DETECTION AND MITIGATION TECHNIQUES AGAINST EAVESDROPPING AND SPOOFING IN GPS

by

Mehmet Özgün Demir

B.S., Telecommunication Engineering, Istanbul Technical University, 2014

M.S., Telecommunication Engineering, Istanbul Technical University, 2017

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Electrical and Electronics Department

Boğaziçi University

2023

ACKNOWLEDGEMENTS

I would also like to extend my deepest gratitude to my supervisors, Prof. Ali Emre Pusane and Prof. Güneş Karabulut Kurt for their valuable advice on all technical aspects, continuous support for keeping me motivated during the research, and patience during my Ph.D. study. Their extensive knowledge and wide experience have encouraged me in all the time of my academic research, daily life, and professional career. More importantly, they always believe in me and my abilities.

I also extend my sincere thanks to my thesis committee for their insightful comments that improve my thesis. Thanks should also go to all the members of the Wireless Communication Laboratory, where I also had the great pleasure of working. I'm extremely grateful to the whole Lifemote Networks family, especially the research team, for being kind colleagues for last more than two years and providing the time for my research when I needed it.

To my mom and my sister, you should know that your support and encouragement were worth more than I can express on paper. I also wish that my dad would also see this day. We miss you every day for all those long years. Last, but not least, I especially thank Deniz for your love, for sharing my joy in good times, and for your companion in the dark times.

ABSTRACT

DETECTION AND MITIGATION TECHNIQUES AGAINST EAVESDROPPING AND SPOOFING IN GPS

As a significant CPS application, V2X networks have their own specific security atmosphere in addition to their latency and energy efficiency requirements. There are two critical attack types against secure V2X deployments, which are eavesdropping and spoofing attacks. In this thesis, we evaluate these attacks in a novel system model, where aerial spoofers, ground-located spoofers, and eavesdroppers are jointly integrated. Against eavesdropper attacks, an error vector is added to the information data to falsify the attackers using FEC codes with respect to and relay-aided McEliece cryptosystem-based methodology. The results show that significantly noisier channels between relays and the eavesdropper enhance information security. Additionally, the impacts of time spoofing attacks on satellite positioning signals are studied. For this purpose, two pseudorange-based spoofing detection algorithms with different performances and complexity levels are proposed to mitigate ground-located spoofers. In order to find the best detection thresholds of these algorithms, Pareto fronts are plotted, and then a decision tree-based detection approach is analyzed. These algorithms are also tested against novel aerial spoofers, which suffer additional mobility and atmospheric errors. Finally, a high-level hybrid decision methodology is applied to improve spoofing detection rates and reduce false alarm rates by integrating a decision fusion module. With this approach, the decision fusion module combines individual outcomes of the proposed algorithm and improves spoofing decision performance. In summary, this thesis proposes extensive strategies against eavesdropper and spoofing attacks in V2X networks by providing secure transmission schemes and detection algorithms in a novel system model.

ÖZET

GPS İÇİN GİZLİ DİNLEME VE YANILMA SALDIRILARINA KARŞI SEZME VE KISITLAMA TEKNİKLERİ

Önemli bir CPS uygulaması olarak V2X ağlarının gecikme süresi ve enerji verimliliği gereksinimlerine ek olarak kendine has güvenlik ortamı bulunur. Güvenli V2X uygulamalarına karşı, gizli dinleme ve yanıltma saldırıları olmak üzere iki kritik saldırı türü vardır. Bu tezde, bu saldırıları havadaki yanıltıcılar, yerde konumlu yanıltıcılar ve gizli dinleyicilerin ortaklaşa entegre edildiği yeni bir sistem modelinde değerlendiriyoruz. Gizli dinleyici saldırılarına karşı olarak, McEliece kriptosistem tabanlı ve röle destekli metodolojiyle FEC kodlarını kullanarak saldırganı yanıltmak için bilgi verilerine bir hata vektörü eklenir. Sonuçlara göre, röleler ve gizli dinleyici arasındaki önemli ölçüde daha gürültülü kanalların bilgi güvenliğini artırır. Ek olarak, zaman bilgisine yönelik yanıltıcı saldırılarının uydu konumlandırma sinyalleri üzerindeki etkileri incelenmiştir. Bu amaçla, yer konumlu yanıltıcılara karşı farklı performanslara ve karmaşıklık seviyelerine sahip sözde mesafe tabanlı iki yanıltıcı saldırı algılama algoritması önerilmiştir. Bu algoritmaların en iyi sezme eşiklerini bulmak için Pareto cepheleri çizilir ve sonrasında karar ağacı tabanlı bir sezme yaklaşımı analiz edilir. Ayrıca bu algoritmalar, hareketlilik ve atmosferik gibi ek hatalara maruz kalan havadaki yanıltıcılara karşı test edilir. Son olarak, yanıltıcı tespit oranlarını iyileştirmek ve yanlış alarm oranlarını azaltmak için bir karar birleştirme modülünü entegre edilerek yukarı seviye bir hibrit karar metodolojisi uygulanır. Bu yaklaşımla, karar birleştirme modülü, önerilen algoritmanın ayrı aryı sonuçlarını birleştirir ve yanıltıcı tespiti performansını geliştirir. Özetle, bu tez, yeni bir sistem modelinde güvenli iletim şemaları ve algılama algoritmaları sağlayarak V2X ağlarında gizli dinleme ve yanıltma saldırılarına karşı kapsamlı stratejiler önermektedir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii
ABSTRACT iv
ÖZET v
LIST OF FIGURES ix
LIST OF TABLES xiii
LIST OF SYMBOLS xiv
LIST OF ACRONYMS/ABBREVIATIONS xv
1. INTRODUCTION 1
1.1. System Model for the Security of V2X Applications 5
1.2. Literature Survey 9
2. THE SECURITY NEEDS OF V2X BASED ON A GARDEN OF CPS . . . 14
2.1. Flowers of CPS 16
2.1.1. Industrial Wireless Sensor Networks (IWSN) 18
2.1.2. Smart Grid (SG) 19
2.1.3. Healthcare Applications 20
2.1.4. V2X 22
2.2. Security Aspects 23
2.2.1. IWSN 23
2.2.2. Smart Grid 24
2.2.3. Healthcare Applications 24
2.2.4. V2X 25
2.3. Implementation Aspects 26
2.3.1. IWSN 26
2.3.2. Smart Grid 27
2.3.3. Healthcare Applications 27
2.3.4. V2X 28
3. GNSS FUNDAMENTALS WITH SPOOFING THREATS 29
3.1. The Fundamentals of GNSS 29

3.2. Transmitting and Receiving Units of GPS 31
3.2.1. GPS Signal Generation 31
3.2.2. GPS Receiver 32
3.2.3. Pseudorange Calculation 34
3.3. Security Threats Against GPS 36
3.3.1. Spoofing Attack Model 37
3.3.2. Pseudorange Observation-based Spoofing Detection 39
4. DEFEATING EAVESDROPPER USING FEC CODES 42
4.1. System Model of FEC-based Eavesdropping Mitigation with the Help of Relays and McEliece Cryptosystem 43
4.2. Calculation of Analytical Error Probabilities 49
4.3. Application Scenarios 53
4.4. Simulation Results 54
5. PSEUDORANGE BASED SPOOFING DETECTION ALGORITHMS USING HYPERBOLA EQUATIONS 60
5.1. System Model 61
5.2. Time-spoofing Attacks in GPS 62
5.3. Derivation of Hyperbola Equations 65
5.3.1. Single Difference Between Two Authentic Pseudoranges (Case 1) 66
5.3.2. Single Difference Between One Authentic Pseudorange and One Spoofed Pseudorange (Case 2) 66
5.3.3. Single Difference of Two Spoofed Pseudorange Values (Case 3) . 67
5.4. Spoofing Detection Strategies 68
5.4.1. Algorithm 1: Sub-optimal Search-based Spoofing Detection Algorithm 70
5.4.2. Algorithm 2: Subset Selection-based Spoofing Detection Algorithm 74
5.4.3. Hybrid Spoofing Detection Methodology 75
5.5. Simulation Results of Spoofing Detection Algorithms 77
5.5.1. Comparison with Detection Thresholds and the Decision-tree Machine Learning Algorithm 80
5.5.2. Simulation Results 81

5.5.3. Decision Fusion for Hybrid Decision 85
6. CONCLUSION 87
REFERENCES 90

LIST OF FIGURES

Figure 1.1. System model for V2X network with possible eavesdropper, ground-located, and aerial spoofers. 6

Figure 2.1. A simplified cycle of CPS. The vulnerable positions in CPS are indicated. 14

Figure 2.2. CPS flower, which shows requirements, challenges, and implementations with some external factors that help to satisfy efficient, robust, and secure CPS deployments and possible security threats. 16

Figure 2.3. Flowers of the CPS applications. Their leaves and length of roots indicate the specific requirements of the application: (a) IWSNs; (b) smart grid; (c) smart healthcare; (d) V2X communications. . . 17

Figure 3.1. The block diagram of a GPS receiver. 32

Figure 3.2. C/A correlation peak of the unspoofed signal during the acquisition process. 33

Figure 4.1. The proposed system model, where Alice injects a spatially distributed random error vector to the codeword and transmits with the aid of relays to Bob. Eve tries to decode successfully, where SNR_{low} valued dashed-lined channels may lead to additional errors. 43

Figure 4.2. The comparison of simulation results of **S1** and analytical results of $P_{\epsilon}^{C_I}$ where $L \in \{0, 1, 2, 3, 4\}$ for Reed-Solomon $(15, 11, 2)$ code. . 55

Figure 4.3. The comparison of the simulation results and analytical results of FER values of Eve for Golay $(23, 12, 3)$ code for AWGN and Rayleigh channels where $L \in \{0, 1, 2\}$. 56

Figure 4.4. The comparison of the simulation results and analytical results of FER values of Eve for Hamming $(7, 4, 1)$ code for AWGN and Rayleigh channels where $L \in \{0, 1, 2\}$. 57

Figure 4.5. The comparison of the simulation results and analytical results of FER values of Eve for Hamming $(15, 11, 1)$ code for AWGN and Rayleigh channels where $L \in \{0, 1, 2\}$. 58

Figure 5.1. A system model of the pseudorange-based spoofing detection in a vehicular communication scenario. At the target vehicle, both spoofing and authentic GPS signals are available. 61

Figure 5.2. Receivers' position errors in meters are given on each plane with respect to the values of $N_S \geq N_A$ for Δt_{tr}. 64

Figure 5.3. Receivers' position errors in meters are given on each plane with respect to the values of $N_S \geq N_G$ for Δt^s. 65

Figure 5.4. The real and calculated vehicle positions using hyperbolic equations with five authentic satellites. 69

Figure 5.5. The real and calculated vehicle positions using hyperbolic equations with four authentic satellites and one spoofer. 69

Figure 5.6. The operation block diagram of spoofing detection module, which may run both Algorithm 1 or Algorithm 2. 71

Figure 5.7. Pseudocode of sub-optimal search-based spoofing detection algorithm (Algorithm 1). 72

Figure 5.8. Hybrid spoofing detection operation of both of the proposed algorithms based on power sources of the devices for $I = 4$. 76

Figure 5.9. Average position error of proposed algorithms, Algorithm 1: sub-optimal search-based detection algorithm and its efficient version, Algorithm 2: subset selection-based detection algorithm for a various number of spoofers and chosen additional spoofer time error. 78

Figure 5.10. Pareto front curves of the spoofing detection algorithms with respect to spoofing detection rates and false alarm rates for additional spoofing time error $\Delta t^s \in \mathcal{T} = \{1s, 0.1s, 10ms, 0.1ms, 10\mu s\}$. . . . 80

Figure 5.11. Spoofing detection rates of the detection algorithms. 81

Figure 5.12. False alarm ratios of the spoofing detection algorithms. 82

Figure 5.13. Spoofing detection rates of the detection algorithms with respect to mobility and spoofing imperfection errors comparing to reference values, which are firstly given in Figure 5.11 for $\Delta t^s \in \{0.1s, 1ms\}$ and $\lambda \in \{7.5 \times 10^5, 10^6\}$. 83

Figure 5.14. False alarm rates of the detection algorithms with respect to mobility and spoofing imperfection errors comparing to reference values, which are firstly given in Figure 5.12 for $\Delta t^s \in \{0.1s, 1ms\}$ and $\lambda \in \{7.5 \times 10^5, 10^6\}$. 84

Figure 5.15. Spoofing detection rates of the detection algorithms with respect to spoofer types, which can be ground located, UAV, HAPS, or LEO satellite spoofers. 85

Figure 5.16. Spoofing detection and false alarm rates of CV and CVR are compared respective sole use of Algorithm 1 and Algorithm 2. Simulations are extended for $I = 2$ and $I = 3$ when three possible BSC errors. 86

LIST OF TABLES

Table 2.1. Security perspective of the chosen CPS applications. 23

Table 2.2. Implementation perspective of the chosen CPS applications. 26

Table 3.1. The list of significant parameters for pseudorange expressions. . . . 35

Table 4.1. Probabilities of number of overlaps for Reed-Solomon codes with respect to the error probability of Eve. 50

Table 5.1. Parameters of time-spoofing attack. 63

Table 5.2. The list of remaining parameters and the notation of Chapter 6 in addition Table 3.1. 70

Table 5.3. m and $\bar{m}$ values for sub-optimal search-based and subset selection-based spoofing detection algorithms for $N = 8$. 73

Table 5.4. Error Parameters for various spoofing attacks. 84

LIST OF SYMBOLS

$\boldsymbol{G}$	Generation matrix
I	Ionospheric error
J	Total number of vehicles and/or RSU
N_G	Number of visible GPS-like signals
N_A	Number of authentic signals
N_S	Number of spoofing signals
P_D	Detection rate
P_F	False alarm rate
P_e	Error probability
SNR	SNR value
T	Trophospheric error
$\boldsymbol{x}$	Information vector
$\boldsymbol{x_e}$	Error vector
Δt^s	Additional time error due to spoofing
ϵ_b	Binary symmetric channel error rate
λ	Detection threshold
ρ	Pseudorange
Ψ_{df}	Decision fusion methodology

LIST OF ACRONYMS/ABBREVIATIONS

AWGN	Additive White Gaussian Noise
BPSK	Binary Phase Shift Keying
BSC	Binary Symmetric Channel
CPS	Cyber-Physical Systems
CR	Counting Rule
CVR	Chair-Varshney Rule
DSRC	Direct Short-Range Communication
FEC	Forward Error Correction
FER	Frame Error Rate
GLONASS	Global Navigation Satellite System in Russian
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HAPS	High Altitude Alatforms
IWSN	Industrial Wireless Sensor Networks
LEO	Low Earth Orbit
LTE	Long Term Evolution
OBU	Onboard Unit
RSU	Roadside Unit
SG	Smart Grid
SNR	Signal-to-Noise Ratio
UAV	Unmanned Aerial Vehicle
V2X	Vehicle-To-Everything
WBAN	Wireless Body Area Network

1. INTRODUCTION

Cyber-physical systems (CPS) are new-generation closed-loop solutions with integrated cyber and physical parts that offer a wide range of applications in diverse sectors. In the literature, the main applications are classified as smart grids, industrial wireless sensor networks (IWSN), vehicle-to-everything (V2X) communication networks, and smart healthcare systems, which are mostly deployed with sensor-based architectures [1]. As a part of CPS, V2X communication networks are also sensitive to a variety of security threats because a wireless communication channel is more prone to risks than a wired communication medium against malicious attacks. These automated networks are highly interconnected with human behavior, daily life, military, and industrial operations; hence, secure deployments are needed to protect individuals in physical and cyber environments while satisfying smooth and efficient system operations [2]. In addition to the security requirements, there are many other application-dependent requirements, such as energy efficiency, computational complexity, latency, and reliability [3].

Possible threats against secure V2X systems can be observed as an active attack, which has direct impacts on the system, or as a passive attack, where the attacker sneakily steals sensitive information. In both cases, consequences may be crucial. For example, there may be threats to the security of lives, private information disclosures, or the waste of costly investments. An attacker can falsify the sensed data during wireless transmission from sensors to controllers in V2X communication systems. As a result, passengers' lives may be at risk if a car crashes due to the wrong decision of the controller within this car [4]. Thus, the need for security is crucial in V2X systems, and it is not possible to deploy these systems without proper security-enhancing mechanisms.

In the scope of this thesis, we first evaluate the eavesdropping attack, which is one of the critical passive attacks against V2X security. The main target of an eavesdrop-

per is to obtain sensitive information about the users. In general, the different types of data can be considered sensitive information in different CPS applications. For example, personal health data in smart health applications and the users' positions in V2X systems. The most recognized solution against eavesdroppers is the encryption of transmitted signals. However, cryptographic methods may not meet all of the requirements of CPS due to the constraints of CPS sensors, such as finite memory, low complexity, and limited battery life. In order to deal with an eavesdropper, physical layer-based security schemes could be used in CPS as an additional layer of security. In addition, it is possible to achieve perfectly secure communication in an information-theoretic sense [5] with efficient power consumption [6] and low complexity [7]. Physical layer security enhancements have emerged with Wyner's wiretap channel model in [8]. In this model, a legitimate transmitter, Alice, transmits information signals to an approved receiver, Bob. At the same time, an unauthorized eavesdropper, Eve, tries to capture the transmitted information signals over the wire-tap channel. In recent years, this model has been vastly studied for various security mechanisms.

In order to enhance security against an eavesdropper, we introduce a new system model for CPS based on the error injection idea of the McEliece cipher and cooperative communication networks with the forward error correction (FEC) codes. The security capability of the proposed system model is analyzed by evaluating the decoding error probability of an eavesdropper while satisfying reliable communication with the legitimate receiver. In order to evaluate the system performance, we analyze the decoding frame error rates (FER) of an eavesdropper, while satisfying reliable communication with the legitimate receiver. Analytical error probability expressions are derived and verified with simulations for the considered FEC codes, which are Reed-Solomon, Hamming, and Golay codes. Our results indicate that this approach is highly beneficial, especially when the transmission environment is dynamic and noisy, e.g., communication channels of V2X systems.

Beyond passive attacks, V2X systems may be vulnerable to active attacks. One possible attack strategy is spoofing the satellite positioning signals. V2X systems highly

rely on positioning systems, where satellite communication systems play a critical role in providing location information. However, the security for these systems is mostly negligible until the recent past; therefore, spoofing may severely impact positioning services. As a result, there are significant threats against secure signaling between transmitter and receiver due to the vulnerability of the use of civilian Global Navigation Satellite System (GNSS) signals. Recently, several incidents have been reported, such as the highjack of an aerial vehicle (AV) of the USA by Iran using spoofing [9]. There are also reported real-world incidents of Global Positioning Systems (GPS) spoofing in Russia and the Black Sea region [10], where spoofing is used as a defense mechanism. With the development of software-defined radios (SDR), generating attack signals against positioning systems is significantly cheaper and more manageable than before; as a result, there are several examples of these attacks utilizing SDR [11].

The target of a spoofer is to mimic the transmitter, whose signals are received and used by the receivers without being detected. There are several versions of spoofing attacks in GNSS; hence, their countermeasures are also varied [9, 12]. Spoofers generally send a combination of legitimate GNSS signals to the receiver. Spoofing attacks are significant threats to aerial vehicles and communication systems, e.g., high-altitude platform systems (HAPS), unmanned aerial vehicles (UAV), and low Earth orbit (LEO) satellites. Among them, HAPS platforms regain their popularity with the possible services as a part of communication and computation infrastructures of many industries, e.g., logistics, navigation, weather services, aerospace, telecommunication [13, 14]. These quasi-stationary platforms are located around 20km above the ground level [15]. Possible low-delay performance and minimum back-haul infrastructures are the main advantages of these systems for various applications. In addition, there are multiple projects with LEO satellites that are currently expanding, such as SpaceX's Starlink, OneWeb, and Amazon's Project Kuiper [14, 16]. With the integration of UAV applications, GNSS satellites, and ground units, vertical heterogeneous networks (VHetNet) can be constructed [17]. Overall, security is also a non-negotiable concept for these networks against mentioned active and passive attacks. However, the main assumption is that the aerial stations are friendly and ready to cooperate to

detect ground spoofers [18, 19]. On the other hand, we believe that these stations can be hijacked [14]; therefore, they may behave as aerial spoofers. As a result, the aerial spoofer should also be accounted for in the design of spoofing detection systems.

In the literature on the detection strategies of spoofing attacks, the majority of the solutions are based on the assumption that all available signals are spoofed. In this thesis, we focus on detecting spoofing attacks in which authentic and spoofing positioning signals coexist in a V2X system. In other words, the target receiver is expected to obtain authenticated GNSS signals in addition to the spoofing signals while under attack. Pseudoranges, which are required observables to obtain position information, are also affected, that in turn, may lead to significant failures in V2X systems due to wrong positioning under a spoofing attack. Besides, many works in the literature aim to detect spoofing attacks with pseudorange-based operations, e.g., single and double-differences [20, 21]. On the other hand, in this thesis, as the first step of spoofing detection, we inversely exploit known single difference operation, which enables us to have only one receive antenna, while multi-antenna spoofing detection is common. This usage provides a hyperbolic equation system whose solution is used for spoofing detection by comparing it with a reference value. The comparison is achieved by utilizing increased GNSS receivers' sensitivities.

Under these circumstances, we propose a sub-optimal search-based algorithm (Algorithm 1) with a search of all possible spoofing signal scenarios for the given available GNSS signal set. In order to reduce the algorithm complexity, Algorithm 1 is improved by intelligently selecting the search subsets of the spoofing attack scenarios in Algorithm 2. The performances of these algorithms are analyzed for a variety of spoofing attack scenarios, e.g., ground spoofer, HAPS spoofer, and LEO satellite spoofer. Due to complexity issues, Algorithm 1 is more suitable for the V2X elements, which have power connections, e.g., roadside units (RSU), since RSUs may have fixed positions and power connections in many deployments. Since GNSS receivers are extremely power-consuming [22], Algorithm 2 fits better to battery-limited devices, which can be considered moving V2X elements, e.g., motorized vehicles, In addition to the power

limitations, the vulnerabilities of these V2X elements against the positioning attacks may also be varied, since RSUs have a fixed deployment position, whereas the target vehicles move nearby RSUs. Beyond complexity issues, these algorithms variously perform in terms of detection and false alarm rate under different spoofing attacking scenarios. In order to improve overall system performance, a collaboration between RSUs and target vehicles is needed, and a high-level hybrid operation can be deployed in V2X systems. This coordination can be done at a power-connected RSU, where it fuses spoofing detection decisions from different sources (e.g., moving vehicles or other RSUs). In this thesis, considered detection fusion methodologies are the counting rule (CR) and Chair–Varshney rule (CVR), which are well-known techniques in the literature [23, 24].

In the next sections, we present the system model for the security of V2X applications and the literature review of the thesis, respectively. In Chapter 2, we present the challenges and requirements of main CPS applications, including V2X, from a security perspective with a novel flower form. This novel approach was published in [25]. In Chapter 3, the fundamentals of GNSS are explained in detail with possible spoofing attacks. In Chapter 4, we present the FEC-based studies against eavesdropper attacks in a relay-aided V2X scenario, where the main content of this chapter was published in [26, 27]. The studies on spoofing detection against satellite-based positioning systems are given in Chapter 5. In this chapter, spoofing detection algorithms against both ground and aerial spoofers and high-level hybrid decision methodology are presented, and the main body of this chapter was published in [28, 29]. In the end, this thesis is concluded with a large overview of chapters in Chapter 6.

1.1. System Model for the Security of V2X Applications

Our system model is given in Figure 1.1 with a V2X communication scenario with possible passive and active attackers. The target vehicle is a part of the V2X network and it communicates with other vehicles with the help of RSUs with a relay-aided structure. Here, we can assume that the V2X network supports DSRC or LTE-V2X, where

DSRC is IEEE 802.11p-based and LTE-V2X is 3rd generation partnership project (3GPP) supported for V2X communications [30]. Both of these V2X communication approaches should satisfy the fundamental requirements of V2X systems, which are very low latency, high reliability, and enhanced security [25, 31]. At the same time, GNSS satellites provide positioning signals to the V2X agents, which may be vehicles or RSUs. Since GNSS signals can be considered as independent inputs to each V2X system element from outside, the V2X communication technology can be arbitrarily selected. The GNSS signals are used by a target vehicle and an RSU to find their position. The target vehicle computes its position on its onboard unit (OBU), and it also communicates with other V2X elements (e.g., other vehicles, RSUs) with the same unit.

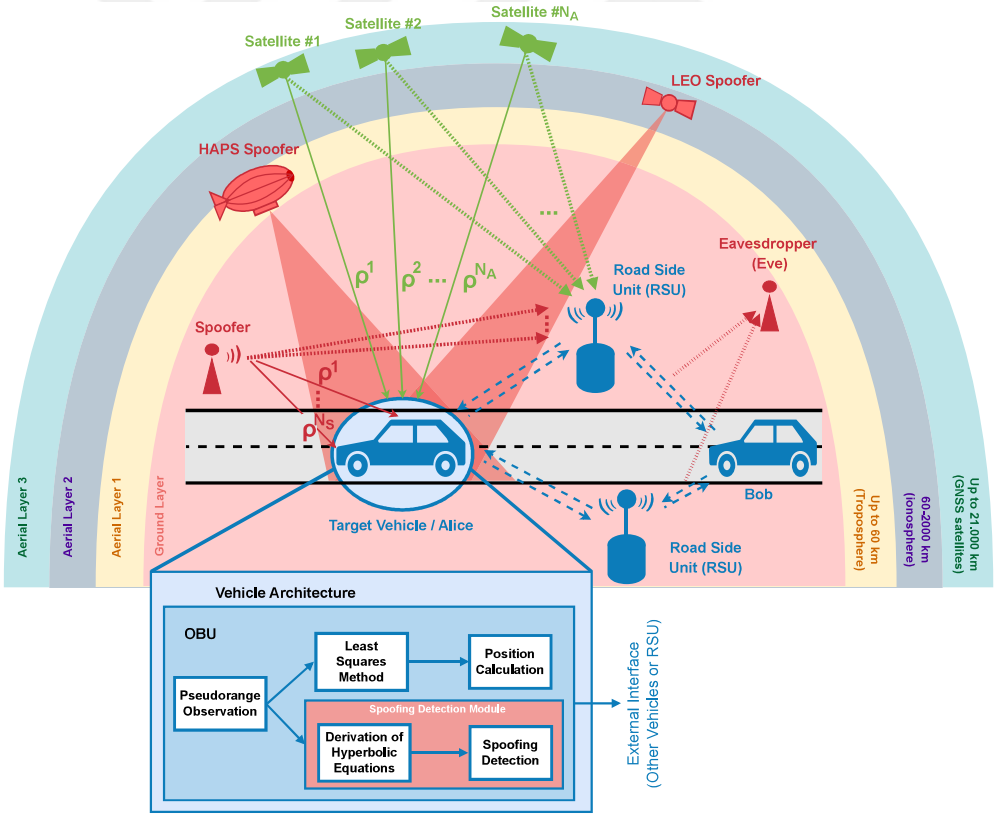

Figure 1.1. System model for V2X network with possible eavesdropper, ground-located, and aerial spoofers.

In case of an active attack against V2X communication security, a spoofer sends malicious signals to the target receiver (e.g., a roadside unit (RSU) or a vehicle) in

order to deceive the receiver. On the other hand, an eavesdropper (Eve), only listens to the communication medium and tries to capture critical information without any interaction with the target receiver [30]. As shown in Figure 1.1, these passive and active attackers are integrated into the same picture. In order to integrate passive attackers, we have been inspired by Wyner's wiretap channel model, which plays a key role in physical layer security studies [8]. In the wiretap model, one legitimate transmitter (Alice) and one legitimate receiver (Bob) communicate over the main channel, while a non-legitimate receiver, an eavesdropper called Eve, monitors the communication between legitimate users over the wiretap channel. This well-known wiretap channel model is vastly studied in various system models, whose requirements differ considerably [32]. By monitoring the channel, Eve could get access to sensitive control systems. This model is adapted when there is no direct main channel. Instead, there may be multiple relays between Alice and Bob. In this case, Eve tries to monitor the channels between the relays and Bob.

In this thesis, we propose a secure communication model against the eavesdropper by injecting error vectors that are inspired by the McEliece cryptosystem [33] in a distributed manner by combining existing security solutions, which are FEC codes and cooperative communication techniques, when Eve is a part of V2X communication network. FEC codes are already used in many applications to enhance communication security when the eavesdropper is a part of the network [34]. In [6], the authors propose that existing security techniques should emerge with secure coding schemes in order to achieve a high level of protection. In this model, we first encode information messages and then introduce random errors to confuse the eavesdropper by using a similar idea as the McEliece cryptosystem. The error vector is determined according to the error correction capability of the receiver node. We use Reed Solomon, Golay, and Hamming codes as FEC codes and encoded messages can be successfully decoded on the receiver side with a high SNR scenario. Due to the channel imperfections between relays and Eve, she may not receive all the codewords correctly; therefore, she cannot successfully decode them. As a result, we can achieve significantly high error rates on the eavesdroppers' side. The proposed model can improve the security performance of

V2X against eavesdroppers since we can achieve a high level of protection by merging existing physical layer methods with the secure coding schemes [6]. Also, the proposed methodology can be utilized as another layer of security in CPS by using the massive amount of sensor nodes that may be placed in these systems [35].

