2,2n-3}$. For a given significance level α, the critical value τ is calculated as

$$
\tau = \frac{4n-4}{2n-3} F_{2,2n-3}(\alpha). \tag{4.19}
$$

The null hypothesis $\mathscr{H}_0$ is rejected when $T_0^2 > \tau$, and the false alarm rate, a.k.a., type I error, is:

$$
P_{\mathscr{H}_0}(T_0^2 > \tau) = \alpha. \tag{4.20}
$$

4.6.3 Formulation of $\mathscr{H}_a$ for Power Analysis

Fig. 4.12 illustrates the power and significance level of a statistical test. Previously, we have determined the distribution of $\mathscr{H}_0$ and the critical value. Next, we must ensure that the test has enough power so that the distribution of $\mathscr{H}_0$ and that of $\mathscr{H}_a$ are sufficiently apart and both type I and type II errors are small. The power of a hypothesis test is the probability that the test correctly rejects the null hypothesis, as illustrated by the red dashed area. It should be noted that statistical power is positively related to the sample size. The larger the sample size, the easier it is to achieve the expected statistical power. There are two possible cases where one fails to reject the null hypothesis: (1) The null hypothesis is really true. (2) The sample size is not large enough to reject the null hypothesis (i.e., statistical power is too low). Additional samples may be needed to either accept or reject the null hypothesis.

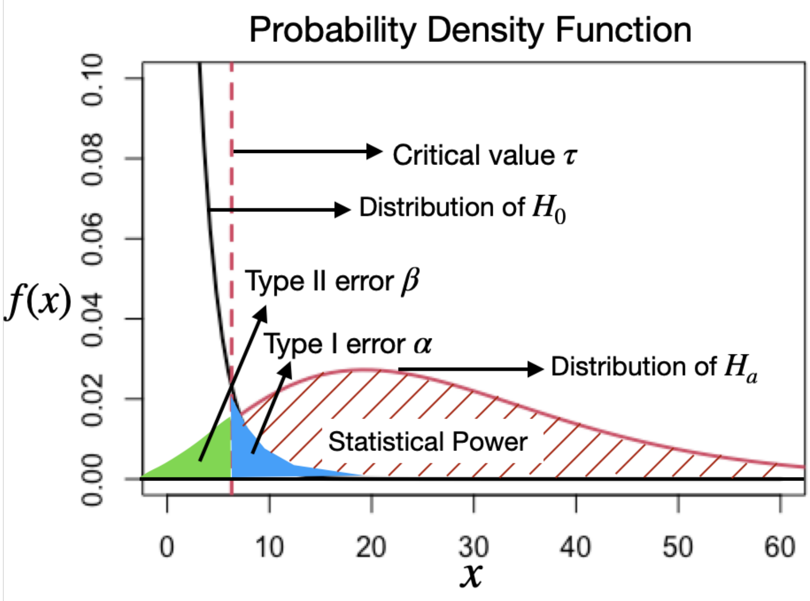


Figure 4.12: Illustration of a Statistical Test.

Now, we will design scenarios of $\mathscr{H}_a$, under which the power analysis can be performed to determine the smallest sample size required to achieve a satisfactory detection performance. When designing $\mathscr{H}_a$, it is impossible to enumerate all possibilities. In fact, the detector is not designed to identify every malicious attack, but rather to identify spoofing attacks that can lead to severe consequences. Specifically, in our study, we focus on two attack goals: (1) **emergency brake** triggered by injecting a fake static object in front of the LiDAR; (2) **failure of the automatic braking system** by injecting a fake object that is relatively stationary to the LiDAR. Specifically, we consider a scenario where the AV is fast-moving towards a static real obstacle, and a brake decision is required to avoid a collision. Note that the braking decision of the AV system is based on the combination consideration of the distance and the relative speed between the AV and the object. Therefore, the attacker launches the attack by sending spoofing signals that mimic a fake object in close range (so the faked signal will be the first to arrive at the AV's LiDAR than that of the real object) to hide the real obstacle from LiDAR detection and is relatively stationary to the AV. Although the distance between the fake object and the AV is small, due to the small relative speed between them, the AV's decision-making system will not trigger a braking decision, as it perceives no immediate collision risk. Consequently, the AV might continue at its current speed and collide with the real obstacle. In addition to the above attack goals, we also consider the attacker to be static or mobile and design three attack scenarios, in which we provide the distribution of $\mathscr{H}_a$.

Attack Scenario 1: an emergency brake triggered by a static attacker

A static attacker wants to trigger an emergency brake by faking a static object in front of the LiDAR. Because both the attacker and the fake object are static, according to Eq. 4.5 and Eq. 4.6, we have $\mathscr{H}_a: \boldsymbol{\mu}_{Dop} = 2\boldsymbol{\mu}_{ToF}$. The test statistic under $\mathscr{H}_a$ is:

$$\begin{aligned} T_a^2 =& [\bar{\boldsymbol{x}}_{Dop} - 2\bar{\boldsymbol{x}}_{ToF} - (\boldsymbol{\mu}_{Dop} - 2\boldsymbol{\mu}_{ToF})]'[\frac{5}{n}S_{pooled}]^{-1} \\ & [\bar{\boldsymbol{x}}_{Dop} - 2\bar{\boldsymbol{x}}_{ToF} - (\boldsymbol{\mu}_{Dop} - 2\boldsymbol{\mu}_{ToF})]. \end{aligned} \tag{4.21}$$

According to [53], the test statistic T_a^2 follows a non-central F distribution $\frac{25(n-1)}{5n-6}F_{2,5n-6}$ with a *non-centrality parameter* ($n.c.p.$) equal to

$$n.c.p. = \frac{n}{\sigma^2}\left[(\bar{\boldsymbol{x}}_{Dop} - \bar{\boldsymbol{\mu}})'(\bar{\boldsymbol{x}}_{Dop} - \bar{\boldsymbol{\mu}}) + (2\bar{\boldsymbol{x}}_{ToF} - \bar{\boldsymbol{\mu}})'(2\bar{\boldsymbol{x}}_{ToF} - \bar{\boldsymbol{\mu}})\right], \tag{4.22}$$

where $\bar{\boldsymbol{\mu}} = \frac{(2\boldsymbol{\mu}_{Dop} + \boldsymbol{\mu}_{ToF})}{2}$ and σ^2 is the mean square error. Given a critical value τ, the type II error is represented as:

$$P_{\mathscr{H}_a}(T_a^2 < \tau) = \beta. \tag{4.23}$$

Attack Scenario 2: an emergency brake triggered by a moving attacker

The attacker is moving at the same speed as the victim LiDAR and wants to trigger an emergency brake by faking a static object in front of the victim LiDAR. In this case, we have $\mathscr{H}_a: \boldsymbol{\mu}_{Dop} = 0$ and $\boldsymbol{\mu}_{ToF} \neq 0$. The test statistic under $\mathscr{H}_a$ is:

$$T_a^2 = [\bar{\boldsymbol{x}}_{Dop} - \bar{\boldsymbol{x}}_{ToF} + \boldsymbol{\mu}_{ToF})]'[\frac{2}{n}S_{pooled}]^{-1}[\bar{\boldsymbol{x}}_{Dop} - \bar{\boldsymbol{x}}_{ToF} + \boldsymbol{\mu}_{ToF})] \tag{4.24}$$

$$\text{with } n.c.p. = \tag{4.25}$$

$$\frac{n}{\sigma^2}[(\bar{\boldsymbol{x}}_{Dop} - \boldsymbol{\mu}_{ToF})'(\bar{\boldsymbol{x}}_{Dop} - \boldsymbol{\mu}_{ToF}) + (\bar{\boldsymbol{x}}_{ToF} - \boldsymbol{\mu}_{ToF})'(\bar{\boldsymbol{x}}_{ToF} - \boldsymbol{\mu}_{ToF})],$$

which follows $\frac{4n-4}{2n-3}F_{2,2n-3}$.

