Spring 5-7-2024

Machine Learning for Intrusion Detection into Unmanned Aerial System 6G Networks

Faisal Alrefaei
Embry-Riddle Aeronautical University, alrefaef@my.erau.edu

Machine Learning for Intrusion Detection into Unmanned Aerial System 6G Networks

By
Faisal Alrefaei

A dissertation submitted to the Faculty of
Embry-Riddle Aeronautical University in partial fulfillment
of the requirements for the degree of Doctor of Philosophy in
Electrical Engineering and Computer Science

Embry-Riddle Aeronautical University
Daytona Beach, Florida
May 2024

Machine Learning for Intrusion Detection into Unmanned Aerial System 6G Networks

By
Faisal Alrefaei

This dissertation was prepared under the direction of the candidate's Dissertation Committee Chair, Dr. Radu F. Babiceanu, and has been approved by the members of the dissertation committee. It was submitted to the College of Engineering and accepted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Electrical Engineering and Computer Science.

Radu F. Babiceanu, Ph.D.
Committee Chair

Houbing Song, Ph.D.
Committee Co-Chair

Laxima Niure Kandel, Ph.D.
Committee Member

Eduardo A. Rojas-Nastrucci, Ph.D.
Committee Member

Kenji Yoshigoe, Ph.D.
Committee Member

Massood Towhidnejad, Ph.D.
Chair, Electrical Engineering and Computer Science — Date

James W. Gregory, Ph.D.
Dean, College of Engineering — Date

Norbert J. Zarb, Ph.D.
Vice Provost for Academic Affairs — Date

Abstract

Progress in the development of wireless network technology has played a crucial role in the evolution of societies and provided remarkable services over the past decades. It remotely offers the ability to execute critical missions and effective services that meet the user's needs. This advanced technology integrates cyber and physical layers to form cyber-physical systems (CPS), such as the Unmanned Aerial System (UAS), which consists of an Unanned Aerial Vehicle (UAV), ground network infrastructure, communication link, etc. Furthermore, it plays a crucial role in connecting objects to create and develop the Internet of Things (IoT) technology. Therefore, the emergence of the CPS and IoT technologies provided many connected devices, generating an enormous amount of data. Consequently, the innovation of 6G technology is an urgent issue in the coming years. The 6G network architecture is an integration of the satellite network, aerial networks, terrestrial networks, and marine networks. These integrated network layers will provide new enabling technologies, for example, air interfaces and transmission technology. Therefore, integrating heterogeneous network layers guarantees an expansion strategy in the capacity that leads to low latency, ultra-high throughput, and high data rates. In the 6G network, Unmanned Aerial Vehicles (UAVs) are expected to densely occupy aerial spaces as UAV flying base stations (UAV-FBS) that comprise the aerial network layer to offer ubiquitous connectivity and enhance the terrestrial network in remote areas where it is challenging to deploy traditional infrastructure, for example, mountain, ocean deserts, and forest. Although the aerial network layer offers benefits to facilitate governmental and commercial missions, adversaries exploit network vulner-

abilities to block intercommunication among nodes by jamming attacks and violating integrity through executing spoofing attacks.

This work offers a practical IDS onboard UAV intrusion detection system to detect unintentional interference, intentional interference jamming, and spoofing attacks. Integrating time series data with machine learning models is the main part of the suggested IDF to detect anomalies accurately. This integration will improve the accuracy and effectiveness of the model. The 6G network is expected to handle a high volume of data where non-malicious interference and congestion in the channel are similar to a jamming attack. Therefore, an efficient anomaly detection technique must distinguish behaviors in the drone's wireless network as normal or abnormal behavior. Our suggested model comprises two layers. The first layer has the algorithm to detect the anomaly during transmission. Then it will send the initial decision to the second layer in the model, including two separated algorithms, confirming the initial decision separately (nonintentional interference such as congestion in the channel, intentional interference jamming attack, and classify the type of jamming attack, and the second algorithm confirms spoofing attack. A jamming attack is a stealthy attack that aims to exhaust battery level or block communication to make wireless UAV networks unavailable. Therefore, the UAV forcibly relies on GPS signals. In this case, the adversary triggers a spoofing attack by manipulating the Global Navigation Satellite System (GNSS) signal and sending a fake signal to make UAVs estimate incorrect positions and deviate from their planning path to malicious zones. Hackers can start their malicious action either from malicious UAV nodes or the terrestrial malicious node; therefore, this work will enhance security and pave the way to start thinking about leveraging the benefit of the 6G network to design robust detection techniques for detecting multiple attacks that happen separately or simultaneously.

Table of Contents

List of Figures

List of Tables

1

Introduction

This section provides an overview of the revolutionary technology in the mobile wireless communication network and how various security threats have been presented and evolved in wireless communication over recent years. The motivation of this research is that the traditional intrusion detection techniques have become ineffective in securing transmission in the intelligent heterogeneous network, specifically the Unmanned Aerial Systems (UAS), which includes of an Unmanned Aerial Vehicle (UAV), ground network infrastructure, and communication link. Furthermore, machine learning is expected to be a suitable method to secure the Unmanned Aerial Vehicle (UAV)- Flaying Base Station (FBS) (UAV-FBS) wireless communication system. This work proposes an onboard Intrusion Detection System (IDS) to tackle expected security issues in the 6G aerial network layer UAV-FBS, such as distinguishing between unintentional jamming caused

by data congestion in the channel and the real jamming attack and spoofing attack. Lastly, this section highlights the importance of this research work in securing the 6G aerial network layer, specifically the security of the UAV-FBS.

In the past four decades, wireless technology and mobile wireless communication systems have been widely developed and improved to meet the growing connectivity demands of users. People can easily connect with each other, chatting, and exchanging information instead of traveling or spending some days to deliver messages. This progress played a vital role in the revolution of modern industry and opened up new valuable applications. The evolution of these technologies initially resulted in an advanced mobile phone system (AMPS), which was called at that time 1G [1]. 1G was the beginning era of the mobile wireless system, followed by different cellular wireless generations, namely 2G, 3G, 4G, 5G, and the next generation cellular network, 6G, which is planned to be deployed by 2030. The researchers named cellular wireless generation G based on the differentiation of speed, frequency, and capacity [1]. For example, 1G used to support voice calls as analog technology. 2G used digital technology to add a new feature to support text message sending. After these two generations, 3G emerged to increase capacity and provide a high data transmission rate, adding new features to voice calls and text messages. 4G used to support the wireless mobile Internet to reduce costs and increase quality of service (QoS) and bandwidth. 5G overcame the limitations of 4G and was designed to bring the world a wireless World Wide Wireless Web WWWW. 6G is expected to continue developing the cellular network and fix the limitations of 5G by integrating the 5G with the satellite network for large-scale coverage.

In 1980, 1G networks were introduced as analog technology. It provided voice services based on the Advanced Mobile Phone System (AMPS), as shown in Figure 1.1. The

AMPS supported a speed of up to 2.4kbp and used Frequency Division Multiple Access FDMA with a channel capacity of 30KHz and frequency band 824-894MHZ [1]. The main drawback was that this technology provided limited capacity and coverage for users [2]. In addition, security and privacy issues were significantly present in this generation. The transmission was insecure, and encryption techniques were not applied to secure communication and phone conversations [3]. Therefore, the transmission was threaded by eavesdropping illegal access to expose critical information. In addition, 1G was exposed to jamming and spoofing attacks where jamming interfered with the analog signal and spoofing mimicked the analog signal of a legitimate user.

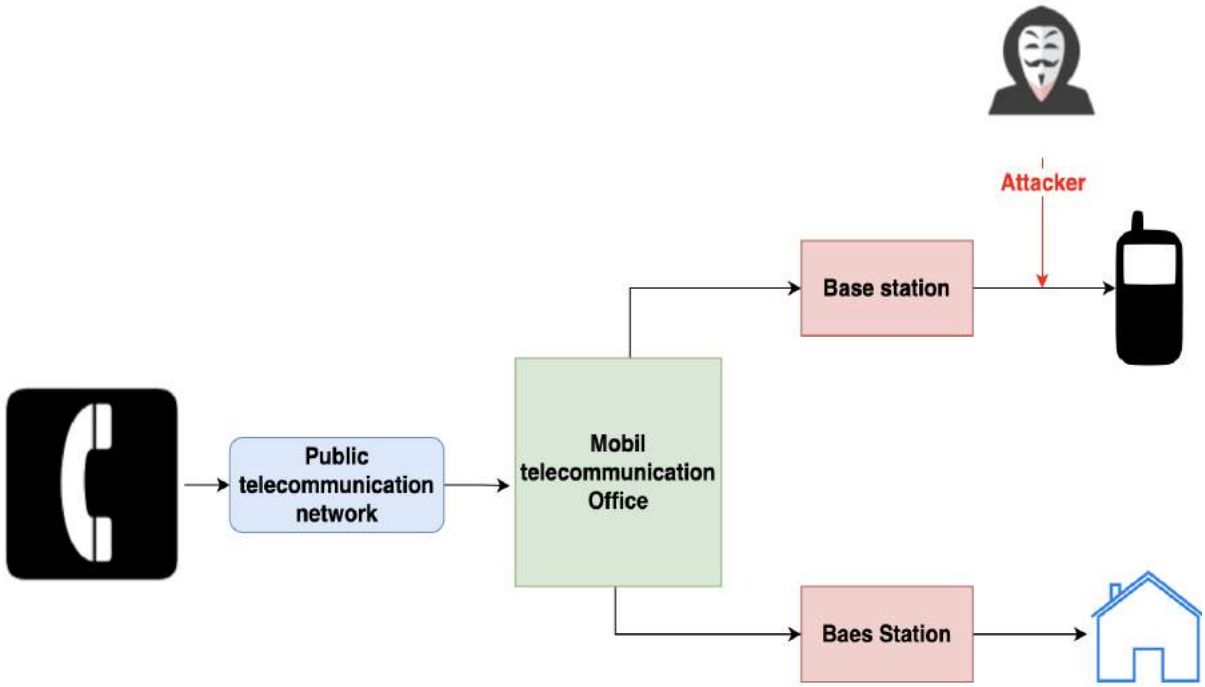

Figure 1.1: 1G network architecture

2G was introduced in 1980 to expand using the Internet. It was presented as digital technology, which used digital modulation techniques, as shown in Figure 1.2. It reached a speed of 64kbps and a bandwidth of 30-200KHz. This technique added new features through supported short messages service SMS, multimedia messages services MMS,

pictures messages, and voice through digital technology [1]. 2G was used to give users a special Code Division Multiple Access (CDMA) for communication. Also, it was used to divide the signal into time slots by the digital modulation schema Time Division Multible Access (TDMA). Some features were improved in this technology, such as increasing the affordability by decreasing the cost and also increasing the coverage [4]. Therefore, developing this technology provided the Global System for Mobile Communication GSM that used to support international roaming. However, security issues were presented in this generation in two forms: first, the authentication were executed in one way that resulted in vulnerabilities in the network. Second, encryption was not an end-to-end technique because of using A5/1[3]. Hence, the 2G was exposed to jamming and spoofing attacks by using an International Mobile Subcarrier Identity (IMSI) catcher.

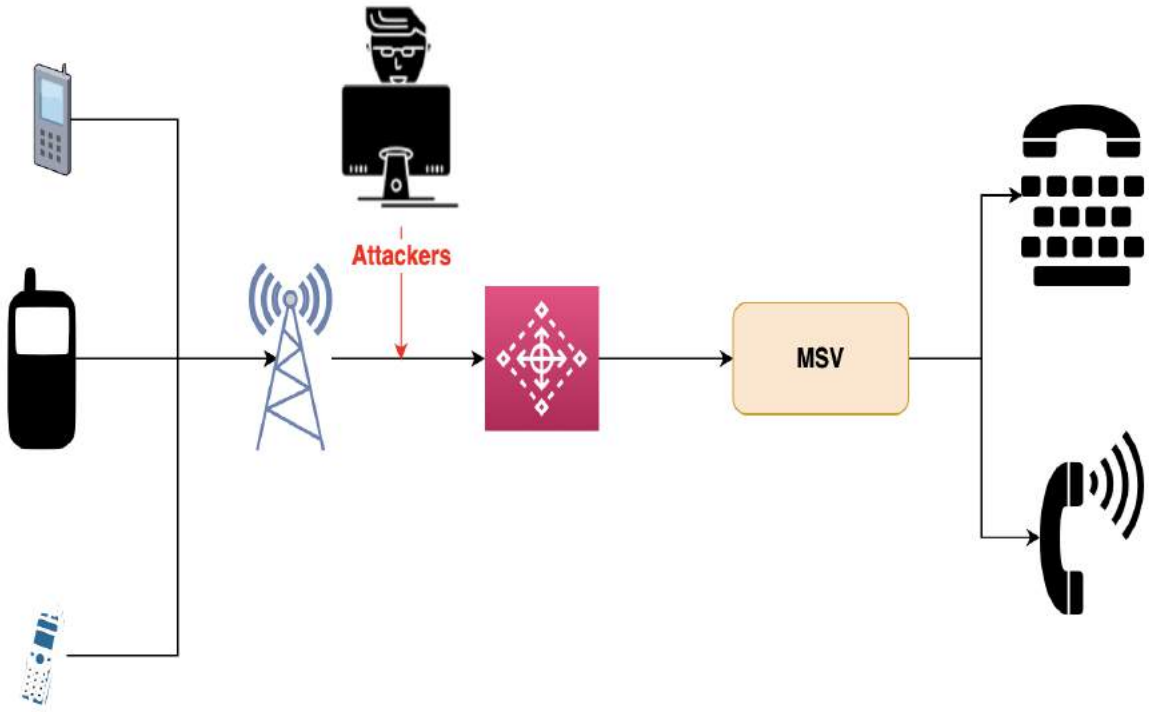

Figure 1.2: 2G network architecture

Year 2000 has seen the emergence of the 3G network. The breakthrough in this technology provided an Internet service on mobile devices. 3G increased the data rate in

the wide range area to 384 kbps, and it used to provide high speed data transmission of up to 2 Mbps in a local coverage area to access the Internet easily [1]. Compared to 1G and 2G, the 3G network added new features to advance this new generation by adding a link to provide the General Packet Radio Service (GPRS), which allowed users to browse the web on mobile devices, TV streaming and video services, and navigation maps and fax as shown in Figure 1.3. Therefore, it rustled in a growing demand for the data rate. However, the limitation of this technology was that the 3G cellular network needed to provide more capacity for communication; therefore, downloading and uploading were slow. In addition, security threats were presented in some form, for example, denied services, illegal access, and violation of transmission integrity [3]. Jamming and spoofing attacks were presented in this technology to exploit vulnerabilities, but were required to have expertise and sophisticated equipment.

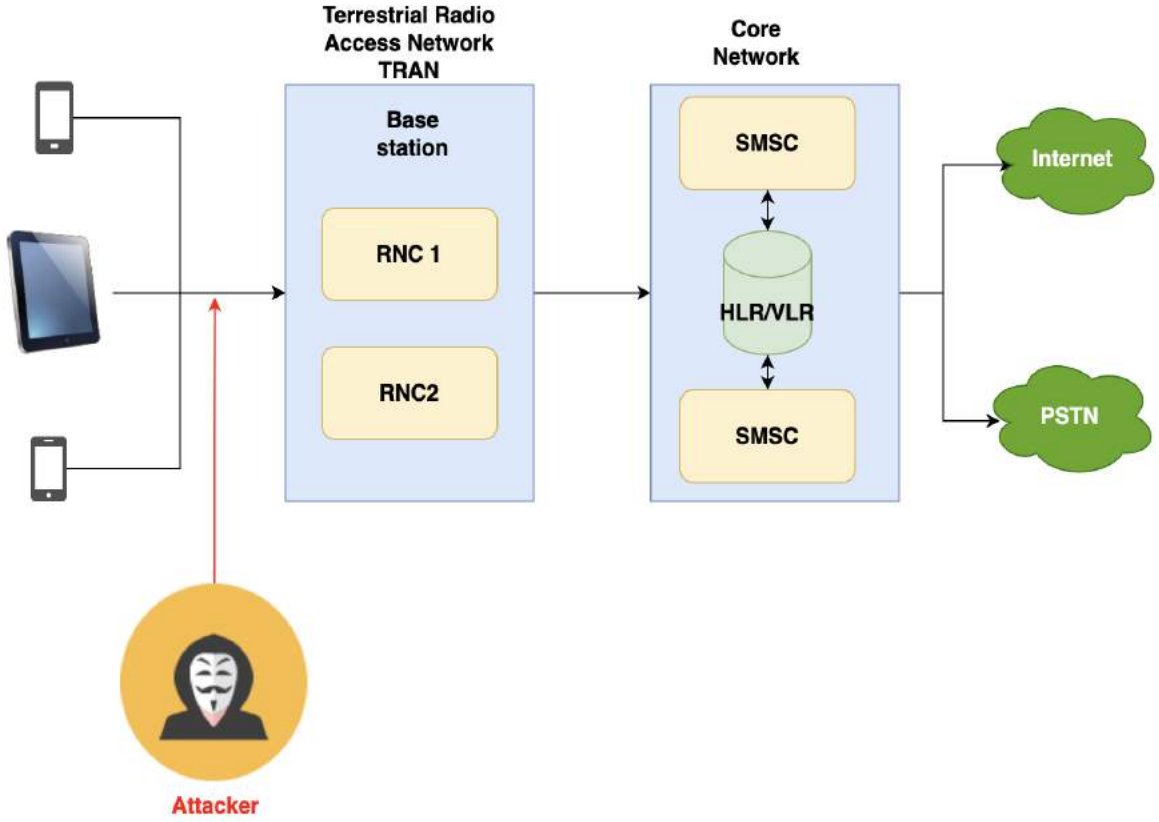

Figure 1.3: 3G network architecture

With continued development in the wireless network architecture, 4G was deployed in 2009. 4G provided different data plans, so multiple mobile devices connected to the network, as shown in Figure 1.4 [5]. 4G introduced long-term Evolution LTE as a standard network that innovated to meet the increase in user internet requests. 4G-LTE provided transmission on the uplink up to 500 Mbits/s and 1 Gbits/s on the downlink [1]. The characteristics presented in this generation were efficient spectrum and low latency, high quality, high speed, and cheap services. Therefore, these characteristics made it possible for advanced technology, for example, multiple inputs and multiple outputs MIMO, Orthogonal Frequency Division Multiplexing (OFDM). [1]. However, multiple limitations presented in 4G LTE that made it impossible to implement it in some critical applications, such as network interruptions, were noticed when it was not guaranteed to provide stable connections during transmission. Furthermore, the average execution time of the handover needed to be increased to transmit all data between a base station [6]. Low latency in end-to-end communication was also presented [7]. Additionally, security concerns and vulnerabilities were presented in this generation, specifically unauthorized access, data integrity, and Denial of Service (DOS) attacks. In addition, a jamming attack was considered to explore the LTE protocol.

As cellular network technology continued to evolve, 5G emerged in 2020 to provide advanced features not offered by previous generations to face the increasing use of devices that required more data. 5G addressed the limitations of the previous generation by increasing the capacity of traffic, the efficiency of the network with high-quality services, and decreasing density [8]. This technology went beyond smartphone devices by supporting IoT. The 5G has the ability to form a complete system at a faster speed than ever, as shown in Figure 1.5. It provides unlimited access and sharing of data whenever the users in the coverage range. 5G is widely used in large-scale IoT applications,

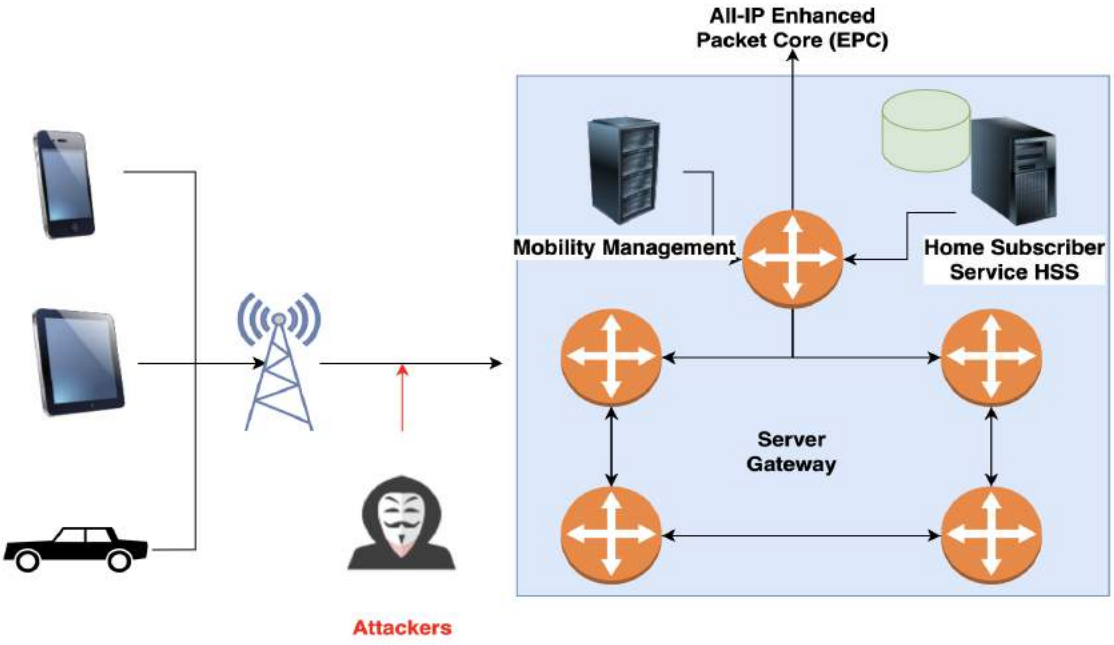

Figure 1.4: 4G network architecture

self-driving vehicles, and UAVs [9]. However, along with the advantages offered by 5G, critical network vulnerability was present in this generation. For example, an enormous number of connected devices is the main reason for DOS resource attacks [1].