On the other hand, spoofing is classified as an active attack among the existing attacks against GNSS-based positioning in the chosen application. During a spoofing attack against a secure V2X system, the transmitter, which may be any corrupted ground-located traffic element, e.g., another vehicle, a traffic sign, or an RSU, aims to deceive the target vehicle with its misleading signals without being detected. In addition to these scenarios, we extend the possible spoofer scenarios in addition to the ground-located spoofers. Hence, possible aerial spoofers, which may be hijacked HAPS stations, a UAV unit, or an LEO satellite, are also integrated into the system model, as shown in Figure 1.1. When aerial spoofers are part of the system, we should also consider the atmospheric impacts of the long-distance positioning and spoofing signals. Among the atmospheric errors, the troposphere is the layer up to $60km$ from the ground, and it may lead to up to $3m$ positioning error based on the weather condition [36–38]. The ionosphere is placed between 60-2000km above ground, and its impact on positioning signals is up to $20m$ based on the solar activity [39–41]. In the case of a UAV spoofer, we add a mobility-based error term to the pseudorange value of the spoofed signal without adding an atmospheric error due to the low operation altitude. On the other hand, we additionally take into account tropospheric errors when there is a HAPS spoofer, which is located at $20km$ above the ground. Finally, we both consider tropospheric and ionospheric error terms in case of the spoofer is an LEO satellite. As a result, during the spoofing detection phase, the pseudorange terms differ for each spoofing scenario based on the additional error sources.

Since the impacts of spoofing attacks may be crucial and unrecoverable in some cases, detection strategies are densely studied in the literature without specifying a particular application. Several approaches can be listed as power level-based detection, pseudorange observation-based detection, and carrier phase-based detection [42–45].

In the mentioned literature, the fundamental approach of the detection algorithms is based on the scenario where all the authentic satellite signals are spoofed. However, with the recent improvements in the receivers' sensitivity and the availability of high antenna gains [46], receivers are able to obtain unspoofed signals even if the majority of the positioning signals are under attack. This fact creates more space and opportunity to detect spoofing attacks with the available side information coming from the legitimate positioning signals transmitted from the satellites. Consequently, an inherent advantage appears for the detection mechanisms due to the existence and availability of unspoofed positioning signals in the wireless medium.

In our system model, we also consider multiple units of V2X, which may be RSUs or vehicles. Due to the different complexity and security levels, these agents can cooperate with existing V2X communication schemes. After running one of the proposed spoofing detection algorithms on their OBUs, these agents can share their spoofing detection decision. In the end, the proposed system model enables high-level hybrid decisions, which can be deployed with decision fusion methodologies.

1.2. Literature Survey

In the literature, FEC-based solutions are studied not only to satisfy reliability but also to enhance security in the physical layer, when Eve is a part of the network [6,34,47]. More importantly, in [48,49], FEC codes are compared with automatic-repeat-request (ARQ) schemes in CPS to ensure ultra-high reliability and low latency requirements. The results show that FEC codes are advantageous over ARQ in terms of latency and energy efficiency issues [48, 49]. As an improvement in security cooperative communication-based techniques can be integrated with FEC codes to enhance security in CPS, since there may be a large number of nodes in CPS [50]. Moreover, together with the advantages of the perfect code, FEC codes are energy efficient and provide low latency, making them suitable security solutions to be used in CPS [48]. As discussed in [51], relays can provide increased security in harsh communication environments, such as industrial deployments, and long-distance communications, such

as smart grids. Moreover, cooperative communication techniques are useful when a noisy transmission and high interference environments may cause deep fades in CPS, and a single transmission link may not satisfy ultra high-reliability target [52]. Also, the McEliece cryptosystem is another well-known tool to provide secure communications [33]. This process is based on error injections with respect to the error correction capability of the receiver. As a result, the security of the system can be improved by using these functionalities.

Various attacks and their countermeasures are studied and listed in [30, 53, 54] against the security of positioning signals. Due to the lack of security solutions in widely-used satellite positioning systems, e.g., the L1 signal in GPS, there are high risks for attacks against secure positioning. In the literature, jamming and spoofing are considered to be the principal attacks against GPS security [43]. These attacks can also be implemented with commercial off-the-shelf products [10]. Jamming is a significant attack to wipe out legitimate signaling in communication systems, and GPS is not an exemption. Due to the very low power level of GNSS signals, a jammer may generate signals with limited energy [55]. In GNSS, many systems only show the last available location after jamming without an alarm [43].

GNSS spoofing is in the group of authenticity/identification attacks in [54], whereas the rest of the attacks are availability attacks, confidentiality/privacy attacks, and data integrity/data trust attacks. In another source, false data injection is the primary attack class of spoofing in [30] in V2X communication systems, while denial of service (DoS) attacks and sybil attacks are the remaining significant groups of V2X attacks. It should be reminded that spoofing attacks against GNSS-based positioning can be easily implemented in SDR. Asymmetric cryptography and the public key infrastructure is the most common approach to provide security in V2X systems; however, this structure may not prevent GNSS spoofing. The reason is that GNSS signals for civilian use are unencrypted, and GNSS signals are the inputs of V2X communication systems. Therefore, the authenticity of GNSS signals should be separately evaluated.

The literature on spoofing detection is broad for communication systems without differentiating the applications due to diverse spoofing attacks [9, 12, 56]. A simple but effective attack is meaconing, during which the attacker listens and repeats the correct signals. As a result, the receiver vehicle may obtain expired data with the wrong time information. Another well-known spoofing attack is based on transmitting more powerful signals than authenticated GPS signals, especially during the acquisition process of the receiver. However, detection of these attacks is straightforward because the received signal power is very low in legitimate signaling. The reason for the low power level is the enormous distance between the satellite and the vehicle. On the other hand, spoofers are expected to be located closer to the receivers than the satellites; hence, there are significant differences in the order of signal powers. As a result, power monitoring is an efficient tool against these basic spoofing attacks [42]. From the attackers' perspective, attacking signal synchronization is also challenging during the receiver's acquisition process. In general, the spoofing signals should be locked to by the victim receiver to be considered a successful attack. In order to guarantee it, the attackers may first jam the legitimate signals. Then they may try to deceive the receivers with their spoofing signals after a restarted searching and acquisition stage. In this case, the chance of a lock to the spoofed signal is higher from the receiver's perspective. Among the advanced spoofing attacks, nulling may severely affect the receiver's perception. In this attack, the attacker generates two signals, the first one is the negative of the authentic signal, and the other one is a fake positioning signal. When the authentic signal is canceled, the receiver only obtains a fake positioning signal [9]. Another frequently used approach is the spreading code correlation-based signal quality monitoring [42]. When a spreading code is multiplied by itself, a peak can be observed due to the structure of the Gold codes and their coarse acquisition (C/A) code practices, which are used in standardized GPS technology. Otherwise, the multiplications tend to go zero. Due to phase shifts during the generation of the spoofing signals, there may be a distortion in correlation values. However, spoofers may not be detected with this approach if they successfully generate C/A with a low phase difference. Therefore, additional analyses are needed. In addition to these methods, there are more complex solutions, such as detection mechanisms with the aid of mobile

services, multiple antenna/receiver-based techniques, and direction-of-arrival (DoA), which are generally based on hypotheses tests [44,45]. DoA, which is based on carrier phases, and pseudorange differences are mostly studied for multiple antenna/receiver scenarios.

There is also some work on the focus of spoofing detection in vehicular communication systems. In [57], the authors focus on spoofing detection in vehicular systems with the help of Doppler shift measurement; however, this methodology requires at least two vehicles, contrary to our proposal. The authors also focus on the mobility of the spoofers in the vehicular system. In another work, Guo *et., al* study a covert spoofing algorithm for uncrewed aerial vehicles [58]. In addition, there are more complex solutions, such as mobile service-aided detection mechanisms, multiple antennas/receivers-based techniques, and DoA, which are generally based on hypotheses tests [44, 45]. DoA is mostly studied for multiple antenna/receiver scenarios on carrier phases and pseudorange differences. However, the literature broadly focuses on only legitimate signals or only attacker signals scenarios. Even if GPS receivers are small enough and can be closely placed inside a vehicle to increase accuracy, the receivers require additional users or antenna beams; hence, the applicability of these solutions may be limited. In the literature, multiple GPS receivers are already studied for autonomous cars and drone applications [59,60]. The deployments of multiple GPS receivers are easier in large-scale vehicles, e.g., commercial airplanes or ships.

The pseudorange-based detection strategies for multiple vehicles are significant for the development of our detection methodology since they have similar steps after obtaining pseudorange values. These techniques rely on the pseudorange differences of the received signals at different vehicles or antennas, whereas the most significant steps are taking single and second differences. Hence, many detection mechanisms are directly correlated with these operations [12,20,21,61–64]. The majority of these papers focus on the spoofing detection technique based on pseudorange differences for multiple receivers scenarios, while the proposed methodology does not require additional receivers due to inversely utilizing the single difference operation. In the literature

on solving hyperbola equations, [65] provides the essential steps to use in positioning systems based on base stations.

2. THE SECURITY NEEDS OF V2X BASED ON A GARDEN OF CPS

CPS, including V2X systems, are the integration of computational (cyber) and physical components with a feedback loop to use in various applications in daily life or industry as an application of the Internet of Things (IoT) networks. In this architecture, the cyber part is composed of computational, communication, and control units, while the physical part consists of sensors and actuators [4]. In the cycle, the actuators may receive commands from control units via wireless communication channels and operate based on the received instructions. The sensors measure the physical environment and monitor the changes. They inform the control units about their measurements over wireless communication channels as well. In the cyber part, control units compute new commands based on the information they capture from the sensors. Finally, they send new commands to the actuators; as a result, the cycle, as shown in Figure 2.1, is completed.

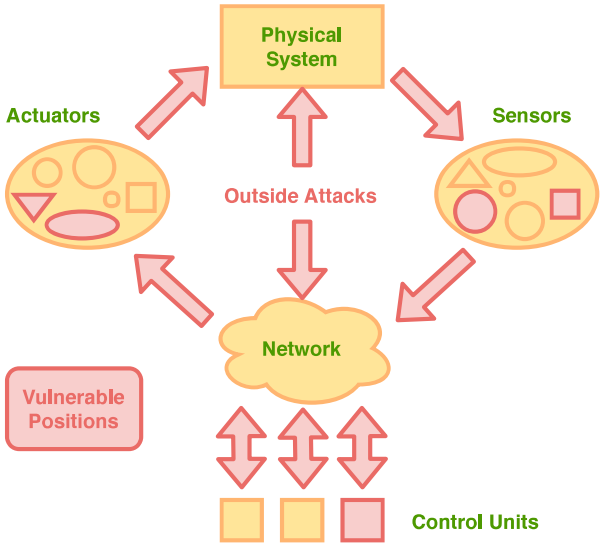

Figure 2.1. A simplified cycle of CPS. The vulnerable positions in CPS are indicated.

With the integration of cyber and physical components with this closed loop, CPS offer a wide range of applications in our daily lives and industrial fields. To

deploy these applications, a diverse set of requirements and challenges should be addressed. As already mentioned, their popularity also introduces severe security threats due to the inherent vulnerabilities of the wireless transmission environment. In this chapter, we focus on the specific requirements and challenges of the main applications of CPS, e.g., industrial wireless sensor networks, smart grids, V2X communication, and smart healthcare, from a security perspective. Elaborating on the 5G flower concept, we extend CPS applications to a garden. We also explain the fundamentals of implementation issues to these applications in real life.

At first, we focus on the visualization of these requirements and challenges concerning the security need. This visualization can be in various shapes (e.g., spider diagram) as an extension of the *5G flower* concept, which appears in [66] to explain the necessities of the distinct fields of the 5G networks. Since 5G and its successor networks will establish the central communication infrastructure of CPS in the near future and beyond [67, 68], we can likewise draw the up-to-date flowers based on the specific needs of CPS. In the original *5G flower* illustration, the flower consists of petals and leaves, symbolizing fundamental requirements of 5G in terms of connectivity, data rate, latency, and challenges, such as efficiency [66]. Another approach is showing these requirements in a spider diagram, as shown in [68], for possible 6G applications in the scope of CPS. However, neither of these illustrations is able to capture the security perspective of CPS.

In the next section, we introduce a garden of the CPS, which consists of the *CPS flowers* in different forms due to various applications with distinct requirements and challenges, including security issues with several metaphors. Our goal is to illustrate CPS in a unified model including all the design aspects, to achieve a robust, secure, efficient, and functional CPS. Due to the variety of actual and future CPS applications, our proposed garden can be easily extended.

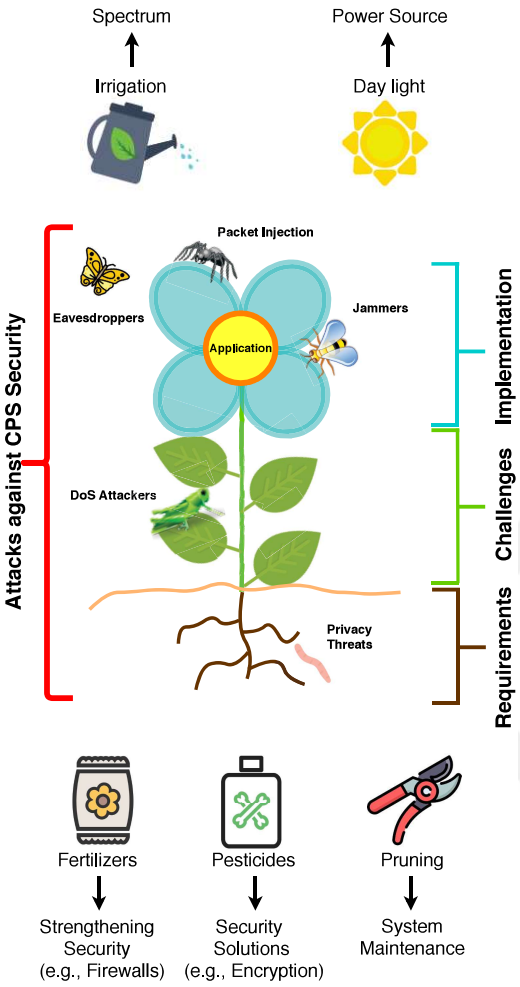

Figure 2.2. CPS flower, which shows requirements, challenges, and implementations with some external factors that help to satisfy efficient, robust, and secure CPS deployments and possible security threats.

2.1. Flowers of CPS

In our CPS flower representations, roots indicate the fundamental requirements of the chosen application, leaves represent the CPS challenges, and petals show the implementation aspects, as shown in Figure 2.2. When we focus on the security concept, there are vulnerabilities in each part of CPS. As a result, attacks and threats may target these distinct parts. The attacks and threats are symbolized as insects, whose species may increase, as there are a variety of attacks against secure CPS. In this figure, there are also external factors that affect system performance. The sun and irrigation

represent the power source and available spectrum, respectively, since CPS need a sufficient spectrum. If we increase the available spectrum, depending on the ambient traffic interference may become stronger. On the other hand, there may be no room on the spectrum for reliable and efficient communications. Other concepts, including fertilizing, using pesticides, and pruning can be considered the fundamental procedures (e.g., improving efficiency, satisfying security, and updating systems) to provide reliable and functional CPS operations. In this structure, fertilizers assist improving the plant's health while increasing resistance against bugs. When we consider CPS, this analogy can be interpreted that fertilizers are the elements of CPS to enhance system security as the first layer, (e.g., firewalls, antivirus software). However, these solutions may not satisfy advanced security requirements. Therefore, we may need pesticides to avoid and eliminate all the malicious cyber physical attackers. These pesticides can be interpreted as encryption techniques in general, as shown in Figure 2.2. Finally, pruning, which is the last metaphor, is considered system maintenance. The readers will agree that CPS should be properly updated with all of the parts to provide that security mechanisms are functional and sufficient.

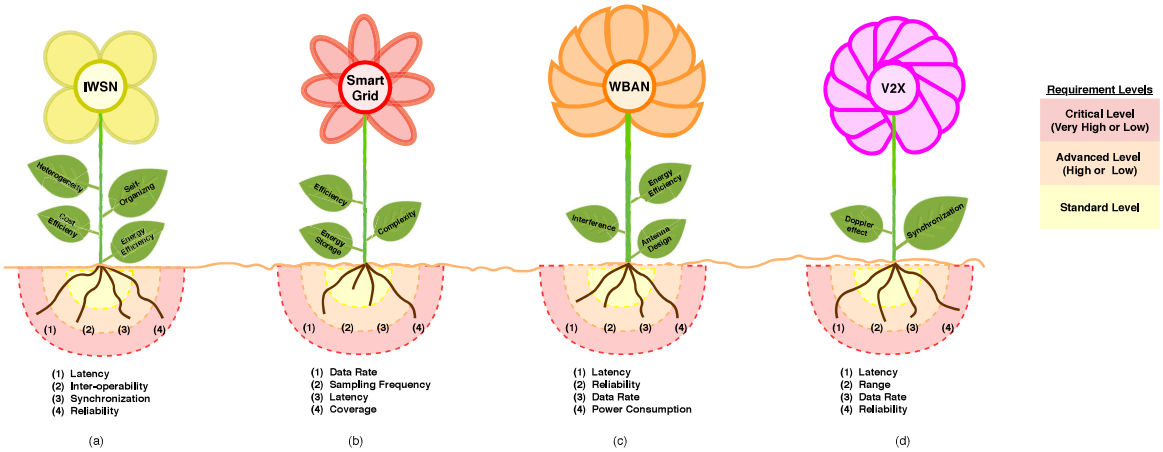

Figure 2.3. Flowers of the CPS applications. Their leaves and length of roots indicate the specific requirements of the application: (a) IWSNs; (b) smart grid; (c) smart healthcare; (d) V2X communications.

In the scope of this section, we focus on industrial wireless sensor networks, smart grids, smart healthcare, and intelligent transportation, although there are many

possible options to implement. Consequently, four different flowers are illustrated for determined applications in Figure 2.3. In this section, we will explain these flowers based on the requirements and challenges of the corresponding CPS application.

2.1.1. Industrial Wireless Sensor Networks (IWSN)

IWSN is a unique and emerging class of wireless sensor networks (WSNs), which is a distributed system of autonomously interacting devices and sensors within the scope of the Industry 4.0 concept. The main target of the deployment of Industry 4.0-based systems is to interconnect and automatize the traditional industries instead of human-centric wireless communication systems with respect to the CPS. In IWSNs, there are sensor-based control and communication units, and each device may autonomously determine its operation, which leads to low cost, fast, reliable, and efficient production processes.

Wireless solutions in factory automation systems have several advantages over wired counterparts in the same environment. Installation and maintenance costs of IWSNs are low when compared to wired solutions, whereas deployment and extension of wired industrial networks are more challenging [69]. Wireless sensors can be deployed on mobile robots or rotating parts of machines. On the other hand, there are some particular requirements, and challenges of IWSNs as visualized in Figure 2.3(a).

In industrial areas, the main challenges arise because of transmission medium characteristics. There may be heavy machinery and several metal surfaces in the factory environment, and these may lead to increased multipath effects, fading, and extreme noise/interference conditions [69]. Additional major challenges of IWSNs are the synchronization of nodes and self-organizing [70], especially in harsh environments in crowded networks. Energy efficiency, which is challenging in many cases, should be carefully considered when there are many power-limited sensor devices. There may be various devices supporting different communication standards, which may lead to an interoperability and heterogeneity problem.

In IWSNs, the most problematic requirements and performance metrics are latency and reliability. Many sub-applications of IWSNs require very low end-to-end latency or very high reliability, while sometimes both constraints must be satisfied at the same time [69]. The challenging delay and reliability requirements are represented with very long roots in Figure 2.3(a). IWSNs also should tolerate communication faults due to a harsh communication environment and should be highly interoperable to run with various devices using different standards [70]. A tight synchronization is needed in case of super crowded networks, which consist of thousands of nodes [69]. Overall high Quality of Service (QoS) is demanded in IWSNs based on the satisfaction of customers [69]. This need is highly based on the deployed application and the network model. Latency conditions may be also vital, and in other circumstances, reliability or scalability may be more critical [70].

2.1.2. Smart Grid (SG)

Due to the inefficiency of conventional power networks and the transmission environment, a change in the power distribution is necessary for cost and energy efficiency. SGs are considered to be used in the near future utilizing CPS to overcome performance issues [71]. In SGs, power plants, which may be conventional or renewable energy-based, and consumers, such as; residential, commercial, and industrial users, are connected in a closed-loop model. In this loop, users measure their power demands via sensors and inform power plants with a communication system. They generate the necessary amount of power to serve the need of users while increasing the efficiency of power distribution. In this section, the flower of SG, which is shown in Figure 2.3(b), is shaped by the following requirements and challenges with security concerns.

SGs are very large and contain some unique requirements based on their operations. The desired latency can vary from sub-seconds to seconds for monitoring, metering, and controlling [72]. Therefore, a medium-length root is placed in the corresponding flower. In addition to latency, $10 - 100$ kbps bandwidth can be supplied in the home area network (HAN), and up to 200 Mbps bandwidth can be achievable

in meter local area network (MLAN) for metering and pricing [72]. In SGs, ultra-high connectivity is demanded, since there may be hundreds of nodes in the system [73]. Any link disconnection between a power plant and a customer can cause a blackout of electricity on the customer side. Or more seriously, large-scale blackouts may be triggered [2]. Network coverage, which may vary from $50m$ to a few kilometers depending on the chosen standard, is also a fundamental design issue shown with a large root. Data frequency sampling can be hourly in a day, which is a relatively low sampling rate and it is not a challenging requirement. Overall, a high QoS is typically needed to perform long-term operations and to satisfy customers alike to the IWSNs.

To deploy SGs with a high QoS, there are some challenges, as shown in Figure 2.3(b). Firstly, the SGs will be very complicated since they are composed of an excessive number of nodes and modeled as large-scale systems. As each meter operates as a router, there is an inherent complexity problem [72]. Other challenges include self-organizing and self-healing structures of SGs [72]. They are conceptually powered by renewable energy sources, and they can also directly transmit electricity to customers. However, the deployments of these systems are expensive; therefore, they should perform for many years and must be profitable for the cost-efficiency concept. Finally, the storage of the generated power should be considered to provide reliability and control to the networks [73]. Further, the balance between the distribution and the storage of energy should be satisfied.

2.1.3. Healthcare Applications

This class of CPS application is the integration of wearable or implanted devices in the human body with a control center. In these systems, doctors can remotely monitor patients' real-time medical data in a hospital or a control center. Another possible application is the robotic surgical systems, which are supervised by doctors. Deploying these solutions to tackle fatal diseases, such as cancer, diabetes, and cardiovascular diseases can enable prevention and early diagnosis [74]. These technologies will be increasingly used in hospitals and other medical centers to improve the health

of patients. They also have some particular requirements to be satisfied. All of these are presented in Figure 2.3(c) and we will explain the details of this figure below.

Sensor devices play the most critical roles in medical applications compared to other CPS applications. They may be implanted into the human body; therefore, they should be very small and highly energy efficient to operate for the long term. Sensor networks in medical CPS can be studied as a wireless body area network (WBAN) [4]. In the literature, WBAN is standardized as IEEE 802.15.6, and the fundamental requirements and challenges of this standard will be overviewed here. The data rate is between 1 kbps and 10 Mbps, as shown in Figure 2.3(c) [74]. Latency should be less than 125 ms for WBANs, whereas the jitter should be less than 50 ms [74]. Reliability can be evaluated as the bit error rate in WBANs, and it should be less than 10^{-3} for the majority of the designs [74]. Since energy efficiency is the inevitable requirement in smart healthcare solutions, desired power consumption should be less than 10 mW to operate for weeks. The devices should consume less than 0.1 mW to perform for years [74]. When these requirements are satisfied, a high QoS can be achieved in the medical applications of CPS.

In the scope of medical CPS, there are two unique challenges, which are interference management and antenna design. For interference, two types of interference can be observed, and these are cross interference and mutual interference. Cross interference exists due to other networks that use the same frequency, while mutual interference emerges from the neighboring WBANs [4]. The sensors may be too close to each other, and it decreases the transmission quality significantly [74]. The other issue is the antenna design, where there is an electromagnetic interaction between the antenna and the human body [74]. Textile and flexible antennas are also needed use in WBANs. Similar to the other CPS, energy efficiency, scalability, and interoperability are required in medical sensor-based applications [4, 74].

2.1.4. V2X

Finally, we focus on intelligent transportation systems in the scope of major CPS applications. In this class, elements of transportation, such as cars, traffic lights, and road signs, communicate with each other. This concept is known as V2X communication. In this concept, cars adapt their speed and position without any intervention from drivers (e.g., autonomous vehicles). They also can assist drivers according to their abilities, such as lane assist, cruise control, and sudden braking. Besides, cars can communicate with pedestrians to prevent accidents. Due to their specific nature, V2X systems have their requirements and challenges in addition to security needs. These conditions are shown in Figure 2.3(d). Especially in a city scenario, a fusion control center can coordinate the trajectories of multiple vehicles.

While dealing with the lives of drivers and passengers, we expect to have very high QoS in smart transportation systems. Latency should be less than 20 ms for some sub-applications, and reliability should be higher than 99% [31], whereas they are represented as long roots. The data rate may fluctuate from a few kilobits per second to $20-25$ Mbps according to the application. This model may be an autonomous driving system, road assistance, a traffic management service, or an alarm to pedestrians or cyclists [75]. In some cases, the effective distance, which may not be the actual distance due to the relative speed of vehicles, is $150m$ for urban areas and $320m$ for freeways [31]. However, a broad communication range is needed for autonomous driving, such as $1000m$ [31]; as a result, the distance is very critical in V2X deployments.

When we focus on challenges in smart transportation systems in CPS, collision, Doppler spread and synchronization are the principal concerns. In the literature, IEEE 802.11p targets V2X communications. It is possible to realize V2V communication systems with this standard while satisfying low latency and being robust against the Doppler effect. But it leads to high congestion due to its random access scheme for high traffic scenarios [75]. Another option, mmWave communication, which increases the network capacity at 60 GHz central frequency. In this case, the Doppler spread is

very high at this frequency compared to the $2-6$ GHz [75]. Due to the many moving nodes at high speeds in V2X communication, synchronization turns into a challenging obstacle since many nodes join or leave the network in a very brief period. As a solution, GNSS can be used with high accuracy for node synchronization [31].

2.2. Security Aspects

Security is an absolute challenge in CPS since there is no advanced security shield yet. Due to the varying CPS applications, there are several security risks, attacks, and protection mechanisms. In this section, we will explain the significant concerns about security in CPS regarding four leading applications with respect to Table 2.1.

Table 2.1. Security perspective of the chosen CPS applications.

		CPS applications			
		IWSN	Smart grid	WBAN	V2X
Security aspects	Attacks	Eavesdropping, traffic analyze, data manipulation	Privacy risks, spoofing, DoS, jamming	Privacy risks, data modification, masquerade	Eavesdropping, DoS, spoofing, falsification, privacy risks,
	Solutions	Lightweight encryption, spread spectrum techniques	Data anonymization, tamper resistant meters	Lightweight encryption	Hardware-based security measures, key shielding

2.2.1. IWSN

In IWSNs, there are many vulnerabilities, especially in the communication units [69]. Cryptographic solutions for security enhancements may not be possible due to the limited processing capabilities of sensors. As a result, secure IWSNs should satisfy both security objectives and communication requirements. Some critical security risks of IWSN appear due to physically vulnerable wireless channels. The intruders may capture, analyze or manipulate the communication between users [2]. Similar to these attacks, an eavesdropper, a passive attacker, may also monitor established

communication [69]. Strong encryption procedures, spread spectrum approaches, locating intruders, and several other methods are introduced as countermeasures. But there are many drawbacks, such as increased power consumption, complexity, or latency. Physical layer security is another countermeasure, where the security can be enhanced directly in the physical layer. This countermeasure may decrease the energy consumption and the latency [76].

2.2.2. Smart Grid

In SGs, there are specific threats, vulnerabilities, and attacks against secure operations, and the critical security requirements are CIA, authentication, nonrepudiation, access control, and accountability [72]. In a home environment, sensors are wireless devices and vulnerable to very straightforward attacks [73]. One foremost threat is the privacy of individuals in SGs. Malicious users may be aware of the consumption model and daily rituals of the customers since their electricity demands of them can be understood from metering data [2, 72]. As a result, private information may be violated or captured during the transmission phase [73]. An adversary may modify the measured data of electricity usage. This modification may cause economic losses for the company or the customers [2]. In the cyber component of SGs, packet flooding attacks, resilience to DoS attacks, and spoofing attacks, which may lead to service delay or performance degradation, are also possible along with jamming attacks [2]. Data anonymization, networking-based, and encryption-based techniques can be performed to overcome the privacy and security risks [72]. Advanced data collection techniques can be studied without revealing private information [72]. In the physical environment, tamper-resistant meters should be applied against tampering attacks.