Attack Scenario 3: Failure of an automatic braking system triggered by a static attacker

The attacker is static and wants to trigger a failure of the automatic braking system of an AV. The attacker sends spoofing signals that mimic a fake object in close range and is relatively stationary to the AV. In this case, we have $\mathscr{H}_a: \boldsymbol{\mu}_{Dop} \neq 0$ and $\boldsymbol{\mu}_{ToF} = 0$. The test statistic under $\mathscr{H}_a$ is:

$$T_a^2 = [\bar{\boldsymbol{x}}_{Dop} - \bar{\boldsymbol{x}}_{ToF} - \boldsymbol{\mu}_{Dop})]'[\frac{2}{n}S_{pooled}]^{-1}[\bar{\boldsymbol{x}}_{Dop} - \bar{\boldsymbol{x}}_{ToF} - \boldsymbol{\mu}_{Dop})] \tag{4.26}$$

$$\text{with } n.c.p. = \tag{4.27}$$

$$\frac{n}{\sigma^2}[(\bar{\boldsymbol{x}}_{Dop} - \boldsymbol{\mu}_{Dop})'(\bar{\boldsymbol{x}}_{Dop} - \boldsymbol{\mu}_{Dop}) + (\bar{\boldsymbol{x}}_{ToF} - \boldsymbol{\mu}_{Dop})'(\bar{\boldsymbol{x}}_{ToF} - \boldsymbol{\mu}_{Dop})],$$

which follows $\frac{4n-4}{2n-3}F_{2,2n-3}$.

4.6.4 Settings for Power Analysis

As mentioned above, the sample size should be large enough to provide the expected statistical power. As a result, both the type I error α from $P_{\mathscr{H}_0}(T_0^2 > \tau) = \alpha$ and the type II error from

$P_{\mathscr{H}_a}(T_a^2 < \tau) = \beta$ are controlled in acceptable ranges. Analysis is carried out in combinations of road conditions, spoofed signal proportion, signal SNR, and attack scenarios to determine the minimum sample size required for the detector to produce satisfactory results for the most practical $\boldsymbol{x}$.

Road condition

Three typical road conditions are considered: (1) highway driving ($v = 33\,m/s$ and $a = 0.5\,m/s^2$); (2) ramp driving ($v = 20\,m/s$ and $a = 1.5\,m/s^2$); (3) city driving ($v = 11\,m/s$ and $a = 5\,m/s^2$). We note that the relative speed implies the distance between the LiDAR and the object. A low relative speed indicates a smooth driving condition, under which any attack can be easily detected due to the sudden change in speed measurements. Rather, a high relative speed may indicate that an abnormal traffic condition is already in place, making the attack less effective. Therefore, we set the relative speed of the victim LiDAR to be 50% of that of each road condition to balance between the difficulty of detection and the consequence of the attack.

Spoofed signal proportion

The high LiDAR sampling rate and the narrow receiver's field-of-view impose stringent constraints on the timing and direction of the spoofed signal. In practice, the attacker hardly has the luxury of continuously spoofing a sequence of signals [141, 101]. It is more practical that the attacker spoofs the LiDAR signals intermittently. The spoofed signal proportion is defined as the ratio of the number of spoofing signal samples to the number of received signal samples. The higher the ratio of the spoofed signal, the easier the attack is detected. In the experiment, we consider the range of the proportion of the spoofed signal p to be 0.1 to 1.

Signal-to-noise ratio (SNR)

The noise level of the signals is affected by weather conditions, ambient light, system error, device noise, etc. Such noises would introduce errors in the velocity estimated from both the

ToF and Doppler shift, and we discuss them separately. Considering the LiDAR measurement error [110] and the disturbance of ambient light, we set the error rate of both a_{ToF} and v_{ToF} to 3%. For the measurement error in Doppler velocity, we follow [2] to calculate the variance of $\boldsymbol{x_{Dop}}$ of MLE:

$$\sigma_v^2 = \frac{1}{SNR}\frac{3}{2\pi^2 N^2}, \sigma_a^2 = \frac{1}{SNR}\frac{45}{2\pi^2 N^4} \tag{4.28}$$

where N is the sampling length of the signal, which is set to 256 in our simulation to trade-off the estimation accuracy and the system burden. The SNR is set to $\{10^{-6}, 10^{-5}, 10^{-4}\}$ according to [70] to fit the real-world scenarios.

4.6.5 Numerical Results of Power Analysis

In our simulation, we follow the convention to set the type I error to $\alpha = 0.05$, and record the least number of samples to achieve the power of 0.9 at each $\mathscr{H}_a$, i.e., type II error is $\beta = 0.1$. The $F1$ score in this setting is 0.923, indicating satisfactory spoofing detection performance. In real-world application scenarios, the type I and type II error settings can be set differently to meet different practical requirements.

Impact of spoofing signal proportion

We set the SNR to 10^{-4} and vary the proportion of the spoofed signal from 0.1 to 1. We record the least number of samples needed to achieve the pre-set significance level under different attack scenarios and road conditions. The results are shown in Fig. 4.13. It can be seen that more samples are needed when the proportion of the spoofed signal is small. When the attacker spoofs only a small proportion of the LiDAR signals, the mean of $\boldsymbol{x}_{ToF}$ is close to that of $\boldsymbol{x}_{Dop}$, therefore, more samples are needed to separate the two distributions. This phenomenon becomes more obvious when the proportion of the spoofed signal is less than 40%, especially for attack 1.

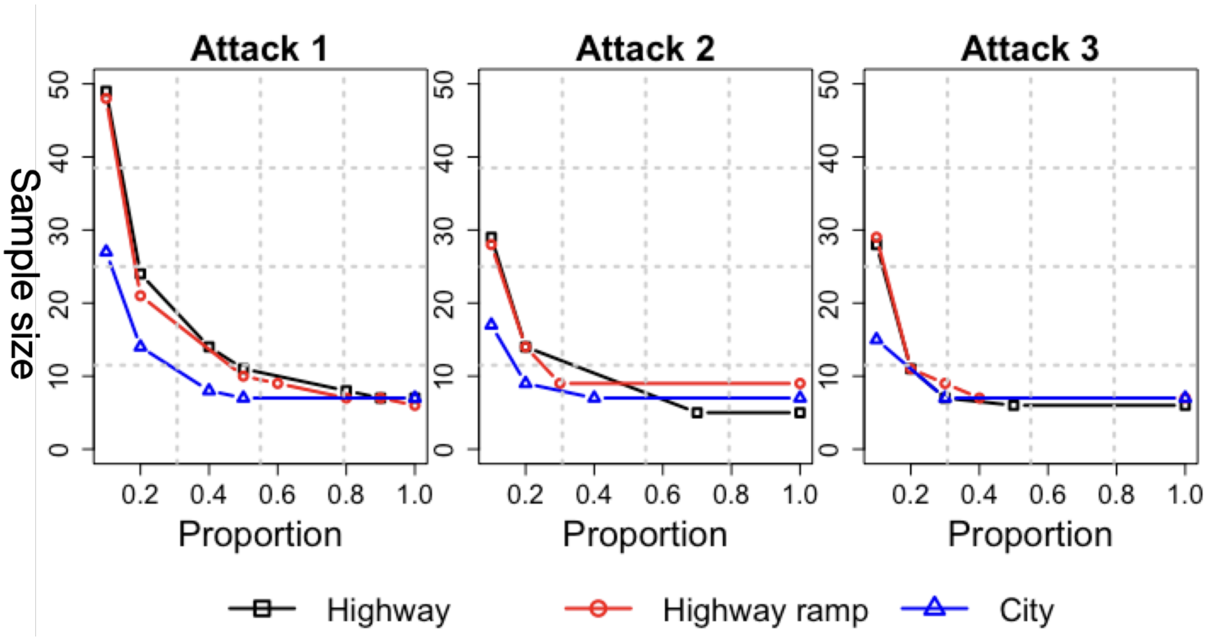


Figure 4.13: The number of samples needed for each road conditions under different spoofed signal proportion.