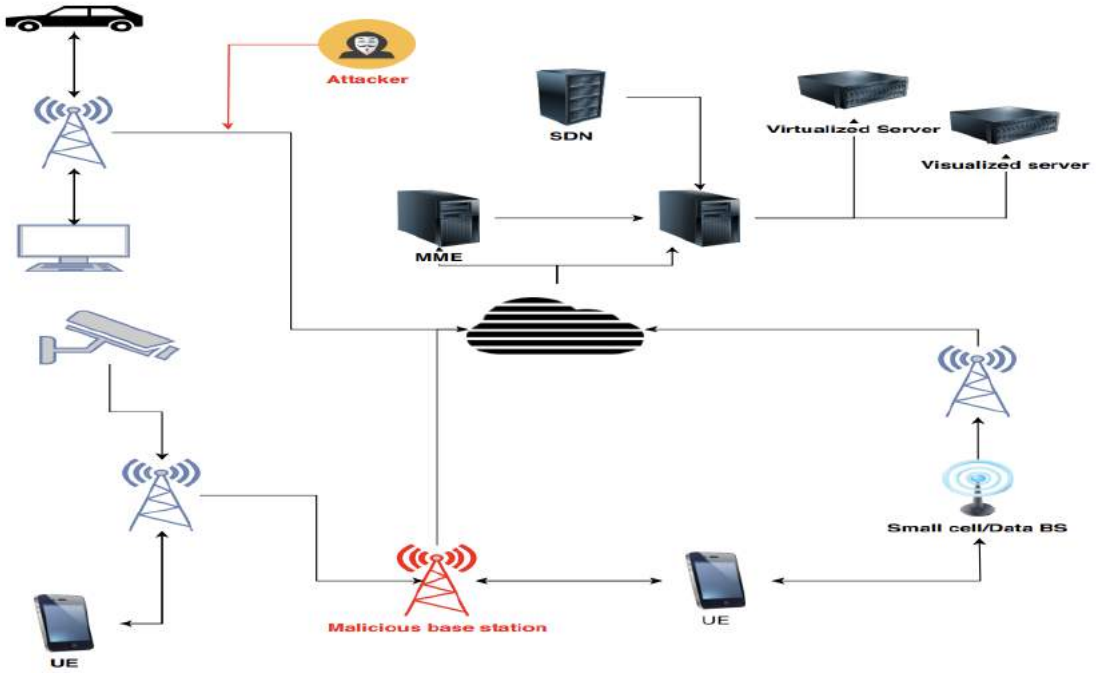

Figure 1.5: 5G network architecture

1.1 Motivation

1.1.1 6G Cellular Communication Network

It is clear that the number of Internet users is growing rapidly, and the massive connectivity of millions of interconnected IOT devices is growing significantly in different places on the planet. Hence, the 5G technology is facing challenges to meet enormously increasing demands in the future. 5G relays on a conventional network infrastructure to provide connectivity that relays on the terrestrial network, deploying a remote antenna or a fixed base station location. Consequently, 5G technology faces obstacles and challenges in providing ubiquitous connectivity to serve the local wireless network such as wireless sensor networks WSN. Therefore, experts and researchers have started to focus on the next generation 6G cellular network as a new era of cellular communication, as shown in Table 1.2. The 6G will be equipped to address and fix the previous limitations of 5G. Therefore, the 6G network architecture is proposed to differ from the previous generation, forming an integration of heterogeneous network layers to extend communication to the space network system, as shown in Figure 1.7. This integration between network layers will provide additional services such as ultra-high speed low latency communication uHSLLC that will provide low latency 1 m and high data speed 1 Tbps, ubiquitous mobile ultra-broadband uMUB to allow 6G to provide wide range coverage 1000 km/h, and ultra-high data density uHDD to provide reliability and data density 10 million / sq km [10]. The 6G wireless network architecture consists of four main layers: space network layer, aerial network layer, terrestrial network layer, and maritime communication [11]. In the aerial layer, UAVs are proposed as a central part to offer ubiquitous connectivity in an area with limited coverage, e.g., mountain, ocean, deserts, and forest. Therefore, UAVs will be deployed densely in the aerial layer to serve

as a UAV-FBS, relay, or access point flying base station UAV. Hence, it will enable low latency access and increase coverage, capacity, and energy efficiency by operating as a cross layer between space and ground.

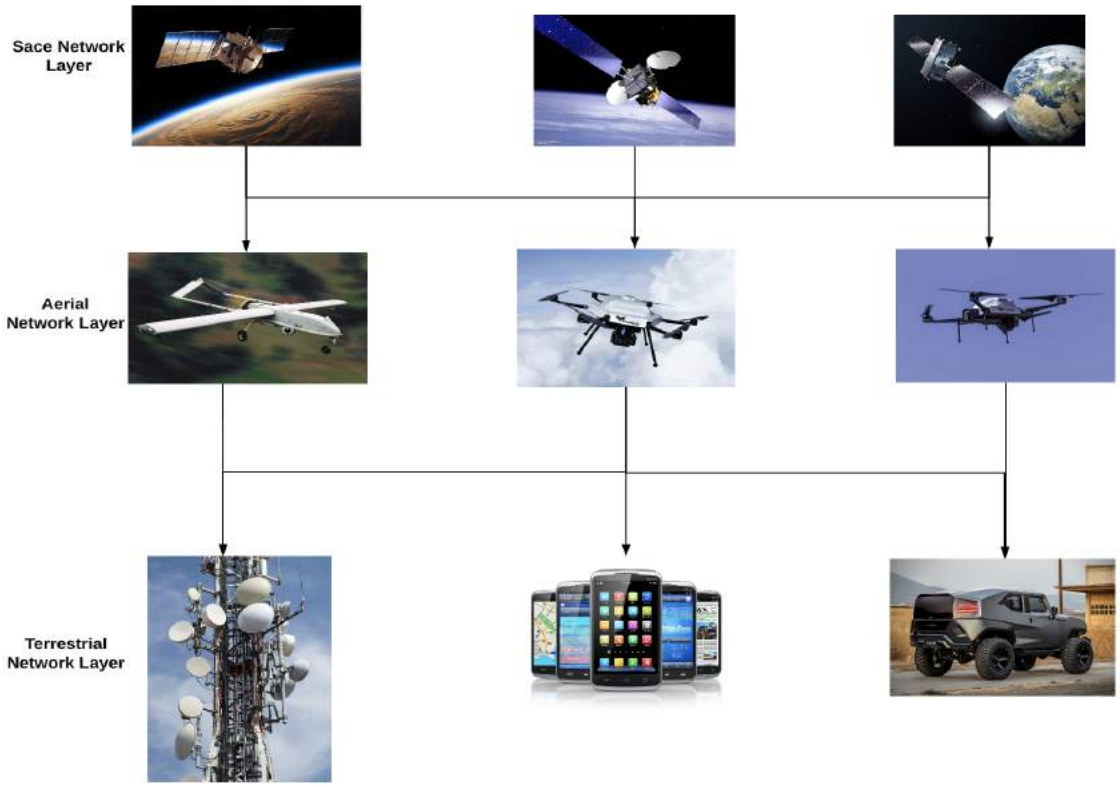

Figure 1.6: UAV deployment scenario

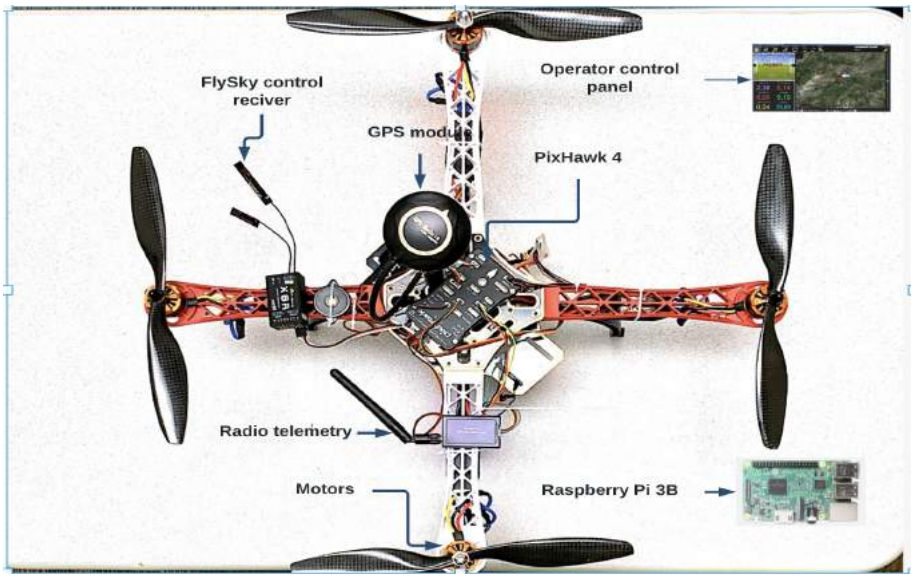

Figure 1.7: UAV archeticture

Year	Network cellular communication genertion	Standard	Core network	frequency band	Mobility distance	Dara rate	Latency
2015	5G	OFDM 5G NR IPv6	IOT	Sub-6 GHz	500 km/h	10 Gbps	5 m
2030	6G	GPS GLONASS COMPASS	IOE	Sub-6 GHz THz band	1000 km/h	1 Tbps	1 ms or less

Table 1.1: Comparison between 5G and 6G

Due to new features and services, the 6G network will provide a high end-to-end reliability level leading to low latency, hence supporting ultra-high mobility. Accordingly, UAVs will be promoted to constitute airborne communication. UAVs, typically known as drones, are considered flying vehicles designed and programmed to fly in predefined aerospace, either remotely by humans or autonomously for specific duties as shown in Figure 1.8 [12], [13]. The UAV is a canonical cyber-physical system (CPS) that integrates the cyber and physical layers by connecting each layer through a wireless network to form an intelligent wireless system that will play a crucial role in multiple applications in the 6G. The 6G will allow the UAV to have various communication simultaneously while the UAV's connection with the terrestrial network is fixed. Hence, linking the satellite with the 6G core network will allow the UAV to locate itself accurately by providing centimeter-level precise, heterogeneous Quality of Service (QOS) provisioning and global coverage [14]. In addition, it will enable autonomous ability at 1000 km/h and through of 1 TBPS for each device [15]. In the aerial layer, the UAV-FBS will function as a service provider to boost connectivity and increase the capacity in the terrestrial IOT networks. Therefore, the UAV-FBS will enhance the terrestrial network and extend connectivity in remote areas where it is challenging to deploy traditional network infrastructure. In this scenario, the UAV-FBS will use physical layer technologies such as cognitive radio, massive MIMO, and mmWave for this purpose [16]. Although all benefits are presented

by using UAV as a cellular-enabled base station, UAV wireless communication networks are susceptible to data congestion in channels and various malicious cyber threats, for example, jamming and spoofing attacks [17].

The natural environment of the spread spectrum is vulnerable to adversary action, violating the availability and integrity of the communication transmission. During the UAV mission, hackers can use shelf hardware that exploits open air transmission to neutralize the UAV and make it useless by launching a jamming or spoofing attack. The jamming attack is one of the dangerous threat attack techniques used against the UAV's network system to disrupt the communication between UAVs and other legitimate entities. It aims to deliberately violate the policies of the media access control protocol (MAC) or the physical layer (PHY) in wireless communication to disrupt data transmissions and degrade the system's performance [18]. It occurs when the hackers send an RF interference as a noise radio signal exploiting the shared natural wireless medium to a subject node in the wireless. It is a version of a DOS attack that leads to catastrophic consequences by hindering ongoing communication and compromising network availability. Therefore, jamming attacks can be classified into four categories: constant, reactive, random, and deceptive. Each type of these attack is executed based on its target behaviors and techniques to perform its malicious action correctly. On the other hand, during the mission, the UAV dependably relies on the navigation and position system, specifically GPS signals in the target zone. These signals are transmitted in the open naturally, so these signals are vulnerable to GPS spoofing attacks. The spoofing attack is an intentionally extreme malicious technique used to target the integrity of the communication channel of the GPS signal to drive the UAV to the hacker's extreme zone. Typically, GPS spoofing attack initiates in two forms, covertly or overtly [19]. In the covert form, the hacker aims to send a high-power signal to mislead the UAV receivers

and receives this fake signal as a legitimate signal. In an overt form, the hacker uses a clever technique by gradually increasing the signal power to make the UAV receiver accept the manipulated signal rather than the authentic one. Therefore, the development and proposal of IDS has become an urgent issue to enhance security transmission in future 6G cell communication.

Network security has gained overwhelming interest after developing communication technology in the recent decades. In 1980, IDS was suggested as a security technique to monitor network traffic, identify abnormal behavior, and meet network security needs [20]. IDS have been used extensively as the primary tool to assess network security and detect suspicious behaviors on networks. However, in the past decades, the continuous development of the internet and wireless communication technology opened up new applications, exploiting the new features such as high data rate, low latency, an increasing number of applications, etc. Therefore, due to the limitations and significant shortage in using traditional IDS to secure 6G application networks, IDS researchers started to improve techniques to work effectively and resist malicious targets [21]. The IDS is divided into three categories: signature-based, specification-based, and anomaly-based [22]. In signature-based detection, the algorithm compares the current network traffic with the attack signature designed [23]. However, these techniques need continual updating in the signature. The specification-based method is designed based on the behavior operation, so it is called the behavioral rule-based, which recognizes malicious behavior based on predefined behavior rules. Once abnormal operations are present, this technique alerts the system administrator of any violation in operation. The limitation of this technique is that there are some factors that a hacker could affect the operation, such as changes in weather. Finally, anomaly-based methods are divided into knowledge-based, statical-based, and machine learning based. In knowledge-based anomaly detection, the

expert designs specific rules to describe how the regular connections are established. In statical-based anomaly detection, it uses the generated stochastic model to compare it with the traffic transmission statistics of normal operation in the network [23]. Machine learning-based anomaly detection uses a training model to monitor objects and identify pattern deviations.

This work focused on anomaly detection, specifically applying machine learning detection techniques. In the previous decades, artificial intelligence emerged as a suitable solution to implement in many applications and enable intelligent services in various technology fields. ML is a kind of artificial intelligence that aims to design an artificial neural network (ANN) model and train the model using specific data to extract and understand the inherent regularity of the information [24]. This process enables ML to make decisions on the input data accurately. The ML is divided into three types: supervised, unsupervised, and reinforcement learning [25]. In the network anomaly detection field, IDS is integrated with ML and has been proven as a suitable technology to ensure transmission and improve detection techniques by detecting the attack and preventing system assessment from being damaged [26]. For example, IDS-based ML is used in critical network systems such as satellites to protect their network through record-specific threats. Therefore, applying ML learning techniques proved that ML is a practical approach to detecting adversary action in UAV networks for multiple reasons: (1) ML can handle unforeseen labels and the dynamic in the UAV network system: (2) the dataset can be easily correlated: (3) ML can replace human interventions: (4) appearance of ML eliminates the need for a mathematical model for the UAV [27].

1.1.2 Methodology

There is no doubt that the benefits of using UAV-FBS in the upcoming years are valuable as it expands the applications to be more efficient, as shown in Figure 1.8. However, some challenges are facing this technology, specifically during data transmission. The UAV in the aerial layer is the primary component, which faces malicious actions and interference that could happen in the channel. Hence, the network security issues in the UAB-FBS in the 6G aerial network layer must be addressed critically.

Figure 1.8: UAV-FBS

This work proposes a novel offline onboard IDS-ML by integrating ML that can be deployed in the UAV-FBS to recognize abnormality during transmission and classify the suspicious behavior, either unintentional interference because of the congestion at the channel or hacker malicious action jamming or spoofing attack. This work develops multiple algorithms and divides them in two layers: the upper layer is training on the unsupervised learning technique that includes autoencoder, the second layer comprises K-means and one class support vector machine OC-SVM. This research focuses on the three main points of this methodology, as depicted in figure 1.9.

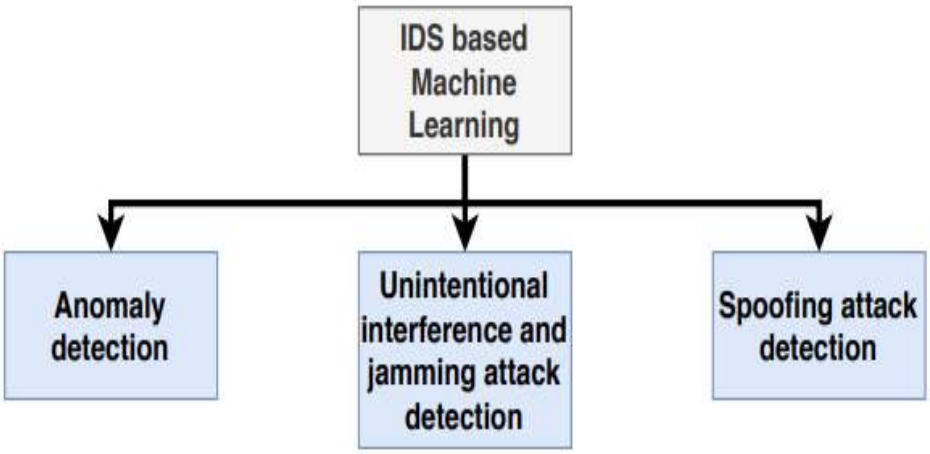

Figure 1.9: Main security issues in this work

1.1.3 Anomaly detection

During the mission of the UAV-FBS, massive data generation is transmitted to the UAV for exchanging or relaying. The UAV-FBS will be equipped to use physical layer techniques, for example, Congitive Radio (CR), Massive Multiple Input Multiple Output (MIMO), and mmWave, which requires a high volume of data. Therefore, channel congestion would frequently occur, causing unintentional interference, in addition to the action of a jamming attack. The literature review shows that providing a detection model to monitor these data patterns and distinguish anomalies was not applied effectively. Previous works have limitations and drawbacks to ensuring accomplished missions securely and without interruption, such as triggering multiple false alarms. Therefore, this work uses some features such as signal strength, signal power, and signal duration to train an effective model in unsupervised techniques, specifically the autoencoder, to recognize anomalies and divide them into three cases: unintentional jamming attack (congestion in the channel), intentional jamming attack, or deviation in the UAV trajectory pattern. Hence, this algorithm in this model provides an initial determination of

the anomaly detected to confirm it by the following two algorithms either as jamming or spoofing attack to increase the assurance in the decision and increase the detection accuracy of the model.

1.1.4 Jamming and data channel congestion discrimination

6G is expected to face a tremendous amount of data that IoT devices will generate. This large number of data at the channel node leads to data congestion and unintentional interference. Therefore, it will be a reason for destruction in the channel, affect packet traffic and Quality of Services (QOS), trigger multiple false alarms, and be slow in decoding and receiving data or blocking the channel. Typically, in the previous work, the authors proposed solutions to distinguish the interference generated by hackers as intentional interference only; hence the previously suggested detection technique would degrade the system performance. In this model, new mechanisms are planned to be added; algorithm trained to distinguish interference into two forms, whether the interference happens due to jamming or congestion in the channel. As an effective algorithm can accurately distinguish between these two issues, the UAV-FBS will be supported, and ensures its performance stable and no interruptions occur during the mission. In this algorithm, multiple features used, such as signal noise to ratio SNR, time domain feature, and signal power, which will be used to distinguish unintentional interference and jamming attacks efficiently. In this scenario, three cases would occur.

- Unintentional interference.
- Jamming attacks occur independently.
- Unintentional interference and jamming attacks happen at the same time.

1.1.5 Spoofing attack detection

The GPS navigation system is a primary component in the UAV system to fly autonomously and securely. While providing service to the IOT terrestrial network, each device needs an entire coverage broadcast to achieve the assigned duty. The efficiency of the UAV-FBS in providing service to end-to-end devices depends on receiving an accurate GPS signal. Therefore, integrating the space network with UAV-FLB to send authentic signals is a requirement. In the previous works, multiple techniques were proposed to defend against spoofing attacks relying on signal characteristics, however these techniques are not valuable in resisting spoofing attacks. Compared to the traditional methods, in this submodel, the solution approach is designed to detect spoofing attacks accurately. Signal characteristics will be incorporated as a feature to define and classify the authentic signal. In this approach, algorithm leveraged the GPS signal features such as distance, speed, and altitude to help determine the fake signal.

2

Contribution and Research Objectives

2.1 Contribution

With the tremendous advancement in cellular technology and the new services planned in the 6G networks, UAV-FBS will become a potential solution to provide ubiquitous connectivity. It will be able to self-organize and have more capability to establish a link for transmission and enhance the terrestrial network when the aerial network layer integrates with the satellite layer. This progress and valuable benefit of the 6G network makes it possible to deploy IoT devices in remote areas for surveillance and monitoring, such as sensing and actuation, since providing wide area coverage was a primary challenge through the use of previous generations of networks in some areas. Most UAV IDS are

conventional; they are deployed on board in UAV or GC and focus on one type of attack. Therefore, they faced challenges in securing UAV missions and will not be effective in the 6G network to secure transmission. The three main goals of this research are:

- Anomaly detection.

  The discovery attack patterns over the network will enhance and support the UAV mission. In previous works, multiple methods focus on one type of suspicion, and these solutions cannot accurately identify the adversary's behavior during the mission if it happens simultaneously. In this research, the abnormal detection algorithm enables discovering the integrity and availability of suspicious violations and classifies them based on attack patterns and behavior.

- Unintentional interference and detection of jamming attacks.

  The use of machine learning in abnormal detection techniques will enhance performance in a different field, making it possible for complex tasks. In this section of the work, K-means is used as a second algorithm to confirm the anomaly detected as unintentional interference or jamming attack. In addition, K-means will be trained to learn four types of jamming attacks, constant, reactive, intermittent, and deceptive jamming attack.

- Detection of spoofing attacks.

  The GPS signal is the primary navigation system on which UAVs rely for their signals to the desired area. In this part, OC-SVM used to confirm GPS spoofing attack model to distinguish and classify the received signal, either authentic or spoofed signals. Detecting the spoofed signal will ensure the performance of the ground nodes achieve their goals. Compared to previous work, this model is an effective technique to detect the deviation of the flight trajectory from the planned

trajectory.

2.1.1 Organization

This work is organized as follows. In Section 3, the security threats directed to UAV networks are presented. Section 4 addresses the literature review and previous related work. Section 5 describes the proposed framework of my work and algorithms developed. Section 6 presents the experimental setup and results analysis of running the proposed algorithms. Section 7 includes the conclusions and future research directions.

3

UAV Network System Security

3.1 Security Threats in UAV Network System

Vulnerabilities and threats make embedded systems prone to security considerations and safety issues. Therefore, this work addresses the expected threats, such as spoofing and jamming attacks, which are present in the layers of the open system interconnection OSI model [28]. The main goal of the OSI model is to explain how data is sent and received in various network devices. Therefore, this model divides the process into seven distinct layers, which describe how the process works when receiving and transmitting data over networks.

- Application layer: In this layer, the user can interact with the UAV system and be served. The Graphical User Interface (GUI) leverages this layer by allowing the user to control the UAV [29]. In addition, it provides communication between the command and controls of C2 center and the satellite network to provide network stability.

- Presentation Layer: In the transmission process, the encoding and formatting of the data is created in this layer to transmit it over the network. Therefore, sensor readings in the UAV is ensured to format and encode the telemetry data generated suitability to be sent over the network to the out application layer [30].

- Data Link Layer: This layer establishes and terminates connections between network nodes. During transmission in the end-to-end connections, collisions are expected when two nodes send data on the same frequency, so this layer is prevalent in this issue. Therefore, the C2 signal will not be received by the UAV. This collision on the network leads to the crash of the vehicle. In addition, this issue will reduce the service where the Media Access Controls (MAC) protocol nodes avoids transmission [31].

- Network Layer: The main objective of the OSI model is to ensure the effective transmission of host to host data. The function of this layer is to route the data packet to the receiver, where this layer encapsulates the Internet protocol address in the packet. Owing to the activities in this layer, several attacks are presented. The hacker can manipulate the nodes to drop and refuse messages. Therefore, the attacked node leads to a sinkhole attack to generate a fake signal [32]. Accordingly, this threat factor creates an unreal node to execute Sybil's attack, which misleads other nodes and degrades the UAV system. In addition, the network layer faces

wormhole attacks, which are used widely against ad-hoc networks. This attack was used to launch malicious activities on other networks.

