Physical Layer Authentication for Wireless Applications

by

Mohammed Aly Abdrabou Mohammed Hammouda
B.Sc., Military Technical College, 2009
M.Sc., Military Technical College, 2016

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Electrical and Computer Engineering

Physical Layer Authentication for Wireless Applications

by

Mohammed Aly Abdrabou Mohammed Hammouda
B.Sc., Military Technical College, 2009
M.Sc., Military Technical College, 2016

Supervisory Committee

Dr. **T. Aaron Gulliver**, Supervisor
(Department of Electrical and Computer Engineering, University of Victoria)

Dr. **Xiaodai Dong**, Departmental Member
(Department of Electrical and Computer Engineering, University of Victoria)

Dr. **Andrew Rowe**, Outside Member
(Department of Mechanical Engineering, University of Victoria)

ABSTRACT

Internet of things (IoT) devices have become ubiquitous and go far beyond smartphones and similar devices. The IoT allows for numerous applications such as smart homes, intelligent healthcare, and intelligent transportation. However, high deployment costs limit cellular network coverage in remote and rural areas, and the reliability of cellular infrastructure during natural disasters is a concern. Thus, space and ground network integration has been proposed to provide global connectivity and support a wide range of IoT applications. Unfortunately, spoofing attacks are problematic due to network complexity and heterogeneity. Authentication for access control is an efficient way to ensure user legitimacy. However, upper layer authentication (ULA) is challenging due to limited computational power, high complexity, and communication overhead. Thus, physical layer authentication (PLA) has been proposed to aid ULA in solving these problems. PLA exploits the fact that legitimate parties and attackers have distinct physical characteristics which are unique between every pair of connected peers based on their spatial locations.

In this dissertation, PLA schemes are presented using wireless attributes. First, an adaptive PLA scheme for IoT applications in urban environments is proposed using machine learning (ML) with antenna diversity to increase the number of features. A one-class classifier support vector machine (OCC-SVM) is employed using the magnitude and real and imaginary parts of the received signal at each receive antenna as features. The sounding reference signal (SRS) in the 5G uplink radio frame is employed for this purpose. Results are presented which show that this scheme provides a high authentication rate (AR) with sufficient antenna diversity. Furthermore, an adaptive PLA scheme is presented for collaboration between distributed IoT devices in multiple-input-multiple-output (MIMO) systems. The performance is evaluated considering two majority voting schemes for practical IoT applications. These schemes may be preferable for IoT devices with limited computing capabilities.

An adaptive PLA scheme for low earth orbit (LEO) satellites is proposed that employs ML with Doppler frequency shift (DS) and received power (RP) features. This scheme is evaluated for fixed and mobile satellite services at different altitudes. Results are presented which show that the proposed scheme provides better authentication performance using DS and RP features together compared to using them separately. Moreover, PLA using a hypothesis test with threshold or ML for satellite authentication is presented. The results show that the AR with DS is higher than with RP at low elevation angles for both schemes, but is higher with RP at high elevation angles. Further, the ML authentication scheme provides a higher AR than the threshold scheme for a small percentage of the training data considered as outliers, but at larger percentages the OR threshold scheme is better. Finally, game-theoretic satellite authentication using physical characteristics for spoofing detection is presented. Results are given to demonstrate the effectiveness of the proposed approach.

Contents

Supervisory Committee **ii**

Abstract **iii**

Contents **v**

List of Figures **x**

List of Tables **xvi**

List of Acronyms **xvii**

List of Publications **xx**

Acknowledgements **xxi**

1 Introduction and Motivation **1**

1.1 Background 1

1.1.1 Internet of Things (IoT) 1

1.1.2 Vertical Heterogeneous Network (VHetNet) 2

1.2 Problem Statement and Motivation 4

1.3 Research Overview 5

1.4 Challenges 6

1.5 Contributions 7

1.6 Outline of the Dissertation 9

2 Physical Layer Authentication for IoT Applications 10

2.1 Related Work . . . 11

2.1.1 Traditional PLA . . . 11

2.1.2 Machine Learning PLA . . . 11

2.2 Contributions . . . 13

2.3 System Model . . . 14

2.4 The Proposed PLA Scheme . . . 16

2.5 Simulation Results . . . 21

2.5.1 Diversity Combining . . . 27

2.5.2 Comparison of PLA Techniques . . . 30

2.5.3 System Complexity . . . 31

2.6 Conclusion . . . 32

3 Adaptive Physical Layer Authentication for IoT in MIMO Communication Systems using Support Vector Machine 33

3.1 Related Work . . . 34

3.2 Contributions . . . 35

3.3 System Model . . . 36

3.4 The Proposed Scheme . . . 38

3.4.1 Proposed Scheme Without Voting . . . 40

3.4.2 Proposed Scheme With Voting . . . 43

3.5 Simulation Results . . . 47

3.5.1 Scenarios 1, 2, and 3 . . . 50

3.5.2 AR for Different SNRs . . . 51

3.5.3 Majority Voting . . . 52

3.5.4 Kernel Comparison . . . 55

3.5.5 Number of Features . . . 55

3.5.6 Effect of Outlier Percentage . . . 57

3.5.7 Multiuser Performance 57
3.6 Conclusion 59

4 Physical Layer Authentication for Satellite Communication Systems using Machine Learning 60
4.1 Related Work 60
4.2 Contributions 62
4.3 System Model 62
4.3.1 Doppler Frequency Shift 64
4.3.2 Received Power 64
4.3.3 Problem Formulation 65
4.4 The Proposed Authentication Scheme 65
4.5 Performance Evaluation 69
4.5.1 Fixed Satellite Services (FSS) 72
4.5.2 Mobile Satellite Services (MSS) 73
4.6 Conclusion 74

5 Authentication for Satellite Communication Systems using Physical Characteristics 76
5.1 Related Work 77
5.2 Contributions 78
5.3 System Model 79
5.3.1 Doppler Frequency Shift 80
5.3.2 Received Power 80
5.4 Authentication Based on Physical Characteristics 82
5.4.1 Estimated Doppler Frequency Shift 82
5.4.2 Estimated Received Power 83
5.4.3 Threshold Authentication Scheme 83
5.4.4 Machine Learning Authentication Scheme 86

5.5 Simulation Results 89
5.5.1 DS and RP Over the Communication Session 89
5.5.2 Threshold Authentication Scheme Performance 93
5.5.3 Machine Learning Authentication Scheme Performance 95
5.5.4 Authentication Scheme Performance Comparison 97
5.6 Conclusion 102

6 Game Theoretic Spoofing Detection for Space Information Networks using Physical Attributes 103
6.1 Related Work 104
6.2 Contributions 105
6.3 System Model 106
6.3.1 Doppler Frequency Spread 108
6.3.2 Received Power 108
6.4 Game Theoretic PLA Scheme 109
6.4.1 Doppler Frequency Spread 109
6.4.2 Received Power 110
6.4.3 Authentication Test 112
6.4.4 Authentication Metrics 113
6.4.5 Game Theoretic Scheme 114
6.5 Performance Evaluation 117
6.5.1 LEO Satellite PLA Without Game Theory 117
6.5.2 LEO Satellite PLA With Game Theory 121
6.6 Conclusion 126

7 Conclusion and Future Work 127
7.1 Conclusion 127
7.2 Future Work 129
7.2.1 PLA with a Reconfigurable Intelligent Surface (RIS) 129

7.2.2 PLA for Optical Communications 130
7.2.3 PLA for Underwater Communications 130

A One-Class Classification Support Vector Machine (OCC-SVM) 131

B Evaluation Metrics 134

Bibliography 136

List of Figures

Figure 1.1 The internet of things (IoT). 2

Figure 1.2 The vertical heterogeneous network (VHetNet) architecture. . 3

Figure 1.3 The PLA scenario. 5

Figure 2.1 System model. 14

Figure 2.2 5G frame structure with configuration 0 and numerology 2. . . 16

Figure 2.3 Authentication scheme flowchart. 18

Figure 2.4 Sliding window for feature updates. 20

Figure 2.5 Urban environment simulation scenario. 22

Figure 2.6 MDR and FAR versus velocity with linear kernel OCC-SVM, SRS every 20 symbols, 1 transmit and 1, 4, and 8 receive antennas, for (a) SNR = 8 dB, (b) SNR = 10 dB, and (c) SNR = 12 dB. 23

Figure 2.7 AR versus velocity with linear kernel OCC-SVM, SRS every 20 symbols, and SNR = 8 dB, 10 dB, and 12 dB for (a) 1 receive antenna, (b) 4 receive antennas, and (c) 8 receive antennas. 24

Figure 2.8 AR versus velocity with linear kernel OCC-SVM, SNR = 8 dB, SRS every 20, 28, and 56 symbols for (a) 1 receive antenna, (b) 4 receive antennas, and (c) 8 receive antennas. 25

Figure 2.9 AR versus velocity with SNR = 8 dB and 15 dB, SRS every 28 symbols, and 4 and 8 receive antennas. 26

Figure 2.10 AR versus velocity with SNR = 15 dB, SRS every 28 symbols, and 8 receive antennas for OCC-SVM using linear, sigmoid, and polynomial kernels. 26
Figure 2.11 MDR and FAR versus velocity with linear kernel OCC-SVM, SRS every 28 symbols, SNR = 8 dB for selection combining (SC), equal gain combining (EGC), and maximum ratio combining (MRC) using 4 receive antennas and without combining using 1 and 4 receive antennas. 28
Figure 2.12 AR versus velocity with linear kernel OCC-SVM, SRS every 28 symbols, SNR = 8 dB for selection combining (SC), equal gain combining (EGC), and maximum ratio combining (MRC) using 4 receive antennas and without combining using 1 and 4 receive antennas. 29
Figure 2.13 AR versus velocity with linear kernel OCC-SVM, SRS every 28 symbols, SNR = 8 dB, and 4 and 8 receive antennas using the real and imaginary parts of the received signals as features [1], the RSS feature [2], and the magnitude and real and imaginary parts of the received signals as features (proposed scheme). 30

Figure 3.1 The system model. 37
Figure 3.2 Flowchart of the proposed authentication scheme. 39
Figure 3.3 Sliding window for feature updates. 41
Figure 3.4 The simulation layout for Scenarios 1, 2, and 3. 49
Figure 3.5 The simulation layout for Scenario 4. 49

Figure 3.6 MDR, FAR, and AR versus the number of receive antennas with linear kernel OCC-SVM, SNR = 8 dB, velocity = 0.4 km/h, and 1, 4, and 8 receive antennas for Scenario 1, 2, and 3 with (a) 1 transmit antenna, (b) 2 transmit antennas, and (c) 4 transmit antennas. 50

Figure 3.7 AR versus the number of transmit antennas for Scenario 2 with linear kernel OCC-SVM, 1, 4, and 8 receive antennas, velocity = 0.4 km/h, and (a) SNR = 8 dB and (b) SNR = 12 dB. 52

Figure 3.8 AR versus SNR with linear kernel OCC-SVM, velocity = 0.4 km/h, and 8 receive antennas for Scenario 2 using MN and N majority voting and without voting with (a) 1 transmit antenna, (b) 2 transmit antennas, and (c) 4 transmit antennas. 53

Figure 3.9 AR versus the number of receive antennas for Scenario 2 with SNR = 12 dB, velocity = 0.4 km/h, 2 transmit antennas, and 1, 4, and 8 receive antennas for OCC-SVM using linear, sigmoid, and polynomial kernels. 54

Figure 3.10 AR versus velocity for Scenario 1 with linear kernel OCC-SVM, SNR = 8 dB, 1 transmit antenna and 4 receive antennas using the magnitude, real and imaginary parts, and the magnitude and real and imaginary parts of the received signals as features. 54

Figure 3.11 MDR, FAR, and AR versus the number of receive antenna with linear kernel OCC-SVM, SNR = 8 dB, velocity = 0.4 km/h, and 1 transmit antenna for Scenario 1 with $\eta = 0.2, 0.5,$ and 0.8. 56

Figure 3.12 MDR, FAR, and AR versus the number of receive antenna with linear kernel OCC-SVM, SNR = 8 dB, velocity = 0.4 km/h, and 1 transmit antenna for Scenarios 1 and 4. 58

Figure 4.1 System model. 63
Figure 4.2 Proposed authentication scheme flowchart. 66
Figure 4.3 Sliding window for feature updates. 67
Figure 4.4 Model layout. 70
Figure 4.5 Scaled DS and scaled RP over the communication session. . . 71
Figure 4.6 MDR, FAR, and AR versus altitude for FSS with (a) linear OCC-SVM kernel, and (b) polynomial OCC-SVM kernel. . . . 72
Figure 4.7 MDR, FAR, and AR versus altitude for MSS with (a) linear OCC-SVM kernel, and (b) polynomial OCC-SVM kernel. . . . 74

Figure 5.1 System model. 79
Figure 5.2 Proposed authentication scheme flowchart. 81
Figure 5.3 Normalized Doppler frequency shift (NDS) and normalized received power (NRP) over the communication session. 89
Figure 5.4 Normalized α_i ($N\alpha_i$) versus θ. 91
Figure 5.5 Normalized β_i ($N\beta_i$) versus θ. 91
Figure 5.6 The trajectories for Eve and Alice where Eve1 corresponds to case 1 and Eve2 to case 2. 92
Figure 5.7 MDR, FAR, and AR for the DS, RP, AND, and OR threshold authentication schemes averaged over the communication session versus (a) α with $\sigma^2 = 0.08$, and (b) σ^2 with $\alpha = 0.4$. . 94
Figure 5.8 AR for the DS and RP threshold authentication schemes versus θ with $\alpha = 0.4$ and $\sigma^2 = 0.03, 0.05$, and 0.06. 96

Figure 5.9 AR for the DS, RP, and DS and RP ML authentication schemes averaged over the communication session with $\eta = 0.5$ and $\ell = 10$ versus (a) α with $\sigma^2 = 0.08$, and (b) σ^2 with $\alpha = 0.4$. . 96

Figure 5.10 AR versus θ for the DS and RP ML authentication schemes with $\eta = 0.5$, $\alpha = 0.4$, $\ell = 10$, and $\sigma^2 = 0.03$ and 0.06. 98

Figure 5.11 AR for the DS, RP, AND, and OR threshold authentication schemes and the DS, RP, and DS and RP ML authentication schemes averaged over the communication session with $\eta =$ 0.1 and 0.5, and $\ell = 10$ versus (a) α with $\sigma^2 = 0.02$, and (b) σ^2 with $\alpha = 0.3$. 98

Figure 5.12 AR versus θ with $\sigma^2 = 0.04$ and $\alpha = 0.3$ for the (a) separate DS and RP threshold and separate DS and RP ML authentication schemes with $\eta = 0.1$ and 0.5 and $\ell = 10$, and (b) AND and OR threshold authentication schemes, and DS and RP ML authentication schemes with $\eta = 0.1$ and 0.5 and $\ell = 10$.100

Figure 5.13 AR for DS and RP ML authentication scheme versus θ with $\sigma^2 = 0.04$, $\eta = 0.1$, and $\alpha = 0.3$ for $\ell = 1, 3, 5$ and 10. 101

Figure 6.1 The PLA system model. The GS selects the optimal threshold to maximize its utility while s tries to select the attack probability to minimize this utility. 107

Figure 6.2 The trajectories for l and s within the GS receive antenna HPBW. 107

Figure 6.3 Normalized values at the GS versus the satellite elevation angle θ (a) $f_d(\theta)$, $m_{d,s,i}$ and $m_{d,l,i}$ and (b) $f_r(\theta)$, $m_{r,s,i}$ and $m_{r,l,i}$. 111

Figure 6.4 (a) $m_{d,l,i}$ and (b) $m_{r,s,i}$ versus θ. 112

Figure 6.5 Average error rate versus τ with $\sigma_l = 0.5$ and 2, and $\sigma_s = 0.5$ and 2 for the DS, RP, and combined authentication schemes (a) P_f and (b) P_m. 118

Figure 6.6 Average AR versus τ with $\sigma_l = 0.5$ and 2, and $\sigma_s = 0.5$ and 2 for the DS, RP, and combined authentication schemes. . . . 119

Figure 6.7 Average AR versus τ with $\sigma_l = 0.5$ and 1, and $\sigma_s = 0.5$ and 1 for the DS [3] and combined DS and RP authentication schemes. 120

Figure 6.8 $U_g(\tau, k)$ versus τ with $\sigma_l = 0.5$, $\sigma_s = 0.5$, $P_{al} = 10$, $P_{rs} = 10$, $C_{rl} = 5$, $C_{as} = 5$, $C_s = 5$, and $k = 0.3, 0.5$ and 0.7 for the DS, RP, and combined authentication schemes. 121

Figure 6.9 $U_g(\tau, k)$ versus τ with $k = 0.5$, $\sigma_l = 0.5$, $\sigma_s = 0.5$, $P_{al} = 10$, $P_{rs} = 10$, $C_{rl} = 5$, and $C_s = 4, 8$, and 12 for the DS, RP, and combined authentication schemes. 123

Figure 6.10 $\frac{\partial U_s(\tau,k)}{\partial k}$ versus τ with $\sigma_l = 1$, $C_s = 5$, and $C_{as} = 8$ and 10 for the DS, RP, and combined authentication schemes. 123

Figure 6.11 $\frac{\partial U_s(\tau,k)}{\partial k}$ versus τ with $\sigma_l = 1$, $C_s = 1, 5$ and 10, and $P_{al} = C_{as} = 8$ for the DS, RP, and combined authentication schemes. 124

Figure 6.12 τ^* and k^* versus σ_s with $C_s = 5$ and $\sigma_a = 1$ for the (a) DS or RP authentication schemes, and (b) combined authentication scheme. 125

Figure B.1 Confusion matrix. 134

List of Tables

Table 2.1 Simulation Parameters 22

Table 3.1 Simulation Parameters 48

Table 4.1 Maximum Distances Between Alice and Eves over the Session at Different Altitudes 70

Table 4.2 Simulation Parameters 71

Table 5.1 Simulation Parameters 90

Table 5.2 Range of Doppler Frequency Shifts at Different Altitudes . . . 90

Table 5.3 Range of Received Power at Different Altitudes 90

List of Acronyms

5G	Fifth Generation
6G	Sixth Generation
APLA	Acoustic-Based Physical Layer Authentication
AR	Authentication Rate
BS	Base Station
CNN	Convolutional Neural Network
CSI	Channel State Information
CANs	Controller Area Networks
CIR	Channel Impulse Response
CM	Confusion Matrix
CNR	Carrier-to-Noise Ratio
CFO	Carrier Frequency Offset
DS	Doppler Frequency Shift
Dyna-PS	Dyna Architecture and Prioritized Sweeping
EGC	Equal Gain Combining
FSS	Fixed Satellite Services
FSPL	Free Space Path Loss
FAR	False Alarm Rate
GNSS	Global Navigation Satellite System
GS	Ground Station
GEO	Geosynchronous Equatorial Orbit
HAPs	High-Altitude Platforms
IoT	Internet of Things

ITU-T	International Telecommunication Union-Telecommunication
IQ	In-Phase and Quadrature
IQI	In-Phase-Quadrature-Phase Imbalance
KNN	K-Nearest Neighbors
LEO	Low Earth Orbit
LAPs	Low-Altitude Platforms
LOS	Line-of-Sight
LMS	Land Mobile Satellite
MISO	Multiple Input Single Output
MIMO	Multiple Input Multiple Output
MEO	Medium Earth Orbit
MANETs	Mobile Ad Hoc Networks
MSS	Mobile Satellite Services
ML	Machine Learning
MDR	Missed Detection Rate
MRC	Maximum Ratio Combining
NTN	Non-Terrestrial Networks
NIST	National Institute of Standards and Technology
OCC-SVM	One-Class Classification Support Vector Machine
OPLA	Optical-Based Physical Layer Authentication
OFDM	Orthogonal Frequency Division Multiplexing
OCC	One-Class Classification
PLA	Physical Layer Authentication
QoS	Quality-of-Service
RSS	Received Signal Strength
RP	Received Power
RF	Radio Frequency
RSSI	Received Signal Strength Indicator

RSS	Received Signal Strength
RIS	Reconfigurable Intelligent Surfaces
SAGIN	Space-Air-Ground Integrated Network
SIMO	Single Input Multiple Output
SVM	Support Vector Machine
STK	System Tool Kit
SRS	Sounding Reference Signal
SC	Selection Combining
TLE	Two-Line Element
TDOA	Time Difference of Arrival
TOA	Time of Arrival
TCC-SVM	Two-Class Classification Support Vector Machine
UAV	Unmanned Aerial Vehicle
ULA	Upper Layer Authentication
USRPs	Universal Software Radio Peripherals
UPLA	Underwater Physical Layer Authentication
VHetNet	Vertical Heterogeneous Network

List of Publications

Journal papers

- M. Abdrabou and T. A. Gulliver, "Adaptive Physical Layer Authentication using Machine Learning with Antenna Diversity," *IEEE Transactions on Communications*, vol. 70, no. 10, pp. 6604-6614, Oct. 2022.
- M. Abdrabou and T. A. Gulliver, "Physical Layer Authentication for Satellite Communication Systems Using Machine Learning," *in IEEE Open Journal of the Communications Society*, vol. 3, pp. 2380-2389, Dec. 2022.
- M. Abdrabou and T. A. Gulliver, "Authentication for Satellite Communication Systems using Physical Characteristics," *in IEEE Open Journal of Vehicular Technology*, vol. 4, pp. 48-60, Nov. 2022.
- M. Abdrabou and T. A. Gulliver, "Adaptive Physical Layer Authentication for IoT in MIMO Communication Systems using Support Vector Machine," *IEEE Internet of Things Journal*, vol. 10, pp. 19861-19873, Nov. 2023.

Conference papers

- M. Abdrabou and T. A. Gulliver, "LEO Satellite Authentication using Physical Layer Features with Support Vector Machine," *in IEEE International Conference on Communication, Networks and Satellite*, Solo, Indonesia, 2022.
- M. Abdrabou and T. A. Gulliver, "Threshold-Based Physical Layer Authentication for Space Information Networks," *in IEEE International Conference on Communication, Networks and Satellite*, Solo, Indonesia, 2022.

ACKNOWLEDGEMENTS

In the name of Allah, the Most Gracious and the Most Merciful. I want to start by expressing my heartfelt **thanks to the Almighty Allah**, Alhamdulillah. Then, I am truly grateful to my supervisor, Prof. **T. Aaron Gulliver**, for his unwavering support, extensive knowledge, and invaluable guidance. I also appreciate the professional advice given by my supervisory committee members, Prof. **Xiaodai Dong** and Prof. **Andrew Rowe** and my external examiner Prof. **Xianbin Wang**. I extend my thanks to my late father, **Aly**, with prayers for mercy and paradise. Words cannot capture my gratitude, appreciation, love, and respect for my mother **Nadia**'s sacrifices, encouragement, and love. A special thanks goes to my wonderful wife **Marwa** and our children, Princess **Habiba** and King **Asser**, for their sacrifices, patience, love, moral support, and constant encouragement throughout the development of this work and my entire life. I also want to express my gratitude to my siblings, **Randa**, **Hassan**, and **Abeer**, and my entire family for their ongoing encouragement, support, and love. Finally, I would like to convey my sincerest appreciation to the **Egyptian government** for supporting my doctoral research.

Mohammed Aly Abdrabou Mohammed Hammouda

Chapter 1

Introduction and Motivation

1.1 Background

Wireless communications has become ubiquitous in the lives of most people worldwide. Moreover, future wireless networks will support applications on devices far beyond smartphones and other mobile devices. These networks have complex and heterogeneous characteristics due to the tremendous number and diversity of devices and applications. Fifth-generation (5G) and beyond wireless networks are being developed to enable a wide range of applications [4, 5, 6, 7]. Recently, the integration of space, air, and ground networks has been considered in what is known as a vertical heterogeneous network (VHetNet) or space-air-ground integrated network (SAGIN) [8].

1.1.1 Internet of Things (IoT)

The rapidly increasing number of low-cost devices and access points, combined with increased mobility and heterogeneity, creates a highly complex and dynamic environment for future wireless networks [9]. The internet of things (IoT) has emerged as a new communication paradigm due to the massive growth in the number of smart

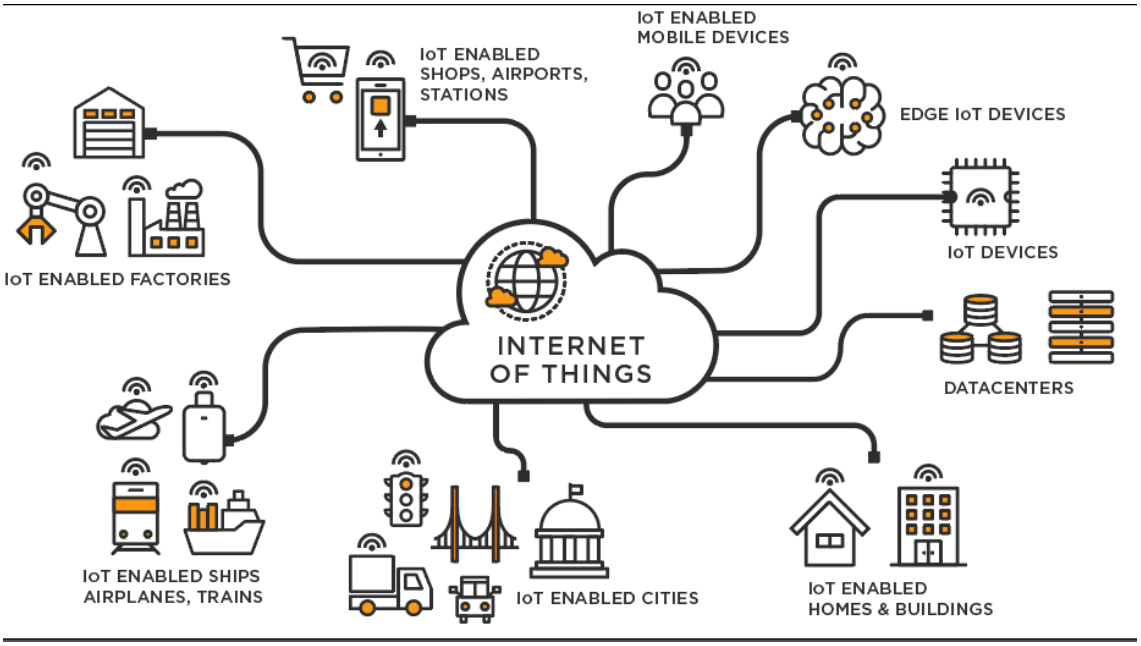

Figure 1.1: The internet of things (IoT).

devices [10]. Figure 1.1 [11] illustrates the diversity of IoT applications such as smart homes and cities, intelligent healthcare, and intelligent transportation [12, 13]. The number of IoT devices is increasing rapidly [14]. According to the International Telecommunication Union (ITU), there will be over 35 billion wireless connected devices in 2025, growing to 97 billion by 2030, which is more than 10 times the projected human population [15].

1.1.2 Vertical Heterogeneous Network (VHetNet)

5G cellular networks have been developed to provide lower latency, higher speeds, and greater capacity than 4G networks. However, the high deployment costs limit cellular network coverage in remote and rural areas. Moreover, natural disasters affect the reliability of cellular infrastructure and can result in isolation in some regions [16]. Future wireless network architectures, e.g. sixth-generation (6G), are being developed to improve coverage and reliability. Research at leading telecom-

munication companies such as Ericsson and Huawei is focused on 6G development [17]. The goal is to provide reliable worldwide connectivity [8]. One approach is to integrate terrestrial networks, e.g. cellular networks, and non-terrestrial aerial networks, e.g. unmanned aerial vehicles (UAVs) and space networks [18].

The VHetNet architecture is composed of space, air, and ground networks as shown in Figure 1.2 [18]. The space network includes geosynchronous equatorial orbit (GEO), medium earth orbit (MEO), and low earth orbit (LEO) satellites as well as inter-satellite links, ground stations, and terminals. The aerial network is composed of high-altitude platforms (HAPs), low-altitude platforms (LAPs), aircraft, UAVs, airships, and balloons, while the ground network includes mobile ad hoc networks (MANETs) and cellular networks [18].

Satellite communication has become important for broadcast and broadband coverage in commercial, emergency, and military applications [19]. In particular,

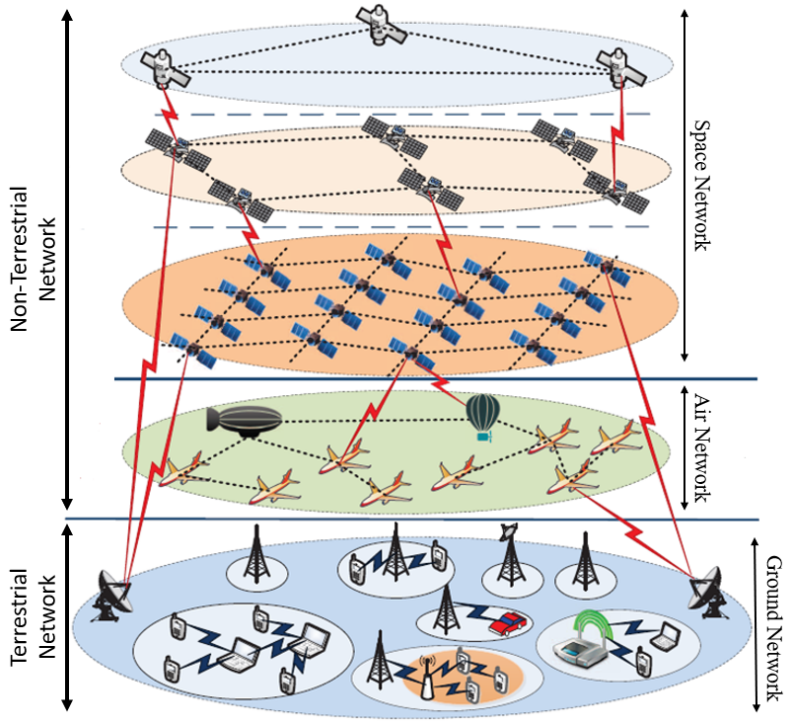

Figure 1.2: The vertical heterogeneous network (VHetNet) architecture.

LEO satellite constellations have gained attention for supporting future networks because they have low cost and low latency, and can provide fixed satellite services (FSS) and mobile satellite services (MSS). The number of LEO satellites is increasing rapidly to provide the services required by VHetNets [16]. LEO constellations now provide global connectivity through thousands of satellites [20]. For example, many constellations are under development such as Kuiper, OneWeb, and Boeing with 3236, 882-1980, and 1396-2956 satellites, respectively [21].

1.2 Problem Statement and Motivation

The open nature and dynamic characteristics of heterogeneous networks raises privacy and security concerns as IoT devices and satellite communication systems are susceptible to spoofing attacks [22]. Spoofing attacks are considered a serious threat as they can allow illegitimate users to impersonate legitimate users to steal information or insert harmful data [3, 19]. Authentication for access control is an efficient way to ensure data security [23, 24, 25, 26]. Despite the existence of numerous solutions, authentication remains crucial yet challenging for communication networks, particularly large-scale and heterogeneous wireless networks, which require robust and flexible authentication [27, 13, 28, 29]. Thus, wireless authentication has gained significant attention [30, 31].

Upper layer authentication (ULA) for future heterogeneous networks is challenging due to the computational power of attackers, which can make it easy to imitate legitimate users, and the complexity and communication overhead of ULA [32, 33]. Physical layer authentication (PLA) can be combined with ULA to strengthen network security [20, 32]. PLA exploits the fact that legitimate parties and attackers have distinct channel attributes with other parties based on their spatial locations as shown in Figure 1.3 [34]. Thus, PLA can provide a high level of security as the corresponding physical layer attributes between spatially located users are unique

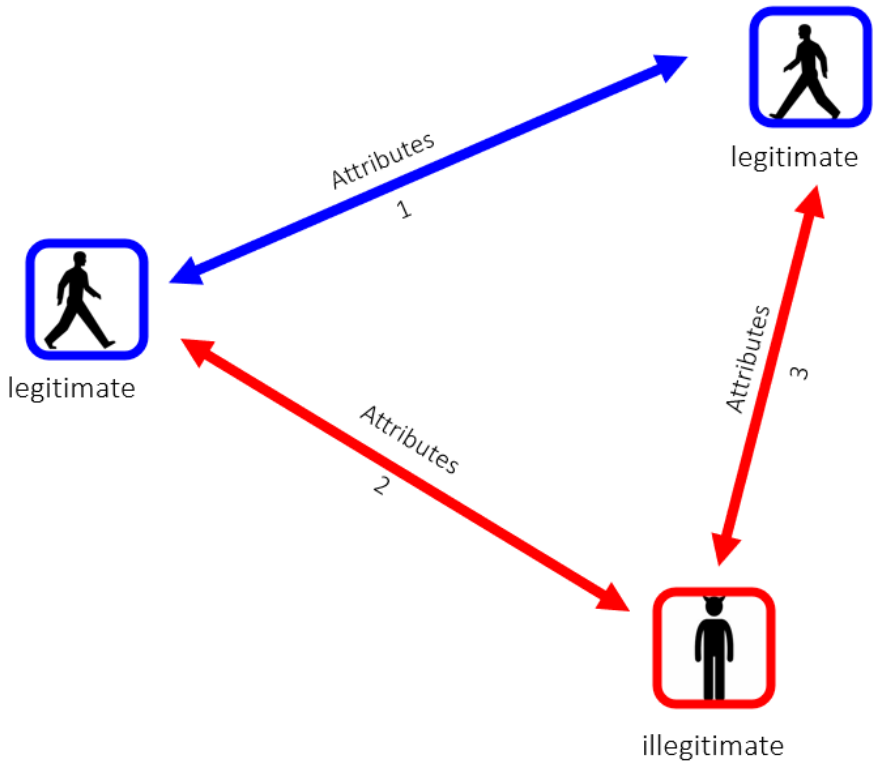

Figure 1.3: The PLA scenario.

[33].

1.3 Research Overview

The goal of this research is to develop adaptive and robust PLA schemes for wireless applications using wireless attributes. First, an efficient PLA scheme is designed to mitigate authentication challenges in IoT applications. The proposed PLA scheme utilizes antenna diversity to increase the number of features. A one-class classification support vector machine (OCC-SVM) which only uses legitimate user data is employed for training and testing. Moreover, PLA for collaboration between distributed IoT devices in multiple-input-multiple-output (MIMO) communication systems is considered. The focus is to develop an authentication scheme using

machine learning (ML) that exploits spatially independent antennas in MIMO systems. Next, an adaptive PLA scheme for LEO satellites is developed using ML with Doppler frequency shift (DS) and received power (RP) as features. Furthermore, hypothesis testing with a threshold and ML are used for satellite authentication to discriminate between legitimate and illegitimate satellites. The estimation error of the DS and RP is employed in evaluating the authentication performance. Finally, game-theoretic satellite authentication based on physical characteristics is considered for spoofing detection using the interaction between the spoofer and legitimate user.

1.4 Challenges

One of the primary challenges in implementing and validating the proposed authentication scheme is the difficulty in obtaining training data from illegitimate users. Collecting such data is often infeasible so to overcome this limitation the proposed schemes rely on training data from legitimate users only. The robustness of the proposed scheme is evaluated under various worst-case scenarios, including situations where illegitimate users start from the same initial location as legitimate users and move in arbitrary directions, and close proximity between legitimate and illegitimate satellites.

Addressing the inherent dynamism of wireless channels, especially in mobile environments, has been a challenge in prior research. An adaptive approach is proposed to mitigate the degradation in authentication performance due to wireless channel fluctuations. The computational complexity of PLA is an issue with previous research. Previous approaches require a large number of training iterations, whereas the proposed schemes provide excellent performance with a small number of iterations. This not only keeps the computational complexity low but also aligns with the demands of practical wireless communication environments.

Most previous PLA schemes increase the overhead in the communication system. The proposed schemes address this concern by utilizing the sounding reference signal (SRS) in the 5G frame structure to obtain features. This improves the practicality, making the proposed approaches more feasible than other PLA solutions. Furthermore, challenges in assessing satellite authentication performance arise from the methodology used for evaluation. The proposed solutions use real satellite data, specifically two-line element (TLE) data, with the system tool kit (STK). This ensures the performance evaluation is realistic and accurate.

1.5 Contributions

In Chapter 2, an adaptive PLA scheme using OCC-SVM is proposed which utilizes antenna diversity. The magnitude and real and imaginary parts of the received signal from each antenna are used as features. The proposed scheme is based on the 5G frame structure and does not require training data from illegitimate users. It requires only a small number of samples for training and they are adaptively updated to adjust to dynamic changes in the wireless channel environment. The proposed scheme is evaluated and shown to be effective in different urban environments. Furthermore, the AR for the selection combining (SC), equal gain combining (EGC), and maximum ratio combining (MRC) diversity combining techniques is obtained.

In Chapter 3, an adaptive PLA scheme using OCC-SVM with antenna diversity in MIMO systems is proposed. This scheme is evaluated using linear, sigmoid, and polynomial OCC-SVM kernels. The sounding reference signal (SRS) in the 5G frame structure is used to obtain the features so it is more practical than other PLA solutions. Further, training data from illegitimate users is not required. The proposed scheme is evaluated and shown to be effective in urban environments. This includes the situation where illegitimate users start from the same initial location as the legitimate user but move in arbitrary directions, which can be considered worst

case. The authentication performance is evaluated considering two majority voting schemes for practical IoT applications.

In Chapter 4, an adaptive physical layer authentication scheme using OCC-SVM is proposed to authenticate LEO satellites utilizing the DS and RP as features. The proposed scheme is validated for FSS and MSS when the illegitimate satellite is within the half power beamwidth (HPBW). The proposed scheme is evaluated using the DS and RP features separately and together. Results are presented using TLE data for real satellites to verify the effectiveness of the proposed schemes, where TLE is orbital data for Earth-orbiting objects [35].

In Chapter 5, an adaptive PLA scheme using DS and RP characteristics to authenticate LEO satellites is proposed. Hypothesis testing using a threshold or ML is used to discriminate between legitimate and illegitimate satellites. The AR is evaluated using DS and RP characteristics separately and together over the communication session and DS and RP estimation errors are considered. Results are presented using TLE data for real satellites to verify the effectiveness of the proposed schemes.

In Chapter 6, a game theoretic PLA scheme is proposed to detect spoofing attacks in LEO satellite communication systems. The DS and RP are used as attributes to discriminate between legitimate and illegitimate satellites at the ground station (GS). Furthermore, a zero-sum PLA game is formulated where the GS selects its optimal detection threshold and the illegitimate satellite selects its optimal attack probability to maximize their respective utilities. TLE data for real satellites is used with the system tool kit (STK) for performance evaluation [36]. It is shown that the mean of the magnitude differences between the current and previous DS and RP values at the GS can be accurately approximated using third-degree polynomials.

1.6 Outline of the Dissertation

The rest of the dissertation is organized as follows. Chapter 2 presents the proposed adaptive PLA scheme for IoT applications. Chapter 3 gives the proposed PLA scheme for IoT in MIMO communication systems. The proposed adaptive PLA scheme for LEO satellites is introduces in Chapter 4. Chapter 5 presents the proposed PLA scheme for LEO satellites using hypothesis testing with a threshold or ML. Chapter 6 gives the proposed game-theoretic satellite authentication based on physical characteristics for LEO satellites. Finally, some concluding remarks and directions for future work are presented in Chapter 7.

Chapter 2

Physical Layer Authentication for IoT Applications

In this chapter, an adaptive physical layer authentication scheme for IoT applications is proposed using ML [37]. Antenna diversity at the receiver is exploited to increase the number of features to achieve a high authentication rate (AR). OCC-SVM is used with the magnitude and real and imaginary parts of the received signal at each receive antenna as features. One-class classification (OCC) is a ML technique for outlier and anomaly detection which uses only legitimate training data. The sounding reference signal (SRS) in the 5G uplink radio frame is employed to obtain the features. The proposed scheme is evaluated in an urban environment under different mobility conditions using linear, sigmoid, and polynomial OCC-SVM kernels. Results are presented which show that this scheme provides a high AR with sufficient antenna diversity. Further it is superior to other approaches in the literature.

2.1 Related Work

2.1.1 Traditional PLA

Two PLA approaches were proposed in [38]. The first is channel state information (CSI) based authentication which estimates the CSI between legitimate users using embedded watermarking codes [39], encrypted CSI using wireless fading channel features, e.g. user locations [40], and robust hypothesis testing [41]. The second approach is authentication using wiretap codes. In this case, a random subset of codewords is chosen to implement authentication-admissible coding [42].

Some challenges to existing PLA techniques were given in [34]. These include low authentication reliability and authentication complexity in heterogeneous networks with a large number of devices and applications. PLA techniques were classified into two categories in [32]. The first includes passive approaches where physical layer features are used for authentication while the second includes active approaches where legitimate users modify their signals using a secret key. A survey of PLA techniques in wireless communication systems was presented in [33]. Channel based, radio frequency (RF) based, and watermark/fingerprint embedding schemes were discussed.

2.1.2 Machine Learning PLA

Traditional PLA schemes may not provide satisfactory performance [43]. Further, the variability and complexity of wireless channels make it difficult for these schemes to distinguish multiple users simultaneously [44]. On the other hand, machine learning (ML) PLA techniques can extract hidden features to improve the performance compared to traditional schemes. Hence, ML has been employed to improve PLA [45] by learning the variations in the channel attributes due to the wireless environment [46]. ML approaches include support vector machine (SVM) [47], one-class

nearest neighbors (OCNN) [1], and k-nearest neighbors (KNN) [48]. The main benefit of using channel attributes for authentication is that they are difficult to imitate. The features employed include channel impulse response (CIR) [49, 50], carrier frequency offset (CFO) [51], and received signal strength (RSS) [52, 53].

In [54], Dyna architecture and prioritized sweeping (Dyna-PS) based and Q-learning based spoofing detection schemes were developed for MIMO systems. The proposed authentication algorithms were implemented using software radio and tested in an indoor setting. It was shown that Dyna-PS reduces the spoofing detection error rate compared to Q-learning. In [2], a two-class classification SVM (TCC-SVM) utilizing RSS, time of arrival (TOA), and correlation of cyclic features was adopted for a stationary indoor environment. A dataset of over 9,000 link signatures in an office environment was used in [55] for evaluation. However, collecting data from illegitimate users for training is not practical. In [56], a PLA technique was proposed based on the AdaBoost algorithm. This technique uses orthogonal frequency division multiplexing (OFDM) sub-carrier amplitudes with or without phase. However, data from illegitimate users is used for training which is not typically available.