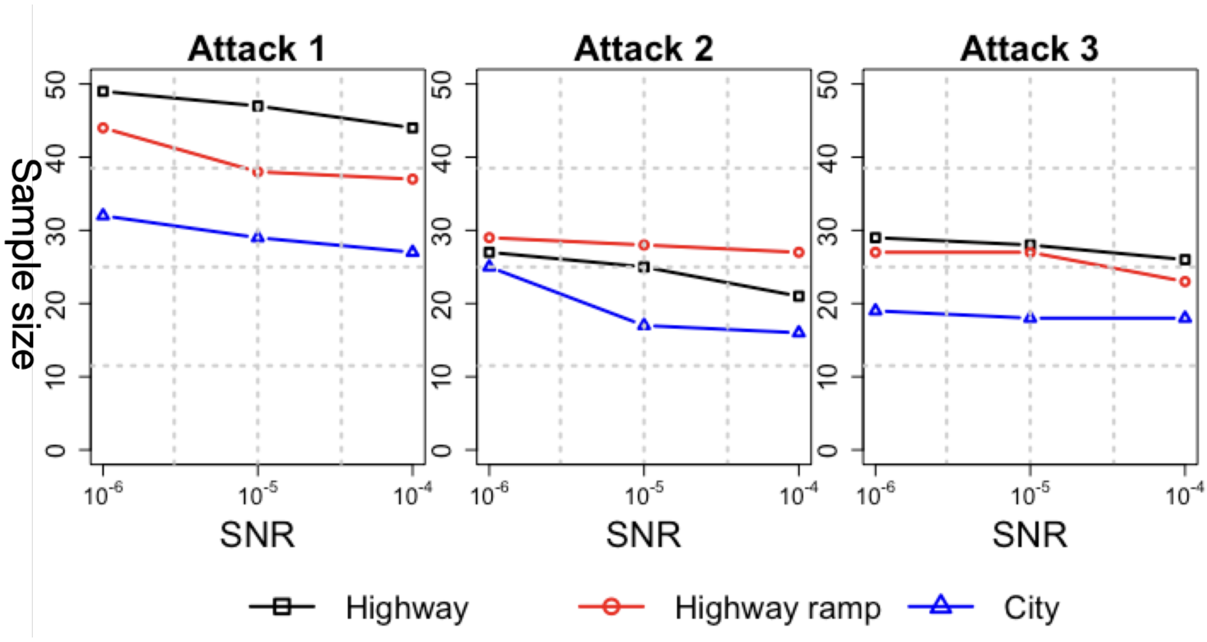


Figure 4.14: The number of samples needed for each road conditions under different SNR.

Impact of signal-to-noise ratio

Then, we fix the proportion of the spoofed signal to 0.1 and vary the SNR from 10^{-6} to 10^{-4}. The minimal sample sizes needed to provide the expected statistical power under various attack scenarios and road conditions are shown in Fig. 4.14. Compared to the spoofed signal proportion that has a greater impact on the difference between the means of $\{\boldsymbol{x}_{ToF}\}_n$ and $\{\boldsymbol{x}_{Dop}\}_n$, the SNR plays a more significant role in affecting their variance. A smaller SNR leads to a larger estimation variance in the velocity and acceleration from the Doppler shift, which increases the uncertainty in detecting spoofing attacks. As a result, more test samples are needed to provide sufficient statistical power.

4.6.6 Discussion on Implementation

After determining the number of samples required, the spoofing attack detection procedure is carried out in the following two steps: (1) Data collection: Assuming that the sample size is 50, 50 samples of $\{\boldsymbol{x}_{dop}\}$ and $\{\boldsymbol{x}_{ToF}\}$ are collected, respectively. (2) Testing: The test statistic is calculated according to Eq. 4.18, then compared with the threshold τ predefined by Eq. 4.19. If the test statistic is greater than the threshold, it suggests the potential presence of a spoofing attack. According to the analysis above, setting the sample size to 50 is sufficient for our test to achieve an $F1$ score of 0.923 in the worst-case scenario. Notably, with sample size of 50, the $F1$-score would be even higher for the remaining cases. For example, under conditions where 40% of the signals are spoofed and the SNR is 10^{-4}, the test produces an impressive $F1$ score of 0.97 in all road conditions.

We then evaluate the time complexity of the proposed method by examining the latency associated with each step above. In the testing step, the calculation of the test statistic directly from the data and the comparison with the predefined threshold incurs negligible time overhead. In the data collection phase, considering a typical 16-beam Velodyne LiDAR system with a rotation speed of 20 Hz [123], 50 samples can be collected in 150 ms. This duration is significantly shorter than the average reaction time of 830 ms for autonomous vehicles [34]. Note that for more advanced AV LiDAR systems with a higher number of laser beams and a faster rotation speed, the data collection time can be further reduced. As a result, our proposed spoofing attack detection mechanism can operate simultaneously with established LiDAR processing algorithms, enhancing the reliability of current AV driving systems without introducing additional time overhead.

4.7 Applicability Discussion, Future Works , and Conclusion

4.7.1 Applicability Discussion

In this section, we discuss the applicability of the proposed method to other sensors, such as cameras and radars. The primary focus of this paper is on addressing the unique problem of safeguarding against LiDAR spoofing attacks, which is distinctive due to the special way of how a LiDAR sensor detects an object and its distance to that object. Therefore, our proposed method cannot be applied to cameras, as cameras lack the capability to measure the Doppler shift of incoming light signals. Specifically, cameras are passive sensors that record natural radiation either emitted or reflected from objects. The resulting signal is represented in terms of pixel intensity and color, and cameras cannot capture any frequency changes in these light signals. As for radars, our proposed method can be used for spoofing attack detection but requires adaptations to address the challenges inherent to radar systems. Notably, while radars are also active sensors and can directly measure the Doppler shift of incoming signals, they present unique challenges when compared to LiDARs. For example, radars typically offer lower spatial resolution and emit signals with a larger spectral bandwidth. This means that the received signal can be influenced by the Doppler effect from several objects simultaneously, each contributing different Doppler frequency shift components. Additionally, the broad spectral bandwidth of radar signals can reduce the precision of Doppler frequency shift measurements. This complexity heightens the challenge of pinpointing spoofing attacks based solely on the Doppler shift and potentially increasing the false positive rate of our proposed method when being applied to radars.

4.7.2 Future Work

We understand that testing our method in a real-world setup, such as on a real autonomous vehicle, would significantly improve the impact and practical relevance of our method. However, as a research lab in a university, we are not capable of fully implementing the proposed methods on a real LiDAR system (note that nearly all LiDAR systems on the market are proprietary and are not open to redevelopment) and then mounting it on a vehicle to perform real-world testing. Realistically, what our capacity allows us to do is the theoretical study of the mathematical models for the spoofing attacks and their detection, and mainly computer-simulation-based performance evaluation for the proposed models. The scope of the paper has to be decided by our capacity above. We acknowledge that there must exist a significant difference between our work and a real-world system that can be directly used by the current autonomous driving vehicles. However, our contribution in this paper is mainly on the modeling aspect of the problem rather than on the system-building/implementation of the model. The theoretical foundation laid in this work could serve as an important reference/guideline for system implementation in the next step, which is out of the scope of this paper and may be conducted in our future work.