- Transport Layer: This layer ensures that the transmission data is completed and retransmitted successfully. Therefore, a jamming attack in the OSI model leads to the consumption of network nodes. In this layer, hackers attack the integrity and synchronization of the stored data by executing a desynchronization attack. In addition, UAV GPS is targeted by GPS spoofing attacks to deviate UAVs from the planned path by sending fake GPS signals. Furthermore, a man-in-the-middle attack is executed in this layer where hackers intercept the communication and modify the signal to complete the GPS spoofing attack action [33].

- Physical layer: This combines wireless communication with physical transmission of the raw data. The malicious action of the three threats, jamming and GPS spoofing attack, executes in this layer [34]. In this layer, a GPS spoofing attack occurs when the signal is stronger than the authentic one to mislead the receiver to take the fake signal. In contrast, the jamming attack is a technique used to noise radio signals to degrade the performance of the communication system.

3.1.1 Wireless Communication System Transmission

The wireless communication system in the UAV establishes two communication links: downlink and uplink. In uplink, GC is equipped to generate the commands that the UAV needs for the mission and send them through a wireless transceiver to be received by the UAV. Once the UAV receives the signal, it is equipped with an antenna to capture the transmitted signal. While in the downlink, the UAV captures a constellation signal of the four satellites to process the GPS signal.

On the transmitter side of the UAV, the UAV sends a modulated signal $s(t)$. It uses the carrier frequency (f_c), phase (10), and amplitude, where $m(t)$ is used to represent the baseband message signal.

$$s(t) = A_c \cdot \cos(2\pi f_c t + \phi) \cdot m(t) \tag{3.1}$$

On the antenna gain side, the transmitted power affects UAV antenna gain ($G_{\_t}$)

$$P_t = \frac{{A_c}^2}{2} G_t \tag{3.2}$$

In the channel propagation, free space path loss presents and affects signal strength. Free Space Path Loss (FSPL) is calculated as follows :

$$P_r = \frac{{A_c}^2 G_t G_r \lambda^2}{(4\pi d)^2} \tag{3.3}$$

The $G_{\_r}$ represents the antenna gain in the Ground Control Station (GCS) and expresses the wavelength. In addition, d represents the distance between the GCS and the UAV.

Multipath fading occurs, and the received signal ($r(t)$) is modeled in the same way as the satellite link.

On the receiver side GCS, the GCS receive a modulated signal, so it needs to demodulate to extract the baseband messages signal.

$$\hat{m}(t) = \mathrm{Re}\{r(t) \cdot e^{-j2\pi f_c t}\} \tag{3.4}$$

In addition to evaluating signal quality, Signal to Noise Ration (SNR) is calculated, in addition to error correction techniques, to get reliable data recovery.

In the other line between satellite communication and UAV, additional factors are considered, such as path loss and different types of fading:

$$P_r = \frac{{A_c}^2 G_t G_r \lambda^2}{(4\pi d_{\text{UAV-Satellite}})^2} \tag{3.5}$$

Once the satellite regains the signal, demodulation is performed to extract the baseband message signal. It is expressed as :

$$r_{\text{UAV-Satellite}}(t) = \sum_{i=1}^{N} h_i \cdot s(t - \tau_i) + n(t) \tag{3.6}$$

Here, the a complex channel is represented, and the delay is expressed by additive white Gaussian noise.

To perform the coherent demodulation, received signals are multiplied by the response signal :

$$\hat{m}_{\text{UAV-Satellite}}(t) = \text{Re}\{r_{\text{UAV-Satellite}}(t) \cdot e^{-j2\pi f_c t}\} \tag{3.7}$$

By using this approach, the carrier frequency is restored, and the baseband message signal is recovered.

In the satellite link signal noise ratio, SNR is expressed as :

$$\text{SNR}_{\text{UAV-Satellite}} = \frac{\text{Signal Power}}{\text{Noise Power}} \tag{3.8}$$

where the data rate R is affected by three factors such as bandwidth, modulation, and SNR.

3.1.2 Formulation of Jamming Attack Problems

A jamming attack is a form of DOS attack that is used to emit unwanted signals to interfere with transmission between the sender and the receiver. It is also used to block the receiver from receiving valid signals. Therefore, the receiver fails to decode the signals when the interference rate in the received signals is high [35]. The hackers have two ways to execute this kind of attack by either corrupting the medium access control (MAC) protocol or sending signals to degrade the MAC protocol [36]. The main result of this attack is to send signals to the receiver or sender in terms of blocking them from sending and/or receiving valid signals. The hackers attempt to block the links between the sensor nodes or the link between sensors to control in order to block the measurement from being transmitted to the controller and then to operate services. The jamming attacks can be divided into four main categories: constant jammer, deceptive jammer, random jammer, and reactive jammer. Each one of these attacks has its own technique and procedure to damage the connection. Therefore, the jamming attack to be effective it must affect Signal Ratio (SR) [37]

$$JSR = \frac{J}{S} = \frac{P_j G_j G^*_{com} R^2_{com}}{P_S G_S G_{com} R^2_J} \tag{1}$$

where P is the transmitter and jammer power and G is the gainer of the receiver and jammer. R represents the distance between jammer and transmitter to the UAV. G_{com} represent UAV gainer and G^R/J.

By considering three use cases, Tg, Rg and J represent transmitter,receiver and jammer respectively. The position of Tg (x_T,y_T,z_T), (x_R,y_R,z_R) represent receiver's location and (x_J,y_J,z_J) is the jammer location. Typically, the Received Signal Strength RSS is impacted by noise signal [38] and can be derived as

$$RSS = P_{\mathrm{TR}} + P_{\mathrm{JR}} + P_{\mathrm{N}}. \tag{3}$$

Where P_T is the transmit power and the power strength represent by P_T_R. The negative of path loss exponent is represented by $-\alpha$.

3.1.3 Constant jammer

A constant jammer is a jammer that continuously sends radio signals to the receiver or the legitimate transmitter for different purposes, for example, degrading signals or making congestion on the link connection, as shown in Figure 3.1. There are several types of attacks the hackers can do, but there are two of them more common. One attempt the hackers use is an arbitrary signal to degrade the original signal quality to make the signal at the receiver's channel undecoded. Another attempt is that hackers can transmit numerous signals on the network channel to make it busy [39]. These numerous signal transmissions cause a delay in fresh signals, where the legitimate transmitter cannot transmit the gain receiver channel as a result of this congestion on the network channel. Thus, this technique is the main reason for degrading the system performance and blocking the link between the sender and the receiver.

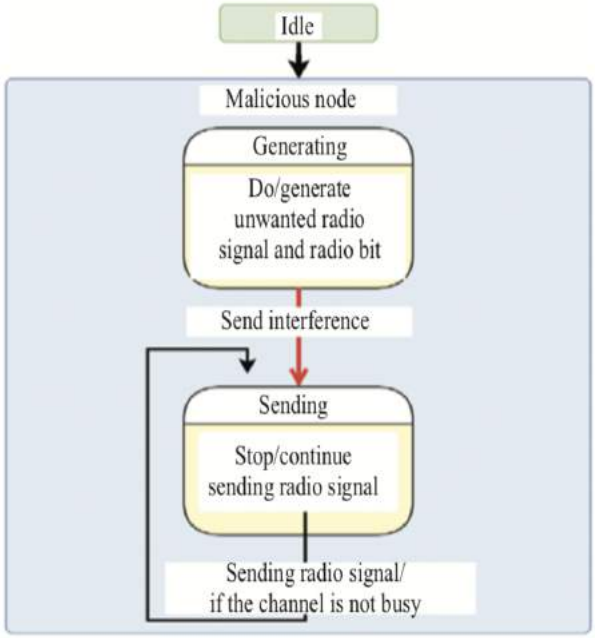

Figure 3.1: Constant jammer model

Intermittent jammer

This is a jamming that sends signals randomly from time to time, as shown in Figure 3.2. This kind of jamming attack preserves its power by consuming rather than continue emitting signals. It is used to send malicious signals for the sake of destructive connection or transmit signals with malicious intent. It is similar to the constant jammer, but its approach is based on saving its energy when it is moving between active or sleeping mode [40].

3.1.4 Deceptive jammer

The deceptive jammer is the most innovative kind of jammer attack. It leads to the data packet being injected into the transmission while the sender sends data packets. Once the gap is presented in the transmission, the deceptive attack can exploit that gap to inject a valid data packet with a useless payload, as shown in Figure 3.3[41].

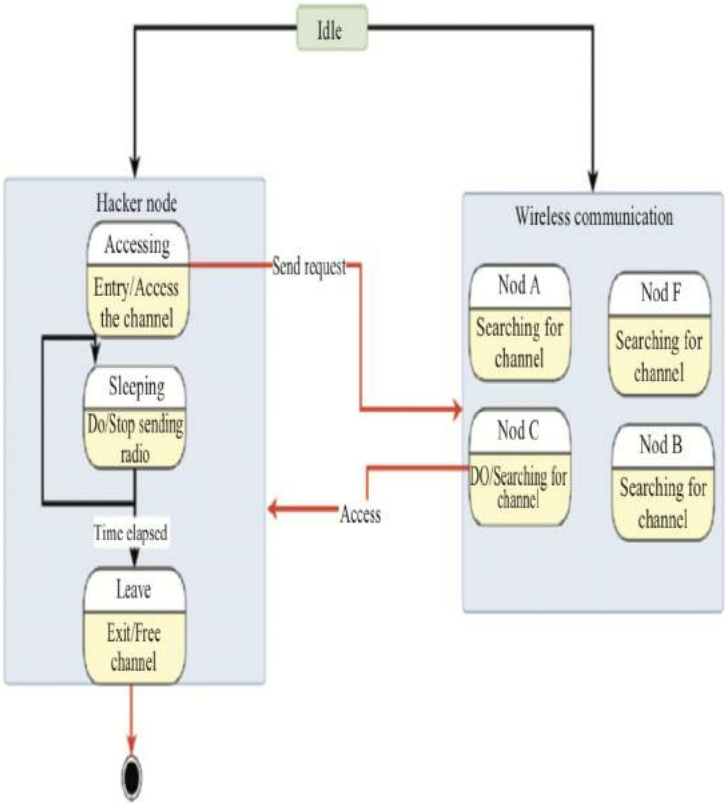

Figure 3.2: Intermittent jammer model

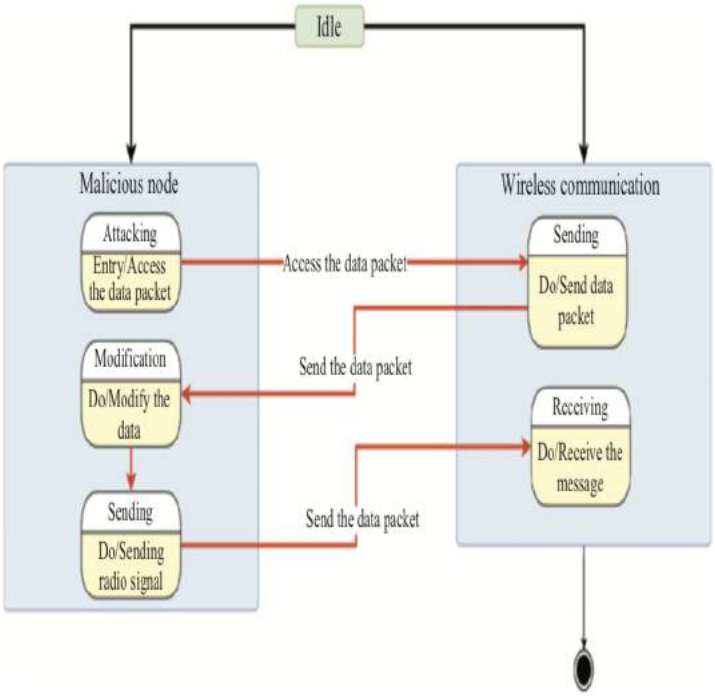

Figure 3.3: Deceptive jammer model

3.1.5 Reactive jammer

A reactive jammer is an attack that is in idle mode until the legitimate node is activated, and then it starts sending a data packet, as shown in Figure 3.4. Additionally, it senses whether the wireless channel is busy to start emitting its malicious signals to degrade the transmitting data at the receiver [42]. The reactive jammer has the ability to distinguish whether the active legitimate node signal is weak or strong in terms of sending signals.

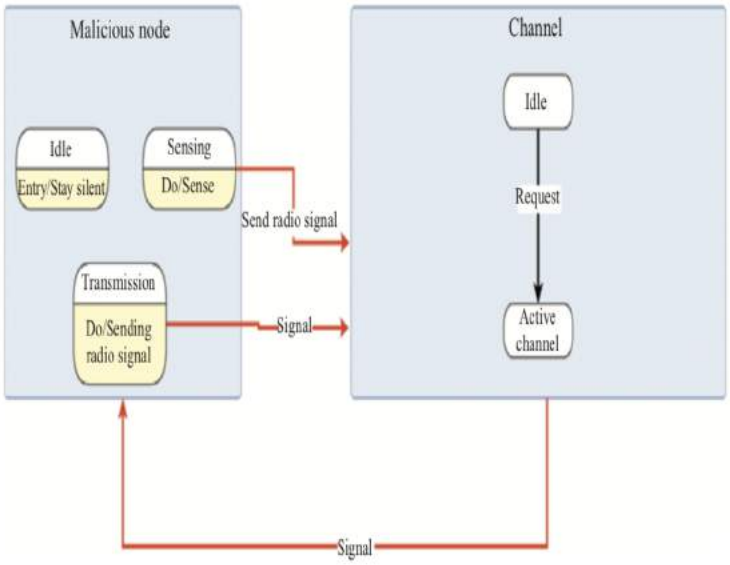

Figure 3.4: Reactive jammer model

3.2 The Global Navigation Satellite System GNSS

The GNSS is a satellite system used widely for multiple purposes such as positioning, navigation, and timing PNT. Various countries use this system for their missions based on coverage and capabilities. The USA uses GPS, Russia uses GLONASS, the European space agency uses Galileo, while China uses the Biedou Navigation satellite system BDS. These four systems operate by different modulation schemes and carrier frequencies [43].

For example, frequency division multiple access FDAM modulation schema is used for the GLONASS signal. On the other hand, GPS uses the code division multiple access CDMA. These characteristics make these two systems different in architecture and the way they transmit signals. However, these systems use standard design and operation to achieve their goals. They use time stamps to transmit radio frequency signals, which they receive to decode these signals to calculate their locations and time. In addition, they can use four satellites in the constellation by time synchronization, 3-dimensional location, and navigation data [44]. All these GNSS systems provide unencrypted signals to the public users. Therefore, it exposes them to hackers who execute a GPS spoofing attack which is addressed later in the GPS spoofing attack section.

3.2.1 Global Position System

GPS was launched in 1979 and was called NAVSTAR for US military purposes. Later, in 1994, the GPS became available for public users for global coverage and became a central component of GNSS. It is deployed in Medium Earth Orbit (MEO) in the 24 satellites orbiting the earth from 22,200 km [45]. They use six equally spaced orbital planes at an inclination of 55 degrees. The GPS architecture is divided into three main components namely user segment, control segment, and space segment:

- User segment: it represents the user receivers and services provided for military or civil missions. These receivers can receive signals and decode them for their position and time. These position and time estimation calculations are performed on the receiver sides, where the devices are equipped with L-band receivers.
- Control segment: it is used to preserve the integrity of the GPS by monitoring the commands and controls. It is formed by users through a global network of ground

facilities to collect valuable information such as telemetry.

- Space segment: it is a constellation of global satellite networks that uses radio frequency RF to send signals, including navigation data and coded information.

GPS Transmission

The GPS uses the L-band frequency f_0 of 10.23 MHz to generate a signal using an onboard atomic clock. L-band is divided into two frequencies, L1 at 1575.42 MHZ and L2 at 1227.60 MHz, obtined by multipling 154 and 120 to generate these two carrier frequencies [46]. The Coarse Acquisition Code C/A and Precise P are used to modulate the signals in the spectrum signals 2.046 MHz and 20.046 MHz bandwidth. The signal generated from each satellite has a unique code called Pseudo-Random Noise PRN, which is a reference for each one. These will improve the SNR, identify each satellite in the GPS constellation, reduce signal interference, and ensure accurate ranging [47].

GPS Receiver

The user segment receives the signal through an antenna equipped with receiver devices. Once the signals are received at the front end, multiple procedures are performed for filtred, digitized, and amplified signals to get baseband signals [48]. After that, the signal processing calculates the navigation information to extract pseudo-range. Also, the rate of pseudo-range determines the difference in the information to estimate Position, Velocity and Timing (PVT) [49]. Multiple stages are presented in this process, such as tracking, extraction, acquisition, and monitoring.

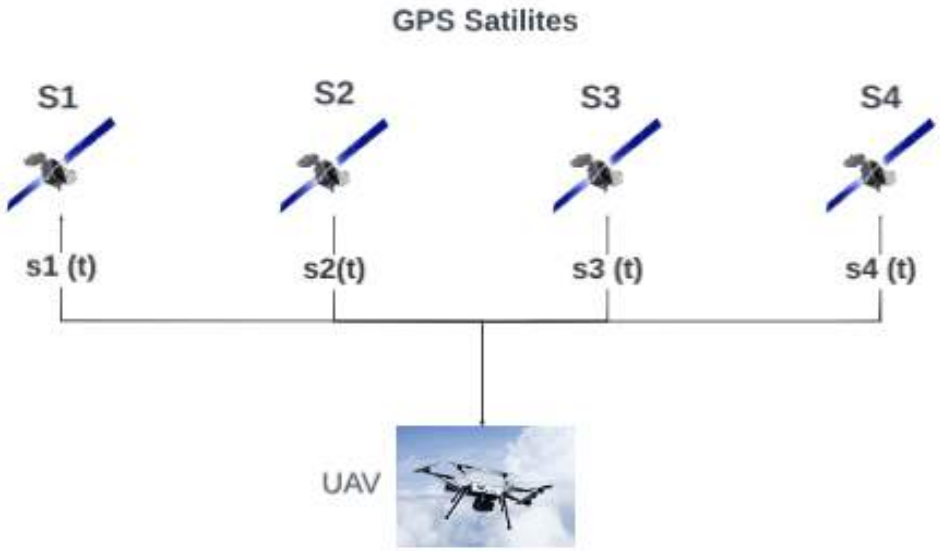

Figure 3.5: UAV GPS receiver observing the generated signal by satilites

3.2.2 Spoofing Attack on GPS

GPS is the main component in driving the UAV in the planned direction. The satellites are located in at $\{P_n^g\}_{n=1}^G$. The satilite posititve vector represent in $p^g = [x^g, y^g]^T$. UAVs depend entirely on GPS data to fly, but the link between GPS and UAVs is vulnerable to adversaries' goals. The spoofing of this link threatens the UAV's civilian or military tasks. For example, the GPS spoofing The attack can be executed in three ways: fake GPS signals, sending signals with higher frequency, or spoofing the high gain antenna [50]. In addition, GPS spoofing can be a form of eavesdropping that listens to the transmission of data between UAVs and GPS signals in space [51].

- Simple spoofing sttack. The hacker does not know the UAV position and sends a fake signal with high-level power, which is unsynchronized with the real signal. This attack leads to significant pseudo-range measurements. Typically, it executes with low-cost hardware and software.

- Intermediate spoofing sttack. In this scenario, the hacker knows the UAV position, leading to code phase alignment between spoofed or benign signals. These attacks are generated simultaneously on the channels. During the mission, the hacker considered the detection system, which relies on the signal characteristics, so it is complex to detect.

- Spoofing with antenna attack. This is a sophisticated technique used against multiple antenna receivers to disrupt the frequency of the other signals. It leads to the gaining of control over the UAV system.

4

Literature Survey

4.1 Review of Current Literature Survey

The main goal of the UAV-FBS is to offer communication infrastructure to exchange sensitive data between nodes during missions and to provide efficient comprehensive coverage whenever they need access to the Internet. However, the UAV-FBS network architecture is vulnerable to malicious threats and anomalies due to open-air radio space and design constraints in UAVs, such as communication capability and computations. This research reviewed the previous works used to detect suspicious abnormalities in the UAV network. The algorithms and models are then used to detect and defend against adversaries in the UAV network. The defense technique is divided into active

and passive defense in Fly Ad-Hoc Network (FANET) [52]. Active protection techniques prevent UAV networks from external attacks before malicious action happens by using encryption. The passive detection technique refers to detecting an anomaly when it occurs on the network to reduce the adversarial impact. This work focuses on the second technique, which is the passive defense.

4.1.1 Intrusion Detection System

IDS is a security schema, which is considered a passive defense and has used broadly to increase accuracy by monitoring network traffic and analyzing object behavior. The IDS's primary goal is to detect illegal activities and abnormal behavior on the networks accurately. Over the past years, traditional IDS techniques have no longer be effective in complex network systems and need improvements [53]. The IDS are categorized into signature-based, specification-based, and anomaly detection as shown in figure 4.1. The signature-based detection's main drawback is that it needs to continue signature updating. In contrast, specification-based detection methods trigger many false alarms and provide low detection accuracy. Machine learning techniques are also applied as IDS, but it faces challenges and needs improvements, such as the author needs enough data. In this section, this work reviewed previous studies used broadly for attack detection on UAV networks and addressed the main drawbacks presented. This work focuses intensely on anomaly-based model learning detection.

4.1.2 Signature-based-IDS

The signature-based technique has extensively used the signature pattern to recognize abnormal behavior and malicious action that happens to the traffic in the wireless net-

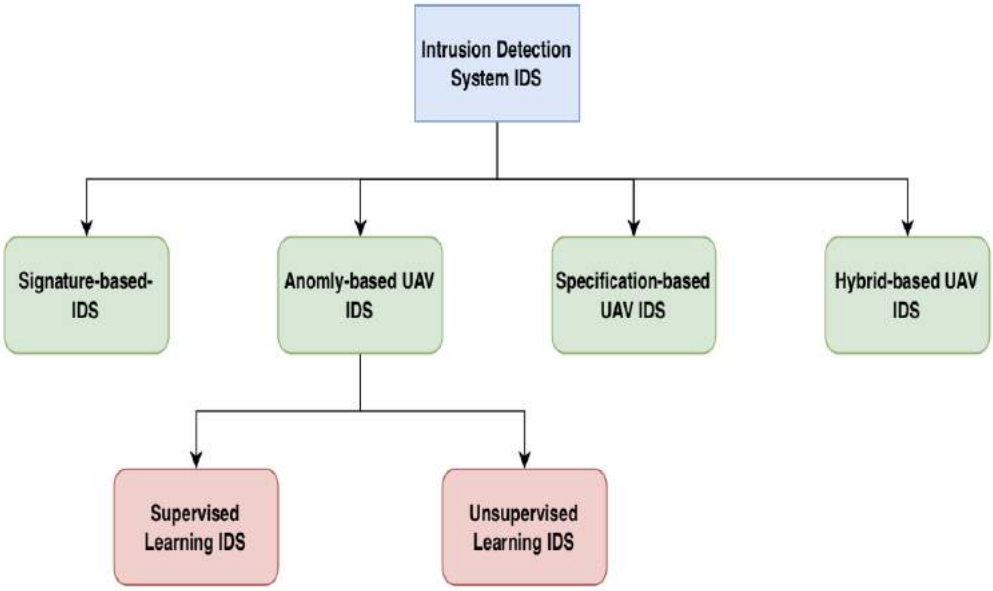

Figure 4.1: The architecture of distributed intrusion detection system

Signature based IDS	Advantages	Disadvantages
	Simple to design	Hard to detect unknown attack
	High detection accuracy	Difficult to keep signature up to date
	Low false alarm rate	High false alarm rate for unknown attack
	Low computation cost	Signature attacks need to update continually

Table 4.1: Advantages and disadvantages of signature-based UAV IDS

works. In this technique, the security engineer identifies a specific signature for each type of attack to be recognized when the hacker launches the attack. They extracted malicious behavior parameters of the attack when it happens in the network during the abnormal behavior [54]. However, as shown in Table 4.1 these techniques faced challenges, such as signature-based detection cannot recognize unknown attacks in which its signature is not predefined and needs to keep the signature up to data [54].