In [48], features for a mobile indoor environment were obtained using the National Institute of Standards and Technology (NIST) industrial dataset [57]. A PLA scheme was proposed utilizing decision trees, SVM, KNN, and ensemble learning algorithms. The channel matrix dimensions were investigated considering the authentication accuracy and computational complexity. In [58], a PLA scheme was proposed for adaptive authentication using the RSS. A neural network was employed with a feature matrix updated to track time-varying channel attributes. However, training data from both legitimate and illegitimate users is used which may not be possible in a real scenario. In [59], a PLA framework was proposed which uses the mean and ratio of the received signals as features. This framework employs TCC-SVM and a one-class classifier support vector machine (OCC-SVM). OCC-SVM was

determined to be more suitable for real scenarios.

In [1], OCC-SVM and OCNN classifiers were used for PLA with the real and imaginary parts of the channel coefficients for each sub-carrier as features. It was determined that the probability of missed detection improves with the number of subcarriers, but the probability of false alarm increases slightly. In [46], an adaptive PLA scheme was proposed utilizing the CIR, CFO, and received signal strength indicator (RSSI) as features. This scheme employs a Gaussian kernel to track variations in these attributes. The authentication performance was shown to improve with the number of features. In [60], a multi-dimensional adaptive PLA scheme was proposed using a fuzzy model for the attributes which have uncertainties due to the nature of wireless channels. The CIR, CFO, RSSI, and in-phase-quadrature-phase imbalance (IQI) were used as features and the scheme was evaluated in a low-mobility urban environment.

2.2 Contributions

The main contributions of this chapter are as follows.

- An adaptive physical layer authentication scheme using OCC-SVM is proposed which utilizes antenna diversity. The magnitude and real and imaginary parts of the received signal from each antenna are used as features.
- The proposed scheme is more practical than other PLA solutions because it is based on the 5G frame structure and does not require training data from illegitimate users.
- Few samples are considered for training and adaptively update them, which is more practical for the dynamic changes of the wireless channel environment.
- The proposed scheme is evaluated and shown to be effective under different urban environments.

- The AR for the selection combining (SC), equal gain combining (EGC), and maximum ratio combining (MRC) diversity combining techniques is evaluated.

2.3 System Model

The system model for the proposed authentication scheme is illustrated in Figure 2.1. In this model, Alice (legitimate user) is trying to achieve authentication at Bob while an arbitrary number of Eves (illegitimate users) are trying to impersonate Alice. Both Alice and the Eves are assumed to be mobile. On the other hand, Bob is stationary, e.g. a base station (BS), and tries to determine the legitimacy of Alice and reject the Eves. Alice and the Eves are assumed to have a single transmit antenna while Bob has N receive antennas. Bob must decide between the two hypotheses

$$\begin{cases} \mathcal{H}_0: & \text{Alice transmitting,} \\ \mathcal{H}_1: & \text{Eve transmitting.} \end{cases} \tag{2.1}$$

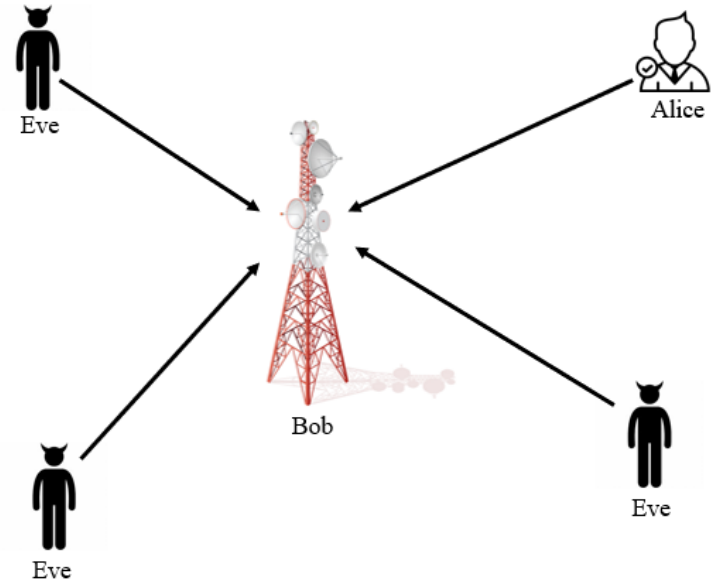

Figure 2.1: System model.

The proposed scheme is employed on the uplink and is divided into an initial phase (T_1) and subsequent phases ($T_2, T_3, \ldots, T_n$). In the initial phase, ULA authentication is performed and OCC-SVM training is conducted at Bob using the received SRS signal from Alice. OCC-SVM is illustrated in Appendix A. Bob obtains the magnitude and real and imaginary parts of the received signal at each antenna. Thus, there are 3 features per antenna and $3N$ features in total where N is the number of receive antennas. In subsequent phases, OCC-SVM testing at Bob is conducted to establish the legitimacy of the corresponding received signals which could be from Alice or Eve. Therefore, the received signals in these phases are considered to be from an unknown user. If the test is passed, the features are updated and OCC-SVM training is repeated. Conversely, if the test is failed, the connection is terminated.

OCC-SVM training in the initial phase determines the authentication boundary for the features from Alice. In subsequent phases, OCC-SVM testing is conducted to determine if the corresponding received signal features are located within this boundary. The user is accepted as legitimate if this test is passed. On the other hand, if the test decision is outside the boundary, the user is rejected.

Due to the spatial independence of Alice and the Eves, the physical layer attributes used for authentication can be considered independent. Thus, the goal of the proposed scheme is to employ wireless channel characteristics as features and obtain a large number of features using antenna diversity. Moreover, these attributes change slowly over time due to factors such as mobility [60, 61, 62, 63], so the OCC-SVM authentication boundary is updated in each subsequent phase to provide reliable authentication.

2.4 The Proposed PLA Scheme

The proposed scheme employs OCC-SVM with antenna diversity using the SRS. The SRS is a signal transmitted by users in the uplink of the 5G frame structure that allows gNodeB to estimate the CSI of users. Figure 2.2 shows the 5G frame structure with slot configuration 0 (14 symbols per slot), and numerology 2 (4 slots per subframe). A 5G radio frame is composed of 10 subframes and has a 10 ms duration so a subframe has duration 1 ms. A subframe contains a number of slots based on the numerology. A SRS is sent over 1, 2, or 4 consecutive OFDM symbols within the last six symbols in the slot and repeats every 2 (28 symbols) or 4 (56 symbols) slots [64]. The SRS is used to obtain the data for OCC-SVM training and testing.

The received signal at Bob can be expressed as

$$\mathbf{y} = x\mathbf{h} + \mathbf{n}, \tag{2.2}$$

where x is the transmitted SRS symbol, $\mathbf{h}$ is the channel vector, and $\mathbf{n}$ is a vector

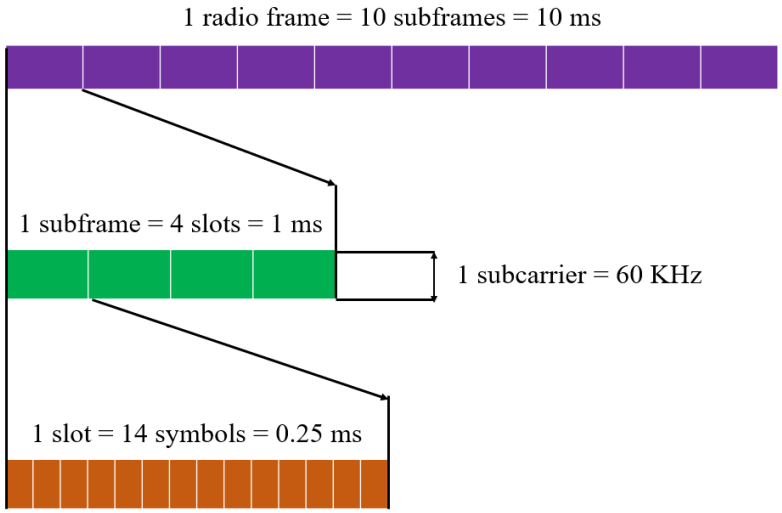

Figure 2.2: 5G frame structure with configuration 0 and numerology 2.

whose elements are independent additive white Gaussian noise (AWGN). For N receive antennas we have

$$\mathbf{y} = y_1, y_2, \ldots, y_N, \tag{2.3}$$

where

$$y_i = xh_i + n_i,$$

and h_i and n_i are the corresponding channel coefficients and AWGN, respectively. A non-line of sight (NLOS) urban environment is considered with independent channels from a user to each receive antenna. This can be assumed due to antenna spacing. The multipath channel model can then be expressed as [65]

$$h(\tau, t) = \sum_{j=1}^{J} a_j(t)\delta(\tau - \tau_j(t)), \tag{2.4}$$

where J is the number of multipath propagation paths, and τ_j and a_j are the propagation delay and attenuation for the jth path, respectively. The magnitude M and real R and imaginary I parts of the elements of $\mathbf{y}$ are used to obtain the features which gives the data vector

$$\mathbf{m} = [R\ I\ M]^1, [R\ I\ M]^2, \ldots, [R\ I\ M]^N, \tag{2.5}$$

Figure 2.3 presents the proposed authentication scheme. In the initial phase T_1, a user is authenticated using ULA. After ULA authentication, data is collected from the legitimate user for OCC-SVM training. The data legitimacy in this phase is guaranteed via the higher layer protocols [1, 60], e.g. 3GPP standard protocols [66]. Then, in subsequent phases data from the SRS signals received from an unknown user is used by Bob for testing and training. If the test is passed in a given phase, the corresponding data is used to update the features for training. A sliding window is used for this update so the oldest data is discarded. On the other hand, if the

test fails, the connection is terminated.

In phase T_1, authentication is first performed with the legitimate user through

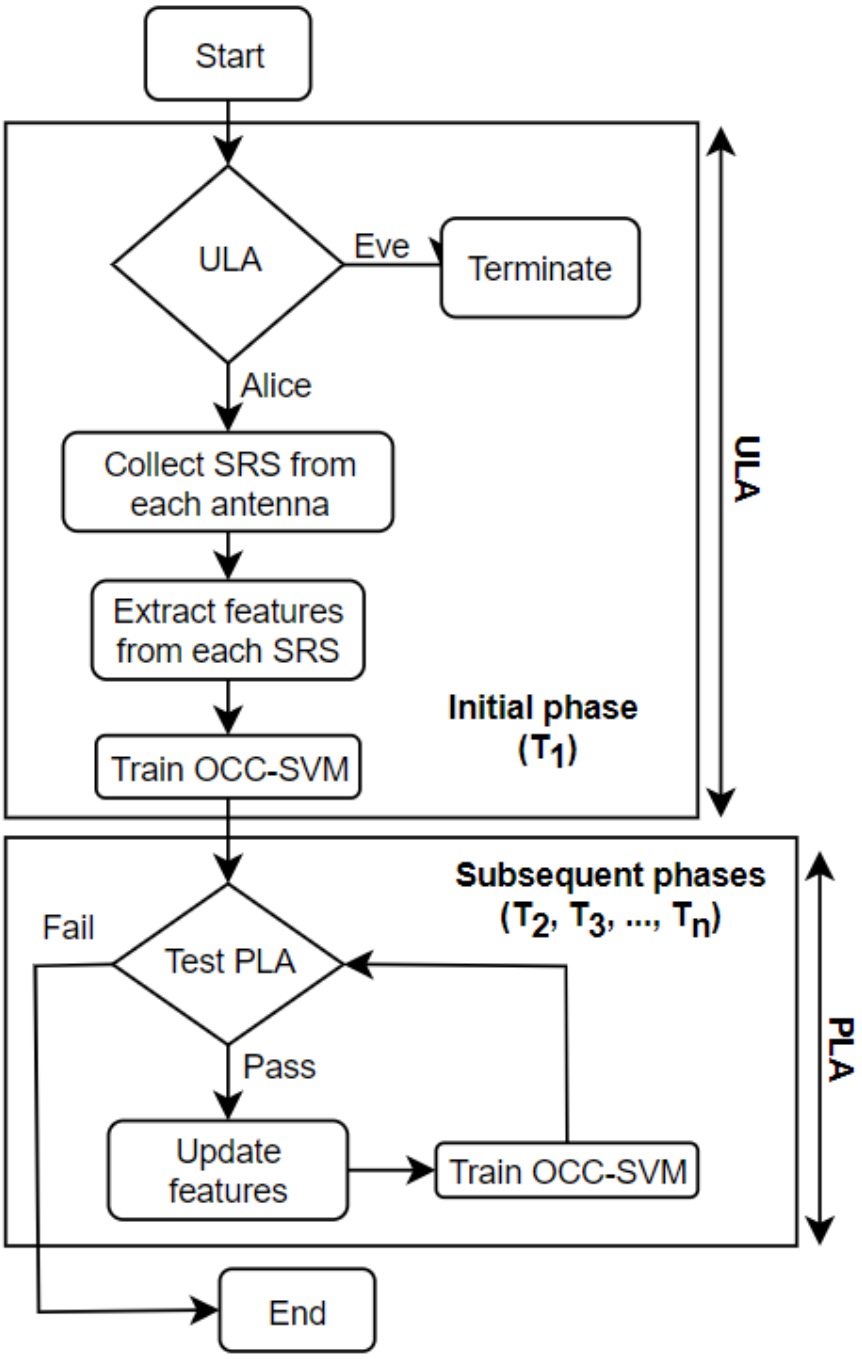

Figure 2.3: Authentication scheme flowchart.

Algorithm 1 The proposed scheme

Authenticate legitimate user through ULA.
Collect the received SRS symbols from each antenna of the legitimate user.
Extract M, R, and I from these symbols.
Perform Min-Max scaling to obtain $\mathbf{g}_i$.
Form the training matrix $\mathbf{G}$ in (2.12).
Train OCC-SVM to obtain the authentication boundary.
Test OCC-SVM with $\mathbf{t}$ which is obtained by scaling $\mathbf{b}$.
while $(f(\mathbf{t}) > 0)$ **do**
    Accept the user.
    Update the feature vectors.
    Retrain OCC-SVM.
    Test OCC-SVM.
end while

ULA. Then, ℓ received SRS symbols from A are used to construct the data vectors

$$\mathbf{d}_i = [R_i I_i M_i]^1, [R_i I_i M_i]^2, \ldots, [R_i I_i M_i]^N, i = 1, 2, \ldots, \ell, \tag{2.6}$$

where N is the number of antennas so there are $3N$ features in each vector. These vectors are scaled and used for OCC-SVM training to determine the decision function that characterizes the authentication boundary. In subsequent phases, OCC-SVM is used to test (after scaling), new data vectors

$$\mathbf{b} = [R\ I\ M]_U^1, [R\ I\ M]_U^2, \ldots, [R\ I\ M]_U^N, \tag{2.7}$$

from an unknown user U using (A.9) where U could be the legitimate user A or an illegitimate user E. If the test is passed the user is accepted, the features are updated, and OCC-SVM is retrained. However, if the test fails, the connection is terminated.

Figure 2.4 shows the sliding window update process for the data where the rows are the data vectors. In phase T_1, the training data from A is a matrix with dimensions $\ell \times 3N$. Then, in phase T_2 a new data vector $\mathbf{b}$ from (3.6) is tested (after

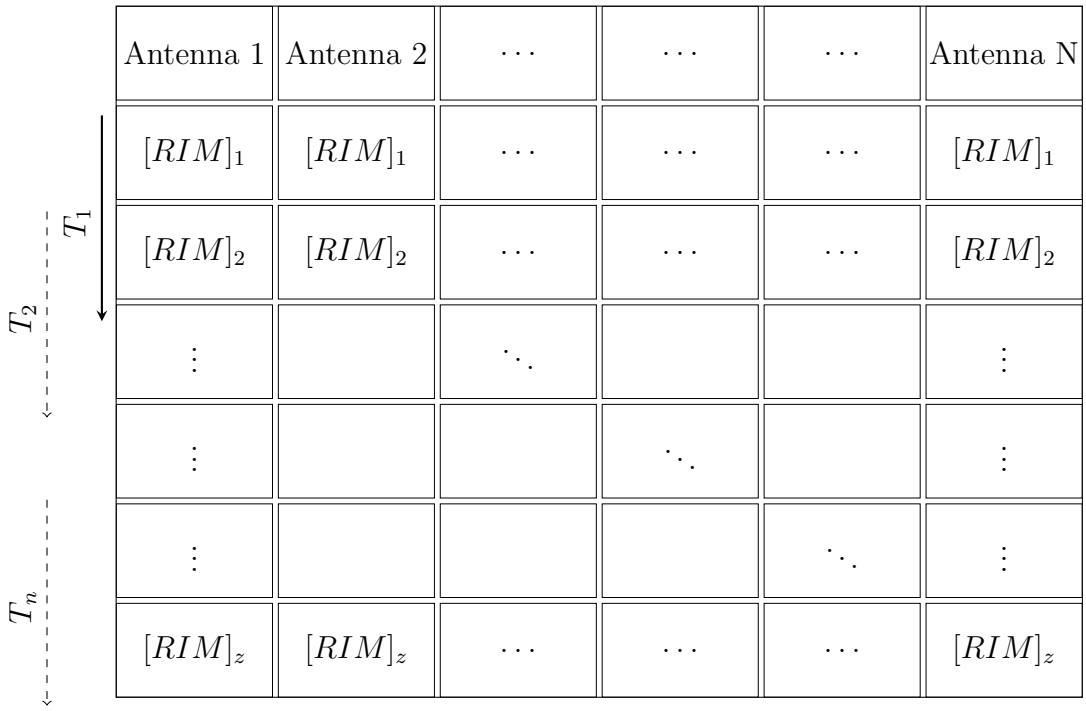

Figure 2.4: Sliding window for feature updates.

scaling), and if accepted the data matrix is updated by discarding the first row $\mathbf{d}_1$ and adding the new data vector as row $\ell + 1$. Thus, if the first e new data vectors are accepted, the training data matrix is

$$\mathbf{M}_e = \begin{bmatrix} \mathbf{d}_{1+e} \\ \mathbf{d}_{2+e} \\ \vdots \\ \mathbf{d}_{\ell+e} \end{bmatrix}, \tag{2.8}$$

as shown in Figure 2.4 so ℓ vectors are used for training. Min-Max scaling is separately applied to each of the $3N$ columns of this matrix

$$\mathbf{m}_r = [m_{1,r}\ m_{2,r}\ \ldots\ m_{\ell,r}]^T, r = 1, 2, \ldots, 3N, \tag{2.9}$$

to obtain

$$\mathbf{g}_r = [g_{1,r}\ g_{2,r}\ \ldots\ g_{\ell,r}]^T, r = 1, 2, \ldots, 3N. \tag{2.10}$$

where

$$g_{i,r} = \frac{m_{i,r} - m_{min,r}}{m_{max,r} - m_{min,r}}, \tag{2.11}$$

$m_{min,r}$ is the minimum value in $\mathbf{m}_r$, and $m_{max,r}$ is the maximum value in $\mathbf{m}_r$. Then, the feature matrix $\mathbf{G}$ used in training is

$$\mathbf{G} = [\mathbf{g}_1\ \mathbf{g}_2\ \ldots\ \mathbf{g}_{3N}] \tag{2.12}$$

The elements of the new data vector $\mathbf{b}$ are scaled using $m_{min,r}$ and $m_{max,r}$ from the training process as follows

$$t_r = \frac{b_r - m_{min,r}}{m_{max,r} - m_{min,r}}, \tag{2.13}$$

to form the test vector $\mathbf{t}$. The proposed scheme is summarized in Algorithm 1.

2.5 Simulation Results

In this section, the proposed scheme is evaluated in multipath fading channels using Monte-Carlo simulation. BPSK modulation is considered and the WINNER II channel model for NLOS urban environments is employed [67]. The simulation scenario is shown in Figure 2.5. The number of symbols used for training is $\ell = 10$. Evaluation metrics used in analysis is illustrated in Appendix B. Seven different velocities are considered ranging from a human walking (0.4 km/h) to a car (60 km/h). OCC-SVM is implemented using the scikit-learn library in Python with linear, sigmoid, and polynomial kernels. The number of trials is 2500, the number of SRS symbols from each user per trial is 500, and there are 3 Eves so $\gamma = \frac{1}{3}$. The simulation parameters are given in Table 2.1.

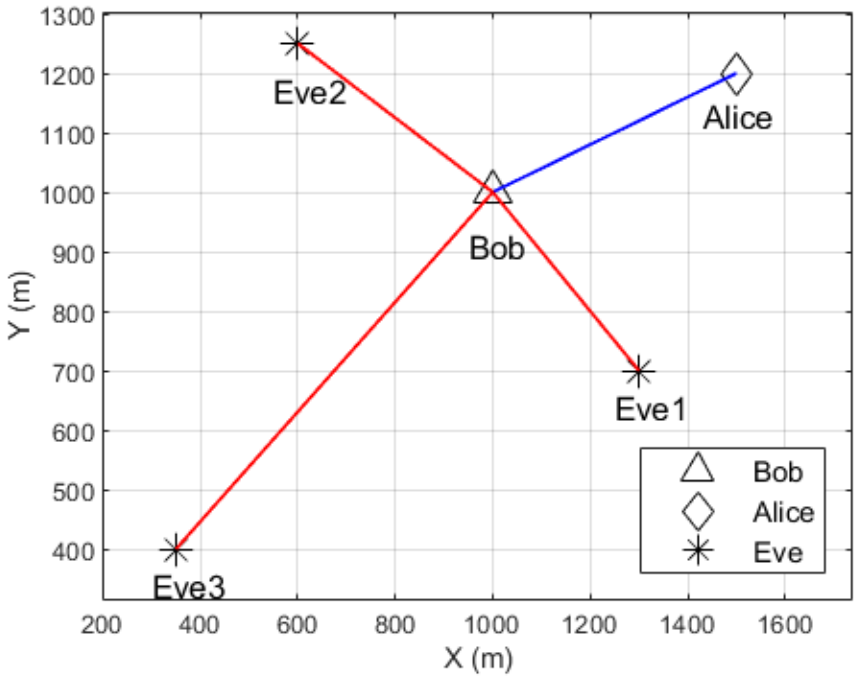

Figure 2.5: Urban environment simulation scenario.

Table 2.1: Simulation Parameters

Parameter	**Value**
Carrier frequency	5 GHz
Number of Tx Antennas for Alice/Eve	1
Number of Rx Antennas for Bob	1, 4, 8
Velocity	[0.4-60] km/h
SNR	8, 10, 12, 15 dB
Position of Bob	1000, 1000 m
Initial position of Alice	1500, 1200 m
Initial position of Eve 1	1300, 700 m
Initial position of Eve 2	600, 1250 m
Initial position of Eve 3	350, 400 m
Number of trials	2500
Number of SRS symbols per user per trial	500
ℓ	10
SRS every	20, 28, 56 symbols

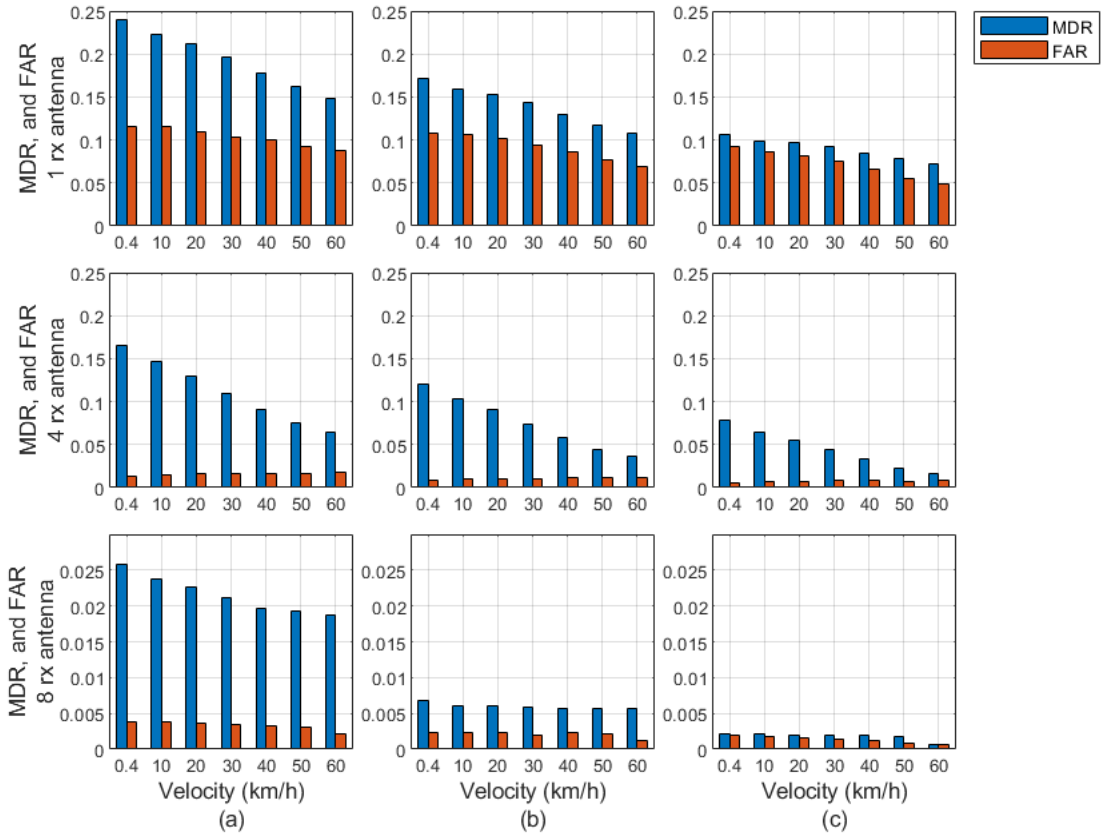

Figure 2.6: MDR and FAR versus velocity with linear kernel OCC-SVM, SRS every 20 symbols, 1 transmit and 1, 4, and 8 receive antennas, for (a) SNR = 8 dB, (b) SNR = 10 dB, and (c) SNR = 12 dB.

The MDR and FAR versus velocity with linear kernel OCC-SVM, SRS every 20 symbols, 1 transmit and 1, 4 and 8 receive antennas, and SNR = 8 dB, 10 dB, and 12 dB are given in Figures 2.6a, 2.6b, and 2.6c, respectively. These results show that the MDR and FAR improve with the number of antennas. Figure 2.6c indicates that at a velocity of 0.4 km/h, the MDR decreases from 0.106 with 1 antenna to 0.078 with 4 antennas and 0.002 with 8 antennas, and the FAR decreases from 0.092 with 1 antenna to 0.006 with 4 antennas and 0.002 with 8 antennas. The MDR and FAR decrease for all other velocities (10, 20, 30, 40, 50, 60 km/h), and SNRs when the number of antennas is increased. Moreover, for a given number of antennas and

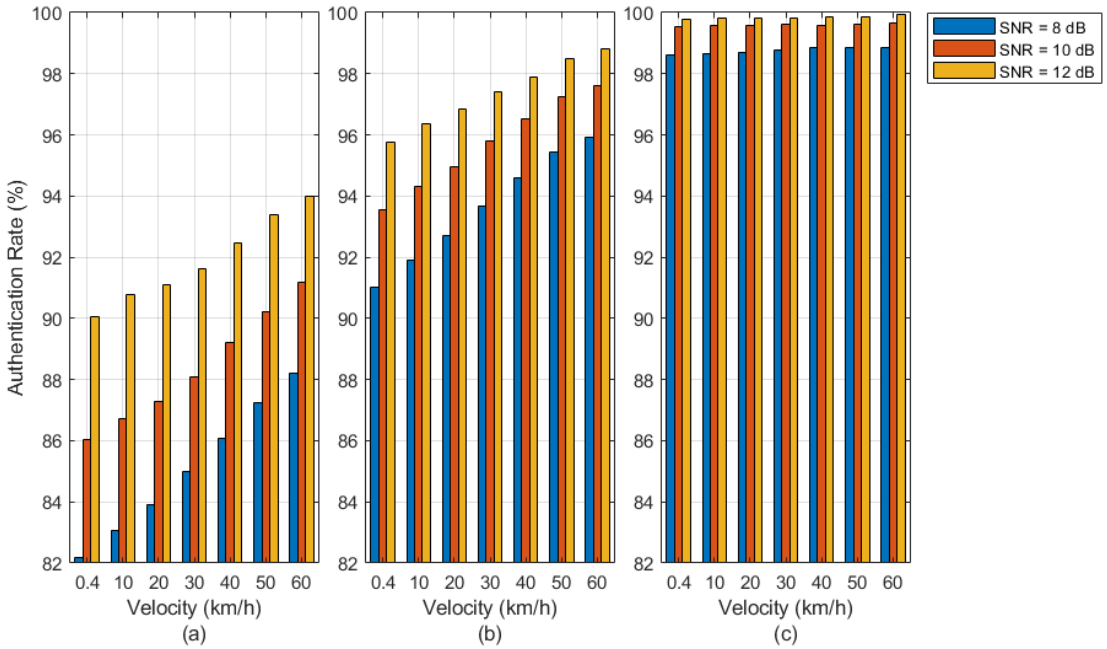

Figure 2.7: AR versus velocity with linear kernel OCC-SVM, SRS every 20 symbols, and SNR = 8 dB, 10 dB, and 12 dB for (a) 1 receive antenna, (b) 4 receive antennas, and (c) 8 receive antennas.

SNR, the performance improves with increasing velocity because this increases the Doppler shift which improves the discrimination between users.

The AR versus velocity with linear kernel OCC-SVM, SRS every 20 symbols, SNR = 8 dB, 10 dB, and 12 dB, and 1 transmit antenna for 1, 4, and 8 receive antennas is given in Figures 2.7a, 2.7b, and 2.7c, respectively. This shows that the AR improves with an increase in the number of receive antennas. For example, at SNR = 12 and 60 km/h, the AR is 94% with 1 antenna and increases to 98.8% with 4 antennas and 99.9% with 8 antennas. This is because the number of features is increased from 3 to 12 to 24. In addition, the AR increases with SNR. For example, at a velocity of 30 km/h, the AR is 93.7% at SNR = 8 dB, 95.8% at SNR = 10 dB, and 97.4% at SNR = 12 dB. This improvement in AR with increasing velocity is due to the greater Doppler shift.

Figure 2.8 presents the AR versus velocity with linear kernel OCC-SVM, SRS

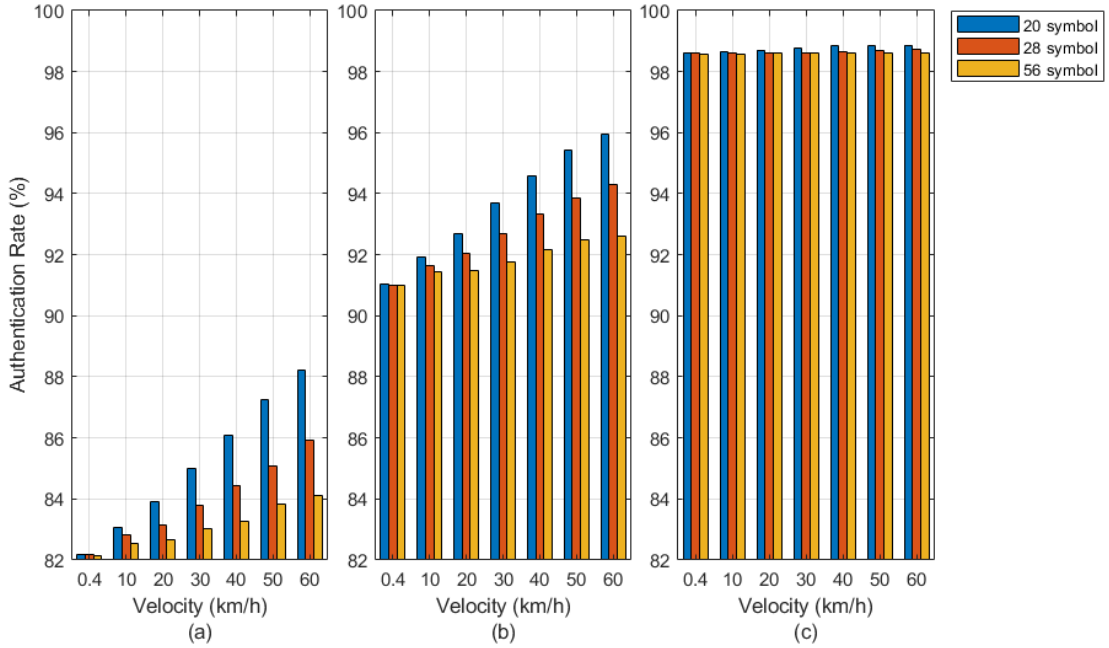

Figure 2.8: AR versus velocity with linear kernel OCC-SVM, SNR = 8 dB, SRS every 20, 28, and 56 symbols for (a) 1 receive antenna, (b) 4 receive antennas, and (c) 8 receive antennas.

every 20, 28, and 56 symbols, SNR = 8 dB, and 1, 4, and 8 receive antennas. Note that the 5G radio frame has an SRS every 28 or 56 symbols. Figure 2.8a shows the AR with 1 receive antenna. This indicates that the AR increases with velocity regardless of the SRS symbol frequency. As the gap between SRS symbols increases, the AR decreases slightly at low velocities. Further, the AR difference is greater at high velocities, e.g. at 60 km/h the AR decreases from 88.2% for SRS every 20 symbols to 85.9% for SRS every 28 symbols and 84.1% for SRS every 56 symbols. The AR improves with the number of receive antennas for all SRS symbol frequencies, and the AR difference between SRS every 20, 28, and 56 symbols decreases with the number of receive antennas, especially at high velocities, e.g. the AR difference between SRS every 20 and every 56 symbols at 60 km/h with 1 receive antenna is 4.1%, 4 antennas is 3.3%, and 8 antennas is 0.26%.

Figure 2.9 presents the AR versus velocity for SRS every 28 symbols, SNR = 8

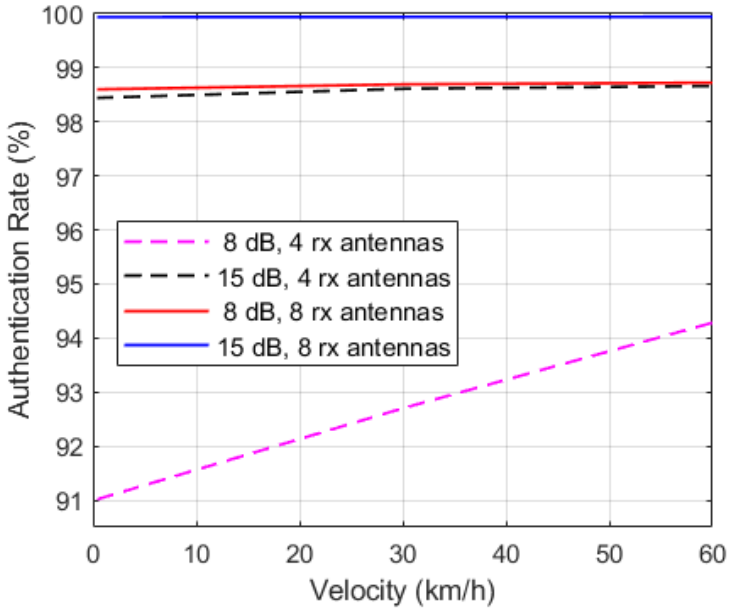

Figure 2.9: AR versus velocity with SNR = 8 dB and 15 dB, SRS every 28 symbols, and 4 and 8 receive antennas.

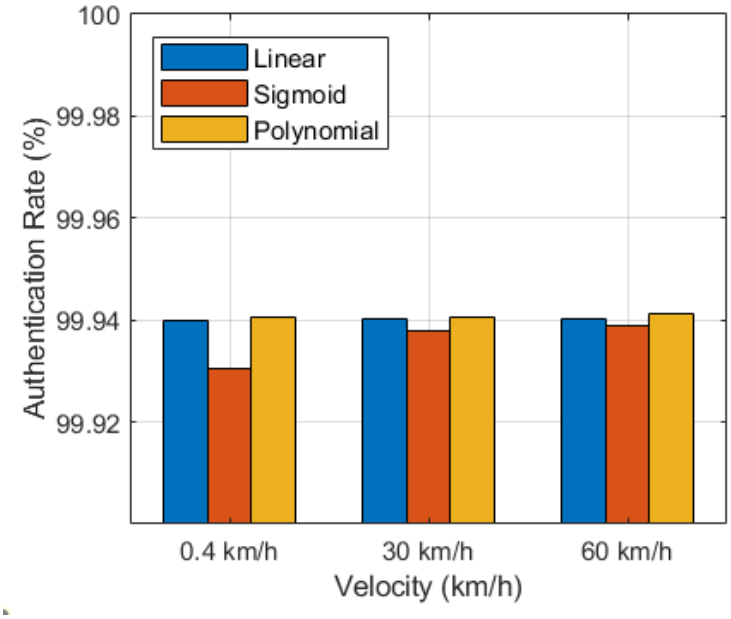

Figure 2.10: AR versus velocity with SNR = 15 dB, SRS every 28 symbols, and 8 receive antennas for OCC-SVM using linear, sigmoid, and polynomial kernels.

dB and 15 dB, and 4 and 8 receive antennas. This shows that the AR improves with increasing SNR, e.g. for 8 receive antennas and 60 km/h, the AR is greater than 99.9% for SNR = 15 dB compared to 98.7% for SNR = 8 dB. Figure 2.10 gives the AR versus velocity for SNR = 15 dB, SRS every 28 symbols, 8 receive antennas, and OCC-SVM using linear, sigmoid, and polynomial kernels. The polynomial kernel provides the highest AR, followed by the linear kernel, and then the sigmoid kernel. The difference in results for the kernels, except for the sigmoid kernel at the lowest velocity, is within 0.005% for the same velocity, which indicates that using a less complex kernel is the best choice.

2.5.1 Diversity Combining

Figure 2.11 presents the MDR and FAR versus velocity with linear kernel OCC-SVM, SRS every 28 symbols, and SNR = 8 dB for SC, EGC, and MRC using 4 receive antennas and without combining using 1 and 4 receive antennas. This shows that diversity combining has a greater effect on the MDR than the FAR for all velocities. The MDR at 30 km/h is 0.864 for MRC, 0.649 for SC and 0.581 for EGC compared to 0.247 for 1 receive antenna and 0.095 for 4 receive antennas without combining. Thus, diversity combining degrades the MDR. Conversely, the FAR for 1 receive antenna is similar to that with combining as the difference is 0.114 to 0.135 at 30 km/h. However, for 4 receive antennas without combining the FAR is much lower at 0.024.

Figure 2.12 presents the AR versus velocity with linear kernel OCC-SVM, SRS every 28 symbols, SNR = 8 dB for SC, EGC, and MRC using 4 receive antennas and without combining using 1 and 4 receive antennas. This shown that the AR without combining is better than that with diversity combining as the difference is at least 17%. The reason for the poor performance with diversity combining is that combining makes the features of the users less different so it is more difficult to discriminate between them. To illustrate this, the values for 500 trials at 0.4 km/h

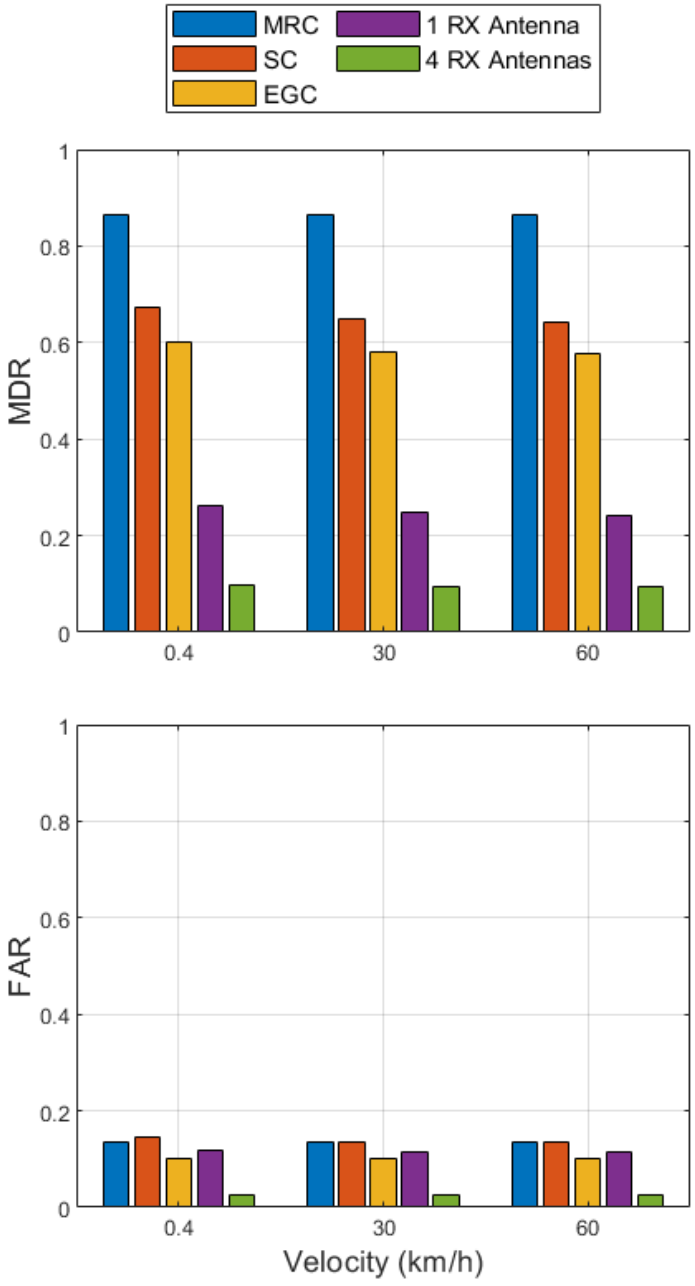

Figure 2.11: MDR and FAR versus velocity with linear kernel OCC-SVM, SRS every 28 symbols, SNR = 8 dB for selection combining (SC), equal gain combining (EGC), and maximum ratio combining (MRC) using 4 receive antennas and without combining using 1 and 4 receive antennas.