4.7.3 Conclusion

In this paper, we investigated the LiDAR security problem in the autonomous driving system. We performed a detailed analysis on the vulnerability of the LiDAR sensors. To better illustrate how to use Doppler shift for spoofing attack detection in different attack scenarios, we considered three attack models, including static attacker, moving attacker without/with control of velocity, and moving attacker with control of both velocity and signal frequency. Under each of these models, we first show how the spoofing attack is performed, and then present our proposed countermeasures. To address the uncertainty caused by vehicle acceleration, we proposed a statistical spoofing detection framework to jointly consider the impact of acceleration on vehicle velocity. Extensive numerical evaluations are conducted to verify the effectiveness and accuracy of the proposed methods in a wide range of test settings.

Chapter 5

Conclusion and Future Work

5.1 Dissertation Conclusion

In this dissertation, we have focused on enhancing the efficiency, reliability, and security of networks and systems operating within the mmWave and beyond frequency bands. Our focus has been on leveraging environmental and contextual network information to develop intelligent, cross-layer optimization strategies. Three works are presented. In the first work, we developed an environment perception-based smart beam switching method for the commercial off-the-shelf (COTS) mmWave product. In the second work, we investigate a network topology optimization problem for RIS-assisted mmWave directional communication networks. In the third work, we propose a physical layer spoofing detection method to fundamentally protect the sensing data of vehicle LiDAR systems from malicious attacks.

By addressing these critical aspects in the high-frequency band, this dissertation not only contributes to the existing body of knowledge but also opens up new avenues for future research in wireless network optimization. We hope that the insights and algorithms presented in this dissertation will inspire further innovation and exploration in the field, leading to the development of more advanced and intelligent wireless network systems in the near future.

5.2 Future Work

5.2.1 Multi-modal Sensing with mmWave Radar

With the rapid development of the Internet of Things (IoT) and the rise of 5G communication networks, mmWave sensing is emerging and beginning to impact our daily lives. Due to the high frequency of the mmWave signal, it is capable of providing high sensing sensitivity and precision [51, 146]. In addition, the short wavelength of mmWave signals further enables antennas to be highly integrated in a small space, enabling beamforming and other techniques that support directional sensing capabilities. As a result, mmWave sensing has great potential in human subtle motion sensing over low-frequency sensing technologies such as Wi-Fi, UWB, and LoRa [145].

The first work explores the sensing potential of COTS mmWave radio, following this idea, my future plan is to explore and maximize the extraordinary sensing potential of mmWave signals by focusing on the development of accurate and robust applications for non-intrusive human activity recognition. Traditional mmWave radar applications have been largely confined to macro-motion identification, such as detecting and tracking large objects [151] or monitoring vehicle movements [126]. However, the ability of mmWave signals to provide millimeter-level accuracy presents a largely untapped opportunity for more refined, fine-grained sensing applications. For example, nuanced gesture recognition [148], vital signs monitoring, and even extending into the realms of speech recognition and eavesdropping. To achieve this goal, a feasible method includes the integration of various types of sensory data. By combining visual data (images), mmWave signals, and acoustic inputs (sound signals), and employing advanced multimodal machine learning models [43, 73], such as Transformer [122], we are able to create a more holistic, accurate, and versatile framework for human activity recognition. The fusion of these diverse data types with multimodal machine learning algorithms is expected to yield insights and capabilities beyond the current scope of individual sensing modalities and make a significant leap forward in the field of smart sensing technology.

5.2.2 UAV-Assisted Reconfigurable Intelligent Surface Deployment

RIS serves as a key technology for next-generation wireless networks, which are made of low-cost metasurfaces that possess the ability to manipulate the propagation of a signal by reflecting or refracting the signal. With massive RIS being deployed, we can easily manipulate the wireless propagation environment according to different working scenarios. In the second work, we propose a novel coverage model called (k,α)-coverage to characterize the impact of path direction differences on path availability for the RIS-aided mmWave network. With the (k,α)-coverage, the system is able to robustness for the path blockage.

In our second work, we considered the static RIS deployment scenario, where RIS are assumed to be deployed in fixed positions to provide the area (k,α)-coverage . Yet, the dynamic deployment scenario stands out as a more desirable approach, particularly with the integration of Unmanned Aerial Vehicles (UAVs). In recent decades, UAV systems have gained considerable attention due to their ability to hover across the area. UAVs, also called drones, can travel to areas that lack infrastructure and are inaccessible to humans. The UAV-assisted RIS network can provide dynamic coverage for the area according to the requirement. A possible research direction is to integrate the dynamic feature of the UAV with the (k,α)-coverage model to provide dynamic (k,α)-coverage. In this case, a joint optimization problem should be considered to optimize the UAV trajectory, RIS phase shift, package scheduling, and power consumption while maintaining area (k,α)-coverage.

5.2.3 Autonomous Vehicle Cooperative Lidar Security

The rapid advancement in autonomous vehicle (AV) technology has created opportunities for smart urban mobility. A typical prediction of the future of autonomous vehicles includes people being relieved from the stress of daily commute driving, which is expected to be achieved by replacing imperfect human drivers with better computer-based autopilots. But how to achieve such fully autonomous vehicles while maintaining high driving safety is still a challenge.

In the third work, a physical layer spoofing attack detection method has been proposed to protect the trustworthiness of a standalone AV LiDAR sensing result. Recently, with the advancement in wireless communication, cooperation can be realized between different vehicles equipped with LiDAR, which can form a cooperative LiDAR system and is expected to achieve synergistic gains in LiDAR sensing performance. In a cooperative LiDAR system, decentralized federated learning(DFL) [142] is a good fit to exchange perception information. DFL enables direct communication between clients, resulting in significant savings in communication resources. In addition, DFL requires only model updates, not the raw data, to be shared among participants, which can protect the privacy of participants. However, recent research revealed that these shared model updates could be exploited to infer sensitive user data, thereby compromising user privacy. A promising research direction in this domain is to enhance privacy in DFL-assisted cooperative LiDAR systems. Addressing this challenge requires a balanced focus on both the security of data and the efficiency of data exchange. A feasible solution could involve the development of sophisticated data masking techniques that combine the geographic location of cooperative AV, which can minimize communication overhead while ensuring data security.

References

[1] 5G Americas. Understanding mmwave for 5g networks. Technical report, 5G Americas, 12 2020. Accessed: 2024-03-06.

[2] T. J. Abatzoglou. Fast maximum likelihood joint estimation of frequency and frequency rate. *IEEE Transactions on Aerospace and Electronic Systems*, AES-22(6):708–715, 1986.

[3] N. V. Abhishek, M. N. Aman, T. J. Lim, and B. Sikdar. Drive: Detecting malicious roadside units in the internet of vehicles with low latency data integrity. *IEEE Internet of Things Journal*, 9(5):3270–3281, 2021.

[4] M. Agiwal, A. Roy, and N. Saxena. Next generation 5g wireless networks: A comprehensive survey. *IEEE communications surveys & tutorials*, 18(3):1617–1655, 2016.

[5] N. Ahmed, S. Kanhere, and S. Jha. Probabilistic coverage in wireless sensor networks. In *The IEEE Conference on Local Computer Networks 30th Anniversary (LCN'05)*, pages 8 pp.–681, 2005.