4.1.3 Specification-based UAV IDS

Specification-based intrusion detection methods use the constraints and specifications to recognize attacks on the UAV network system. Pre-flying the UAV, users typically define the planned path that the UAV must follow toward the target direction during the mission by identifying starting and ending points. During flight, the UAV is monitored, so any deviation in the planned path is considered out of the course, and the detection technique triggers alarms immediately. For example, the GPS spoofing attack deviates the UAV from its planned routes, so this technique can detect it easily. However, as shown in table 4.2 the limitations of this technique were identified in previous works. Multiple factors would affect UAV operation and impact its behavior during missions, such as weather, generating false alarms frequently. In addition, it is not sustained in the UAV application where the configuration is updated.

In some early studies, reference [55] used the specifications of the normal mode of the UAV during flying to detect malicious onboard systems. The authors identified some features which were used to recognize abnormal behavior in the networks. They extracted specific rules based on the attack behavior, such as randomness, recklessness, and opportunistic characteristics, to analyze and detect attacks. Based on the model results, their method resulted in high detection accuracy and a low false positive rate of 0.05, 6.0 opportunistic, and 7.0 random attacks. In another work, author in [56] proposed a technique deployed on the UAV and the GCS. Their method was not only offered as a detection technique but was also used to respond to the attack action. The authors addressed five categories of lethal attack used to degrade UAV network systems, such as GPS spoofing, jamming, gray hole attack, black hole attack, and false data injection attack. Hence, to recognize the attack, the paper used features that indicated

Specification based IDS	Advantages	Disadvantages
	Detection of unknown attack	Needs high processing overhead
	Addresses overall behavior of the UAV behavior	Normal behavior is challenging
	Low false alarm rate	High false alarm rate for unknown attack
	Can be integrated with signature based IDS to detect attacks from authorized users	Signature attacks need to be updated

Table 4.2: Advantages and disadvantages of specification based IDS

the abnormality in the networks, such as the number of packets sent (NPS) and signal strength intensity (SSI). They analyzed how these attacks target the number of packets sent or signal strength intensity. These four attacks were evaluated and investigated, and the model detection showed low false positives and had a high detection accuracy.

4.1.4 Anomly-based UAV IDS

Anomaly detection is an IDS approach widely used to monitor observations that deviate from normal behavior or detect rare events to ensure safety and security in UAV operations. It uses specific parameters related to the object to detect abnormality and deviation during the mission. Recently, the anomaly detection technique has become a research hot spot. It is divided into two methods used to detect abnormal behavior during UAV communication: model-based approach (supervised learning based IDS), data-driven strategy (unsupervised learning-based IDS)

4.1.5 Model-based learning IDS

In recent decades, researchers have started thinking of new techniques to solve the drawbacks of the previous methods as shown in table 4.3. Model learning based has attracted

researchers' attention to overcome the issues presented in the traditional detection techniques. With the rapid developments in artificial intelligence, deep learning algorithms, and machine learning algorithms, researcher have investigated the effectiveness of these techniques. It was discovered that these techniques are an effective way to enhance security and solve detection problems. Artificial intelligence methods are divided into sub-branch machine learning ML and deep learning DL. They have been used widely as detection techniques to detect and recognize malicious behavior over networks.

Supervised Learning IDS

The supervised learning algorithm is trained on the predefined data set to achieve a specific output and precision. The predefined data set consists of two tuple labels and attribute [57]. The label represents the output, while the attribute represents the input. The machine learning algorithm discovers suspicious behavior by learning complex patterns in the data set. Therefore, it performs better when applied to detect malicious action over the network. It uses network traffic to extract valuable information based on feature engineering. This algorithm predicts labels corresponding to the trained model on the data set. Therefore, it only detects the trained label and neglects false positives.

Reference [58] proposed an intrusion detection schema designed based on five machine learning models to detect an advanced external cyber attack against the UAV network system in real time. The paper trained multiple machine learning models such as Naive Bayes, support vector machine, decision tree, k-nearest neighbors, and deep learning multi-layer perceptron. After that, the work tested these models to evaluate their performance by detecting sophisticated malicious activity in the drone network. The author designed a testbed environment that includes nodes and a virtual machine.

They used Kali Linux to launch attacks, such as DOS and DDOS. In addition, they simulated TCP and UDP to mimic real UAV networks. The author leveraged the Onion platform and Argus tool to collect the needed packet and label two scenarios as normal and under attack events. The result of these models showed that the detection accuracy was high and the false alarm was low. With continued development in the IDS, it is clear that the system avoiding collision is not sufficient these days when the UAV is deployed as an autonomous system. The hackers can manipulate the data transmission to crash the drone while the avoid collision system can not protect this accident from happening. The author in [50] showed that the connection among UAV components could be intermittent intentionally, so they proposed two algorithms to defend against jamming and spoofing attacks. First, they suggested a model, such as self-taught learning STL, to extract features by knowing the network parameters. Then, they propose a support vector machine (SVM) to classify the attack as jamming or spoofing. This proposed IDS showed that it was efficient and reached height detection accuracy. Reference [59] proposed a combined methodology consisting of a machine learning technique and a multi-agent system to detect DOS cyber attacks that target UAV communication systems. The proposed model was able to detect known and unknown DOS attacks. The input goes through multiple components, starting from gaining the packet by an antenna in the sniffer agent to the final decision made by unsupervised machine learning. The known attack was recognized in the match checker agent if the attack parameter matched the signature database; otherwise, unsupervised machine learning would address an unknown attack to form a signature and store it in the database. The author of this technique, used unsupervised learning results to update the knowledge base signature to increase detection accuracy in the knowledge base module. The author [60] proposed a technique to detect and classify the jamming attack against a link between

the UAV and GCS. The author implemented attack action against the control link IEEE 802.11 orthogonal frequency division multiplexing OFDM at 2.4 GHz. They used GNU-Radio to trigger four types of attacks on the link ,such as signal tone, barrage, protocol awareness, and successive pulse and they used features, for example OFDM parameter, energy parameter, and signal-to-noise ratio. Three features extracted from OFDM are cyclic prefix length, subcarrier length, and subcarrier spacing. Two features extracted from energy are the average received power and threshold. Lastly, they extracted three features from the SNR estimator: average signal power, signal-to-noise ratio SNR, and average noise power. This work showed that the detection rate reached 93, and the false alarm rate was 1.1. The author in [61] proposed combining two strategies to identify the jamming attack. The first strategy used statistics to decompose the signal block when the receivers receive the signals. The second strategy used was a deep neural network. The paper integrated the statistical model since it did not require heavy computation and used a deep neural network to achieve high Accuracy in classifying attacks. Their methods showed that the statistical methods identified accuracy 84.38 when the attack happened in a range of 30 m close to the UAV, while deep network accuracy was 99 when the jamming distance was close to 200 m. Another study [18] proposed a system consisting of two different techniques, decision trees and multi-layer perception. The paper's evaluation relied on simulated and actual data set types. The paper addressed the reactive jammer and focused on detecting this sophisticated attack. During the experiments, the author used features such as signal strength indicator RSSI, packet delivery ratio PDR, and throughput as a predefined matrix to train the model and recognize the jamming attack. The paper shows that both models were evaluated, and based on the results, the detection accuracy of the MLP was superior to the decision trees.

The authors in [62] clarified how the hypothesis test works and why it is unsuitable

for detecting GPS spoofing attacks. In the hypothesis test, setting a threshold for the path losses to detect spoofing attacks is affected by various factors such as cloud, vapor, and temperature. Therefore, the hypothesis threshold is facing challenges. The changing of the environments influences the path loss, increasing the error and decreasing the detection accuracy. Secondly, a false alarm is expected if the threshold value is not determined appropriately. Therefore, the author integrated deep learning algorithms with statistical methods to propose effective GPS spoofing attack detection. The author in [63] proposed an IDS of multiple techniques. It integrates a decision tree, random forest, naive Bayesian linear regression, and support vector machine to detect actuator GPS spoofing attacks. In this work, the author used k-flood to increase the Accuracy before the implementation. They leveraged signal features such as frequency modulation, jitter RAP, jitter local, jitter PPQ5,shimmer APQ3, shimmer, shimmer local, and shimmer dB.

Reference [64] proposed a new combination technique by applying data-driven methods and digital twin architecture. The digital twin architecture represents the real system. Thus, the authors digitized the model and trained it to use UAV flight data as a reference to design the detection model. In addition, multiple algorithms were proposed for analysis and evaluation, such as one-class support vector machine OC-SVM, isolation forest IF, local outlier factor, and deep neural network DNN. This work as in [65] did not target specific attacks in the detection techniques. They used different types of machine learning by monitoring the network traffic to detect an anomaly: minimum packet sizes, the maximum number of packets, and flow duration as features. The author confirmed that decision tree algorithms are the best technique to detect anomalies on the UAV network. The other algorithms used and compared with decision trees are logistic regression, linear discriminant analysis, K-nearest neighbors algorithm, Gaussian naive,

Table 4.3: Advantages and Disadvantages of Model-based learning IDS

Model-based learning IDS	Advantages	Disadvantages
	Automatic detection process	Computation resource requirement is high
	The Quality of the data affect detection accuracy	Requires time to train the model
	Uses data to process it in real-time	Training data overfitting is present
	Detection accuracy is high	Exhibits complexity

Bayes algorithm, stochastic gradient descent, and K-mean algorithm.

Unsupervised Learning IDS

Unsupervised learning algorithms have been widely used to identify patterns and define correlation without training in the predefined data set. For example, it includes clustering in the Unsupervised Leaning (USL) to determine the similarity between data groups. Some suggested works used this technique to detect abnormalities in the UAV networks. In [66], the author proposed Long Short-Term Memory (LSTM) to improve IDS in the network. They modeled the time series as a problem to detect an anomaly on the UAV network. The authors trained their model on a data set that contained only normal sensor data. In their suggested solution, the authors expected that the prediction model would face difficulties with uncertainty intervals. Therefore, they calculated the residual variance to identify the anomaly detection's point and deviation in the normal pattern data mode. Also, the authors used pneumatic lifting and north direction speeds to evaluate the proposed method to detect anomaly points. They evaluate the anomaly detection performance based on three matrices false positive rate, false negative rate, and accuracy. In another work [67], the author proposed a deep learning approach to detect outlier behavior in the UAV network based on monitoring the time series of the sensor data in the UAV. The detection system was a combination of Convolutional Neural

Network (CNN) and Convolutional Long Short-Term Memory (ConvLSTM). The CNN extracted the features from the dataset and fed it to LSTM. The author used multiple time windows to evaluate the model from 0.5 to 5.0 s. They assessed the performance of their proposed model, which resulted in high detection accuracy when the time window was at the maximum value of 5.0 s. The paper [17] proposed a GPS detection method using the long short-term memory LSTM algorithm. They used latitude, longitude, velocity, and acceleration as a parameters in the method; hence, the method predicted the location based on the given flight path. To detect the anomaly, the author depended on the difference between the position provided by the GPS and the predicted position.

Reference [13] proposed a combination of detection techniques of artificial neural networks and Kullback-Leibler divergence. They calculated KLD in two types, forward and backward, to detect the deviation in the data generated by the UAV. To accumulate value and increase the detection accuracy, the author identified the divergence value for each particular time interval. After that, they added the value to the entire time series. The author in [55] proposed a technique to detect anomalies in the cell networks. The method divided into two parts: LSTM algorithm to detect overload in the base station and deploying drones as flying base stations for the backing up. The author used a real dataset from the Call Data Records (CDR) of Milan. The algorithm monitored the deployed cells, and once it detected any overload on the cell, it triggered the UAV to fly next to cells to assist in relaying data to maximize cell coverage and minimize energy consumption during the mission. The author in [68] proposed a deep learning technique to detect intrusion in the communication links by suggesting an intrusion detection system. They applied the detection technique to the scenario in the smart city. They proposed UAV to UAV link, UAV to Road Side Unit (RSU), and vehicle to RSU. The UAV connects with another UAV to exchange sensitive data. At the same

time, this UAV loads the data to RSU, which sends the necessary data to the vehicle navigation system for auto driving. In [69], the author proposed a neural network to recognize the fault during a UAV mission. They proposed embedded deep learning to detect faults in the UAV system. They presented Bi-LSTM and CNN to classify the fault in an encoder-decoder logic. They used the temporal data generated from onboard sensors such as accelerometers, Inertial Measurement Unit (IMU), etc. They evaluated their method, which got a high detection accuracy of 99 and 85.00 in real-time data. The authors in [70] addressed the security issues in cellular-connected UAV networks. They suggested IDS-based ML using a dataset generated by a Canadian cybersecurity laboratory. They, in this technique, used logistic regression LR, k-mean, decision tree DT, stochastic gradient descent SGD, and linear discriminant analysis LDA. They have concluded that the decision tree supered the other ML where it reached 99.99 accuracy and false negative. Also, in [71], the author provided anomaly detection-based IDS to detect abnormal patterns in the data of the Internet of flying things. The researchers used the ECU-IOFT dataset [72] to train and test five algorithms, such as Histogram Based Outlier Score (HBOS), Local Outlier Factor, K-Nearest Neighbors (KNN) Local Density Cluser Outlier Factor (LDCOF), and Cluster-Based Local Outlier Factor (CBLOF). The other part of this work focused on cracking attack de-authentication, including API exploits in the WI-FI. Their experimental results showed that the Accuracy was between 21.42 and 84.69 for KNN, and based on these results, this work was considered one of the suggested works used to detect anomalies in the IoFT. In another work [73], the authors addressed a monitoring approach to the anomaly detection of fleet drones. They proposed machine learning techniques to detect abnormal behavior during the mission. The system operates in two forms: to detect strange behavior in the overall drone and to identify which drone exhibits abnormal behavior. The authors validated

their method using real flight data in online mode.

In this section, the primary technique used to detect anomalies over the network were addressed. The previous works in the IDS were classified into signature-based, knowledge-based, and anomaly-based. This section showed that each one of these techniques has drawbacks and some limitations related to usability, expandability, and scalability. First, in the signature-based, we have seen that the pattern's signature needs to be updated frequently. Therefore, this technique leads to an increase in cost and time. The specification based on the main drawback faced in this technique is that the UAV mission path is unreceived and changeable, and the bounds platform is a software-operated flight control that affects the environmental impact. Therefore, these multiple drawbacks presented in the technique made applying it in the UAV domain challenging. Third, anomaly detection is divided into supervised learning detection techniques that need to obtain enough data to train the model and get high detection accuracy performance. Therefore, this technique is costly and a reason for wasting time. On the other hand, anomaly detection is an unsupervised technique. The unsupervised technique does not require time to train the model. Still, it has limitations in the previous works, where it challenged classifying the issues into multiple attacks simultaneously and faced challenges to classify them accurately.

5

Framework

5.1 Motivation

In the coming years, 6G will play a critical role in handling large amounts of data and providing a fully covered area through UAV-FBS. The 6G will offer a fast data rate and low latency compared to previous generation networks, such as 5G, 4G, 3G, 2G, and 1G. Due to the significant advantages that will be provided by 6G, such as ultra-high--speed, low latency communication uHSLL, ultra-broadband uMUB, ultra-high data density uHDD, reliable network access, and a wide coverage area, some new practical applications will be provided, for example, surveillance. Therefore, various sensors and cameras will be deployed for critical missions, monitoring, and collecting sensitive data

in harsh environments. UAVs will, hence, play a key role in enhancing connectivity with these deployed components on the ground by deploying the UAV-FBS. However, while the UAV is classified as the main component of the aerial layer of the 6G network, it will face security threats that will lead to discontinuity with the nodes or receiving a false GPS signal to deviate from the planned area.

Sophisticated malicious threats, such as jamming and spoofing attacks, are the main challenges that cause performance degradation on the communication part of the UAV system. UAVs in some critical areas will be deployed as flying space stations to enhance terrestrial networks, handle high user demands, and provide coverage connectivity on a large scale area. However, adversaries exploit natural medium transmission to block transmitted signals by sending RF noise or mislead the UAV path by injecting a false GPS signal. During the jamming attack, this work expects hackers to know the modulation spectrum and frequency during transmission. They intentionally target command and control links to block communication by sending multiple radio noise signals. In the spoofing attack, the adversaries deliberately sends high-power signals to mislead the GPS sensor to receive a fake GPS signal. Hackers can execute their malicious adversaries by receiving signals from the GPS target and regenerating them to the target UAV. The target UAV will be misled and deviate from the planned path to the attacker's zone. Hence, preventing and securing communication is urgent to avoid this extreme action. Therefore, the IDS-based model in the drone system is an efficient way to ensure reliability in transmission and ensure that the necessary data are available and received without any modification. Traditional IDS effectively prevented data transmission in a centralized network system, so they ensured the integrity and availability of the systems. However, due to the possibility of deploying conventional IDSs in the UAV network system, they need to be more suitable for the decentralized system. Therefore, researchers

started thinking cleverly about how to innovate new IDS techniques. This work increases security and reliability, and it avoids triggering a false alarm. Also, it can distinguish between unintentional and intentional interference compared to previous works.

In this work, an effective detection technique is designed to improve transmission security, ensuring reliability, integrity, and availability simultaneously. As shown in 5.1, it applies the unsupervised ML model to detect network traffic anomalies such as jamming or spoofing attacks.

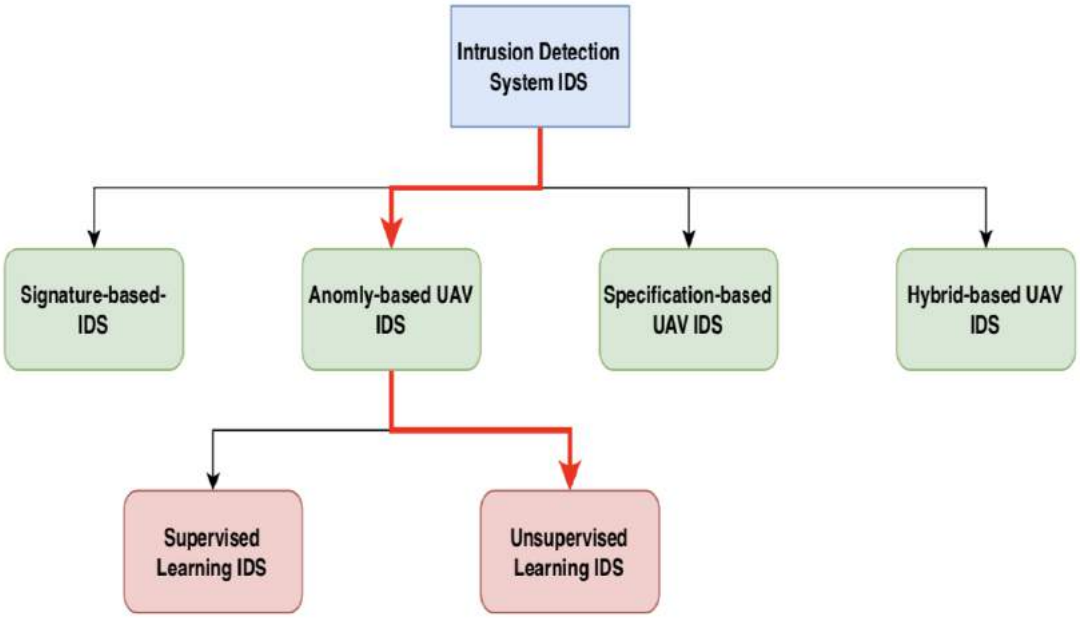

Figure 5.1: The architecture of distributed intrusion detection system

The abnormal detection algorithm starts once it recognizes suspicious activity in network traffic. Once the anomaly is detected, the first algorithm sends the detected anomaly to the second algorithm to confirm the detected abnormality and analyze the abnormal behavior. The second layer of algorithms confirm the initial detection of the anomaly as intentional interference jamming attack, or deviation of the UAV path because of the spoofing attacks.

5.2 System Model

This section describes the system model components connected through a network for communication and information exchange. Some obstacles prevent the transmitted signals from reaching the deployed sensor and camera. The existing obstacles can include trees, a high mountain, etc.

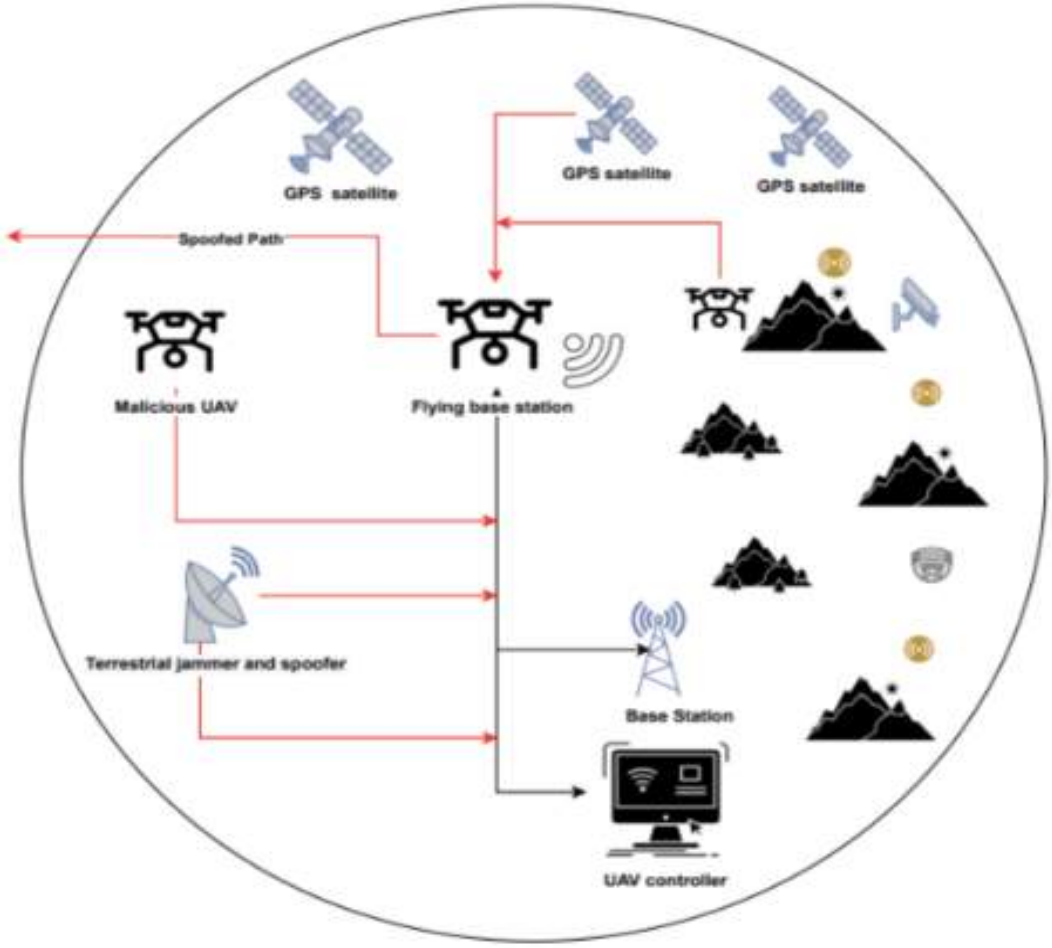

Figure 5.2: The architecture of distributed intrusion detection system

Figure 5.2 shows the security threats present in the network scenario and how they act to damage communication. The system model consists of the base station BS, satellite S, UAV, controller UC, and cellular connected UAV U_u with multiple jammers and spoofers on the ground M_j, s and malicious UAV U_m along the UAV base station

U_{Bs}. The adversaries attack communication links, control, and command to block links between nodes, and if they plan to deviate UAV from its initial plan, they inject false GPS signals.