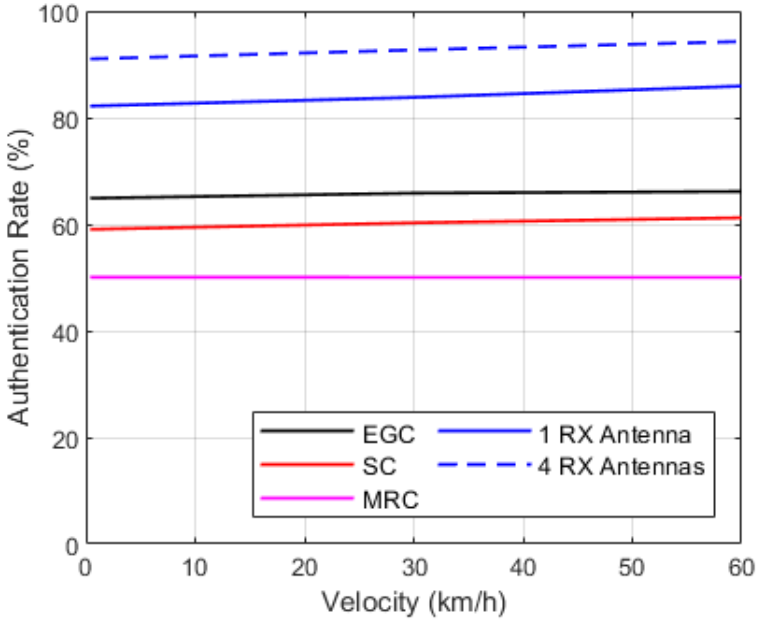

Figure 2.12: AR versus velocity with linear kernel OCC-SVM, SRS every 28 symbols, SNR = 8 dB for selection combining (SC), equal gain combining (EGC), and maximum ratio combining (MRC) using 4 receive antennas and without combining using 1 and 4 receive antennas.

each with 500 feature vectors for Alice $\mathbf{t}_A$ and 1500 for Eve $\mathbf{t}_E$ were obtained using MRC, SC, EGC, and no diversity combining with 4 receive antennas. The averages of the absolute values of the differences between the feature vector elements were determined using

$$|\mathbf{t}_A - \mathbf{t}_E| = \frac{|t_{1_A} - t_{1_E}| + |t_{2_A} - t_{2_E}| + \ldots + |t_{p_A} - t_{p_E}|}{p},$$

where $p = 3$ with diversity combining and $p = 12$ without diversity combining. The averages for the 500×1500 vector differences for MRC, SC, EGC, and without diversity combining are 0.19, 0.33, 0.35, and 0.61, respectively. Thus, the smallest difference is with MRC indicating that the feature vectors are the most similar while the greatest difference is with no diversity combining indicating that the feature vectors are the least similar. These results clearly show the disadvantage of using

diversity combining in the proposed scheme.

2.5.2 Comparison of PLA Techniques

Figure 2.13 presents AR versus velocity with linear kernel OCC-SVM, SNR = 8 dB, SRS every 28 symbols, and 4 and 8 receive antennas using the real and imaginary parts of the received signals as features [1], the RSS feature [2], and the magnitude and real and imaginary parts of the received signals as features (proposed scheme). This shows that the AR with the proposed scheme is the highest for all velocities. For example, the AR at 0.4 km/h with 4 receive antenna using the magnitude and real and imaginary parts of the received signals as features is 91%, while the AR using the real and imaginary parts of the signals as features is 89.2%, and it is only 56.7% using the RSS. Moreover, the AR at 60 km/h with 8 receive antenna using the magnitude and real and imaginary parts of the received signals as features is 98.7%,

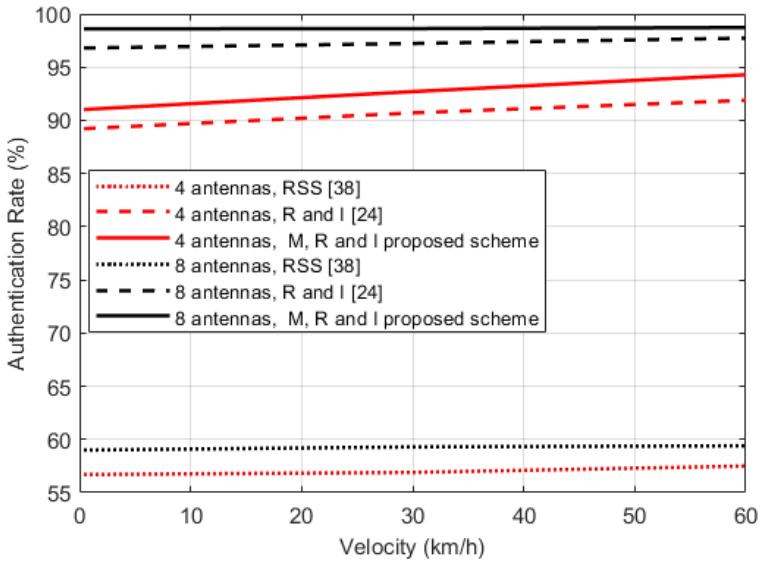

Figure 2.13: AR versus velocity with linear kernel OCC-SVM, SRS every 28 symbols, SNR = 8 dB, and 4 and 8 receive antennas using the real and imaginary parts of the received signals as features [1], the RSS feature [2], and the magnitude and real and imaginary parts of the received signals as features (proposed scheme).

while the AR using real and imaginary parts of the received signals as features is 97.7% and it is only 59.4% using the RSS feature. Although the magnitude is not independent of the real and imaginary parts of the received signal, using it as an additional feature provides a small improvement in performance.

2.5.3 System Complexity

The system complexity is evaluated considering the average number of iterations, the number of samples used for training, and the execution time [59]. Previous work considered a large number of iterations for training, which is impractical for a mobile wireless channel environment. The authentication scheme in [46] requires 30 iterations to converge while the approach in [60] needs about 50 iterations. On the other hand, the average number of iterations required for the proposed scheme to converge is only 2. The proposed scheme employs OCC-SVM in Python using the scikit-learn library and takes on average 5.6 us, 63 us, and 70 us for training and testing using linear, sigmoid, and polynomial kernels, respectively. These times are low because only 10 samples are used for training. However, the scheme in [1] uses training sets of sizes 100 and 1000 samples while the approach in [59] uses 200 samples for training.

The proposed scheme has no communication overhead because it utilizes the SRS symbols in the 5G frame structure, so extra symbols are not required for PLA. As in the related literature, OCC-SVM is used as an efficient and practical ML model for PLA. For example, the approaches in [1, 59] employ OCC-SVM, but the proposed scheme has lower computational complexity because it uses fewer training samples and requires fewer iterations to converge. Furthermore, this model is simpler than the convolution neural network (CNN) in [58], the deep reinforcement learning in [68], and the deep neural network in [69], and these models have numerous layers with many nodes per layer that require a large amount of data (and thus time), for training. One reason for the good performance is the use of a sliding window

for feature updates, which allows the proposed scheme to adapt to time-varying environments.

2.6 Conclusion

An adaptive lightweight physical layer authentication (ALPLA) scheme using machine learning (ML) was proposed. This scheme exploits antenna diversity to improve the authentication rate (AR). One-class classifier support vector machine (OCC-SVM) with linear, sigmoid, and polynomial kernels was used. The magnitude and real and imaginary parts of the received sounding reference signal (SRS) at each receive antenna are used as features. Thus, the number of features is $3N$, where N is the number of receive antennas. Results were presented which show that the AR increases with the number of antennas. A higher velocity increases the Doppler shift which provides better discrimination between users. For example, the AR with the proposed scheme using eight antennas exceeds 99.9% which indicates robust authentication performance. Moreover, diversity combining was shown to degrade the performance as it makes it more difficult to discriminate between users. This confirms the benefits of using antenna diversity. Finally, the proposed scheme which employs the magnitude and the real and imaginary parts of the received signals as features was shown to be superior to using just the real and imaginary parts of the received signals and only the RSS.

Chapter 3

Adaptive Physical Layer Authentication for IoT in MIMO Communication Systems using Support Vector Machine

MIMO technology employing multiple antennas is used to increase the throughput of broadband wireless access [70]. It is a fundamental technology for current and future wireless networks, e.g. IoT, fifth-generation (5G), and sixth-generation (6G) [5, 6, 7]. The antennas in MIMO systems can be considered to be spatially independent due to their spacing [22]. In this chapter, this independence is exploited for IoT PLA. In this chapter, an adaptive PLA scheme is proposed which exploits the antenna diversity inherent in MIMO systems [71]. This scheme employs a OCC-SVM with the magnitude and real and imaginary parts of the received signal as features. Results are presented which show that this scheme provides robust authentication. The authentication performance is evaluated considering two majority voting schemes for IoT applications.

3.1 Related Work

Channel state information (CSI) is commonly used to obtain features for PLA. This is due to the fact that the wireless channels between different locations can be considered independent [47]. In [1], a PLA scheme was proposed which employs the real and imaginary parts of the channel coefficients for each OFDM subcarrier as features for machine learning (ML). This scheme employs a one-class classifier support vector machine (OCC-SVM) and the results obtained show that the probability of missed detection improves with the number of subcarriers, but the probability of false alarm increases slightly. In [59], a PLA scheme was proposed which uses the mean and ratio of the received signals as features. This scheme employs a two-class classifier support vector machine (TCC-SVM) and a OCC-SVM. The OCC-SVM was considered more suitable for realistic scenarios.

In [46], an adaptive PLA scheme was proposed using the channel impulse response (CIR), carrier frequency offset (CFO), and received signal strength indicator (RSSI) as features. This scheme employs a Gaussian kernel to track variations in the features. The authentication performance was shown to improve with the number of features. In [60], an adaptive PLA scheme was proposed using a fuzzy model for the features which have uncertainties due to the nature of wireless channels. The CIR, CFO, RSSI, and in-phase-quadrature-phase imbalance (IQI) were used as features.

In [72], channel-based PLA for MIMO systems was proposed. This scheme considers hardware impairments such as power amplifier nonlinearities and the aggregate effect of these impairments on authentication performance was analyzed. In [73], a PLA scheme for non-coherent single-input multiple-output (SIMO) Industrial Internet of Things (IIoT) communication systems was proposed. This scheme embeds an authentication signal into the message signal for PLA. The power allocated to the message and authentication signals was optimized considering the tradeoff in error performance. However, data from illegitimate users is used for training, but

obtaining this data is not practical.

In [74], a PLA scheme for IoT systems was proposed to identify illegitimate users. First, the angle of arrival and path gain from all devices are obtained as features. Then, ML is employed with these features to identify illegitimate users. However, data from illegitimate users is used for training which is not typically available. In [75], a PLA scheme for user authorization in an edge-computing system was proposed. A weighted voting technique is employed with ML using CSI features. This scheme requires training data from all devices.

In [1, 46, 60], it was determined that the authentication performance improves with the number of features. Moreover, fewer iterations were required for the ML algorithms to converge as the number of features increases, which is important for practical wireless networks. Thus, the goal in this paper is to exploit antenna diversity in MIMO system to obtain a large number of features. In this case, non ML based PLA schemes may not provide satisfactory performance [43].

3.2 Contributions

The contributions of this chapter are as follows.

- An adaptive PLA scheme using OCC-SVM with antenna diversity in MIMO systems is proposed.
- The proposed scheme is evaluated using linear, sigmoid, and polynomial OCC-SVM kernels.
- The magnitude and real and imaginary parts of the received signals are used as features.
- The SRS in the 5G frame structure is used to obtain the features so it is more practical than other PLA solutions. Further, training data from illegitimate users is not required.

- The proposed scheme is evaluated and shown to be effective in urban environments. This includes the situation where illegitimate users start from the same initial location as the legitimate user but move in arbitrary directions, which can be considered worst case.
- The authentication performance is evaluated considering two majority voting schemes for practical IoT applications.

3.3 System Model

The system model for the proposed authentication scheme is shown in Figure 3.1. Alice is a legitimate user trying to achieve authentication at Bob while Eve is an illegitimate user trying to impersonate Alice. Bob, represented by a network of IoT devices, employs an authentication process to determine the legitimacy of users. Consequently, Bob must decide between the two hypotheses

$$\begin{cases} \mathcal{H}_0: & \text{Alice is transmitting,} \\ \mathcal{H}_1: & \text{Eve is transmitting.} \end{cases} \tag{3.1}$$

Thus, $\mathcal{H}_0$ denotes that the signal is from Alice while $\mathcal{H}_1$ means that it is from Eve. Both Alice and Eve are assumed to be mobile and Bob is fixed. Alice and Eve are equipped with M transmit antennas, while Bob has N receive antennas.

Sounding reference signals (SRSs) are transmitted by users in the uplink of the 5G frame structure to facilitate channel estimation [64]. A 5G radio frame has a 10 ms duration and is composed of 10 subframes each with duration 1 ms. Each subframe contains a number of slots based on the numerology according to the 3GPP standard. Each slot contains 14 symbols for slot configuration 0. SRSs are sent over 1, 2, or 4 consecutive symbols within the last six symbols in the slot and repeated every 2 (28 symbols) or 4 (56 symbols) slots [64]. SRS symbols are sent every 0.5

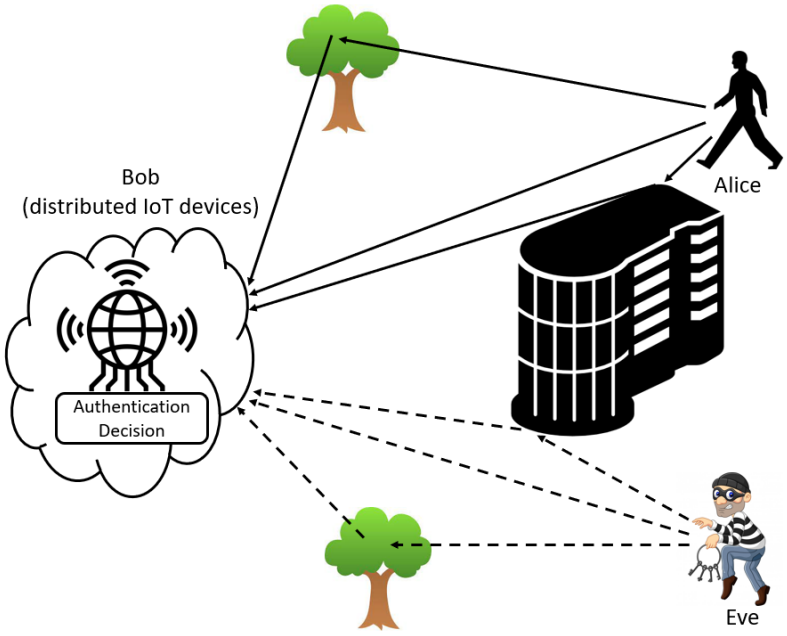

Figure 3.1: The system model.

ms for 28 symbols frequency and this case is considered here to obtain the data for OCC-SVM training and testing.

The received signal corresponding to the SRS symbol at the nth receive antenna from the mth transmit antenna can be expressed as

$$y_{nm} = x_m h_{nm} + n_{nm}, \tag{3.2}$$

where x_m is the transmitted SRS symbol, and h_{nm} and n_{nm} are the corresponding channel coefficients and additive white Gaussian noise (AWGN), respectively. The WINNER II channel model for non-line-of-sight (NLOS) urban environments is employed [67]. The received signal matrix for the MIMO system can be expressed

as

$$\mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1M} \\ y_{21} & y_{22} & \cdots & y_{2M} \\ & & \vdots & \\ y_{N1} & y_{N2} & \cdots & y_{NM} \end{bmatrix}. \tag{3.3}$$

$\mathbf{Y}$ is used to obtain the features for OCC-SVM training and testing.

The proposed scheme is employed on the uplink and is divided into an initial phase (T_1) and subsequent phases ($T_2, T_3, \ldots, T_n$). In the initial phase, upper-layer authentication (ULA) is performed and OCC-SVM training is conducted at Bob using the received signal from Alice. Bob obtains the magnitude and real and imaginary parts of $\mathbf{Y}$ so there are $3MN$ features in total. In subsequent phases, OCC-SVM testing at Bob is conducted to establish the legitimacy of the received signal which could be from Alice or Eve. Therefore, the received signals in these phases are considered to be from an unknown user. If the test is passed, the features are updated and OCC-SVM training is repeated. Conversely, if the test is failed, the connection is terminated.

The spatial independence of Alice and Eve means the physical layer characteristics used for authentication can be considered independent. The goal of the proposed scheme is to employ wireless channel characteristics as features. Moreover, these features change over time due to factors such as mobility [60, 61, 76, 62, 63], so the OCC-SVM authentication boundary is updated in each subsequent phase to provide robust authentication.

3.4 The Proposed Scheme

Figure 3.2 presents a flowchart of the proposed authentication scheme. OCC-SVM is employed with features obtained from the received signal corresponding to SRS symbols in a MIMO system. In the initial phase T_1, a user is authenticated using

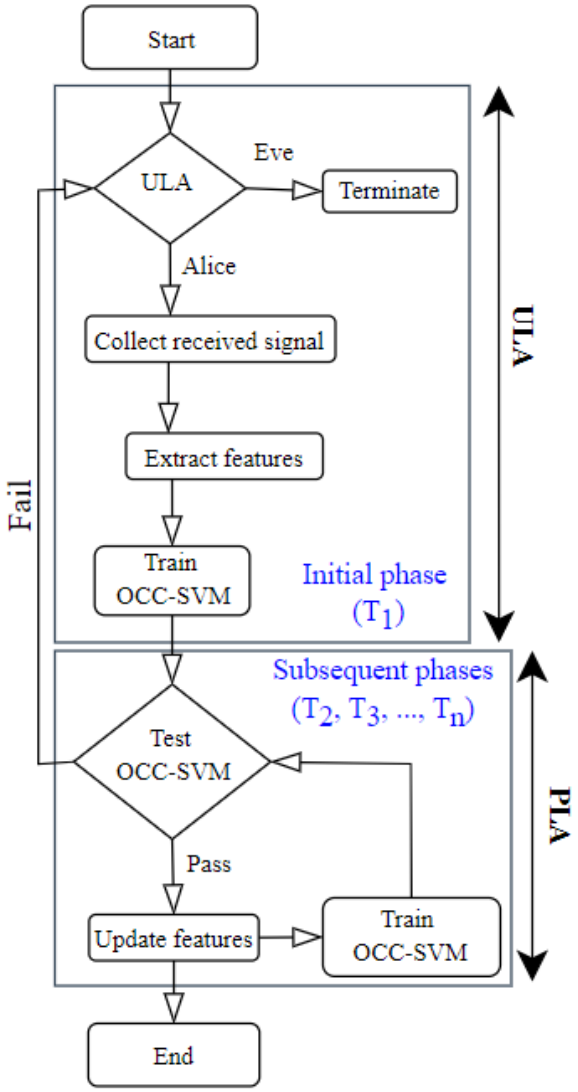

Figure 3.2: Flowchart of the proposed authentication scheme.

ULA so user legitimacy in this phase is guaranteed using an upper layer protocol [66]. If the authentication is successful, the received signal data is collected for OCC-SVM training. Then, in subsequent phases the received signal from an unknown user is used by Bob for testing. If the test is passed in a given phase, the corresponding data is used to update the features for training. A sliding window is used for this update so the oldest data is discarded. On the other hand, if the test fails, the

connection is terminated.

The proposed authentication scheme is considered with and without voting. When voting is not employed all the features obtained from the elements of $\mathbf{Y}$ in (3.3) are used in one location. However, collecting this data when there is a large distance between IoT devices may not be feasible. Hence, majority voting approach is adopted, where the features are either used separately (MN majority voting scheme) or in groups of M (N majority voting scheme). Thus, with MN majority voting, the features from each element of $\mathbf{Y}$ are used separately, while with N majority voting, the features from each row of $\mathbf{Y}$ are used together. In the following subsections, the proposed authentication scheme without voting is introduced, and then with voting.

3.4.1 Proposed Scheme Without Voting

The magnitude M and real R and imaginary I parts of the elements of $\mathbf{Y}$ in (3.3) are used to obtain the features which gives the data vector

$$\begin{aligned}\mathbf{m} = [&[R\ I\ M]^{11}, \ldots, [R\ I\ M]^{1M}, [R\ I\ M]^{21}, \ldots, \\ &[R\ I\ M]^{N1}, \ldots, [R\ I\ M]^{NM}].\end{aligned} \tag{3.4}$$

Thus, there are $3MN$ features in each vector for M transmit antennas and N receive antennas.

In phase T_1, ℓ data vectors from the legitimate user are used to construct the training data vectors

$$\begin{aligned}\mathbf{d}_i = [&[R_i\ I_i\ M_i]^{11}, \ldots, [R_i\ I_i\ M_i]^{1M}, [R_i\ I_i\ M_i]^{21}, \ldots, \\ &[R_i\ I_i\ M_i]^{N1}, \ldots, [R_i\ I_i\ M_i]^{NM}], i = 1, 2, \ldots, \ell.\end{aligned} \tag{3.5}$$

These vectors are scaled and used for OCC-SVM training to determine the decision function f. In subsequent phases, OCC-SVM is used to test (after scaling), new

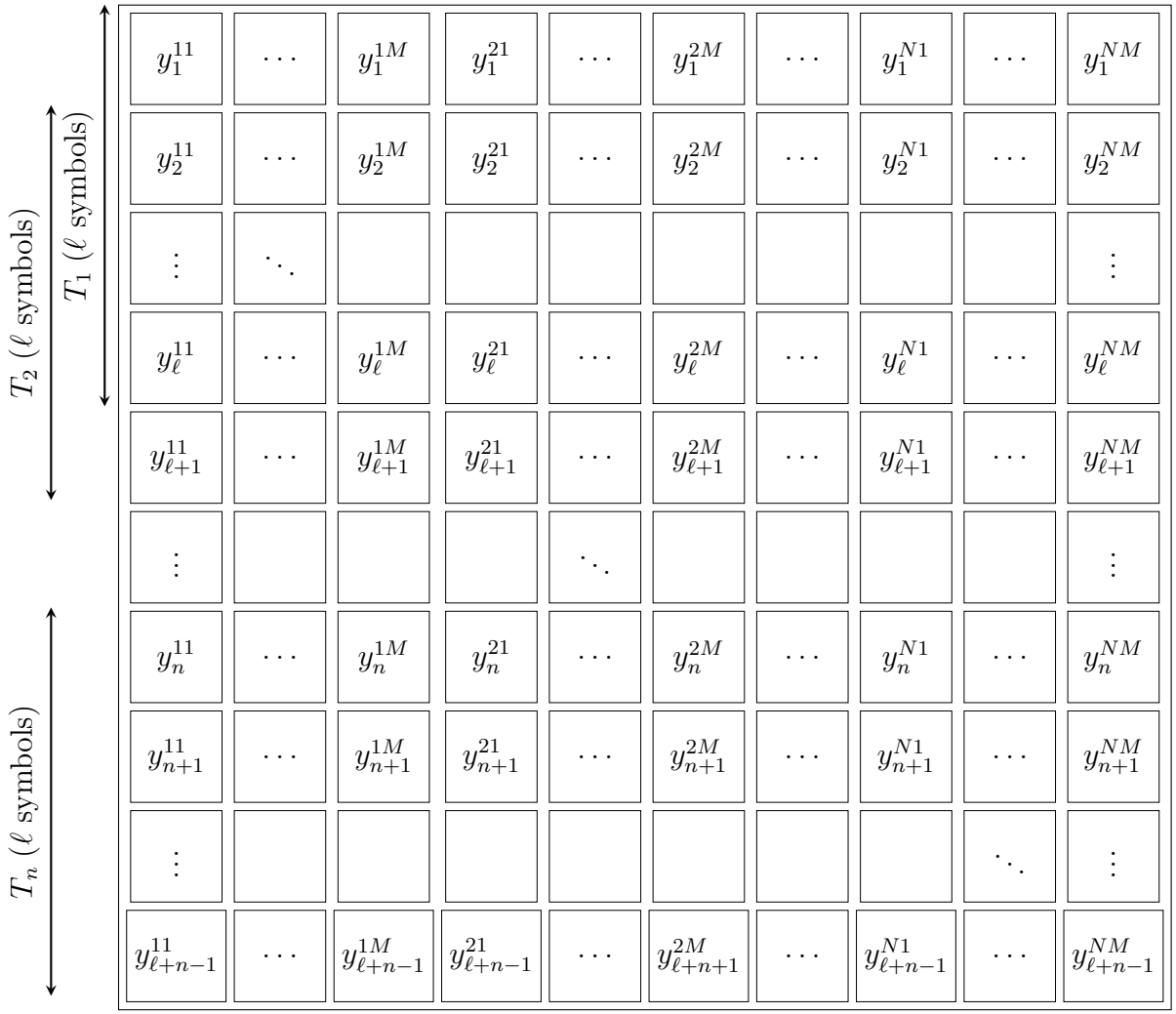

Figure 3.3: Sliding window for feature updates.

data vectors

$$\begin{aligned}\mathbf{b} = [&[R\ I\ M]_U^{11}, \ldots, [R\ I\ M]_U^{1M}, [R\ I\ M]_U^{21}, \ldots, \\ &[R\ I\ M]_U^{N1}, \ldots, [R\ I\ M]_U^{NM}],\end{aligned} \tag{3.6}$$

from an unknown user U using (A.9) where U could be Alice or Eve. If the test is passed the user is accepted, the features are updated, and OCC-SVM is retrained. However, if the test fails, the connection is terminated.

Figure 3.3 shows the sliding window update process for the data where the rows are the data vectors. In phase T_1, the training data from Alice is a matrix with

Algorithm 2 Proposed scheme without voting

Authenticate legitimate user through ULA.
Collect the received signals corresponding to the SRS symbols from the legitimate user.
Extract M, R, and I from the received signals.
Perform Min-Max scaling to obtain $\mathbf{g}_i$.
Form the training matrix $\mathbf{G}$ in (3.11).
Train OCC-SVM to obtain f.
Test the new received signal using f.
if $(f(\mathbf{t}) > 0)$ **then**
    Accept the user.
    Update the feature vectors.
    Retrain OCC-SVM.
else if $(f(\mathbf{t}) \leq 0)$ **then**
    Terminate the connection.
end if

dimensions $\ell \times 3MN$. Then, in phase T_2 a new data vector $\mathbf{b}$ is tested (after scaling), and if accepted the data matrix is updated by discarding the first row $\mathbf{d}_1$ and adding the new data vector as row $\ell + 1$. Thus, if the first e new data vectors are accepted, the training data matrix is

$$\mathbf{M}_e = \begin{bmatrix} \mathbf{d}_{1+e} \\ \mathbf{d}_{2+e} \\ \vdots \\ \mathbf{d}_{\ell+e} \end{bmatrix}, \tag{3.7}$$

so ℓ vectors are used for training. Min-Max scaling is separately applied to each of the $3MN$ columns of this matrix

$$\mathbf{m}_r = [m_{1,r}\ m_{2,r}\ \ldots\ m_{\ell,r}]^T, r = 1, 2, \ldots, 3MN, \tag{3.8}$$

to obtain

$$\mathbf{g}_r = [g_{1,r}\ g_{2,r}\ \ldots\ g_{\ell,r}]^T, r = 1, 2, \ldots, 3MN. \tag{3.9}$$

where

$$g_{i,r} = \frac{m_{i,r} - m_{min,r}}{m_{max,r} - m_{min,r}}, \tag{3.10}$$

$m_{min,r}$ is the minimum value in $\mathbf{m}_r$, and $m_{max,r}$ is the maximum value in $\mathbf{m}_r$. Then, the feature matrix used for training is given by

$$\mathbf{G} = [\mathbf{g}_1 \ \mathbf{g}_2 \ \ldots \ \mathbf{g}_{3MN}], \tag{3.11}$$

and the elements of the new data vector $\mathbf{b}$ are scaled using $m_{min,r}$ and $m_{max,r}$ from the training data as follows

$$t_r = \frac{b_r - m_{min,r}}{m_{max,r} - m_{min,r}}, \tag{3.12}$$

to form the test vector $\mathbf{t}$. The proposed scheme is summarized in Algorithm 2. Each row of Figure 3.3 is a vector used for training and testing.

3.4.2 Proposed Scheme With Voting

The proposed authentication scheme without voting (Algorithm 2) uses all $3MN$ features in the rows of the feature matrix in Figure 3.3 for training and testing at one location. However, collecting this data when there is a large distance between IoT devices may not be feasible. In this case, OCC-SVM training and testing can be performed separately at each device and then a majority vote conducted.

Two majority voting schemes are considered. In the MN majority voting scheme, OCC-SVM training and testing is performed for each feature separately so the columns in Figure 3.3 are used individually. In the N majority voting scheme, OCC-SVM training and testing is performed separately on the data at each device so groups of M columns in Figure 3.3 are used. Then, a majority vote on the MN or N testing results is used to accept or reject the user. These schemes are described below.

Algorithm 3 MN majority voting scheme

Authenticate legitimate user through ULA.
Collect the received signals corresponding to the SRS symbols from the mth antenna at the nth antenna represented by the columns in Figure 3.3.
Extract M, R, and I features from these signals.
Perform Min-Max scaling.
Form the training matrices $\mathbf{G}_j$ in (3.16).
Train OCC-SVM to obtain $f_j, j = 1, 2, \ldots, MN$.
Test the new features $\mathbf{t}_j$ using f_j.
if $(\sum_j f_j(\mathbf{t}_j) > \frac{MN}{2})$ **then**
    Accept the user.
    Update the feature vectors.
    Retrain OCC-SVM.
else if $(\sum_j f_j(\mathbf{t}_j) \leq \frac{MN}{2})$ **then**
    Terminate the connection.
end if

MN Majority Voting Scheme

The ℓ training data vectors used for the MN majority voting scheme are

$$\begin{aligned} \mathbf{d}_i^1 =& [R_i\ I_i\ M_i]^1, \\ \mathbf{d}_i^2 =& [R_i\ I_i\ M_i]^2, \\ &\vdots \\ \mathbf{d}_i^{MN} =& [R_i\ I_i\ M_i]^{MN}, \end{aligned} \quad , i = 1, 2, \ldots, \ell, \tag{3.13}$$

and the corresponding test data vectors are

$$\begin{aligned} \mathbf{b}^1 =& [R\ I\ M]_U^1, \\ \mathbf{b}^2 =& [R\ I\ M]_U^2, \\ &\vdots \\ \mathbf{b}^{MN} =& [R\ I\ M]_U^{MN}, \end{aligned} \tag{3.14}$$

Thus, if the first e new data vectors are accepted, the training data matrices are

$$\mathbf{M}_e^j = \begin{bmatrix} \mathbf{d}_{1+e}^j \\ \mathbf{d}_{2+e}^j \\ \vdots \\ \mathbf{d}_{\ell+e}^j \end{bmatrix}, j = 1, 2, \ldots, MN. \tag{3.15}$$

Min-Max scaling is applied separately to each of the 3 columns of these matrices so the feature matrices used for training are given by

$$\mathbf{G}_j = [\mathbf{g}_1 \ \mathbf{g}_2 \ \mathbf{g}_3], j = 1, 2, \ldots, MN. \tag{3.16}$$

and the elements of the new data vectors $\mathbf{b}^j, j = 1, 2, \ldots, MN$, are scaled as in (3.12) to form the test vectors $\mathbf{t}_j, j = 1, 2, \ldots, MN$. The user is accepted if the following condition is satisfied

$$\sum_j f_j(\mathbf{t}_j) > \frac{MN}{2}, j = 1, 2, \ldots, MN, \tag{3.17}$$

The MN majority voting scheme is summarized in Algorithm 3.

N Majority Voting Scheme

The ℓ training data vectors used for the N majority voting scheme are

$$\begin{aligned} \mathbf{d}_i^1 =& [[R_i \ I_i \ M_i]^1 \ldots [R_i \ I_i \ M_i]^M]^1, \\ \mathbf{d}_i^2 =& [[R_i \ I_i \ M_i]^1 \ldots [R_i \ I_i \ M_i]^M]^2, \\ & \vdots \\ \mathbf{d}_i^N =& [[R_i \ I_i \ M_i]^1 \ldots [R_i \ I_i \ M_i]^M]^N, \end{aligned} \quad , i = 1, 2, \ldots, \ell. \tag{3.18}$$

Algorithm 4 N majority voting scheme

Authenticate legitimate user through ULA.
Collect the received signals corresponding to the SRS symbols at the N antennas separately.
Extract M, R, and I from the groups of M columns in Figure 3.3.
Perform Min-Max scaling.
Form the training matrices $\mathbf{G}_k$ in (3.21).
Train OCC-SVM to obtain $f_k, k = 1, 2, \ldots, N$.
Test the new features $\mathbf{t}_k$ using f_k.
if $(\sum_k f_k(\mathbf{t}_k) > \frac{N}{2})$ **then**
    Accept the user.
    Update the feature vectors.
    Retrain OCC-SVM.
else if $(\sum_k f_k(\mathbf{t}_k) \leq \frac{N}{2})$ **then**
    Terminate the connection.
end if

and the corresponding test data vectors are

$$\begin{aligned} \mathbf{b}^1 =& [[R\ I\ M]^1 \ldots [R\ I\ M]^M]_U^1, \\ \mathbf{b}^2 =& [[R\ I\ M]^1 \ldots [R\ I\ M]^M]_U^2, \\ & \vdots \\ \mathbf{b}^N =& [[R\ I\ M]^1 \ldots [R\ I\ M]^M]_U^{MN}, \end{aligned} \tag{3.19}$$

Thus, if the first e new data vectors are accepted, the training data matrices are

$$\mathbf{M}_e^k = \begin{bmatrix} \mathbf{d}_{1+e}^k \\ \mathbf{d}_{2+e}^k \\ \vdots \\ \mathbf{d}_{\ell+e}^k \end{bmatrix}, k = 1, 2, \ldots, N. \tag{3.20}$$

Min-Max scaling is applied separately to each of the $3M$ columns of these matrices so the feature matrices used for training are given by

$$\mathbf{G}_k = [\mathbf{g}_1\ \mathbf{g}_2\ \ldots\ \mathbf{g}_M], k = 1, 2, \ldots, N. \tag{3.21}$$

and the elements of the new data vectors $\mathbf{b}^j, j = 1, 2, \ldots, N$, are scaled as in (3.12) to form the test vectors $\mathbf{t}_k, j = 1, 2, \ldots, N$. The user is accepted if the following condition is satisfied

$$\sum_k f_k(\mathbf{t}_k) > \frac{N}{2}, k = 1, 2, \ldots, N, \tag{3.22}$$

The N majority voting scheme is summarized in Algorithm 4.

3.5 Simulation Results

In this section, the proposed scheme is evaluated in multipath fading channels using Monte-Carlo simulation. BPSK modulation is considered and the WINNER II channel model for NLOS urban environments is employed [67]. Four simulation scenarios with three Eves are considered as shown in Figures 3.4 and 3.5. Bob is fixed in all scenarios while Alice moves from a fixed initial location and the Eves move from initial locations that differ in the scenarios. In Scenario 1, the Eves are far from Alice, while in Scenario 2 they are close to Alice. In Scenario 3, the Eves start from the same initial location as Alice but move in arbitrary directions, so this scenario can be considered worst case. Scenario 4 is the same as Scenario 1 except there are three Alice with different initial locations. OCC-SVM is implemented using the scikit-learn library in Python with linear, sigmoid, and polynomial kernels. The number of trials is 2500, the number of SRS symbols from each user per trial is 500, and $\gamma = \frac{1}{3}$ in Scenarios 1, 2, and 3 and $\gamma = 1$ in Scenario 4 to account for the numbers of Alice and Eves. The number of symbols used for training is $\ell = 10$. The simulation parameters are given in Table 3.1.

Table 3.1: Simulation Parameters

Parameter	**Value**
The urban model frequency band	2-6 GHz
Carrier frequency	5 GHz
Sampling rate	20 MHz
Shadow fading standard deviation	4 dB
Number of Tx Antennas for Alice/Eve	1, 2, 4
Number of Rx Antennas for Bob	1, 4, 8
Antenna height for Alice/Eve	1 m
The urban model mobility range	[0-70] km/h
Velocity	[0.4-60] km/h
SNR	8, 12 dB
Position of Bob	1000 m, 1000 m
Initial position of Alice 1	1500 m, 1200 m
Initial position of Eve 1 in Scenario 1	1300 m, 700 m
Initial position of Eve 2 in Scenario 1	600 m, 1250 m
Initial position of Eve 3 in Scenario 1	350 m, 400 m
Initial position of Eve 1 in Scenario 2	1550 m, 1100 m
Initial position of Eve 2 in Scenario 2	1300 m, 1250 m
Initial position of Eve 3 in Scenario 2	1600 m, 1000 m
Initial position of Eves in Scenario 3	1500 m, 1200 m
Initial position of Alice 2 in Scenario 4	500 m, 900 m
Initial position of Alice 3 in Scenario 4	1000 m, 600 m
Initial position of Eve 1 in Scenario 4	1300 m, 700 m
Initial position of Eve 2 in Scenario 4	600 m, 1250 m
Initial position of Eve 3 in Scenario 4	350 m, 400 m
Number of trials	2500
Number of SRS symbols per user per trial	500
SRS every	28 symbols
ℓ	10

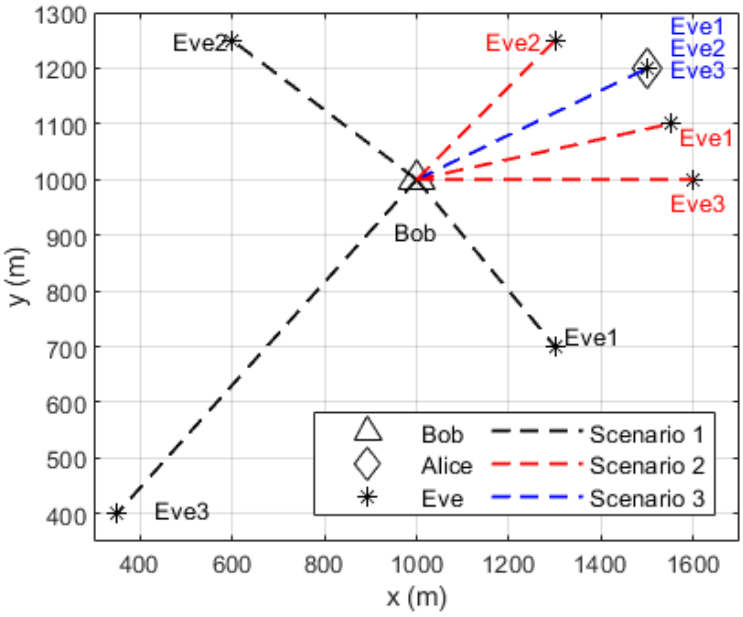

Figure 3.4: The simulation layout for Scenarios 1, 2, and 3.

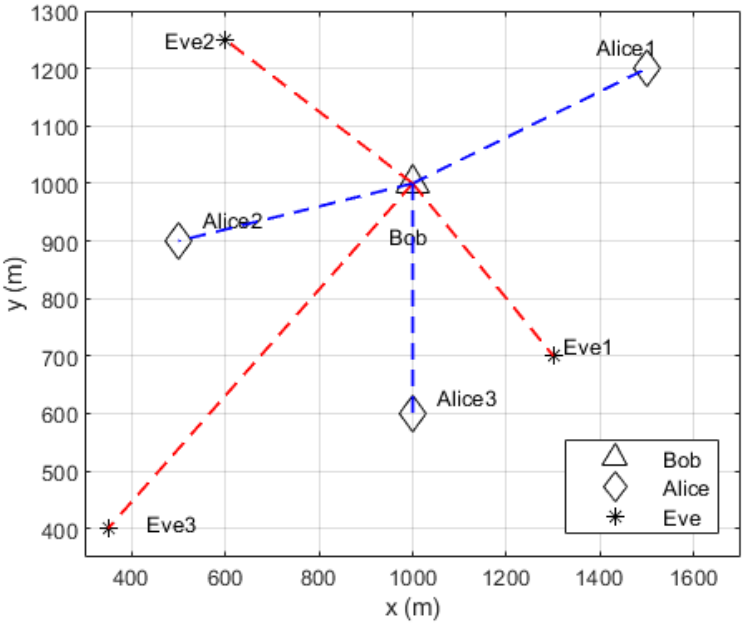

Figure 3.5: The simulation layout for Scenario 4.

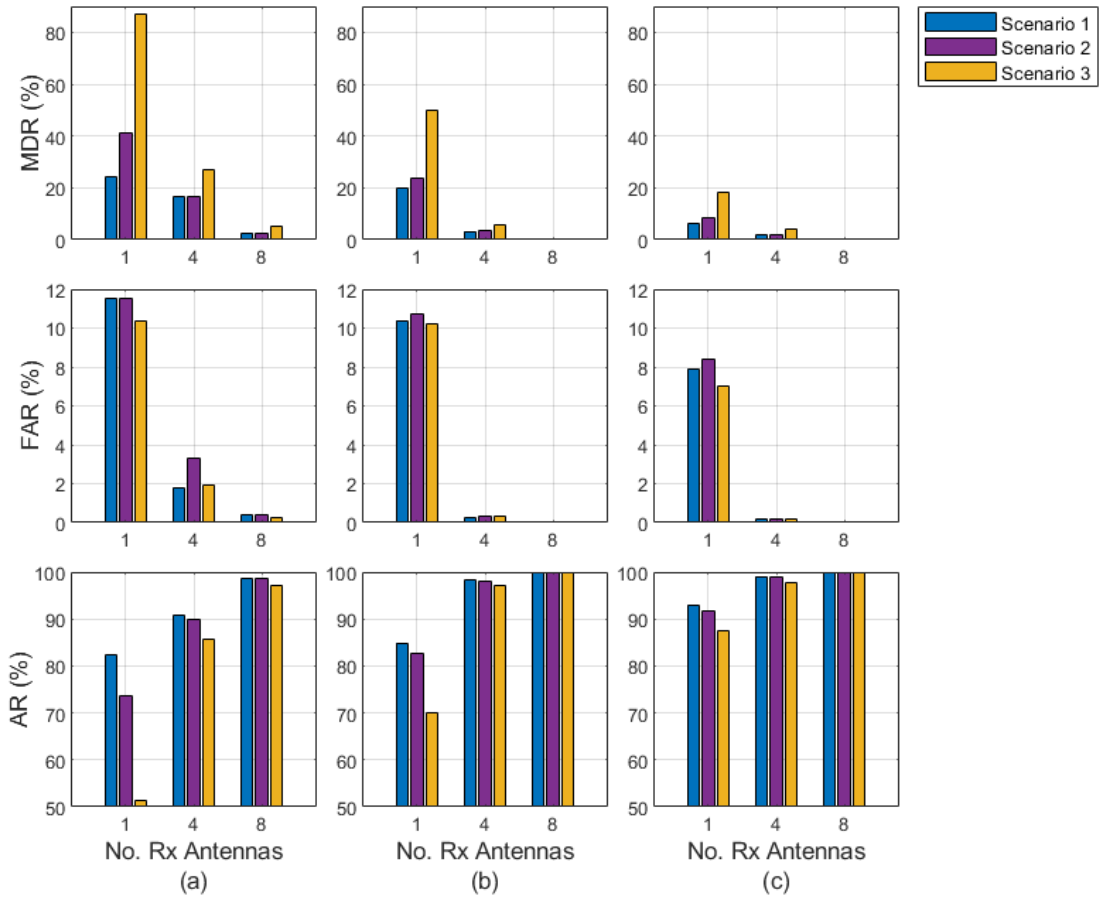

Figure 3.6: MDR, FAR, and AR versus the number of receive antennas with linear kernel OCC-SVM, SNR = 8 dB, velocity = 0.4 km/h, and 1, 4, and 8 receive antennas for Scenario 1, 2, and 3 with (a) 1 transmit antenna, (b) 2 transmit antennas, and (c) 4 transmit antennas.