2.2.3. Healthcare Applications

As these systems operate with the medical data of patients and their locations, the most significant threats commonly target users' privacy. The attackers can eavesdrop or monitor the data transmission on the communication channel to violate patients'

privacy [4]. Furthermore, medical records are stored in different locations, and the attacker may target capturing these data [2]. Beyond privacy, secure deployments should be applied against attackers. In medical CPS, an attacker may modify data (data modification attack), change legitimate nodes with fake ones (masquerade attack), and totally or partially paralyze the network with a DoS attack [4]. An encryption-based solution can be applied, but it should be lightweight and energy efficient due to the constraints of sensor devices. In WBANs, physical layer privacy solutions can also be attractive for future research due to their application in the first two layers of the communication model.

2.2.4. V2X

When we deal with smart transportation and V2X communication systems, security is remarkably significant. In the absence of an advanced security mechanism, an attacker can quietly capture the private location information of users. More importantly, they can hack an autonomous car, which may lead to severe consequences [4]. Privacy leakages may appear especially for the location data of users. Attackers may manage various attacks to violate security, including DoS, packet injection, or falsification [2]. One significant attack in V2X systems is jamming. The jammers can distort the control information generated in a control unit for the coordination of multiple vehicles in a city scenario, and the sensor information from external sensors [4]. Physical layer attacks on the traffic elements are also significant threats to V2X communications. If some of the links are broken anyhow, for example by natural disasters or external attacks, V2X systems tend to be dysfunctional. Therefore, the system may not work precisely possibly leading to severe consequences [2]. As a solution, hardware-based security procedures and key shielding mechanisms can be executed to provide robust V2X deployments [2]. Another critical attack is spoofing against secure V2X deployments. As discussed before, spoofing attacks against authentic GNSS positioning signals can disable whole V2X networks, because positioning information in V2X networks is irreplaceable. Since we have already discussed the impacts of spoofing attacks and some of the solutions before, we prefer to continue with the implementation

aspects of the considered CPS applications.

2.3. Implementation Aspects

As CPS are highly application-based real-life concepts, their implementation issues should also be discussed. There are varying requirements according to the application; therefore, corresponding standards, network topology, and system models can vary significantly. Existing or future implementations are also given for each chosen class of CPS along with the standards and network specifications. The overview of the implementation aspects is given in Table 2.2 with respect to considered applications.

Table 2.2. Implementation perspective of the chosen CPS applications.

		CPS applications			
		IWSN	Smart grid	WBAN	V2X
Implementation aspects	Standards & organizations	Bluetooth, ZigBee, ISA100.11a, WirelessHART	WiMAX, ZigBee, MobileFi, PLC, IEC 61850,	IEEE 802.15.6, ZigBee, UWB, Bluetooth,	IEEE 802.11p, NGMN, 5GAA
	Network type	Mesh or star topology	HAN, NAN, WAN	Single or multihop networks	5G V2X RAN, LTE, mmWave
	Exiting & future applications	UPS' ORION, LLVs	AMI, BPL, SMI	CodeBlue, UbiMon, LifeGuard,	CACC, lane merging, automated parking

2.3.1. IWSN

IEEE 802.15.4 (Bluetooth), ZigBee, WirelessHART, and ISA 100.11a [69, 70] can be used to deploy IWSNs. Both star and mesh networks can be used during implementation, but each has its advantages and drawbacks. As an example, star networks provide lower latency, while mesh networks ensure higher reliability. Time division multiple access (TDMA) has several benefits for multiple accessing in IWSNs, such as energy efficiency [77]. Some principal requirements (e.g., reliability, latency) can be satisfied with TDMA-based solutions. However, synchronization and efficient slot allocations are open issues in TDMA-based IWSNs [77]. Modulation and coding schemes should be carefully selected to generate robust signals, which should survive

in harsh and high-interference factory environments [69]. The ongoing application of IWSNs is the postal services, where IoT and CPS systems will boost the process efficiency and critically reduce the service time. The offered solutions include UPS' ORION (On-Road Integrated Optimization and Navigation).

2.3.2. Smart Grid

Since SGs consist of many subbranches such as power engineering, communication, control, and information technologies, there are many standards in literature, such as IEC 61850 for the substation automation processes [4] or ZigBee, IEEE 802.16 (WiMAX), IEEE 802.20 (MobileFi), and PLC for the communication of smart meters [72]. When we focus on the network models, advanced metering infrastructure (AMI) and smart metering infrastructure (SMI) are the essentials during the deployments of SGs. The collection and analysis of the power usage in real-time are the main functions of AMIs, in addition to establishing the connection to measuring devices, [4]. Thanks to AMIs, two-way communication between smart meters is possible. These types of networks can be modeled as a HAN, neighborhood area network (NAN), or wide-area network (WAN). The gathered data by AMIs may be forwarded to fixed networks such as Broadband over Power Line (BPL) for pricing and demanding more energy from power distributors.

2.3.3. Healthcare Applications

As we already discussed, WBANs offer a great variety of medical applications. In [74], a list of actual applications of WBANs, and this list is expanding every day. In addition to the IEEE 802.15.6 standard, there are other standards for the medical deployments of CPS, such as ultra-wide band (UWB) technologies, ZigBee, 802.11.b, and Bluetooth Low Energy [4, 74]. The sensor types should be different based on the monitoring application. The WBAN networks are mostly connected to the Internet for long-distance communications. In these systems, both single hop or multihop networks are possible concerning application type. This feature affects the performance

of medical cyber-physical systems (MCPS). In [4], some existing medical applications, such as *CodeBlue, UbiMon, LifeGuard* are listed in detail. In some cases, open security problems, which should be addressed, arise.

2.3.4. V2X

Today automation is one of the key components in the automotive industry, and many manufacturers put on the market their vehicles from the most economical class with superior technologies like autonomous park assistance, emergency brakes, or maneuvers. Among the current research topics, cooperative adaptive cruise control (CACC), lane merging, and connected automated parking are well-known scenarios [75]. For large-scale deployments, LTE-V2X systems are already tested with the aid of the Next Generation Mobile Networks Alliance (NGMN). Another organization is the 5G Automotive Association (5GAA), which offers solutions for connected mobility, road safety applications, remote access, and integration with smart cities and smart transportation [31]. As already mentioned, two main technology are already accepted to deploy V2X in common. These are DSRC or LTE-V2X technologies [30], and it should be noted that mentioned security-enhancing mechanisms for V2X systems are suitable to be deployed with DSRC or LTE-V2X technologies.

Cyber-physical systems are application-oriented, and they have their control, communication, and computation units in closed-loop mechanisms. In this chapter, we have analyzed four main applications of CPS, which are IWSNs, smart grids, WBAN, and V2X, with the concept of a CPS garden. These flowers can be considered as the design guides for the corresponding application. Due to the recent spoofing-related security incidents as mentioned earlier [9, 10], the security of the V2X application from the mentioned CPS application group gained attention. In the rest of the thesis, we focus on V2X networks and their security against chosen passive and active attacks as a main research area.

3. GNSS FUNDAMENTALS WITH SPOOFING THREATS

Before going into details of attacks against the security of V2X systems, we prefer to give an overview of satellite-based positioning technologies since this overview presents the required information to understand Chapter 5. In this chapter, the fundamentals of GNSS structure are presented with the critical steps of transmitting, receiving, and positioning with the help of pseudorange values. In the scope of this thesis, we extend the definition of pseudorange values by adding the mobility term of the target receiver as a part of the V2X problem. Similarly, we present the pseudorange expression under a spoofing attack with possible spoofing imperfections due to the challenges of deploying such an attack, e.g., synchronization. Beyond that, the updated pseudorange expression for a spoofing attack also includes atmospheric errors because of the extended system model with aerial spoofers. Since there are many parameters with these additional error terms, the generalized notation is listed in Table 3.1 to provide ease of readiness.

3.1. The Fundamentals of GNSS

Today, various types of GNSS-based communication are used with different satellites. The most popular is GPS program, which has been started in the 70s by the USA. This system has currently 32 satellites, which run Standard Positioning Services (SPS) for public use, and Precise Positioning Service (PPS) for mostly military-based usage. The orbits are arranged so that at least 6 satellites are visible from everywhere on the Earth's surface. It offers encrypted signals for military use and non-encrypted signals for civilian use. In this proposal, we principally consider GPS as the primary platform, and the operational details, which are given in the rest of this proposal focus on GPS platforms.

Beyond GPS, other satellite communication systems are still in use. The first one is GLONASS, developed by the Soviet Union during the cold war era, and it

is now operated by Russia with 26 active satellites in orbit. It uses FDMA-based communication instead of the CDMA-based structure of GPS. Similar to GPS, it offers services for civilians in addition to military use services. The second one is Galileo, which is a joint program of the European Union, but the main contributors are Germany and Italy. This program has 24 satellites in orbit, but it is still not fully operable. This program is planned to offer open services for public use, paid high-accuracy services, encrypted services for civilians, and search and rescue services. Last but not least, BeiDou is run by Chinese authorities with global coverage with 33 active satellites, and they claim that this system will be the most accurate one compared to other systems with millimeter-level accuracy. The main structure of the satellite communication systems is very similar, and they have 4 segments, which are ground, control, space, and user segments. The ground units, such as monitor stations, ground antennas, and control centers, are called the control segment of GNSS. Another division is the space segments, consisting of a constellation of satellites, which send their ephemeris data to users.

As discussed, GPS, GLONASS, BeiDou, and Galileo are well-known positioning services as a part of GNSS. These technologies may have significant differences in their operations, which are generally in coding schemes, frequency, and accuracy levels, as well as their ground units. However, the basics of the positioning idea remain the same [78], whereas a pseudorange should be calculated from the received data. In the standard GNSS-based positioning process, pseudoranges are different from the exact range between the available GNSS satellites and the receiver due to long-distance and signal travel time. The pseudoranges are the observables from the transmitted satellite ephemeris data. At the receiver vehicle, pseudoranges can directly be calculated from the received GNSS signals with the broadcasting satellite ephemeris information, and then these values are responsible for position calculation by using the least-squares approach. Due to this similarity in utilizing pseudorange values, we focus on GPS as the primary positioning source in V2X in this thesis in order to develop a spoofing detection tool. This system tends to suffer from spoofing due to unencrypted signals (e.g., L1 signals in GPS) [43], while many open-source dictionaries of GPS are available

to build a solid detection algorithm.

3.2. Transmitting and Receiving Units of GPS

There are different versions of GPS signals, which can be used in civil and military applications with various levels of security mechanisms. The GPS signals include 3 components, which are carrier wave, navigation data, and spreading sequence [79]. Even if there are 3 main types of GPS signals, which are L1, L2, and L5 signals, there are also sub-classes of these types of signals. The carrier wave is a sine wave with $f_{L1} = 1575.42$ MHz for the L1 channel, whose fundamental frequency is $f_0 = 10.23$ MHz. The L1 signal is mostly used for civil applications and there is no encryption for this signal. The $L2$ signal is transmitted with a carrier frequency of $f_{L2} = 1227.60$ MHz, and it has the same fundamental frequency. The latest pattern of GPS is using L5 signals for aviation applications with the carrier frequency of $f_{L5} = 1176.45$ MHz. Navigation data consists of the information about satellite orbit, and its data rate is 50 bps [79–81].

3.2.1. GPS Signal Generation

There are two types of spreading sequences that are used in GPS, which are coarse/acquisition (C/A) codes and precise (P(Y)) codes. These codes are unique for each satellite Therefore, this feature can be used for the identification of the satellites. The C/A code has 1023 chips and it repeats each millisecond at 1.023 Mbps chip rate [80]. This code is only used in L1 signals. On the other hand, P(Y) codes can be used in both signals, while it also offers encryption-based security for L2 signals. Hence, it can be used in anti-spoofing (A/S) mode. It has a much longer code with 23500 chips with a chipping rate of 10.23 Mbps. [79,81]. Beyond that navigation signals are transmitted very slow rate of 50 bps.

During GPS signaling, the main feature is that the power level of the transmitted signals is less than the noise levels due to spread spectrum codes, C/A and P(Y)

codes [79]. Therefore, this type of communication can be considered as covert communication in the literature. This covertness of GPS signals can be evaluated as the most basic covert communication scheme due to spreading that is similar to CDMA-based communication [79,82]. Since we only use C/A and P(Y) codes, as shown in the literature without a change, we omit to give the details of the code generation process.

3.2.2. GPS Receiver

As shown in Figure 3.1, the receiver has several steps before finding its location, and these steps will be explained briefly with their contribution to security. In order to complete the positioning at least 4 GPS signals are required due to the used trilateration techniques for 4 unknown parameters for positioning. In this case, trilateration is used with pseudoranges for at least 4 visible satellites to locate the GPS receivers with the position, velocity, and time (PVT) information. The incoming "raw" GPS signals are

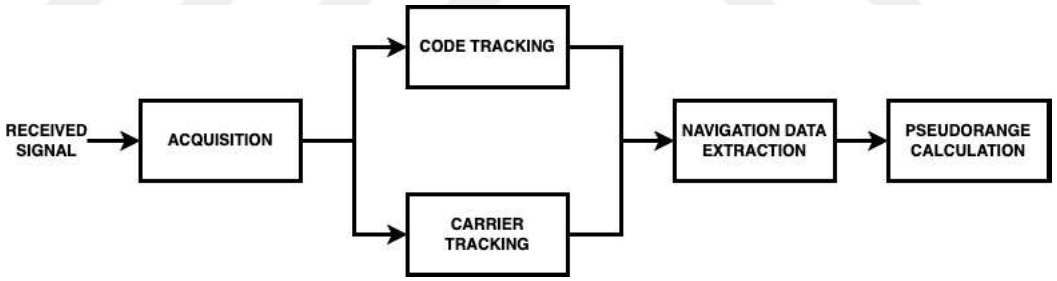

Figure 3.1. The block diagram of a GPS receiver.

transmitted as the combination of navigation data, carrier wave and spreading codes, and demodulation and despreading should be done to obtain navigation data [83]. GPS signal can be written as

$$y(t) = \mathrm{Re}\Big\{ \sum_{i=1}^{N_A} A_i D_i[t-\tau_i(t)] C_i[t-\tau_i(t)] e^{j[w_c t - \phi_i(t)]} \Big\} \tag{3.1}$$

where N_A is the spreading-code-specific signals, in other words, the number of authentic satellites, A_i is the carrier amplitude of ith signal, $D_i(t)$ is a data bit stream of the ith signal, $C_i(t)$ is the spreading code, $\tau_i(t)$ is the code phase of the ith signal, w_c is the carrier frequency, and $\phi_i(t)$ the ith beat carrier [9, 12]. Following steps should be

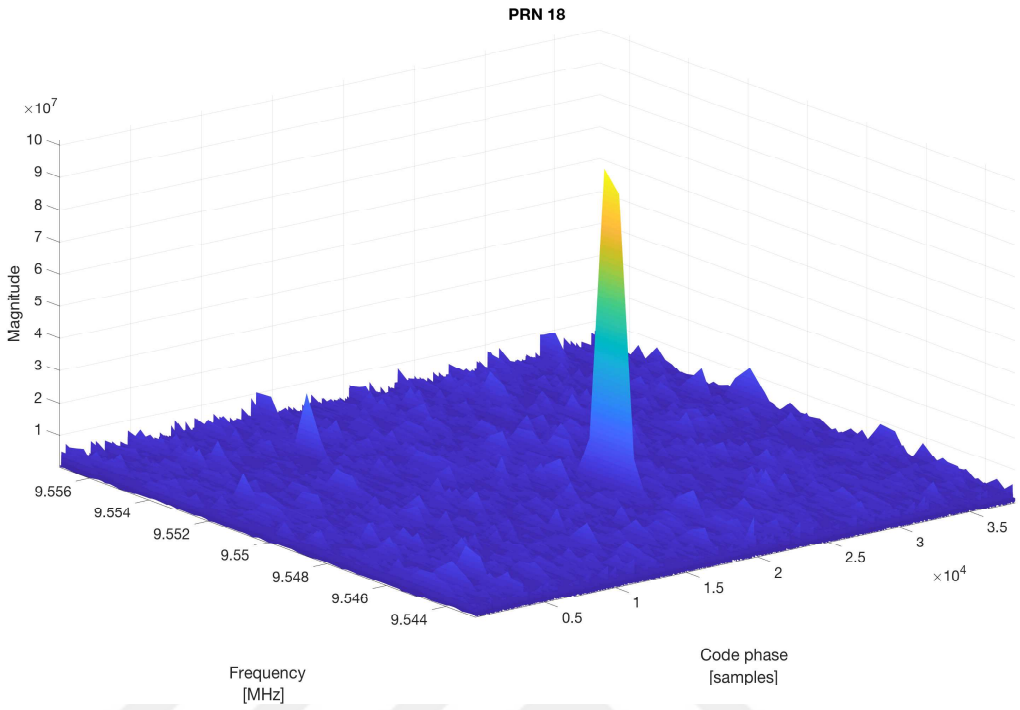

Figure 3.2. C/A correlation peak of the unspoofed signal during the acquisition process.

completed to find the position of the receiver. These processes are signal acquisition, tracking, and navigation data extraction, which are explained in details as

(i) The acquisition process is necessary to clarify which satellites are visible, then the frequency and the code phase should be available at the receiver side after this process [79]. In the signal acquisition process, the C/A code correlation properties are very significant. As shown in Equation (3.1), the received signal is the combination of N_G visible satellite signals. During the acquisition process, $y(t)$ is multiplied with the C/A code of kth satellite to acquire this satellite. Since the cross-correlation of C/A codes of different satellites is nearly zero, there is a peak for the acquired satellite in the frequency-code phase plane, as shown in Figure 3.2 [79]. In the literature, there are 2 main techniques that use correlation of C/A codes, which are serial and parallel search algorithms. The advantage of serial search is its implementation simplicity for hardware, while it takes too

much time for software-based analyses [83].

(ii) After the acquisition, the next step is tracking to reduce the effects of time variations. The accuracy of the code phase is highly correlated with the estimation of the pseudorange value, which is very significant in calculating the receiver's position, as shown in the next sections [12,79]. There are two parts of tracking, which are code tracking and frequency/phase tracking. Code tracking has a closed-loop structure; a delay lock loop (DLL), and three local code replicas are generated and correlated with the incoming GPS signal. Frequency/phase tracking can be performed separately for either frequency or phase. Tracking continuously follows the incoming signal. If the track of a satellite signal is lost, the acquisition should be repeated [12,79].

(iii) After a successful acquisition and tracking, the C/A code and the carrier wave can be extracted from the received signal to obtain navigation data. These data bits include ephemeris data consisting of satellite position, satellite clock bias, and other information broadcasted by satellites [12,79].

3.2.3. Pseudorange Calculation

Overall, finding the vehicle position is highly based on the pseudorange values, which can be observed with the help of ephemeris data. In the process of the calculation of the pseudorange, two time-related expressions, which are the receive time of the ith vehicle, t_i, and the signal transmission time from the kth satellite, t^k, are required. Due to this time difference, the pseudorange term is different from the true range between the transmitter (a satellite or a spoofer) and the vehicle in satellite-based positioning systems. Hence, the pseudorange is fundamentally calculated as

$$\rho_i^k = c \times (t_i - t^k) \tag{3.2}$$

$$t_i = T_i + \tau_i \tag{3.3}$$

$$t^k = T^k + \tau^k, \tag{3.4}$$

where c is the speed of light, T_i and T^k are, respectively, the actual receive and the actual transmission time with respect to the atomic clock. The clock biases of the vehicle and satellite are τ_i and τ^k, respectively. After a rearrangement, the pseudorange can be rewritten as

$$\rho_i^k = d_i^k + c(\tau_i - \tau^k) + m_i^k + I_i^k + T_i^k + M_i^k + \epsilon_i^k, \tag{3.5}$$

where (x^k, y^k, z^k) is the kth satellite position, (x_i, y_i, z_i) is the vehicle position. d_i^k is the true range between the ith vehicle and the kth satellite as

$$d_i^k = \sqrt{(x^k - x_i)^2 + (y^k - y_i)^2 + (z^k - z_i)^2}. \tag{3.6}$$

Table 3.1. The list of significant parameters for pseudorange expressions.

Notation	Definition	Condition
N_G	Number of visible GPS-like signals	$N_G \geq 4$
N_A, N_S	Number of authentic and spoofing signals	$N_A, N_S \leq N_G$
J	Total number of vehicles and/or RSU	
i, j	Indices of target vehicles or RSU	$i, j \leq J$
k, l	Indices of authentic GPS satellites	$k, l \in \mathcal{N}_A$, $k \neq l$
s, p	Indices of spoofing signal source	$s, p \in \mathcal{N}_S$, $s \neq p$
n, u	Indices of a transmitter (an authentic satellite or a spoofer)	$n, u \in \{k, l, s, p\}$
d_i^n	Real distance between i^{th} receiver and n^{th} transmitter	
ρ_i^n	Pseudorange between i^{th} receiver and n^{th} transmitter	$\rho_i^n > d_i^n$
Δt^s	Additional time error due to spoofing	$\Delta t^s \in \mathcal{T}$
$\hat{d}^k$	Additional distance error due to spoofing with Δt^s	
m_i^n	Mobility error with variance σ_m^2 for n^{th} transmitter	$\sigma_m^2 \in \mathcal{P}$
s_i^s	Spoofing imperfection with variance σ_{si}^2	$\sigma_{si}^2 \in \mathcal{S}$
I_i^n	Ionosphere error with variance σ_t^2 for n^{th} transmitter	$\sigma_t^2 \in \mathcal{I}$
T_i^n	Troposphere error with variance σ_i^2 for n^{th} transmitter	$\sigma_i^2 \in \xi$
M_i^n	Multipath error with variance σ_{mp}^2 for n^{th} transmitter	$\sigma_{mp}^2 \in \zeta$
$\triangle$	Single difference operation	
$\nabla\triangle$	Double difference operation	

In Equation (3.5), the atmospheric error terms are I_i^k and T_i^k, which are ionospheric and tropospheric errors with distributions $\mathcal{N}(0, \sigma_i^2)$ and $\mathcal{N}(0, \sigma_t^2)$, respectively. M_i^k is the multipath error term, which is distributed $\mathcal{N}(0, \sigma_{mp}^2)$, and ϵ_i^k is the obser-

vational noise, which is omitted in the rest of the thesis. In addition to the standard definition, the impact of mobility is added to the pseudorange expression and it is denoted as m_i^k, and it is distributed as $\mathcal{N}(0, \sigma_m^2)$. The transmitted satellite ephemeris data provides τ^k, which is satellite clock bias, and the satellite position. In Equation (3.5), at least four pseudorange observations are required to calculate the three unknown parameters of the vehicle position and the receiver clock bias τ_i [79, 81]. Therefore, we assume that $N \geq 4$. In the end, pseudoranges are calculated from broadcasted GPS ephemeris data at the receiver side, and their value is highly dependent on time information at the receiver. In order to find the receiver position, the least-square method is generally applied after utilizing Taylor's expansion theorem with observed pseudorange values.

3.3. Security Threats Against GPS

Due to the lack of security solutions in most of the satellite positioning systems; e.g., L1 signal in GPS, there are high risks for attacks against secure positioning. In the literature, jamming and spoofing are considered fundamental attacks against GPS security [43]. These attacks can also be implemented with commercial off-the-shelf products [10]. Jamming is a basic attack to wipe out legitimate signaling in communication systems, and GPS is not an exception. In general, jamming signals create power noise in the system frequency, Therefore, the system units cannot use this frequency to communicate. Due to the very low power level of GPS signals, a jammer may generate signals with low power [55]. In GPS, many systems only show the last available location after jamming without an alarm [43]. There are also some jamming techniques by using GPS pseudorandom (PRN) signals as noise to pass antijamming filters.

When compared to jamming, spoofing is more advance. Since the nature of a spoofing attack is about the similarity of the authentic signal, spoofing signal models are presented very similarly to the authentic GPS signals in the literature [9]. The reason is that spoofing attacks should not be easily distinguished from authentic signals

to be considered a successful attack. There are several versions of spoofing attacks in GNSS; hence, their countermeasures are also varied [9, 12, 56]. One of the spoofing attacks is called a meaconing attack, during which the meaconer estimates the true signals, and repeats the collected signals. After a successful attack, the receiver may get old signal data with the wrong time information. Another basic spoofing attack is sending more powerful signals than legitimate GPS signals during the acquisition stage of the receiver. This type of basic attack can be easily detected with power monitoring [42], and it is hard to synchronize this attack with the receiver's acquisition process. Therefore, in some spoofing attacks, there are some general techniques to be locked by the receiver. The first one, the attackers may first jam the GPS signaling to force the receiver into the searching stage and then apply their spoofing attack. In this case, the chance of locking to the spoofed signal is higher from the receiver's perspective. Another strategy is that the spoofer may primarily reproduce correct signals with a very low power level, then it increases their power level slowly until successful locking by the receiver. After being locked, the spoofer may shift receivers' frequency, power level, and spreading codes.Among the advanced spoofing attacks, nulling may affect the receiver's perception severely. In this attack, the attacker generates 2 signals, one is a negative of the true signal and the other is a fake positioning signal. In this case, the true signal is canceled, and the receiver only obtains a fake positioning signal.

3.3.1. Spoofing Attack Model

Within the scope of the thesis, a malicious GNSS spoofer exists in the V2X communication model in our adversary model. The spoofer broadcasts attacking signals at the same frequency as GNSS signals (e.g., 1,575.42 MHz for GPS). These signals are assumed to have the same identifiers and correct coding format as legitimate satellites with higher power than authenticated GNSS signals [84]. Similar to legitimate GNSS signals, the spoofing signals can be captured by both the target vehicle and RSU at the same time. Therefore, various V2X elements may simultaneously detect the same spoofer after running the spoofing detection algorithm.

Our attacker model includes a spoofer, which does not have power or computational limitations with the knowledge of legitimate GNSS signal structure. It is able to receive authentic signals and then replicate them as a spoofing attack by adjusting the amount of time error Δt^s. On the spoofer side, the decoding and replicating errors of legitimate GNSS signals and hardware-related errors should be considered. Therefore, we have also modeled the spoofer with additional spoofing imperfections. The spoofing imperfection is shown as an additive error source on spoofed pseudorange values and the imperfections are modeled as Gaussian distribution with zero mean and σ_s^2 variance. The number of spoofing signals is denoted N_S, where $N_S \leq N_G$. The following GPS spoofing signal is sent after a successful acquisition by the receiver

$$y_s(t) = \mathrm{Re}\Big\{ \sum_{i=1}^{N_S} A_{si} \hat{D}_i[t - \tau_{si}(t)] C_i[t - \tau_{si}(t)] e^{j[w_c t - \phi_{si}(t)]} \Big\}. \tag{3.7}$$

In this expression, A_{si} is the signal amplitude for the corresponding spoofing signal, $\hat{D}_i(t)$ is the estimated data bitstream, $\tau_{si}(t)$, and $\phi_{si}(t)$ are the code phase and the beat carrier of spoofing signals, respectively. In the signal medium, the generated spoofing signal is very similar to the legitimate GPS signal.

As shown in Equation (3.7), they transmit fake data signals with the copied spreading codes at the correct frequency with small phase errors. After an attack, the received signal is the collection of the actual signal, the spoofed signal, and noise as

$$y_{tot}(t) = y(t) + y_s(t) + \nu(t), \tag{3.8}$$

where $\mathbf{H}$ is the $N_r \times N_t$ dimensional channel matrix and $\mathbf{y}$ and $\boldsymbol{\nu}$ are the $N_r \times 1$ dimensional received signal and channel noise vectors, respectively.

In case of the existence of a spoofing signal, we can rewrite Equation (3.2) for the spoofing signal as

$$\rho_i^s = \hat{d}^k + d_i^s + c(\tau_i - \tau^s) + m_i^s + s_i^s + I_i^s + T_i^s + M_i^s + \epsilon_i^s, \tag{3.9}$$

where d_i^s is the distance between the ith vehicle and spoofer, and $\hat{d}^k$ is the false location data added by the spoofer for the kth satellite due to additional time error added by spoofer Δt^s. In Equation (3.9), m_i^s represents the mobility impacts of the target vehicles, and it has distributed as $\mathcal{N}(0, \sigma_m^2)$. s_i^s represents the spoofing imperfections.