5.3 Attack Model

This work concentrates on the two most common dangerous attacks on the UAV wireless network: jamming and GPS spoofing attacks. The jamming attack starts its malicious plan by sending RF jamming signals, such as constant jamming, intermittent, reactive, and deceptive jamming attacks. Each technique has a specific approach to executing its target on the UAV wireless network. They emit an interference signal at the same radio frequency the UAV uses to disrupt communication. In addition, they respect characteristics to ensure their jamming signal efficiency to achieve their goals, such as strategy, duration, time, location, and target. However, the second malicious action is to inject a false GPS signal into the transmission to mislead and deviate the UAV from its planned trajectory. This attack is executed when the jamming attack blocks communication with the controller, and the UAV switches to self-flying mode and relies on a GPS signal. Hence, the fake signals lead the UAV to crash or land in the hacker's extreme zone.

In the suggested scenario, a ground-to-air UAV link transfers command to the UAV and observation to the GC. As shown in Figure 5.1, the jammer attempte to disrupt the transmission link by launching the attack. Three use cases are considered: Tg, Rg, and J represent the transmitter on the ground, the receiver on the ground, and the jammer, respectively. The position of Tg (x_T,y_T,z_T), (x_R,y_R,z_R) represents the transmitter and receiver's location, and (x_J,y_J,z_J) is the jammer location. The GC periodically sends

a specific command to the receiver UAV. It provides the power needed to transfer the signal simultaneously and from a close distance; the jammer is located and emits a jamming attack J in the same frequency over communication with the high-power signal to interrupt the transmission. In this scenario, this work did not consider the exact position and power.

Typically, the RSS is affected by the noise signal, where it can be derived as:

$$RSS = P_{\mathrm{TR}} + P_{\mathrm{JR}} + P_{\mathrm{N}}. \tag{3}$$

.

P_T is the transmit power and the power strength represented by $P_T R$. The path loss exponent is represented by $-\alpha$.

Once the UAV receives the packet, it starts investigating the preamble to determine whether the received packet is valid or corrupted. In the case of the jamming signal, interference exists in the packet and decreases the SNR which is calculated as :

$$SNR = \frac{P_{\mathrm{TR}}}{P_{\mathrm{JR}} + P_{\mathrm{N}}} \tag{4}$$

The BER is related to the SNR in the digital modulation system, so it relies on the jamming signal, which is a noise signal; the hacker attempts to decrease the SNR. Hence, it leads to an increase in the BER. BER is defined as:

$$BER = f(SINR) \tag{5}$$

where f represents the decreasing function determined by the modulation system.

The packet delivery rate represents the packet delivered successfully to the receiver. Therefore, it represents the packet as multiple bits. It is identified as

$$PSR = (1 - BER)^{nbits} \tag{6}$$

In the equation bits are represented as N bits in the packet, so the packet error rate was identified as:

$$PER = 1 - PSR. \tag{7}$$

The GPS spoofing signal can effectively disrupt the UAV's mission. The hacker creates fake signals about false drone positions to achieve their goals. The general manipulated signal has a higher power; therefore, the hacker's signal is higher than the genuine signal. Hackers expect that not all targets fall into their coverage area; hence, they follow a repeater-based spoofer technique to collect all G satellite and regenerate them to mislead the UAV through processing times and clock offset [74]. Therefore, receiver j receives a false pseudo-range measurement such as $Z_{fj} = \{z_{fn,j}\}_{n=1}^{G}$ and false pseudo-ranges for the satellite n represented in

$$z_{n,j}^{t} = \sqrt{(x_n^g - x_j^t)^2 + (y_n^g - y_j^t)^2} + c(dt_n - dt) + \mathcal{N}(0, \sigma_t^2) \tag{1}$$

In the presence of the spoofing attack, the received signal will be modeled as follows

[75]:

$$r(t) = s_{\text{gn}}(t - \tau_{\text{gn}}; A_{\text{gn}}, f_{c,\text{gn}}, \phi_{\text{gn}}) + s_{\text{sp}}(t - \tau_{\text{sp}}; A_{\text{sp}}, f_{c,\text{sp}}, \phi_{\text{sp}}) + n(t) \tag{2}$$

The $s_{\text{gn}}(t)$ represents the genuine GNSS signal, and the $s_{\text{sp}}(t)$ is the fake signal generated by hackers. τ is the delay in signal propagation. $n(t)$is the additive noise.

5.4 Simulation Environment

This section shows the steps to simulate anomaly detection and the necessary scenarios to collect the dataset in the 6G communication network using Matlab, Python, and other open software resources.

- The goals and scope of the simulation were defined and planned before the simulation was run. These goals and scope included expected behavior and attacks needed to be detected, the dataset to use, and the performance metrics for evaluation and improvement.

- Processing and filtering the datasets is needed. In the simulation experiments, this work used public datasets; hence, pre-processing and filtering were performed to use valuable data and exclude outliers or noise data. Furthermore, ML needs suitable data in a specific format so that data conversion is used in addition to data separation into training, testing, and validation data.

- Machine learning algorithms such as autoencoder,oc-svm, and K-means were selected. Python libraries, such as sci-kit-Learn and TensorFlow, were used to implement and train these algorithms.

- Train and test models. Once the model was built and designed, four datasets were used to train and test these models. Various processes were used, such as splitting the dataset into suitable data for training, testing, and validation. Multi-variate matrices were used to measure and evaluate the model's performance, such as F1, recall, and accuracy.

- Improving the model. This was achieved after validation of the model. Three factors were considered to enhance performance: models, clustering models, and hyperparameters.

- Result visualization. The visualization process was performed to gain insight and better understand the results.

5.5 Dataset

A high-quality dataset to train a machine learning model is essential for consistent and trustworthy results. Fortunately, some authors have worked on collecting datasets in different areas, such as UAV data, and have made them accessible to the public. Therefore, two published datasets were used to train and evaluate the designed model in this work. However, these selected two datasets have a lack of data and included only some of the features needed for this work. This work compensated for the lack of data by simulating various scenarios to collect two more datasets to assess the models accurately. Typically, during the collection of normal data in real-time simulations, the process does not face

obstacles in collecting normal observation events. On the other hand, researchers must address some challenges when implementing abnormal scenarios, where some expected consequences must be considered. For example, implementing hacker scenarios leads to crashes and damage to the experimental vehicle, which results in high costs. Additionally, abnormal scenarios are comprehensive, and all possible malicious scenarios must be covered. Therefore, the simulation included two types of operation, normal and abnormal, to collect the target parameters in the 6G UAV communication link and the GPS sensor to gather data that was not included in the published dataset, with features related to parameters RSS, SNR, PDR and Throughput. In addition, new datasets collected included GPS correlation such as alt, lat, long, velocity, and acceleration, to use more datasets and enhance the model's accuracy in detecting spoofing attacks. The tools and open source systems such as Matlab, Omint++, Gazebo, and QGroundcontrol were used to measure and collect the needed data by simulating two scenarios normal and under attacks, to have an efficient dataset.

This work used the UAV attack dataset [76] to extract specific features to train and evaluate the performance of the algorithms in the model. This dataset was collected based on the simulation of three types of flights such as flight in normal mode, flight under jamming attack, and flight under spoofing attack. Therefore, this dataset is divided into benign, jammed, and spoofing data. Flight data was extracted after completing each flight, so the number of records contained in the dataset is 7516 and is classified into:

- Normal flight: the attack in this scenario was absent, so dataset included the normal data pattern of the UAV.
- Jamming attack: the jamming attack happened, and the data was collected under the scenario of a jamming attack.

Alt	Lat	Long
40981	362048146	1382529220
40989	362048146	1382529219
41074	362048149	1382529222
41124	362048153	1382529221
41112	362048158	1382529221

Table 5.1: UAV normal positioning dataset

- Spoofing attack: the spoofing attack was presented in this scenario, and the system logs were collected under the scenario of a spoofing attack.

The various operations of the real-world behavior of UAVs were collected in dataset, as shown in Table 5.1 and 5.2. In this dataset, two UAV behaviors were simulated as normal and abnormal incidents, which provided comprehensive information, including variability in UAV status, factors that affected flight, and various data points. Features such as alt, lat, long, velocity, and acceleration are included in this dataset, allowing trained algorithms to recognize significant changes in the parameters of the selected features or sudden increases in velocity and acceleration which are signs of a GPS spoof attack. Therefore, the algorithm learned the malicious scenario and discerned patterns in the data pattern under both conditions, normal or deviation, so once a GPS spoofing attack was launched, the algorithms recognized it. These scenarios facilitated the learning process by exposing the algorithms to abnormal events, and these algorithms were trained to identify abnormalities in the patterns in the absence of labels. Therefore, this work particularly benefited from recognizing the evolving GPS spoofing techniques, giving the trained algorithms a comprehensive view to identify UAV behavior and allowing the algorithm to learn the deviation in the parameters in a different dimension.

The controlled environment was built to capture data in a simulated environment and in a live scenario. In the simulated environment, multiple open sources were used, such

VX	VY	VZ	ACCX	ACCY	ACCZ
-0.001238745	0.009944092	-0.00264930	-0.02669428	-0.07598478	-9.812940949
-0.00154938	0.0120948	-0.002639220	-0.030594320	0.0242958	-9.79204958
-0.00189847	0.0190450	-0.00340058	0.02729456	0.0269584	-9.79620499
-0.00284928	0.02405949	-0.00449203	0.007594430	0.07529489	-9.78922094
-0.003015042	0.010945903	-0.00445920	-0.69439509	-0.038892061	-9.772394590

Table 5.2: UAV normal dynamic behavior dataset

as PX4 and Gazebo, for the virtual environment. Various UAVs were used during the simulation to offer various UAV models. In addition, the data collected utilized hardware and software in the loop for multiple UAV behaviors. Keysight EXG N5172B was used for signal generation to offer accurate coordination during live experiments. The Great Scott gadget hackRF was used to conduct the GPS spoofing attack. Additionally, to set up a control system for the experiment, the PX4 autopilot with version 1.11.3 ran on the Pixhawk 4 flight controller. Holybro S5OO UAV model was used due to its versatility and stability in various flight scenarios. Additionally, QgroundControl version 4.0.9 provided an interface for users to control and monitor UAVs to configure the GPS. All data recorded in this experiment stored in ULOG files for post-flight data analysis. The ulog2csv was used to convert ULOG files to CSV format.

Various attack scenarios are included in this dataset, ranging from simulated to live experiments, as shown in Table 5.3 and 5.4. The GPS message was generated during the simulation through Gazebo by the hooking configured for this mission. This hook helped to inject fabricated GPS messages into the UAV navigation system, including a wrong location. Hence, it resulted in a controlled environment to generate and collect real and manipulated GPS UAV messages. On the other hand, during the live simulation, the manipulated GPS signal was conducted through hardware. The HackRF software used in GPS-SDR-SIM allowed incorrect GPS coordinates to be generated in the receiver. Also, HackRF was used to jam the GPS by adding Gaussian white noise to disrupt

alt	lat	lon
49204	362048117	1382529164
49187	362048117	1382529164
49172	362048115	1382529165
49122	362048115	1382529162
49210	362048108	1382529166

Table 5.3: UAV attack positioning dataset

VX	VY	VZ	ACCX	ACCY	ACCZ
-0.006483048	-0.06854320	-0.01945429	-0.05843200	-0.0194589	-9.828454030
-0.002945684	-0.03845440	-0.0955900	-0.05944832490	0.00394660	-9.8294540
-0.00945690	-0.02845590	-0.00295842	-0.08558439	-0.04329445	-9.8685390
-0.00294569	-0.03945520	-0.002495520	0.05284449	0.092945669	-9.78458550
-0.00584294	-0.00945580	-0.00585439	0.053922944	-0.05843950	-9.8686539

Table 5.4: UAV attack dynamic behavior dataset

the GPS signal. Keysight EXG N55172b was used in the live experiment to generate messages with accurate coordination to succeed in the GPS scenario.

The CRAWDAD dataset [77] was also used for testing and validation. The Federal Aviation Administration and the tutorial sheet on Unmanned Aircraft [78] explained that the altitude in the UAV is stable during the mission, so the third dimension of the movement of the UAV is fixed during the flight. Also, there is a shortage in published jamming attacks providing UAV communication under jamming attacks, so this dataset is used to evaluate the performance of the two layer algorithms in addition to the simulated 6G UAV dataset, explained in the following subsection. Due to the comprehensive scenarios included, this dataset was used to train and evaluate the model to identify the anomaly in the signal parameters and recognize it as a jamming attack. As

RSSI	SNR	PDR
-48.38	22.35	0.999999895
-49.14	21.62	0.999999465
-49.92	20.85	0.999997415
-50.69	20.07	0.999988835
-51.46	19.3	0.999958016

Table 5.5: Jamming attack dataset

shown in Table 5.5, this dataset included signal parameters such as RSS, SNR, PDR, and throughput, and introduced various scenarios of RF jamming attacks. Therefore, the autoencoder and K-means used both normal and abnormal data to train and evaluate the network's performance behavior based on the metrics included in the dataset. Hence, this dataset makes algorithms in unsupervised learning facing development in the attack evolving and are suitable for securing a 6G UAV network.

This dataset is not just a theoretical construct but a practical representation of real-world scenarios in the VANET vehicle ad hoc network. It was collected in various environments, such as highways and suburban areas, using Software Defined Radio (SDR) to capture and collect the dataset. This device measured RSS, SNR, PDR, and throughput parameters in the RF system. Additionally, the dataset was combined with a mobility model that simulated dynamic movement to ensure the data was captured correctly. This practical approach ensured the relevance and reliability of the dataset in real-world applications. The jamer attack was fixed during the attack scenario, and concurrency introduced signal noise to a specific target. This jammer continually increased its plan to send more noise signals, interrupting communication between the receiver and the sender. Once the jammer achieved its goals closely, it reduced the jamming signal to move back to a safe distance to mislead the detector for a while. This action was frequently repeated, covering distances between 5 and 10 meters with frequencies between 5 and 15 times per unit.

5.5.1 Dataset Creation

This research compensated for the lack in the datasets by simulating multiple scenarios to collect new data. By using Matlab, two UAV scenarios were performed: normal

behavior and operation under jamming attack. These scenarios led to generating two 6G UAV network datasets. Each of these dataset was made up of 23,938 samples. The normal UAV dataset was simulated without external attack to collect signal parameters such as RSS, SNR, PDR, and Throughput. On the other hand, the abnormal dataset was generated when jamming scenarios were executed to interrupt the communication. UAVs in this scenario faced various levels of attack. Hence, these noise signals disrupted the transmission and affected the link.

Both scenarios, normal and jamming attacks, were implemented to evaluate the network performance in the 6G UAV link. In the typical scenario, the parameters of the UAV and GC model were determined using the directional beamforming antenna pattern for both the UAV and the GC. The power of the transmitter and receiver was identified as 35 dBm for the UAV and 40 dBm for the GC. The simulation also modeled data transmission as channel condition, coding, and modulation. The data transmission parameters were identified as package size 1500, polar code was used for advanced error correction, and modulation was used as qam256 for high data rate. Also, the signal propagation model was designed with the highest precision for UAV and GC, so it took into account multiple factors, including shading where it was set to 5 as effect. The path loss used terahertz, and fading effects used Rayleigh with maximum doppler shift to 250 Hz. The position of the UAV and GC were represented as $(x_{\mathrm{U}}, y_{\mathrm{U}}, z_{\mathrm{U}})$and $(x_{\mathrm{G}}, y_{\mathrm{G}}, z_{\mathrm{G}})$ with the UAV coordinates in (90,40,5) and GC was fixed at (0,0,0). In the signal propagation, the distance between the UAV and GC was identified as $d = \sqrt{(x_{\mathrm{uav}} - x_{\mathrm{gc}})^2 + (y_{\mathrm{uav}} - y_{\mathrm{gc}})^2 + (z_{\mathrm{uav}} - z_{\mathrm{gc}})^2}$, and considered factors such as path fading effects, path loss and shadowing, to measure the quality and strength of the signal. Hence, the communication in the link was normal and the signal parameters are stable during transmission as shown in Figure 5.3.

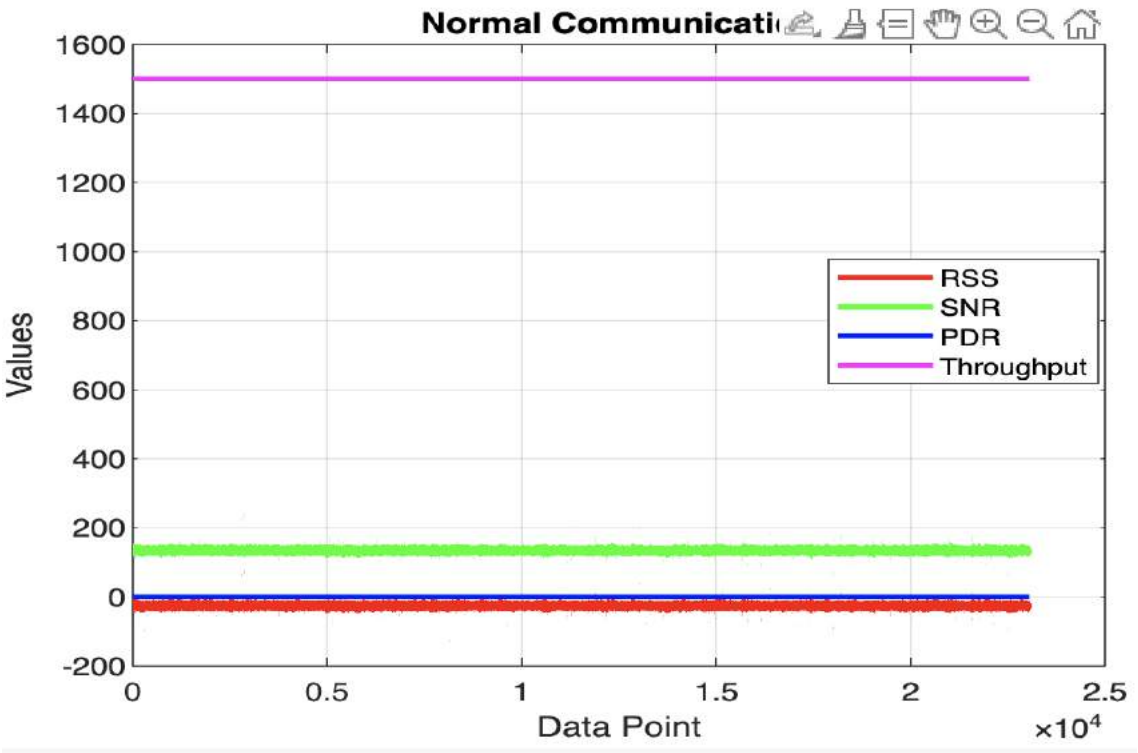

Figure 5.3: Normal scenario

The main jamming attack scenario involved sending a noise signal to disrupt communication. The simulation of communication function model considered the jamming attack and incorporated it into data transmission and signal propagation. The signal propagation function model considered four types of jamming attacks in the link between UAV and GC with factors such as path loss, fading, and shadowing. In addition, the data transmission function integrated jamming attacks into data transmission and reception. Therefore, the jamming scenario in the communication affected the received signal power on the receiver side, presenting the noise attenuation caused by malicious action. Thus, the measurement of the RSS, SNR, PDR, and throughput recalculated signal metrics to reflect the effect of the noise signal. RSS and SNR were directly affected by noise signals, affecting signal strength and decreasing SNR. The PDR was calculated based on the received power and the difference between the success of packet reception and the damage caused by the noise signal. Finally, the throughput affected the PDR,

so it was recalculated based on the new PDR.