[6] N. Al-Falahy and O. Y. Alani. Technologies for 5g networks: Challenges and opportunities. *It Professional*, 19(1):12–20, 2017.

[7] Q. An, Y. Shi, and Y. Zhou. Reconfigurable intelligent surface assisted non-orthogonal unicast and broadcast transmission. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, pages 1–5, 2020.

[8] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang. What will 5g be? *IEEE Journal on selected areas in communications*, 32(6):1065–1082, 2014.

[9] C. K. Anjinappa, F. Erden, and I. Güvenç. Base station and passive reflectors placement for urban mmwave networks. *IEEE Transactions on Vehicular Technology*, 70(4):3525–3539, 2021.

[10] G. E. Athanasiadou, P. Fytampanis, D. A. Zarbouti, G. V. Tsoulos, P. K. Gkonis, and D. I. Kaklamani. Radio network planning towards 5g mmwave standalone small-cell architectures. *Electronics*, 9(2):339, 2020.

[11] K. Bahirat, U. Shah, A. A. Cardenas, and B. Prabhakaran. Alert: Adding a secure layer in decision support for advanced driver assistance system (adas). In *Proceedings of the 26th ACM international conference on Multimedia*, pages 1984–1992, 2018.

[12] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak. Compressed channel sensing: A new approach to estimating sparse multipath channels. *Proceedings of the IEEE*, 98(6):1058–1076, 2010.

[13] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang. Wireless communications through reconfigurable intelligent surfaces. *IEEE Access*, 7:116753–116773, 2019.

[14] E. Bastug, M. Bennis, M. Médard, and M. Debbah. Toward interconnected virtual reality: Opportunities, challenges, and enablers. *IEEE Communications Magazine*, 55(6):110–117, 2017.

[15] G. Bielsa, J. Palacios, A. Loch, D. Steinmetzer, P. Casari, and J. Widmer. Indoor localization using commercial off-the-shelf 60 ghz access points. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pages 2384–2392. IEEE, 2018.

[16] K. Bigdely-Shamloo. How to get raw (positional) data from htc vive?, Aug 2017.

[17] L. Bouchard. Combine lidar and cameras for 3d object detection - waymo, Mar 2022.

[18] B. Cao, J. Zhao, P. Yang, P. Yang, X. Liu, and Y. Zhang. 3-d deployment optimization for heterogeneous wireless directional sensor networks on smart city. *IEEE Transactions on Industrial Informatics*, 15(3):1798–1808, 2019.

[19] Y. Cao, C. Xiao, B. Cyr, Y. Zhou, W. Park, S. Rampazzi, Q. A. Chen, K. Fu, and Z. M. Mao. Adversarial sensor attack on lidar-based perception in autonomous driving. In *Proceedings of the 2019 ACM SIGSAC conference on computer and communications security*, pages 2267–2281, 2019.

[20] R. Changalvala and H. Malik. Lidar data integrity verification for autonomous vehicle. *IEEE Access*, 7:138018–138031, 2019.

[21] Y. Chen, C. Lu, and W. Chu. A cooperative driving strategy based on velocity prediction for connected vehicles with robust path-following control. *IEEE Internet of Things Journal*, 7(5):3822–3832, 2020.

[22] J. Choi. Beam selection in mm-wave multiuser mimo systems using compressive sensing. *IEEE Transactions on Communications*, 63(8):2936–2947, 2015.

[23] K.-Y. Chow, K.-S. Lui, and E. Y. Lam. Maximizing angle coverage in visual sensor networks. In *2007 IEEE International Conference on Communications*, pages 3516–3521, 2007.

[24] M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang. 6g wireless communication systems: Applications, requirements, technologies, challenges, and research directions. *IEEE Open Journal of the Communications Society*, 1:957–975, 2020.

[25] B. Coll-Perales, M. Gruteser, and J. Gozalvez. Evaluation of ieee 802.11 ad for mmwave v2v communications. In *2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pages 290–295. IEEE, 2018.

[26] M. Cudak, A. Ghosh, T. Kovarik, R. Ratasuk, T. A. Thomas, F. W. Vook, and P. Moorut. Moving towards mmwave-based beyond-4g (b-4g) technology. In *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2013.

[27] J. Cui, X. Chen, J. Zhang, Q. Zhang, and H. Zhong. Toward achieving fine-grained access control of data in connected and autonomous vehicles. *IEEE Internet of Things Journal*, 8(10):7925–7937, 2020.

[28] M. Cui, G. Zhang, and R. Zhang. Secure wireless communication via intelligent reflecting surface. *IEEE Wireless Communications Letters*, 8(5):1410–1414, 2019.

[29] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini. What should 6g be? *Nature Electronics*, 3(1):20–29, 2020.

[30] F. Devoti and I. Filippini. Planning mm-wave access networks under obstacle blockages: A reliability-aware approach. *IEEE/ACM Transactions on Networking*, 28(5):2203–2214, 2020.

[31] S. Dhillon and K. Chakrabarty. Sensor placement for effective coverage and surveillance in distributed sensor networks. In *2003 IEEE Wireless Communications and Networking, 2003. WCNC 2003.*, volume 3, pages 1609–1614 vol.3, 2003.

[32] M. Di Renzo, K. Ntontin, J. Song, F. H. Danufane, X. Qian, F. Lazarakis, J. De Rosny, D.-T. Phan-Huy, O. Simeone, R. Zhang, M. Debbah, G. Lerosey, M. Fink, S. Tretyakov, and S. Shamai. Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison. *IEEE Open Journal of the Communications Society*, 1:798–807, 2020.

[33] A. Y. Ding and M. Janssen. 5g applications: Requirements, challenges, and outlook. *arXiv preprint arXiv:1810.06057*, 2018.

[34] V. V. Dixit, S. Chand, and D. J. Nair. Autonomous vehicles: disengagements, accidents and reaction times. *PLoS one*, 11(12):e0168054, 2016.

[35] M. Elhoseny, A. Tharwat, A. Farouk, and A. E. Hassanien. K-coverage model based on genetic algorithm to extend wsn lifetime. *IEEE Sensors Letters*, 1(4):1–4, 2017.

[36] M. Elhoseny, A. Tharwat, X. Yuan, and A. E. Hassanien. Optimizing k-coverage of mobile wsns. *Expert Systems with Applications*, 92:142–153, 2018.

[37] L. Fan, X. Xiong, F. Wang, N. Wang, and Z. Zhang. Rangedet: In defense of range view for lidar-based 3d object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2918–2927, 2021.

[38] R. Ford, M. Zhang, M. Mezzavilla, S. Dutta, S. Rangan, and M. Zorzi. Achieving ultra-low latency in 5g millimeter wave cellular networks. *IEEE Communications Magazine*, 55(3):196–203, 2017.

[39] G. Fusco and H. Gupta. ε-net approach to sensor k-coverage. In *International Conference on Wireless Algorithms, Systems, and Applications*, pages 104–114. Springer, 2009.

[40] S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang. Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey. *IEEE Communications Surveys Tutorials*, 22(4):2283–2314, 2020.

[41] A. Gupta and R. K. Jha. A survey of 5g network: Architecture and emerging technologies. *IEEE access*, 3:1206–1232, 2015.

[42] C. Han and S. Duan. Impact of atmospheric parameters on the propagated signal power of millimeter-wave bands based on real measurement data. *IEEE Access*, 7:113626–113641, 2019.

[43] B. L. Harrison, S. Consolvo, and T. Choudhury. Using multi-modal sensing for human activity modeling in the real world. In *Handbook of ambient intelligence and smart environments*, pages 463–478. Springer, 2010.