These imperfections may result from hardware-related issues, erroneously decoding legitimate GNSS signals, and erroneously replicating GNSS signals during the spoofing attack. s_i^s is modeled as $\mathcal{N}(0, \sigma_{si}^2)$. In general, it is assumed that the spoofer locates on earth similar to the model given in [57]; however, in our system model, we also consider aerial spoofers, e.g., a HAPS station or LEO satellite. Therefore, ionosphere or troposphere-related terms should be included in the pseudorange expression. As a result, the atmospheric error terms are I_i^s and T_i^s, which are ionospheric and tropospheric errors with distributions $\mathcal{N}(0, \sigma_i^2)$ and $\mathcal{N}(0, \sigma_t^2)$, respectively. M_i^s is the multipath error term, which is distributed $\mathcal{N}(0, \sigma_{mp}^2)$.

In the threat model for V2X communication systems, the spoofer may have the credentials of the public key infrastructure to communicate with other vehicles and RSUs. However, our spoofer attack model does not rely on communicating with the other V2X elements, instead, the attacker broadcasts its spoofing signals. In addition, RSUs are assumed as trustworthy similar to the threat model in [30].

3.3.2. Pseudorange Observation-based Spoofing Detection

In the literature on spoofing detection based on pseudorange values, the most important steps are taking single and second differences. Hence, many detection mechanisms are directly correlated with these operations [12,20,21,61–64,85,86]. The single difference between i^{th} and j^{th} receivers for the kth satellite is given as

$$\triangle_{ij}^k = \rho_i^k - \rho_j^k. \tag{3.10}$$

The main reason to use the single difference is to eliminate satellite clock bias. Similarly, the second difference, whose purpose is eliminating receiver clock bias, is stated as

$$\nabla\triangle_{ij}^{kl} = \triangle_{ij}^k - \triangle_{ij}^l. \tag{3.11}$$

Before we apply single and second difference operations for the spoofing case, pseudo-range expressions of legitimate signal and spoofing signal can be combined as

$$\rho_i^{n\in\{k,l,s,p\}} = \begin{cases} \rho_i^s = \hat{d}^k + d_i^s + c(\tau_i - \tau^s) + \epsilon_i^s, & H_0, n = s \\ \rho_i^k = d_i^k + c(\tau_i - \tau^k) + \epsilon_i^k, & H_1, n = k \end{cases} \tag{3.12}$$

where H_0 denotes the spoofing case and H_1 denotes the legitimate signaling. It should be noted that the atmospheric and other error sources (except observational error terms) are neglected in Equation (3.12) for simplicity. In this case, the single difference can be written for the received signal after omitting the terms due to ionosphere and troposphere as

$$\triangle\rho_{ij}^{n\in\{k,l,s,p\}} = \begin{cases} d_i^s - d_j^s + c(\tau_i - \tau_j) + \triangle e_{ij}^s, & H_0, n = s \\ d_i^k - d_j^k + c(\tau_i - \tau_j) + \triangle e_{ij}^k, & H_1, n = k. \end{cases} \tag{3.13}$$

When we assume that receivers obtain only spoofing signals or legitimate signals with s, the double difference of pseudoranges can be written as

$$\nabla\triangle\rho_{ij}^{n,u\in\{k,l,s,p\}} = \begin{cases} \nabla\triangle e_{ij}^{s,p}, & H_0 \cap H_1, n = s, u = p \\ d_i^k - d_j^k - d_i^l + d_j^l + \nabla\triangle e_{ij}^{k,l}, & H_1 \cap H_1, n = k, u = l, \end{cases} \tag{3.14}$$

where $d_i^s - d_j^s = d_i^p - d_j^p$ since spoofing signals are generated from one source when $d_i^s = d_i^p$ and $d_j^s - d_j^p$. The H_1 terms can be approximately zero or may take other values for the different time indices. There are some drawbacks to this methodology. The first one is that at least two receivers (e.g., i^{th} and j^{th} receivers) are required to detect the existence of spoofing signals. However, this condition may not be satisfied all the time. Therefore, spoofing detection methodologies, which are not dependent on multiple receivers, should be studied. The other drawback is that it is assumed that all signals are modeled as spoofing signals at the receiver. However, this assumption may not be true, and in this case, the double difference term does not provide straightforward

spoofing detection terms as

$$\nabla\triangle\rho_{ij}^{n,u\in\{k,l,s,p\}} = \begin{cases} d_i^k - d_j^k - d_i^s + d_j^s + \nabla\triangle e_{ij}^{k,s}, & H_0 \cap H_1, n = k, u = s \\ d_i^k - d_j^k - d_i^l + d_j^l + \nabla\triangle e_{ij}^{k,l}, & H_1 \cap H_1, n = k, u = l. \end{cases} \tag{3.15}$$

As shown in Equation (3.15), it is harder to detect spoofing attacks by using the double differences approach when all the signals are not spoofed. Hence, another methodology should be developed for the detection of spoofing signals while some of the authentic satellite positioning signals are still available at the receiver. The proposed methodology will be detailed in Chapter 5, while we continue with the eavesdropping problem for V2X systems in Chapter 4.

4. DEFEATING EAVESDROPPER USING FEC CODES

As discussed before, security in V2X is a non-negotiable concept, since without a proper security mechanism that may risk human lives, the privacy of individuals, and system operations. In this chapter, we focus on physical layer security approaches in V2X to prevent passive eavesdropping attacks, and we propose an integration of physical layer operations to enhance security. Physical layer security techniques are promising solutions to assist cryptographic methods in the presence of an eavesdropper in V2X setups. In this chapter, we present a physical layer security scheme, which is based on both insertions of a random error vector to FEC codewords and transmission over decentralized relay nodes.

In the scope of this chapter, the intended physical layer solution can be considered as another layer of security and it should also satisfy the reliability, energy efficiency, and latency requirements. In the physical layer, FEC codes and ARQ schemes are already proposed in the literature to overcome ultra-high reliability issues in V2X [31]. Between these two techniques, FEC codes are more suitable than ARQ schemes to deploy in V2X systems due to better power consumption and latency performances. Because of the same reasons, Reed-Solomon, Golay, and Hamming codes are advantageous to other FEC codes [48, 49]. Even if the FEC codes are used, transmission over one link may not satisfy the ultra-high reliability constraints; therefore, cooperative communication techniques should be used for cases when there is no direct path in V2X systems. As a result, FEC codes should be combined with cooperative communication schemes and the obtained model should also be secure against eavesdropping attacks.

The system model of this chapter is given in Figure 4.1, which includes decentralized relay nodes. Thanks to the McEliece cryptosystem, error injection is first applied to information bits, which are encoded with the FEC codes. Reed-Solomon, Golay, and Hamming codes are selected as FEC codes to satisfy power and computational

efficiency. Then obtained codewords are transmitted across reliable intermediate relays to the legitimate receiver. As a performance metric, the decoding FER of the eavesdropper is analytically obtained for the fragmentary existence of significant noise between relays and Eve. The simulation results validate the analytical calculations, and the obtained results show that the number of low-quality channels and the selected FEC code affects the performance of the proposed model. In other words, the security level is highly based on the location of the eavesdropper and secure communication can be achieved when some of the channels between the eavesdropper and relay nodes are significantly noisier.

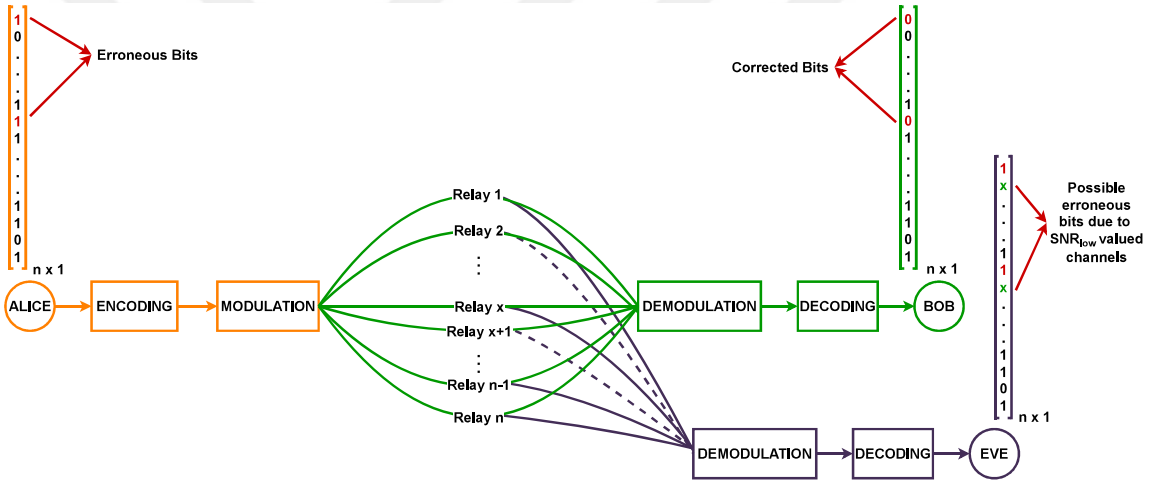

Figure 4.1. The proposed system model, where Alice injects a spatially distributed random error vector to the codeword and transmits with the aid of relays to Bob. Eve tries to decode successfully, where SNR_{low} valued dashed-lined channels may lead to additional errors.

4.1. System Model of FEC-based Eavesdropping Mitigation with the Help of Relays and McEliece Cryptosystem

Similar to most of the system models proposed in the literature on physical layer security, there are three end nodes, a legitimate transmitter (Alice), a legitimate receiver (Bob), and an eavesdropper (Eve). In this model, Eve tries to capture the information transmission between Alice and Bob (exactly as in the wiretap channel

model [8]). In addition to these three end-nodes, N intermediate relay nodes are placed in the scope of the V2X setup between Alice and the receivers, which are Bob and Eve, as shown in Figure 4.1. We assume that the distance between Alice and Bob is large enough to avoid a direct communication path between them and all of the relays, which are trustworthy, operate based on the amplify and forward protocols. As another assumption, Eve has full knowledge of the agreed coding scheme between Alice and Bob, and Eve also has limitless computational resources. These conditions represent the worst-case scenarios in terms of communication secrecy. Alice has only the state information of the channels between Alice and the trustworthy relays, whereas she does not know the channel information of corresponding channels between relays and Eve, who is a passive attacker and is not transmitting any packets (including pilots). On the receiver side, both Bob and Eve use hard decoding algorithms.

In communication systems, FEC is a broad class and has many sub-branches in the literature, whereas we exploit FEC codes in order to increase the security performance of a V2X system at the same time satisfying the reliability of the legitimate user. FEC codes are generally denoted as (n, k, t), where n is the codeword length, k is the information bit length, t is the error correction capability and the coding rate is calculated as $R = k/n$. In general, $n \leq N$, where N is the number of relay nodes in the system, and $n = N$ is chosen for simplicity in the scope of this chapter. In order to satisfy the specific requirements of V2X, such as latency and reliability, Reed-Solomon, Golay, and Hamming codes are chosen to improve the security of the proposed system model, since these FEC codes are energy, memory, and time-efficient structures in comparison with other FEC codes, such as low-density parity check (LDPC) and Turbo codes.

In order to increase data security, Alice first encodes information bit vectors $\boldsymbol{x}$ using the chosen (n, k, t) FEC code by multiplying code generation matrix $\boldsymbol{G}$, where n is the code length, k is the information bit length and t is the error correction capability of the receiver as

$$\boldsymbol{x_c} = \boldsymbol{G} \cdot \boldsymbol{x}, \tag{4.1}$$

where $\boldsymbol{x_c}$ is the encoded bit vector. In this thesis, Reed-Solomon, Golay, and Hamming codes are utilized as FEC codes with the generation matrices, $\boldsymbol{G_R}$, $\boldsymbol{G_G}$, and $\boldsymbol{G_H}$, respectively. Among the determined coding schemes, Reed-Solomon codes are symbol-wise codes, and they can be successfully decoded in the presence of up to $(n-k)/2$ errors or $(n-k)$ erasures [87]. Since Reed-Solomon codes are symbol-wise, the number of symbols in a block cannot exceed 2^m-1, where m is the symbol length in terms of bits. On the other hand, Golay and Hamming codes are perfect codes, which are modeled as spheres with radius $(d_{min}-1)/2$ based on the sphere packing problem [87]. Here, d_{min} is the minimum Hamming distance between codewords. The most significant property of perfect codes is that there is precisely one codeword within t distance, Therefore,errors up to t are certainly corrected at the receiver, and more than t errors cannot be corrected under no circumstance.

Before the transmission of the encoded bitstream, Alice also adds t errors onto the encoded data vector as

$$\boldsymbol{x_A} = \boldsymbol{x_c} \oplus \boldsymbol{x_e}, \tag{4.2}$$

where $\oplus$ is the modulo 2 summation operator, and the $\boldsymbol{x_e}$ is the error vector of weight t, and $\boldsymbol{x_A}$ is the vector that is transmitted to the intermediate relays after modulating using the BPSK modulation scheme. During transmission, we assume that there is no direct path between Alice and Bob, or Alice and Eve. This assumption is realistic in cases when many V2X agents operate in long ranges with the help of relays. The obtained vectors after encoding and modulation are transmitted with the help of intermediate relays by applying a specific routing scheme, in which each group of transmitted vector bits is sent over n distinct relays, where $n \leq N$. The transmission mode of each relay is assumed as amplify and forward. The transmission mediums between Alice and the relays can be assumed as error-free channels since Alice can choose the best relays to get their assistance during the transmission of vectors to Bob [88].

In this thesis, channels between each three end nodes and relays are classified into two groups, where these channels may have SNR_{high} or SNR_{low} values. Since these values indicate the channel quality, channels denoted as SNR_{low} tend to create more

transmission errors in comparison with SNR_{high} valued channels, where $SNR_{low} \ll SNR_{high}$. In the scope of this study, we assume that all of the relay nodes are perfectly available for Alice and Bob, where the channels between relays and Alice or Bob have SNR_{high} values. Since SNR_{high} may be large enough not to cause transmission errors, it may mean that the channels with SNR_{high} values may be error-free channels. In this case, the channel responses of these SNR_{high} valued channels between relays and Bob can be denoted by $\boldsymbol{H_B}$, which is an $n \times n$ channel state matrix. Therefore, received codewords at Bob can be expressed as

$$\boldsymbol{y_B} = q(\boldsymbol{x_A} \cdot \boldsymbol{H_B} + \boldsymbol{w_b}), \tag{4.3}$$

where $q(\cdot)$ is a quantization function due to hard decision at Bob, and it quantizes its input to the possible BPSK levels. In the same expression, $\boldsymbol{w_b}$ is the noise vector, whose elements are generated with respect to SNR_{high} values.

From Eve's point of view, there are very low SNR values, which are denoted as SNR_{low}, on some channels, which are indicated with dashed purple lines in Figure 4.1, even if the majority of the channels between intermediate relays and Eve have SNR_{high} values. This assumption is realistic when the relays utilize beamforming techniques, which may lead to low SNR values on the Eve side due to the directional transmission pattern. Under these circumstances, there are L channels with SNR_{low} between the relays and Eve, whereas $n - L$ channels have SNR_{high} as the SNR value. As a result, the received vector at Eve can be stated as

$$\boldsymbol{y_E} = q(\boldsymbol{x_A} \cdot \boldsymbol{H_E} + \boldsymbol{w_b}), \tag{4.4}$$

where $\boldsymbol{H_E}$ is the $n \times n$ channel state matrix of the channels between intermediate nodes and Eve, and $\boldsymbol{w_e}$ is the noise vector. The elements of the noise vector are generated based on the numbers and the positions of the SNR_{high} and the SNR_{low} valued channels. In the rest of the chapter, all of SNR_{high} and SNR_{low} valued channels are respectively modeled as AWGN or Rayleigh fading for further analyses. It should also be noted that both $\boldsymbol{H_B}$ and $\boldsymbol{H_E}$ are diagonal matrices due to the transmission of each bit of $\boldsymbol{x_A}$ over distinct relay nodes.

The motivation of the proposed model is to satisfy the target communication reliability with Bob while increasing the security against a passive eavesdropping attack on Eve. Under these circumstances, Eve only listens to the messages sent from relays and tries to capture as many data bits as possible. In this model, we assume that Eve knows all the transmission processes including the selected FEC code, and she has limitless computational resources. Hard decoding algorithms are used by Eve and Bob, while the channel states between relays and Eve are not known by Alice. As expected, the channels with SNR_{high} become error-free for high SNR values, and erroneous bits are observed with very low probability. As a result, Bob can successfully decode the received vector to the original information message with a very high probability, when SNR_{high} is large enough. This situation is required for reliable transmission between Alice and Bob.

On the other hand, the security performance of the proposed model is based on the error correction capability of Eve. To achieve a high level of security against Eve, we expect to observe that there may be additional bit errors due to transmission over SNR_{low} valued channels, in addition to the injected t errors. In this thesis, two possible channel models, which may create additional errors or erasures, are considered for SNR_{low} valued channels between Eve and relay nodes. The first channel model denoted as C_I is that SNR_{low} valued channels create erasures at the receiver side with a certain probability (i.e., $SNR_{low} \rightarrow -\infty$). Therefore, only coding schemes, which support erasure decoding schemes, can correctly decode transmitted codewords in case of a limited number of erasures. Here, we also assume that SNR_{high} valued channels are error-free during analytical calculations. This assumption is realistic since we expect to observe very high reliability, and Alice can select appropriate relays by running a relay selection algorithm with respect to channel coefficients. In the scope of this thesis, the erasure channel C_I is applicable for Reed-Solomon coding schemes up to $n-k$ erasures with respect to the considered channel model. An important assumption is that Eve has the knowledge of all positions of erased bits in the received codeword.

The second considered channel model denoted as C_{II} is an additive white Gaussian noise (AWGN) or Rayleigh fading channel with a very low SNR value (e.g., $SNR_{low} = -50$ dB). The error probability of channels with SNR_{low} and SNR_{high} are indicated as P_e^{low} and P_e^{high}, respectively. Due to the considered models, which are chosen as AWGN and Rayleigh fading within the scope of this thesis, we distinguish error terms indicating a lower index. These error terms can be expressed for BPSK in the AWGN channel as

$$P_{e,AWGN}^{low} = Q\big(\sqrt{2SNR_{low}}\big) \tag{4.5}$$

$$P_{e,AWGN}^{high} = Q\big(\sqrt{2SNR_{high}}\big), \tag{4.6}$$

where $SNR_{low} << SNR_{high}$ and $Q(\cdot)$ is the tail distribution function of the standard normal distribution [89]. Therefore, the transmission errors tend to be observed on SNR_{low} valued channels. Both of these channel models can also be modeled as binary symmetric channels (BSC) with bit-flip error probabilities calculated as Equation (4.5) and Equation (4.6). Under these circumstances, there is a risk of correction injected error bits. If there is a bit flip on SNR_{low} valued channels while transmitting injected error bits over these channels, Eve may successfully decode the information bitstream. This scenario and further analyses on the error probability of Eve will be deeply studied in the next section for the AWGN channel model.

Similar to the AWGN channel model, error probabilities of Rayleigh fading channels can be calculated as [89]

$$P_{e,Rayleigh}^{low} = \frac{1}{2}\sqrt{\frac{SNR_{low}}{SNR_{low}+1}} \tag{4.7}$$

$$P_{e,Rayleigh}^{high} = \frac{1}{2}\sqrt{\frac{SNR_{high}}{SNR_{high}+1}}, \tag{4.8}$$

where, again, $SNR_{low} << SNR_{high}$ [89]. These error probabilities demonstrate a similar behavior as in the AWGN channel case, such that increasing error leads to a decreased channel quality. Similar risks about error correction and security leakage in

AWGN channels appear for channels with Rayleigh fading. The error terms, given in Equation (4.5)-Equation (4.8), are utilized in the decoding error probability of Eve, which is used as the principal general performance metric and explained in detail in the next section.

4.2. Calculation of Analytical Error Probabilities

The security gap [47] is defined as the channel quality difference between Eve and Bob for the required security level. It is highly dependent on the error probabilities of receiver nodes. Therefore, within the scope of this chapter, the error probability can be used as an appropriate metric when an attacker exists, even if general error probability expressions are considered to evaluate the reliability of communication schemes [90]. We measure the decoding FER of the receiver side, which may be Bob or Eve, to understand the performance of the proposed system model regarding security concerns. In this case, the FER of Bob indicated as P_{ϵ}^{Bob}, and it should be small enough to satisfy reliable communication with Alice regarding the condition $P_{\epsilon}^{Bob} \to 0$, while we aim to achieve $P_{\epsilon}^{Eve} \to 1$ to observe secure communication between legitimate users, where P_{ϵ}^{Eve} denotes the FER of Eve.

In reality, amounts and values of the SNR_{low} and SNR_{high} valued channels are based on the positions of the end nodes and the relays. However, channel states are generically given to compute the error performance of the proposed system model. System performance, which will be evaluated with P_{ϵ}, is based on the overlapping condition that is observed when some of the injected error bits or all of them are transmitted over SNR_{low} valued channels. $P_{overlap}$, which indicates the probability of occurrence of at least one overlap, can be found with the expression $P_{overlap} = 1 - P_0$, which is the probability of zero overlaps in total and it can be calculated as:

$$P_0 = \binom{n-t}{L} \Big/ \binom{n}{L}, \tag{4.9}$$

where t denotes the number of injected errors, and L is the number of channels with SNR_{low}. Due to the nature of the Reed-Solomon code, erasures can count as half of the

error bits; therefore, 2 erasures have the same impact as 1 error on the receiver side. As a result, the received codeword can be successfully decoded when each erased error bit is replaced with up to 2 erasures. For instance, the 1 error bit, which is intentionally injected by Alice, is erased, and the receiver correctly decodes up to 1 more erasure. Under that specific condition, if the value of L is 1 or 2, $P_\epsilon^{C_I} = P_0$ for the corresponding L value, where $P_\epsilon^{C_I}$ is error probability of Eve with the channel model C_I. However, the probability of 1 overlap or more overlaps should also be calculated in order to find $P_\epsilon^{C_I}$, when L is larger than 2. In general, error probability can be calculated as $P_\epsilon^{C_I} = P_0 + P_1 + P_2 + \cdots + P_i + \cdots$, where P_i is the probability of i overlaps. The final term of this summation is decided based on the error correction capability of Eve under the condition $t - i + \frac{L}{2} \leq t$. In other words, Eve cannot correctly decode when $2i < L$. As a result, the decoding error probability of Eve can be calculated as follows

$$P_\epsilon^{C_I} = \sum_{i=0}^{\min(t,L)} \binom{n-t}{L-i}\binom{t}{i} \Big/ \binom{n}{L}, \qquad \text{for} \quad 2i < L \tag{4.10}$$

when $SNR_{low} \to -\infty$ and where i may take the values from zero up to the maximum number of possible overlaps, which is the minimum of t and L.

Table 4.1. Probabilities of number of overlaps for Reed-Solomon codes with respect to the error probability of Eve.

Reed Solomon Code	L	P_0	P_1	$P_\epsilon^{C_I}$
$(15, 11, 2)$	1	0.8667	-	0.8667
$(15, 11, 2)$	2	0.7429	-	0.7429
$(15, 11, 2)$	3	0.6286	0.3429	0.9714
$(15, 11, 2)$	4	0.5238	0.4190	0.9429

In Table 4.1, numerical values of P_0 and P_1, and respective analytical values of $P_\epsilon^{C_I}$ for the chosen Reed-Solomon code, which will be used as main coding scheme in next sections, are given in order to verify the derivation of Equation (4.10). Here, the columns denoted with " - " indicate that P_1 is not taken into account during the calculation of the error probability of Eve, due to the $2i < L$ condition not being satisfied. When this table is extended with larger values of L, additional terms, such

as P_2, P_3, will be needed to calculate $P_\epsilon^{C_I}$. As a result, the derivation of Equation (4.10) from Equation (4.9) can be verified with observed values given in the table.

When C_{II} is considered as a real valued low SNR channel model between relays and Eve, the calculation of the decoding error probability of Eve becomes more complicated. Before calculating the error probability of Eve, we should explain all the possible scenarios in which Eve cannot decode the information bitstream. Therefore, Eve must observe at least $t+1$ errors. This condition is highly dependent on the channel qualities, and their error probabilities are given in equations Equation (4.5)-Equation (4.8). The possible necessities for erroneous decoding at Eve are listed as follows

- No error correction due to bit flip on SNR_{low} or SNR_{high} valued channels and at least one error occurs on SNR_{low} or SNR_{high} valued channels,
- Correction of t_c errors due to bit flip on SNR_{low} or SNR_{high} valued channels and at t_c+1 errors on SNR_{low} or SNR_{high} valued channels, where $t_c \leq t$.

Since $P_e^{low} > P_e^{high}$, we expect to observe that there are more injected error bit corrections and additional bit errors on SNR_{low} valued channels. Under these circumstances, we can distinguish the best and the worst case scenarios for our aims to have high security in CPS. The best-case scenario is transmitting injected errors on SNR_{high} valued channels to Eve, where we anticipate observing more transmission errors on SNR_{low} valued channels. This scenario is realistic most of the time, since SNR_{low} valued channels may be observed in the noisy and dynamic environment of V2X networks. On the other hand, the worst-case scenario is that there is an overlap of injected errors and SNR_{low} valued channels. In other words, injected errors are transmitted over SNR_{low} valued channels. Therefore, these bits tend to be corrected due to bit flips with a considerable probability. In addition, the desired additional bit errors are hard to be observed in the worst-case scenario since SNR_{high} valued channels are less erroneous. Beyond the best and worst-case scenarios, there are many possible states that the system model may encounter. When we consider all of them, we can calculate the error probability on the receiver side to analyze the system performance.

The error probability of Eve is given for input tuple $(n, L, t, SNR_{high}, SNR_{low})$ with sub-functions $\mathcal{G}(\cdot)$ and $\mathcal{H}(\cdot)$, which are used to simplify the notation, for the proposed system model when perfect FEC (n, k, t) schemes are derived as follows

$$P_{\epsilon}^{C_{II}} = 1 - \left(\left(\sum_{i=0}^{\min(t,L)} \binom{n-t}{L-i} \binom{t}{i} \sum_{j=0}^{t-i} \sum_{k=0}^{\min(t-j,i)} \sum_{l=0}^{j+k} \sum_{m=0}^{\min(L-i,j+k+l)} \mathcal{G}(\cdot)\mathcal{H}(\cdot) \right) \Big/ \binom{n}{L} \right) \tag{4.11}$$

$$\mathcal{G}(\cdot) = \left(P_e^{low}\right)^{k+m} \left(1 - P_e^{low}\right)^{L-k-m} \binom{i}{k} \binom{L-i}{m} \tag{4.12}$$

$$\mathcal{H}(\cdot) = \left(P_e^{high}\right)^{l+j} \left(1 - \left(P_e^{high}\right)\right)^{N-L-l-j} \binom{t-i}{j} \binom{n-L-t+i}{l}, \tag{4.13}$$

where i indicates the number of possible overlaps, where k and j are the numbers of corrected errors, which are actually injected by Alice, on SNR_{low} and SNR_{high} valued channels, respectively for $k + j = t_c$. In other words, there is at least one overlap if $k > 0$, since an error cannot be corrected on a channel with SNR_{low} without an overlap. In addition to that, l and m denote transmission bit errors on the corresponding transmission channels, which have SNR_{high} and SNR_{low}, respectively. $\mathcal{G}(\cdot)$ and $\mathcal{H}(\cdot)$, which are given in Equation (4.12) and Equation (4.13), respectively, are the functions with respect to the probability of correct decoding for the combination of the corresponding (i, j, k, l, m) tuple. These functions are used to simplify the whole error probability term, but they can be considered as the impacts of SNR_{low} and SNR_{high} valued channels to P_{ϵ}. By considering all possible values of this tuple, the summation of multiplication of $\mathcal{G}(\cdot)$ and $\mathcal{H}(\cdot)$ helps to calculate the probability of correct decoding at Eve.

In order to calculate the error probability of Bob, Equation (4.11) can straightforwardly be used for $L = 0$, where there is no SNR_{low} valued AWGN channels. Therefore, this expression is very beneficial to analyze the overall system performance based on error ratios of the legitimate and the non-legitimate receivers. Another feature of this expression is that it can be easily reformulated in the existence of more legitimate or non-legitimate receivers by directly changing the channel error probability terms

for legitimate users and introducing more sub-functions for additional non-legitimate receivers.

4.3. Application Scenarios

In this study, we consider four different error control coding schemes, which are Reed-Solomon $(15, 11, 2)$, Golay $(23, 12, 3)$, Hamming $(15, 11, 1)$, and Hamming $(7, 4, 1)$ codes. Also, three scenarios are taken into account with respect to channel models, which are detailed as follows

(i) In the first scenario (**S1**), Reed-Solomon $(15, 11, 2)$ codes are utilized as FEC code with calculated FER of Eve given in Equation (4.10). Due to the existence of high noise in V2X communication environments, erasures can be encountered on the receiver side. Thus, Reed-Solomon $(15, 11, 2)$ code is used as the FEC code, where Eve is able to detect erased bits, which are observed due to $SNR_{low} \to -\infty$ valued erasure channel in the first scenario (**S1**). During the performance analysis of this code, $t = 2$ weighted error vectors are added to information bits and obtained codewords are transmitted over the proposed system model. In this case, Eve may correctly solve received vectors with up to 4 erasures due to the nature of Reed-Solomon codes. The performance of this scenario will be evaluated with $P_e^{C_I}$ as given in Equation (4.10) as security outages occur due to the erasure effects of the channel. As discussed before, Eve should know the exact placements of erasure bits shown as X in Figure 4.1 in order to decode received bits.