3.5.1 Scenarios 1, 2, and 3

Figure 3.6 gives the MDR, FAR, and AR versus the number of receive antennas with linear kernel OCC-SVM, SNR = 8 dB, 1, 4, and 8 receive antennas, and velocity = 0.4 km/h for Scenarios 1, 2, and 3 using 1, 2 and 4 transmit antennas. These results show that the MDR, FAR, and AR improve with an increase in the number of transmit or receive antennas for all scenarios. For example, Figure 3.6a indicates that the MDR for Scenario 1 decreases from 24.0% with 1 receive antenna to 16.6%

with 4 antennas and 2.6% with 8 antennas, compared to 41.0%, 16.5%, and 2.6% for Scenario 2, respectively. Moreover, the FAR for Scenario 1 decreases from 11.5% with 1 antenna to 1.8% with 4 antennas and 0.4% with 8 antennas, compared to 11.5%, 3.0%, and 0.4% for Scenario 2, respectively. The MDR for Scenario 3 is worse than for Scenarios 1 and 2, particularly with 1 receive antenna. For example, in Figure 3.6a the MDR with 1 receive antenna is 87.0%, but improves to 5.2% with 8 receive antennas, which is slightly higher than for the other scenarios.

Figure 3.6a indicates that the AR for Scenario 1 using 1 receive antenna is greater than in Scenarios 2 and 3. However, the AR increases with the number of receive antennas. For example, the AR for Scenario 1 increases from 82.2% with 1 antenna to 90.8% with 4 antennas and 98.5% with 8 antennas, compared to 73.6%, 90.1%, and 98.5% for Scenario 2, and 51.3%, 85.6%, and 97.3% for Scenario 3, respectively. With 8 receive antennas, the AR is 97.3% for Scenario 3, which is 1.2% lower than Scenario 1. Furthermore, the AR improves with the number of transmit antennas as shown in Figures 3.6b and 3.6c. For example, the AR is greater than 99.9% with 4 transmit antennas and 8 receive antennas for all scenarios.

3.5.2 AR for Different SNRs

Figure 3.7 presents the AR versus the number of transmit antennas for Scenario 2 with velocity = 0.4 km/h, linear kernel OCC-SVM, and 1, 4, and 8 receive antennas for SNR = 8 dB and 12 dB. Figure 3.7a shows that increasing the number of transmit antennas increases the AR. For example, the AR with 4 receive antennas and 4 transmit antennas is 98.9% compared to 98.0% with 2 transmit antennas and 90.0% with 1 transmit antenna. Furthermore, the AR with 1 transmit antenna and 4 receive antennas is 90.0% compared to 91.6% with 4 transmit antennas and 1 receive antenna. Thus, there is a greater increase in the AR when the number of transmit antennas is increased compared to increasing the number of receive antennas. The AR with 4 transmit antennas and 8 receive antennas exceeds 99.9%. Figure 3.7b

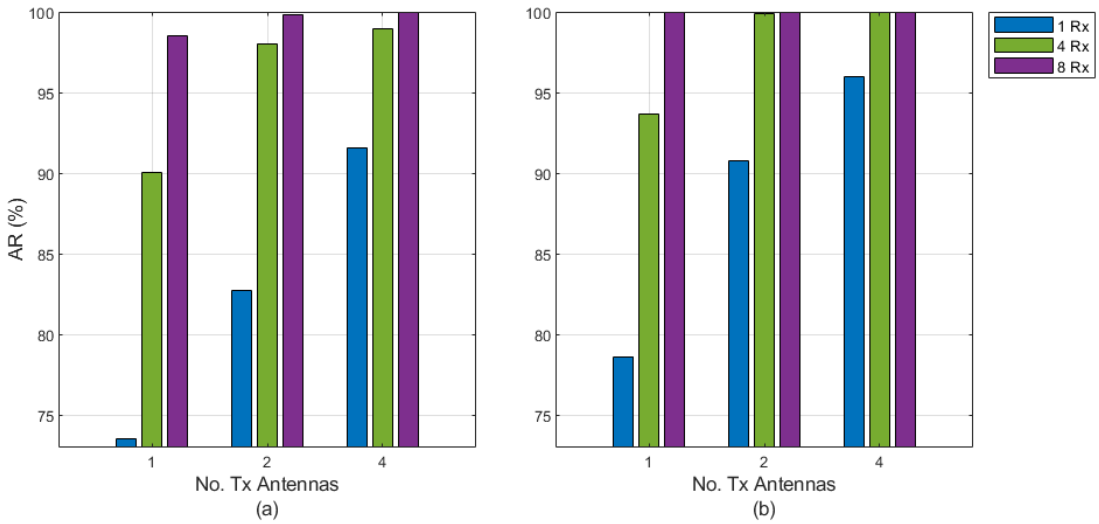

Figure 3.7: AR versus the number of transmit antennas for Scenario 2 with linear kernel OCC-SVM, 1, 4, and 8 receive antennas, velocity = 0.4 km/h, and (a) SNR = 8 dB and (b) SNR = 12 dB.

shows that with SNR = 12 dB, the AR is 99.9% with 4 transmit antennas and 4 receive antennas compared to 99.8% with 2 transmit antennas and 93.7% with 1 transmit antenna. In addition, the AR with 1 transmit antenna and 4 receive antennas is 93.7% compared to 96% with 4 transmit antennas and 1 receive antenna.

3.5.3 Majority Voting

Figure 3.8 presents the AR versus SNR with velocity = 0.4 km/h, linear kernel OCC-SVM, and 8 receive antennas for Scenario 2 using MN and N majority voting and without voting with 1, 2, and 4 transmit antennas. These results show that not voting provides a higher AR compared to majority voting with 1 and 2 transmit antennas. For example, Figure 3.8a indicates that the AR at SNR = 8 dB without voting and using MN voting is 98.5% and 97.3%, respectively, and increasing the SNR to 12 dB improves these values to 99.9% and 99.8%, respectively. Figure 3.8b shows that with 2 transmit antennas, the AR without voting and using MN and N voting exceeds 99.7% and 99.9% for SNR = 8 dB and 12 dB, respectively. In all

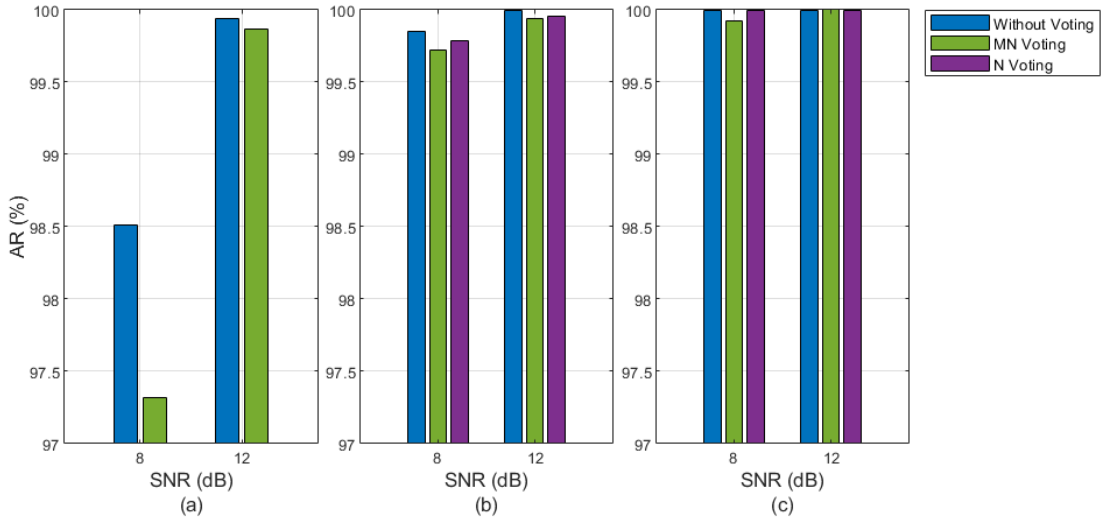

Figure 3.8: AR versus SNR with linear kernel OCC-SVM, velocity = 0.4 km/h, and 8 receive antennas for Scenario 2 using MN and N majority voting and without voting with (a) 1 transmit antenna, (b) 2 transmit antennas, and (c) 4 transmit antennas.

cases, AR without voting is the highest followed by N voting and then MN voting. However, with 4 transmit antennas and SNR = 8 dB, the AR without voting and using N voting is similar and greater than with MN voting, and with SNR = 12 dB the AR for all three schemes is almost identical. For example, Figure 3.8c shows that with 4 transmit antennas and SNR = 8 dB, the AR without voting and using N voting is almost 100% which is slightly greater than the 99.9% using MN voting. With SNR = 12 dB, the AR of the three schemes is almost 100%. These results indicate that voting is a practical option if it is not feasible to collect all the data in one location. Further, collecting the data requires that one device conduct all the training and testing tasks, which may not be fair. A distributed approach to making authentication decisions is beneficial when there are IoT devices with limited computing capabilities as each device must conduct training and testing on a small portion of the data. Then, the decisions are collected and only a simple majority vote is required. In summary, the topology and device capabilities must be considered when determining an appropriate authentication scheme.

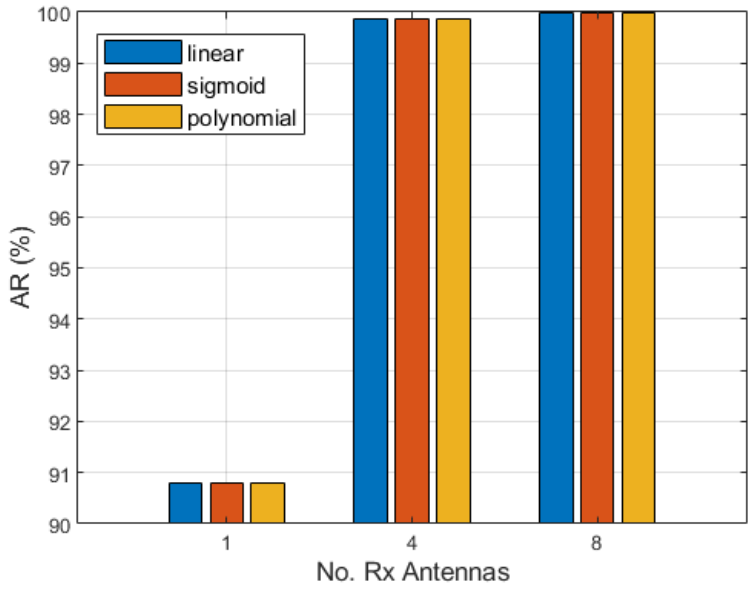

Figure 3.9: AR versus the number of receive antennas for Scenario 2 with SNR = 12 dB, velocity = 0.4 km/h, 2 transmit antennas, and 1, 4, and 8 receive antennas for OCC-SVM using linear, sigmoid, and polynomial kernels.

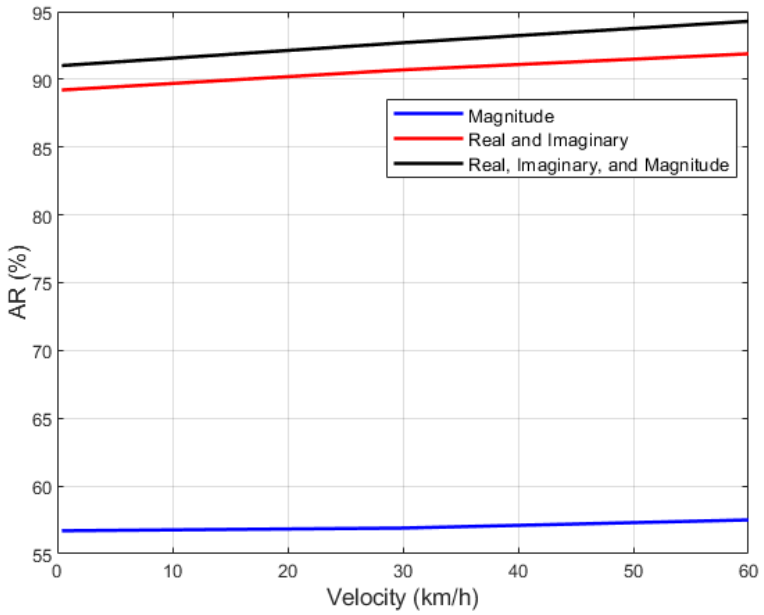

Figure 3.10: AR versus velocity for Scenario 1 with linear kernel OCC-SVM, SNR = 8 dB, 1 transmit antenna and 4 receive antennas using the magnitude, real and imaginary parts, and the magnitude and real and imaginary parts of the received signals as features.

3.5.4 Kernel Comparison

Figure 3.9 presents the AR versus the number of receive antennas for Scenario 2 with velocity = 0.4 km/h, SNR = 12 dB, 2 transmit antennas, and 1, 4, and 8 receive antennas for OCC-SVM using linear, sigmoid, and polynomial kernels. The AR using 1, 4, and 8 receive antennas is approximately 90.8%, 99.8%, and 99.9% for all three kernels, respectively. There is a very small difference in all cases with the polynomial kernel providing the highest AR followed by linear and then sigmoid. This indicates that using a less complex kernel is the best choice.

3.5.5 Number of Features

Figure 3.10 presents the AR versus velocity for Scenario 1 with linear kernel OCC-SVM, SNR = 8 dB, 1 transmit antenna and 4 receive antennas using the magnitude, the real and imaginary parts, and the magnitude and real and imaginary parts of the received signals as features. This shows that the AR is improved with more features for all velocities. For example, the AR at velocity = 0.4 km/h with 4 receive antennas using the magnitude and real and imaginary parts of the received signals as features is 91.0%, while the AR using the real and imaginary parts is 89.2%, and using just the magnitude is only 56.7%. Although the magnitude is not independent of the real and imaginary parts of the received signal, using it as an additional feature provides a small improvement in performance. Further, these results show that the AR improves with velocity. For example, the AR using the magnitude and real and imaginary parts of the received signals at velocity = 0.4 km/h, 30 km/h, and 60 km/h is 91.0%, 92.7%, and 94.3%, respectively. This is because a higher velocity increases the Doppler frequency shift which provides better discrimination between users.

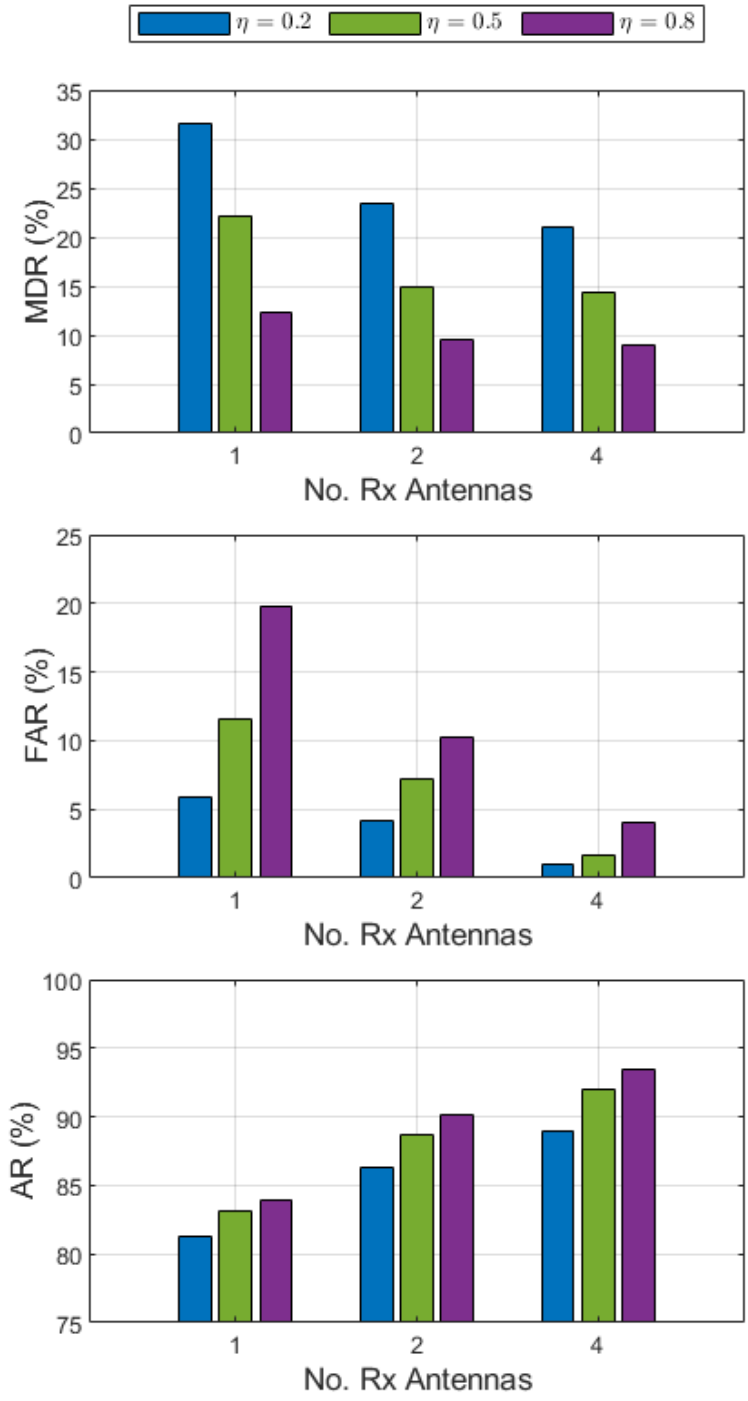

Figure 3.11: MDR, FAR, and AR versus the number of receive antenna with linear kernel OCC-SVM, SNR = 8 dB, velocity = 0.4 km/h, and 1 transmit antenna for Scenario 1 with $\eta = 0.2, 0.5$, and 0.8.

3.5.6 Effect of Outlier Percentage

Figure 3.11 presents the MDR, FAR, and AR versus the number of receive antennas with linear kernel OCC-SVM, SNR = 8 dB, velocity = 0.4 km/h, and 1 transmit antenna for Scenario 1 with $\eta = 0.2, 0.5$, and 0.8. This shows that increasing η, which is the percentage of data considered as outliers, decreases the MDR which lowers the probability of accepting Eves. For example, the MDR with 2 receive antennas is 23.3%, 14.9%, and 9.6% for $\eta = 0.2, 0.5$, and 0.8, respectively. Moreover, increasing η results in a greater FAR which increases the probability of rejecting Alice. For example, the FAR with 1 receive antenna is 5.9%, 11.6%, and 19.8% for $\eta = 0.2, 0.5$, and 0.8, respectively. Finally, these results indicate that the AR increases with η. For example, the AR with 4 receive antennas is 88.9%, 92.0%, and 93.5% for $\eta = 0.2, 0.5$, and 0.8, respectively.

3.5.7 Multiuser Performance

The proposed scheme is now evaluated for Scenario 4 which has three Alice. This is the same as Scenario 1 except Alice 2 and Alice 3 are added. Thus, there are three legitimate users and three illegitimate users, all with velocity = 0.4 km/h. Figure 3.12 presents the MDR, FAR, and AR versus the number of receive antennas with linear kernel OCC-SVM, SNR = 8 dB, and 1 transmit antenna for Scenarios 1 and 4. This shows that the authentication performance is decreased with more legitimate users. For example, the MDR is 16.6% and 39.1% with 4 receive antennas for Scenarios 1 and 4, respectively, and the corresponding FAR is 1.8% and 2.7% with 4 receive antennas. Finally, the AR is 90.8% and 79.1% with 4 receive antennas for Scenarios 1 and 4, respectively, so the AR decreases with the number of users.

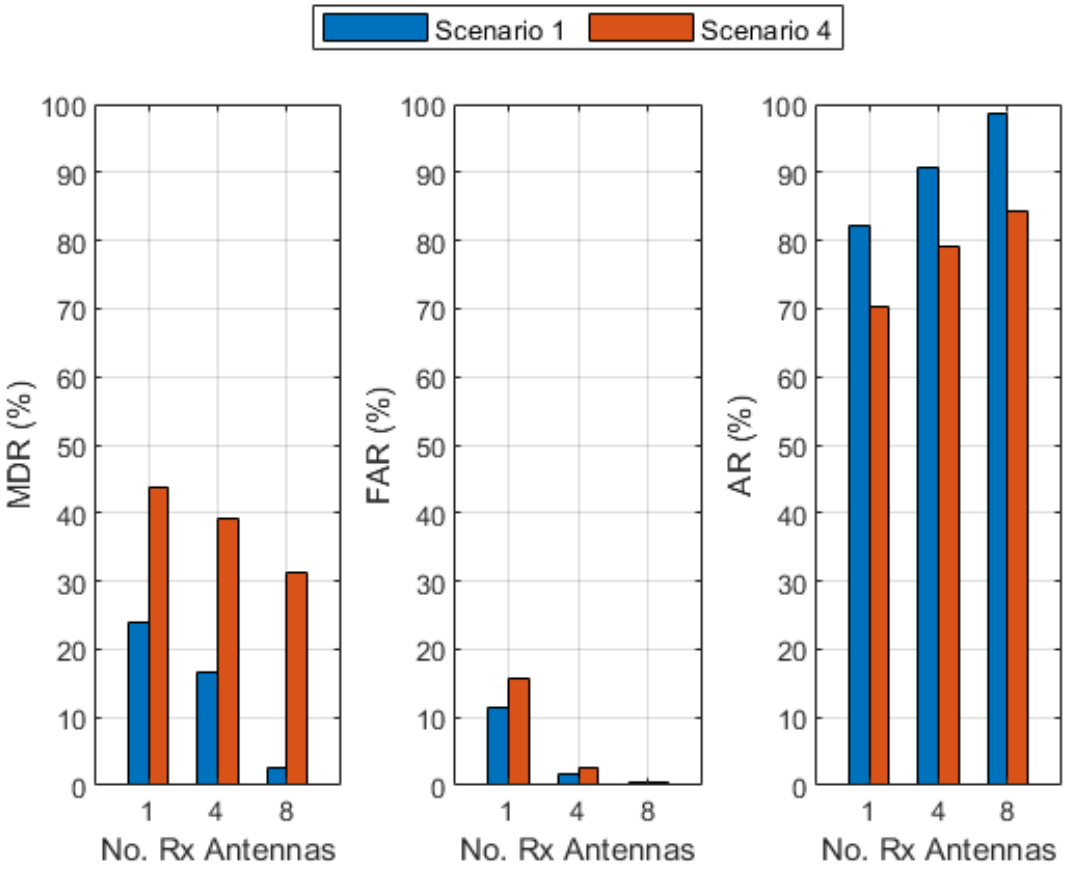

Figure 3.12: MDR, FAR, and AR versus the number of receive antenna with linear kernel OCC-SVM, SNR = 8 dB, velocity = 0.4 km/h, and 1 transmit antenna for Scenarios 1 and 4.

3.6 Conclusion

An adaptive physical layer authentication (PLA) scheme using machine learning (ML) was proposed for Internet of Things (IoT) applications. This scheme exploits the antenna diversity in multi-input-multi-output (MIMO) systems based on the 5G frame structure. One-class classification support vector machine (OCC-SVM) with linear, sigmoid, and polynomial kernels is employed. The proposed scheme is adaptive as a sliding window is used to update features to provide robust authentication. The magnitude and real and imaginary parts of the received signals corresponding to the 5G sounding reference signals (SRSs) are used as features. Thus, the number of features is $3MN$ for each transmitted symbol where M and N are the number of transmit and receive antennas, respectively. Results were presented for an urban environment which show that the performance improves with the number of antennas. The AR using two majority vote schemes was presented. These schemes may be preferable for distributed IoT devices with limited computing capabilities. The AR using all $3MN$ features at one location was shown to be slightly higher than using $3M$ features separately at each device and N voting. In future work, the proposed scheme can be evaluated for indoor and rural environments. In addition, other machine learning algorithms can be employed with the proposed scheme.

Chapter 4

Physical Layer Authentication for Satellite Communication Systems using Machine Learning

In this chapter, an adaptive physical layer authentication scheme using Doppler frequency shift (DS) and received power (RP) features for LEO satellites is proposed [77, 36]. The proposed scheme employs OCC-SVM which uses only legitimate training data. Also, the scheme is evaluated for fixed satellite services (FSS) and mobile satellite services (MSS) at different altitudes. The proposed scheme is evaluated using linear and polynomial OCC-SVM kernels. Results are presented which show that the proposed scheme provides a higher authentication performance using DS and RP features together compared to using them separately.

4.1 Related Work

A comprehensive survey on the security challenges for satellite communication systems was presented in [78]. Many satellite systems currently send unauthenticated messages or messages that have been authenticated at the application layer using

symmetric key (implicit authentication) or public key solutions. Thus, simple and effective solutions to detect spoofing attacks are required. Several anti-spoofing schemes have been developed, e.g. global navigation satellite system (GNSS) spoofing detection [79, 80], received signal correlation using multiple antennas at the receiver [81], examining physical information such as received power, carrier-to-noise ratio (CNR), and angle of arrival [82], leveraging the ad-hoc network infrastructure [83, 84], and dedicated hardware [85].

It was shown in [3] that satellite communications is vulnerable to spoofing attacks, especially the downlink satellite system information signalling. Thus, a PLA scheme was proposed to validate satellites using the DS. It is used prior to initial access to the land mobile satellite system so an attacker cannot imitate the real-time DS of a user. The DS can be estimated either through signal observations or user calculations from satellite broadcast ephemeris. In terrestrial networks, physical layer attributes such as the channel state information (CSI) can be used for PLA [1, 9, 60, 62, 86]. However, CSI-based schemes may not be suitable for satellite authentication because of the strong line-of-sight (LOS) channel which does not provide sufficiently unique features.

In [87], Iridium LEO satellite signatures were obtained using signal IQ values. The signals from these satellites exhibit unique attenuation and fading characteristics due to the high mobility of up to 25000 km/h. The proposed scheme employs a convolutional neural network (CNN) for authentication, and pattern recognition techniques are used to generate synthetic images from the IQ values. An orbit-based authentication scheme for satellite communications was proposed in [88]. Satellites orbiting the Earth on a fixed trajectory provide a priori information for security purposes and time difference of arrival (TDOA) measurements from multiple receivers are used for authentication.

In [89, 90], the DS of spacecraft links was used to generate symmetric keys. A PLA scheme using DS was proposed in [20] for LEO satellites. Since velocity

and location information for all satellites is available, reference DS values for any satellite can easily be calculated. Thus, each satellite in a constellation can compare the measured DS value with the reference value for the satellite in the constellation to decide whether it is legitimate. Then, a majority vote is taken at a fusion center to make the final authentication decision.

4.2 Contributions

The main contributions of this chapter are as follows.

- An adaptive physical layer authentication scheme using OCC-SVM is proposed to authenticate LEO satellites utilizing DS and RP are used as features.
- The proposed scheme is validated for FSS and MSS when the illegitimate satellites are within the half power beamwidth (HPBW).
- The proposed scheme is evaluated using MDR, FAR, and AR metrics utilizing DS and RP features separately and together.
- Results are presented using two-line element (TLE) data for real satellites to verify the effectiveness of the proposed schemes, where TLE is orbital data for Earth-orbiting objects [35].

4.3 System Model

The system model for the LEO satellite PLA scheme is illustrated in Figure 4.1. The FSS or MSS station (FMS) must authenticate the legitimate satellite (Alice) over the entire session while preventing spoofing attacks from illegitimate satellites (Eves). The communication session starts from the lowest elevation angle where the DS is maximum and the RP is minimum. A LEO satellite communication session is

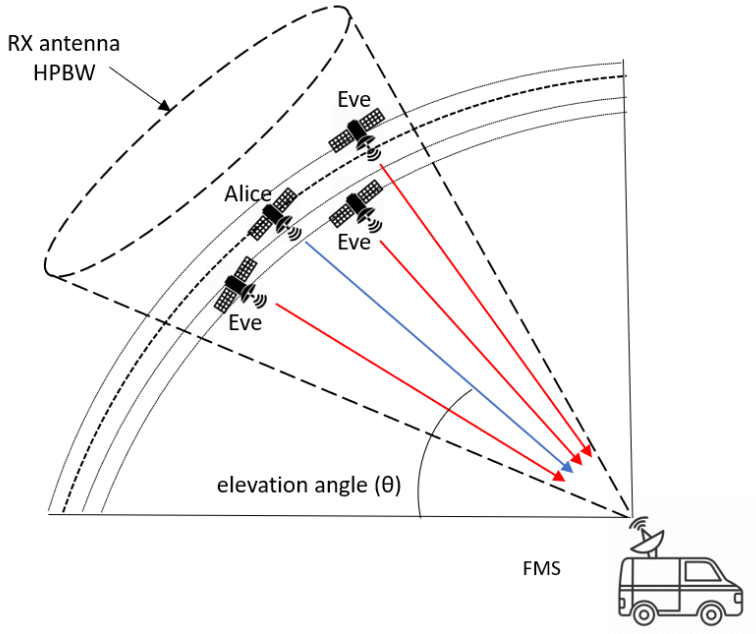

Figure 4.1: System model.

the time over which the satellite is continuously serving a given ground service [91]. A session is assumed to have $2n-1$ phases (time instances). Figure 4.1 shows the first n phases. Eves try to imitate Alice in order to send incorrect information to users. We assume an attack scenario such that Eve satellites are very close to Alice, where Eves are at the same altitude as Alice and within the FMS receive antenna HPBW. The FMS authenticates Alice using PLA over the communication session after upper-layer authentication (ULA).

The proposed scheme is employed on the downlink and is divided into an initial phase and subsequent phases. In the initial phase, ULA authentication is performed and OCC-SVM training is conducted using the DS and RP features of Alice. In subsequent phases, OCC-SVM testing is conducted to establish the legitimacy of the received signals which could be from Alice or Eve. Therefore, the features in these phases are considered to be from an unknown satellite (U), where U = {Alice, Eve}. If the test is passed, the FMS decides the signal is from Alice so the training

features are updated and OCC-SVM training is repeated. Conversely, if the test is failed, the FMS decides the signal is from Eve and the connection is terminated.

4.3.1 Doppler Frequency Shift

The received signal at the FMS will have a DS given by [92]

$$f_d = \frac{v \times f_c}{c} \times \cos(\phi), \tag{4.1}$$

where v is the velocity of the satellite, c is the speed of light, f_c is the center frequency, and ϕ is the angle between the satellite to FMS link and the direction of motion of the satellite. Consequently, for the same v and f_c, at a given time ϕ will differ between satellites so the DS is unique to a satellite.

4.3.2 Received Power

The power received at the FMS in watts is given by [93]

$$p_r = \frac{p_t \; g_t \; g_r}{(4\pi d/\lambda)^2}, \tag{4.2}$$

where p_t is the transmitted power, g_t is the gain of the transmit antenna, g_r is the gain of the receive antenna, d is the distance between transmitter and receiver, and λ is the wavelength. The term $(4\pi d/\lambda)^2$ is known as the free space path loss (FSPL). The gain of the receiver antenna in the direction θ is defined as $g_r(\theta)$ [94]. The angle θ is usually in the direction of the maximum gain, called the boresight direction of the antenna. It is used for FMS antenna tracking the trajectory of Alice. Eve will not have the same trajectory and location as Alice at a given time. Thus, the RP from Eve at the FMS will differ from the corresponding RP from Alice due to the difference in θ. Further, the FSPL for different satellites will differ due to the distance from the satellite to the FMS. It is assumed that the Eves have the same

values of p_t and g_t as Alice which can be considered worst case.

4.3.3 Problem Formulation

In the initial phase, the DS and RP of Alice are obtained by the ULA at the start of the session. Then, PLA scheme is employed utilizing the DS and the RP obtained at subsequent phases. Over this session, the signals of Alice and the Eves are directed towards the FMS. The Eves do this to imitate Alice. However, the FMS is following the trajectory of Alice over the communication session which is known. The FMS must decide between the two hypotheses

$$\begin{cases} \mathcal{H}_0: & \text{Alice transmitting,} \\ \mathcal{H}_1: & \text{Eve transmitting,} \end{cases} \tag{4.3}$$

where the null hypothesis ($\mathcal{H}_0$) denotes that the signal is from Alice while the alternative hypothesis ($\mathcal{H}_1$) means that it is from Eve. OCC-SVM training in the initial phase determines the authentication boundary for the features from Alice. In subsequent phases, OCC-SVM testing is conducted to determine if the corresponding features are located within this boundary. The satellite is accepted as legitimate if this test is true. On the other hand, if the test decision is outside the boundary, the satellite is rejected. Furthermore, OCC-SVM authentication boundary is updated in each subsequent phase to provide robust authentication.

4.4 The Proposed Authentication Scheme

In a real system, Alice will have a deviation from the reference trajectory [95] which will affect the signal received at the FMS station [96]. It is impossible for Eve to determine this deviation and this will cause errors if Eve tries to manipulate her RP and DS values to imitate Alice. The proposed authentication scheme employs

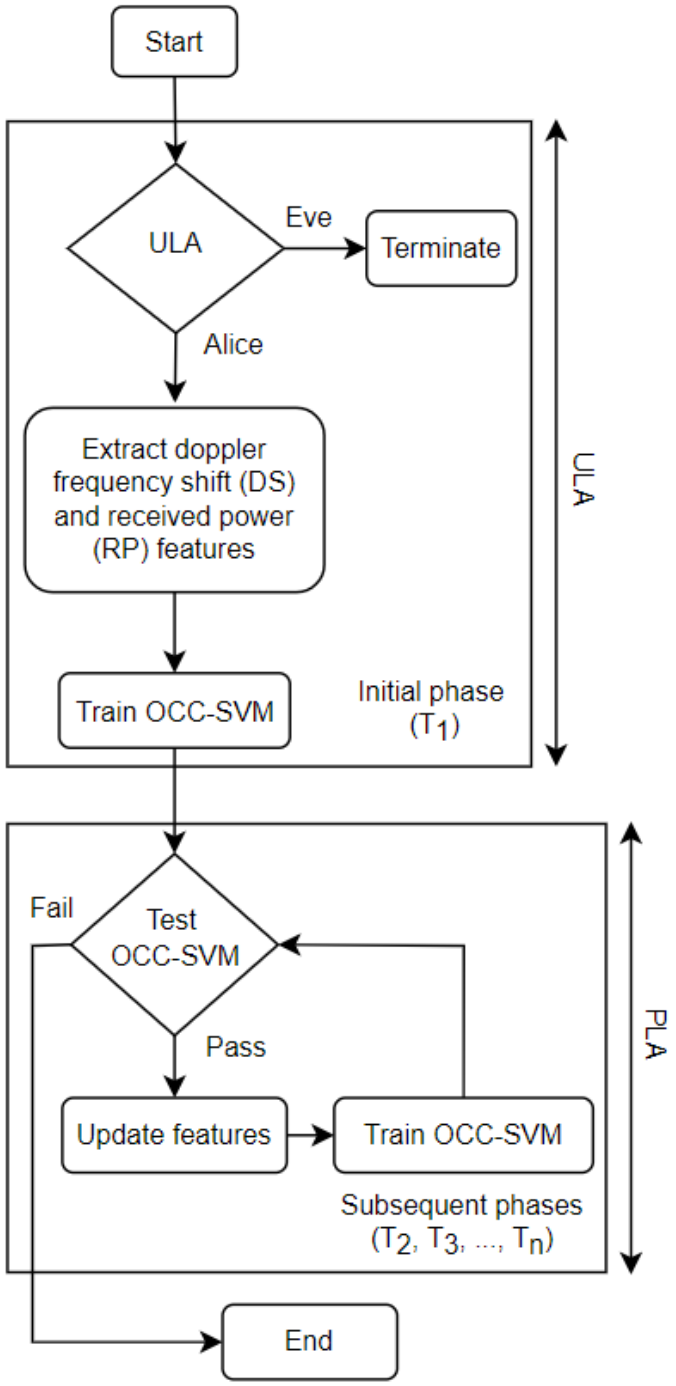

Figure 4.2: Proposed authentication scheme flowchart.

OCC-SVM using the DS and RP as features for training and testing. Figure 4.2 presents the proposed authentication scheme. In the initial phase T_1, after ULA authentication, data is collected from the legitimate satellite for OCC-SVM training. Then, in subsequent phases data from U is used by FMS for testing and training. If the test is passed in a given phase, the corresponding data is used to update the features for training. A sliding window is used for this update so the oldest data is discarded. On the other hand, if the test fails, the connection is terminated.

The data vectors have the form

$$\mathbf{m} = [s\ p] \tag{4.4}$$

where s and p are the DS and RP, respectively. In phase T_1, authentication is first performed with Alice through ULA. Then, ℓ data vectors from Alice

$$\mathbf{d}_i = [s_i\ p_i], i = 1, 2, \ldots, \ell, \tag{4.5}$$

are scaled and used for OCC-SVM training to determine the boundary for authen-

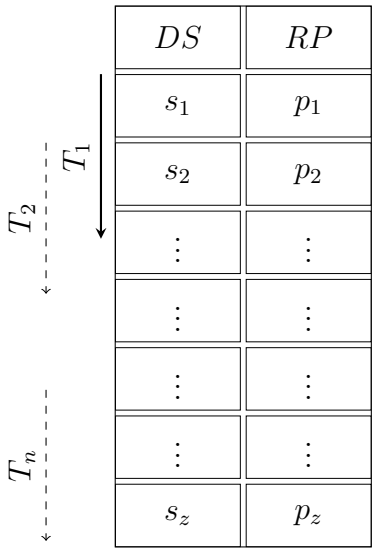

Figure 4.3: Sliding window for feature updates.

Algorithm 5 The proposed scheme

Authenticate Alice through ULA.
Extract s and p.
Perform Min-Max scaling to obtain $\mathbf{g}_i$ (4.10).
Form the training matrix $\mathbf{G}_e$ in (4.11).
Train OCC-SVM.
Test OCC-SVM with $\mathbf{t}$ (4.12).
while $(f(\mathbf{t}) > 0)$ **do**
    Update the feature matrix $\mathbf{G}_e$.
    Retrain OCC-SVM.
    Test OCC-SVM with $\mathbf{t}$ (4.12).
end while

tication. In subsequent phases, OCC-SVM is used to test (after scaling), new data vectors

$$\mathbf{b} = [s\ p]_{\mathrm{U}}, \tag{4.6}$$

from U using (A.9). If the test is passed the satellite is accepted, the features are updated, and OCC-SVM is retrained. However, if the test fails, the connection is terminated.

Figure 4.3 shows the sliding window update process for the data where the rows are the data vectors. In phase T_1, the training data from Alice is a matrix with dimensions $\ell \times 2$. Then, in phase T_2 a new data vector $\mathbf{b}$ is tested (after scaling), and if accepted the data matrix is updated by discarding the first row $\mathbf{d}_1$ and adding the new data vector as row $\ell + 1$. Thus, if the first e new data vectors are accepted, the training data matrix is

$$\mathbf{M}_e = \begin{bmatrix} \mathbf{d}_{1+e} \\ \mathbf{d}_{2+e} \\ \vdots \\ \mathbf{d}_{\ell+e} \end{bmatrix}, \tag{4.7}$$

as shown in Figure 4.3 so ℓ vectors are used for training in each phase.

Min-Max scaling is separately applied to each feature. At the lowest elevation

angle θ, DS is maximum s_{max} and RP is minimum p_{min}, and at the highest θ DS is minimum s_{min} and RP is maximum p_{max}. These values are based on satellite orbit and trajectory and are available using the system tool kit (STK) [35, 97]. The data vectors $\mathbf{d}_i$ are scaled as follows

$$s'_i = \frac{s_i - s_{min}}{s_{max} - s_{min}}, \tag{4.8}$$

$$p'_j = \frac{p_j - p_{min}}{p_{max} - p_{min}}, \tag{4.9}$$

so the corresponding feature vectors are

$$\mathbf{g}_i = [s'_i \; p'_i]. \tag{4.10}$$

The matrix of training vectors is then

$$\mathbf{G}_e = \begin{bmatrix} \mathbf{g}_{1+e} \\ \mathbf{g}_{2+e} \\ \vdots \\ \mathbf{g}_{\ell+e} \end{bmatrix} \tag{4.11}$$

The new data vectors $\mathbf{b}$ are scaled to obtain the testing vectors

$$\mathbf{t} = [s' \; p']_U. \tag{4.12}$$

The proposed scheme is summarized in Algorithm 5.