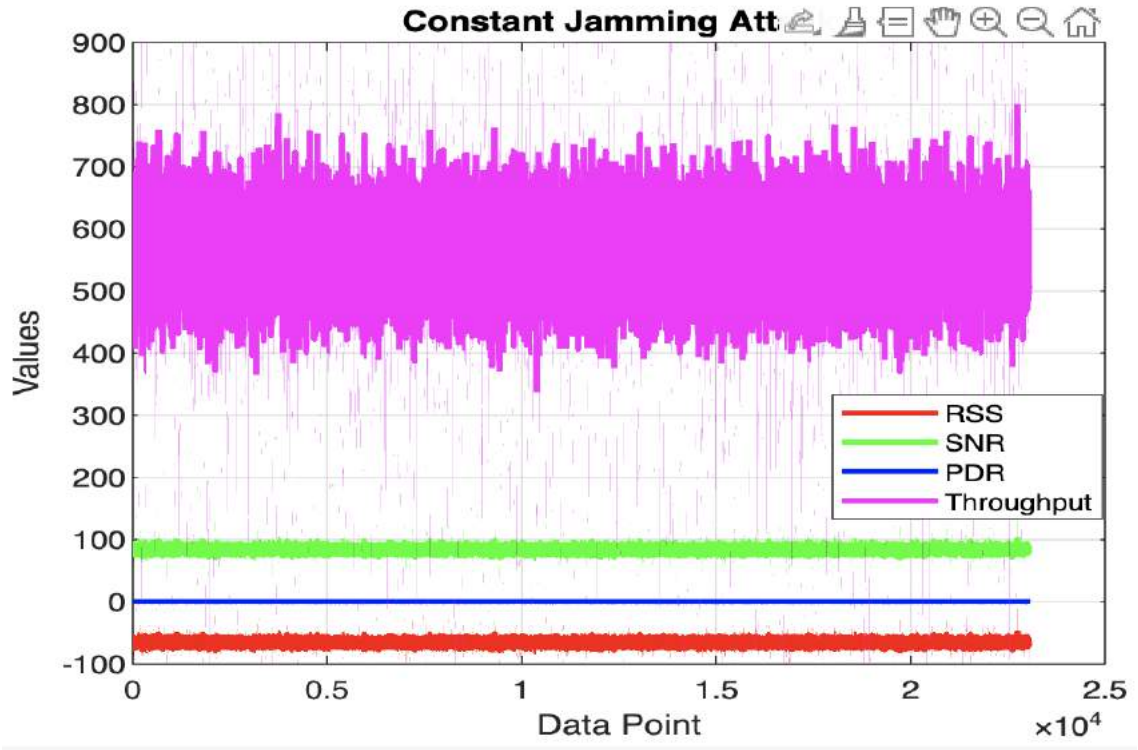

Figure 5.4: Constant jamming attack

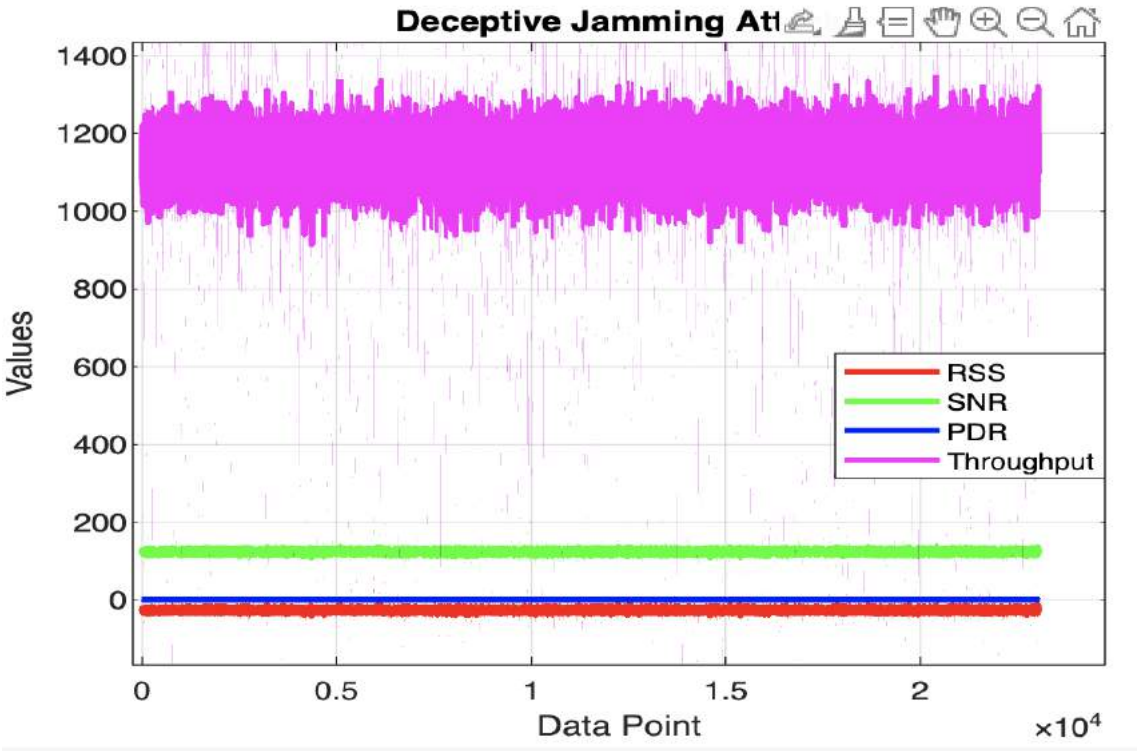

Figure 5.5: Deceptive jamming attack

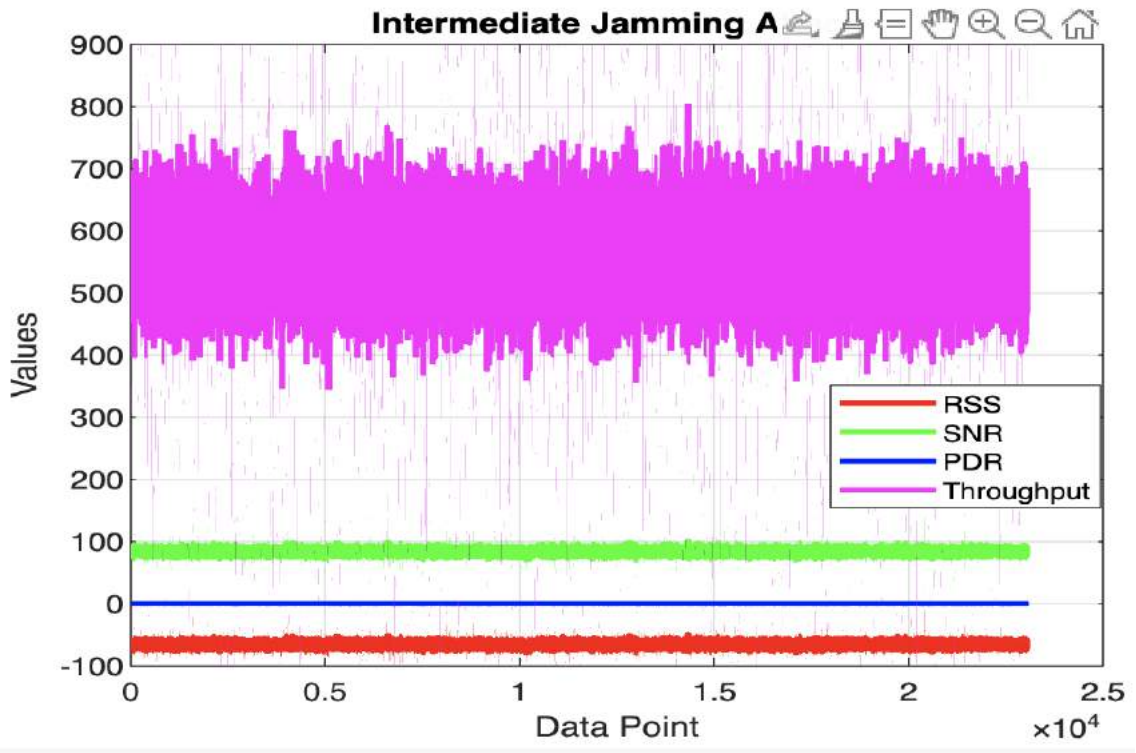

Figure 5.6: Intermediate jamming attack

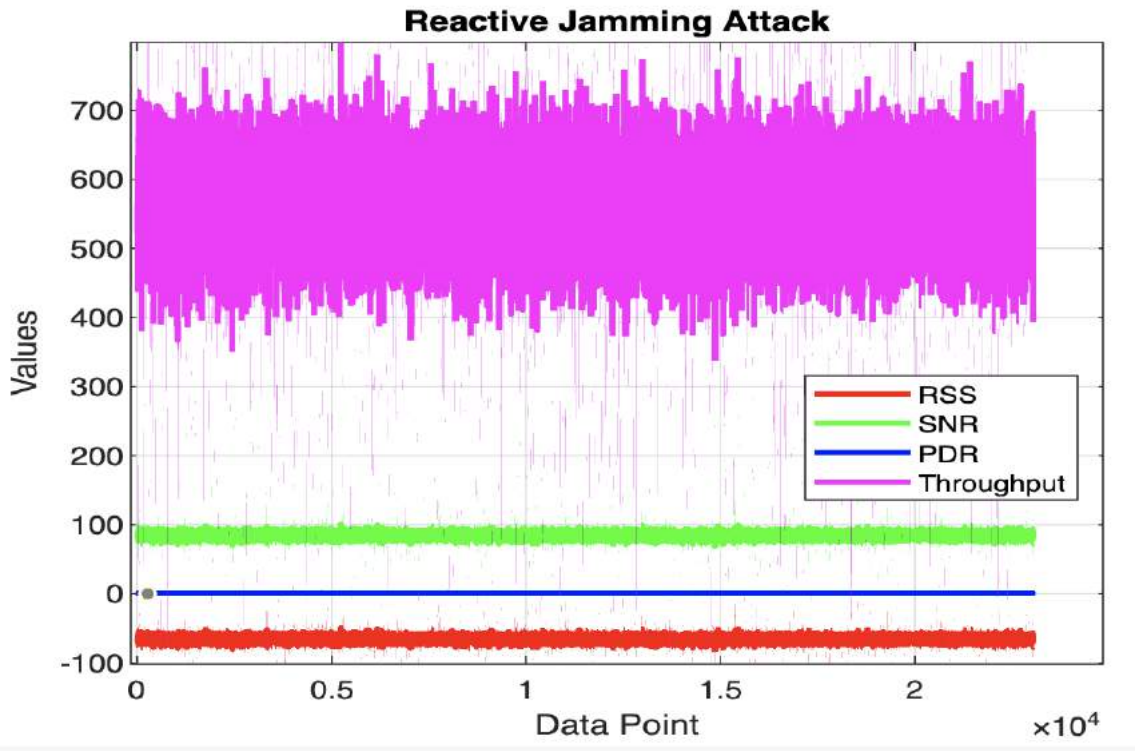

Figure 5.7: Reactive jamming attack

The simulated 6G UAV dataset provided valuable measurements of signal parameters in the link to detect anomalies, specifically jamming attacks, as shown in Figures 5.4 - 5.7. These data provided a comprehensive view of the signal of normal and jamming attacks to train the model effectively. The data included the signal parameters in need, such as RSS, SNR, PDR, and throughput measuring signal parameters of the UAV and GC. Therefore, presenting both measurements made the dataset comprehensive, captured various scenarios, and allowed algorithms to learn and recognize different data patterns related to ordinary status or noise signals. Therefore, this dataset met the diversity characteristic, enhancing the generality in the algorithms by simulating multiple situations to minimize False Positive (FP) or False Negative (FN). Furthermore, the dataset's size supports the algorithms' validation by allowing for comprehensive events.

The simulated data is represented in RSS, SNR, PDR, and throughput columns, representing a set of observations of different scenarios as shown in Table 5.6 and 5.7. These metrics give insight into the system's performance in various sections, leading to investigation and optimization efforts. In the RSS column, the value shows how the received signal looked, providing insight into attenuation and signal propagation. In the SNR column, the observation shows the received signal quality measurements with the noise signal presence to evaluate the effect of jamming of this signal. In the PDR column, the observation shows the rate of successfully delivered packets and how the rate decreased when the jamming attack affected the communication. The throughput column shows the rate of the data transmitted successfully, in addition to the rate during the jamming attack. Therefore, this data gives a comprehensive view of the communication line for optimization and identifies the deviation that occurred in the target metrics.

RSS	SNR	PDR	Throughput
-23.3860247	136.6139753	0.9	1500
-18.2011532	141.7988468	0.9	1500
-34.5720807	125.4279193	0.9	1500
-22.08799998	137.912	0.9	1500
-24.2616323	135.7383677	0.9	1500

Table 5.6: Simulated normal signals

RSS	SNR	PDR	Throughput
-65.26654777	84.73345223	0.378413831	567.6207462
-68.84881092	81.15118908	0.344297039	516.4455583
-62.26878152	87.73121848	0.406963986	610.4459783
-57.60502758	92.39497242	0.45138069	677.0710346
-65.80742805	84.19257195	0.37326259	559.893885

Table 5.7: Simulated abnormal signals

The simulation to collect and increase the dataset, included GPS normal data and spoofing attack, implemented by leveraging PX4, Gazebo, and QGroundControl App QGC. During the simulation, the scenarios represented the normal instances of the UAV and GPS messages. It described the coordination and dynamic behavior of the UAV related to velocity and acceleration. Once the spoofing attack was launched, the feature observation changes were recorded and stored in the file. Hence, the file included the UAV coordination and dynamic movement dataset. In this work, PX4 was used to run in a simulated environment. The selected UAV model was the default Quadrotor. The QGC was used to control flights and missions. GPS spoofing was achieved by modifying the original GPS signal using the Gazebo shared library. Therefore, it manipulated and created an incorrect location message to transmit to the simulation environment. The algorithm depicted in Figure 5.8 shows how the fabricated GPS spoofing was generated in Gazebo. The shift from the planned path was represented by (ϵ), which represents the current location from the starting point. When the malicious spoofing action started, the algorithms stored the current location (pos) in a variable while (x) of the UAV was stored. At this moment, the UAV received the fabricated GPS messages and moved away

from its planned route as shown in Figure 5.9. Therefore, the simulation environment shows how a UAV was affected by a manipulated GPS message and how it deviated from its intended path. The dataset used includes the main features needed, for example alt, lat, long, velocity, and acceleration, in normal and abnormal scenarios.

```
Impact caused by spoofing attack on each parameter B(ε_altitude, ε_latitude, ε_longitude, ε_velocity, ε_acceleration)
X is the distance of spoofing attack
Set init to 1 to indicate the start of the flight

Output: In the Gazebo environment, the deviation occurred by ε

1 initialization;
2 while spoofing is active
3     get local position and parameters (B) of UAV,
        (Position, altitude, latitude, longitude, velocity, acceleration)
4     if init = 1 then
5         store current position and parameters (Z) of UAV,
           (A, Z_alt, Z_lat, Z_long, Z_velocity, Z_acceleration)
6         set init to 0
7     end
8     if (Distance(Z, Position) > X) then
9         Transmit (Position + ε_longitude,
                    altitude + ε_altitude,
                    latitude + ε_latitude,
                    longitude + ε_longitude,
                    velocity + ε_velocity,
                    acceleration + ε_acceleration) to UAV
10    else
11        Transmit the accurate location and parameters,
           (Y, altitude, latitude, longitude, velocity, acceleration)
12    end
13 end
```

Figure 5.8: Creating GPS spoofing messages

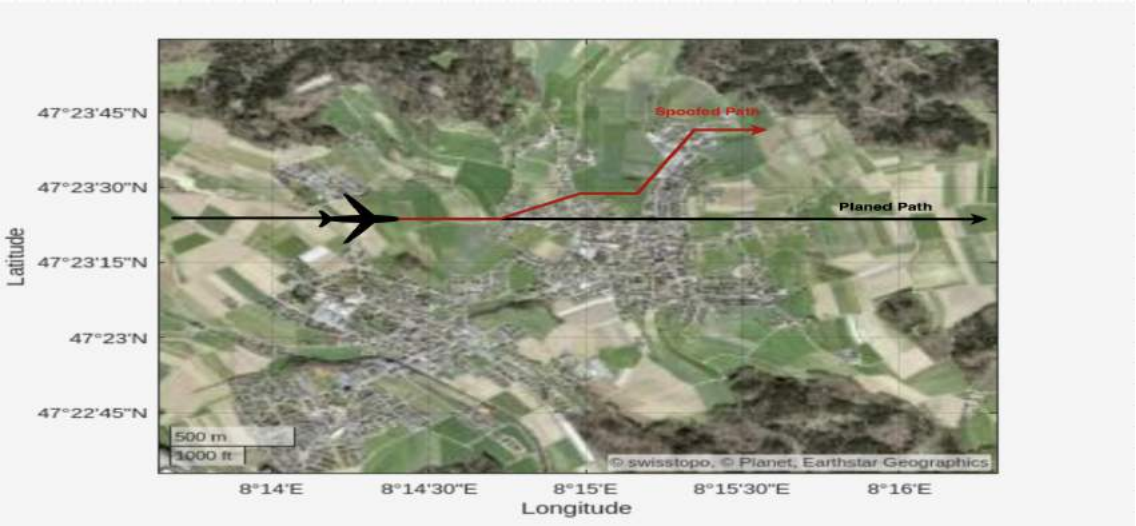

Figure 5.9: Example of flying and spoofed path

Each row in this dataset includes columns representing the UAV location as alt for altitude, lat for latitude, and long for longitude and its dynamic status such as velocity and acceleration, as shown in Tables 5.8 - 5.11. The columns represent the UAV status in the integer value recorded at the data point. The velocity shows how fast the UAV

Altitude	Latitude	Longitude
49980	439451550	-788968509
49979	439451550	-788968510
49993	439451549	-788968510
49999	439451548	-788968510
49998	439451548	-788968509

Table 5.8: GPS coordination simulated normal data

VX	VY	VZ	ACCX	ACCY	ACCZ
-0.001260226	0.009861966	-0.002602668	-0.026609369	-0.075475879	-9.810001373
-0.001533535	0.010381611	-0.002698347	-0.030851338	0.024228677	-9.79118824
-0.001887404	0.010645169	-0.003754747	0.027984818	0.026756179	-9.796936989
-0.002844293	0.010396368	-0.004532666	0.007293397	0.075241081	-9.782605171
-0.003015042	0.010762365	-0.004381465	-0.139361888	-0.038892061	-9.772027016

Table 5.9: UAV movement simulated normal data

Altitude	Latitude	Longitude
49982	439451550	-788968509
49989	439451551	-788968511
50001	439451550	-788968510
49990	439451549	-788968511
50004	439451550	-788968509

Table 5.10: GPS coordination simulated abnormal data

VX	VY	VZ	ACCX	ACCY	ACCZ
-0.006000763	-0.001113015	-0.010537908	-0.010475708	-0.015321761	-9.811365128
-0.006954138	-0.001167268	-0.010231952	-0.062327698	0.001425983	-9.853741646
-0.006469611	-0.00118432	-0.0094188	-0.056902725	-0.0677993	-9.822723389
-0.006058747	-0.001686975	-0.008144341	0.051815644	0.090789631	-9.716182709
-0.00597062	-0.002140084	-0.007122496	0.059347354	-0.00251192	-9.868725777

Table 5.11: UAV movement simulated normal data

is during flight, with its acceleration showing the change in speed per time. Since, all features were not represented in the previous public dataset, as shown in Table 5.12, this section showed how this work simulated and collected new dataset used in this research.

5.6 Feature extraction

Feature extraction is a process of extracting sufficient features that describe UAV status in the system log to recognize the UAV's dynamic behavior and identify any deviation of the data parameters during transmission. Typically, UAV manufacturers produce various models and configurations to build and design UAVs. To overcome this issue in selecting features, this work selected features that exist in any UAV regardless of its model, components, and configuration.

For the UAV sensors, this work focused on selecting features based on the reliable data generated. Therefore, the selected feature was based on the non-fault sensor consistency. In addition, some factors that affect these data were considered, such as unchanged data or data-based statistics. However, these led to follow the variability and consistency of the data generated for the all-flight scenario. On the other hand, in this process, this research conducted a comprehensive analysis of the components of the UAV following the generality concept to ensure that the result of the study achieved features commonly deployed in UAVs.

This research also targets the model's applicability and adaptability, neglecting the hardware's characteristics. Hence, it has led to an increase in the ability to address multiple scenarios. Furthermore, control features were excluded because control inputs vary from one piece of equipment to another. Therefore, these features were not prac-

tical in terms of hardware compatibility. Hence, this work considered that most UAV manufacturers produce different types of UAV models and, from there, it was decided to select records related to UAV coordination and physical status during the mission. These features represent metrics related to the location and spatial position. Based on these metrics of these features, this can ensure that detection works well regardless of the type of UAV and its configuration.

This work also categorized the features from the dataset into:

- Received Signal Strength : RSS in 6G UAV communication is a critical parameter for measuring the signal radio power strength in the UAV receiver. It includes the expected noise present in the signal and the potential radio jamming signal. When the jamming attack is present in the communication link, the RSS might not provide the strength of the original signal where it presents the strength of the jamming signal. Hence, these signals are high and evaluate the receiver about the actual quality of the link.

$$\mathrm{RSS}_{\mathrm{total}} = P_t + G_t + G_r - L_{\mathrm{path}} + G_{\mathrm{b}} - L_{\mathrm{j}} - L_{\text{other factors}} \tag{5.1}$$

Where:

  - P_t represents the transmit power of the original signal.
  - G_t represents the measurement of the transmitting antenna.
  - G_r represents the UAV antenna's gain measurement.
  - L_{path} represents the path loss of the original signal.
  - G_{b} represents the beamforming gain to reduce jamming.
  - L_{j} represents the impact loss during the jamming attack.

- $L_{\text{other factors}}$ includes other factors such as fading, obstruction, etc.

- Signal-to-noise SNR: In the 6G UAV network, SNR is expected to measure the noise present in the signal compared to the actual signal strength. SNR may affect the data rate in the expected communications when the jamming attack is presented. Jamming attacks increase the SNR in the channel to challenge the distinction between the actual signal and noise.

  SNR in the presence of a jamming attack is given by the equation:

$$\text{SNR} = \frac{P_{\text{s}}}{P_{\text{n}} + P_{\text{j}}} \tag{5.2}$$

  Where:

  - P_{s} represents the power in the original received signal.
  - P_{n} represents the noise.
  - P_{j} represents power of the jamming signal attack.

- Packet Error Rate PER: low latency and ultra-reliability are expected in the 6G UAV network. PER represents the ratio of the signal received from the receiver compared to the sender. This ratio is expected to be high in the 6G network due to advanced features such as the error correction technique. The main goal of the jamming attack is to decrease this PER through interference to decrease correctly received data packets.

$$\text{PER} = \frac{N \times P_{\text{s}}}{N} \tag{5.3}$$

Where:

- N represents the number of the packets.
- P_s represents the number of packets retrieved successfully during jamming.

- Throughput: represents the data rate received successfully from the sender over the communication channel. The throughput is measured by calculating a bit per second.

$$\text{Throughput} = \frac{\text{Packet Size}}{\text{Transmission Duration}} \tag{5.4}$$

- UAV GPS Altitude: represents the height of the UAV from sea level. In the presence of the GPS spoofing attack, the hacker sends fake altitude data to the UAV. Therefore, the UAV moves to the wrong altitude data based on the fake signal received. Hence, it leads the UAV to fly on different levels, causing a crash or collision with obstacles or terrain or flying at a high level.

- UAV GPS latitude: represents the accurate geographic coordination of the UAV on the Earth's surface. In the presence of the spoofing attack, the UAV is driven in the wrong direction. Therefore, this leads the UAV to an incorrect flight path, which causes it to be moved from its target area.

- UAV GPS longitude: It represents the geographic point of the UAV on specific coordinates east-west. Mistaking the UAV's longitude leads to deviation and is a reason to lose the UAV's path and disrupt its mission.

- UAV velocity: represents the change of the UAV in the position with a corresponding time.

- UAV acceleration: represents the change in the velocity of the UAV considering the time.

5.6.1 Proposed methodology

The performance of 6G depends on the efficiency of the aerial network layer during signal transmission. The UAV in this layer can cover large scale areas to provide an entire coverage area based on the user's demands and the mission's requirements. However, the adversary's actions to disrupt the transmission are the main challenges, where they launch their malicious actions to violate the integrity and availability of the transmission. The transmission link between legitimate nodes in the aerial layer needs to be more secure and reliable during transmission. The legitimate node faces threats such as noise frequency signals and spoofed signals, which blocks communication and consume power resources in the receiver. Additionally, they aim to mislead the UAV trajectory by spoofed signals. This malicious action, when it happens, will target the transmission and lead to a degradation of the performance of the UAV system by injecting a false GPS signal to deviate the UAV from its planned path, leading to loss of coverage area until the system is recovered.

This section focus first on detecting abnormal behavior in network traffic, as shown in Figure 5.10. Two levels of algorithms are used in the model: one to detect an anomaly and the second to confirm anomaly detection in the data received. The detected anomaly is classified into two types: jammed signal or deviation from the planned route of the UAV due to a fake GPS signal. The main components in this model are presented below.