(ii) In the scope of the second scenario (**S2**), Golay $(23, 12, 3)$, Hamming $(15, 11, 1)$, and Hamming $(7, 4, 1)$ codes are studied, when SNR_{low} takes very low real values as discussed with the C_{II} definition. The channel is modeled as an AWGN channel with error probabilities Equation (4.5) and Equation (4.6). The interpretation of the system performance will be done with Equation (4.11) for $SNR_{low} = -50$ dB, which leads to $Q\left(\sqrt{2SNR}\right) = 0.5$ error probability on the AWGN channel. Since these codes are perfect codes, the simulation results are expected to well match analytical calculations. In this scenario, injected errors will be corrected

with probability 0.5, if there are overlaps with SNR_{low} valued channels. It should be remembered that there are no erasure bits at Eve, and the bits shown as X at Eve given in Figure 4.1 take value from the set $\{0, 1\}$ with probability $p = 0.5$.

(iii) Third scenario (**S3**) has the same features except for the channel model. In this scenario, the channel is modeled as Rayleigh with channel error probabilities Equation (4.7) and Equation (4.8) where $SNR_{low} = -50$ dB. Similar to the **S2** FER calculation of Eve is completed via Equation (4.11) for the perfect codes Golay $(23, 12, 3)$, Hamming $(15, 11, 1)$, and Hamming $(7, 4, 1)$.

4.4. Simulation Results

In this section, the security capabilities of the scenarios **S1**, **S2**, and **S3** are evaluated by considering analytical error probabilities, which are given in Equation (4.10) and Equation (4.11), and their respective simulation results. The set of the possible number of SNR_{low} valued channels are denoted as $\mathcal{L}$, and the sets are trained with respect to the corresponding t values of coding schemes. In order to understand the consequences of the selected L value, which may be larger or smaller than t, L is determined from $\{1, 2, 3, 4\}$ for **S1**, and determined from, but not limited to $\{0, 1, 2\}$ for **S2** and **S3**. This set enables a fair comparison of Hamming and Golay FEC codes, and we can see the impacts of various t/L rates. In simulations, proposed transmission schemes are run for 10^5 times for each L value to follow the long-term performance of the proposed system model. Moreover, SNR_{high} takes on the values from 0 to 50 dB. The obtained results are given in Figure 4.2, Figure 4.3, Figure 4.4, and Figure 4.5, where burgundy lines show the performance of the legitimate receiver Bob with $L = 0$ value. It can be seen that Bob can successfully decode received vectors for the increased signal power as we expected. When we focus on **S2** and **S3**, we aim at validating analytical the error probability expression given in Equation (4.11) with simulations for various FEC codes. Another target is to find the appropriate perfect FEC code to be utilized in the noisy communication environment of V2X. It should also be emphasized that the length of each frame is equal to the codelength of encoded vectors.

In Figure 4.2, simulation results of $P_{\epsilon}^{C_I}$ based on the Reed-Solomon $(15, 11, 2)$ code is given with reference analytical calculations, which are found with Equation (4.10). These figures are shown for the set $L \in \{0, 1, 2, 3, 4\}$ with respect to the first scenario **S1**. In Figure 4.2, it can be easily seen that the simulation results for the Reed-Solomon $(15, 11, 2)$ code are highly correlated with analytical calculations when SNR_{high} is larger than 10 dB. In Figure 4.2, results show an interesting behavior, such that there is no sequential order with respect to L values. The results show that the condition of $P_{\epsilon}^{C_I}(L = 2) < P_{\epsilon}^{C_I}(L = 1) < P_{\epsilon}^{C_I}(L = 4) < P_{\epsilon}^{C_I}(L = 3)$, where $P_{\epsilon}^{C_I}(\cdot)$ is a function of the decoding error correction probability of Eve for given L. This non-sequential order leads to an implementation difficulty for the proposed system model. As an example, at least two additional channels should have SNR_{low} instead of SNR_{high} in order to observe higher $P_{\epsilon}^{C_I}$ values when $L = 1$. Therefore, this type of erasure FEC models do not guarantee increasing security for incremental values of L.

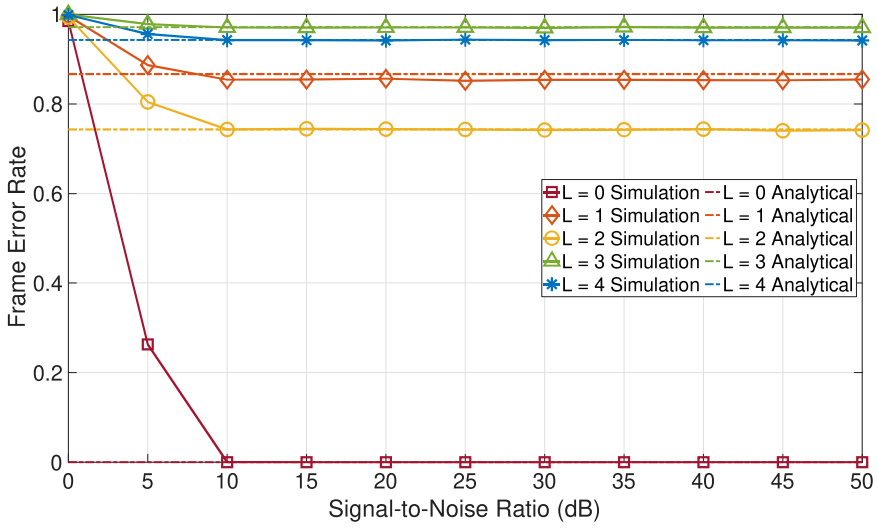

Figure 4.2. The comparison of simulation results of **S1** and analytical results of $P_{\epsilon}^{C_I}$ where $L \in \{0, 1, 2, 3, 4\}$ for Reed-Solomon $(15, 11, 2)$ code.

In the scope of **S2** and **S3**, the analytical results calculated with the equations Equation (4.11), Equation (4.12), and Equation (4.13) are given with the simulation outcomes in Figure 4.3, Figure 4.4, and Figure 4.5. It should be noted that we target to achieve $P_{\epsilon} \to 1$. In this chapter, the simulations are run for 10^5 times for each FEC

code and L value, where $SNR_{low} = -50$dB and SNR_{high} take on values between 0 and 50 dB. In each figure, the simulation results are shown with various markers and compared with corresponding analytical curves, which are calculated with Equation (4.11), Equation (4.12), and Equation (4.13). It can be clearly seen that the simulation results and the analytical calculations are perfectly matched for all considered L values in the target SNR region for the selected FEC codes.

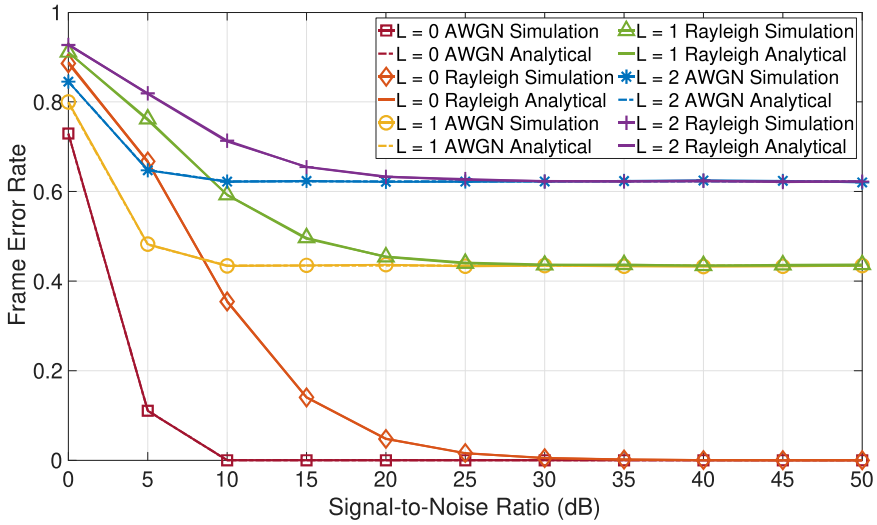

Figure 4.3. The comparison of the simulation results and analytical results of FER values of Eve for Golay $(23, 12, 3)$ code for AWGN and Rayleigh channels where $L \in \{0, 1, 2\}$.

In these figures, the decoding performances of Bob are shown in these figures with the green lines, where $L = 0$. These results demonstrate that the decoding performance of Bob improves as expected since Bob receives fewer erroneous bits thanks to the better communication channels. The results show that the Hamming $(7, 4, 1)$ code shows the best performance for satisfying a reliable transmission between Alice and Bob for $SNR_{high} > 25$ dB. This shows that Hamming $(7, 4, 1)$ code is a suitable candidate in noisy environments for reliability between Alice and Bob in comparison with Golay $(23, 15, 3)$ and Hamming $(15, 11, 1)$ codes.

When we focus on the performance of Eve, a Golay $(23, 12, 3)$ code is investigated, and the analytical and simulation results of this code are given in Figure 4.3. We

can easily evaluate the results by splitting them into two cases, which are results for $SNR_{high} < 25$ dB and $SNR_{high} \geq 25$ dB. For each case, one possible way can be considered to increase the FER of Eve. In the first condition where $SNR_{high} < 25$ dB, worsening channels between Bob and relays with fading effects can obtain more security. In this case, Bob cannot correctly decode all the received codewords. As a result, this method is not helpful to satisfy both secure and reliable communication. In the second case with $SNR_{high} \geq 25$ dB, we observe increased security as L increases, which means the number of SNR_{low} valued channels should be higher. However, locating Eve and worsening the channel is not a minor problem. We also study two different Hamming codes to compare with the security performance of the Golay code.

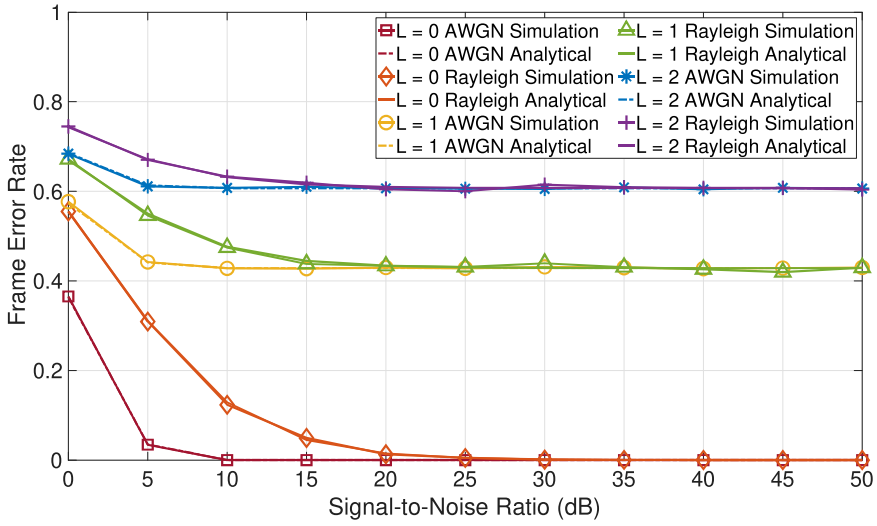

Figure 4.4. The comparison of the simulation results and analytical results of FER values of Eve for Hamming $(7, 4, 1)$ code for AWGN and Rayleigh channels where $L \in \{0, 1, 2\}$.

In Figure 4.4 and Figure 4.5, the analytical calculations and simulation results of Hamming $(7, 4, 1)$ and Hamming $(15, 11, 1)$ are given for AWGN and Rayleigh channel models. These figures show that security performances of the Hamming $(7, 4, 1)$ and the Golay $(23, 12, 3)$ code are very similar for various L values, where $SNR_{high} > 25$ dB. However, the Hamming $(15, 11, 1)$ code gives the best results when compared to the two other codes for the increased SNR_{high}. On the other hand, the results of the Golay

$(23, 12, 3)$ code and the Hamming $(15, 11, 1)$ code perform very similarly to each other and much better than the results of the Hamming $(7, 4, 1)$ code for $SNR_{high} < 25$ dB. Therefore, a trade-off appears for low SNR values with respect to the performances of Eve and Bob. As we have discussed, Hamming $(7, 4, 1)$ demonstrates the best performance for Bob and the worst performance against Eve, and the other two codes perform in the opposite manner.

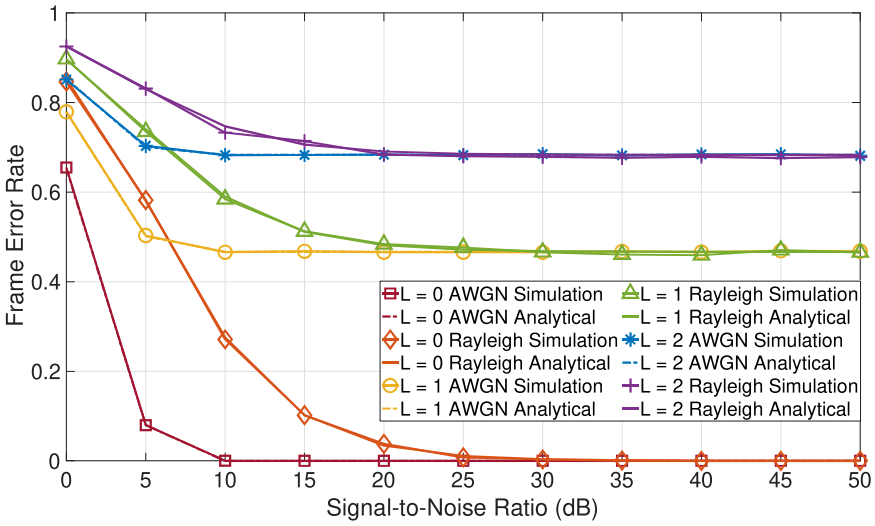

Figure 4.5. The comparison of the simulation results and analytical results of FER values of Eve for Hamming $(15, 11, 1)$ code for AWGN and Rayleigh channels where $L \in \{0, 1, 2\}$.

Here, we also discuss a third parameter based on the coding rates of the chosen FEC codes. When we compare the coding rates, the Hamming $(15, 11, 1)$ code is advantageous with the coding rate $R_{Hamming(15,11,1)} = 0.73$, where the coding rate of the Hamming $(15, 11, 1)$ code is $R_{Hamming(7,4,1)} = 0.57$ and the coding rate of Golay $(23, 12, 3)$ code is $R_{Golay} = 0.52$. Another point of view is based on the deployment concerns such as latency and computational complexity. Since there will be additional headers for each parallel transmission over distinct channels, there are more header bits in a frame due to an increased number of relays. The Hamming codes are good candidates to be used in CPS with their simple structure [49].

The main consequence is that significantly noisy channels between relays and the eavesdropper are helpful in securing communication between legitimate users. The results verify that a Hamming $(15, 11, 1)$ code can be used instead of a Golay $(23, 12, 3)$ code or a Hamming $(7, 4, 1)$ code due to its better overall performance against an eavesdropper. Besides, the coding rate of a Hamming $(15, 11, 1)$ code is much higher than the other contenders so information efficiency can be increased with the usage of this scheme. Another outcome is that an increased number of bit erasures enhances system performance non-linearly for Reed-Solomon coding scheme, while perfect codes perform better for a large number of bit errors.

5. PSEUDORANGE BASED SPOOFING DETECTION ALGORITHMS USING HYPERBOLA EQUATIONS

Spoofing attacks have become a significant issue in the security of navigation systems in recent years due to incidents against existing location services. In this chapter, we first focus on the impacts of time-spoofing attacks, including the falsification of the GPS time stamp and signal travel time information on GPS platforms. The analyses are completed while taking into account the number of authenticated GPS signals and the quantity of artificial GPS-like signals within the spoofing attack. The results show that the GPS time stamp attacks are more effective than signal travel time-based attacks. Also, the precision of GPS time stamp attacks can be adjusted with additional costs in attack design while staying incognito.

After understanding the impacts of time spoofing attacks, we propose spoofing detection algorithms for the joint existence of spoofing and authentic GPS signals, while the literature widely focuses on only legitimate signals or only attacker signals scenarios. Within this process, we derive hyperbola equations with the help of inversely utilized single difference operation without requiring multiple antennas for available pseudorange values, which may be either spoofed or authentic in a V2X system. In order to detect spoofing attacks against secure positioning signals, we propose an algorithm, which is named the sub-optimal search-based spoofing detection algorithm (Algorithm 1), and it considers all possible numbers of spoofing attacking signals, but not all spoofing scenarios with the same amount of spoofing signals. To address the complexity-related issues due to the increased number of search scenarios of this approach, we propose another algorithm, which is called subset selection-based spoofing detection algorithm (Algorithm 2), with a smart selection of the search subsets. Both of these algorithms are first compared with fixed detection thresholds, which are determined with the Pareto front approach. A supervised learning decision-tree approach is also studied to compare both algorithms. After this phase, these algorithms are also evaluated for a variety of error terms, for example, spoofing imperfections, at-

mospheric errors, multipath, and mobility in order to represent V2X networks better. Finally, a high-level decision fusion-supported hybrid operational model is proposed for both algorithms in V2X deployments to improve security at a network level.

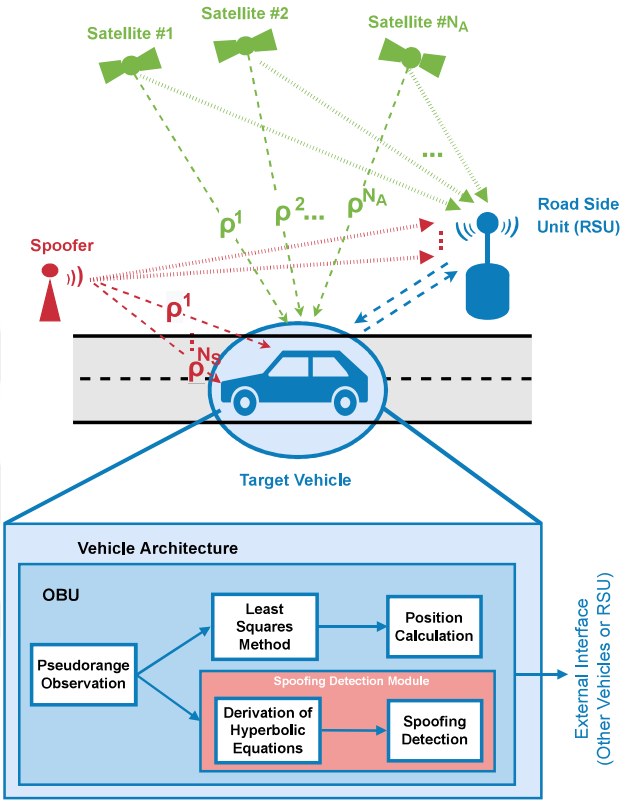

Figure 5.1. A system model of the pseudorange-based spoofing detection in a vehicular communication scenario. At the target vehicle, both spoofing and authentic GPS signals are available.

5.1. System Model

In Chapter 1.1 and Chapter 3.1, the overall system model and the fundamentals of the GNSS-based positioning with possible spoofing attacks are already detailed. When we focus on spoofing attacks against V2X communication, the system model can be degraded, as shown in Figure 5.1. In the scope of this section, we assume that the number of available satellite signals is N_G, where $N_G \geq 4$. In this case, the target vehicle can obtain broadcasted ephemeris data of available satellites. The local clock and position information are part of the ephemeris data; hence, the receiver calculates

its position based on this information. As mentioned before, our proposed algorithms are based on pseudorange values, which are independent of the coding schemes and transmission frequency, and GPS is chosen as the development environment for the algorithm design due to the high amount of studies, availability of open-source data, and open-source system model. During the design of the spoofing detection algorithms, we focus on the combination of spoofed and authentic GPS signals since we believe that most V2X applications can be better represented with the joint analysis of the spoofing and authentic GPS signals. After increasing GPS receivers' sensitivity, the existence of both signals on the wireless medium is realistic

In the system model, there is a target receiver, which is using unencrypted L1 GPS services, and this service relies on N_A authentic signals, where $N_A \leq N_G$. At the same time, a spoofer, which may be ground-located or aerial, sends a combined signal of N_S signals to deceive the target receiver, where $N_A + N_S = N_G$. As a part of the V2X communication network, one RSU or multiple RSUs may be located on the travel trajectory of the target receiver, which communicates these RSUs simultaneously. The deployed V2X communication technologies can be DSRC or LTE-V2X systems, whereas interconnection of the V2X elements is out of the scope of this thesis. The target device finds its position and communicates with other V2X agents, such as an RSU or another vehicle, with the help of its OBU, which also runs one of the proposed spoofing detection algorithms in the spoofing detection module. It should be noted that RSUs may have a fixed position and power connection in many deployments.

5.2. Time-spoofing Attacks in GPS

Even if there are intensive studies to find effective detection and mitigation techniques to provide secure positioning signaling, the effects of the number of spoofing signals are not completely understood yet. In this section, we focus on the distance-based consequences of GPS time-spoofing attacks. We analyze the impacts of falsification of GPS timestamp and signal travel time, where there are still authenticated signals in addition to the spoofing signals.

In time-spoofing attacks, the spoofer may violate the GPS time stamp by injecting a time error Δt^s, which directly influences d_i^k. An advance spoofer adjusts $\hat{d}^k$ regarding the real distance between the corresponding satellite and the receiver. In other time-spoofing attacks, the spoofer may change the signal travel time by adding t_{tr}. This falsified information can also be shown with the $\hat{d}^k$ term since the false pseudorange is adjusted with fake location data regarding signal travel time.

During our simulations, we utilize the SoftGNNS v3.0 MATLAB toolbox provided in [79], where the GPS signals are recorded as raw data, where the position of the receiver is $45°3'55.2708''N$, $7°39'31.9896''E$ at $183.970m$ height without any spoofing attack. Both of the attacks are studied for the parameters given in Table 5.1 for one simulation due to the lack of the probabilistic parameters, while $N_S \leq N = 6$. As a quality metric for a successful spoofing attack without detection, horizontal distance error, and vertical distance error should be less than $1km$ and $150m$, respectively [11]. Since we reconsider with ENU coordinate system, distance errors on X and Y axes should be less than $700m$, which are the approximate values for the circle with a radius of $1km$. For this section, $700m$ and $150m$ are the reference detection levels for X and Y-planes, and the Z-plane, respectively. We present the distance errors for each ENU coordinate in X, Y, Z planes for the attacks given in Table 5.1.

Table 5.1. Parameters of time-spoofing attack.

Δt_{tr} (milliseconds)	Δt^s (seconds)
Signal Travel Time Attack	GPS Time Stamp Attack
$10, 1, 0.1, 0.01, 0, 001$	$6, 5, 4, 3, 2, 1$

In Figure 5.2, the position errors for each plane are given for various values of N_S with respect to t_{tr} with a reference detection level. These results demonstrate that the increased value of t_{tr} leads to very high distance errors in each coordinate. When t_{tr} is larger than 10μs, the attack will be detected in X and Y-planes, even if the attacks, which are larger than 1μs, can be detected in Z-plane. The results for $N_S = 4$ and $N_S = 3$ are significant on X and Y-planes, respectively. It is harder to detect these

cases due to fewer position errors when t_{tr} is smaller. When the spoofer generates its signals as the combination of all available satellites, e.g., $N = N_S$, the spoofer may successfully violate the receivers' position.

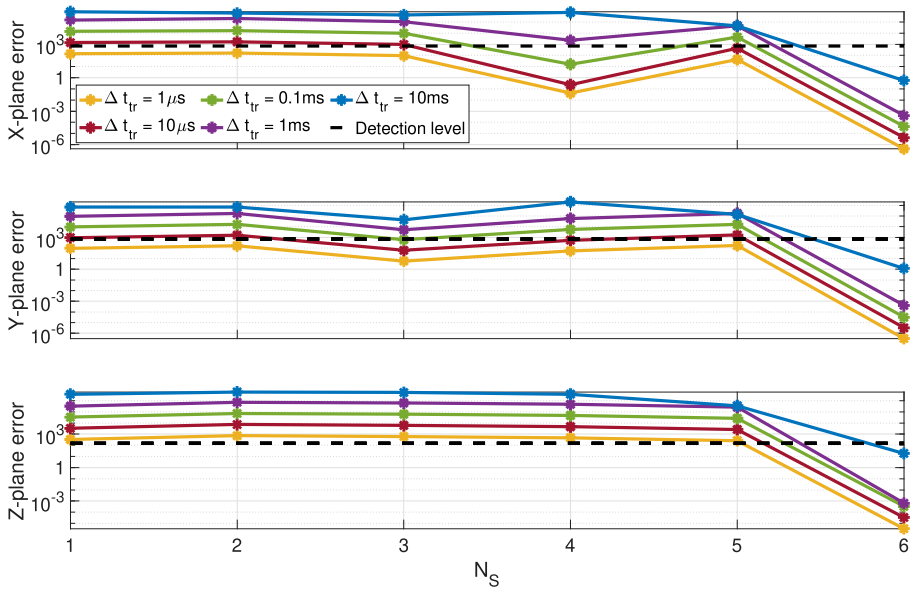

Figure 5.2. Receivers' position errors in meters are given on each plane with respect to the values of $N_S \geq N_A$ for Δt_{tr}.

After a GPS satellite time stamp attack is generated, the results are demonstrated in Figure 5.3 with respect to the chosen Δt^s in Table 5.1 for $N_S \geq N_G$. As the main difference of the results for Δt^s and t_{tr} based attacks, many scenarios violate the receivers' location without detecting in t^k based GPS time stamp attacks. We can perceive that the most threatening attack is the $N_S = 1$ scenario due to its simplification since only a single copy of GPS-like signals can be very effective for spoofing instead of sending a combination of various GPS-like signals. This type of attack may effectively falsify the users' location even for increased t^s errors, e.g., $\Delta t^s = 4$. This type of attack can be used as the initial step for convincing a receiver, then the number of spoofing signals and the amount of Δt^s can be slowly increased. Spoofers can significantly falsify receivers' location information with the incremental number of spoofed signals even if there are some reductions for the cases $N_S = 4$ and $N_S = 5$ for Y and Z-planes, respectively.

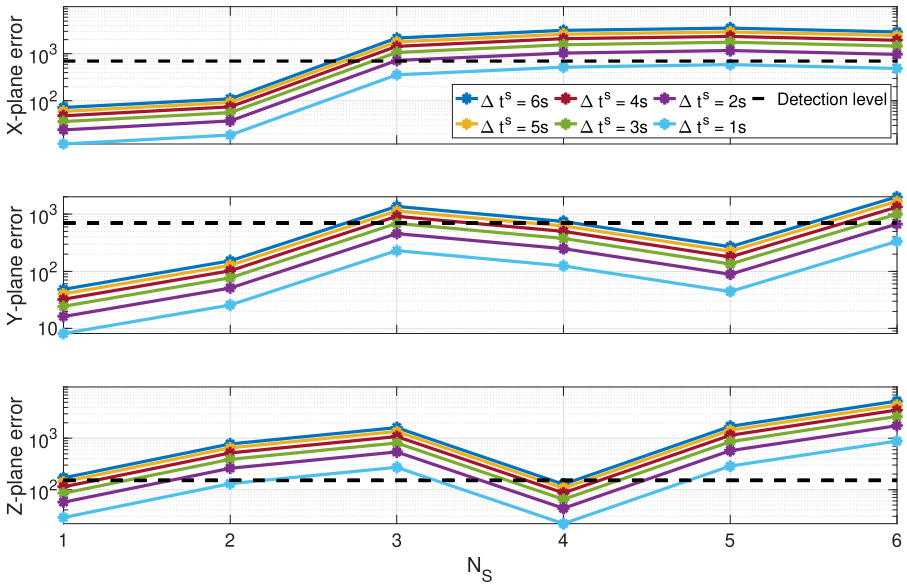

Figure 5.3. Receivers' position errors in meters are given on each plane with respect to the values of $N_S \geq N_G$ for Δt^s.