4.5 Performance Evaluation

In this section, the proposed scheme is evaluated using the STK to obtain the DS and RP values and the scikit-learn library in Python is used for linear and polynomial

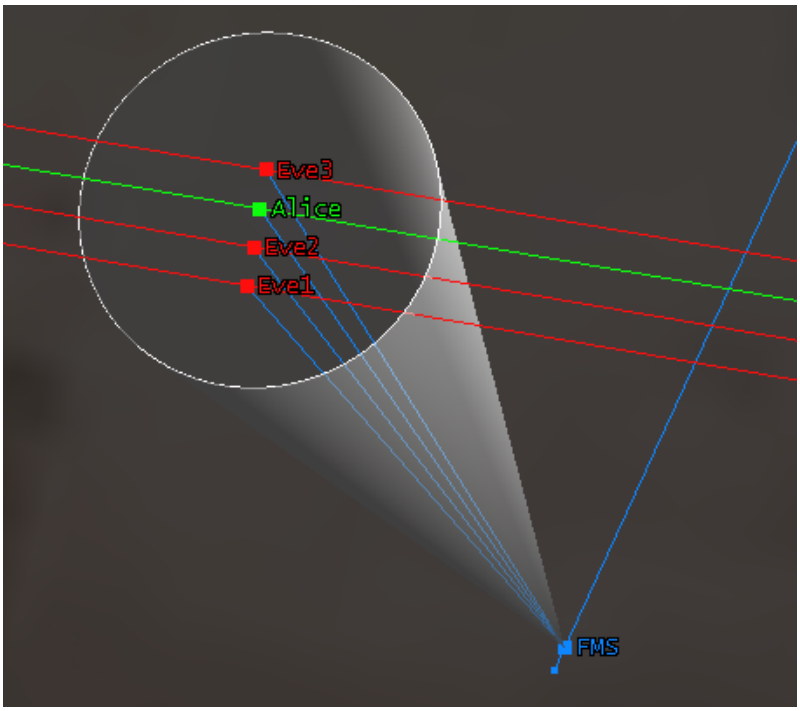

Figure 4.4: Model layout.

Table 4.1: Maximum Distances Between Alice and Eves over the Session at Different Altitudes

Altitude	Alice - Eve1	Alice - Eve2	Alice - Eve3
500 km	10 km	6 km	6 km
1000 km	12 km	7 km	7 km
1500 km	13 km	8 km	8 km
2000 km	14 km	8.5 km	8.5 km

OCC-SVM. STK is used to obtain the DS and RP values along the trajectory of Alice and Eves using (4.1) and (4.2). The model layout is shown in Figure 4.4, where two scenarios are considered. The first is on-pause communications which refers to FSS where a fixed ground station authenticates Alice using a receive antenna 1.5 m in diameter. The second is on-move communications which refers to MSS where a mobile vehicle authenticates Alice using a receive antenna 0.5 m in diameter. A worst-case situation is considered where all Eves satellites are very close to Alice

Table 4.2: Simulation Parameters

Parameter	**Value**
Center frequency	7.5 GHz
FSS antenna diameter	1.5 m
MSS antenna diameter	0.5 m
Satellite altitudes	500, 1000, 1500, and 2000 km
Tx power for all satellites	10 dBW
ℓ	10

with the same altitude and within the FMS receive antenna HPBW as shown in Figure 4.4. The maximum distances between Alice and Eves over the session in both scenarios are given in Table 4.1. The simulation parameters are given in Table 4.2 and $\gamma = \frac{1}{3}$ as we have 3 Eves. The number of symbols used for training is $\ell = 10$. Figure 4.5 presents the scaled DS and the scaled RP for Alice at an altitude of 2000 km obtained from STK. This shows that the scaled DS and the scaled RP values in the first half of the communication session are similar to those in the second half.

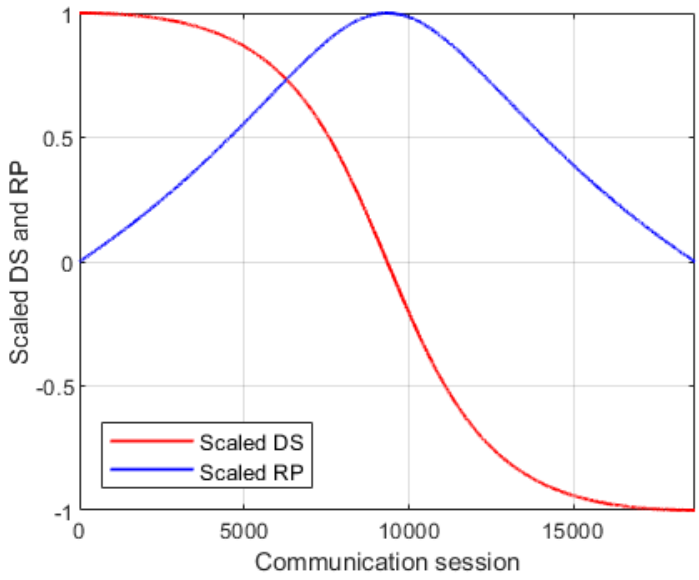

Figure 4.5: Scaled DS and scaled RP over the communication session.

Thus, only DS and RP values for the first half of the communication session are considered in the simulations.

4.5.1 Fixed Satellite Services (FSS)

Figure 4.6 presents the MDR, FAR, and AR versus altitude for FSS and OCC-SVM with linear and polynomial kernels. DS and RP denote that only those features are used while DS & RP means both features are employed. The FAR of DS and DS & RP for linear and polynomial OCC-SVM kernels is 0 and so is not shown in the figure. Figure 4.6a gives the MDR, FAR, and AR for the linear kernel. The MDR

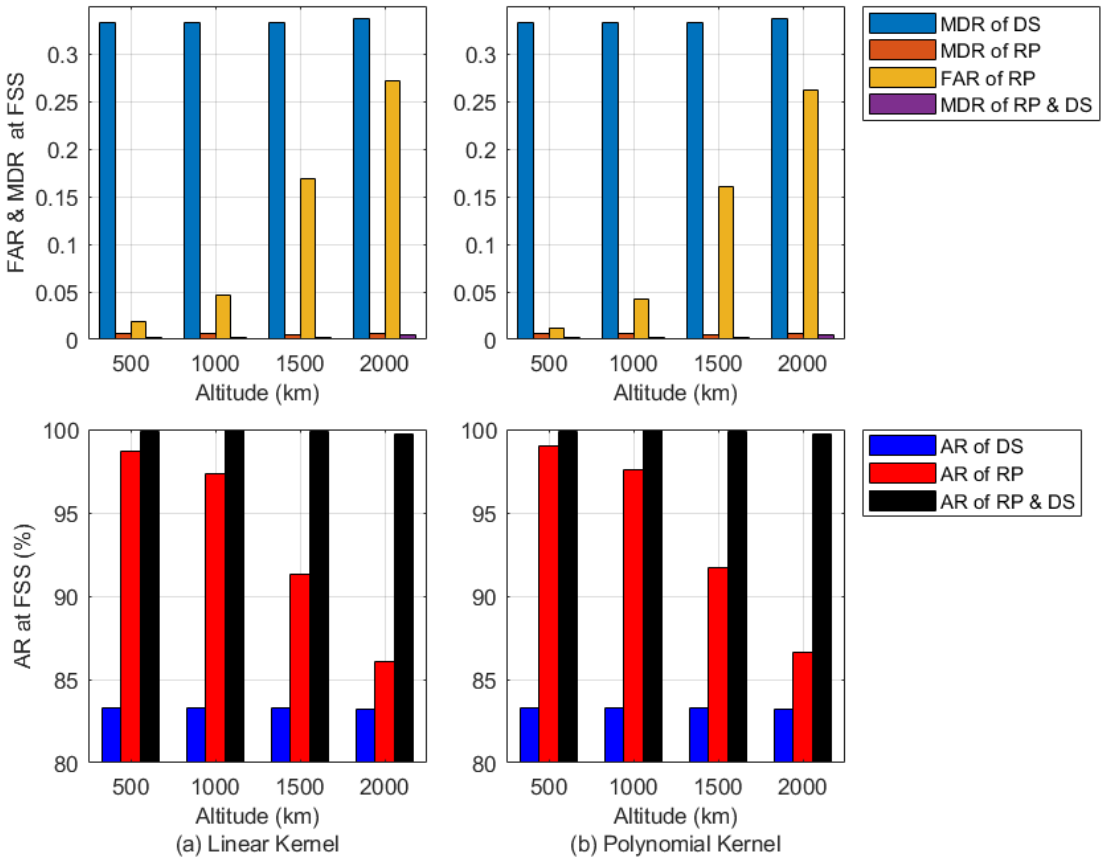

Figure 4.6: MDR, FAR, and AR versus altitude for FSS with (a) linear OCC-SVM kernel, and (b) polynomial OCC-SVM kernel.

of DS is around 33.3% at all the altitudes and the AR is approximately 83.3% at all the altitudes. The MDR of RP is around 0.6% at all altitudes while the FAR of RP is 1.9% at 500 km altitude and increases with altitude to 27.1% at 2000 km. Thus, the AR of RP decreases with increasing altitude from 98.7% at 500 km to 86.1% at 2000 km. The MDR with both DS and RS is between 0.2% and 0.5% for all altitudes, and the corresponding AR is between 99.7% and 99.9%. Figure 4.6b presents the MDR, FAR, and AR for the polynomial kernel. This shows that the variations are smaller than with the linear kernel with an AR between 99.7% and 99.9%. However, the performance with the linear and polynomial kernels is similar for all altitudes. Further, using both DS and RP in the proposed scheme provides an AR which exceeds 99.7% for FSS.

4.5.2 Mobile Satellite Services (MSS)

Figure 4.7 gives the MDR, FAR, and AR versus altitude for MSS and OCC-SVM with linear and polynomial kernels. The receive antenna diameter is 0.5 m so the HPBW is wider than in the FSS case. The FMS is moving in the direction shown in Figure 4.4. The FAR of DS and DS & RP for linear and polynomial kernels is 0 as in the FSS case, and so is not shown in the figure. Figure 4.7a gives the MDR, FAR, and AR with the linear kernel. The MDR with DS is slightly higher than in the FSS case. The AR is about 83% for all altitudes. Also, it is slightly different than FSS for MDR of RP, which becomes around 0.45% over all altitudes. Consequently, the FAR of RP is also increased with increasing altitude. It is 1.4% at 500 km altitude and increased to 27.1% at 2000 km altitude. The AR of RP decreases with increasing altitude, from 99% to 86.1% at 500 and 2000 km altitude, respectively. Finally, the MDR when using DS and RS features, slightly higher than in the FSS case, around 0.6% for all altitudes so the AR exceed 99.6%. These results show that the performance with the linear and polynomial kernels in Figures 4.7a and 4.7b, respectively, is similar for all altitudes. Moreover, using both DS & RP

in the proposed scheme provides an AR which is greater than 99.6%. In summary, the proposed scheme achieves an AR greater than 99.6% for FSS and MSS at all altitudes with OCC-SVM using both DS and RP, and linear or polynomial kernels.

4.6 Conclusion

The increase in the number of LEO satellite constellations makes VHetNets a solution to provide worldwide wireless coverage. However, LEO satellites are vulnerable to spoofing attacks so an efficient and effective authentication scheme is needed. An adaptive PLA scheme using ML was proposed to solve this problem using the DS

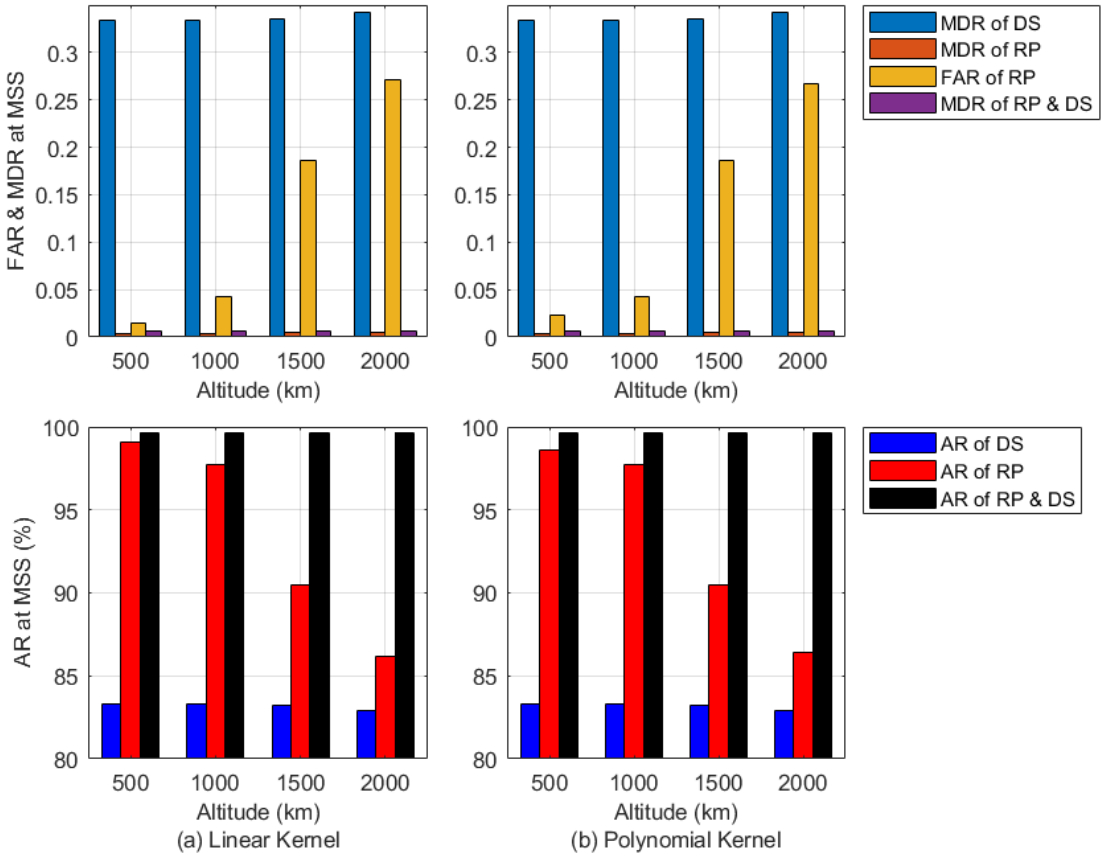

Figure 4.7: MDR, FAR, and AR versus altitude for MSS with (a) linear OCC-SVM kernel, and (b) polynomial OCC-SVM kernel.

and RP as features. A OCC-SVM with linear and polynomial kernels was employed. Results were presented which show that a high AR can be achieved for both fixed and mobile satellite services using the DS and RP as features. In particular, the AR in this case exceeds 99.6% for FSS and MSS scenarios. Finally, the proposed scheme employing both the DS and RP as features was shown to be superior to using only the DS or RP as features as in [20] and [98], respectively.

Chapter 5

Authentication for Satellite Communication Systems using Physical Characteristics

In this chapter, PLA is proposed for low earth orbit (LEO) satellites using the Doppler frequency shift (DS) and received power (RP) characteristics [99, 100]. Hypothesis testing using a threshold or machine learning (ML) is considered to discriminate between legitimate and illegitimate satellites. For ML, a one-class classification support vector machine (OCC-SVM) is employed which uses training data from only legitimate users. The performance is evaluated using real satellite data from the system tool kit (STK). Results are presented which show that the authentication rate (AR) with DS is higher than with RP at low elevation angles for both schemes, but is higher with RP at high elevation angles. Further, the ML authentication scheme provides a higher AR than the threshold scheme for a small percentage of the training data considered as outliers, but at larger percentages the OR threshold scheme is better.

5.1 Related Work

In [78], the security challenges for satellite communication systems were considered. Several anti-spoofing schemes have been developed, e.g. global navigation satellite system (GNSS) spoofing detection [79, 101], received signal correlation using multiple antennas at the receiver [81], examining physical information such as received power, carrier-to-noise ratio (CNR), and angle of arrival [82, 102], and leveraging ad-hoc network infrastructure [83, 84] and dedicated hardware [103, 85]. It was shown in [3] that satellite communications is vulnerable to spoofing attacks. Thus, a PLA scheme was proposed to validate satellites using the Doppler frequency shift (DS). It is used prior to initial access to the land mobile satellite (LMS) system so an attacker cannot imitate the real-time DS of a user. The DS can be estimated either through signal observations or user calculations from satellite broadcast ephemeris.

In terrestrial networks, physical layer attributes such as the channel state information (CSI) can be used for PLA [86, 1, 37, 60, 49, 104, 62, 40, 50, 105, 9]. However, CSI-based schemes may not be suitable for satellite authentication because of the line-of-sight (LOS) channel which does not provide sufficiently unique features. In [68], a PLA framework was proposed for controller area networks (CANs) to mitigate spoofing attacks. This scheme utilizes the arrival intervals and magnitudes of the received signals as features. Moreover, reinforcement learning (RL) is employed for authentication using the Dyna architecture.

In [87], Iridium LEO satellite signatures were obtained using in-phase and quadrature (IQ) signal values. The signals from these satellites exhibit unique attenuation and fading characteristics due to the high mobility of up to 25000 km/h. The proposed scheme employs a convolutional neural network (CNN) for authentication, and pattern recognition techniques are used to generate synthetic images from the IQ values. In [89, 90], the DS of spacecraft links was used to generate symmetric keys. An orbit-based authentication scheme for satellite communications was pro-

posed in [88]. Satellites orbiting the Earth on a fixed trajectory provide a priori information for security purposes and time difference of arrival (TDOA) measurements from multiple receivers are used for authentication. A PLA scheme using DS was proposed in [20] for LEO satellites. Since velocity and location information for all satellites is available, reference DS values for any satellite can easily be calculated. Thus, each satellite in a constellation can compare the measured DS value with the reference value for the satellite in the constellation to decide whether it is legitimate. Then, a majority vote is taken at a fusion center to make the final authentication decision.

5.2 Contributions

The main contributions of this chapter are as follows.

- An adaptive PLA scheme using DS and RP characteristics to authenticate LEO satellites is proposed.
- Hypothesis testing using a threshold or ML is used to discriminate between legitimate and illegitimate satellites.
- The AR is evaluated using DS and RP characteristics separately and together over the communication session and DS and RP estimation errors are considered.
- Results are presented using two-line element (TLE) data for real satellites to verify the effectiveness of the proposed schemes. TLE data is orbital data for Earth-orbiting objects and is available at: https://celestrak.com/.

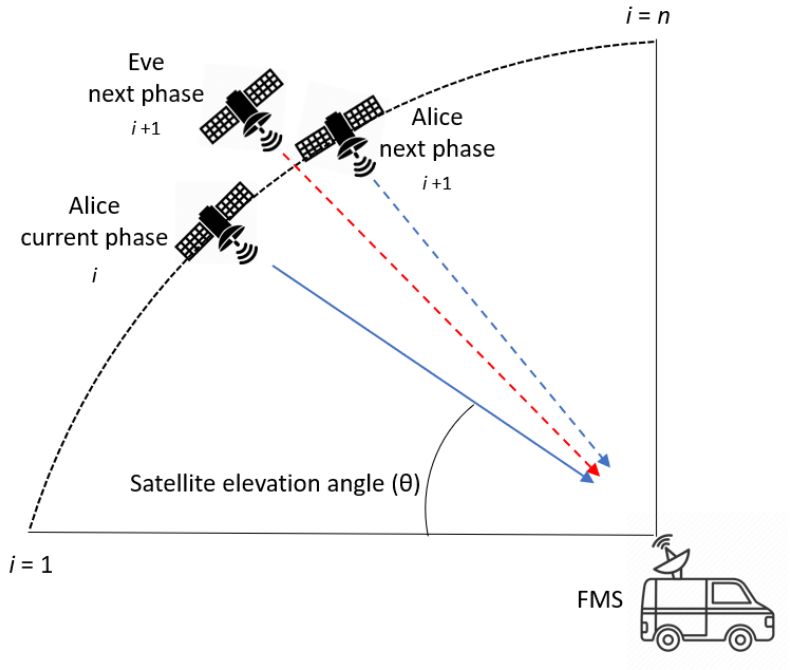

Figure 5.1: System model.

5.3 System Model

The system model for the LEO satellite PLA scheme is illustrated in Figure 5.1. The fixed or mobile satellite services station (FMS) must authenticate the legitimate satellite (Alice) over the communication session while preventing spoofing attacks from an illegitimate satellite (Eve). Eve tries to imitate Alice in order to send incorrect or malicious data to users. The FMS first authenticates Alice using ULA and then PLA is performed over the session. A LEO satellite communication session is the time over which the satellite is continuously serving a given ground user [91]. A session is assumed to have $2n - 1$ phases (time instances). Figure 5.1 shows the first n phases. ULA is performed in the initial phase and DS and RP values are obtained. In subsequent phases, PLA is performed at the FMS which must decide

between the two hypotheses

$$\begin{cases} \mathcal{H}_0: & \text{Alice is transmitting,} \\ \mathcal{H}_1: & \text{Eve is transmitting.} \end{cases} \tag{5.1}$$

Thus, $\mathcal{H}_0$ denotes that the signal is from Alice while $\mathcal{H}_1$ means it is from Eve. If the test is passed, then the current DS and RP values are kept and used to test new DS and RP values in the next phase.

5.3.1 Doppler Frequency Shift

The received signal at the FMS will have a DS given by [92]

$$f_d = \frac{v \times f_c}{c} \times \cos(\phi), \tag{5.2}$$

where v is the velocity of the satellite, c is the speed of light, f_c is the center frequency, and ϕ is the angle between the satellite to FMS link and the direction of motion of the satellite. Consequently, for the same v and f_c at a given time, ϕ will differ between satellites so the DS is unique to a satellite.

5.3.2 Received Power

The power received at the FMS in watts is given by [93]

$$p_r = \frac{p_t \; g_t \; g_r}{(4\pi l/\lambda)^2}, \tag{5.3}$$

where p_t is the transmit power, g_t is the gain of the transmit antenna, g_r is the gain of the receive antenna, l is the distance between the satellite and FMS, and λ is the wavelength. The term $(4\pi l/\lambda)^2$ is known as the free space path loss (FSPL). Figure 5.2 shows the authentication flowchart. In the initial phase, ULA is performed

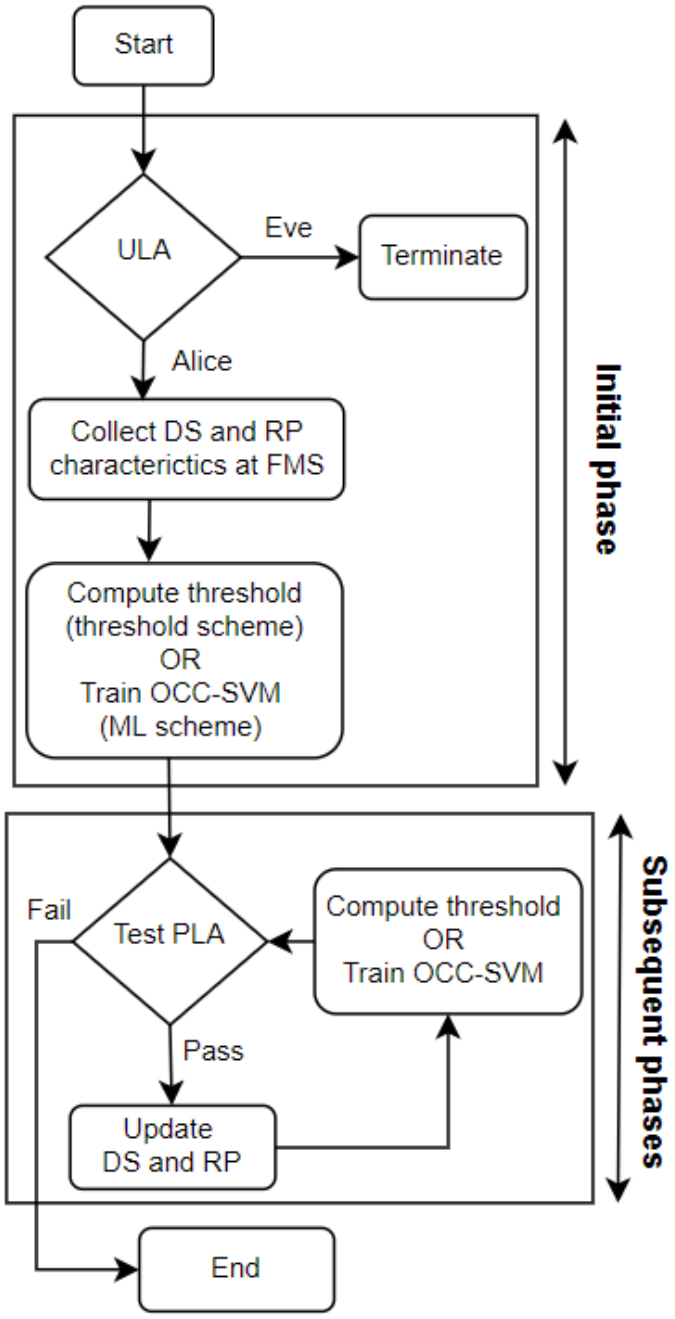

Figure 5.2: Proposed authentication scheme flowchart.

and DS and RP values are obtained at the FMS. Then, the threshold is computed for the threshold authentication scheme or the OCC-SVM is trained for the ML authentication scheme. In subsequent phases, a threshold or OCC-SVM test is performed using new DS and RP values. If the test is passed, these values are kept and used to test the DS and RP values in the next phase, otherwise the connection is terminated. Note that it is intractable for Eve to reproduce the exact DS and RP values at the ground station corresponding to Alice.

5.4 Authentication Based on Physical Characteristics

5.4.1 Estimated Doppler Frequency Shift

Let $\hat{f}_{d,a,i}$, $\hat{f}_{d,a,i+1}$, and $\hat{f}_{d,e,i+1}$ be the estimated DS at the FMS for Alice in the current phase, and Alice and Eve in the next phase, respectively. Then

$$\hat{f}_{d,a,i} = f_{d,a,i} + \varepsilon_{d_1}, \tag{5.4}$$

$$\hat{f}_{d,a,i+1} = f_{d,a,i+1} + \varepsilon_{d_2}, \tag{5.5}$$

$$\hat{f}_{d,e,i+1} = \alpha_{i+1} f_{d,a,i+1} + \varepsilon_{d_3}, \tag{5.6}$$

where $f_{d,a,i}$ and $f_{d,a,i+1}$ are the exact DS for Alice in the current and next phase, respectively, α_{i+1}, $0 < \alpha_{i+1} < 1$, is the ratio between the DS of Alice and Eve, and ε_{d_1},ε_{d_2}, and ε_{d_3} are the DS estimation errors at the FMS due to factors such as approximation and receiver noise. The DS estimation errors can be modeled as Gaussian random variables with $\varepsilon_{d_1} \sim N(0, \sigma_{d_1}^2)$, $\varepsilon_{d_2} \sim N(0, \sigma_{d_2}^2)$, and $\varepsilon_{d_3} \sim N(0, \sigma_{d_3}^2)$ [3].

5.4.2 Estimated Received Power

Let $\hat{p}_{r,a,i}$, $\hat{p}_{r,a,i+1}$, $\hat{p}_{r,e,i+1}$ be the estimated RP at the FMS for Alice in the current phase, and Alice and Eve in the next phase, respectively. Then

$$\hat{p}_{r,a,i} = p_{r,a,i} + \varepsilon_{r_1}, \tag{5.7}$$

$$\hat{p}_{r,a,i+1} = p_{r,a,i+1} + \varepsilon_{r_2}, \tag{5.8}$$

$$\hat{p}_{r,e,i+1} = \beta_{i+1} p_{r,a,i+1} + \varepsilon_{r_3}, \tag{5.9}$$

where $p_{r,a,i}$ and $p_{r,a,i+1}$ are the exact RP for Alice in the current and next phase, respectively, β_{i+1}, $0 < \beta_{i+1} < 1$, is the ratio between the RP of Alice and Eve, and ε_{r_1}, ε_{r_2}, and ε_{r_3} are the RP estimation errors. The RP estimation errors can be modeled as Gaussian random variables with $\varepsilon_{r_1} \sim N(0, \sigma_{r_1}^2)$, $\varepsilon_{r_2} \sim N(0, \sigma_{r_2}^2)$, and $\varepsilon_{r_3} \sim N(0, \sigma_{r_3}^2)$ [106].

In a real system, Alice will have a deviation from the reference trajectory [95] which will affect the signal received at the ground station [96]. It is impossible for Eve to determine this deviation and this will cause errors if Eve tries to manipulate her RP and DS values to imitate Alice. Further, the errors in the estimated DS and RP values at both Eve and Alice will make this task even more difficult.

5.4.3 Threshold Authentication Scheme

In the threshold authentication scheme, the magnitudes of the differences between two consecutive DS and RP values are employed [1]. For Alice, these magnitudes are

$$T_{d,a,i} = |\hat{f}_{d,a,i+1} - \hat{f}_{d,a,i}|, \tag{5.10}$$

$$T_{r,a,i} = |\hat{p}_{r,a,i+1} - \hat{p}_{r,a,i}|, \tag{5.11}$$

respectively, and can be expected to be within small thresholds ϵ_d and ϵ_r. Therefore, the DS hypothesis test can be expressed as

$$\begin{cases} \mathcal{H}_0 : |T_{d,u,i+1} - T_{d,a,i}| \leq \epsilon_d, \\ \mathcal{H}_1 : |T_{d,u,i+1} - T_{d,a,i}| > \epsilon_d, \end{cases} \tag{5.12}$$

where $T_{d,u,i+1}$ is the DS magnitude from an unknown satellite which could be Alice or Eve. Similarly, the RP hypothesis test can be expressed as

$$\begin{cases} \mathcal{H}_0 : |T_{r,u,i+1} - T_{r,a,i}| \leq \epsilon_r, \\ \mathcal{H}_1 : |T_{r,u,i+1} - T_{r,a,i}| > \epsilon_r, \end{cases} \tag{5.13}$$

where $T_{r,u,i+1}$ is the RP magnitude from an unknown satellite which could be Alice or Eve. The AND hypothesis test for the DS and RP magnitudes is given by

$$\begin{cases} \mathcal{H}_0 : |T_{d,u,i+1} - T_{d,a,i}| \leq \epsilon_d \text{ AND} \\ \qquad\quad |T_{r,u,i+1} - T_{r,a,i}| \leq \epsilon_r, \\ \mathcal{H}_1 : |T_{d,u,i+1} - T_{d,a,i}| > \epsilon_d \text{ AND} \\ \qquad\quad |T_{r,u,i+1} - T_{r,a,i}| > \epsilon_r, \end{cases} \tag{5.14}$$

and the OR hypothesis test for these magnitudes is given by

$$\begin{cases} \mathcal{H}_0 : |T_{d,u,i+1} - T_{d,a,i}| \leq \epsilon_d, \text{ OR} \\ \qquad\quad |T_{r,u,i+1} - T_{r,a,i}| \leq \epsilon_r, \\ \mathcal{H}_1 : |T_{d,u,i+1} - T_{d,a,i}| > \epsilon_d \text{ OR} \\ \qquad\quad |T_{r,u,i+1} - T_{r,a,i}| > \epsilon_r. \end{cases} \tag{5.15}$$

The AND and OR authentication schemes provide lower and upper bounds, respectively, on the performance so the performance of other schemes such as soft-

Algorithm 6 Threshold authentication scheme

Authenticate using ULA.
Collect DS and RP values from Alice.
Compute $T_{d,a,i}$ and $T_{r,a,i}$ using (5.10) and (5.11), respectively.
Test $T_{d,u,i+1}$ and $T_{r,u,i+1}$.
while Alice **do**
    Update DS and RP values from Alice.
    Compute $T_{d,a,i}$ and $T_{r,a,i}$ using (5.10) and (5.11), respectively.
    Test $T_{d,u,i+1}$ and $T_{r,u,i+1}$.
end while

decision fusion will lie between them. The threshold authentication scheme is summarized in Algorithm 6. The false alarm rate (FAR), missed detection rate (MDR), and authentication rate (AR) are used to evaluate this scheme. The FAR for the DS and RP threshold authentication schemes is defined as

$$FAR_{d,t} = P(|T_{d,u,i+1} - T_{d,a,i}| > \epsilon_d|\mathcal{H}_0), \tag{5.16}$$

$$FAR_{r,t} = P(|T_{r,u,i+1} - T_{r,a,i}| > \epsilon_r|\mathcal{H}_0), \tag{5.17}$$

respectively. The MDR for the DS and RP threshold authentication schemes is defined as

$$MDR_{d,t} = P(|T_{d,u,i+1} - T_{d,a,i}| \leq \epsilon_d|\mathcal{H}_1), \tag{5.18}$$

$$MDR_{r,t} = P(|T_{r,u,i+1} - T_{r,a,i}| \leq \epsilon_r|\mathcal{H}_1), \tag{5.19}$$

respectively. The FAR and MDR for the AND threshold authentication scheme are defined as

$$\begin{aligned} FAR_{\text{AND},t} = P(&|T_{d,u,i+1} - T_{d,a,i}| > \epsilon_d \text{ AND} \\ &|T_{r,u,i+1} - T_{r,a,i}| > \epsilon_r|\mathcal{H}_0), \end{aligned} \tag{5.20}$$

$$\begin{aligned} MDR_{\text{AND},t} = P(&|T_{d,u,i+1} - T_{d,a,i}| \leq \epsilon_d \text{ AND} \\ &|T_{r,u,i+1} - T_{r,a,i}| \leq \epsilon_r|\mathcal{H}_1), \end{aligned} \tag{5.21}$$

respectively. The FAR and MDR for the OR threshold authentication scheme are defined as

$$FAR_{\mathrm{OR},t} = P(|T_{d,u,i+1} - T_{d,a,i}| > \epsilon_d \text{ OR } |T_{r,u,i+1} - T_{r,a,i}| > \epsilon_r | \mathcal{H}_0), \tag{5.22}$$

$$MDR_{\mathrm{OR},t} = P(|T_{d,u,i+1} - T_{d,a,i}| \leq \epsilon_d \text{ OR } |T_{r,u,i+1} - T_{r,a,i}| \leq \epsilon_r | \mathcal{H}_1), \tag{5.23}$$

respectively. The AR for the DS, RP, AND, and OR threshold authentication schemes is given by

$$\mathrm{AR}_{d,t} = \frac{1}{2} \times \left[(1 - FAR_{d,t}) + (1 - MDR_{d,t}) \right], \tag{5.24}$$

$$\mathrm{AR}_{r,t} = \frac{1}{2} \times \left[(1 - FAR_{r,t}) + (1 - MDR_{r,t}) \right], \tag{5.25}$$

$$\mathrm{AR}_{\mathrm{AND},t} = \frac{1}{2} \times \left[(1 - FAR_{\mathrm{AND},t}) + (1 - MDR_{\mathrm{AND},t}) \right], \tag{5.26}$$

$$\mathrm{AR}_{\mathrm{OR},t} = \frac{1}{2} \times \left[(1 - FAR_{\mathrm{OR},t}) + (1 - MDR_{\mathrm{OR},t}) \right], \tag{5.27}$$

respectively.

5.4.4 Machine Learning Authentication Scheme

In the ML scheme, OCC-SVM is employed using DS, RP, or DS and RP as features for training and testing. In the initial phase, DS and RP values are collected from Alice for OCC-SVM training. Then, DS and RP values from an unknown satellite u, which could be Alice or Eve, are used for testing at the FMS. If the test is passed, the corresponding DS and RP values are used to update the features for training. On the other hand, if the test fails, the connection is terminated.

The DS and RP data vector has the form

$$\mathbf{m} = [\hat{f}_{d,a,i} \quad \hat{p}_{r,a,i}]. \tag{5.28}$$

Algorithm 7 Machine learning authentication scheme

Authenticate using ULA.
Collect DS and RP values from Alice.
Form the training matrix **M**.
Train using OCC-SVM.
Test using OCC-SVM.
while Alice **do**
  Update the training matrix **M**.
  Train using OCC-SVM.
  Test using OCC-SVM.
end while

After Alice is authenticated via ULA, ℓ data vectors corresponding to ℓ samples from Alice

$$\mathbf{d}_j = [\hat{f}_{d,a,i,j} \quad \hat{p}_{r,a,i,j}], j = 1, 2, \ldots, \ell, \tag{5.29}$$

are used for OCC-SVM training. Then, OCC-SVM is used to test ℓ data vectors from u

$$\mathbf{b}_j = [\hat{f}_{d,u,i+1,j} \quad \hat{p}_{r,u,i+1,j}], j = 1, 2, \ldots, \ell. \tag{5.30}$$

If the test is passed the satellite is accepted, the features are updated, and OCC-SVM is retrained. Otherwise, the connection is terminated. The data matrix in phase $i-1$ is

$$\mathbf{M} = \begin{bmatrix} \mathbf{d}_{i+1} \\ \mathbf{d}_{i+2} \\ \vdots \\ \mathbf{d}_{i+\ell} \end{bmatrix}. \tag{5.31}$$

In the initial phase ($i = 1$), this data is from Alice and is used for training. The matrix in subsequent phases ($i > 1$) is first tested, and if the test is passed, the matrix is used for training in the next phase. The ML authentication scheme is summarized in Algorithm 7.

The confusion matrix is used to evaluate the performance of the proposed schemes. True positive (TP) denotes correctly accepting a legitimate satellite, true negative (TN) denotes correctly rejecting an illegitimate satellite, false negative (FN) denotes incorrectly rejecting a legitimate satellite, and false positive (FP) denotes incorrectly accepting an illegitimate satellite. The FAR for DS, RP, and DS and RP is given by

$$FAR_{d,l} = \frac{FN_d}{P}, \tag{5.32}$$

$$FAR_{r,l} = \frac{FN_r}{P}, \tag{5.33}$$

$$FAR_{d,r,l} = \frac{FN_{d,r}}{P}, \tag{5.34}$$

respectively, where FN_d, FN_r, and $FN_{d,r}$ are the FN for DS, RP, and DS and RP, respectively, and $P = TP + FN$. The MDR for DS, RP, and DS and RP is given by

$$MDR_{d,l} = \frac{FP_d}{N}, \tag{5.35}$$

$$MDR_{r,l} = \frac{FP_r}{N}, \tag{5.36}$$

$$MDR_{d,r,l} = \frac{FP_{d,r}}{N}, \tag{5.37}$$

respectively, where FP_d, FP_r, and $FP_{d,r}$ are the FP for DS, RP, and DS and RP, respectively, and $N = TN + FP$. The AR for the DS, RP, and DS and RP when $P = N$ (equal amounts of data from Alice and Eve), is given by

$$AR_{d,l} = \frac{1}{2} \times \left[(1 - FAR_{d,l}) + (1 - MDR_{d,l})\right], \tag{5.38}$$

$$AR_{r,l} = \frac{1}{2} \times \left[(1 - FAR_{r,l}) + (1 - MDR_{r,l})\right], \tag{5.39}$$

$$AR_{d,r,l} = \frac{1}{2} \times \left[(1 - FAR_{d,r,l}) + (1 - MDR_{d,r,l})\right], \tag{5.40}$$

respectively.

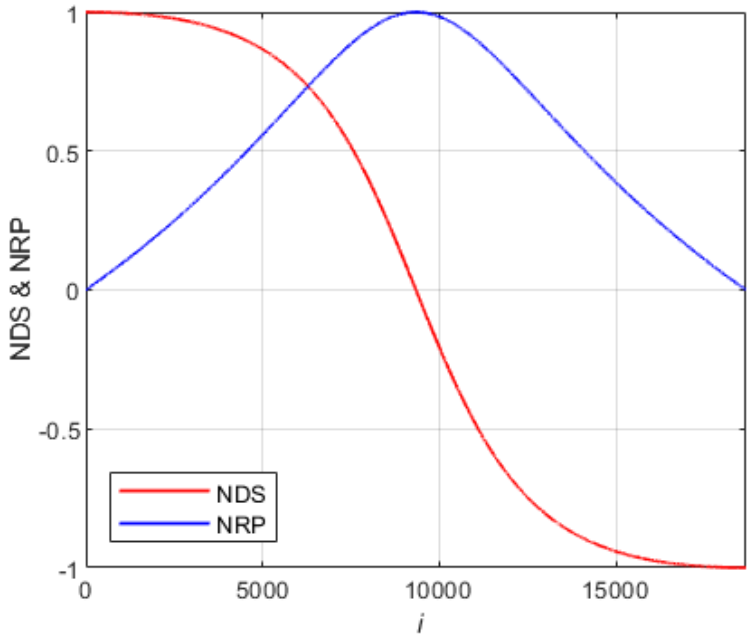

Figure 5.3: Normalized Doppler frequency shift (NDS) and normalized received power (NRP) over the communication session.

5.5 Simulation Results

In this section, the proposed authentication schemes are evaluated using real DS and RP values obtained using STK during the communication session along the satellite trajectories. DS and RP values are obtained every 0.01 s and $n = 9330$. The number of symbols used for training is $\ell = 10$. The proposed ML authentication scheme employs OCC-SVM using the scikit-learn library in Python. The simulation parameters are given in Table 5.1.

5.5.1 DS and RP Over the Communication Session

Tables 5.2 and 5.3 give the range of DS and RP values, respectively, at different altitudes. The normalized Doppler frequency shift (NDS) and normalized received

Table 5.1: Simulation Parameters

Parameter	**Value**
Center frequency	7.5 GHz
Antenna diameter	0.5 m
Modulation	BPSK
Bandwidth	10 MHz
Data rate	10 Mbps
Satellite altitude	2000 km
Tx power for all satellites	10 dBW
ℓ	10

Table 5.2: Range of Doppler Frequency Shifts at Different Altitudes

Altitude	Minimum	Maximum
500 km	0 Hz	165 kHz
1000 km	0 Hz	145 kHz
1500 km	0 Hz	130 kHz
2000 km	0 Hz	120 kHz

Table 5.3: Range of Received Power at Different Altitudes

Altitude	Minimum	Maximum
500 km	-110 dBW	-98 dBW
1000 km	-113 dBW	-102 dBW
1500 km	-115 dBW	-105 dBW
2000 km	-117 dBW	-108 dBW

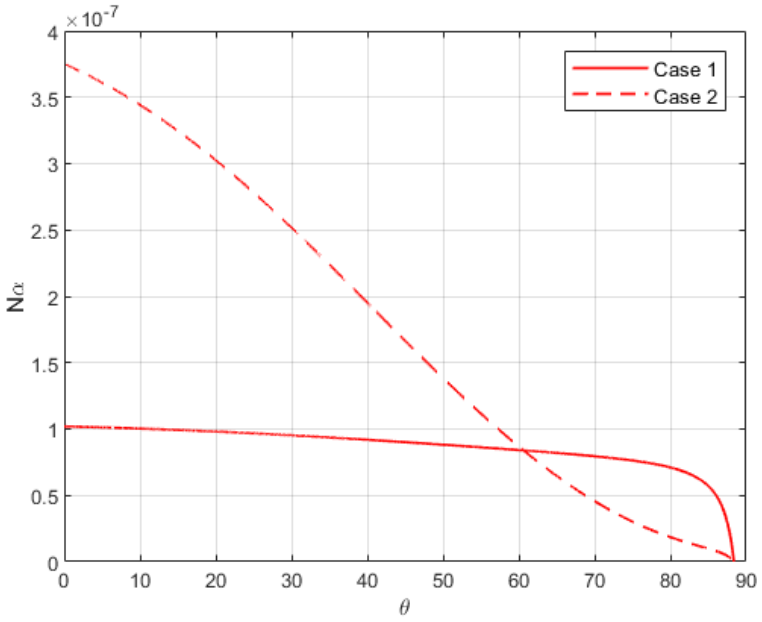

Figure 5.4: Normalized α_i ($N\alpha_i$) versus θ.