Feature Engineering and Scaling

In the designed approach, this work considered unjust data patterns and non-essential data, so feature engineering was used to address these obstacles. The model was designed to take advantage of the plotted data patterns based on common features in all types of

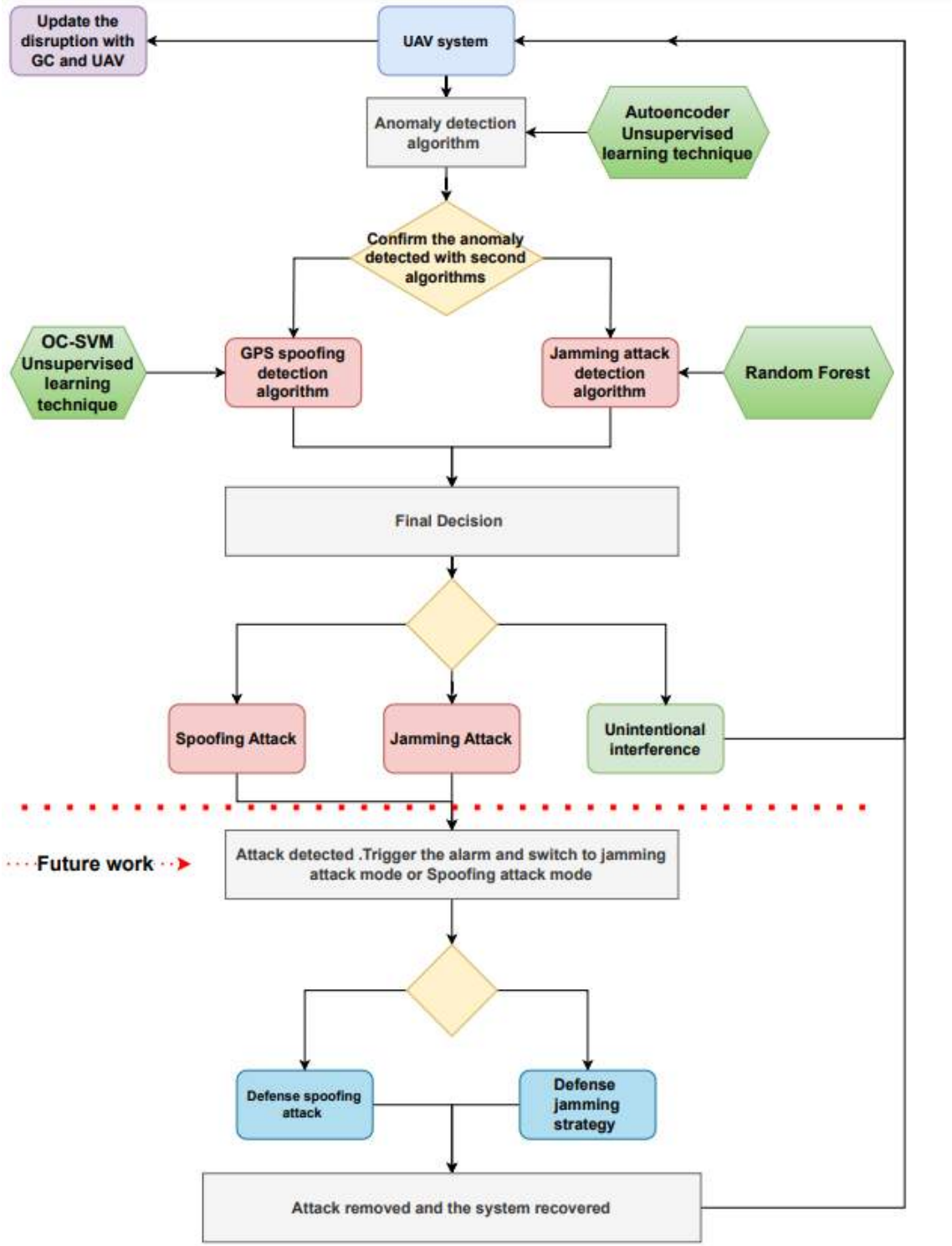

Figure 5.10: The architecture of distributed intrusion detection system

UAVs. Therefore, obtaining these characteristics was essential to building and deploying a neural network model to detect jamming and spoofing attacks. To recall data and obtain a unified range of values, this work used the min-max normalization equation to

adjust the characteristics between [0, 1].

$$X_{\text{norm}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \tag{5.5}$$

5.6.2 Splitting of the Dataset

Once the dataset processing was completed, the dataset was separated into training and testing data. The training data is used to train the designed model, while the testing dataset is used to evaluate the proposed model. Therefore, the proposed system separated data into multiple splits to start this process, and suitable splits were used. Once the data and data separation processes completed, the dataset was ready to detect a proposed anomaly. Therefore, the three algorithms, Autoencoder, k-means, and OC-SCM, were prepared to be employed in the model.

- Anomaly detection algorithm:

  The first algorithm in this model was proposed to use the autoencoder model. The autoencoder is a deep neural network that takes an input data vector and decodes it to reconstruct the input vector. The encoding process aims to optimize the parameters to introduce a vector representation that includes the characteristic. At the same time, the decoding part takes the produced vector through the reconstruction parameter. The autoencoder measures the difference between the input and reconstruction vectors produced by the loss function. This work used a stacked autoencoder in which various layers of neurons were stacked in the encoder and decoder processes. Therefore, each signal layer used the activation function ReLU for non-linearity and computes the linear operation. The models e() and

d() represent the encoder and decoder. In contrast, x represents the input vector to the encoding, and r is used for the representation vector in the decoder. W represents weight, and b represents biases in the linear operation. ReLU and the sigmoid function are present in the model architecture, and ReLU represents the nonlinear activation function.

The encoding phase was the first phase of training the autoencoder to detect anomalies. In this phase, the autoencoder trained on the normal dataset parameters. Therefore, various levels of reconstruction were produced between normal and abnormal data. This process allowed the algorithm to learn normal data parameters. Thus, the parameter of the autoencoder was handled with standard data so that these features were optimized for encoding and decoding the benign data. This means that the autoencoder produced a low reconstruction loss of the normal data set and a high reconstruction loss representing the abnormal data. Therefore, this work addressed and analyzed these results to identify normal or abnormal data. Thus, the only log used in the encoding phase was the normal data status. Furthermore, the validation phase identified the threshold by selecting a particular threshold level. The threshold allowed the model to determine an anomaly when the reconstruction loss of the input feature vector was more significant than the threshold. Therefore, choosing the threshold of the trained model should recognize abnormal data effectively, where the model performance depends entirely on selecting the appropriate threshold.

In the model, the autoencoder was trained as the first algorithm on some features to detect anomalies during data transmission. It used multiple features from the dataset to train this algorithm, such as signal noise SNR, receive signal strength RSS, bit error rate (BER), packet error rate (PER), and GPS coordinates. These

features were trained in the first algorithm to detect an abnormality in the signal and data received and then identify the initial anomaly detected. Finally, the initial result of the first algorithm will be sent to the second algorithm to increase the accuracy of the initial detection by confirming the suspicious result.

- Jamming attack detection model:

  K-means clustering is an unsupervised clustering algorithm that aims to divide the data point into clustering containing unlabeled data, such as data not included in the groups. It can handle the enormous size of the dataset [79]. The K-means technique's primary goal is to find a cluster in the data, including a variable named K. The clustering is executed based on the similarity of the data by reducing the distance between the data point and the assigned cluster centroid point X in the model [80]. In abnormality detection, K-means can identify the abnormality through the deviation in the data point that is differentiated from the point in the train clustering, such as increasing the distance between the data point and the assigned cluster centroid point X.

  In this model, an algorithm that includes the K-mean integrated with an autoencoder was created. This integration extracted deviation in the signal pattern presented in the input data used during the training model. We used K-means and autoencoders in the model because K-means know noise patterns and variables that describe complex relationships. Hence, this model's architecture enhanced the accuracy of an autoencoder in the overall model performance. Based on the result received from the first algorithm, k-mean executed the received log again to confirm the suspicious case. If the model confirms a jamming attack, the model sends the decision to the final decision model for the next step. Therefore, the anomaly will be categorized in tow cases:

- Unintentional interference: The second algorithm confirmed that the anomaly was detected as unintentional interference due to channel congestion, so the result will be sent to the UAV system for updating.

- Jamming Attack: This means that the previous first and second algorithms confirmed the jamming attack, which will be classified according to the type of jamming attack that occurred.

- GPS spoofing detection model:

  The OC-SVM one-class support vector machine is widely used in the detection technique to detect the attack and classify it according to the trained model [81]. It is similar to SVM; it is considered a semi-supervised method [82]. OC-SVM was trained only on normal data and classified the new data, which differs from the training data, as abnormal. This model receives the result of the first algorithm related to the probability of a GPS spoofing attack. OC-SVM algorithm executed the log to confirm the decision. If the model confirms the spoofing attack, it will send the result to the final decision model for the next step.

  In this work, OC-SVM was trained on normal data to learn the boundary presented in the normal dataset. Then, it received the number of anomalies detected as a spoofing attack from the autoencoder to confirm them. The OC-SVM creates an auto-function with gamma parameters. This function automatically adjusts the model to the gamma value, affecting the SVM's decision boundary. After that, it predicts the data related to the abnormal or normal data by using -1 as a label for the anomaly and 1 as a normal label. Hence, the confirmation in the algorithm was designed with an autoencoder. This algorithm confirmed that the autoencoder detected data point was abnormal and related to a spoofing attack. It is identified

as an anomaly if both methods consider the detected data point an anomaly.

5.7 Performance Metric and Evaluation result

This section applied the evaluation to the proposed methodology using the receiver operating characteristic curve. ROC curve is used to measure the accuracy of the proposed technique with true positive rate TPR and false negative rate FNR. The proposed system was mapped as a classification problem with normal or abnormal classification. The following two algorithms trained to identify anomalies, such as jamming or spoofing attacks. To prove the model's effectivity, this work aimed to increase TPR and decrease FPR, which were measured by using the aerea under the ROC curv AUC [83]. The main reason for using AUC is that it can recognize normal and abnormal classes. The higher value of the AUC leads to better performance.

The confusion matrix is performed to obtain prominent information about the predicted output. It is also widely called the error matrix and is used to show the prediction of the output of the classification model during the test and validation data. The classification result, either correct or incorrect classes, is displayed in the Table 5.12. The table shows the confusion matrix of intrusion detection. The main goal of the confusion matrices is to present the performance of the ML algorithms.

- : True Positive (TP): it measures the capability of the model in classifying normal

Confusion Matrix		
	Positive Class Predicted	**Negative Class Predicted**
Normal Class	TP	FN
Attack Class	FP	TN

Table 5.12: Confusion matrix in binary class

action correctly as:

$$\text{TP Rate} = \frac{TP}{FN + TP} \times 100\% \tag{5.6}$$

- : True Negative (TN): it measures the capability of the model in classifying abnormal action correctly:

$$\text{TN Rate} = \frac{TP}{TN + FP} \times 100\% \tag{5.7}$$

- : False Positive (FP): it measures the error of the model in classifying normal action correctly as:

$$\text{FP Rate} = \frac{FP}{TN + FP} \times 100\% \tag{5.8}$$

- : Flase Negative (FN): It measures the error of the model of classifying abnormal action correctly as:

$$\text{FN Rate} = \frac{FN}{TP + FN} \times 100\% \tag{5.9}$$

- :Accuracy: In the shown equations, accuracy is the ratio of the proposed methodology able to correctly detect anomalies among the records used in the dataset. The best performance, is the highest accuracy achieved where accuracy ranges from 0 to 1.The main goals of the accuracy is make balances of the data collected and used.

$$\text{Accuracy} = \frac{\text{TP + TN}}{\text{Total Population}} \tag{5.10}$$

- :F1-score: It is called F1-measures and is an effective performance measure used to identify the relation between recall measures and harmonic mean of the precision. The highest number of F1 score means that the methods used were effective.

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5.11}$$

- :Precision: It represents the correct prediction from the overall prediction event.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{5.12}$$

- :Recall: It represents the proportion of the positive rate.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{5.13}$$

6

Experimental Setup and Analysis of Results

6.1 Experimental Setup

In this section, the experiment is addressed and discussed. To complete the evaluation of the experiment, Python was used as a programming language. The decision to use Python in the experiment was made for multiple reasons. It is the easiest and has been used widely as a programming language in numerous application areas. It has various libraries used in the multiappearance area, such as web server tools, operating systems, internet protocol, and string operation. One of the main open-source libraries used in this work was Scikit Learn, which performs and assembles multiple algorithms for

clustering and classification. Also, some libraries used for dataset processing, including Sklearn, Numpay, and Panda. In addition, model n-fold cross-validation and train test split were used for evaluation.

The primary goal of the experiment was to evaluate whether the proposed methodology adequately identified the anomaly in the autoencoder algorithm and confirmed it using the following two algorithms. During the experiment, the focus was on two activities: first, the validation of the proposed model found an extensive reconstruction when the abnormal dataset was used. Second, it checked the difference between normal and abnormal data in reconstruction loss. The following subsections show how the configuration of the experiment and the result of the suggested method are effective.

6.1.1 Setup

The three logs, normal flight, jamming, and spoofing attacks from the dataset, were used. The suggested algorithms were built and designed based on an unsupervised learning technique; therefore, they were trained only to observe the features when the flight is in normal data and validated in the dataset, including both normal and abnormal behavior of events.

6.1.2 Experiment Results

For the IDS, confusion matrices such as FN and FP play a crucial role in performance evaluation, recall, and precision. Reducing FN and FP in IDS is necessary, as their increase in the model leads to classifying the attack as a legitimate event. The conclusion of the precedence showed that a low result means that the false positive rate is high. The low level of the recall parameter indicated that the model classifies the attack event

in the network as normal, which leads to an increase in the number of FN. Furthermore, increasing F1 results leads to less incorrect classification, such as correctly classifying normal events as normal and abnormal data as attack events. In addition, accuracy indicates the correct classification of normal and abnormal data.

Autoencoder

The linear autoencoder trained on a training test that contains normal observation. The data pattern must be fitted before model training; therefore, several techniques have been applied to fit the normal data pattern. First, batch normalization was applied to execute the encoder and decoder. Second, to avoid an overfitting problem, both the L1 and L2 regularizers were applied. The Adam optimizer used to optimize the parameter to increase the model's efficiency. Once the model was fully trained, the test was executed. Therefore, the two experiment result showed in the two forms jamming and spoofing attacks.

The red part of the line in the Figure 6.1 shows how the construction loss increased when the attack data was started. From the figure, we can extract that the reconstruction loss increased extensively when the malicious action of the UAV launched a jamming or spoofing attack scenario. Figure 6.1 shows how the reconstruction loss changed when the hacker started its malicious plan. These significant behavior imply changes in the pattern. Therefore, the first detection algorithm extracted and learned the dynamic of the normal data pattern and can identify the malicious pattern in the hacker's presence.

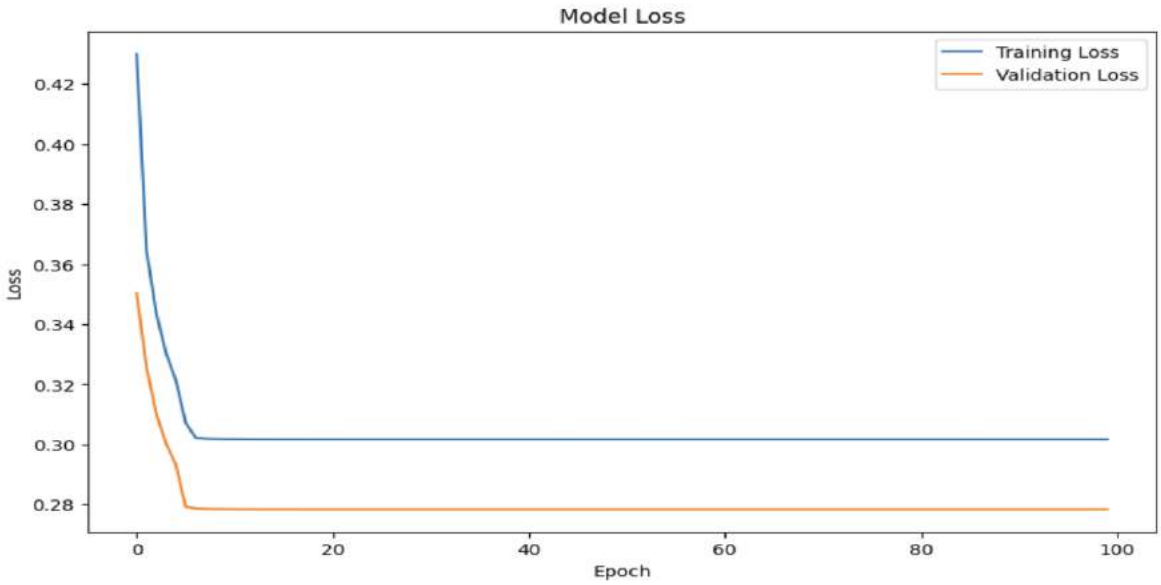

Figure 6.1: Reconstruction loss

Attack Type	Precision	Recall	F1 Score	Accuracy
Jamming	$\frac{25844}{25844+350} \approx 0.986$	$\frac{25844}{25844+1364} \approx 0.949$	$\frac{2\times(0.986\times0.949)}{0.986+0.946} \approx 0.962$	$\frac{25844+2286}{29844} \approx 0.942$
Spoofing	$\frac{34621}{34621+1853} \approx 0.949$	$\frac{34621}{34621+2348} \approx 0.935$	$\frac{2\times(0.949\times0.935)}{0.949+0.935} \approx 0.941$	$\frac{34621+5212}{41428} \approx 0.961$

Table 6.1: Metrics performance public dataset

Autoencoder Performance on UAV dataset

The first experiment was performed to identify the accuracy of the first algorithm and showed that the autoencoder achieved good accuracy. This algorithm was investigated and evaluated by the public, and the simulated dataset included all features such as RSS, SNR, PDR, and throughput, alt, lat, long, acceleration, and velocity of normal and abnormal parameters. The results presented in Tables 6.1 and 6.2 show that the autoencoder did not change significantly in different datasets. Therefore, the accuracy was between 94 and 98.6, and the FNR was between 0.06 and 0.05, so it shows good performance in correctly detecting abnormalities in the communication link and UAV movements.

Attack Type	Precision	Recall	F1 Score	Accuracy
Jamming	$\frac{23674}{23674+1594} \approx 0.940$	$\frac{23674}{23674+1425} \approx 0.946$	$\frac{2\times(0.940\times0.946)}{0.940+0.946} \approx 0.942$	$\frac{23674+4891}{31584} \approx 0.904$
Spoofing	$\frac{32795}{32795+1348} \approx 0.959$	$\frac{32584}{32584+1148} \approx 0.965$	$\frac{2\times(0.959\times0.964)}{0.959+0.964} \approx 0.961$	$\frac{32584+6348}{41428} \approx 0.955$

Table 6.2: Metrics performance SIM-Dataset

Attack Type	TP	FN	FNR	Accuracy
Spoofing	34621	2348	2348/(2348+34620=0.06)	34621+1853/44034=0.986
Jamming	25844	1364	1364/(1364+25844=0.0501	25844+2286/29844=0.94

Table 6.3: Autoencoder accuracy

Table 6.3 shows that the autoencoder performed on two datasets separately, including spoofing and jamming attack parameters. During the spoofing attack, the first algorithms reached a high accuracy of 0.98. Additionally, the FNR was very low, rising by 0.06, and classifies 34,621 as a normal pattern. However, during the jamming attack, the autencoder was trained on various features such as SNR, PDR, RSS, and Throughput, the autencoder detected changes in the signal parameters during transmission. As shown in Table 6.3 the autoencoder worked well, reaching 94 with the lowest FNR rate, which was 0.05. Furthermore, this algorithm classified the 25,844 events as normal behaviors.

OC-SVM with Autoencoder

Combining the two algorithms, autoencoder, and oc-svm, to detect anomalies such as spoofing attacks proved to be effective. The main characteristics of the autoencoder are the ability to handle complexity in the data relationship and deal with linear patterns that may not be captured by OC-SVM, where it works effectively to identify the boundary presented in the normal data. Therefore, combining both algorithms increased the accuracy in determining the real anomaly and reduced the FPR.

In the second experiment, an evaluation was performed to obtain the autoencoder detection result and confirm the detection of abnormalities by oc-svm. These two algorithms are also trained only on features related to the spoofing attack, such as alt, lat, long, velocity, and acceleration. The validated datasets used in this algorithm included all normal and abnormal parameters related to the features. These algorithms showed that oc-svm shows a significant result with the autoencoder detection results as shown in Figure 6.2. The accuracy of these algorithms reached 96, and the FNR 0.08.

Tables 6.4, 6.5, and 6.6 show that accuracy, F1 score, recall, and precision detection in two algorithms, combined and separately, confirmed the spoof attack. When the spoofing attack occurred, OC-SVM enhanced the transmission and the accuracy detection ratio to 94. Also, the FNR was low at 0.08.

Attack Type	Precision	Recall	F1 Score	Accuracy
Spoofing	$\frac{31748}{31748+2374} \approx 0.93$	$\frac{31748}{31748+2943} \approx 0.91$	$\frac{2\times(0.93\times0.91)}{0.93+0.91} \approx 0.919$	$\frac{34748+6895}{44034} \approx 0.94$

Table 6.4: Metrics performance on public dataset

Attack Type	Precision	Recall	F1 Score	Accuracy
Spoofing	$\frac{21521}{21521+2045} \approx 0.91$	$\frac{21521}{21521+1396} \approx 0.939$	$\frac{2\times(0.91\times0.939)}{0.91+0.939} \approx 0.924$	$\frac{21521+4495}{29458} \approx 0.88$

Table 6.5: Metrics performance on SIM-Dataset

Attack Type	Precision	Recall	F1 Score	Accuracy
Algorithm-combined	$\frac{69369}{69369+3722} \approx 0.94$	$\frac{69369}{69369+4091} \approx 0.94$	$\frac{2\times(0.94\times0.94)}{0.94+0.94} \approx 0.919$	$\frac{82612+13243}{90625} \approx 0.91$

Table 6.6: Metrics performance combined

In the first detection process, the autoencoder detected 31,621 anomalies in the deviation of GPS coordination, as shown in Figure 6.3 and Table 6.7. This number of anomalies detected shows data points that deviate from the trained data of the model.