5.3. Derivation of Hyperbola Equations

The main structure of our detection strategies of spoofing attacks for V2X systems lies in solving hyperbola equations, which are generated by utilizing available pseudorange values with the single difference operation. Before deriving hyperbola equations, the main operation, which is the first step of the derivation, should be explained. In the literature, the pseudorange-based spoofing detection schemes include two fundamental operations, which are single and second differences [61]. However, in this thesis, we redefine the single difference as the difference of two pseudorange values of the kth and l^{th} satellites with respect to the ith receiver to utilize in hyperbolic equations as

$$\Delta_i^{kl} = \rho_i^k - \rho_i^l. \tag{5.1}$$

It should be noted that this usage of the single difference enables us to detect spoofing signals without the aid of an additional vehicle or an RSU. Due to the existence of s_i^s in Equation (3.9), single difference expression will be varied with respect to possible spoofing scenarios. For that reason, three possible scenarios should be investigated separately. These scenarios are the single difference between two authentic pseudoranges

(Case 1), the single difference between one authentic pseudorange and one spoofed pseudorange (Case 2), and the single difference between two spoofed pseudoranges (Case 3).

5.3.1. Single Difference Between Two Authentic Pseudoranges (Case 1)

When there is no spoofing attack, Equation (5.1) expression can be rewritten based on the pseudorange expression given in Equation (3.5) as follows

$$\Delta_i^{kl} = d_i^k - d_i^l + c(\tau^l - \tau^k) + \Delta m_i^{kl} + \Delta T_i^{kl} + \Delta I_i^{kl} + \Delta M_i^{kl}, \tag{5.2}$$

where (x^k, y^k, z^k) is the kth satellite position, (x_i, y_i, z_i) is the target vehicle position, $\Delta m_i^{kl} = m_i^k - m_i^l$, $\Delta T_i^{kl} = T_i^k - T_i^l$, $\Delta I_i^{kl} = I_i^k - I_i^l$, and $\Delta M_i^{kl} = M_i^k - M_i^l$. When we rearrange and take the square of both sides, the hyperbola equation for the ith vehicle with the kth and l^{th} satellite signals can be written as

$$\begin{aligned}(d_i^l)^2 = (\Delta_i^{kl} - c\Delta\tau^{lk} - \Delta m_i^{kl} - \Delta T_i^{kl} - \Delta I_i^{kl} - \Delta M_i^{kl})^2 + (d_i^k)^2 \\ - 2(\Delta_i^{kl} - c\Delta\tau^{lk} - \Delta m_i^{kl} - \Delta T_i^{kl} - \Delta I_i^{kl} - \Delta M_i^{kl})d_i^k,\end{aligned} \tag{5.3}$$

where $\Delta\tau^{lk} = \tau^l - \tau^k$ and (x^l, y^l, z^l) is the position of the lth satellite. After rearranging Equation (5.3), finally, the hyperbola equation can be rewritten for Case 1 as

$$\begin{aligned}&\sqrt{(x^k - x_i)^2 + (y^k - y_i)^2 + (z^k - z_i)^2} = \\ &\frac{(x^k)^2 - (x^l)^2 - 2(x^k - x^l)x_i + (y^k)^2 - (y^l)^2 - 2(y^k - y^l)y_i + (z^k)^2 - (z^l)^2 - 2(z^k - z^l)z_i}{2(\Delta_i^{kl} - c\Delta\tau^{lk} - \Delta m_i^{kl} - \Delta T_i^{kl} - \Delta I_i^{kl} - \Delta M_i^{kl})} \\ &+ \frac{(\Delta_i^{kl} - c\Delta\tau^{lk} - \Delta m_i^{kl} - \Delta T_i^{kl} - \Delta I_i^{kl} - \Delta M_i^{kl})}{2}.\end{aligned} \tag{5.4}$$

5.3.2. Single Difference Between One Authentic Pseudorange and One Spoofed Pseudorange (Case 2)

When we derive the hyperbola equation under spoofing attack, there are two possible cases for the pseudorange expressions because of the spoofing imperfection

term. The first case is deriving a hyperbola equation of a single difference of one spoofed pseudorange and one unspoofed pseudorange value with respect to Equation (3.5) and Equation (3.9), respectively. In this case, the single difference can be written as

$$\Delta_i^{ks} = d_i^k - d_i^s + c(\tau^s - \tau^k) - \hat{d}^k + \Delta m_i^{ks} + \Delta T_i^{ks} + \Delta I_i^{ks} + \Delta M_i^{ks} + \epsilon_{i,si}^s, \tag{5.5}$$

where $\Delta m_i^{ks} = m_i^k - m_i^s$, $\Delta T_i^{ks} = T_i^k - T_i^s$, $\Delta I_i^{ks} = I_i^k - I_i^s$, and $\Delta M_i^{ks} = M_i^k - M_i^s$. Similar to Case 1, the single difference is rearranged, and taking the square of both sides, the expression can be written as

$$\begin{aligned}(d_i^s)^2 = (\Delta_i^{ks} - c\Delta\tau^{sk} + \hat{d}^k - \Delta m_i^{ks} - \Delta T_i^{ks} - \Delta I_i^{ks} - \Delta M_i^{ks} - \epsilon_{i,si}^s)^2 + (d_i^k)^2 \\ - 2(\Delta_i^{ks} - c\Delta\tau^{sk} + \hat{d}^k - \Delta m_i^{ks} - \Delta T_i^{ks} - \Delta I_i^{ks} - \Delta M_i^{ks} - \epsilon_{i,si}^s)d_i^k.\end{aligned} \tag{5.6}$$

Finally, the hyperbola equation can be rewritten after rearranging Equation (5.6) for Case 2 as

$$\begin{aligned}&\sqrt{(x^k - x_i)^2 + (y^k - y_i)^2 + (z^k - z_i)^2} = \\ &\frac{(x^k)^2 - (x^s)^2 - 2(x^k - x^s)x_i + (y^k)^2 - (y^s)^2 - 2(y^k - y^s)y_i + (z^k)^2 - (z^s)^2 - 2(z^k - z^s)z_i}{2(\Delta_i^{ks} - c\Delta\tau^{sk} - \Delta m_i^{ks} - \Delta T_i^{ks} - \Delta I_i^{ks} - \Delta M_i^{ks} - \epsilon_{i,si}^s)} \\ &+ \frac{(\Delta_i^{ks} - c\Delta\tau^{sk} - \Delta m_i^{ks} - \Delta T_i^{ks} - \Delta I_i^{ks} - \Delta M_i^{ks} - \epsilon_{i,si}^s)}{2}.\end{aligned} \tag{5.7}$$

5.3.3. Single Difference of Two Spoofed Pseudorange Values (Case 3)

In this case, two pseudoranges are spoofed, whereas their expression is given in Equation (3.9). Under this condition, the single difference can be written as

$$\Delta_i^{sp} = d_i^s + d_i^p + c(\tau^p - \tau^s) + \Delta m_i^{sp} + \Delta T_i^{sp} + \Delta I_i^{sp} + \Delta M_i^{sp} + \Delta\epsilon_{i,si}^{sp}, \tag{5.8}$$

where s and p are the indices of spoofed pseudoranges, $\Delta m_i^{sp} = m_i^s - m_i^p$, $\Delta T_i^{sp} = T_i^s - T_i^p$, $\Delta I_i^{sp} = I_i^s - I_i^p$, $\Delta M_i^{sp} = M_i^s - M_i^p$, and $\Delta\epsilon_{i,si}^{sp} = \epsilon_{i,si}^s - \epsilon_{i,si}^p$. Similar to Case 2, after rearranging and taking the square of both sides, the single differences between

two pseudoranges becomes

$$\begin{aligned}(d_i^p)^2 = (\Delta_i^{sp} - c\Delta\tau^{ps} - \Delta m_i^{sp} - \Delta T_i^{sp} - \Delta I_i^{sp} - \Delta M_i^{sp} - \Delta\epsilon_{i,si}^{sp})^2 + (d_i^s)^2 \\ - 2(\Delta_i^{sp} - c\Delta\tau^{ps} - \Delta m_i^{sp} - \Delta T_i^{sp} - \Delta I_i^{sp} - \Delta M_i^{sp} - \Delta\epsilon_{i,si}^{sp})d_i^s.\end{aligned} \tag{5.9}$$

Finally, the hyperbola equation can be rewritten in the open form for Case 3 as

$$\begin{aligned}&\sqrt{(x^s - x_i)^2 + (y^s - y_i)^2 + (z^s - z_i)^2} = \\ &\frac{(x^s)^2 - (x^p)^2 - 2(x^s - x^p)x_i + (y^s)^2 - (y^p)^2 - 2(y^s - y^p)y_i + (z^s)^2 - (z^p)^2 - 2(z^s - z^p)z_i}{2(\Delta_i^{sp} - c\Delta\tau^{ps} - \Delta m_i^{sp} - \Delta T_i^{sp} - \Delta I_i^{sp} - \Delta M_i^{sp} - \Delta\epsilon_{i,si}^{sp})} \\ &+ \frac{(\Delta_i^{sp} - c\Delta\tau^{ps} - \Delta m_i^{sp} - \Delta T_i^{sp} - \Delta I_i^{sp} - \Delta M_i^{sp} - \Delta\epsilon_{i,si}^{sp})}{2}.\end{aligned} \tag{5.10}$$

In the scope of this chapter, Equation (5.4), Equation (5.7), Equation (5.10) are the fundamental equations in the design process of the proposed spoofing detection algorithms based on amounts of N_S and N_A. Once a set of hyperbola equations is created with H equations, their joint solution with respect to (x_i, y_i, z_i) provides a position of a receiver in the (x, y, z) axis. In Figure 5.4 and Figure 5.5, exemplary position outcomes are given when $N = 5$, and $N_A = 5$ in (a) and $N_A = 4$ (in other words $N_S = 1$) in (b). In these figures, the impact of a spoofing attack due to the additional time error is clear in terms of the distance error.

5.4. Spoofing Detection Strategies

Since the impact of a spoofing signal is apparent in Figure 5.4 and Figure 5.5, this difference can be utilized for designing a spoofing algorithm. In this thesis, we focus on developing search-based spoofing detection algorithms with the help of hyperbola equations for the V2X systems. In the case of a spoofing attack, the solution of a set of hyperbola equations is notably different from the case when there is no spoofing, as already explained. It should also be noted that the amount of the position error is highly correlated with the additional time information error due to spoofing. In this section, we primarily explain the sub-optimal search-based spoofing detection algorithm

(Algorithm 1), which searches all possible numbers of spoofing signals. After that, a more efficient version of this algorithm, which is named subset selection-based spoofing detection algorithm (Algorithm 2), with the smart selection of the spoofing signal subsets is given to reduce the complexity of Algorithm 1. Corresponding parameters are given in Table 5.2.

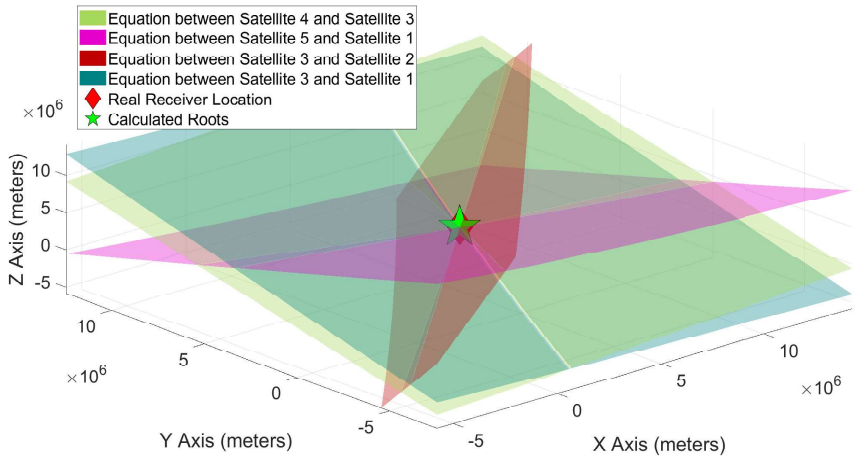

Figure 5.4. The real and calculated vehicle positions using hyperbolic equations with five authentic satellites.

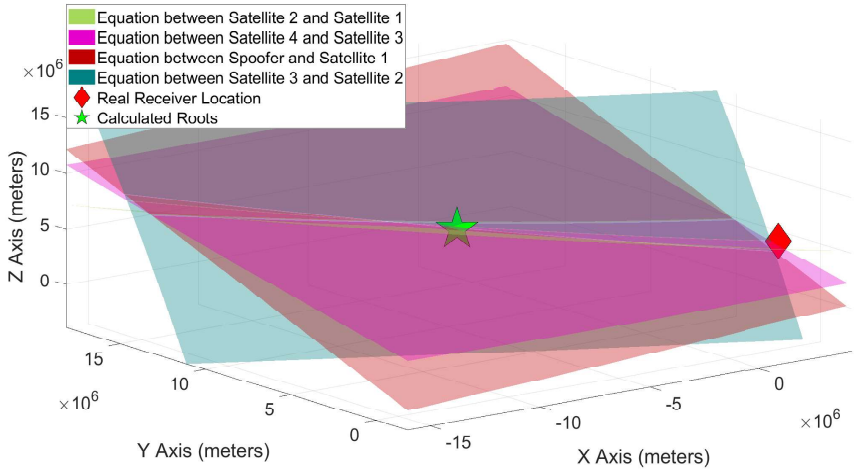

Figure 5.5. The real and calculated vehicle positions using hyperbolic equations with four authentic satellites and one spoofer.

Table 5.2. The list of remaining parameters and the notation of Chapter 6 in addition Table 3.1.

Notation	Definition	Condition
$\mathcal{N}^M$	Subset of $\mathcal{N}$ with M elements	
m	Required number of equations to detect N_S spoofer (possible spoofing attacks)	$m = \sum_{N_S=N}^{0} \binom{N}{N_S}$
$\bar{m}$	Required number of equations to detect N_S spoofer for sub-optimal spoofing detection algorithm	Equation (5.11), where Equation (5.12) is applied
g	Index of subsets with M elements	$g = 1 \dots m/\bar{m}$
x_0, y_0, z_0	Reference vehicle coordinates	(0,0,0)
x_c^g, y_c^g, z_c^g	Calculated coordinates of gth subset	
e_g	Error between (x_0, y_0, z_0) and (x_c^g, y_c^g, z_c^g)	
$\mathbf{e}$	Error vector of the subset	$\mathbf{e} = [e_1, \dots, e_m]$
H	Number of hyperbola equations	$H \geq 4$
λ	Detection threshold	From the set Λ
d_i	Local decision of the ith vehicle or RSU	$d_i \in \{0, 1\}$
$P_{D,i}/P_{F,i}$	Detection / False alarm rate of the ith vehicle or RSU	$0 \leq P_{D,i}/P_{F,i} \leq 1$
Ψ_{df}	Decision fusion methodology	$\Psi_{df} \in \{CV, CVR\}$
ϵ_b	Binary symmetric channel error rate	$\epsilon_b \in [0, 1]$

5.4.1. Algorithm 1: Sub-optimal Search-based Spoofing Detection Algorithm

The first proposed algorithm is based on a sub-optimally exhaustive search for each number of possible spoofing signals. The working principle of Algorithm 1 is creating subsets with M elements of available signals and deriving hyperbola equations for each subset, as given in Figure 5.6. The pseudocode of the proposed sub-optimal search-based spoofing detection algorithm is given in Figure 5.7, where the complexity of this algorithm is calculated as $\mathcal{O}(n^3\sqrt{n})$.

As the first step, subset selection with M elements should be completed from the N_G available signals. After that, a set of hyperbola equations can be constructed with H equations. Each equation from this set is derived for two randomly chosen pseudoranges. The correct modeling of the hyperbola equations should be made by

utilizing Equation (5.4), Equation (5.7), and Equation (5.10) based on the spoofed and unspoofed pseudorange scenarios. Due to the randomness, we prefer to call this algorithm a sub-optimal search-based spoofing detection algorithm. The detection process is based on the comparison of the solutions of H equations for the subset size M. The spoofing attack can be detected after thresholding a specific ratio. This ratio equals dividing the mean of all solutions of the chosen subset to the minimum of the same solutions. If any spoofer is not detected, the subset size is reduced as $M = M-1$, and the algorithm operates in the same manner with smaller subsets until at least one of the two following conditions is satisfied. The first condition is that spoofing is detected for $M \geq 4$. The second condition is $M < 4$, which means that there are not enough signals in the subset to be solved with hyperbola equations.

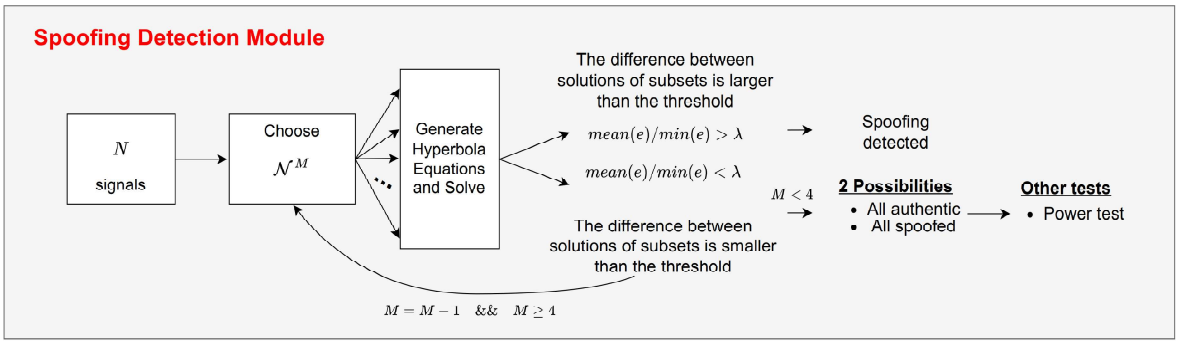

Figure 5.6. The operation block diagram of spoofing detection module, which may run both Algorithm 1 or Algorithm 2.

When Algorithm 1 ends without a spoofing detection, two possible outcomes can be encountered due to the nature of the algorithm. The first one is that there is no spoofing attack labeled as "all authentic". The second possible outcome is that all the possible signals are spoofed. The main reason behind this contrast is based on the subtraction of the two pseudorange values given in Equation (5.1). When two spoofed pseudorange values are subtracted, the additional spoofing errors in pseudorange values given in Equation (3.9) are mostly disappeared. Hence, a positioning error is not encountered. It can be considered a major drawback for our algorithm, but we assume that at least one authentic GNSS signal is available thanks to other trusted V2X

elements, e.g., an RSU, or improved receivers' sensitivity, which provides the existence of both authentic and spoofed signals in the wireless medium.

It should also be noted that symmetric spoofing attacker scenarios are observed. These scenarios are with the symmetric amount of authentic and spoofing signals, e.g., $N_{A,i} = N_{S,j}$ and $N_{S,i} = N_{A,j}$. Since these cases may result in a similar amount of pseudorange time errors, they provide similar position information in terms of spoofing detection. Hence, these symmetric spoofing tuples, e.g., $(1,7),(2,6)$ are jointly analyzed, as given in Table 5.3.

$M \leftarrow N-1$
while (mean(**e**) / min(**e**) $< \lambda$) & $M \geq 4$ **do**
  for $g = 1, \ldots, \binom{N}{M}$ **do**
    for $h = 1, \ldots, H$ **do**
      $k \leftarrow$ choose from $\mathcal{N}^M$
      $l \leftarrow$ choose from $\{\mathcal{N}^M - \{k\}\}$
      $Eq(h) \leftarrow$ hyperbola equation for $(\rho_k - \rho_l)$
    Eq $\leftarrow [Eq(1), \ldots, Eq(H)]$
    $x_c^g, y_c^g, z_c^g \leftarrow solve(\mathbf{Eq})$
    $e_g = |\sqrt{(x_c^g - x_0)^2 + (y_c^g - y_0)^2 + (z_c^g - z_0)^2}|$
  e $= [e_1, e_2, \ldots, e_{\binom{N}{M}}]$
  if (mean(**e**) / min(**e**)) $> \lambda$ **then**
    Spoofing detection = True
  else
    $M = M - 1$
break

Figure 5.7. Pseudocode of sub-optimal search-based spoofing detection algorithm (Algorithm 1).

An example operation of Algorithm 1 for $N = 8$ can be given as follows. If there is a spoofer with the satellite index 1, e.g., $s = 1$, the solution of the hyperbola equations for all the subsets with the satellite index 1 provides approximately similar positioning coordinates. However, one subset, which is $\{2,3,4,5,6,7,8\}$, results in a

distinct solution from the previous solutions in terms of magnitude. As a result, after searching eight subsets with seven elements, we can realize that there is an additional error due to a spoofing attack, as listed in Table 5.3. In another case, where there are two spoofing signals, where $s = 7$ and $p = 8$, all the results for $M = 7$ provide similar positioning information based on the distance from the reference vehicle position. When we decrease the size of the subset as $M = 6$, the solution of one of the subsets, which is $\{1, 2, 3, 4, 5, 6\}$ to be precise, is highly different from the other solutions. Hence, we can again detect the spoofing attack after analyzing available 28 subsets with 6 elements. This procedure can continue until $M < 4$. If all of the solutions for each subset have the same degree of error to the reference vehicle position, no spoofing attack is detected. In this case, further investigations are required, such as one of the power tests that are proposed in the literature.

Table 5.3. m and $\bar{m}$ values for sub-optimal search-based and subset selection-based spoofing detection algorithms for $N = 8$.

Detection Capability (Number of Spoofer)	**Algorithm 1 (Sub-optimal search-based Spoofing Detection)**		**Algorithm 2 (Subset selection-based Spoofing Detection)**	
	Subsets with M element	m	**Subsets with M element**	$\bar{m}$
1 or 7	All possible subsets with 7 elements ($\{1,2,3,4,5,6,7\}, \ldots,$ $\{2,3,4,5,6,7,8\}$)	8	$\{1,2,3,4,5,6\}, \{1,2,4,5,7,8\},$ $\{1,3,5,6,7,8\}, \{2,3,4,6,7,8\}$	4
2 or 6	All possible subsets with 6 elements	28	$\{3,4,5,7,8\}, \{2,4,5,6,8\},$ $\{2,3,5,6,7\}, \{1,3,4,6,8\},$ $\{1,4,5,6,7\}, \{1,2,6,7,8\},$ $\{1,2,3,5,8\}, \{1,2,3,4,7\}$	8
3 or 5	All possible subsets with 5 elements	56	$\{1,2,3,6\}, \{1,2,4,7\}, \{1,2,5,8\},$ $\{1,3,4,5\}, \{1,3,7,8\}, \{1,4,6,8\}$ $\{1,5,6,7\}, \{2,3,4,8\}, \{2,3,5,7\}$ $\{2,4,5,6\}, \{2,6,7,8\}, \{3,4,6,7\}$ $\{3,5,6,8\}, \{4,5,7,8\}$	14
4	All possible subset with 4 elements	70	Not capable to detect 4 spoofers	

Algorithm 1 is not simple in terms of computational complexity; hence, it may be considered that it is not suitable for lightweight traffic elements due to searching all possible spoofing signal numbers. To be clear, m values, which are the required number of equations to understand the existence of spoofing signals are given in Table 5.3. As a result, Algorithm 1 is more suitable for devices with power connections or large batteries. Since m values are increasing for such cases, the algorithm complexity incrementally rises. In order to avoid this situation, a smart subset selection-based version of this algorithm is proposed in the next section as Algorithm 2. It should be reminded that Algorithm 1 does not search all possible pairs for the hyperbola equations. Instead, it searches all possible subsets with M elements but selects pseudorange indices randomly.

5.4.2. Algorithm 2: Subset Selection-based Spoofing Detection Algorithm

Since there are many repetitive subsets and derivations of the same hyperbola equations, efficient reusing the same subsets for multiple spoofing attack scenarios can significantly fasten our detection algorithm, while decreasing the subset size. This performance increase is also required to deploy in battery-limited devices and sensors.

In order to find the optimal usage of subsets, we first decrease the size of the subsets from $M-1$ to $M-2$. After that, similar to Algorithm 1, a group of the subset is listed. Instead of listing all the subsets with $M-2$ elements, we try to find the best group of subsets with $\bar{m}$ elements as

$$\bar{m} = \sum_{N_S=0}^{N} \bar{m}(N_S), \tag{5.11}$$

where $\bar{m}(N_S)$ is bounded as

$$\begin{cases} \lceil \frac{\binom{M}{N_S}}{N_S+1} \rceil \geq \bar{m}(N_S) \geq \frac{\binom{M}{N_S} - \bar{m}(N_S-1)}{N_S+1} & \text{for } N_S \leq \lceil \frac{N}{2} \rceil \\ \bar{m}(N_S) = \bar{m}(N-N_S) & \text{otherwise.} \end{cases} \tag{5.12}$$

In Table 5.3, the proposed subsets are listed, while reusing the same spoofing attack scenarios is tried to be minimized. It should also be noted that a different subset selection can also be made. To avoid reusing, the subsets should be as orthogonal as possible in the multi-dimensional plane.

When a previous example is studied for Algorithm 2, four subsets with six elements are sufficient to detect a spoofer. As shown in Table 5.3, the solution of the set of hyperbola equations for the subset $\{1, 2, 3, 4, 5, 6\}$ is highly different from the other three subsets, when $s \in \{7, 8\}$. This subset can also be used when two spoofers are indicated as $s = 7$ or $p = 8$. On the other hand, subsets with $M = 5$ elements are required for most of the cases, e.g., $s = 1$ and $p = 6$. When compared with Algorithm 1, the number of subsets is dramatically reduced with a cost of non-capability of the detecting spoofing attacks in case 4 attacker signals. Beyond this complexity reduction, predefined sets can be utilized or similar sets with the same features can be easily selected for real deployments of V2X applications to further performance enhancements. A footnote is that the pseudocode of Algorithm 2 is almost identical to Algorithm 1, except for the changes $M = N - 2$ at the first line, and $\bar{m}$ instead of m.

5.4.3. Hybrid Spoofing Detection Methodology

Previously introduced spoofing detection algorithms are based on searching subsets of the GNSS signals for the spoofing scenarios. As mentioned before, Algorithm 1 is power-hungry with an increased number of subsets compared to Algorithm 2, whereas GNSS receivers are already extremely power-consuming [22]. Consequently, these algorithms are diverse in terms of power and battery concerns. In addition, their detection and false alarm rates of them vary, whereas the performance discussion will be given in the next section. Hence, each algorithm has specific advantages and drawbacks. In this case, a joint and a high-level hybrid operation of both algorithms can be considered. Since RSUs are able to take advantage of running Algorithm 1 in terms of a high detection rate, they can also inform nearby vehicles about the existence of spoofers in the deployment area. Algorithm 2 should be preferred for the moving target

vehicles with a low detection rate; however, RSUs in the close range should inform the targets in a high-level hybrid operation model.

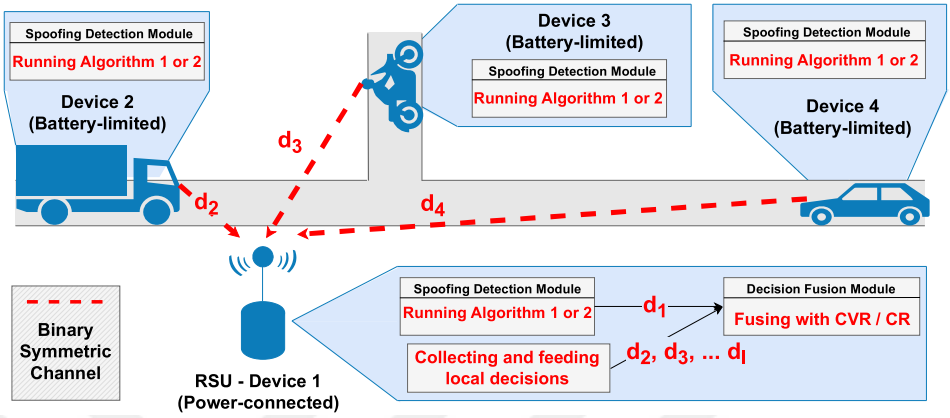

Figure 5.8. Hybrid spoofing detection operation of both of the proposed algorithms based on power sources of the devices for $I = 4$.

The proposed high-level hybrid spoofing detection methodology is based on fusing local decisions as given in Figure 5.8. After running one of the spoofing detection algorithms, each battery-limited vehicle sends its decision d_i to the power-connected RSU over an erroneous transmission channel, where the number of vehicles is $I-1$. The interconnection between V2X elements can be completed with any of the existing V2X communication technologies, which can be DSRC or LTE-V2X-based. At this unit, the RSU fuses individual spoofing detection decisions including itself at the decision fusion module, as shown in Figure 5.8. In the scope of this thesis, counting rule (CR) and Chair–Varshney rule (CVR) are considered detection fusion methodologies, which are well-known techniques in the literature [23, 24], given as

$$\begin{aligned} \Psi_{CR} &= \sum_{i=1}^{I} d_i \\ \Psi_{CVR} &= \sum_{i=1}^{I} \left[d_i \ln\left(\frac{P_D, i}{P_F, i}\right) - (1 - d_i) \ln\left(\frac{1 - P_D, i}{1 - P_F, i}\right) \right], \end{aligned} \tag{5.13}$$

where $P_{D,i}$ and $P_{F,i}$ are the detection and false alarm rates of ith vehicle or RSU, respectively. After running the fusion methodologies on a decision fusion module, spoofing attacks may be effectively detected. In this case, the RSUs should also inform the

nearby vehicles, which are located in the communication range with the deployed V2X communication system. With this hybrid methodology, improvements in the detection rate with a low false alarm ratio are possible. In this methodology, the friendly and trusted network elements, e.g., vehicles and RSUs, run the spoofing detection algorithms based on their power connections or battery sizes in spoofing detection modules. The communication between V2X elements can be provided with DSRC or LTE-V2X technologies with encrypted signals.