[44] Z. Hau, S. Demetriou, L. Munoz-González, and E. C. Lupu. Shadow-catcher: Looking into shadows to detect ghost objects in autonomous vehicle 3d sensing. In *European Symposium on Research in Computer Security*, pages 691–711. Springer, 2021.

[45] S. He, D.-H. Shin, J. Zhang, J. Chen, and Y. Sun. Full-view area coverage in camera sensor networks: Dimension reduction and near-optimal solutions. *IEEE Transactions on Vehicular Technology*, 65(9):7448–7461, 2016.

[46] X. Hu, T. Liu, and T. Shu. Fast and high-resolution nlos beam switching over commercial off-the-shelf mmwave devices. *IEEE Transactions on Mobile Computing*, pages 1–1, 2021.

[47] C. Huang, G. C. Alexandropoulos, C. Yuen, and M. Debbah. Indoor signal focusing with deep learning designed reconfigurable intelligent surfaces. In *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5, 2019.

[48] C.-F. Huang and Y.-C. Tseng. The coverage problem in a wireless sensor network. *Mobile networks and Applications*, 10(4):519–528, 2005.

[49] Y. Huo, X. Dong, W. Xu, and M. Yuen. Enabling multi-functional 5g and beyond user equipment: A survey and tutorial. *IEEE Access*, 7:116975–117008, 2019.

[50] K. Iehira, H. Inoue, and K. Ishida. Spoofing attack using bus-off attacks against a specific ecu of the can bus. In *2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, pages 1–4. IEEE, 2018.

[51] C. Iovescu and S. Rao. The fundamentals of millimeter wave sensors. *Texas Instruments*, pages 1–8, 2017.

[52] J. Jia, C. Dong, Y. Hong, L. Guo, and Y. Yu. Maximizing full-view target coverage in camera sensor networks. *Ad Hoc Networks*, 94:101973, 2019.

[53] R. A. Johnson, D. W. Wichern, et al. *Applied multivariate statistical analysis*, volume 5. Prentice hall Upper Saddle River, NJ, 2002.

[54] M. Kamal, G. Srivastava, and M. Tariq. Blockchain-based lightweight and secured v2v communication in the internet of vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 22(7):3997–4004, 2020.

[55] D. T. Kanapram, F. Patrone, P. Marin-Plaza, M. Marchese, E. L. Bodanese, L. Marcenaro, D. M. Gómez, and C. Regazzoni. Collective awareness for abnormality detection in connected autonomous vehicles. *IEEE Internet of Things Journal*, 7(5):3774–3789, 2020.

[56] W.-C. Ke, B.-H. Liu, and M.-J. Tsai. Constructing a wireless sensor network to fully cover critical grids by deploying minimum sensors on grid points is np-complete. *IEEE Transactions on Computers*, 56(5):710–715, 2007.

[57] S. Kutty and D. Sen. Beamforming for millimeter wave communications: An inclusive survey. *IEEE Communications Surveys & Tutorials*, 18(2):949–973, 2015.

[58] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. A. Zhang. The roadmap to 6g: Ai empowered wireless networks. *IEEE Communications Magazine*, 57(8):84–90, 2019.

[59] S. Li, S. Wang, Y. Zhou, Z. Shen, and X. Li. Tightly coupled integration of gnss, ins, and lidar for vehicle navigation in urban environments. *IEEE Internet of Things Journal*, 9(24):24721–24735, 2022.

[60] Y. Li, C. Wen, F. Juefei-Xu, and C. Feng. Fooling lidar perception via adversarial trajectory perturbation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 7898–7907, 2021.

[61] C.-K. Liang, C.-H. Tsai, and M.-C. He. On area coverage problems in directional sensor networks. In *The International Conference on Information Networking 2011 (ICOIN2011)*, pages 182–187, 2011.

[62] J. Liu and J.-M. Park. “seeing is not always believing”: Detecting perception error attacks against autonomous vehicles. *IEEE Transactions on Dependable and Secure Computing*, 18(5):2209–2223, 2021.

[63] Y. Liu, Y. Wang, and G. Chang. Efficient privacy-preserving dual authentication and key agreement scheme for secure v2v communications in an iov paradigm. *IEEE Transactions on Intelligent Transportation Systems*, 18(10):2740–2749, 2017.

[64] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang. An overview of massive mimo: Benefits and challenges. *IEEE journal of selected topics in signal processing*, 8(5):742–758, 2014.

[65] G. R. MacCartney, S. Deng, and T. S. Rappaport. Indoor office plan environment and layout-based mmwave path loss models for 28 ghz and 73 ghz. In *2016 IEEE 83rd vehicular technology conference (VTC Spring)*, pages 1–6. IEEE, 2016.

[66] G. R. MacCartney and T. S. Rappaport. 73 ghz millimeter wave propagation measurements for outdoor urban mobile and backhaul communications in new york city. In *2014 IEEE international conference on communications (ICC)*, pages 4862–4867. IEEE, 2014.

[67] P. Makris, D. N. Skoutas, and C. Skianis. A survey on context-aware mobile and wireless networking: On networking and computing environments’ integration. *IEEE communications surveys & tutorials*, 15(1):362–386, 2012.

[68] E. E. Marvasti, A. Raftari, A. E. Marvasti, Y. P. Fallah, R. Guo, and H. Lu. Cooperative lidar object detection via feature sharing in deep networks. In *2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall)*, pages 1–7. IEEE, 2020.

[69] R. Matsumura, T. Sugawara, and K. Sakiyama. A secure lidar with aes-based side-channel fingerprinting. In *2018 Sixth International Symposium on Computing and Networking Workshops (CANDARW)*, pages 479–482. IEEE, 2018.

[70] R. Meshcheryakov, A. Iskhakov, M. Mamchenko, M. Romanova, S. Uvaysov, Y. Amirgaliyev, and K. Gromaszek. A probabilistic approach to estimating allowed snr values for automotive lidars in "smart cities" under various external influences. *Sensors*, 22(2):609, 2022.

[71] Mikrotik. Manual:interface/w60g, 2018.

[72] Mikrotik. Mikrotik wap60g, March 2, 2020.

[73] A. A. Moamen and N. Jamali. Modesens: An approach for multi-modal mobile sensing. In *Companion Proceedings of the 2015 ACM SIGPLAN International Conference on Systems, Programming, Languages and Applications: Software for Humanity*, pages 40–41, 2015.

[74] E. Moro, I. Filippini, A. Capone, and D. De Donno. Planning mm-wave access networks with reconfigurable intelligent surfaces. In *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1401–1407. IEEE, 2021.

[75] G. Motors. Gm advances self-driving vehicle deployment with acquisition of lidar developer, Oct 2017.

[76] A. Newell, K. Akkaya, and E. Yildiz. Providing multi-perspective event coverage in wireless multimedia sensor networks. In *IEEE Local Computer Network Conference*, pages 464–471, 2010.

[77] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos. A survey of millimeter wave communications (mmwave) for 5g: opportunities and challenges. *Wireless networks*, 21(8):2657–2676, 2015.

[78] K. Ntontin, D. Selimis, A.-A. A. Boulogeorgos, A. Alexandridis, A. Tsolis, V. Vlachodimitropoulos, and F. Lazarakis. Optimal reconfigurable intelligent surface placement in millimeter-wave communications. In *2021 15th European Conference on Antennas and Propagation (EuCAP)*, pages 1–5, 2021.

[79] S. J. Orfanidis. Electromagnetic waves and antennas. *Book*, 2002.

[80] J. Palacios, D. De Donno, and J. Widmer. Tracking mm-wave channel dynamics: Fast beam training strategies under mobility. In *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, pages 1–9. IEEE, 2017.