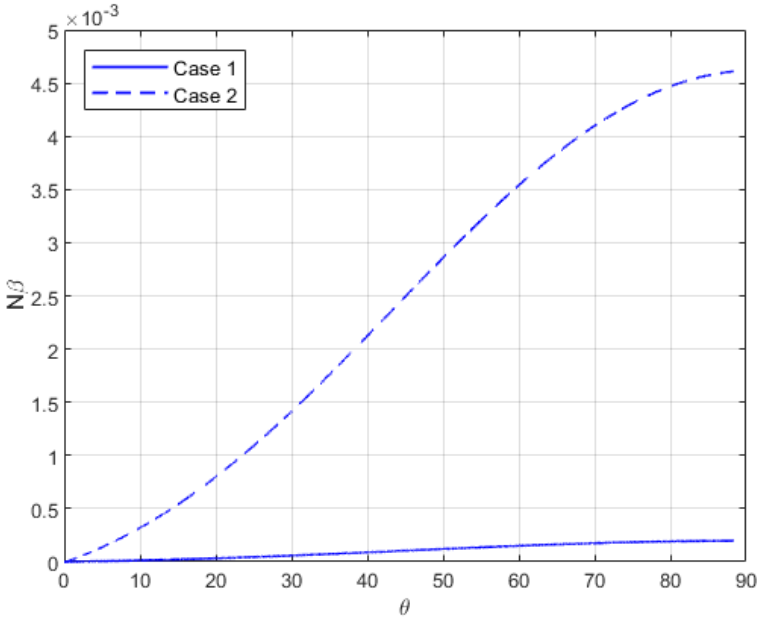

Figure 5.5: Normalized β_i ($N\beta_i$) versus θ.

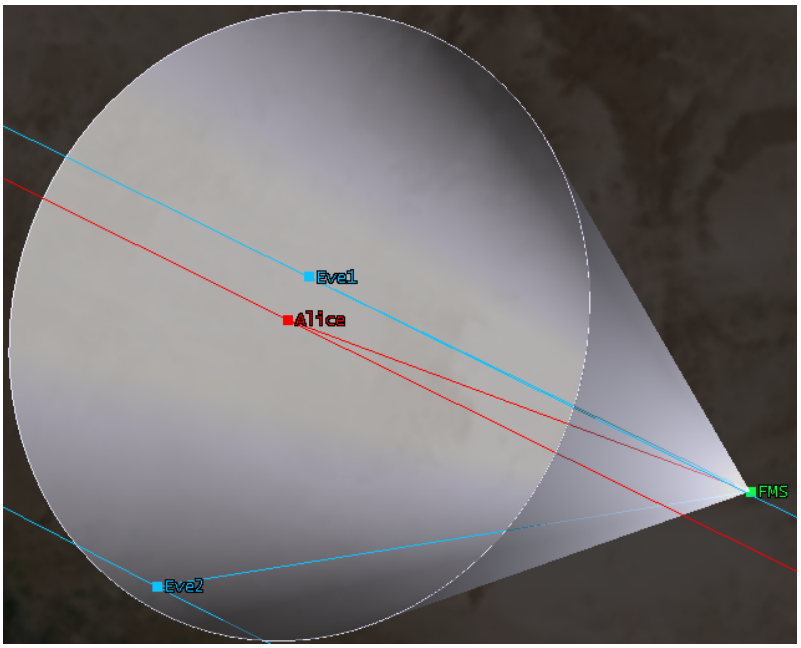

Figure 5.6: The trajectories for Eve and Alice where Eve1 corresponds to case 1 and Eve2 to case 2.

power (NRP) over the communication session are defined as

$$NDS_i = \frac{DS_i}{\max(DS)}, i = 1, 2, \ldots, 2n-1, \tag{5.41}$$

$$NRP_i = \frac{RP_i}{\max(RP)}, i = 1, 2, \ldots, 2n-1, \tag{5.42}$$

respectively, where $\max(DS)$ and $\max(RP)$ are the corresponding maximum values. Figure 5.3 presents the NDS and NRP for Alice at an altitude of 2000 km. This shows that the NDS and NRP values in the first half of the communication session are similar to those in the second half. Thus, only DS and RP values for phases 1 to n are considered in the simulations. The DS and RP ratios between Alice and Eve are given by

$$\alpha_i = \frac{f_{d,a,t_i}}{f_{d,e,t_i}}, i = 1, 2, \ldots, n, \tag{5.43}$$

$$\beta_i = \frac{p_{r,a,t_i}}{p_{r,e,t_i}}, i = 1, 2, \ldots, n, \tag{5.44}$$

respectively, and the corresponding normalized values are

$$N\alpha_i = \frac{\alpha_i}{\max(\alpha_i)}, i = 1, 2, \ldots, n, \tag{5.45}$$

$$N\beta_i = \frac{\beta_i}{\max(\beta_i)}, i = 1, 2, \ldots, n, \tag{5.46}$$

where $\max(\alpha_i)$ and $\max(\beta_i)$ are the maximum α_i and β_i, respectively. Figures 5.4 and 5.5 show $N\alpha_i$ and $N\beta_i$, respectively, versus θ for two cases. In case 1, the DS and RP values are obtained for Alice and Eve when the trajectories are very close, while in case 2, the trajectories are far apart. However, in both cases Eve is within the half power beam width (HPBW) of the FMS receive antenna as indicated in Figure 5.6. Figures 5.4 and 5.5 show that the variations in $N\alpha_i$ and $N\beta_i$ are negligible in both cases. For example, the difference between the largest and smallest values of $N\alpha_i$ is 4×10^{-7} while the corresponding difference in $N\beta_i$ is less than 5×10^{-3}. Thus, it is assumed in the following that $\alpha_i = \alpha$ and $\beta_i = \beta$ over the communication session. In the simulations, $\sigma_{d_1}^2 = \sigma_{d_2}^2 = \sigma_{d_3}^2 = \sigma_{r_1}^2 = \sigma_{r_2}^2 = \sigma_{r_3}^2 = \sigma^2$, $\alpha_i = \beta_i = \alpha$, $\epsilon_d = 0.1 \times T_{d,a,i}$, and $\epsilon_r = 0.1 \times T_{r,a,i}$.

5.5.2 Threshold Authentication Scheme Performance

Figures 5.7a and 5.7b present the MDR, FAR, and AR for the DS, RP, AND, and OR threshold authentication schemes averaged over the communication session versus α with $\sigma^2 = 0.08$ and σ^2 with $\alpha = 0.4$, respectively. Figure 5.7a shows that the FAR for DS is lower than the FAR for RP for all values of α, and the minimum FAR is 48.9% for DS and 53.2% for RP. Figure 5.7b indicates that the FAR for DS is lower than the FAR for RP for all values of σ^2, but there is a small increase with σ^2. For example, the FAR at $\sigma^2 = 0.02$ is 45.6% for DS and 46.3% for RP, while at $\sigma^2 = 0.09$ the AR is 48.9% for DS and 54.9% for RP. Figure 5.7a shows that the MDR for DS is lower than the MDR for RP at low α, but at high α the converse is

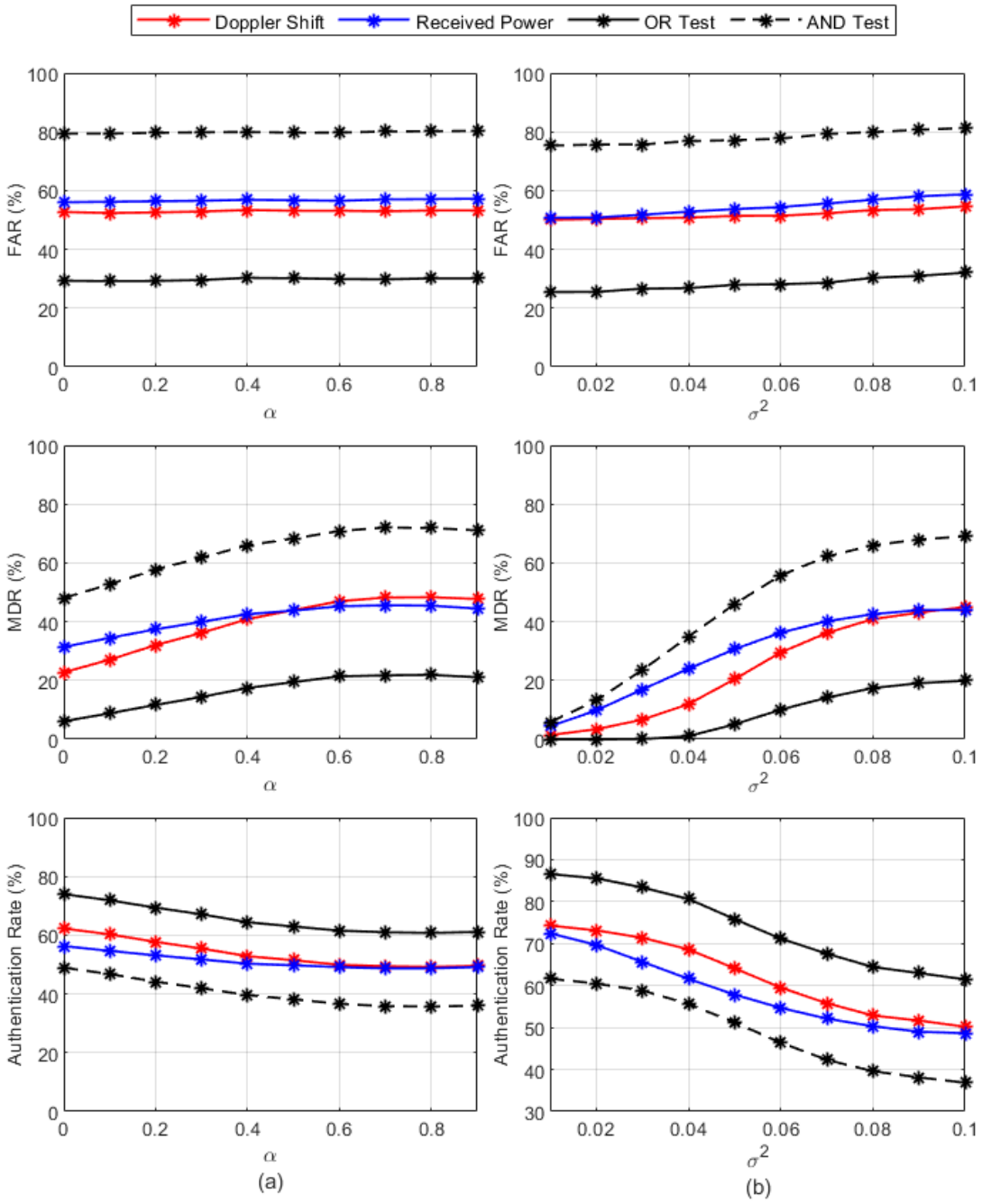

Figure 5.7: MDR, FAR, and AR for the DS, RP, AND, and OR threshold authentication schemes averaged over the communication session versus (a) α with $\sigma^2 = 0.08$, and (b) σ^2 with $\alpha = 0.4$.

true. For example, the MDR with $\alpha = 0.1$ is 25.2% for DS and 35.4% for RP. On the other hand, Figure 5.7b indicates that the MDR for DS is lower than for RP for most values of σ^2, but both increase with σ^2. For example, the MDR at $\sigma^2 = 0.03$ is 5.4% for DS and 16.2% for RP, while at $\sigma^2 = 0.08$ the corresponding values are 42.0% and 44.4%.

Figures 5.7a and 5.7b show that the AR for DS is higher than for RP for all values of α and σ^2. For example, in Figure 5.7a the AR at $\alpha = 0.1$ for DS is 63.1% versus 57.6% for RP. However, this difference decreases with increasing α and is less than 1% at $\alpha = 0.8$. The AR with AND is 48.4% at $\alpha = 0.1$, while for OR it is 74.5%. Both decrease with increasing α so at $\alpha = 0.9$ the AR with AND is 35.2% and with OR is 60.6%. Further, Figure 5.7b shows that the AR for DS is higher than for RP for all values of σ^2. For example, the AR at $\sigma^2 = 0.02$ for DS is 75.5% versus 71.4% for RP. The AR with AND is 64.2% at $\sigma^2 = 0.01$ while for OR it is 88.9%. Both decrease with increasing σ^2, so at $\sigma^2 = 0.1$ the AR with AND is 36.6% and with OR is 61.3%.

Figure 5.8 presents the AR for the DS and RP threshold authentication schemes versus θ with $\alpha = 0.4$ and $\sigma^2 = 0.03$, 0.05, and 0.06. This shows that the AR for DS is better than for RP at low θ, but the converse is true at high θ. For example, at $\theta = 20^\circ$ and $\sigma^2 = 0.06$, the AR for DS is 71.4% versus 51.1% for RP. However, at $\theta = 80^\circ$ and $\sigma^2 = 0.06$, the AR for DS is 48.0% versus 71.1% for RP. This shows that using DS at low θ and switching to RP at high θ can provide good authentication performance. For example, at $\sigma^2 = 0.03$, $\sigma^2 = 0.05$, and $\sigma^2 = 0.06$ the minimum AR with this approach is 76.5%, 71.6%, and 64.0%, respectively.

5.5.3 Machine Learning Authentication Scheme Performance

Figures 5.9a and 5.9b present the AR for the DS, RP, and DS and RP ML authentication schemes with $\ell = 10$ and $\eta = 0.5$ averaged over the communication session versus α with $\sigma^2 = 0.08$ and σ^2 with $\alpha = 0.4$, respectively. Figure 5.9a shows that

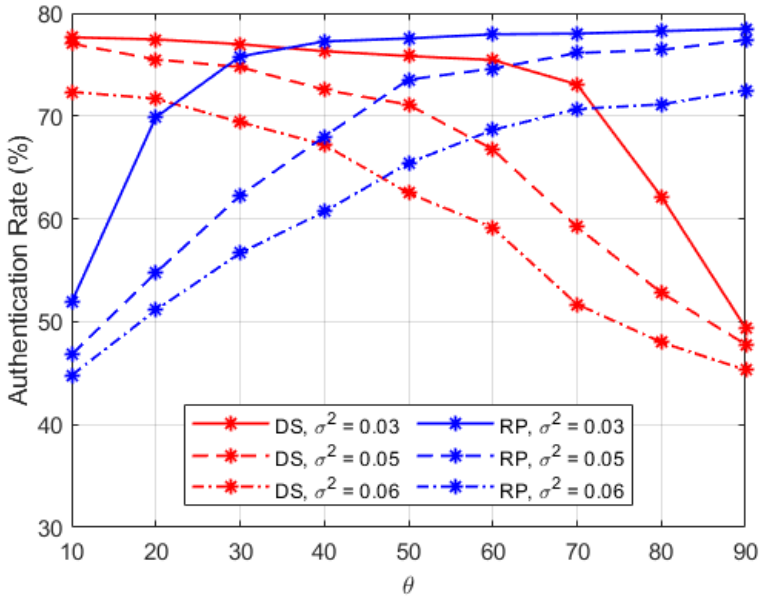

Figure 5.8: AR for the DS and RP threshold authentication schemes versus θ with $\alpha = 0.4$ and $\sigma^2 = 0.03, 0.05$, and 0.06.

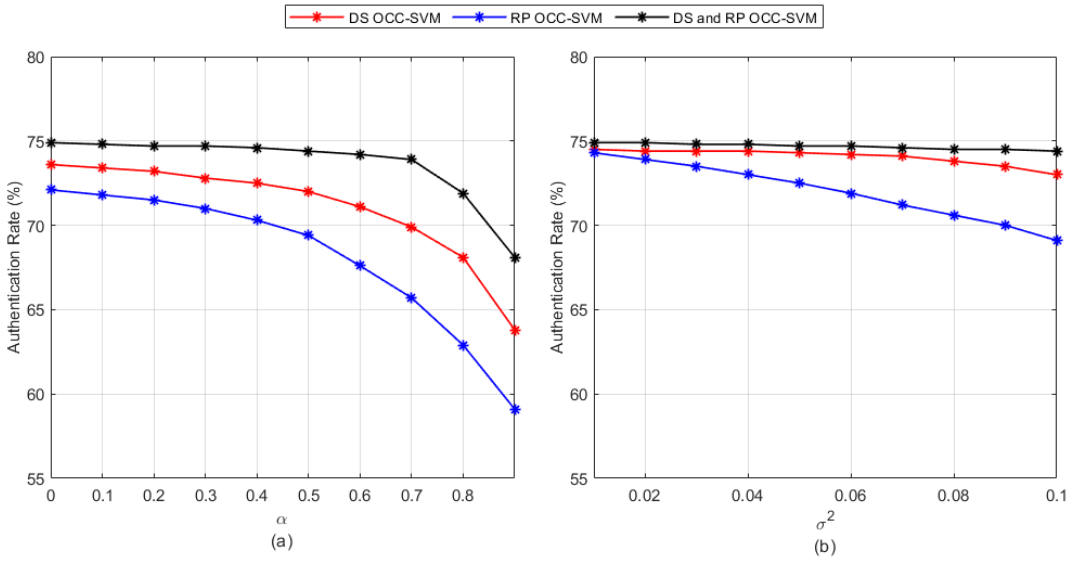

Figure 5.9: AR for the DS, RP, and DS and RP ML authentication schemes averaged over the communication session with $\eta = 0.5$ and $\ell = 10$ versus (a) α with $\sigma^2 = 0.08$, and (b) σ^2 with $\alpha = 0.4$.

the AR for DS is higher than for RP for all values of α. For example, the AR at $\alpha = 0.2$ for DS is 73.3% versus 71.6% for RP, and the AR at $\alpha = 0.8$ for DS is 68.4% versus 63.1% for RP. However, the AR for DS and RP is higher than with DS or RP separately. For example, the AR at $\alpha = 0.6$ for DS and RP is 74.2% versus 71.2% for DS and 67.8% for RP. Figure 5.9b also shows that the AR for DS and RP is higher than with DS or RP separately for all values of σ^2. For example, the AR at $\sigma^2 = 0.06$ for DS and RP is 74.7% versus 74.2% for DS and 71.9% for RP.

Figure 5.10 presents the AR for the separate DS and RP ML authentication schemes versus θ with $\alpha = 0.4$, $\ell = 10$, $\sigma^2 = 0.03$ and 0.06, and $\eta = 0.5$. This shows that the AR for DS is higher than for RP at low θ, but the converse is true at high θ. For example, at $\theta = 20^\circ$ and $\sigma^2 = 0.06$, the AR for DS is 74.7% versus 70.5% for RP, and at $\theta = 80^\circ$ and $\sigma^2 = 0.06$, the AR for DS is 69.0% versus 74.9% for RP. This shows that using DS at low θ and switching to RP at high θ can provide good authentication performance. For example, at $\sigma^2 = 0.03$ and $\sigma^2 = 0.06$ the minimum AR in this case is 74.3% and 74.1%, respectively.

5.5.4 Authentication Scheme Performance Comparison

Figures 5.11a and 5.11b present the AR for the DS, RP, AND, and OR threshold authentication schemes and the DS, RP, and DS and RP ML authentication schemes averaged over the communication session with $\eta = 0.1$ and 0.5 and $\ell = 10$ versus α with $\sigma^2 = 0.02$ and σ^2 with $\alpha = 0.3$, respectively. Figure 5.11a shows that the AR for DS is higher than for RP for all values of α, and the AR decreases with α. Further, the AR for the DS or RP ML authentication schemes with $\eta = 0.5$ is lower than the AR for the DS or RP threshold authentication schemes at low α and higher at high α. However, the AR for DS or RP ML authentication with $\eta = 0.1$ is higher than the AR for DS or RP threshold authentication for all values of α. For example, the AR at $\alpha = 0.1$ for the DS and RP threshold authentication schemes is 76.3% and 74.7%, respectively, but the AR for the corresponding ML

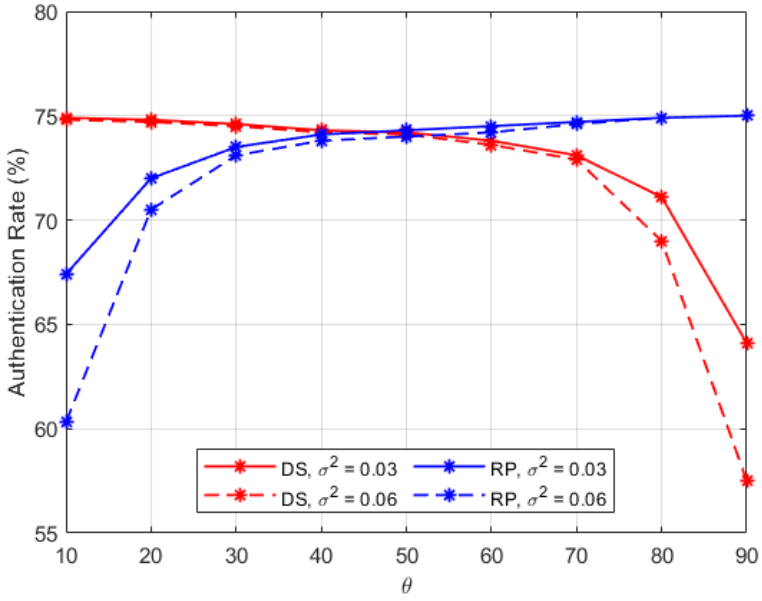

Figure 5.10: AR versus θ for the DS and RP ML authentication schemes with $\eta = 0.5$, $\alpha = 0.4$, $\ell = 10$, and $\sigma^2 = 0.03$ and 0.06.

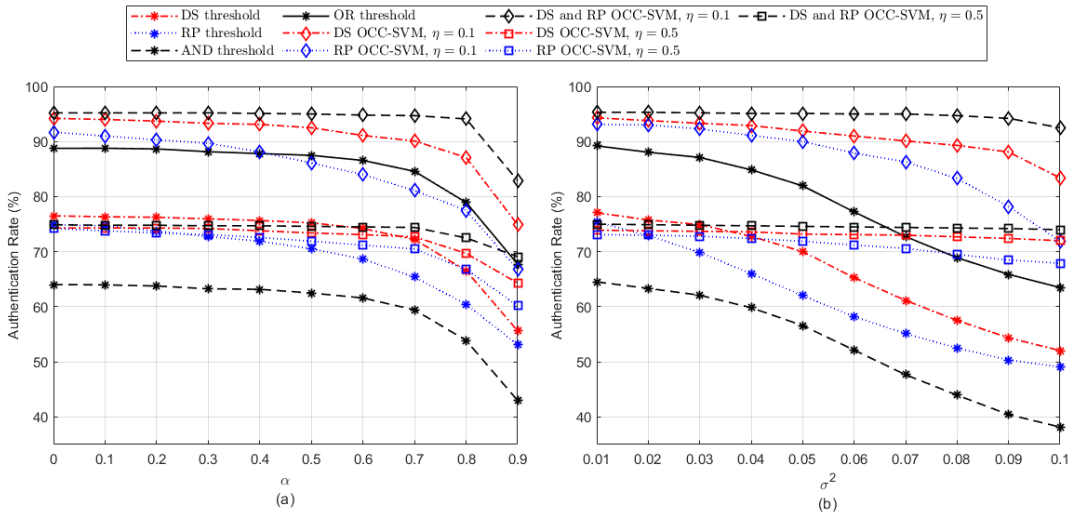

Figure 5.11: AR for the DS, RP, AND, and OR threshold authentication schemes and the DS, RP, and DS and RP ML authentication schemes averaged over the communication session with $\eta = 0.1$ and 0.5, and $\ell = 10$ versus (a) α with $\sigma^2 = 0.02$, and (b) σ^2 with $\alpha = 0.3$.

authentication schemes is 94.0% and 91.2% with $\eta = 0.1$ and 74.3% and 73.8% with $\eta = 0.5$, respectively. In addition, the AR at $\alpha = 0.9$ for the DS and RP threshold authentication schemes is 55.6% and 53.1%, respectively, while the AR for the corresponding ML authentication schemes is 74.9% and 66.8% with $\eta = 0.1$ and 64.3% and 60.2% with $\eta = 0.5$, respectively. However, using both DS and RP for ML authentication with $\eta = 0.1$ provides the highest AR followed by OR threshold authentication, DS and RP ML authentication with $\eta = 0.5$, and then AND threshold authentication. For example, the AR at $\alpha = 0.1$ is 95.2%, 88.7%, 74.8%, and 64% for DS and RP ML authentication with $\eta = 0.1$, OR threshold authentication, DS and RP ML authentication with $\eta = 0.5$, and AND threshold authentication, respectively, and the corresponding AR at $\alpha = 0.8$ is 94.0%, 78.9%, 72.5%, and 53.9%, respectively.

Figure 5.11b shows that the AR for DS or RP ML authentication with $\eta = 0.5$ is lower than the corresponding threshold authentication schemes at low σ^2 and higher at high σ^2. The AR for DS or RP ML authentication with $\eta = 0.1$ is higher than the AR for DS or RP threshold authentication for all values of σ^2. For example, the AR at $\sigma^2 = 0.01$ for DS and RP threshold authentication is 77.1% and 75.4%, respectively, but the corresponding values for ML authentication are 94.3% and 93.1% with $\eta = 0.1$ and 73.9% and 73.1% with $\eta = 0.5$, respectively. The AR at $\sigma^2 = 0.1$ for DS and RP threshold authentication is 52.0% and 49.0%, respectively. However, the AR for the ML authentication schemes with DS and RP separately is 83.4% and 71.9% with $\eta = 0.1$ and 72.0% and 67.9% with $\eta = 0.5$, respectively. The highest AR is achieved with both DS and RP ML authentication with $\eta = 0.1$. For example, the AR at $\sigma^2 = 0.01$ is 95.3%, 89.2%, 75%, and 64.5% for DS and RP ML authentication with $\eta = 0.1$, OR threshold authentication, DS and RP ML authentication with $\eta = 0.5$, and AND threshold authentication, respectively. Furthermore, the AR at $\sigma^2 = 0.1$ is 92.5%, 74%, 63.5%, and 38.1% for DS and RP ML authentication with $\eta = 0.1$, OR threshold, DS and RP ML authentication with

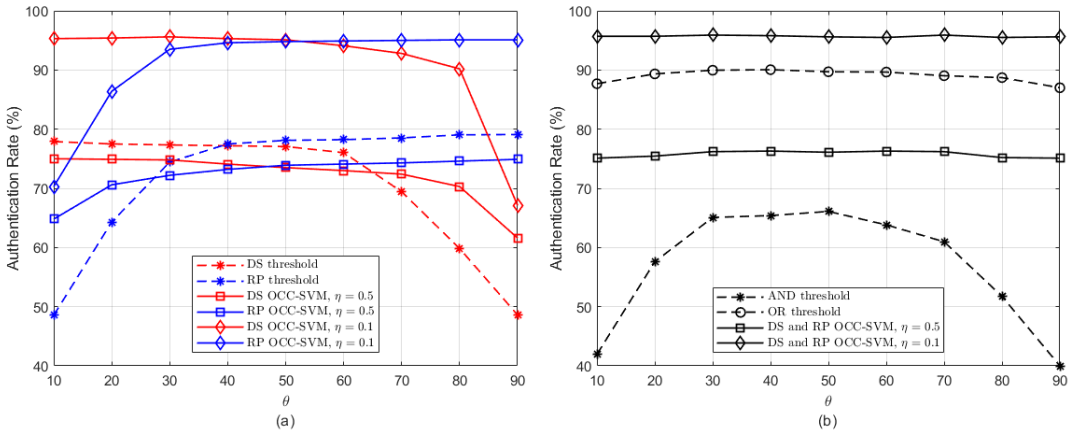

Figure 5.12: AR versus θ with $\sigma^2 = 0.04$ and $\alpha = 0.3$ for the (a) separate DS and RP threshold and separate DS and RP ML authentication schemes with $\eta = 0.1$ and 0.5 and $\ell = 10$, and (b) AND and OR threshold authentication schemes, and DS and RP ML authentication schemes with $\eta = 0.1$ and 0.5 and $\ell = 10$.

$\eta = 0.5$, and AND threshold authentication, respectively.

Figures 5.12a and 5.12b present the AR versus θ with $\sigma^2 = 0.04$ and $\alpha = 0.3$ for the separate DS and RP threshold and separate DS and RP ML authentication schemes with $\eta = 0.1$ and 0.5 and $\ell = 10$, and the AND threshold, OR threshold, and DS and RP ML authentication schemes with $\eta = 0.1$ and 0.5 and $\ell = 10$, respectively. Figure 5.12a shows that the AR for DS is higher than with RP at low θ, but at high θ the AR for RP is higher than with DS. For example, at $\theta = 20°$ the AR for DS and RP threshold authentication is 77.5% and 64.3%, respectively, and the AR for the corresponding ML authentication is 95.4% and 86.4% with $\eta = 0.1$ and 74.9% and 70.6% with $\eta = 0.5$, respectively. However, at $\theta = 80°$, the AR for DS and RP threshold authentication is 59.8% and 79.0%, respectively, and the corresponding values for ML authentication are 90.2% and 95.1% with $\eta = 0.1$ and 70.3% and 74.6% with $\eta = 0.5$, respectively. Thus, it can be concluded that when the DS and RP are used separately for authentication, DS should be considered at low θ and RP at high θ. For example, in this case the minimum AR is 94.8%, 77.3%, and 73.5% for ML authentication with $\eta = 0.1$, threshold authentication,

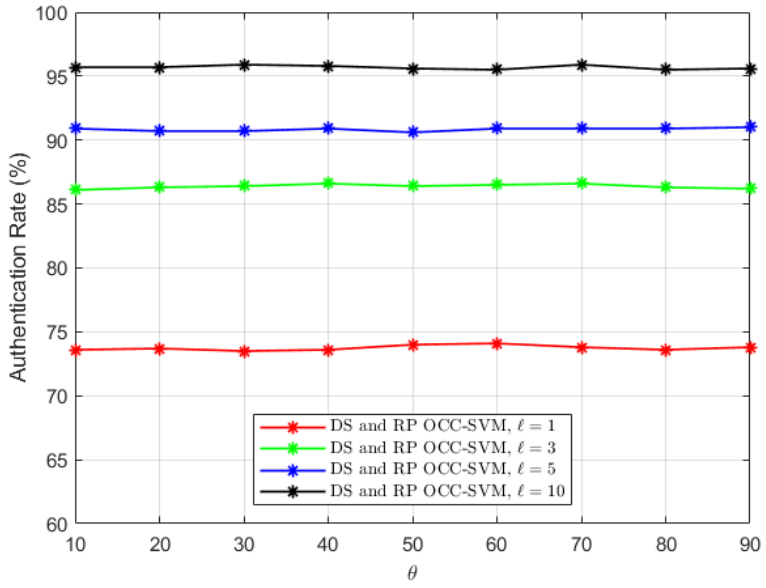

Figure 5.13: AR for DS and RP ML authentication scheme versus θ with $\sigma^2 = 0.04$, $\eta = 0.1$, and $\alpha = 0.3$ for $\ell = 1, 3, 5$ and 10.

and ML authentication with $\eta = 0.5$, respectively. Finally, threshold authentication provides better performance than ML authentication with large η when DS and RP are used separately, but the converse is true with small η. Figure 5.12b presents the AR when DS and RP are both used for authentication. This shows that the AR in this case with small η is greater than with AND threshold authentication. For example, at $\theta = 50°$ the AR is 95.6%, 89.7%, 76.1%, and 66.1% for DS and RP ML authentication with $\eta = 0.1$, OR threshold authentication, DS and RP ML authentication with $\eta = 0.5$, and AND threshold authentication, respectively.

Figure 5.13 presents the AR for DS and RP ML authentication versus θ with $\sigma^2 = 0.04$, $\eta = 0.1$, and $\alpha = 0.3$ for $\ell = 1, 3, 5$ and 10. This shows that the AR increases with ℓ. For example, at $\theta = 80°$ the AR is 73.6%, 86.3%, 90.9%, and 95.5% for $\ell = 1, 3, 5$, and 10, respectively.

5.6 Conclusion

Physical layer authentication (PLA) has emerged as an alternative paradigm that uses physical characteristics to achieve authentication. A PLA scheme was proposed for low earth orbit (LEO) satellites using Doppler frequency shift (DS) and received power (RP) characteristics. This scheme employs hypothesis testing using a threshold or machine learning (ML) to discriminate between legitimate and illegitimate satellites. Estimation errors in the DS and RP values were considered and the performance was evaluated based on real satellite data from the system tool kit (STK). Results were presented which show that DS provides a high authentication rate (AR) at small elevation angles (θ) and decreases with θ, while RP provides a low AR at small θ and increases with θ. Further, ML authentication with a small percentage of outliers η in the training data provides the highest AR. Finally, the AR for the ML authentication scheme increases with the amount of training data ℓ.

Chapter 6

Game Theoretic Spoofing Detection for Space Information Networks using Physical Attributes

In this chapter, game theoretic PLA based on Doppler frequency spread (DS) and received power (RP) attributes is proposed to provide effective authentication for these satellites. Hypothesis testing with a threshold is used to distinguish between legitimate and illegitimate (spoofer) satellites. Then, a zero-sum PLA game in which the ground station (GS) chooses the optimal detection threshold (τ^*) to maximize its utility and a spoofing satellite (s) chooses the optimal attack probability (k^*) to maximize its utility. Results are presented to demonstrate the effectiveness of the proposed approach.

6.1 Related Work

Recently, game theory has been used in wireless communication systems. In [107], optimal anti-jamming communications for cognitive radio with perfect channel state information (CSI) was formulated as a zero-sum game. A non-cooperative random access game for jamming attacks in a wireless network with unknown jamming was considered in [108]. In [109] a Q-learning spoofing detection PLA scheme was proposed which employs radio channel information to detect spoofing attacks in multiple-input multiple-output (MIMO) systems. A zero-sum game was formulated for the interactions between a receiver and spoofer and Q-learning was used to obtain the optimal test threshold for spoofing detection. The impact of the number of antennas on the performance of the dynamic authentication game was investigated.

In [110], spoofing detection based on Q-learning and Dyna-Q was proposed using received signal strength indicator (RSSI) values. Interactions between a receiver and spoofer were modeled as a zero-sum authentication game. The receiver selects the hypothesis test threshold to maximize their utility in spoofing detection based on Bayesian risk whereas the spoofer selects their attack frequency to decrease this utility. The optimal test threshold for spoofing detection was obtained using reinforcement learning.

In [54], a zero-sum game was formulated between a receiver and spoofer in a MIMO system and Q-learning was used to obtain the optimal test threshold at the receiver. A Dyna architecture and prioritized sweeping (Dyna-PS) were used to improve spoofing detection in time-varying radio channels. The proposed approach was implemented using universal software radio peripherals (USRPs) and evaluated in an indoor environment. Experimental results show that the Dyna-PS-based spoofing detection algorithm reduces the spoofing detection error rates and increases the utility of the receiver compared with the Q-learning-based algorithm. The performance of the Dyna and Q-learning schemes improves with the number of transmit

and receive antennas. Results were presented which show that both the spoofing detection error rate and spoofing rate decrease with the number of transmit and receive antennas.

In [98], a game theoretic PLA framework was proposed using the spatial correlation of the received signal strength (RSS) in an unmanned aerial vehicle (UAV) system. A zero-sum game was formulated in which the receiver and spoofer select their RSS detection threshold and attack probability, respectively, to maximize their respective utilities based on a hypothesis test.

6.2 Contributions

In [3], a PLA scheme for validating satellites using the Doppler frequency spread (DS) was proposed. It is used prior to initial access to the land mobile satellite (LMS) system to prevent an attacker from impersonating a user. The DS can be calculated using either signal observations or calculations based on satellite broadcast ephemeris. In [99, 100], DS and received power (RP) attributes were used for LEO satellite PLA. Hypothesis testing with a threshold or machine learning (ML) was used to distinguish between legitimate and illegitimate satellites. Motivated by existing game theoretic PLA schemes, in this paper the PLA schemes proposed for LEO satellites in [99, 100] are extended by using a zero-sum game for the interaction between the ground station (GS) and spoofer satellite (s). In particular, a game theoretic PLA scheme is proposed for detecting spoofing satellites in LEO satellites communication systems. The DS and RP are used as attributes to discriminate between legitimate and illegitimate (spoofer) satellites at the GS. Interactions between the GS and s are formulated as a zero-sum game in which s endeavors to deceive the GS with the optimal attack probability and the GS attempts to authenticate the legitimate satellite (l) using the optimal detection threshold. The contributions of this chapter are as follows.

- A game theoretic PLA scheme is proposed to detect spoofing attacks in LEO satellite communication systems.
- The DS and RP are used as attributes to discriminate between l and s at the GS.
- A zero-sum PLA game is formulated where the GS selects their optimal detection threshold and s selects its optimal attack probability to maximize their respective utilities.
- Two-line element (TLE) data for real satellites is used for performance evaluation. TLE data is orbital data for Earth-orbiting objects [36].
- It is shown that the mean of the magnitude differences between the current and previous DS and RP values at the GS can be accurately approximated using third-degree polynomials.

6.3 System Model

The system model for LEO satellite PLA is shown in Figure 6.1. The GS must authenticate l over the communication session while preventing spoofing attacks from s. s tries to imitate l in order to send incorrect or malicious data to users. Assume that the GS has authenticated l in phase i of the session using the current and previous DS and RP values. This is achieved using the normalized differences in magnitude between the ith and $(i-1)$th values from l, denoted by $\hat{m}_{l,i}$ [111]. In phase $i+1$, the GS computes the normalized differences in magnitude between the $(i+1)$th DS and RP values from an unknown satellite $u = \{l, s\}$ and the ith values from l, denoted by $\hat{m}_{u,i+1}$. To determine the legitimacy of u, the GS must decide between the two hypotheses

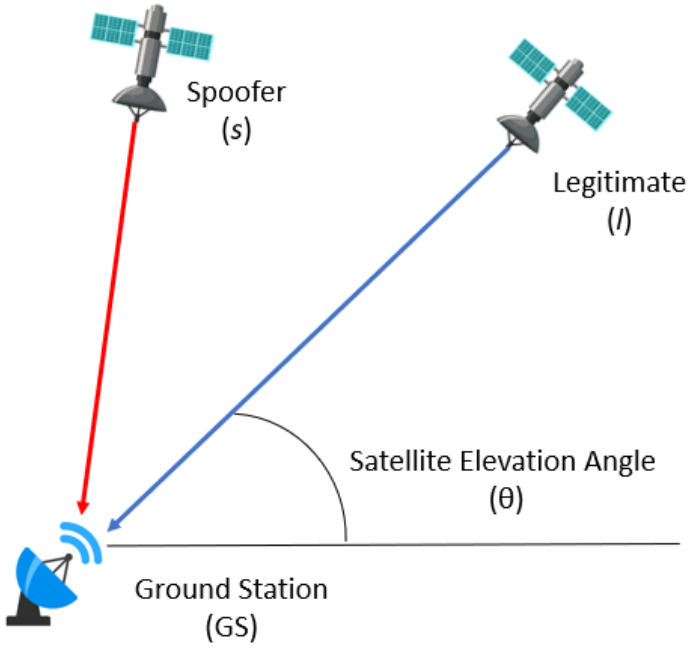

Figure 6.1: The PLA system model. The GS selects the optimal threshold to maximize its utility while s tries to select the attack probability to minimize this utility.

Figure 6.2: The trajectories for l and s within the GS receive antenna HPBW.

$$\begin{cases} H_0 : |\hat{m}_{u,i+1} - \hat{m}_{l,i}| \leq \tau, \\ H_1 : |\hat{m}_{u,i+1} - \hat{m}_{l,i}| > \tau. \end{cases} \tag{6.1}$$

where τ is the authentication threshold. Thus, the null hypothesis H_0 denotes that l is transmitting while the alternative hypothesis H_1 means that s is transmitting. If the null hypothesis is true, then the new normalized differences in magnitude are kept for use in phase $i + 2$. Each time step, the GS tests the legitimacy of u and if legitimate directs its receive antenna to the trajectory of l which is known to the GS. Thus, the satellite elevation angle (θ) considered is for l. A worst-case scenario is considered where s satellite is very close to l with a maximum distance between them of 5 km along the trajectory. In this case, s is within the GS receive antenna half power beamwidth (HPBW).

6.3.1 Doppler Frequency Spread

The received signal DS at the GS is [92]

$$f_d = \frac{v \times f_c}{c} \times \cos(\phi), \tag{6.2}$$

where f_c is the carrier frequency, c is the speed of light, v is the satellite velocity, and ϕ is the angle between the satellite trajectory and the satellite to GS link.

6.3.2 Received Power

The RP at the GS is given by [93]

$$p_r = \frac{p_t \ g_t \ g_r}{(4\pi d/\lambda)^2}, \tag{6.3}$$

where p_t is the transmit power, g_t is the transmit antenna gain, g_r is the receive antenna gain, d is the distance between the receiver and transmitter, and λ is the

wavelength.

6.4 Game Theoretic PLA Scheme

In a real system, l will deviate from the reference trajectory [95] which will affect the signal received at the GS [96]. It is impossible for s to determine this deviation and this will cause errors if s tries to manipulate their RP and DS values to imitate l. Further, the variations in the DS and RP values at both s and l will make this task even more difficult.