5.5. Simulation Results of Spoofing Detection Algorithms

The impacts of mobility in V2X communication systems are significant; hence, these impacts should be considered in our spoofing detection methodology, as the Reviewer suggested. In the literature on GNSS spoofing, mobility is not yet deeply studied. Therefore, as a reference model, we prefer to pursue the multipath error models in the urban environment for pseudoranges when the target vehicle moves at 50 km/h speed [91]. In this study, the position errors are mostly observed in terms of a few meters. In another study [92], which also utilizes [91] as the reference model for pseudorange error, they prefer the chosen error variances as $20m$ in the Gaussian distribution error model with non-zero mean. In our pseudorange model, we have modeled the mobility-based errors are also in Gaussian distribution with zero means and σ_m^2 variance. In the extended simulations, the value of σ_m^2 is chosen between $0m$ and $20m$ to understand the impact of mobility on the performance of our proposed spoofing detection algorithms.

The performances of the discussed spoofing detection algorithms are investigated with MATLAB simulations for $N = 8$. Hence, the number of spoofing signals is possible for the following condition $N_S \in \{0, 1, 2, \ldots, 8\}$. During the simulations, following parameters are diversified as $\Delta t^s \in \mathcal{T} = \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$, $\sigma_{si}^2 \in \mathcal{S} = \{0m, 50m, 100m, 200m\}$, and $\sigma_m^2 \in \mathcal{P} = \{0m, 5m, 10m, 20m\}$ with respect to pseudorange error models of moving vehicles [91, 92]. The set of atmospheric and multipath errors are $\sigma_i^2 \in \mathcal{I} = \{0m, 3m\}$, $\sigma_t^2 \in \xi = \{0m, 20m\}$, and

$\sigma^2_{mp} \in \zeta = \{0m, 60m, 150m\}$ based on the observational errors in the literature, where tropospheric, ionospheric, and multipath errors are modeled up to $3m$, $20m$, and $150m$, respectively [37–40,93]. The simulations are completed for at least 4×10^5 sets of equations for each Δt^s. These sets of equations include both spoofing and unspoofing scenarios for each algorithm. In the simulations, the set of GPS satellites is predefined; hence, GPS signal selection is random.

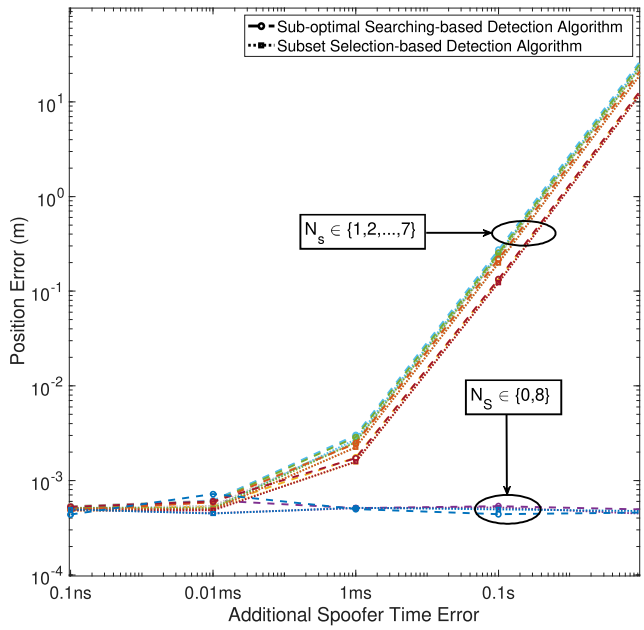

Figure 5.9. Average position error of proposed algorithms, Algorithm 1: sub-optimal search-based detection algorithm and its efficient version, Algorithm 2: subset selection-based detection algorithm for a various number of spoofers and chosen additional spoofer time error.

The performance of the algorithms is four-fold. In the first part, we compare Algorithms 1 and 2 in terms of their capabilities. Secondly, these algorithms are compared in terms of spoofing detection and false alarm rates for chosen detection threshold, λ, values. We evaluate the optimum value of λ for both algorithms by investigating the Pareto front with respect to spoofing detection and false alarm rates. In the third part, these algorithms are compared for chosen λ values with the detection-tree ML algorithm. Finally, the impacts of mobility and spoofing imperfections are discussed

in our spoofing detection algorithms. Before going into details of the simulation results, comparison motivation with the selected thresholds and with the decision-tree ML algorithm is explained in the next subsection.

Since our main strategy to detect spoofing attacks in V2X systems is based on searching for possible attack scenarios, machine learning algorithms may also be utilized. In the proposed spoofing detection algorithms, one of the main challenges is the value of the detection threshold λ. In Figure 5.9, the distance-based distinction can be intuitively made. However, this approach may not be sufficient when the attack scenario is not known. Hence, the selection of the λ for the detection is required for the generalized attack scenarios. In the literature on spoofing detection, there are various metrics that allow authors to evaluate the performance of the proposed algorithms. Some of these metrics can be listed as root mean square, detection delay, detection rate, and positioning error. In the submitted manuscript, we prefer to evaluate our spoofing detection algorithms with detection rate, and false alarm rates, which are very common in the literature [61, 94]. In order to find the best detection thresholds that optimize these metrics, the Pareto front is plotted with respect to detection and false alarm rates. With this approach, we have chosen better detection thresholds, and then the performance improvement of the spoofing detection algorithms is observed, especially in terms of false alarm rates. However, it should be reminded that the selection of the detection threshold is one of the parts of applicability in V2X. In real systems, the proper spoofing detection algorithms should be selected based on the power and complexity requirements of V2X elements. After the selection of the algorithm, the threshold can be fine-tuned to improve performance.

The Pareto front curves of the proposed algorithms are given in Figure 5.10. In this figure, the spoofing detection and false alarm rates of algorithms are plotted with respect to each other to find the best λ value with high detection and low false alarm rates. As readers may observe, the curves move to the left side of the figure when Δt^s decreases. The reason is that the spoofing error leads to less positioning error with low Δt^s, as shown in Figure 5.9. Since there is no perfect threshold for each value

of Δt^s, we choose the λ values, which are close to the lower turning points with high detection and low false alarm rates on the figure. As a result, the detection threshold λ is chosen in the set $\lambda = \{7.5 \times 10^5, 10^6, 2.5 \times 10^6\}$ based on the average distances between spoofed and unspoofed cases.

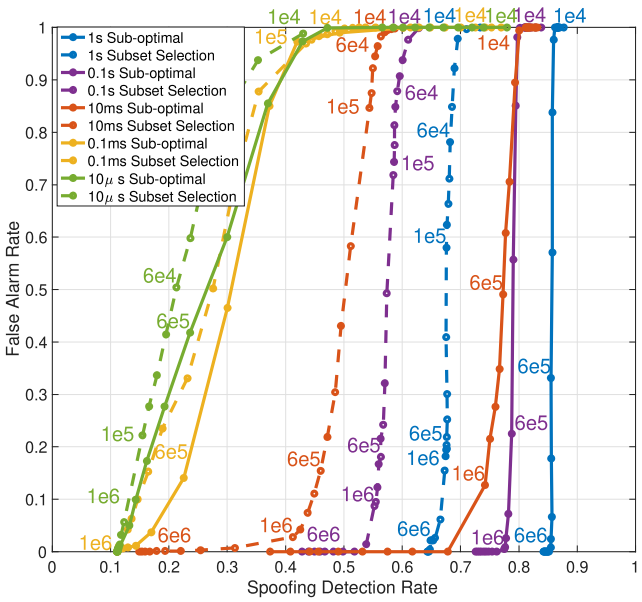

Figure 5.10. Pareto front curves of the spoofing detection algorithms with respect to spoofing detection rates and false alarm rates for additional spoofing time error $\Delta t^s \in \mathcal{T} = \{1s, 0.1s, 10ms, 0.1ms, 10\mu s\}$.

5.5.1. Comparison with Detection Thresholds and the Decision-tree Machine Learning Algorithm

With the help of decision-tree algorithms, proposed detection algorithms can be deployed with trained data for multiple attacking scenarios. This type of algorithm is well-suited to overcome the adjustment of the detection threshold issue since the aim is to classify the solution of the hyperbola equation as *spoofed* or *unspoofed* based on the training data. Another challenge is the number and size of subsets for the increasing number of N_G. In this case, the algorithmic complexity is a challenge for deployments in lightweight devices. With an ML-based decision tree algorithm, calculated solutions of the hyperbola equations can be classified without assigning the best decision threshold.

In the scope of this thesis, the training and ML model development is completed with the classification learner application of MATLAB. During the training process, the half of solutions of the hyperbola equations, which are obtained with MATLAB simulations, are utilized as the training dataset. Hence, the other half of the dataset is used as the control group.

5.5.2. Simulation Results

In Figure 5.9, the error performances of the proposed algorithms are compared. It can be seen that the error performances of these algorithms are very similar. As we already discussed, these algorithms are not capable of distinguishing the cases all signals are spoofed; or there is no spoofing attack. Besides, calculated average position errors are not distinct when the spoofing error is less than 10^{-5}s. As a result, the proposed algorithms are not able to detect spoofing attacks with very small spoofing time errors. It should also be noted that these figures do not guarantee that these algorithms are replaceable since they are compared for only average position error. Therefore, spoofing detection rates should be analyzed.

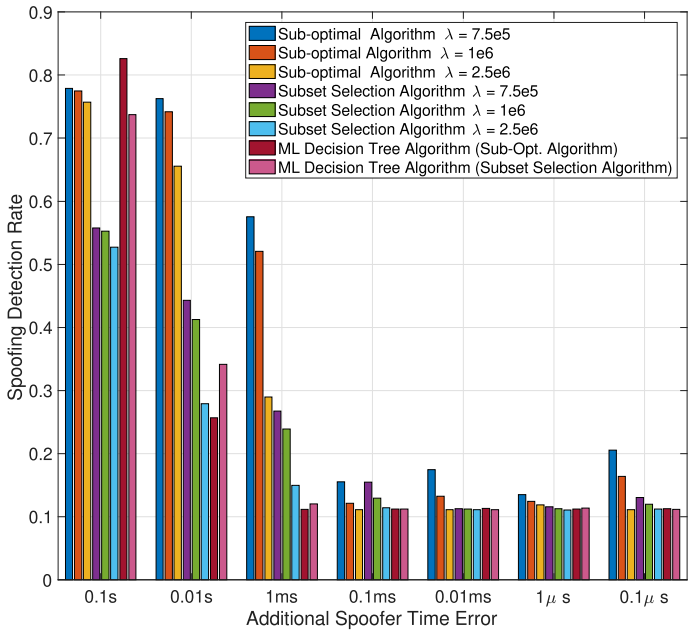

Figure 5.11. Spoofing detection rates of the detection algorithms.

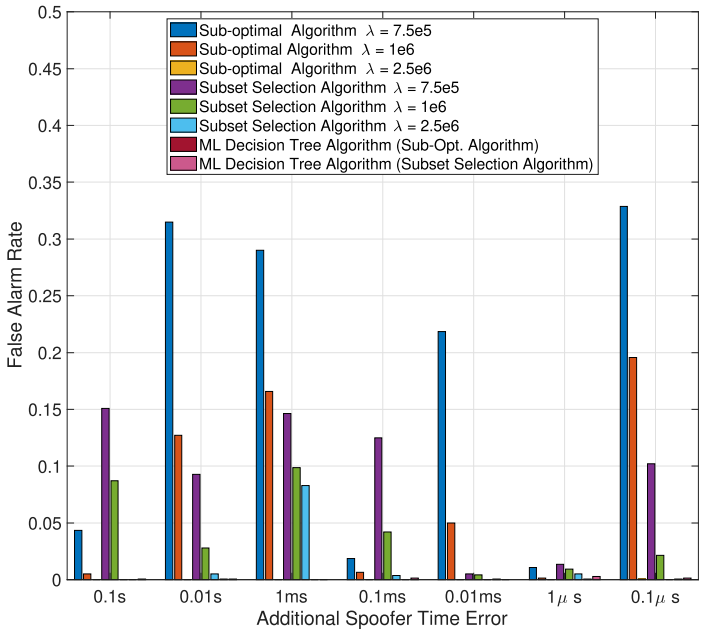

Figure 5.12. False alarm ratios of the spoofing detection algorithms.

The performance of the algorithms is given for a variety of the detection thresholds in Figure 5.11 with their false alarm ratios in Figure 5.12. In these figures, determined λ values are also compared with the ML-based detection tree algorithms. It can be easily observed that Algorithm 1 provides a better detection rate compared to Algorithm 2 when Δt^s is large enough. These two algorithms perform similarly when Δt^s decreases in terms of detection rate. These outcomes are contrary to the primary observations given in Figure 5.9. The reason is that the proposed set of subsets given in Table 5.3 is not able to solve all the attack scenarios with 4 spoofers. However, it tends to have a considerable false alarm ratio. The false alarm rate of Algorithm 2 is far better than Algorithm 1 in most cases thanks to the smart selection of the subsets during the detection process. It should be remembered that both of these algorithms are not able to distinguish all-spoofed signals and no spoofing signal cases that may reduce the performance of the algorithms. Beyond these two proposed algorithms, the performance of the ML algorithm provides the best false alarm rates, while satisfying high detection rates for high Δt^s and similar performance with the other algorithms for low Δt^s values with the subset selection-based algorithm.

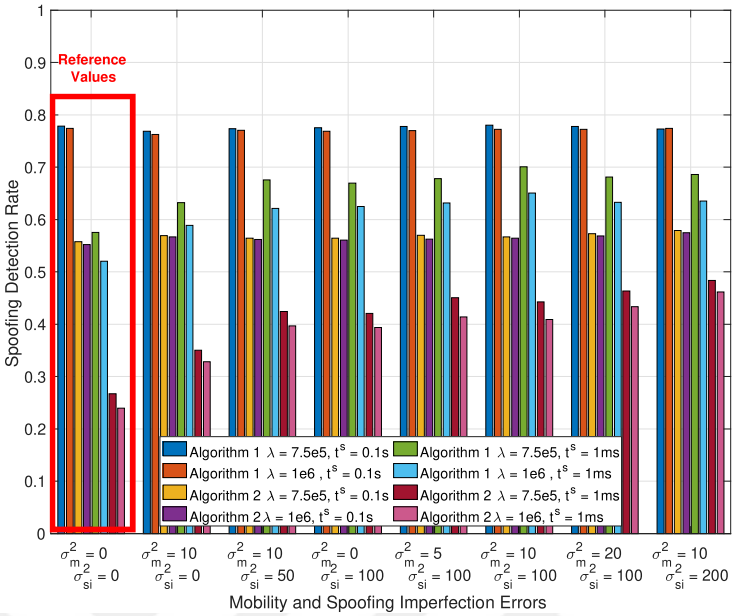

Figure 5.13. Spoofing detection rates of the detection algorithms with respect to mobility and spoofing imperfection errors comparing to reference values, which are firstly given in Figure 5.11 for $\Delta t^s \in \{0.1s, 1ms\}$ and $\lambda \in \{7.5 \times 10^5, 10^6\}$.

In Figure 5.13, and Figure 5.14, the impacts of mobility and spoofing imperfection errors on spoofing detection and false alarm rates of both algorithms are given, respectively. When the spoofing time error is high, e.g., $\Delta t^s = 10^{-1}$s, the spoofing detection rate does not change drastically, as shown in Figure 5.13. On the other hand, the impacts of these errors are visible in case the spoofing time error is low, e.g., $\Delta t^s = 10^{-3}$s. Compared to the reference values of $\sigma_m^2 = \sigma_{si}^2 = 0$, the spoofing detection rate increases with higher error variances. In addition, the false alarm rate is negatively affected when the error variance of mobility increases, as shown in Figure 5.14. This fashion is expected to be observed because unspoofed scenarios may be considered spoofed due to additional mobility terms, which leads to the erroneous receiver position. Besides, when only spoofing imperfections are considered, the false alarm rates are not negatively influenced.

The performances of spoofing detection algorithms are evaluated in Figure 5.15 with respect to considered spoofing types in the system model, which are ground-

located, UAV, HAPS, and LEO satellite spoofers. In the simulations, the additional atmospheric, multipath, and mobility-related errors are studied based on the spoofers' position, where the error parameters are given in Table 5.4. The results show that the spoofing detection rates increase when the attacker moves away from the target receiver. In these cases, the received spoofing signals are noisier compared to the ground spoofer; therefore, it is easier to detect spoofing signals. Contrary, the false alarm rates tend to rise due to the high amount of total noise in the system. In summary, the detection mechanisms lean to be completed with spoofing decisions because of the increased amount of total noise, and it is harder to distinguish unspoofed cases.

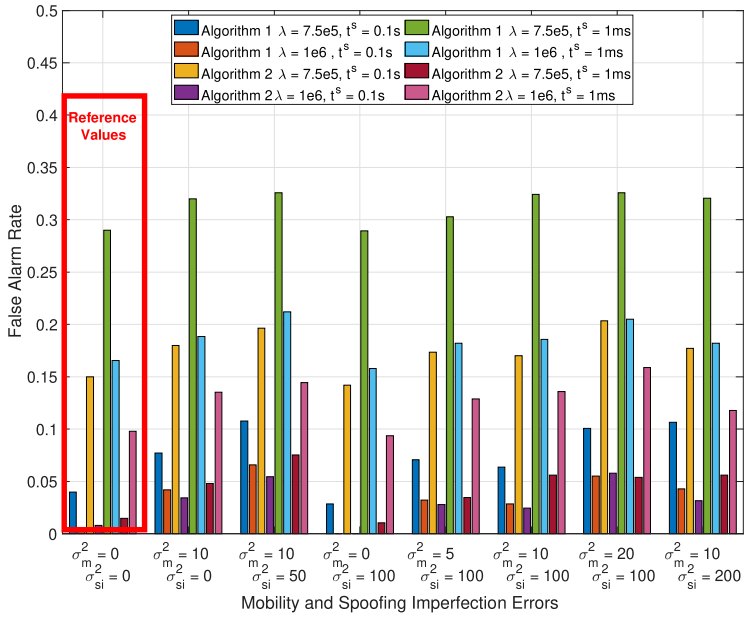

Figure 5.14. False alarm rates of the detection algorithms with respect to mobility and spoofing imperfection errors comparing to reference values, which are firstly given in Figure 5.12 for $\Delta t^s \in \{0.1s, 1ms\}$ and $\lambda \in \{7.5 \times 10^5, 10^6\}$.

Table 5.4. Error Parameters for various spoofing attacks.

Error Parameter	Ground-located	UAV	HAPS	LEO Satellite
σ_t^2	$0m$	$0m$	$3m$	$3m$
σ_i^2	$0m$	$0m$	$20m$	$20m$
σ_m^2	$0m$	$20m$	$20m$	$20m$
σ_{mp}^2	$0m$	$0m$	$60m$	$200m$

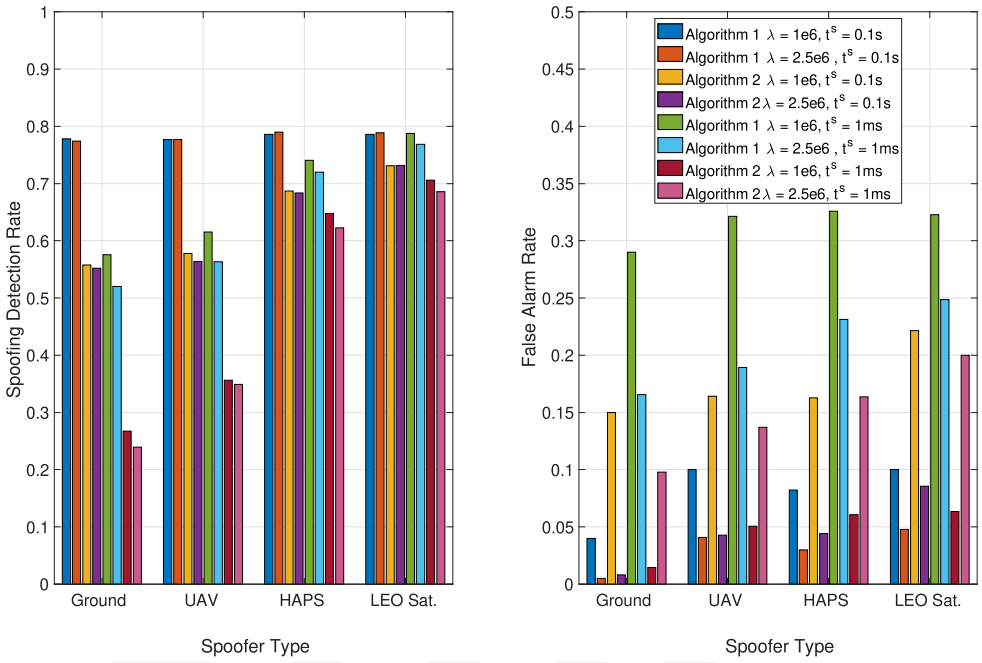

Figure 5.15. Spoofing detection rates of the detection algorithms with respect to spoofer types, which can be ground located, UAV, HAPS, or LEO satellite spoofers.

5.5.3. Decision Fusion for Hybrid Decision

As previously explained, a high-level hybrid approach can be applied to our system model with the help of decision fusion methodologies. In order to evaluate the efficiency of the hybrid approach, two different decision fusion methodologies are applied. The determined methodologies, which are CR and CVR, respectively, given in Equation (5.13), Equation (5.13), are simulated for $I \in \{2, 3\}$. In these simulations, $i = 1$ represents a power-connected RSU with a decision fusion module by running Algorithm 1 for $\Delta t^s = 0.01$ ms. The transporting vehicles are indexed as $i \in \{2, 3\}$, while these vehicles run Algorithm 2. Since vehicles suffer from mobility impacts on positioning, $\sigma_m^2 \in \{10\text{m}, 20\text{m}\}$ is assumed when $i \geq 2$. The erroneous transmission medium is modeled as a BSC with error probability ϵ_b since BSC can be easily applied for binary information. The simulations are repeated for various error rates, where $\epsilon_b \in \{0, 0.15, 0.3\}$. In Figure 5.16, the simulation results of hybrid spoofing detection are given. As a reminder, in this section, we only assume ground-located spoofers.

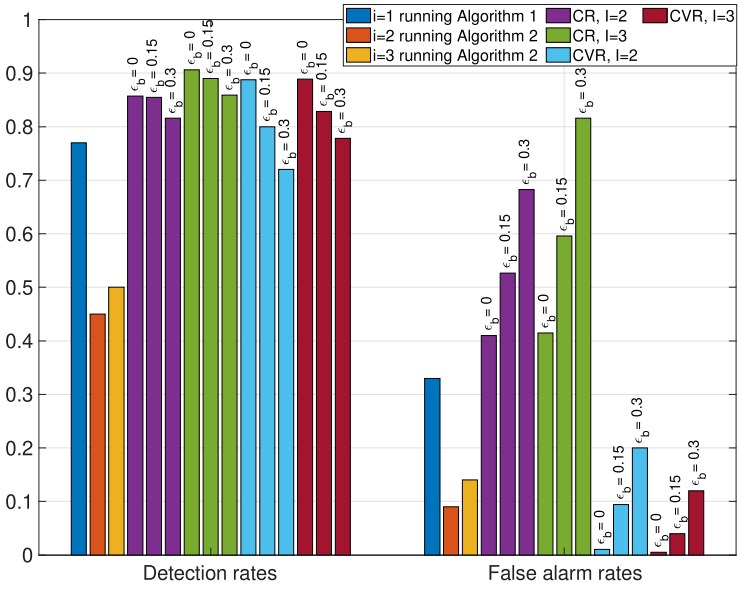

Figure 5.16. Spoofing detection and false alarm rates of CV and CVR are compared respective sole use of Algorithm 1 and Algorithm 2. Simulations are extended for $I = 2$ and $I = 3$ when three possible BSC errors.

The simulations on the performance of decision fusion methodologies show that detection rates tend to increase when multiple decisions are fused with the help of CR and CVR. As expected, the detection rates decrease when ϵ_b increases. On the other hand, the performance of the false alarm rates does not behave similarly to the decision fusion methodologies. In terms of false alarm rates, CR does not provide additional benefits, especially when ϵ_b increases. Fortunately, even if ϵ_b increases, CVR performs more effectively, where CVR satisfies lower false alarm rates than CV. Another observation is that CVR performs better with $I = 3$ agents than $I = 2$. This demonstrates that spoofing detection performance improves when more agents are part of the hybrid spoofing detection implementation. Overall, the readers may observe that fusing decisions with CVR are beneficial compared to running solely Algorithm 1 or 2 to detect GNSS spoofing attacks.

6. CONCLUSION

Within the IoT concept, CPS offer a great variety of applications, and these applications can be implemented with 5G and its successors. In the future, 6G will offer a new class of CPS applications (e.g., augmented reality, virtual reality, or holographic telepresence), but today's main applications of CPS are V2X, industrial wireless sensor networks, smart grids, and smart healthcare applications. In this thesis, we first evaluate the challenges, requirements, and security aspects of these applications with a novel flower approach. Each flower is a graphical illustration of diverse applications with various needs. The length of the roots is determined based on the extremeness of the operational requirements. As we also focus on the security aspect, the roots also represent security needs against attackers, which are harmful insects. There are numerous distinct sub-applications of each primary use of CPS; therefore, this garden may have thousands of distinguished flowers.

From this large garden, we focus on security issues in V2X applications in our system model. At this point, our focal points are spoofing and eavesdropping attacks, which may be very critical and lead to privacy leaks, system interruptions, and even life-threatening conditions. In order to fail the eavesdropper, a relay-aided error vector injection-based security enhancement mechanism is proposed for chosen FEC codes, which are Reed-Solomon, Golay, and Hamming codes. Security capability of the proposed system model is analyzed by evaluating the decoding error probability of an eavesdropper. Observed results on system performance, which are found via both simulations and analytical expressions, show that the number of significantly noisy channels between relays and Eve helps to achieve advanced security. Moreover, FEC codes help to provide error-free communication with Bob for an increased number of bit errors within the limit of decoding capability of FEC codes.

In order to mitigate spoofing attacks against satellite-based positioning systems, we consider a novel V2X system model, where not only ground spoofers but also aerial

spoofers are part of the network. Another significant contribution is that there are authentic GPS signals in addition to spoofing signals, while it is generally assumed that all the positioning signals are spoofed in the literature. Under these circumstances, we propose two spoofing detection algorithms. Firstly, we design Algorithm 1, which sub-optimally searches for subsets of the available signals at target vehicles. Due to the complexity and power consumption issues, this algorithm is improved by intelligently selecting subsets by reusing derived hyperbola equations in Algorithm 2. The results show that discussed algorithms have strong and weak features. Algorithm 1 provides a sufficient amount of detection rate; however, it has high complexity and false alarm ratio-based issues. Algorithm 2 has the lowest successful detection rate, but it has a low false alarm rate in most cases compared to Algorithm 1. It also has a complexity and power consumption advantage against Algorithm 1. Therefore, Algorithm 2 is more suitable for battery-limited lightweight devices in V2X systems. Besides, a supervised learning-based ML-supported detection algorithm is studied, and compared with the proposed detection thresholds. Even though an ML-based decision tree algorithm provides similar detection rates to Algorithm 2 with low Δt^s, it has very low false alarms. Hence, the selection of the algorithm is based on the V2X applications, the number of trusted authentic GPS signals, and the transmission environment. Besides, the mobility of a vehicle improves the detection performance of the algorithms, while it leads to higher false alarm rates. On the other hand, spoofing imperfections do not negatively affect false alarm rates in addition to improved detection performance. Beyond the conventional approach, we also consider aerial spoofers in our system model. These attack types are analyzed with additional atmospheric errors due to the high altitude of HAPS and LEO satellite spoofers. The results indicate that the proposed spoofing algorithms have high detection rates against these spoofers; however, they also result in increased false alarm rates due to significant noise.

In the scope of this thesis, fundamental hybrid spoofing detection methodology is discussed based on decision fusion methodologies. As discussed before, the performance of CV and CVR methodologies are compared with the sole use of Algorithm 1 and Algorithm 2, and it is shown that deploying CVR is more beneficial than single

algorithm operations. We believe that the implementation of CVR in dense deployments provides more success in terms of detection and false alarm rates. However, a hybrid implementation may not be possible all the time. Hence, Algorithm 1 may have to run on an agent without the cooperation of other agents on remote and less frequently-used roads. Similarly, each roadway may not have RSU infrastructure. As a result, Algorithm 2 should solely operate in battery-limited devices without RSU cooperation. Hence, the sole operations of the proposed algorithms are significant even if hybrid spoofing detection can be utilized in more dense deployments.