[81] N. Palizban, S. Szyszkowicz, and H. Yanikomeroglu. Automation of millimeter wave network planning for outdoor coverage in dense urban areas using wall-mounted base stations. *IEEE Wireless Communications Letters*, 6(2):206–209, 2017.

[82] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei. An overview of mimo communications-a key to gigabit wireless. *Proceedings of the IEEE*, 92(2):198–218, 2004.

[83] S. Peng and Y. Xiong. An area coverage and energy consumption optimization approach based on improved adaptive particle swarm optimization for directional sensor networks. *Sensors*, 19(5):1192, 2019.

[84] J. Petit, B. Stottelaar, M. Feiri, and F. Kargl. Remote attacks on automated vehicles sensors: Experiments on camera and lidar. *Black Hat Europe*, 11(2015):995, 2015.

[85] S. Poduri and G. Sukhatme. Constrained coverage for mobile sensor networks. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, volume 1, pages 165–171 Vol.1, 2004.

[86] A. Prasad, M. A. Uusitalo, D. Navrátil, and M. Säily. Challenges for enabling virtual reality broadcast using 5g small cell network. In *2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pages 220–225. IEEE, 2018.

[87] Y. Qi, M. Hunukumbure, and Y. Wang. Millimeter wave los coverage enhancements with coordinated high-rise access points. In *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6. IEEE, 2017.

[88] S. Rajagopal, S. Abu-Surra, and M. Malmirchegini. Channel feasibility for outdoor non-line-of-sight mmwave mobile communication. In *2012 IEEE vehicular technology conference (VTC Fall)*, pages 1–6. IEEE, 2012.

[89] K. Ramachandran, N. Prasad, K. Hosoya, K. Maruhashi, and S. Rangarajan. Adaptive beamforming for 60 ghz radios: Challenges and preliminary solutions. In *Proceedings of the 2010 ACM international workshop on mmWave communications: from circuits to networks*, pages 33–38, 2010.

[90] T. S. Rappaport. 5g millimeter wave wireless: Trials, testimonies, and target rollouts. In *IEEE Infocom*, 2018.

[91] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez. Millimeter wave mobile communications for 5g cellular: It will work! *IEEE access*, 1:335–349, 2013.

[92] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang. Overview of millimeter wave communications for fifth-generation (5g) wireless networks—with a focus on propagation models. *IEEE Transactions on Antennas and Propagation*, 65(12):6213–6230, 2017.

[93] M. E. Rasekh and U. Madhow. Noncoherent compressive channel estimation for mm-wave massive mimo. In *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, pages 889–894. IEEE, 2018.

[94] M. E. Rasekh, Z. Marzi, Y. Zhu, U. Madhow, and H. Zheng. Noncoherent mmwave path tracking. In *Proceedings of the 18th International Workshop on Mobile Computing Systems and Applications*, pages 13–18, 2017.

[95] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar. Millimeter-wave beamforming as an enabling technology for 5g cellular communications: Theoretical feasibility and prototype results. *IEEE communications magazine*, 52(2):106–113, 2014.

[96] R. Roy and T. Kailath. Esprit-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on acoustics, speech, and signal processing*, 37(7):984–995, 1989.

[97] W. Saad, M. Bennis, and M. Chen. A vision of 6g wireless systems: Applications, trends, technologies, and open research problems. *IEEE network*, 34(3):134–142, 2019.

[98] E. Saeedi and Y. Kong. Side-channel vulnerabilities of automobiles. *Transaction on IoT and Cloud Computing*, 2(2):1–8, 2014.

[99] K. Sakaguchi, T. Haustein, S. Barbarossa, E. C. Strinati, A. Clemente, G. Destino, A. Pärssinen, I. Kim, H. Chung, J. Kim, et al. Where, when, and how mmwave is used in 5g and beyond. *IEICE Transactions on Electronics*, 100(10):790–808, 2017.

[100] D. Sarkar, S. Mikki, and Y. Antar. An electromagnetic framework for the deployment of reconfigurable intelligent surfaces to control massive mimo channel characteristics. In *2020 14th European Conference on Antennas and Propagation (EuCAP)*, pages 1–4, 2020.

[101] T. Sato, Y. Hayakawa, R. Suzuki, Y. Shiiki, K. Yoshioka, and Q. A. Chen. Revisiting lidar spoofing attack capabilities against object detection: Improvements, measurement, and new attack. *arXiv preprint arXiv:2303.10555*, 2023.

[102] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE transactions on antennas and propagation*, 34(3):276–280, 1986.

[103] C. Seker, M. T. Güneser, and T. Ozturk. A review of millimeter wave communication for 5g. In *2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, pages 1–5. Ieee, 2018.

[104] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder. 5g: A tutorial overview of standards, trials, challenges, deployment, and practice. *IEEE journal on selected areas in communications*, 35(6):1201–1221, 2017.

[105] H. Shin, D. Kim, Y. Kwon, and Y. Kim. Illusion and dazzle: Adversarial optical channel exploits against lidars for automotive applications. In *International Conference on Cryptographic Hardware and Embedded Systems*, pages 445–467. Springer, 2017.

[106] H. Shin, Y. Son, Y. Park, Y. Kwon, and Y. Kim. Sampling race: Bypassing timing-based analog active sensor spoofing detection on analog-digital systems. In *10th USENIX Workshop on Offensive Technologies (WOOT 16)*, 2016.

[107] Y. Shoukry, P. Martin, Y. Yona, S. Diggavi, and M. Srivastava. Pycra: Physical challenge-response authentication for active sensors under spoofing attacks. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pages 1004–1015, 2015.

[108] S. Singh, F. Ziliotto, U. Madhow, E. Belding, and M. Rodwell. Blockage and directivity in 60 ghz wireless personal area networks: From cross-layer model to multihop mac design. *IEEE Journal on Selected Areas in Communications*, 27(8):1400–1413, 2009.

[109] H. Solomon. *Geometric probability*. SIAM, 1978.

[110] U. N. G. P. Standards and Specifications. Lidar base specification 2023 rev a. `https://www.usgs.gov/media/files/lidar-base-specification-2023-rev-a`, 2023.

[111] D. Steinmetzer, D. Wegemer, M. Schulz, J. Widmer, and M. Hollick. Compressive millimeter-wave sector selection in off-the-shelf ieee 802.11 ad devices. In *Proceedings of the 13th International Conference on emerging Networking EXperiments and Technologies*, pages 414–425, 2017.

[112] J. Sun, Y. Cao, Q. A. Chen, and Z. M. Mao. Towards robust lidar-based perception in autonomous driving: General black-box adversarial sensor attack and countermeasures. In *29th {USENIX} Security Symposium ({USENIX} Security 20)*, pages 877–894, 2020.

[113] S. Sur, X. Zhang, P. Ramanathan, and R. Chandra. Beamspy: Enabling robust 60 ghz links under blockage. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, pages 193–206, Santa Clara, CA, Mar. 2016. USENIX Association.

[114] S. S. Szyszkowicz, A. Lou, and H. Yanikomeroglu. Automated placement of individual millimeter-wave wall-mounted base stations for line-of-sight coverage of outdoor urban areas. *IEEE Wireless Communications Letters*, 5(3):316–319, 2016.

[115] X. Tan, Z. Sun, D. Koutsonikolas, and J. M. Jornet. Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pages 270–278. IEEE, 2018.

[116] D. A. E. Technology. Ydlidar. `https://www.ydlidar.com/`. Accessed: 2021-06-30.

[117] J. Tian, B. Wang, R. Guo, Z. Wang, K. Cao, and X. Wang. Adversarial attacks and defenses for deep learning-based unmanned aerial vehicles. *IEEE Internet of Things Journal*, 2021.