The system tool kit (STK), which is a link budget analysis tool, is used with real satellite TLE orbital data to obtain the DS and RP values for l and s. STK provides a comprehensive set of tools for designing, analyzing, and optimizing satellite systems. It enables users to model and simulate satellite orbits, predict coverage, and evaluate system performance. Complex link budgets can be developed and interference analyses performed. Advanced visualization features allow for accurate representations of satellites and ground stations to aid in system modeling. STK is used by satellite professionals to optimize operations, reduce risk, and maximize mission success. A worst-case situation is considered where s is very close to l and both have altitude 500 km. Figure 6.2 presents the trajectories for l and s which are assumed to be within the GS HPBW. The DS and RP values for l and s are obtained along these trajectories.

6.4.1 Doppler Frequency Spread

Let $f_{d,l,i-1}$, $f_{d,l,i}$, and $f_{d,s,i}$ be the DS at the GS for l in the previous phase, l in the current phase, and s in the current phase, respectively. Then the differences in magnitude between the current and previous DS values at the GS for l and s are given by

$$m_{d,l,i} = |f_{d,l,i} - f_{d,l,i-1}|, \tag{6.4}$$

$$m_{d,s,i} = |f_{d,s,i} - f_{d,l,i-1}|, \tag{6.5}$$

respectively. Due to approximation errors and receiver noise, the actual values for l and s can be expressed as

$$m_{d,l,i} = m_{n,d,l,i} + \varepsilon_{d_l}, \tag{6.6}$$

$$m_{d,s,i} = m_{n,d,s,i} + \varepsilon_{d_s}, \tag{6.7}$$

respectively, where ε_{d_l} and ε_{d_s} can be modeled as zero mean Gaussian random variables with variances $\sigma^2_{d_l}$ and $\sigma^2_{d_s}$ [3].

Curve fitting is employed to obtain the best fit for $m_{d,s,i}$ and this is used in the analysis. Figure 6.3a gives the values of $f_d(\theta)$, $m_{d,l,i}$ and $m_{d,s,i}$ versus the satellite elevation angle θ. The best fit third-degree polynomial for $m_{d,s,i}$ is

$$f_d(\theta) = a_1\theta^3 + a_2\theta^2 + a_3\theta + a_4, \tag{6.8}$$

where the coefficients a_1, a_2, a_3, and a_4 are -2.12×10^{-6}, 9.25×10^{-5}, -0.0025, and 1.011, respectively. The time step is 0.01 s and the values are normalized for comparison purposes. The corresponding root mean squared error (RMSE) is 0.0036. These results show that the values for s can be modeled using $f_d(\theta)$ while the values for l can be modeled as 0. Thus, $m_{d,l,i} \sim N(0, \sigma^2_{d_l})$ and $m_{d,s,i} \sim N(f_d(\theta), \sigma^2_{d_s})$.

6.4.2 Received Power

Let $p_{r,l,i-1}$, $p_{r,l,i}$, and $p_{r,s,i}$ be the RP at the GS for l in the previous phase, l in the current phase, and s in the current phase, respectively. Then the differences in magnitude between the current and previous RP values at the GS for l and s are given by

$$m_{r,l,i} = |p_{r,l,i} - p_{r,l,i-1}|, \tag{6.9}$$

$$m_{r,s,i} = |p_{r,s,i} - p_{r,l,i-1}|, \tag{6.10}$$

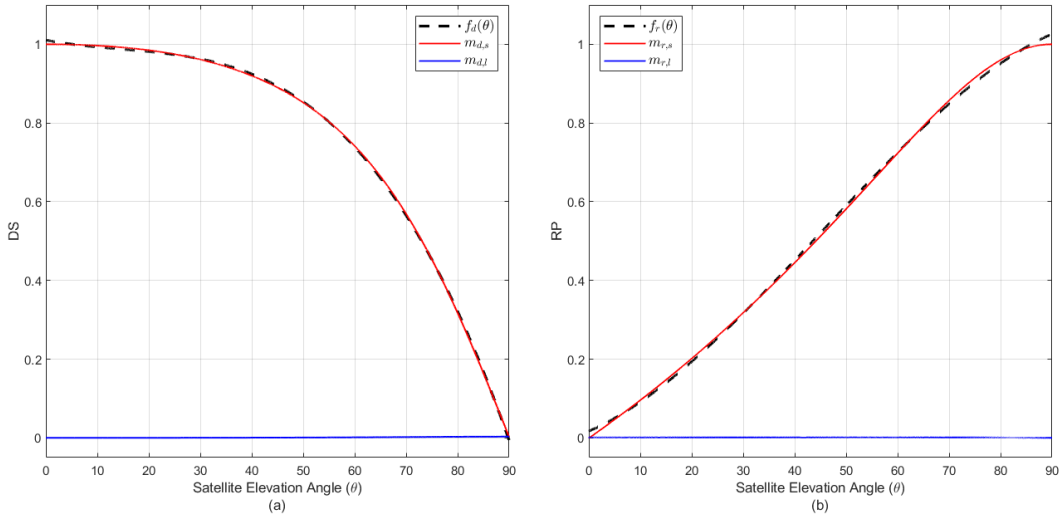

Figure 6.3: Normalized values at the GS versus the satellite elevation angle θ (a) $f_d(\theta)$, $m_{d,s,i}$ and $m_{d,l,i}$ and (b) $f_r(\theta)$, $m_{r,s,i}$ and $m_{r,l,i}$.

respectively. Due to approximation errors and receiver noise, the actual values for l and s can be expressed as

$$m_{r,l,i} = m_{n,r,l,i} + \varepsilon_{r_l}, \tag{6.11}$$

$$m_{r,s,i} = m_{n,r,s,i} + \varepsilon_{r_s}, \tag{6.12}$$

where ε_{r_l} and ε_{r_s} can be modeled as zero mean Gaussian random variables with variances $\sigma^2_{r_l}$ and $\sigma^2_{r_s}$ [106]. Curve fitting is employed to obtain the best fit for $m_{r,s,i}$ and this is used in the analysis. Figure 6.3b gives the values of $f_r(\theta)$, $m_{r,l,i}$ and $m_{r,s,i}$ versus the satellite elevation angle θ. The best fit third-degree polynomial for $m_{r,s,i}$ is

$$f_r(\theta) = b_1\theta^3 + b_2\theta^2 + b_3\theta + b_4, \tag{6.13}$$

where the coefficients b_1, b_2, b_3, and b_4 are -1.3×10^{-6}, 1.8×10^{-4}, 0.0057, and 0.017, respectively. The time step is 0.01 s and the values are normalized for comparison purposes. The corresponding RMSE is 0.0077. These results show that the values for s can be modeled as a the polynomial function while the values for l can be modeled

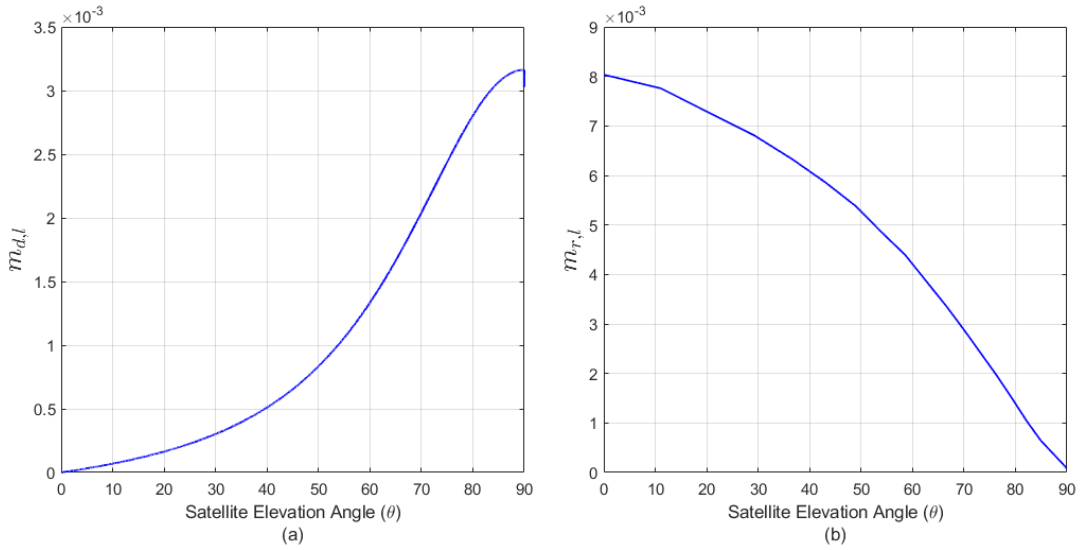

Figure 6.4: (a) $m_{d,l,i}$ and (b) $m_{r,s,i}$ versus θ.

as 0. Thus, $m_{r,l,i} \sim N(0, \sigma_{r_l}^2)$ and $m_{r,s,i} \sim N(f_r(\theta), \sigma_{r_s}^2)$. Figure 6.4 presents $m_{d,l,i}$ and $m_{r,s,i}$ versus θ. This shows that the values are very small for the legitimate satellite so they seem unchanged in Figure 6.3.

6.4.3 Authentication Test

Hypothesis testing using a threshold is employed to discriminate between l and s. The DS hypothesis test is

$$\mathcal{L}_d : \begin{cases} \mathcal{H}_0 : |m_{d,u,i+1} - m_{d,l,i}| \leq \tau, \\ \mathcal{H}_1 : |m_{d,u,i+1} - m_{d,l,i}| > \tau, \end{cases} \tag{6.14}$$

where τ is the threshold. The RP hypothesis test is

$$\mathcal{L}_r : \begin{cases} \mathcal{H}_0 : |m_{r,u,i+1} - m_{r,l,i}| \leq \tau, \\ \mathcal{H}_1 : |m_{r,u,i+1} - m_{r,l,i}| > \tau. \end{cases} \tag{6.15}$$

Figure 6.3a shows that at low θ, the normalized DS mean for s is a maximum and decreases with increasing θ to a value near 0, and it follows a $f_d(\theta)$ functions. Further, Figure 6.3b shows that the RP mean for s begins near 0 and increases with θ, and it follows a $f_r(\theta)$ function. Consequently, switching between DS and RP is employed in the combined authentication scheme so DS is utilized in the session from $\theta = 0^\circ$ to $\theta = 45^\circ$ and RP from $\theta = 45^\circ$ to $\theta = 90^\circ$. This ensures the that the mean for s is far from that for l for all values of θ. The combined hypothesis test for the DS and RP magnitudes is given by

$$\mathcal{L}_c : \begin{cases} \mathcal{H}_0 : |m_{d,u,i+1} - m_{d,l,i}| \leq \tau \ \text{ or} \\ \qquad |m_{r,u,i+1} - m_{r,l,i}| \leq \tau, \\ \mathcal{H}_1 : |m_{d,u,i+1} - m_{d,l,i}| > \tau \ \text{ or} \\ \qquad |m_{r,u,i+1} - m_{r,l,i}| > \tau. \end{cases} \tag{6.16}$$

6.4.4 Authentication Metrics

The false alarm rate P_f which denotes incorrectly rejecting l, the missed detection rate P_m which denotes incorrectly accepting s, and authentication rate AR are the metrics used to evaluate the authentication performance. P_f for the DS, RP, and combined hypothesis test is

$$\begin{aligned} P_{f,d}(\tau) =& P(|m_{d,u,i+1} - m_{d,l,i}| > \tau | \mathcal{H}_0) \\ =& 2\Phi\left(\frac{-\tau}{\sqrt{2}\sigma_l}\right), \end{aligned} \tag{6.17}$$

$$\begin{aligned} P_{f,r}(\tau) =& P(|m_{r,u,i+1} - m_{r,l,i}| > \tau | \mathcal{H}_0) \\ =& 2\Phi\left(\frac{-\tau}{\sqrt{2}\sigma_l}\right), \end{aligned} \tag{6.18}$$

$$P_{f,c}(\tau) = P_{f,d}(\tau) = P_{f,r}(\tau) = 2\Phi\left(\frac{-\tau}{\sqrt{2}\sigma_l}\right), \tag{6.19}$$

respectively, where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$ is the cumulative distribution function (CDF) of the normal distribution. P_m for the DS, RP, and combined hypothesis test is

$$\begin{aligned} P_{m,d}(\tau) =& P(|m_{d,u,i+1} - m_{d,l,i}| \leq \tau | \mathcal{H}_1) \\ =& \Phi\left(\frac{\tau - f_d(\theta)}{\sqrt{\sigma_l^2 + \sigma_s^2}}\right) - \Phi\left(\frac{-\tau - f_d(\theta)}{\sqrt{\sigma_l^2 + \sigma_s^2}}\right), \end{aligned} \tag{6.20}$$

$$\begin{aligned} P_{m,r}(\tau) =& P(|m_{r,u,i+1} - m_{r,l,i}| \leq \tau | \mathcal{H}_1) \\ =& \Phi\left(\frac{\tau - f_r(\theta)}{\sqrt{\sigma_l^2 + \sigma_s^2}}\right) - \Phi\left(\frac{-\tau - f_r(\theta)}{\sqrt{\sigma_l^2 + \sigma_s^2}}\right), \end{aligned} \tag{6.21}$$

$$P_{m,c}(\tau) = \begin{cases} \Phi\left(\frac{\tau - f_d(\theta)}{\sqrt{\sigma_l^2 + \sigma_s^2}}\right) - \Phi\left(\frac{-\tau - f_d(\theta)}{\sqrt{\sigma_l^2 + \sigma_s^2}}\right) \forall \theta \in [0°, 45°], \\ \Phi\left(\frac{\tau - f_r(\theta)}{\sqrt{\sigma_l^2 + \sigma_s^2}}\right) - \Phi\left(\frac{-\tau - f_r(\theta)}{\sqrt{\sigma_l^2 + \sigma_s^2}}\right) \forall \theta \in [45°, 90°], \end{cases} \tag{6.22}$$

respectively. Using these results, the AR for the DS, RP, and combined hypothesis test is defined as

$$AR_d(\tau) = \frac{1}{2} \times (1 - P_{f,d}(\tau)) \times (1 - P_{m,d}(\tau)), \tag{6.23}$$

$$AR_r(\tau) = \frac{1}{2} \times (1 - P_{f,r}(\tau)) \times (1 - P_{m,r}(\tau)), \tag{6.24}$$

$$AR_c(\tau) = \frac{1}{2} \times (1 - P_{f,c}(\tau)) \times (1 - P_{m,c}(\tau)), \tag{6.25}$$

respectively.

6.4.5 Game Theoretic Scheme

The interactions between the GS and s can be modeled as a PLA game in which the GS chooses the DS and RP detection thresholds $\tau \in [0, \infty]$ to maximize its utility while s chooses its attack probability $k \in [0, 1]$ to minimize this utility. A zero-sum game is considered such that the GS and s utilities satisfy $U_g(\tau, k) = -U_s(\tau, k)$.

The payoffs for the GS to accept l or reject s are P_{al} and P_{rs}, respectively, while the costs for the GS to reject l or accept s are C_{rl} and C_{as}, respectively The cost for s to launch an attack is C_s. The GS and s utilities are then formulated as [112, 54, 98]

$$\begin{aligned} U_g(\tau, k) &= -U_s(\tau, k) \\ &= (P_{al}(1 - P_f(\tau)) - C_{rl}P_f(\tau))(1 - k) \\ &+ (P_{rs}(1 - P_m(\tau)) - C_{as}P_m(\tau) + C_s)k. \end{aligned} \tag{6.26}$$

The Nash equilibrium (NE) in the proposed zero-sum game is defined as (τ^*, k^*) which indicates that neither the GS or s can increase their respective utilities by choosing any other strategy which can be expressed as [98]

$$U_g(\tau^*, k^*) \geq U_g(\tau, k^*), \forall\, 0 \leq \tau, \tag{6.27}$$

$$U_s(\tau^*, k^*) \geq U_s(\tau^*, k), \forall\, 0 \leq k \leq 1. \tag{6.28}$$

From (6.17), (6.18), and (6.19), as the threshold τ increases P_f decreases because the probability that l will pass the authentication hypothesis test increases [98], so that

$$P_f(\tau = 0) = 1, \tag{6.29}$$

$$\lim_{\tau -> +\infty} P_f(\tau) = 0, \tag{6.30}$$

$$\frac{\partial P_f(\tau)}{\partial \tau} = -\frac{e^{\left(-\frac{\tau^2}{4\sigma_l^2}\right)}}{\sqrt{\pi}\sigma_l} \leq 0. \tag{6.31}$$

From (6.20), (6.21), and (6.22), as the threshold increases P_m increases because the probability that s will pass the authentication hypothesis test decreases, so that

$$P_m(\tau = 0) = 0, \tag{6.32}$$

$$\lim_{\tau->+\infty} P_m(\tau) = 1, \tag{6.33}$$

$$\frac{\partial P_m(\tau)}{\partial \tau} = \frac{e^{\left(-\frac{(\tau-\eta)^2}{2(\sigma_l^2+\sigma_s^2)}\right)} + e^{\left(-\frac{(-\tau-\eta)^2}{2(\sigma_l^2+\sigma_s^2)}\right)}}{\sqrt{2\pi(\sigma_l^2+\sigma_s^2)}} \geq 0. \tag{6.34}$$

where η is the mean which is $f_d(\theta)$ or $f_r(\theta)$.

First, to obtain τ^* take the derivative of $U_s(\tau, k)$ with respect to k which gives [98]

$$\begin{aligned}\frac{\partial U_s(\tau,k)}{\partial k} =& P_{al} - P_{rs} - C_s - (P_{al} + C_{rl})P_f(\tau) + \\ & (P_{rs} + C_{as})P_m(\tau).\end{aligned} \tag{6.35}$$

Substituting (6.29), (6.30), (6.32), and (6.33) in (6.35) gives

$$\frac{\partial U_s(\tau=0,k)}{\partial k} = -C_{rl} - P_{rs} - C_s < 0, \tag{6.36}$$

$$\frac{\partial U_s(\tau=+\infty,k)}{\partial k} = P_{al} + C_{as} - C_s > 0. \tag{6.37}$$

This shows that there exists a unique positive threshold τ^* that satisfies $\frac{\partial U_s(\tau^*,k)}{\partial k} = 0$ because $\frac{\partial U_s(\tau=0,k)}{\partial k} < 0$ and $\frac{\partial U_s(\tau=+\infty,k)}{\partial k} > 0$. From (6.37), a unique τ^* exists if the following condition is achieved

$$P_{al} + C_{as} > C_s. \tag{6.38}$$

Since $\frac{\partial U_s(\tau^*,k)}{\partial k} = 0$, $U_s(\tau^*, k)$ is a constant and (6.28) is satisfied for any $k^* \in [0, 1]$. Then τ^* can be expressed as

$$\tau^* = \{\tau | \frac{\partial U_s(\tau,k)}{\partial k} = 0\}, \tag{6.39}$$

where τ^* is independent of k.

Next, to obtain k^* take the derivative of $U_g(\tau, k)$ with respect to τ which gives

$$\begin{aligned}\frac{\partial U_g(\tau,k)}{\partial \tau} =& -(1-k)(P_{al}+C_{rl})\frac{\partial P_f(\tau)}{\partial \tau} \\ & -k(P_{rs}+C_{as})\frac{\partial P_m(\tau)}{\partial \tau}\end{aligned}. \tag{6.40}$$

Then k^* is the solution of $\frac{\partial U_g(\tau,k^*)}{\partial \tau}=0$ from (6.31) and (6.34) given by [98]

$$k^* = \frac{P_{al}+C_{rl}}{P_{al}+C_{rl}+\frac{(P_{rs}+C_{as})h(\tau^*)\sigma_l}{\sqrt{2(\sigma_l^2+\sigma_s^2)}}}, \tag{6.41}$$

where

$$\begin{aligned}h(\tau) =& \exp\left(\frac{(\tau)^2(\sigma_s^2-\sigma_l^2)+4\tau\eta\sigma_l^2-2\,\eta^2\sigma_l^2}{4\sigma_l^2(\sigma_l^2+\sigma_s^2)}\right) \\ &+\exp\left(\frac{(\tau)^2(\sigma_s^2-\sigma_l^2)-4\tau\eta\sigma_l^2-2\,\eta^2\sigma_l^2}{4\sigma_l^2(\sigma_l^2+\sigma_s^2)}\right),\end{aligned} \tag{6.42}$$

and k^* is dependant on τ^*.

6.5 Performance Evaluation

In this section, we evaluate the proposed game theoretic PLA approach for LEO satellites. Average results are given for $0^\circ \leq \theta \leq 90^\circ$. First, we evaluate the performance without game theory and determine the average P_f, P_m, and AR of the proposed approach versus τ. Then the interaction between l and s is evaluated via the zero-sum game. The average values of $\frac{\partial U_s(\tau,k)}{\partial k}$, τ^*, and k^* are determined.

6.5.1 LEO Satellite PLA Without Game Theory

Equations (6.17), (6.18), and (6.19) indicate that the expressions for $P_{f,d}(\tau)$, $P_{f,r}(\tau)$, and $P_{f,c}(\tau)$ are similar, while (6.20), (6.21), and (6.22) show that the expressions for $P_{m,d}(\tau)$, $P_{m,r}(\tau)$, and $P_{m,c}(\tau)$ differ based on whether the mean is $f_d(\theta)$ or $f_r(\theta)$. Figures 6.5a and 6.5b presents P_f and P_m, respectively, versus τ with $\sigma_l = 0.5$ and

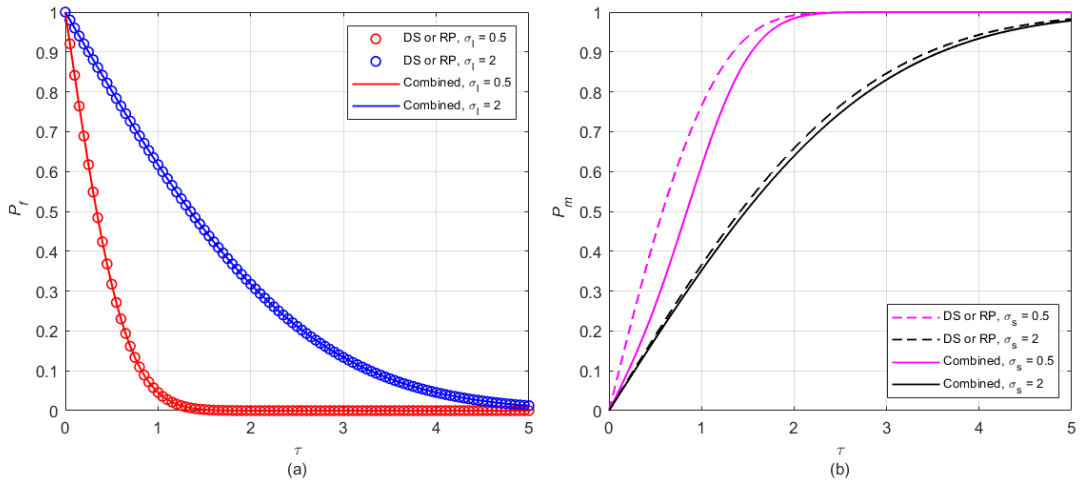

Figure 6.5: Average error rate versus τ with $\sigma_l = 0.5$ and 2, and $\sigma_s = 0.5$ and 2 for the DS, RP, and combined authentication schemes (a) P_f and (b) P_m.

2, and $\sigma_s = 0.5$ and 2 for the DS, RP, and combined authentication schemes. Figure 6.5a shows that P_f for the DS, RP, and combined authentication schemes is the same for all values of τ. For example, P_f is 61.7% with $\sigma_l = 2$ and $\tau = 1$. Moreover, P_f decreases with τ and is near 0 for $\tau > 1$ and $\sigma_l = 0.5$. P_f for $\sigma_l = 2$ is 80.3% at $\tau = 0.5$ and 31.7% at $\tau = 2$. P_f for the DS, RP, and combined authentication schemes increases with σ_l. For example, P_f for $\tau = 0.5$ is 31.7% with $\sigma_l = 0.5$ and 80.3% with $\sigma_l = 2$.

Figure 6.5b shows that P_m for the DS and RP authentication schemes is the same and is higher than P_m for the combined authentication scheme at all values of τ. For example, P_m for the DS and RP authentication schemes is 76.4% with $\sigma_l = 0.5$ at $\tau = 1$ and 61.3% for the combined authentication scheme. Moreover, P_m increases with τ. For example, P_m for the DS and RP authentication schemes with $\sigma_l = 0.5$ is 43.9% at $\tau = 0.5$ and 26.4% for the combined authentication scheme, while at $\tau = 1.5$ it is 93.9% and 88.7%, respectively. Furthermore, P_m decreases with σ_s. For example, P_m for the DS and RP authentication schemes at $\tau = 2$ is 65.8% with $\sigma_l = 0.5$ and 63.9% for the combined authentication scheme, while for $\sigma_l = 2$ it is 99.2% and 98.4, respectively.

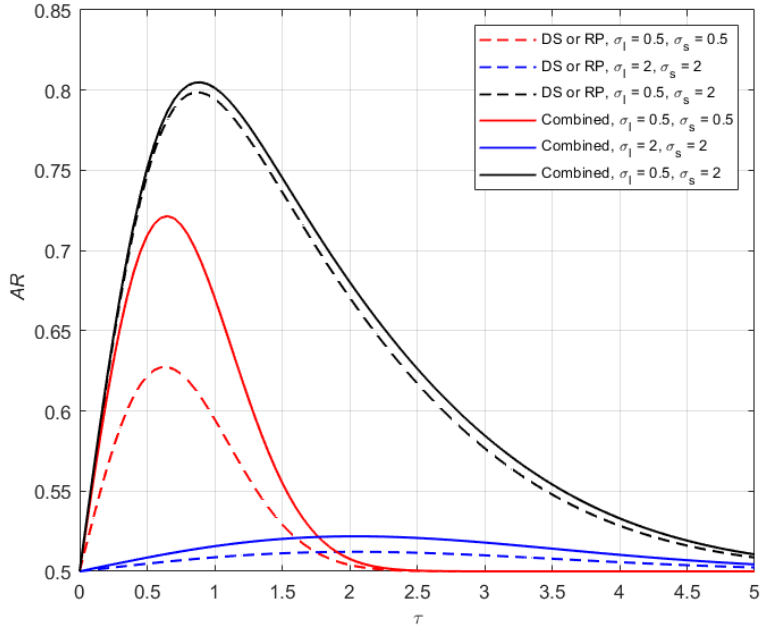

Figure 6.6: Average AR versus τ with $\sigma_l = 0.5$ and 2, and $\sigma_s = 0.5$ and 2 for the DS, RP, and combined authentication schemes.

Figure 6.6 presents the AR versus τ with $\sigma_l = 0.5$ and 2, and $\sigma_s = 0.5$ and 2 for the DS, RP, and combined authentication schemes. This shows that the AR for the combined authentication scheme outperforms the AR for the DS and RP authentication schemes at all values of τ and the AR for the DS and RP authentication schemes is the same. For example, the AR for the DS and RP authentication schemes at $\tau = 1$ is 50.9%, 59.4%, and 79.5% with $\sigma_l = 2$ and $\sigma_s = 2$, $\sigma_l = 0.5$ and $\sigma_s = 0.5$, and $\sigma_l = 0.5$ and $\sigma_s = 2$, respectively, while the corresponding AR for the combined authentication scheme is 51.6%, 67.1%, and 80.1%, respectively. Moreover, the AR for the DS, RP, and combined authentication schemes initially increases with τ and after reaching a maximum decreases to a value close to 50.0%. For example, the AR for the combined authentication scheme with $\sigma_l = 0.5$ and $\sigma_s = 0.5$ is 50.0%, 72.2%, and 50.0% at $\tau = 0, 0.65$, and 3, respectively. Furthermore, the AR for the DS, RP, and combined authentication schemes with $\sigma_l < \sigma_s$ is higher. This is a practical situation as s must estimate the DS and RP values between l and the GS

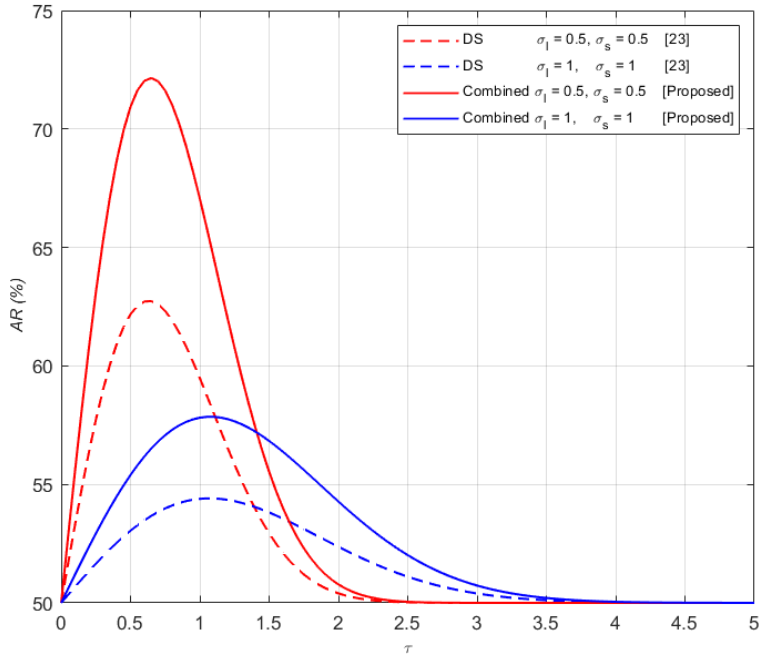

Figure 6.7: Average AR versus τ with $\sigma_l = 0.5$ and 1, and $\sigma_s = 0.5$ and 1 for the DS [3] and combined DS and RP authentication schemes.

to launch an attack and there will be estimation errors. For example, the maximum AR for the DS and RP authentication schemes with $\sigma_l = 0.5$ and $\sigma_s = 0.5$, $\sigma_l = 2$ and $\sigma_s = 2$, and $\sigma_l = 0.5$ and $\sigma_s = 2$ is 62.7%, 52.2%, and 79.9%, respectively, and the corresponding maximum AR for the combined authentication scheme is 72.2%, 51.2%, and 80.9%, respectively.

Figure 6.7 presents the average AR versus τ with $\sigma_l = 0.5$ and 1, and $\sigma_s = 0.5$ and 1 for the DS [3] and combined DS and RP authentication schemes. These results show that the combined DS and RP authentication scheme outperforms the scheme in [3] that only employs DS attributes. For example, the AR at $\tau = 1$ with $\sigma_l = 0.5$ and $\sigma_l = 0.5$ for the approach in [3] and the proposed combined scheme is 54.5% and 67.1%, respectively.

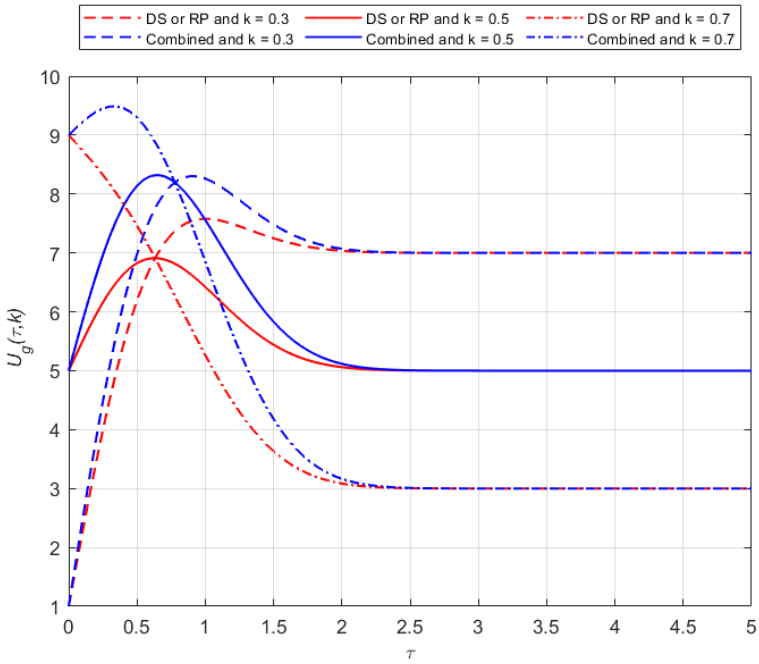

Figure 6.8: $U_g(\tau, k)$ versus τ with $\sigma_l = 0.5$, $\sigma_s = 0.5$, $P_{al} = 10$, $P_{rs} = 10$, $C_{rl} = 5$, $C_{as} = 5$, $C_s = 5$, and $k = 0.3, 0.5$ and 0.7 for the DS, RP, and combined authentication schemes.

6.5.2 LEO Satellite PLA With Game Theory

In this subsection the interaction between s and GS using game theory is evaluated. A worst-case scenario is considered, where s is very close to l with a maximum distance between them of 5 km along the trajectory so they are both within the GS receive antenna HPBW. Figure 6.8 presents $U_g(\tau, k)$ versus τ with $\sigma_l = 0.5$, $\sigma_s = 0.5$, $P_{al} = 10$, $P_{rs} = 10$, $C_{rl} = 5$, $C_{as} = 5$, and $C_s = 5$ for the DS, RP, and combined authentication schemes. This shows that $U_g(\tau, k)$ for the combined authentication scheme is greater than for the DS and RP authentication schemes for all values of k and τ, and $U_g(\tau, k)$ for the DS and RP authentication schemes is the same. For example, $U_g(\tau, k)$ for the DS and RP authentication schemes with $k = 0.3$ at $\tau = 1$ is 7.58 and the corresponding value for the combined authentication scheme is 8.26. Moreover, for small τ $U_g(\tau, k)$ for the DS, RP, and combined authentication

schemes is highest with $k = 0.7$ followed by $k = 0.5$, and is lowest with $k = 0.3$. For example, the $U_g(\tau, k)$ for the combined authentication scheme with $k = 0.7$, $k = 0.5$, and $k = 0.3$ at $\tau = 0$ is 9, 5, and 1, respectively. However, $U_g(\tau, k)$ for the DS, RP, and combined authentication schemes with $k = 0.7$ decreases with τ and converges to 3.00, while with $k = 0.5$ it converges to 5.00 and with $k = 0.3$ it converges to 7.00.

Figure 6.9 presents $U_g(\tau, k)$ versus τ with $k = 0.5$, $\sigma_l = 0.5$, $\sigma_s = 0.5$, $P_{al} = 10$, $P_{rs} = 10$, $C_{rl} = 5$ for different values of C_s for the DS, RP, and combined authentication schemes. This shows that $U_g(\tau, k)$ for the combined authentication scheme is greater than for the DS and RP authentication schemes for all values of C_s and τ. For example, $U_g(\tau, k)$ for the DS and RP authentication schemes at $\tau = 1$ with $C_s = 4$, $C_s = 8$, and $C_s = 12$ is 5.92, 7.92, and 9.92, respectively, and the corresponding values for the combined authentication scheme are 7.06, 9.06, and 11.06, respectively. Moreover, $U_g(\tau, k)$ for all three schemes increases with C_s for all values of τ. For example, $U_g(\tau, k)$ for the DS and RP authentication schemes at $\tau = 0.5$ with $C_s = 4$, $C_s = 8$, and $C_s = 12$ is 6.33, 8.33, and 10.33, respectively, and the corresponding values for the combined authentication scheme are 7.64, 9.64, and 11.64, respectively.

Figure 6.10 presents $\frac{\partial U_s(\tau,k)}{\partial k}$ versus τ with $\sigma_l = 1$, and $C_s = 5$ for the DS, RP, and combined authentication schemes. This shows that $\frac{\partial U_s(\tau,k)}{\partial k}$ increases with τ. For example, $\frac{\partial U_s(\tau,k)}{\partial k}$ for the DS or RP authentication schemes with $\sigma_s = 1$ and $P_{al} = C_{as} = 8$ at $\tau = 2$, $\tau = 3$, and $\tau = 4$ is 4.26, 8.20, and 9.97, respectively. Furthermore, $\frac{\partial U_s(\tau,k)}{\partial k}$ decreases with σ_s for all values of τ. For example, $\frac{\partial U_s(\tau,k)}{\partial k}$ for the DS or RP authentication schemes with $\sigma_s = 1$ and $P_{al} = C_{as} = 8$ at $\tau = 3.5$ is 9.27 while with $\sigma_s = 2$ and $P_{al} = C_{as} = 8$ at $\tau = 3.5$ it is 2.28. Moreover, $\frac{\partial U_s(\tau,k)}{\partial k}$ with $P_{al} > C_{as}$ is the highest followed by $P_{al} < C_{as}$ and is lowest with $P_{al} = C_{as}$ for all values of σ_s and τ. For example, $\frac{\partial U_s(\tau,k)}{\partial k}$ for the DS or RP authentication schemes with $\sigma_s 1$ at $\tau = 2.5$ for $P_{al} > C_{as}$, $P_{al} < C_{as}$, and $P_{al} = C_{as}$ is 8.59, 8.15, and 6.62,

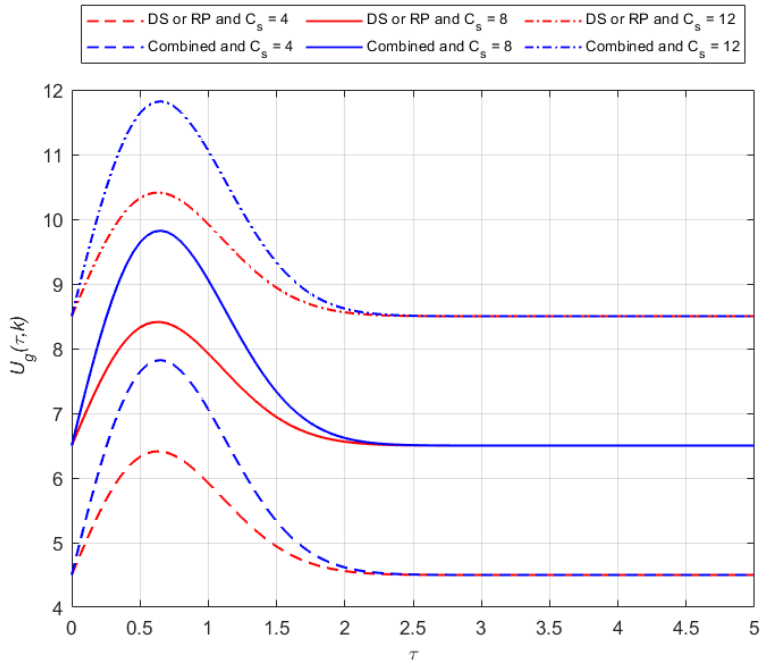

Figure 6.9: $U_g(\tau, k)$ versus τ with $k = 0.5$, $\sigma_l = 0.5$, $\sigma_s = 0.5$, $P_{al} = 10$, $P_{rs} = 10$, $C_{rl} = 5$, and $C_s = 4, 8$, and 12 for the DS, RP, and combined authentication schemes.

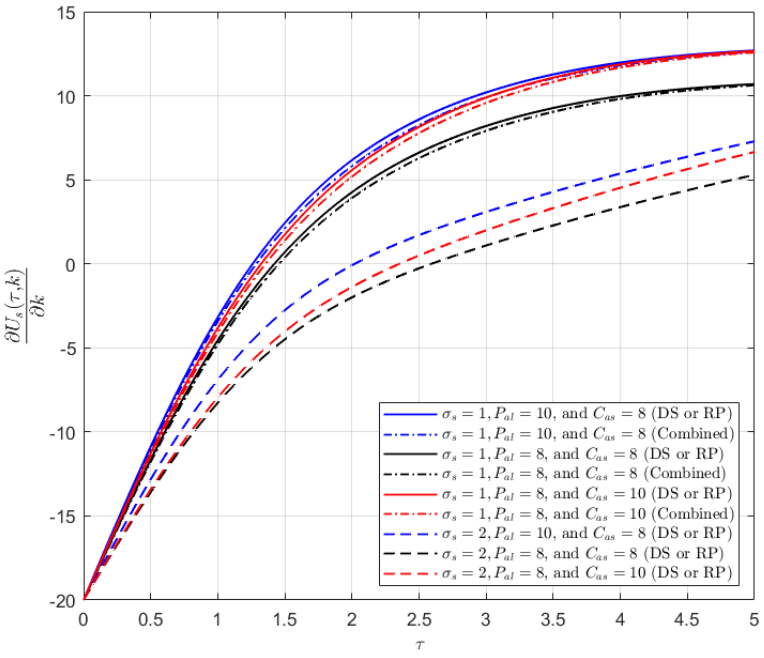

Figure 6.10: $\frac{\partial U_s(\tau,k)}{\partial k}$ versus τ with $\sigma_l = 1$, $C_s = 5$, and $C_{as} = 8$ and 10 for the DS, RP, and combined authentication schemes.

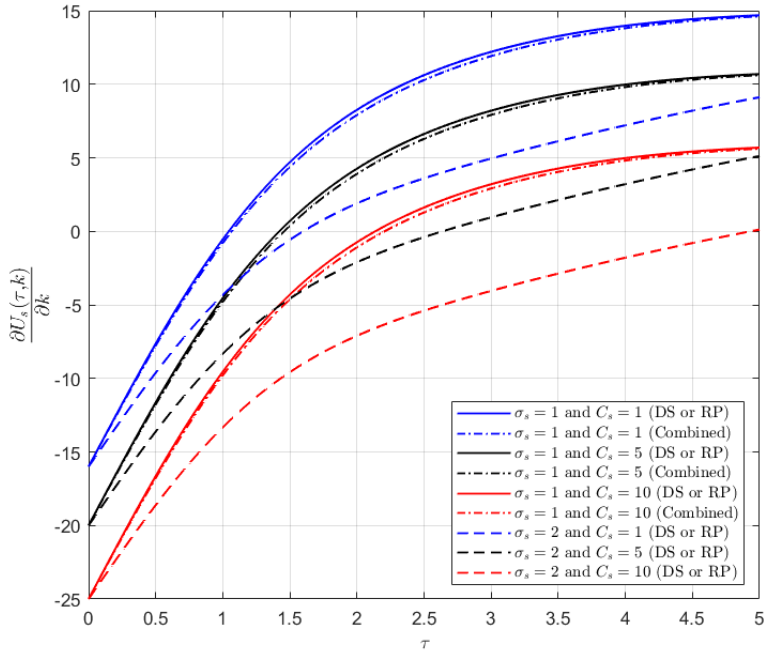

Figure 6.11: $\frac{\partial U_s(\tau,k)}{\partial k}$ versus τ with $\sigma_l = 1$, $C_s = 1, 5$ and 10, and $P_{al} = C_{as} = 8$ for the DS, RP, and combined authentication schemes.

respectively. Moreover, the DS and RP authentication schemes provide slightly higher $\frac{\partial U_s(\tau,k)}{\partial k}$ compared to the combined authentication scheme for all values of τ.