Beyond the hybrid approach, the individual performance of the proposed algorithms may not be sufficient for attack scenarios, e.g., all signals are spoofed cases. Our algorithm needs improvement, especially for all-spoofed signaling scenarios. In this case, a suitable detection mechanism should be integrated, e.g., pseudorange-based spoofing detection with multiple receivers. Besides, Algorithm 2 has lower spoofing detection capability due to the reduced size of subsets. The loss of this capability significantly impacts the spoofing detection rate; the design of the subsets can be adjusted without the loss of detection capability.

REFERENCES

1. Ashibani, Y. and Q. H. Mahmoud, "Cyber Physical Systems Security: Analysis, Challenges and Solutions", *Elsevier Computers & Security*, Vol. 68, pp. 81–97, 2017.

2. Humayed, A., F. Lin, J.and Li and B. Luo, "Cyber-Physical Systems Security—A Survey", *IEEE Internet of Things Journal*, Vol. 4, No. 6, pp. 1802–1831, 2017.

3. Wang, Z., H. Song, D. W. Watkins, K. G. Ong, P. Xue, Q. Yang and X. Shi, "Cyber-Physical Systems for Water Sustainability: Challenges and Opportunities", *IEEE Communications Magazine*, Vol. 53, No. 5, pp. 216–222, 2015.

4. Song, H., G. A. Fink and S. Jeschke, *Security and Privacy in Cyber-physical Systems: Foundations, Principles, and Applications*, John Wiley & Sons, New Jersey, 2017.

5. Zhu, J., Y. Zou and B. Zheng, "Physical-Layer Security and Reliability Challenges for Industrial Wireless Sensor Networks", *IEEE Access*, Vol. 5, pp. 5313–5320, 2017.

6. Harrison, W. K., J. Almeida, M. R. Bloch, S. W. McLaughlin and J. Barros, "Coding for Secrecy: An Overview of Error-Control Coding Techniques for Physical-Layer Security", *IEEE Signal Processing Magazine*, Vol. 30, No. 5, pp. 41–50, 2013.

7. Yang, N., 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.

8. Wyner, A. D., "The Wire-tap Channel", *Bell Labs Technical Journal*, Vol. 54, No. 8, pp. 1355–1387, 1975.

9. Psiaki, M. L. and T. E. Humphreys, "GNSS Spoofing and Detection", *Proceedings of the IEEE*, Vol. 104, No. 6, pp. 1258–1270, 2016.

10. Jansen, K., *Detection and Localization of Attacks on Satellite-based Navigation Systems*, Ph.D. Thesis, Ruhr-Universität Bochum, 2019.

11. Tippenhauer, N. O., C. Pöpper, K. B. Rasmussen and S. Capkun, "On the Requirements for Successful GPS Spoofing Attacks", *Proceedings of the 18th ACM Conference on Computer and Communications Security*, pp. 75–86, Chicago, Illinois, USA, 2011.

12. Stenberg, N., *Spoofing Mitigation Using Multiple GNSS-Receivers*, M.S. Thesis, Linköping University, 2019.

13. Karabulut Kurt, G. and H. Yanikomeroglu, "Communication, Computing, Caching, and Sensing for Next-Generation Aerial Delivery Networks: Using a High-Altitude Platform Station as an Enabling Technology", *IEEE Vehicular Technology Magazine*, Vol. 16, No. 3, pp. 108–117, 2021.

14. Karabulut Kurt, G., M. G. Khoshkholgh, S. Alfattani, A. Ibrahim, T. S. J. Darwish, M. S. Alam, H. Yanikomeroglu and A. Yongacoglu, "A Vision and Framework for the High Altitude Platform Station (HAPS) Networks of the Future", *IEEE Communications Surveys & Tutorials*, Vol. 23, No. 2, pp. 729–779, 2021.

15. Shibata, Y., N. Kanazawa, M. Konishi, K. Hoshino, Y. Ohta and A. Nagate, "System Design of Gigabit HAPS Mobile Communications", *IEEE Access*, Vol. 8, pp. 157995–158007, 2020.

16. Harris, M., "Tech Giants Race to Build Orbital Internet", *IEEE Spectrum*, Vol. 55, No. 6, pp. 10–11, 2018.

17. Mershad, K., H. Dahrouj, H. Sarieddeen, B. Shihada, T. Al-Naffouri and M.-S. Alouini, "Cloud-Enabled High-Altitude Platform Systems: Challenges and Op-

portunities", *Frontiers in Communications and Networks*, Vol. 2, p. 30, 2021.

18. Clements, Z., P. Ellis, M. Psiaki and T. E. Humphreys, "Geolocation of Terrestrial GNSS Spoofing Signals from Low Earth Orbit", *Proceedings of the 35th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+)*, pp. 3418–3431, Denver, Colorado, USA, 2022.

19. Dang, Y., C. Benzaïd, B. Yang, T. Taleb and Y. Shen, "Deep-ensemble-learning-based GPS Spoofing Detection for Cellular-Connected UAVs", *IEEE Internet of Things Journal*, Vol. 9, No. 24, pp. 25068–25085, 2022.

20. Borio, D. and C. Gioia, "A Sum-of-squares Approach to GNSS Spoofing Detection", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 52, No. 4, pp. 1756–1768, 2016.

21. Liu, K., W. Wu, Z. Wu, L. He and K. Tang, "Spoofing Detection Algorithm Based on Pseudorange Differences", *Sensors*, Vol. 18, No. 10, p. 3197, 2018.

22. Egea-Roca, D., M. Arizabaleta-Diez, T. Pany, F. Antreich, J. A. López-Salcedo, M. Paonni and G. Seco-Granados, "GNSS User Technology: State-of-the-Art and Future Trends", *IEEE Access*, Vol. 10, pp. 39939–39968, 2022.

23. Al-Jarrah, M. A., M. A. Yaseen, A. Al-Dweik, O. A. Dobre and E. Alsusa, "Decision Fusion for IoT-based Wireless Sensor Networks", *IEEE Internet of Things Journal*, Vol. 7, No. 2, pp. 1313–1326, 2019.

24. Tabella, G., N. Paltrinieri, V. Cozzani and P. S. Rossi, "Wireless Sensor Networks for Detection and Localization of Subsea Oil Leakages", *IEEE Sensors Journal*, Vol. 21, No. 9, pp. 10890–10904, 2021.

25. Demir, M. Ö., A. E. Pusane, G. Dartmann, G. Ascheid and G. Karabulut Kurt, "A Garden of Cyber Physical Systems: Requirements, Challenges, and Implementation Aspects", *IEEE Internet of Things Magazine*, Vol. 3, No. 3, pp. 84–89,

2020.

26. Demir, M. Ö., G. Karabulut Kurt, G. Dartmann, G. Ascheid and A. E. Pusane, "Security Analysis of Forward Error Correction Codes in Relay Aided Networks", *IEEE Global Information Infrastructure and Networking Symposium (GIIS)*, pp. 1–5, Thessaloniki, Greece, 2018.

27. Demir, M. Ö., O. A. Topal, G. Dartmann, A. Schmeink, G. Ascheid, G. Karabulut Kurt and A. E. Pusane, "Using Perfect Codes in Relay Aided Networks: A Security Analysis", *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pp. 1–6, Barcelona, Spain, 2019.

28. Demir, M. Ö., G. Karabulut Kurt and A. E. Pusane, "On the Limitations of GPS Time-Spoofing Attacks", *43rd IEEE International Conference on Telecommunications and Signal Processing (TSP)*, pp. 313–316, Milano, Italy, 2020.

29. Demir, M. Ö., G. Karabulut Kurt and A. E. Pusane, "A Pseudorange-based GPS Spoofing Detection Using Hyperbola Equations", *IEEE Transactions on Vehicular Technology*, Vol. 72, No. 8, pp. 10770–10783, 2023.

30. Hasan, M., S. Mohan, T. Shimizu and H. Lu, "Securing Vehicle-to-Everything (V2X) Communication Platforms", *IEEE Transactions on Intelligent Vehicles*, Vol. 5, No. 4, pp. 693–713, 2020.

31. Chen, S., J. Hu, Y. Shi, Y. Peng, J. Fang, R. Zhao and L. Zhao, "Vehicle-to-Everything (V2X) Services Supported by LTE-based Systems and 5G", *IEEE Communications Standards Magazine*, Vol. 1, No. 2, pp. 70–76, 2017.

32. Bloch, M. and J. Barros, *Physical-Layer Security: from Information Theory to Security Engineering*, Cambridge University Press, Cambridge, 2011.

33. McEliece, R. J., "A Public-key Cryptosystem Based on Algebraic", *Coding The-*

ory, Vol. 4244, pp. 114–116, 1978.

34. Bloch, M., M. Hayashi and A. Thangaraj, "Error-Control Coding for Physical-Layer Secrecy", *Proceedings of the IEEE*, Vol. 103, No. 10, pp. 1725–1746, 2015.

35. Low, K. S., W. N. N. Win and M. J. Er, "Wireless Sensor Networks for Industrial Environments", *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06)*, pp. 271–276, Vienna, Austria, 2005.

36. Mousa, A. E.-K., N. Aboualy, M. Sharaf, H. Zahra and M. Darrag, "Tropospheric Wet Delay Estimation Using GNSS: Case Study of A Permanent Network in Egypt", *NRIAG Journal of Astronomy and Geophysics*, Vol. 5, No. 1, pp. 76–86, 2016.

37. Osah, S., A. A. Acheampong, C. Fosu and I. Dadzie, "Evaluation of Zenith Tropospheric Delay Derived from Ray-traced VMF3 Product over the West African Region Using GNSS Observations", *Advances in Meteorology*, Vol. 2021, pp. 1–14, 2021.

38. Subirana, J. S., J. J. Zornoza and M. Hernández-Pajares, "Tropospheric Delay", `https://gssc.esa.int/navipedia/index.php/Tropospheric_Delay/`, 2011, accessed on May 13, 2023.

39. Subirana, J. S., J. J. Zornoza and M. Hernández-Pajares, "Ionospheric Delay", `https://gssc.esa.int/navipedia/index.php/Ionospheric_Delay/`, 2011, accessed on May 13, 2023.

40. Mallika, L., D. V. Ratnam, S. Raman and G. Sivavaraprasad, "Machine Learning Algorithm to Forecast Ionospheric Time Delays Using Global Navigation Satellite System Observations", *Acta Astronautica*, Vol. 173, pp. 221–231, 2020.

41. Sunehra, D., "Real-time Estimation of Ionospheric Delay Using Dual Frequency GPS Observations", *European Scientific Journal*, Vol. 9, No. 15, 2013.

42. Manfredini, E. G., D. M. Akos, Y.-H. Chen, S. Lo, T. Walter and P. Enge, "Effective GPS Spoofing Detection Utilizing Metrics from Commercial Receivers", *Proceedings of the Institute of Navigation International Technical Meeting*, pp. 672–689, Reston, Virginia, USA, 2018.

43. Schmidt, D., K. Radke, S. Camtepe, E. Foo and M. Ren, "A Survey and Analysis of the GNSS Spoofing Threat and Countermeasures", *ACM Computing Surveys (CSUR)*, Vol. 48, No. 4, p. 64, 2016.

44. Psiaki, M. L., B. W. OHanlon, S. P. Powell, J. A. Bhatti, K. D. Wesson and T. E. Schofield, "GNSS Spoofing Detection Using Two-antenna Differential Carrier Phase", *Proceedings of the 27th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2014)*, pp. 2776–2800, Tampa, Florida, USA, 2014.

45. Magiera, J. and R. Katulski, "Detection and Mitigation of GPS Spoofing Based on Antenna Array Processing", *Journal of Applied Research and Technology*, Vol. 13, No. 1, pp. 45–57, 2015.

46. Capuano, V., P. Blunt, C. Botteron and P.-A. Farine, "Orbital Filter Aiding of a High Sensitivity GPS Receiver for Lunar Missions", *NAVIGATION, Journal of the Institute of Navigation*, Vol. 64, No. 3, pp. 323–338, 2017.

47. Baldi, M., M. Bianchi and F. Chiaraluce, "Non-Systematic Codes for Physical Layer Security", *IEEE Information Theory Workshop*, pp. 1–5, Cairo, Egypt, 2010.

48. Yu, K., M. Gidlund, J. Åkerberg and M. Björkman, "Reliable and Low Latency Transmission in Industrial Wireless Sensor Networks", *Elsevier Procedia Com-*

puter Science, Vol. 5, pp. 866–873, 2011.

49. Yitbarek, Y. H., K. Yu, J. Åkerberg, M. Gidlund and M. Björkman, "Implementation and Evaluation of Error Control Schemes in Industrial Wireless Sensor Networks", *IEEE International Conference on Industrial Technology (ICIT)*, pp. 730–735, Busan, Korea (South), 2014.

50. Mukherjee, A., S. A. A. Fakoorian, J. Huang and A. L. Swindlehurst, "Principles of Physical Layer Security in Multi User Wireless Networks: A Survey", *IEEE Communications Surveys & Tutorials*, Vol. 16, No. 3, pp. 1550–1573, 2014.

51. Mo, J., M. Tao and Y. Liu, "Relay Placement for Physical Layer Security: A Secure Connection Perspective", *IEEE Communications Letters*, Vol. 16, No. 6, pp. 878–881, 2012.

52. Swamy, V. N., S. Suri, P. Rigge, M. Weiner, G. Ranade, A. Sahai and B. Nikolić, "Real-time Cooperative Communication for Automation over Wireless", *IEEE Transactions on Wireless Communications*, Vol. 16, No. 11, pp. 7168–7183, 2017.

53. Van Huynh, L., J. den Hartog and N. Zannone, "Security and Privacy for Innovative Automotive Applications: A Survey", *Computer Communications*, Vol. 132, pp. 17–41, 2018.

54. Cui, J., L. S. Liew, G. Sabaliauskaite and F. Zhou, "A Review on Safety Failures, Security Attacks, and Available Countermeasures for Autonomous Vehicles", *Ad Hoc Networks*, Vol. 90, p. 101823, 2019.

55. Faria, L. d. A., C. A. Silvestre, M. A. F. Correia and N. A. Roso, "GPS Jamming Signals Propagation in Free-space, Urban and Suburban Environments", *Journal of Aerospace Technology and Management*, Vol. 10, 2018.

56. Van der Merwe, J. R., X. Zubizarreta, I. Lukčin, A. Rügamer and W. Felber, "Classification of Spoofing Attack Types", *IEEE European Navigation Conference*

(ENC), pp. 91–99, Gothenburg, Sweden, 2018.

57. Sanders, C. and Y. Wang, "Localizing Spoofing Attacks on Vehicular GPS Using Vehicle-to-Vehicle Communications", *IEEE Transactions on Vehicular Technology*, Vol. 69, No. 12, pp. 15656–15667, 2020.

58. Guo, Y., M. Wu, K. Tang, J. Tie and X. Li, "Covert Spoofing Algorithm of UAV Based on GPS/INS-integrated Navigation", *IEEE Transactions on Vehicular Technology*, Vol. 68, No. 7, pp. 6557–6564, 2019.

59. Schrader, D. K., B.-C. Min, E. T. Matson and J. E. Dietz, "Combining Multiple, Inexpensive GPS Receivers to Improve Accuracy and Reliability", *IEEE Sensors Applications Symposium Proceedings*, pp. 1–6, Brescia, Italy, 2012.

60. Gowda, M., J. Manweiler, A. Dhekne, R. R. Choudhury and J. D. Weisz, "Tracking Drone Orientation with Multiple GPS Receivers", *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, pp. 280–293, New York City, New York, USA, 2016.

61. Wang, F., H. Li and M. Lu, "GNSS Spoofing Detection Based on Unsynchronized Double-antenna Measurements", *IEEE Access*, Vol. 6, pp. 31203–31212, 2018.

62. Xiao, L., X. Li and Y. Zeng, "GNSS Spoofing Detection Using Pseudo-range Single Difference between Two Receivers", *2nd Atlantis Press International Conference on Machinery, Electronics and Control Simulation (MECS)*, pp. 292–298, Taiyuan, China, 2017.

63. Jahromi, A. J., A. Broumandan and G. Lachapelle, "GNSS Signal Authenticity Verification Using Carrier Phase Measurements with Multiple Receivers", *8th IEEE ESA Workshop on Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing (NAVITEC)*, pp. 1–11, Noordwijk, Netherlands, 2016.

64. Zhang, Z. and X. Zhan, "GNSS Spoofing Network Monitoring Based on Differential Pseudorange", *Sensors*, Vol. 16, No. 10, p. 1771, 2016.

65. Potluri, S., *Hyperbolic Position Location Estimator with TDOA's from Four Stations*, M.S. Thesis, New Jersey Institute of Technology, Department of Electrical and Computer Engineering, 2001.

66. IMT 2020 (5G) Promotion Group, ""5G Vision"", `https://www.itu.int/dms_pub/itu-r/oth/0a/06/R0A0600005D0001PDFE.pdf`, 2015, accessed on May 20, 2023.

67. Atat, R., L. Liu, H. Chen, J. Wu, H. Li and Y. Yi, "Enabling Cyber-Physical Communication in 5G Cellular Networks: Challenges, Spatial Spectrum Sensing, and Cyber-Security", *IET Cyber-Physical Systems: Theory & Applications*, Vol. 2, No. 1, pp. 49–54, 2017.

68. Giordani, M., M. Polese, M. Mezzavilla, S. Rangan and M. Zorzi, "Toward 6G Networks: Use Cases and Technologies", *IEEE Communications Magazine*, Vol. 58, No. 3, pp. 55–61, 2020.

69. Islam, K., W. Shen and X. Wang, "Wireless Sensor Network Reliability and Security in Factory Automation: A Survey", *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, Vol. 42, No. 6, pp. 1243–1256, 2012.

70. Queiroz, D. V., M. S. Alencar, R. D. Gomes, I. E. Fonseca and C. Benavente-Peces, "Survey and Systematic Mapping of Industrial Wireless Sensor Networks", *Elsevier Journal of Network and Computer Applications*, Vol. 97, pp. 96–125, 2017.

71. Yu, X. and Y. Xue, "Smart Grids: A Cyber-Physical Systems Perspective", *Proceedings of the IEEE*, Vol. 104, No. 5, pp. 1058–1070, 2016.

72. Kumar, P., Y. Lin, G. Bai, A. Paverd, J. S. Dong and A. Martin, "Smart Grid Metering Networks: A Survey on Security, Privacy and Open Research Issues", *IEEE Communications Surveys & Tutorials*, Vol. 21, No. 3, pp. 2886–2927, 2019.

73. Yan, Y., Y. Qian, H. Sharif and D. Tipper, "A Survey on Smart Grid Communication Infrastructures: Motivations, Requirements and Challenges", *IEEE Communications Surveys & Tutorials*, Vol. 15, No. 1, pp. 5–20, 2013.

74. Movassaghi, S., M. Abolhasan, J. Lipman, D. Smith and A. Jamalipour, "Wireless Body Area Networks: A Survey", *IEEE Communications Surveys & Tutorials*, Vol. 16, No. 3, pp. 1658–1686, 2014.

75. Boban, M., A. Kousaridas, K. Manolakis, J. Eichinger and W. Xu, "Connected Roads of the Future: Use Cases, Requirements, and Design Considerations for Vehicle-to-Everything Communications", *IEEE Vehicular Technology Magazine*, Vol. 13, No. 3, pp. 110–123, 2018.

76. Li, H., Z. Liang, G. K. Kurt, G. Ascheid and G. Dartmann, "Secure Probability Map: Transmission Policy Design for Passive Eavesdroppers in Correlated Channels", *IEEE 82nd Vehicular Technology Conference (VTC Fall)*, pp. 1–5, Boston, Massachusetts, USA, 2015.

77. Kumar S., A. A., K. Ovsthus and L. M. Kristensen, "An Industrial Perspective on Wireless Sensor Networks—A Survey of Requirements, Protocols, and Challenges", *IEEE Communications Surveys & Tutorials*, Vol. 16, No. 3, pp. 1391–1412, 2014.

78. Kaplan, E. D. and C. Hegarty, *Understanding GPS/GNSS: Principles and Applications*, Artech House, London, 2017.

79. Borre, K., D. M. Akos, N. Bertelsen, P. Rinder and S. H. Jensen, *A Software-Defined GPS and Galileo Receiver: A Single-Frequency Approach*, Springer Sci-

ence & Business Media, Boston, Massachusetts, USA, 2007.

80. Sadeghi, M. and M. Gholami, "Time Synchronizing Signal by GPS Satellites", *WSEAS Transactions on Communications*, Vol. 7, No. 5, pp. 521–530, 2008.

81. Blewitt, G., "Basics of the GPS Technique: Observation Equations", *Geodetic applications of GPS*, Vol. 16, pp. 10–54, 1997.

82. Bash, B. A., D. Goeckel, D. Towsley and S. Guha, "Hiding Information in Noise: Fundamental Limits of Covert Wireless Communication", *IEEE Communications Magazine*, Vol. 53, No. 12, pp. 26–31, 2015.

83. Manandhar, D., Y. Suh and R. Shibasaki, "GPS Signal Acquisition and Tracking-An Approach Towards Development of Software-based GPS Receiver", *Technical Report of IEICE*, 2004.

84. Oligeri, G., S. Sciancalepore, O. A. Ibrahim and R. Di Pietro, "Drive Me Not: GPS Spoofing Detection via Cellular Network: (Architectures, Models, and Experiments)", *Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks*, pp. 12–22, Miami, Florida, USA, 2019.

85. Radin, D., P. F. Swaszek, K. C. Seals and R. J. Hartnett, "GNSS Spoof Detection Based on Pseudoranges from Multiple Receivers", *Proceedings of the 2015 International Technical Meeting of the Institute of Navigation*, pp. 657–671, Dalian, China, 2015.

86. Mosavi, M., S. Azarshahi, I. Emamgholipour and A. Abedi, "Least Squares Techniques for GPS Receivers Positioning Filter Using Pseudo-range and Carrier Phase Measurements", *Iranian Journal of Electrical and Electronic Engineering*, Vol. 10, No. 1, pp. 18–26, 2014.

87. Lin, S. and D. J. Costello, *Error Control Coding, Second Edition*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, 2004.

88. Jing, Y. and H. Jafarkhani, "Single and Multiple Relay Selection Schemes and Their Achievable Diversity Orders", *IEEE Transactions on Wireless Communications*, Vol. 8, No. 3, pp. 1414–1423, 2009.

89. Goldsmith, A., *Wireless Communications*, Cambridge University Press, Cambridge, 2005.

90. He, B., X. Zhou and A. L. Swindlehurst, "On Secrecy Metrics for Physical Layer Security over Quasi-static Fading Channels", *IEEE Transactions on Wireless Communications*, Vol. 15, No. 10, pp. 6913–6924, 2016.

91. Salós, D., C. Macabiau, A. Martineau, B. Bonhoure and D. Kubrak, "Nominal GNSS Pseudorange Measurement Model for Vehicular Urban Applications", *IEEE/ION Position, Location and Navigation Symposium*, pp. 806–815, Indian Wells, California, USA, 2010.

92. Milaat, F. A. and H. Liu, "Decentralized Detection of GPS Spoofing in Vehicular Ad Hoc Networks", *IEEE Communications Letters*, Vol. 22, No. 6, pp. 1256–1259, 2018.

93. Bidikar, B., B. P. Chapa, M. V. Kumar and G. S. Rao, "GPS Signal Multipath Error Mitigation Technique", *Satellites Missions and Technologies for Geosciences*, IntechOpen, 2020.

94. Broumandan, A. and G. Lachapelle, "Spoofing Detection Using GNSS/INS/Odometer Coupling for Vehicular Navigation", *Sensors*, Vol. 18, No. 5, p. 1305, 2018.

95. Ashraf, S. A., I. Aktas, E. Eriksson, K. W. Helmersson and J. Ansari, "Ultra-Reliable and Low-Latency Communication for Wireless Factory Automation: From LTE to 5G", *IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, pp. 1–8, Berlin, Germany, 2016.

96. Broumandan, A., A. Jafarnia-Jahromi, S. Daneshmand and G. Lachapelle, "Overview of Spatial Processing Approaches for GNSS Structural Interference Detection and Mitigation", *Proceedings of the IEEE*, Vol. 104, No. 6, pp. 1246–1257, 2016.

97. Cheminod, M., L. Durante and A. Valenzano, "Review of Security Issues in Industrial Networks", *IEEE Transactions on Industrial Informatics*, Vol. 9, No. 1, pp. 277–293, 2013.

98. Dong, L., Z. Han, A. P. Petropulu and H. V. Poor, "Improving Wireless Physical Layer Security via Cooperating Relays", *IEEE Transactions on Signal Processing*, Vol. 58, No. 3, pp. 1875–1888, 2010.

99. Frotzscher, A., U. Wetzker, M. Bauer, M. Rentschler, M. Beyer, S. Elspass and H. Klessig, "Requirements and Current Solutions of Wireless Communication in Industrial Automation", *2014 IEEE International Conference on Communications (ICC) workshops*, pp. 67–72, Sydney, Australia, 2014.

100. Güngör, V. and G. Hancke, *Industrial Wireless Sensor Networks: Applications, Protocols, and Standards*, Industrial Electronics, Taylor & Francis, Boca Raton, 2013.

101. Kumar, S. A., T. Vealey and H. Srivastava, "Security in Internet of Things: Challenges, Solutions and Future Directions", *49th IEEE Hawaii International Conference on System Sciences (HICSS)*, pp. 5772–5781, Koloa, Hawaii, USA, 2016.

102. Lee, J., B. Bagheri and H.-A. Kao, "A Cyber-Physical Systems Architecture for Industry 4.0-based Manufacturing Systems", *Manufacturing Letters*, Vol. 3, pp. 18–23, 2015.

103. Poor, H. V. and R. F. Schaefer, "Wireless Physical Layer Security", *Proceedings*

of the National Academy of Sciences, Vol. 114, No. 1, pp. 19–26, 2017.

104. Sadeghi, A.-R., C. Wachsmann and M. Waidner, "Security and Privacy Challenges in Industrial Internet of Things", *52nd ACM/EDAC/IEEE Design Automation Conference (DAC)*, pp. 1–6, San Francisco, California, USA, 2015.

105. Si, H., O. O. Koyluoglu and S. Vishwanath, "Polar Coding for Fading Channels: Binary and Exponential Channel Cases", *IEEE Transactions on Communications*, Vol. 62, No. 8, pp. 2638–2650, 2014.

106. Varghese, A. and D. Tandur, "Wireless Requirements and Challenges in Industry 4.0", *IEEE International Conference on Contemporary Computing and Informatics (IC3I)*, pp. 634–638, Mysore, India, 2014.

107. Yılmaz, O. N., Y.-P. E. Wang, N. A. Johansson, N. Brahmi, S. A. Ashraf and J. Sachs, "Analysis of Ultra-Reliable and Low-Latency 5G Communication for A Factory Automation Use Case", *IEEE International Conference on Communication Workshop (ICCW)*, pp. 1190–1195, London, United Kingdom, 2015.

108. Zou, Y., J. Zhu, X. Wang and L. Hanzo, "A Survey on Wireless Security: Technical Challenges, Recent Advances, and Future Trends", *Proceedings of the IEEE*, Vol. 104, No. 9, pp. 1727–1765, 2016.

109. United States Department of Defense, "Global Positioning System Standard Positioning Service Signal Specification", `https://https://rosap.ntl.bts.gov/view/dot/16930/`, accessed on May 23, 2008.

110. Nguyen, V. H., G. Falco, M. Nicola and E. Falletti, "A Dual Antenna GNSS Spoofing Detector Based on the Dispersion of Double Difference Measurements", *9th IEEE ESA Workshop on Satellite NavigationTechnologies and European Workshop on GNSS Signals and Signal Processing (NAVITEC)*, pp. 1–8, Noordwijk, Netherlands, 2018.

111. Wei, X. and B. Sikdar, "Impact of GPS Time Spoofing Attacks on Cyber Physical Systems", *IEEE International Conference on Industrial Technology (ICIT)*, pp. 1155–1160, Melbourne, Australia, 2019.

112. Crustal Dynamics Data Information System (CDDIS DAAC), "International GNSS Service", `https://cddis.nasa.gov/Data_and_Derived_Products/GNSS/GNSS_data_and_product_archive.html/`, accessed on May 23, 2023.

113. Pan, L. and F. Guo, "Real-time Tropospheric Delay Retrieval with GPS, GLONASS, Galileo and BDS Data", *Scientific reports*, Vol. 8, No. 1, p. 17067, 2018.

114. Wang, Z., R. Liu, Q. Liu, L. Han and J. S. Thompson, "Feasibility Study of UAV-assisted Anti-jamming Positioning", *IEEE Transactions on Vehicular Technology*, Vol. 70, No. 8, pp. 7718–7733, 2021.