[118] Y. M. Tsang, A. S. Poon, and S. Addepalli. Coding the beams: Improving beamforming training in mmwave communication system. In *2011 IEEE Global Telecommunications Conference-GLOBECOM 2011*, pages 1–6. IEEE, 2011.

[119] Y.-C. Tseng, P.-Y. Chen, and W.-T. Chen. k-angle object coverage problem in a wireless sensor network. *IEEE Sensors Journal*, 12(12):3408–3416, 2012.

[120] H. Ueda, R. Kurachi, H. Takada, T. Mizutani, M. Inoue, and S. Horihata. Security authentication system for in-vehicle network. *SEI technical review*, 81:5–9, 2015.

[121] J. Van Brummelen, M. O'Brien, D. Gruyer, and H. Najjaran. Autonomous vehicle perception: The technology of today and tomorrow. *Transportation research part C: emerging technologies*, 89:384–406, 2018.

[122] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[123] Velodyne LiDAR. Puck datasheet. `https://velodynelidar.com/wp-content/uploads/2019/12/PuckDatasheet-Web.pdf`, 2019. Accessed: November 10, 2023.

[124] Waymo. Waymo driver. `https://waymo.com/intl/zh-cn/waymo-driver/`, May 2009.

[125] T. Wei, A. Zhou, and X. Zhang. Facilitating robust 60 ghz network deployment by sensing ambient reflectors. In *14th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 17)*, pages 213–226, 2017.

[126] Z. Wei, F. Zhang, S. Chang, Y. Liu, H. Wu, and Z. Feng. Mmwave radar and vision fusion for object detection in autonomous driving: A review. *Sensors*, 22(7):2542, 2022.

[127] W. Whyte, A. Weimerskirch, V. Kumar, and T. Hehn. A security credential management system for v2v communications. In *2013 IEEE Vehicular Networking Conference*, pages 1–8. IEEE, 2013.

[128] P.-F. Wu, F. Xiao, C. Sha, H.-P. Huang, R.-C. Wang, and N.-X. Xiong. Node scheduling strategies for achieving full-view area coverage in camera sensor networks. *Sensors*, 17(6), 2017.

[129] Q. Wu and R. Zhang. Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network. *IEEE Communications Magazine*, 58(1):106–112, 2020.

[130] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang. Intelligent reflecting surface-aided wireless communications: A tutorial. *IEEE Transactions on Communications*, 69(5):3313–3351, 2021.

[131] W. Wu, Q. Shen, M. Wang, and X. Shen. Performance analysis of ieee 802.11. ad downlink hybrid beamforming. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2017.

[132] Y. Wu and X. Wang. Achieving full view coverage with randomly-deployed heterogeneous camera sensors. In *2012 IEEE 32nd International Conference on Distributed Computing Systems*, pages 556–565, 2012.

[133] Y. Xiong, M. Lu, and W. Chen. Modeling and maximizing angle coverage in visual sensor networks. In *2018 37th Chinese Control Conference (CCC)*, pages 2406–2409, 2018.

[134] B. Xu, Y. Zhu, D. Li, D. Kim, and W. Wu. Minimum (k, ω)-angle barrier coverage in wireless camera sensor networks. *International Journal of Sensor Networks*, 2016.

[135] Y. Yaman and P. Spasojevic. An intra-cluster model with diffuse scattering for mmwave communications: Rt-icm. *arXiv preprint arXiv:1905.08295*, 2019.

[136] P. Yang, Y. Xiao, M. Xiao, and S. Li. 6g wireless communications: Vision and potential techniques. *IEEE Network*, 33(4):70–75, 2019.

[137] Z. Yang, P. H. Pathak, J. Pan, M. Sha, and P. Mohapatra. Sense and deploy: Blockage-aware deployment of reliable 60 ghz mmwave wlans. In *2018 IEEE 15th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pages 397–405. IEEE, 2018.

[138] W. Yi and G. Cao. On full-view coverage in camera sensor networks. In *2011 Proceedings IEEE INFOCOM*, pages 1781–1789, 2011.

[139] I. Yildirim, A. Uyrus, and E. Basar. Modeling and analysis of reconfigurable intelligent surfaces for indoor and outdoor applications in future wireless networks. *IEEE Transactions on Communications*, 69(2):1290–1301, 2021.

[140] E. Yildiz, K. Akkaya, E. Sisikoglu, and M. Sir. An exact algorithm for providing multi-perspective event coverage in wireless multimedia sensor networks. In *2011 7th International Wireless Communications and Mobile Computing Conference*, pages 382–387, 2011.

[141] C. You, Z. Hau, and S. Demetriou. Temporal consistency checks to detect lidar spoofing attacks on autonomous vehicle perception. In *Proceedings of the 1st Workshop on Security and Privacy for Mobile AI*, pages 13–18, 2021.

[142] L. Yuan, L. Sun, P. S. Yu, and Z. Wang. Decentralized federated learning: A survey and perspective. *arXiv preprint arXiv:2306.01603*, 2023.

[143] W. Yuan, S. M. Armour, and A. Doufexi. An efficient and low-complexity beam training technique for mmwave communication. In *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 303–308. IEEE, 2015.

[144] S. Zeng, H. Zhang, B. Di, Z. Han, and L. Song. Reconfigurable intelligent surface (ris) assisted wireless coverage extension: Ris orientation and location optimization. *IEEE Communications Letters*, 25(1):269–273, 2021.

[145] F. Zhang, Z. Chang, K. Niu, J. Xiong, B. Jin, Q. Lv, and D. Zhang. Exploring lora for long-range through-wall sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(2):1–27, 2020.

[146] J. Zhang, R. Xi, Y. He, Y. Sun, X. Guo, W. Wang, X. Na, Y. Liu, Z. Shi, and T. Gu. A survey of mmwave-based human sensing: Technology, platforms and applications. *IEEE Communications Surveys & Tutorials*, 2023.

[147] J. Zhang, Y. Zhang, K. Lu, J. Wang, K. Wu, X. Jia, and B. Liu. Detecting and identifying optical signal attacks on autonomous driving systems. *IEEE Internet of Things Journal*, 8(2):1140–1153, 2020.

[148] R. Zhang and S. Cao. Real-time human motion behavior detection via cnn using mmwave radar. *IEEE Sensors Letters*, 3(2):1–4, 2018.

[149] X. Zhang, A. Zhang, J. Sun, X. Zhu, Y. E. Guo, F. Qian, and Z. M. Mao. Emp: Edge-assisted multi-vehicle perception. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*, pages 545–558, 2021.

[150] Y. Zhang, X. Sun, and B. Wang. Efficient algorithm for k-barrier coverage based on integer linear programming. *China Communications*, 13(7):16–23, 2016.

[151] P. Zhao, C. X. Lu, J. Wang, C. Chen, W. Wang, N. Trigoni, and A. Markham. mid: Tracking and identifying people with millimeter wave radar. In *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pages 33–40. IEEE, 2019.

[152] A. Zhou, X. Zhang, and H. Ma. Beam-forecast: Facilitating mobile 60 ghz networks via model-driven beam steering. In *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, pages 1–9. IEEE, 2017.

[153] Z. Zhou, S. Das, and H. Gupta. Connected k-coverage problem in sensor networks. In *Proceedings. 13th International Conference on Computer Communications and Networks (IEEE Cat. No. 04EX969)*, pages 373–378. IEEE, 2004.

[154] C. Zhu, C. Zheng, L. Shu, and G. Han. A survey on coverage and connectivity issues in wireless sensor networks. *Journal of Network and Computer Applications*, 35(2):619–632, 2012.