Figure 6.11 presents $\frac{\partial U_s(\tau,k)}{\partial k}$ versus τ with $\sigma_l = 1$, different values of C_s, and $P_{al} = C_{as} = 8$ for the DS, RP, and combined authentication schemes. This shows that $\frac{\partial U_s(\tau,k)}{\partial k}$ increases with τ. For example, $\frac{\partial U_s(\tau,k)}{\partial k}$ for the DS or RP authentication schemes with $\sigma_s = 1$ and $C_s = 1$ at $\tau = 2$, $\tau = 2.5$, and $\tau = 3$ is 8.26, 10.62, and 12.20, respectively. Furthermore, $\frac{\partial U_s(\tau,k)}{\partial k}$ decreases with σ_s for all values of τ. For example, $\frac{\partial U_s(\tau,k)}{\partial k}$ for the DS or RP authentication schemes at $\tau = 3$ with $C_s = 5$ for $\sigma_s = 1$ and $\sigma_s = 2$ is 8.20 and 0.95, respectively. Moreover, $\frac{\partial U_s(\tau,k)}{\partial k}$ decreases with C_s. For example, $\frac{\partial U_s(\tau,k)}{\partial k}$ for the DS or RP authentication schemes with $\sigma_s 1$ at $\tau = 4$ for $C_s = 1$, $C_s = 5$, and $C_s = 10$ is 13.97, 9.97, and 4.97, respectively. The DS and RP authentication schemes provide slightly higher $\frac{\partial U_s(\tau,k)}{\partial k}$ compared to the combined authentication scheme for all values of τ.

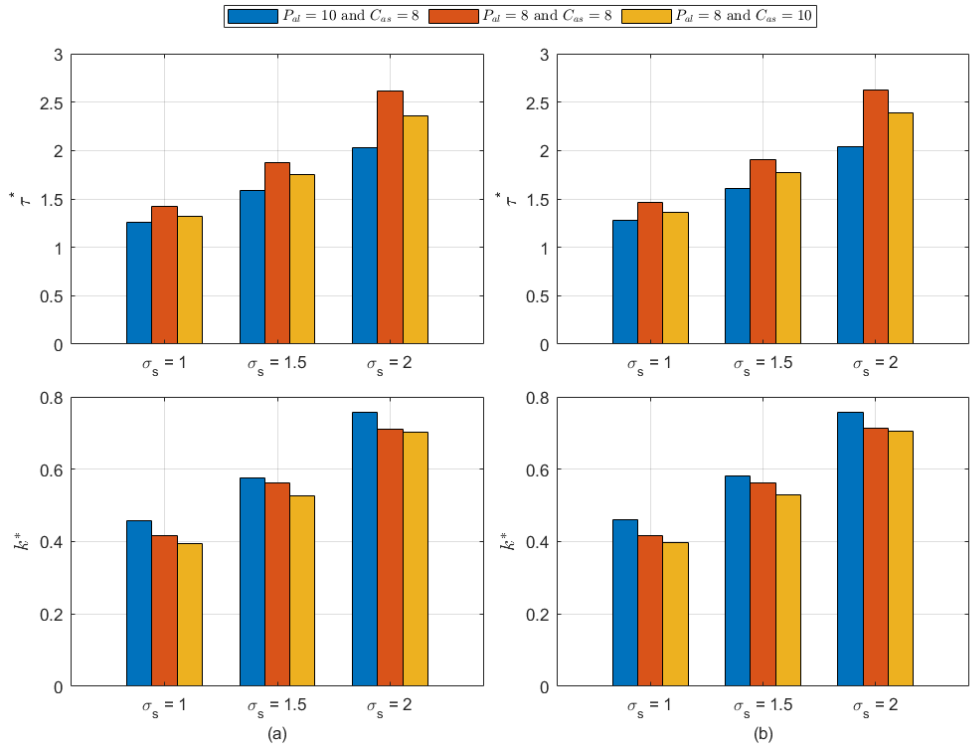

Figure 6.12: τ^* and k^* versus σ_s with $C_s = 5$ and $\sigma_a = 1$ for the (a) DS or RP authentication schemes, and (b) combined authentication scheme.

Figures 6.12a and 6.12b present τ^* and k^* versus σ_s with $C_s = 5$ for the DS or RP authentication schemes and combined authentication scheme, respectively. Figure 6.12a shows that τ^* increases with σ_s for all values of P_{al} and C_{as}. For example, τ^* for the DS or RP authentication schemes with $P_{al} = 10$ and $C_{as} = 8$ for $\sigma_s = 1$, $\sigma_s = 1.5$, and $\sigma_s = 2$ is 1.26, 1.59, and 2.03, respectively. Furthermore, τ^* with $P_{al} > C_{as}$ is the lowest followed by τ^* with $P_{al} < C_{as}$ and is highest with $P_{al} = C_{as}$ for all values of σ_s. For example, τ^* for the DS or RP authentication schemes with $\sigma_s = 2$ for $P_{al} > C_{as}$, $P_{al} < C_{as}$, and $P_{al} = C_{as}$ is 2.03, 2.36, and 2.61, respectively. Figures 6.12b shows that k^* increases with σ_s for all values of P_{al} and C_{as}. For example, k^* for the combined authentication scheme with $P_{al} = 8$ and $C_{as} = 8$ for $\sigma_s = 1$, $\sigma_s = 1.5$, and $\sigma_s = 2$ is 0.42, 0.56, and 0.71, respectively. Furthermore, k^* with $P_{al} < C_{as}$ is the lowest followed by k^* with $P_{al} = C_{as}$ and

is highest with $P_{al} > C_{as}$ for all values of σ_s. For example, k^* for the combined authentication scheme with $\sigma_s = 1.5$ for $P_{al} < C_{as}$, $P_{al} = C_{as}$, and $P_{al} > C_{as}$ is 0.58, 0.56, and 0.53, respectively. Finally, τ^* and k^* for the combined authentication scheme is slightly higher than τ^* and k^* for the DS or RP authentication schemes for all values of σ_s, P_{al}, and C_{as}. For example, τ^* for the DS and RP, and combined authentication schemes, with $\sigma_s = 1$ for $P_{al} > C_{as}$ is 1.26 and 1.36, respectively and the corresponding values of k^* are 0.45 and 0.46, respectively.

6.6 Conclusion

The huge increase in the number of low earth orbit (LEO) satellites in space information networks (SINs) makes these networks more vulnerable to spoofing attack which is considered as serious threat. Hence, a game theoretic spoofing detection scheme for LEO satellites using physical layer attributes was proposed. Doppler frequency spread (DS) and received power (RP) attributes are employed as they are considered unique based on the satellite trajectories. A zero-sum PLA game was formulated where the ground station (GS) chooses its optimal detection threshold (τ^*) to maximize its utility, while the spoofer satellite (s) chooses its optimal attack probability (k^*) to minimize this utility. Results were presented which show that using a combination of the DS and RP provides the highest authentication performance. Moreover, the optimal threshold τ^* and optimal attack probability k^* increase with σ_s. The evaluation of the proposed approach was specifically focused on LEO satellites. However, it should be noted that our model may be inadequate for geosynchronous equatorial orbit (GEO) satellites due to the absence of the DS attribute in these orbits. As a result, further research is warranted to incorporate additional attributes and generalize our approach to encompass LEO, medium Earth orbit (MEO), and GEO satellites. This extension would enhance the applicability and robustness of the proposed model across different satellite orbits.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

The open nature and dynamic characteristics of heterogeneous networks makes IoT devices and satellite communication systems are susceptible to spoofing attacks, which raises privacy and security concerns. Thus, wireless communications authentication has gained significant attention [30, 31]. Implementing authentication for access control presents an effective approach to ensure the security of data. However, traditional ULA techniques have drawbacks such as increased delay, computational complexity, and bandwidth consumption due to protocols involving key generation, distribution, and updates [113]. To address these concerns, PLA has been introduced to complement ULA [114]. PLA in conjunction with ULA has been shown to provide reliable and robust authentication for wireless communications [20]. Despite the existence of various PLA solutions, authentication is still challenging and complex [27, 13, 28]. Thus, simple and effective PLA solutions are required to improve network security [78].

In Chapter 2, an adaptive PLA scheme that leverages ML was proposed. This scheme exploits antenna diversity to improve the AR. OCC-SVM with linear, sigmoid, and polynomial kernels was used. The magnitude and real and imaginary

parts of the received SRS at each receive antenna are used as features. Thus, the number of features is $3N$ where N is the number of receive antennas. Results were presented which show that the AR increases with the number of antennas. Further, higher velocities increase the DS which improves differentiation between users and thus performance. Moreover, diversity combining was shown to degrade the performance as it makes it more difficult to discriminate between users. This confirms the benefits of using antenna diversity.

In Chapter 3, an adaptive PLA scheme using ML was proposed for IoT applications in MIMO communication systems. This scheme exploits the antenna diversity in MIMO systems based on the 5G frame structure. A sliding window is used to update features to provide robust authentication. The magnitude and real and imaginary parts of the received signals corresponding to the 5G SRSs are used as features. Thus, the number of features is $3MN$ for each transmitted symbol where M and N are the number of transmit and receive antennas, respectively. Results were presented for an urban environment which show that the performance improves with the number of antennas. Two majority vote schemes were also presented. These schemes may be preferable for distributed IoT devices with limited computing capabilities. The AR using all $3MN$ features at one location was shown to be slightly higher than using $3M$ features separately at each device and N voting.

In Chapter 4, an adaptive PLA scheme using ML was proposed for LEO satellites using the DS and RP as features. An OCC-SVM with linear and polynomial kernels was employed. Results were presented which show that a high authentication rate (AR) can be achieved for both fixed and mobile satellite services using the DS and RP as features. In particular, the AR using both DS and RP exceeds 99.6% for fixed satellite services (FSS) and mobile satellite services (MSS) scenarios, and is superior to using only DS or RP as features as in [20] and [98], respectively.

In Chapter 5, a PLA scheme was proposed for LEO satellites using DS and RP characteristics. This scheme employs hypothesis testing using a threshold or ML to

discriminate between legitimate and illegitimate satellites. Estimation errors in the DS and RP values were considered and the performance was evaluated based on real satellite data from the system tool kit (STK). Results were presented which show that DS provides a high AR at small elevation angles (θ) that decreases with θ, while RP provides a low AR at small θ that increases with θ. Further, ML authentication with a small percentage of outliers η in the training data provides the highest AR. Finally, the AR for the ML authentication scheme increases with the amount of training data ℓ.

In Chapter 6, a game theoretic spoofing detection scheme for LEO satellites using physical layer attributes was proposed. DS and RP attributes are employed as they are considered unique based on the satellite trajectories. A zero-sum PLA game was formulated where the ground station (GS) chooses the optimal detection threshold (τ^*) to maximize its utility, while the spoofer satellite (s) chooses the optimal attack probability (k^*) to minimize this utility. Results were presented which show that using both DS and RP provides the highest authentication performance. Moreover, the optimal threshold τ^* and optimal attack probability k^* increase with σ_s.

7.2 Future Work

Some directions for future work are given in this section.

7.2.1 PLA with a Reconfigurable Intelligent Surface (RIS)

A reconfigurable intelligent surface (RIS) is a programmable two-dimensional structure composed of numerous elements that can be configured to control the propagation of electromagnetic waves. An RIS can provide multiple wireless signal paths for users. ML algorithms can be trained using these signals which are difficult for an attacker to mimic. Furthermore, RIS elements can be designed to have a high degree of randomness to improve the robustness and reliability of PLA.

7.2.2 PLA for Optical Communications

Optical communication networks are used in many applications such as data centers, high-performance computing, and telecommunications. Quantum-based techniques can be used to authenticate optical signals. Furthermore, ML techniques can be trained using optical signal features. Interference-based techniques can be used to detect and authenticate optical signals using the interference patterns generated when signals interact with the environment to improve the security of optical links.

7.2.3 PLA for Underwater Communications

The proposed approaches to authenticate can be extended to underwater communications. Underwater communications employ optical and acoustics channels. Underwater PLA (UPLA) can be used to ensure secure communications in underwater environments. Acoustic-based PLA (APLA) relies on the unique properties of sound waves in water such as the time delay and phase shift. For example, an APLA scheme can use the time delay of a signal reflected off an underwater object for authentication APLA can be useful in underwater environments with low visibility where optical signals may not be reliable. Optical-based PLA (OPLA) uses the properties of light waves in water such as polarization and absorption for authentication. For example, an OPLA scheme can use the polarization of a light signal reflected off an underwater object for authentication. OPLA can be useful in underwater environments with high visibility where acoustic signals may be affected by ambient noise or reverberation. A hybrid approach that combines both acoustic and optical signals may provide better authentication. For example, the time delay and phase shift of an acoustic signal and the polarization of an optical signal can be used to provide authentication in environments with variable visibility and ambient noise levels. Thus, APLA and OPLA can be used separately or in combination to provide authentication depending on the environment.

Appendix A

One-Class Classification Support Vector Machine (OCC-SVM)

One-class classification (OCC) is a ML technique that can be used to solve problems such as authentication. OCC is used here to distinguish between illegitimate user features and legitimate user features using training data from only the legitimate user. The proposed authentication framework employs the OCC-SVM [115] algorithm which is an extension of the TCC-SVM algorithm [116]. The goal with OCC-SVM is to find the optimal authentication boundary that surrounds most of the training data from the legitimate user [1]. The method in [115] is used to solve the OCC problem using SVM. OCC-SVM computes a decision function f, which encloses most of the training data [1]. A test sample $\mathbf{t}$ is accepted if $f(\mathbf{t}) > 0$ which indicates it is within the authentication boundary.

The following optimization problem is first solved [59, 115]

$$\begin{aligned} \min_{\mathbf{w},\mathbf{s},\rho} \quad & \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{\eta\ell}\sum_{i=1}^{\ell} s_i - \rho, \\ \text{subject to} \quad & (\mathbf{w}\cdot\Phi(\mathbf{g}_i)) \geq \rho - s_i, \quad s_i \geq 0 \end{aligned} \tag{A.1}$$

where $\mathbf{w}$ is the weight vector, ρ is the distance from the origin to the boundary, ℓ is the number of training samples, Φ is the feature mapping, $\mathbf{g}_i$ is the ith feature vector, s_i is the corresponding slack variable, and η is the percentage of data considered as outliers [1]. Using Lagrange multipliers $p_i, q_i \geq 0$ to solve (A.1) gives [115]

$$L(\mathbf{w}, \mathbf{s}, \mathbf{p}, \mathbf{q}, \rho) = \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{\eta\ell}\sum_{i=1}^{\ell} s_i - \rho - \sum_{i=1}^{\ell} p_i((\mathbf{w} \cdot \Phi(\mathbf{g}_i)) - \rho + s_i) - \sum_{i=1}^{\ell} q_i s_i. \tag{A.2}$$

Setting the derivatives with respect to $\mathbf{w}$, $\mathbf{s}$ and ρ equal to zero results in [115]

$$p_i = \frac{1}{\eta\ell} - q_i \leq \frac{1}{\eta\ell}, \tag{A.3}$$

$$\sum_{i=1}^{\ell} p_i = 1, \tag{A.4}$$

$$\mathbf{w} = \sum_{i=1}^{\ell} p_i \Phi(\mathbf{g}_i). \tag{A.5}$$

The decision function used to test a new vector $\mathbf{t}$ is [1, 59]

$$f(\mathbf{t}) = \operatorname{sgn}\left((\mathbf{w} \cdot \Phi(\mathbf{t})) - \rho\right), \tag{A.6}$$

and substituting $\mathbf{w}$ from (A.5) gives

$$f(\mathbf{t}) = \operatorname{sgn}\left(\sum_i (p_i \Phi(\mathbf{g}_i) \cdot \Phi(\mathbf{t})) - \rho\right). \tag{A.7}$$

The kernel expansion is defined as [115]

$$k\left(\mathbf{g}_i, \mathbf{t}\right) = \Phi(\mathbf{g}_i) \cdot \Phi(\mathbf{t}) \tag{A.8}$$

so the decision function is

$$f(\mathbf{t}) = \operatorname{sgn}\left(\sum_i p_i k\left(\mathbf{g}_i, \mathbf{t}\right) - \rho\right). \tag{A.9}$$

The test is passed if $f(\mathbf{t}) > 0$ and fails otherwise. Linear, sigmoid, and polynomial kernels can be considered. The linear kernel is simply

$$k\left(\mathbf{g}_i, \mathbf{t}\right) = \mathbf{g}_i \cdot \mathbf{t}. \tag{A.10}$$

The sigmoid kernel is

$$k\left(\mathbf{g}_i, \mathbf{t}\right) = \tanh(\lambda \mathbf{g}_i \cdot \mathbf{t} + c), \tag{A.11}$$

where λ is the slope and c is constant, and the polynomial kernel is [117]

$$k\left(\mathbf{g}_i, \mathbf{t}\right) = (\mathbf{g}_i \cdot \mathbf{t} + r)^d, d > 1, \tag{A.12}$$

where d and r are the degree and coefficient of the polynomial, respectively.

Appendix B

Evaluation Metrics

The confusion matrix shown in Figure B.1 is used to evaluate the performance. True positive (TP) denotes correctly accepting a legitimate user

$$f(\mathbf{t}|A) > 0,$$

Confusion Matrix	Predict Negative	Predict Positive
Actual Negative (N)	TN $\mathcal{H}_1$	FP Type II error
Actual Positive (P)	FN Type I error	TP $\mathcal{H}_0$

Figure B.1: Confusion matrix.

true negative (TN) denotes correctly rejecting an illegitimate user

$$f(\mathbf{t}|E) \leq 0,$$

false negative (FN) denotes incorrectly rejecting a legitimate user

$$f(\mathbf{t}|A) \leq 0,$$

and false positive (FP) denotes incorrectly accepting an illegitimate user

$$f(\mathbf{t}|E) > 0.$$

The goal of PLA is to make the number of FN and FP low.

The metrics used for performance evaluation are missed detection rate (MDR), false alarm rate (FAR), and authentication rate (AR) which are given by

$$\text{MDR} = \frac{FP}{N}, \tag{B.1}$$

$$\text{FAR} = \frac{FN}{P}, \tag{B.2}$$

$$\text{AR} = \frac{TP + \gamma \times TN}{P + \gamma \times N}, \tag{B.3}$$

respectively, where TP, TN, FN, and FP are the number of TP, TN, FN, and FP, respectively, $N = TN + FP$, $P = TP + FN$ and $\gamma = \frac{P}{N}$ is used to balance between legitimate and illegitimate users.

Bibliography

[1] L. Senigagliesi, M. Baldi, and E. Gambi, "Comparison of statistical and machine learning techniques for physical layer authentication," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1506–1521, 2020.

[2] C. Pei, N. Zhang, X. S. Shen, and J. W. Mark, "Channel-based physical layer authentication," in *IEEE Global Communications Conference*, Austin, TX, USA, 2014, pp. 4114–4119.

[3] Q.-Y. Fu, Y.-H. Feng, H.-M. Wang, and P. Liu, "Initial satellite access authentication based on Doppler frequency shift," *IEEE Wireless Communications Letters*, vol. 10, no. 3, pp. 498–502, 2020.

[4] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2016.

[5] M. Lou, J. Jin, H. Wang, D. Wu, L. Xia, Q. Wang, Y. Yuan, and J. Wang, "Performance analysis of sparse array based massive MIMO via joint convex optimization," *China Communications*, vol. 19, no. 3, pp. 88–100, 2022.

[6] E. Bjornson, L. Van der Perre, S. Buzzi, and E. G. Larsson, "Massive MIMO in sub-6 GHz and mmWave: Physical, practical, and use-case differences," *IEEE Wireless Communications*, vol. 26, no. 2, pp. 100–108, 2019.

[7] S. Chen, J. Zhang, Y. Jin, and B. Ai, "Wireless powered IoE for 6G: Massive access meets scalable cell-free massive MIMO," *China Communications*, vol. 17, no. 12, pp. 92–109, 2020.

[8] M. Alzenad and H. Yanikomeroglu, "Coverage and rate analysis for vertical heterogeneous networks (VHetNets)," *IEEE Transactions on Wireless Communications*, vol. 18, no. 12, pp. 5643–5657, 2019.

[9] H. Fang, X. Wang, and S. Tomasin, "Machine learning for intelligent authentication in 5G and beyond wireless networks," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 55–61, 2019.

[10] N. Alhussien, "Resource allocation in cellular machine-to-machine networks," Ph.D. dissertation, University of Victoria, 2021.

[11] Tibco, *What is the Internet of Things (IoT).* [Online], Available: https://www.tibco.com, 2022.

[12] Z. Yan, P. Zhang, and A. V. Vasilakos, "A survey on trust management for Internet of Things," *Journal of Network and Computer Applications*, vol. 42, pp. 120–134, 2014.

[13] V. Hassija, V. Chamola, V. Saxena, D. Jain, P. Goyal, and B. Sikdar, "A survey on IoT security: Application areas, security threats, and solution architectures," *IEEE Access*, vol. 7, pp. 82 721–82 743, 2019.

[14] H. Tschofenig and E. Baccelli, "Cyberphysical security for the masses: A survey of the internet protocol suite for internet of things security," *IEEE Security & Privacy*, vol. 17, no. 5, pp. 47–57, 2019.

[15] International Telecommunication Union, "IMT traffic estimates for the years 2020 to 2030," *Report M.2370*, 2015.

[16] A. Guidotti, A. Vanelli-Coralli, M. Caus, J. Bas, G. Colavolpe, T. Foggi, S. Cioni, A. Modenini, and D. Tarchi, "Satellite-enabled LTE systems in LEO constellations," in *IEEE International Conference on Communications Workshops*, Paris, France, 2017, pp. 876–881.

[17] H. Guo, J. Li, J. Liu, N. Tian, and N. Kato, "A survey on space-air-ground-sea integrated network security in 6G," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 53–87, 2021.

[18] J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, "Space-air-ground integrated network: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2714–2741, 2018.

[19] I. Altaf, M. A. Saleem, K. Mahmood, S. Kumari, P. Chaudhary, and C.-M. Chen, "A lightweight key agreement and authentication scheme for satellite-communication systems," *IEEE Access*, vol. 8, pp. 46 278–46 287, 2020.

[20] O. A. Topal and G. K. Kurt, "Physical layer authentication for LEO satellite constellations," in *IEEE Wireless Communications and Networking Conference*, Austin, TX, USA, 2022, pp. 1952–1957.

[21] Wikipedia, *Satellite constellation*. [Online], Available: https://en.wikipedia.org/wiki/Satellite_constellation, 2022.

[22] P. Zhang, J. Liu, Y. Shen, and X. Jiang, "Exploiting channel gain and phase noise for phy-layer authentication in massive MIMO systems," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 4265–4279, 2020.

[23] O. Günlü, K. Kittichokechai, R. F. Schaefer, and G. Caire, "Controllable identifier measurements for private authentication with secret keys," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 8, pp. 1945–1959, 2018.

[24] M. Abdrabou, A. Prince, and S. Hosny, "Implementation for EEA3 algorithm based on 3GPP standard," in *International Computer Engineering Conference*, Cairo, Egypt, 2018, pp. 137–140.

[25] M. A. Abdrabou, A. D. E. Elbayoumy, and E. Abd El-Wanis, "LTE authentication protocol (EPS-AKA) weaknesses solution," in *International Conference on Intelligent Computing and Information Systems*, Cairo, Egypt, 2015, pp. 434–441.

[26] M. A. Abdrabou, "Robust pre-authentication protocol for wireless network," in *International Computer Engineering Conference*, Cairo, Egypt, Dec. 2017, pp. 331–336.

[27] M. Trnka, J. Svacina, T. Cerny, E. Song, J. Hong, and M. Bures, "Securing internet of things devices using the network context," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 4017–4027, 2019.

[28] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: A survey of the emerging 5G network edge cloud architecture and orchestration," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1657–1681, 2017.

[29] T. O. Olwal, K. Djouani, and A. M. Kurien, "A survey of resource management toward 5G radio access networks," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1656–1686, 2016.

[30] X. Yao, Z. Chen, and Y. Tian, "A lightweight attribute-based encryption scheme for the Internet of Things," *Future Generation Computer Systems*, vol. 49, pp. 104–112, 2015.

[31] T. Bose, S. Bandyopadhyay, A. Ukil, A. Bhattacharyya, and A. Pal, "Why not keep your personal data secure yet private in IoT: Our lightweight approach,"

in *IEEE International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, Singapore, 2015.

[32] N. Xie, Z. Li, and H. Tan, "A survey of physical-layer authentication in wireless communications," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 282–310, 2020.

[33] L. Bai, L. Zhu, J. Liu, J. Choi, and W. Zhang, "Physical layer authentication in wireless communication networks: A survey," *Journal of Communications and Information Networks*, vol. 5, no. 3, pp. 237–264, 2020.

[34] X. Wang, P. Hao, and L. Hanzo, "Physical-layer authentication for wireless security enhancement: Current challenges and future developments," *IEEE Communications Magazine*, vol. 54, no. 6, pp. 152–158, 2016.

[35] Celestrak, *Celestrak Orbit Visualization.* [Online], Available: https://celestrak.org/, 2022.

[36] M. Abdrabou and A. Gulliver, "Physical layer authentication for satellite communication systems using machine learning," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 2380–2389, 2022.

[37] M. Abdrabou and T. A. Gulliver, "Adaptive physical layer authentication using machine learning with antenna diversity," *IEEE Transactions on Communications*, vol. 70, no. 10, pp. 6604–6614, 2022.

[38] Y. Liu, H.-H. Chen, and L. Wang, "Physical layer security for next generation wireless networks: Theories, technologies, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 347–376, 2016.

[39] L. Y. Paul, J. S. Baras, and B. M. Sadler, "Physical-layer authentication," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 38–51, 2008.

[40] D. Shan, K. Zeng, W. Xiang, P. Richardson, and Y. Dong, "PHY-CRAM: Physical layer challenge-response authentication mechanism for wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1817–1827, 2013.

[41] F. J. Liu, X. Wang, and H. Tang, "Robust physical layer authentication using inherent properties of channel impulse response," in *IEEE Military Communications Conference*, Baltimore, MD, USA, 2011, pp. 538–542.

[42] E. Martinian, G. W. Wornell, and B. Chen, "Authentication with distortion criteria," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2523–2542, 2005.

[43] M. Soltani, V. Pourahmadi, A. Mirzaei, and H. Sheikhzadeh, "Deep learning-based channel estimation," *IEEE Communications Letters*, vol. 23, no. 4, pp. 652–655, 2019.

[44] R.-F. Liao, H. Wen, J. Wu, F. Pan, A. Xu, H. Song, F. Xie, Y. Jiang, and M. Cao, "Security enhancement for mobile edge computing through physical layer authentication," *IEEE Access*, vol. 7, pp. 116 390–116 401, 2019.

[45] N. Wang, W. Li, P. Wang, A. Alipour-Fanid, L. Jiao, and K. Zeng, "Physical layer authentication for 5G communications: Opportunities and road ahead," *IEEE Network*, vol. 34, no. 6, pp. 198–204, 2020.

[46] H. Fang, X. Wang, and L. Hanzo, "Learning-aided physical layer authentication as an intelligent process," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2260–2273, 2018.

[47] H. Liu, Y. Wang, J. Liu, J. Yang, Y. Chen, and H. V. Poor, "Authenticating users through fine-grained channel information," *IEEE Transactions on Mobile Computing*, vol. 17, no. 2, pp. 251–264, 2017.

[48] F. Pan, Z. Pang, H. Wen, M. Luvisotto, M. Xiao, R.-F. Liao, and J. Chen, "Threshold-free physical layer authentication based on machine learning for industrial wireless CPS," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6481–6491, 2019.

[49] M. Rezaee, P. J. Schreier, M. Guillaud, and B. Clerckx, "A unified scheme to achieve the degrees-of-freedom region of the MIMO interference channel with delayed channel state information," *IEEE Transactions on Communications*, vol. 64, no. 3, pp. 1068–1082, 2016.

[50] Z. Jiang, J. Zhao, X.-Y. Li, J. Han, and W. Xi, "Rejecting the attack: Source authentication for Wi-Fi management frames using CSI information," in *IEEE INFOCOM*, Turin, Italy, 2013, pp. 2544–2552.

[51] O. H. Salim, A. A. Nasir, H. Mehrpouyan, and W. Xiang, "Multi-relay communications in the presence of phase noise and carrier frequency offsets," *IEEE Transactions on Communications*, vol. 65, no. 1, pp. 79–94, 2016.

[52] H. Lohrasbipeydeh, T. A. Gulliver, and H. Amindavar, "Unknown transmit power RSSD based source localization with sensor position uncertainty," *IEEE Transactions on Communications*, vol. 63, no. 5, pp. 1784–1797, 2015.

[53] J. Yang, Y. Chen, W. Trappe, and J. Cheng, "Detection and localization of multiple spoofing attackers in wireless networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 1, pp. 44–58, 2012.

[54] L. Xiao, T. Chen, G. Han, W. Zhuang, and L. Sun, "Game theoretic study on channel-based authentication in MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 7474–7484, 2017.

[55] N. Patwari and S. K. Kasera, "Temporal link signature measurements for location distinction," *IEEE Transactions on Mobile Computing*, vol. 10, no. 3, pp. 449–462, 2010.

[56] S. Chen, H. Wen, J. Wu, J. Chen, W. Liu, L. Hu, and Y. Chen, "Physical-layer channel authentication for 5G via machine learning algorithm," *Wireless Communications and Mobile Computing*, vol. 2018, art. no. 6039878, 2018.

[57] R. Candell, K. A. Remley, J. T. Quimby, D. Novotny, A. Curtin, P. B. Papazian, J. Diener, and M. Kashef, "Industrial wireless systems: Radio propagation measurements," *National Institute of Standards and Technology*, 2017.

[58] X. Qiu, J. Dai, and M. Hayes, "A learning approach for physical layer authentication using adaptive neural network," *IEEE Access*, vol. 8, pp. 26 139–26 149, 2020.

[59] T. M. Hoang, T. Q. Duong, H. D. Tuan, S. Lambotharan, and L. Hanzo, "Physical layer security: Detection of active eavesdropping attacks by support vector machines," *IEEE Access*, vol. 9, pp. 31 595–31 607, 2021.

[60] H. Fang, X. Wang, and L. Xu, "Fuzzy learning for multi-dimensional adaptive physical layer authentication: A compact and robust approach," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5420–5432, 2020.

[61] W. Hou, X. Wang, J.-Y. Chouinard, and A. Refaey, "Physical layer authentication for mobile systems with time-varying carrier frequency offsets," *IEEE Transactions on Communications*, vol. 62, no. 5, pp. 1658–1667, 2014.

[62] A. Ferrante, N. Laurenti, C. Masiero, M. Pavon, and S. Tomasin, "On the error region for channel estimation-based physical layer authentication over Rayleigh fading," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 5, pp. 941–952, 2015.

[63] W. Wang, Z. Sun, S. Piao, B. Zhu, and K. Ren, "Wireless physical-layer identification: Modeling and validation," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 2091–2106, 2016.

[64] 3GPP, "NR; Physical channels and modulation," *3rd Generation Partnership Project, Technical Specification 38.211*, version 15.2.0, 2018.

[65] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, Cambridge, UK, 2005.

[66] 3GPP, "Security architecture," *3rd Generation Partnership Project, Technical Specification 33.102*, version 17.0.0, 2022.

[67] P. Kyösti, J. Meinilä, T. Jämsä *et al.*, "WINNER II channel models," *Information Society Technologies, Technical Report IST-4-027756 WINNER II D1.1.2 V1.2*, 2007.

[68] L. Xiao, X. Lu, T. Xu, W. Zhuang, and H. Dai, "Reinforcement learning-based physical-layer authentication for controller area networks," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 2535–2547, 2021.

[69] K. St. Germain and F. Kragh, "Physical-layer authentication using channel state information and machine learning," in *International Conference on Signal Processing and Communication Systems*, Adelaide, SA, Australia, 2020.

[70] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. Di Renzo, "Safeguarding 5G wireless communication networks using physical layer security," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 20–27, 2015.

[71] M. Abdrabou and T. A. Gulliver, "Adaptive physical layer authentication for IoT in MIMO communication systems using support vector machine," *IEEE Internet of Things Journal*, vol. 10, no. 22, pp. 19 861–19 873, Nov. 2023.

[72] P. Zhang, T. Taleb, X. Jiang, and B. Wu, "Physical layer authentication for massive MIMO systems with hardware impairments," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 1563–1576, 2019.

[73] Z. Gu, H. Chen, P. Xu, Y. Li, and B. Vucetic, "Physical layer authentication for non-coherent massive SIMO-enabled industrial IoT communications," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3722–3733, 2020.

[74] N. Wang, L. Jiao, P. Wang, M. Dabaghchian, and K. Zeng, "Efficient identity spoofing attack detection for IoT in mm-Wave and massive MIMO 5G communication," in *IEEE Global Communications Conference*, Abu Dhabi, United Arab Emirates, 2018.

[75] F. Xie, Z. Pang, H. Wen, W. Lei, and X. Xu, "Weighted voting in physical layer authentication for industrial wireless edge networks," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 4, pp. 2796–2806, 2021.

[76] W. Wang, Y. Chen, and Q. Zhang, "Privacy-preserving location authentication in Wi-Fi networks using fine-grained physical layer signatures," *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1218–1225, 2015.

[77] M. Abdrabou and T. A. Gulliver, "LEO satellite authentication using physical layer features with support vector machine," in *IEEE International Conference on Communication, Networks and Satellite*, Solo, Indonesia, 2022.

[78] P. Tedeschi, S. Sciancalepore, and R. Di Pietro, "Satellite-based communications security: A survey of threats, solutions, and research challenges," *Computer Networks*, vol. 216, p. 109246, 2022.

[79] E. Schmidt, N. Gatsis, and D. Akopian, "A GPS spoofing detection and classification correlator-based technique using the LASSO," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 6, pp. 4224–4237, 2020.

[80] T. E. Humphreys, "Detection strategy for cryptographic GNSS anti-spoofing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 2, pp. 1073–1090, 2013.

[81] L. Heng, D. B. Work, and G. X. Gao, "GPS signal authentication from cooperative peers," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 1794–1805, 2014.

[82] S. Bhamdipati, T. Y. Mina, and G. X. Gao, "GPS time authentication against spoofing via a network of receivers for power systems," in *IEEE/ION Position, Location and Navigation Symposium*, Monterey, CA, USA, 2018, pp. 1485–1491.

[83] G. Oligeri, S. Sciancalepore, and R. Di Pietro, "GNSS spoofing detection via opportunistic IRIDIUM signals," in *ACM Conference on Security and Privacy in Wireless and Mobile Networks*, Linz (Virtual Event), Austria, 2020, pp. 42–52.

[84] E. Axell, E. G. Larsson, and D. Persson, "GNSS spoofing detection using multiple mobile COTS receivers," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, South Brisbane, QLD, Australia, 2015, pp. 3192–3196.

[85] K. D. Wesson, J. N. Gross, T. E. Humphreys, and B. L. Evans, "GNSS signal authentication via power and distortion monitoring," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 2, pp. 739–754, 2017.

[86] P. Baracca, N. Laurenti, and S. Tomasin, "Physical layer authentication over MIMO fading wiretap channels," *IEEE Transactions on Wireless Communications*, vol. 11, no. 7, pp. 2564–2573, 2012.

[87] G. Oligeri, S. Sciancalepore, S. Raponi, and R. Di Pietro, "PAST-AI: Physical-layer authentication of satellite transmitters via deep learning," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 274–289, 2022.

[88] E. Jedermann, M. Strohmeier, M. Schäfer, J. Schmitt, and V. Lenders, "Orbit-based authentication using TDOA signatures in satellite networks," in *ACM Conference on Security and Privacy in Wireless and Mobile Networks*, Abu Dhabi, United Arab Emirates, 2021, pp. 175–180.

[89] O. A. Topal, G. K. Kurt, and H. Yanikomeroglu, "Securing the inter-spacecraft links: Doppler frequency shift based physical layer key generation," in *IEEE International Conference on Wireless for Space and Extreme Environments*, Vicenza, Italy, 2020, pp. 112–117.

[90] O. A. Topal, K. Kurt, and H. Yanikomeroglu, "Securing the inter-spacecraft links: Physical layer key generation from Doppler frequency shift," *IEEE Journal of Radio Frequency Identification*, vol. 5, pp. 232–243, 2021.

[91] A. Al-Hourani, "Session duration between handovers in dense LEO satellite networks," *IEEE Wireless Communications Letters*, vol. 10, no. 12, pp. 2810–2814, 2021.

[92] A. Goldsmith, *Wireless Communications*. Cambridge University Press, Cambridge, UK, 2005.

[93] T. Pratt and J. E. Allnutt, *Satellite Communications*. John Wiley & Sons, New Delhi, India, 2020.

[94] S. Silver, *Microwave Antenna Theory and Design.* McGraw Hill, New York, NY, USA, 1949.

[95] M. Murata, I. Kawano, and K. Inoue, "Precision onboard navigation for LEO satellite based on precise point positioning," in *IEEE/ION Position, Location and Navigation Symposium*, Portland, OR, USA, 2020, pp. 1506–1513.

[96] A. Hauschild, J. Tegedor, O. Montenbruck, H. Visser, and M. Markgraf, "Precise onboard orbit determination for LEO satellites with real-time orbit and clock corrections," in *International Technical Meeting of the Satellite Division of The Institute of Navigation*, Portland, OR, USA, 2016, pp. 3715–3723.

[97] J. Zhang, G. Yang, Q. Xu, and Y. Zhao, "Application in radar simulation of STK/connect module," in *WRI World Congress on Computer Science and Information Engineering*, Los Angeles, CA, USA, 2009, pp. 274–276.

[98] Y. Zhou, P. L. Yeoh, K. J. Kim, Z. Ma, Y. Li, and B. Vucetic, "Game theoretic physical layer authentication for spoofing detection in UAV communications," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 6750–6755, 2022.

[99] M. Abdrabou and T. A. Gulliver, "Threshold-based physical layer authentication for space information networks," in *IEEE International Conference on Communication, Networks and Satellite*, Solo, Indonesia, Nov. 2022, pp. 289–293.

[100] M. Abdrabou and Gulliver, "Authentication for satellite communication systems using physical characteristics," *IEEE Open Journal of Vehicular Technology*, vol. 4, pp. 48–60, 2022.

[101] J. N. Gross, C. Kilic, and T. E. Humphreys, "Maximum-likelihood power-distortion monitoring for GNSS-signal authentication," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 1, pp. 469–475, 2018.

[102] K. D. Wesson, B. L. Evans, and T. E. Humphreys, "A combined symmetric difference and power monitoring GNSS anti-spoofing technique," in *IEEE Global Conference on Signal and Information Processing*, Austin, TX, USA, 2013, pp. 217–220.

[103] D. Borio, "PANOVA tests and their application to GNSS spoofing detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 1, pp. 381–394, 2013.

[104] S. Mathur, R. Miller, A. Varshavsky, W. Trappe, and N. Mandayam, "Proximate: Proximity-based secure pairing using ambient wireless signals," in *International Conference on Mobile Systems, Applications, and Services*, Washington, DC, USA, 2011, pp. 211–224.

[105] F. Formaggio and S. Tomasin, "Authentication of satellite navigation signals by wiretap coding and artificial noise," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, art. no. 98, 2019.

[106] N. Sabri, S. Aljunid, M. Salim, R. Kamaruddin, R. Ahmad, and M. Malek, "Path loss analysis of WSN wave propagation in vegetation," *Journal of Physics: Conference Series*, vol. 423, art. no. 012063, 2013.

[107] C. Chen, M. Song, C. Xin, and J. Backens, "A game-theoretical anti-jamming scheme for cognitive radio networks," *IEEE Network*, vol. 27, no. 3, pp. 22–27, Jun. 2013.

[108] Y. E. Sagduyu and A. Ephremides, "A game-theoretic analysis of denial of service attacks in wireless random access," in *Wireless Networks*, Limassol, Cyprus, Apr. 2007, pp. 651–666.

[109] L. Xiao, T. Chen, G. Han, W. Zhuang, and L. Sun, "Channel-based authentication game in MIMO systems," in *IEEE Global Communications Conference*, Washington, DC, USA, Dec. 2016.

[110] L. Xiao, Y. Li, G. Han, G. Liu, and W. Zhuang, "PHY-layer spoofing detection with reinforcement learning in wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10 037–10 047, Dec. 2016.

[111] Y. Chen, J. Yang, W. Trappe, and R. P. Martin, "Detecting and localizing identity-based attacks in wireless and sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2418–2434, 2010.

[112] T. Jing, Y. Wu, Y. Huo, and Q. Gao, "A Stackelberg game based physical layer authentication strategy with reinforcement learning," in *IEEE International Conference on Communications*, Seoul, Korea, May 2022, pp. 3322–3327.

[113] M. A. Aygül, S. Büyükçorak, D. B. da Costa, H. F. Ateş, and H. Arslan, "Signal relation-based physical layer authentication," in *IEEE International Conference on Communications*, Dublin, Ireland, Jun. 2020.

[114] Y.-S. Shiu, S. Y. Chang, H.-C. Wu, S. C.-H. Huang, and H.-H. Chen, "Physical layer security in wireless networks: A tutorial," *IEEE wireless Communications*, vol. 18, no. 2, pp. 66–74, Apr. 2011.

[115] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, 2001.

[116] A. Fernández, S. García, M. Galar, R. C. Prati, B. Krawczyk, and F. Herrera, *Learning from Imbalanced Data Sets*. Springer, Cham, Switzerland, 2018.

[117] Z. Jiang, J. Zhao, X.-Y. Li, J. Han, and W. Xi, "Comparative performance analysis of kernel functions in support vector machines in the diagnosis of pneumonia using lung sounds," in *International Conference on Computing and Information Technology*, Tabuk, Saudi Arabia, 2022, pp. 320–324.