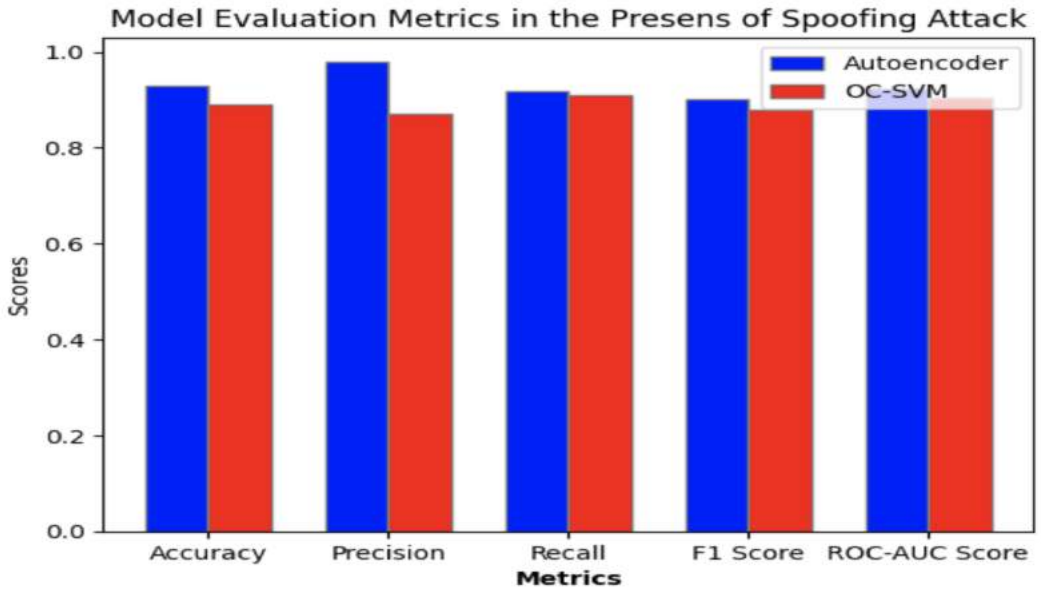

Figure 6.2: Metrics performance oc-svm and autoencoder

Table 6.7: Spoofing attack ROC curve

Attack Type	**Number of Normal Logs**	**Number of Attack Logs**	**AUC**
Spoofing Attack	45835	13524	0.9037

Hence, the non-linearity pattern and ability to capture complexity in the normal data point played a crucial role in detecting the deviation from the normal data pattern by using the OC-SVM. This result indicated that this number is considered out of the ordinary where the main work of the oc-svm, any point presented out of the encapsulated normal data, is an anomaly. Finally, the confirmed number of the anomaly was 31,748, close to the number of anomalies detected by the autoencoder. Therefore, this result represents the anomalies detected by autoencoder and oc-svm algorithms.

K-means with Autoencoder

Using their unique characteristics, the combination of K-means and Autoencoder provided a good detection technique. The autoencoder was designed to learn benign data

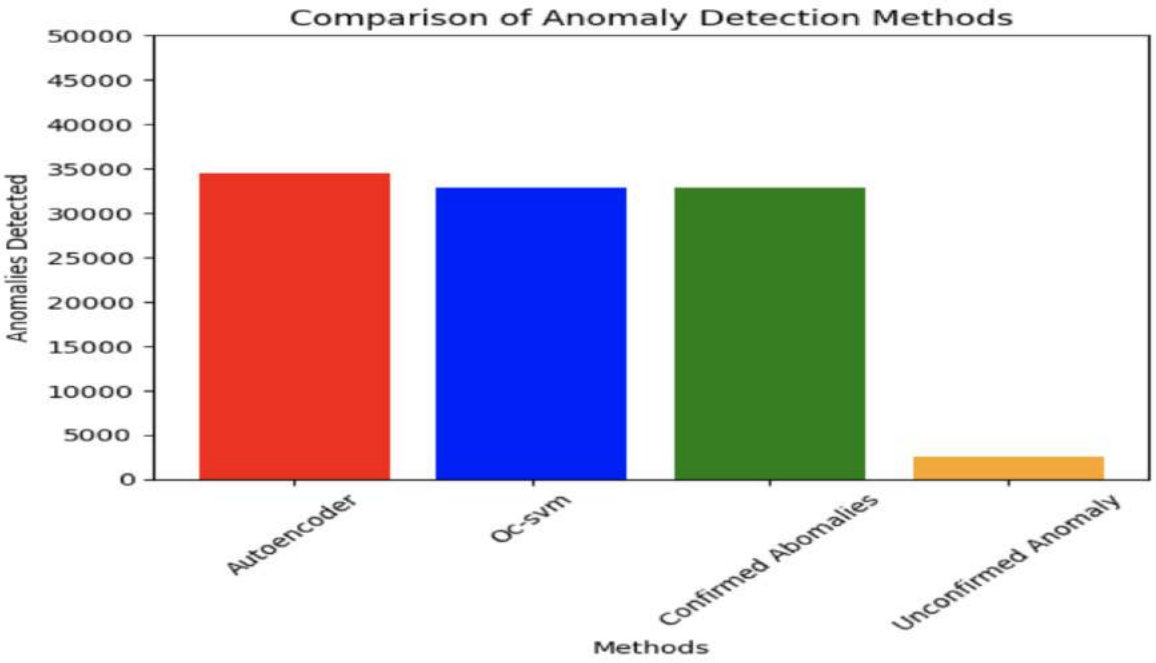

Figure 6.3: K-means with autoencoder

codings in an unsupervised learning technique, so it identified 34,583. This detection relies on the difference between reconstruction errors in the data point. Therefore, the number of anomalies detected is due to the increase in these abnormal data reconstruction errors. Following the autoencoder, K-means received the detected anomalies and analyzed them according to the deviation of the jamming attack data pattern.

In the first session, to evaluate the model in detecting jamming attacks and distinguishing them from unintentional noise, K-means received the anomaly detected by the autoencoder to confirm the deviation in the signal parameters. This algorithm was evaluated based on two datasets: a public dataset and a simulated dataset with 6G signal parameters. Different features were considered in these algorithms, such as SNR, RSS, PDR, and through. The results showed that these algorithms controlled the abnormality event detected by the autencoder. Therefore, as shown in Figure 6.4, the accuracy of these algorithms reached 0.92 and the overall accuracy results were between 91 and 96. Also, it showed that the FNR was low and between 0.03 and 0.11.

The below Tables 6.8 - 6.11 showed the percent of accuracy achieved in these algorithms in k-means and overall of the two algorithms. Under the jamming attack, these algorithms evaluated the transmission parameters and detected and confirmed the detection accuracy at 0.92. In addition, this algorithm correctly classified 25,583 as a normal event. Therefore, based on these results, these algorithms worked effectively.

The idea behind using two algorithms was to increase accuracy and reduce FNR, which these algorithms achieved. The evaluation of combining these two algorithms was performed, and it showed a good result by confirming all detected events in the first algorithms; hence, it worked efficiently with high accuracy and low FNR, as shown in Figure 6.5 and, Tables 6.11 and 6.12.

The overall evaluation of the model, as shown in Figures 6.6 and 6.7, and Table 6.13, achieved high accuracy detection alongside precision and other matrices. Hence, this high percentage proved efficient for the model in the presence of malicious action jamming and spoofing attacks.

Attack Type	Precision	Recall	F1 Score	Accuracy
Jamming	$\frac{25844}{25844+1284} \approx 0.95$	$\frac{25844}{25844+2184} \approx 0.92$	$\frac{2\times(0.95\times0.92)}{0.95+0.92} \approx 0.93$	$\frac{2358+3134}{28845} \approx 0.92$

Table 6.8: Metrics performance K-mean

Attack Type	Precision	Recall	F1 Score	Accuracy
Jamming	$\frac{14321}{14321+937} \approx 0.938$	$\frac{14321}{14321+336} \approx 0.977$	$\frac{2\times(0.938\times0.977)}{0.938+0.977} \approx 0.93$	$\frac{14321+2541}{18322} \approx 0.957$

Table 6.9: Metrics performance K-mean

Attack Type	Precision	Recall	F1 Score	Accuracy
Algorithm-combined	$\frac{37995}{37995+2531} \approx 0.937$	$\frac{37995}{37995+1761} \approx 0.955$	$\frac{2\times(0.937\times0.955)}{0.937+0.955} \approx 0.945$	$\frac{45427}{49619} \approx 0.915$

Table 6.10: Metrics performance combined

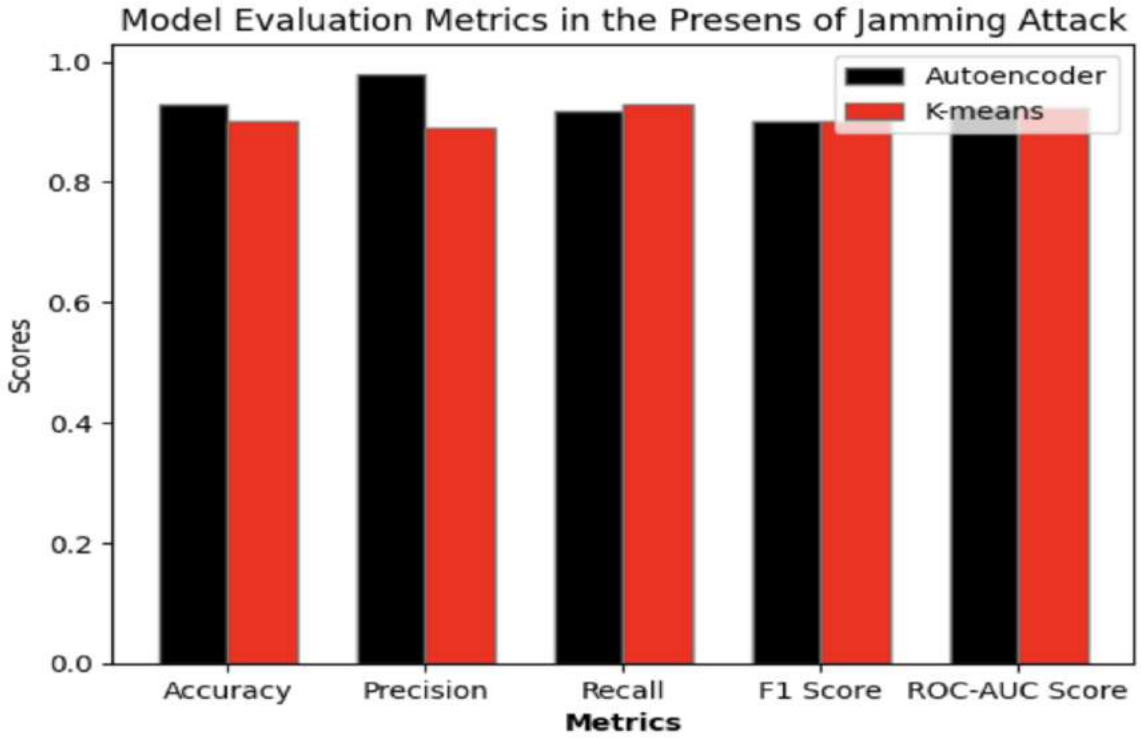

Figure 6.4: Overall evaluation

Algorithms	Accuracy	FNR	FN	TP
Autoencoder	0.98	0.06	2348	34621
K-means	2358+3134/28845=0.92	1394/(1394+23844)=0.05	1394	25844
Algorithms-combined	0.95	0.11	4722	60465

Table 6.11: Metrics performance K-means

6.2 Discussion

Recently, with the significant development of artificial intelligence, security mechanisms have become modern, and adversary techniques have constantly evolved. Unlike previous network technology, 6G is expected to have a high volume of data, resulting in noise in the channel, similar to a jamming attack. Therefore, robust and effective security techniques are essential, so designing efficient anomaly detection mechanisms considering these characteristics is required. Owing to supporting the 6G network to Non-Terrestrial Network (NTN) to deploy the UAV to enhance terrestrial network in-

Table 6.12: Attack confirmed

Attack Type	Number of Normal Logs	Number of Attack Logs	AUC
Jamming Attack	25437	5836	0.9168

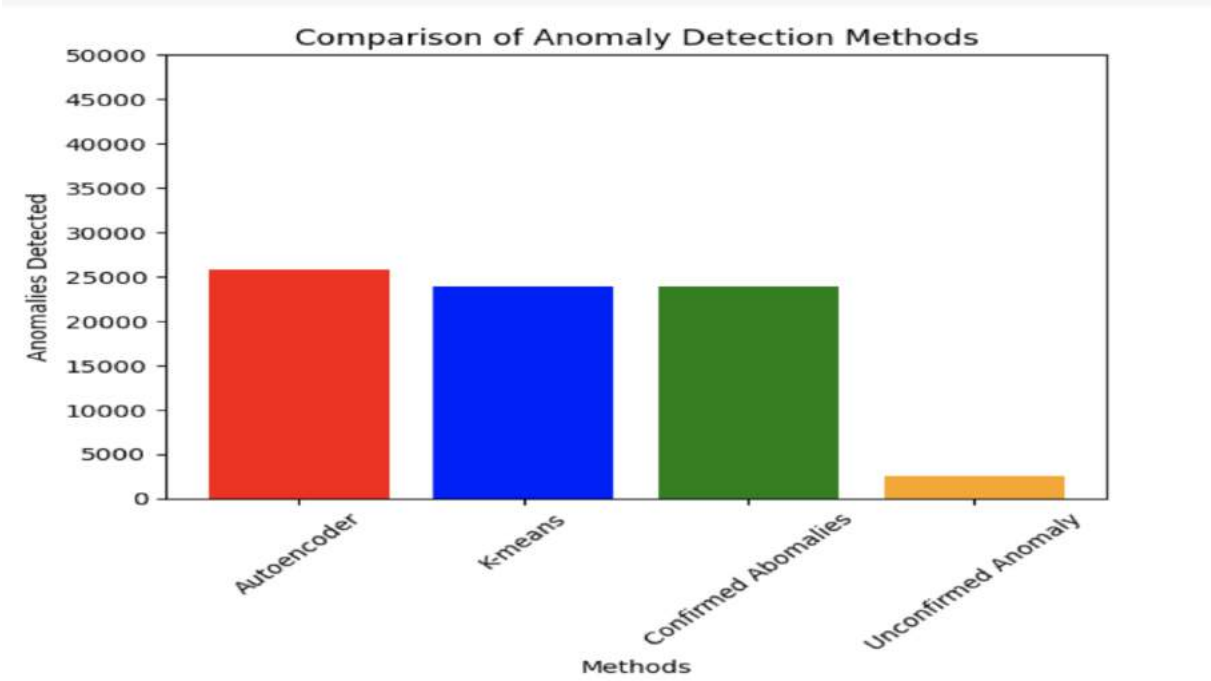

Figure 6.5: K-meanse with Autoencoder

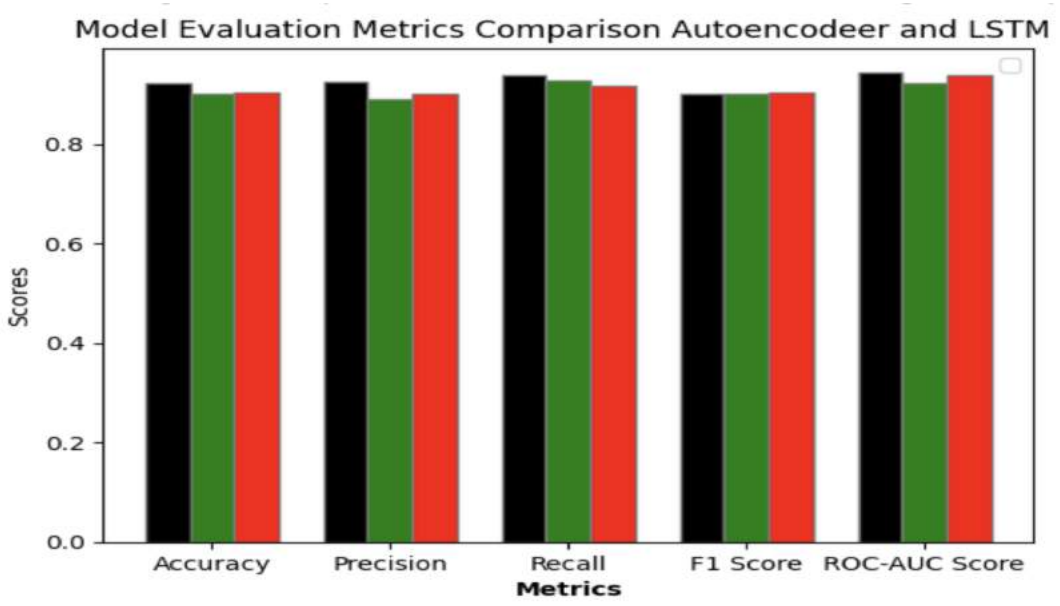

Figure 6.6: Overall evaluation

Algorithms	Accuracy	FNR	FN	TP
Autoencoder	0.98	0.06	2348	34621
Oc-svm	0.94	0.08	2374	31748
K-means	0.92	0.05	1394	25844
Algorithms-combined	0.94	0.06	6116	92213

Table 6.13: Overall evaluation

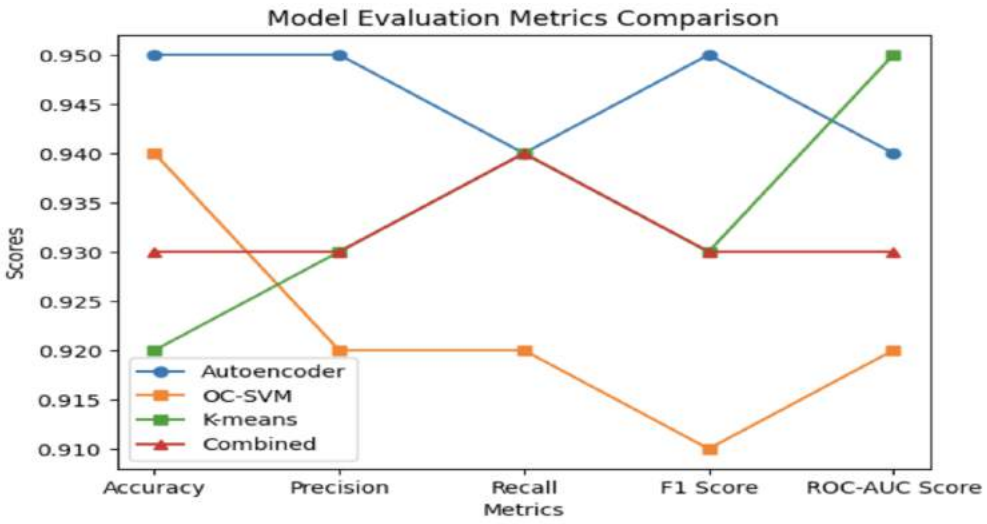

Figure 6.7: Overall evaluation

frastructure for critical missions, this work designed a new security mechanism as two layers of algorithms to enhance security in the UAS. This technique improves system security by monitoring network traffic and UAV behavior to identify changes in the signal communication parameters, geospatial coordinates, and dynamic behavior of the UAV. The first algorithms used in the unsupervised technique were trained on normal data to detect anomalies during the mission; therefore, this model plays a crucial role in improving the accuracy of the suggested technique by giving initial detection decisions. In the second layer, two algorithms are designed, and each was trained separately to confirm the anomaly detected received from the first layer algorithm and identify the deviation as either a jamming attack or a spoofing attack. Therefore, the model improves performance in the second layer by reducing false alarms that cause mission destruction, as shown in Figure 6.8 and Table 6.13.

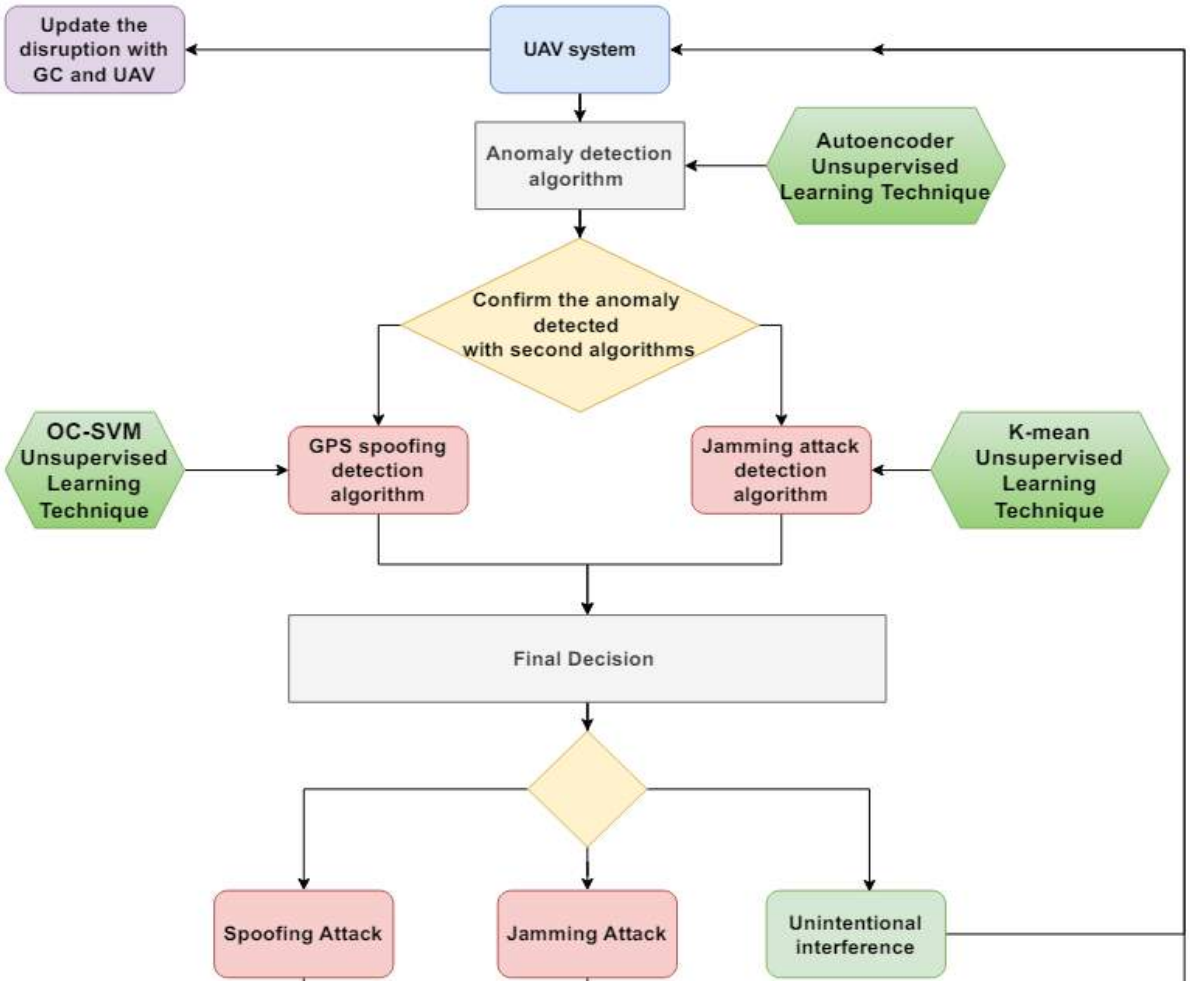

Figure 6.8: Anomaly detection methodology

The 6G network is expected to handle high data rates because of the significantly increasing number of connected devices. Hence, it leads to presenting unintentional noise signals because of the congestion at the channel. The two-layer algorithms in this work provide the ability to distinguish jamming attacks and congestion in the channel represented in a 6G UAV environment, starting from the autoencoder to detect anomaly detection on the network-related signal parameters features to the second algorithm k-means to confirm anomaly as a jamming attack. Therefore, these two algorithms exploited signal parameters such as RSS, SNR, PDR, and throughput to identify the jamming attack from noise caused by congestion in the channel. Hence, the approach can identify maliciously intentional jamming accurately, providing reliability in the com-

munication.

Additionally, this work handles the main security issue of securing UAVs from being hijacked or disrupted. GPS spoofing attack is an obstacle facing UAVs during missions, which leads to interrupting the connectivity in the coverage area when it is used to enhance the terrestrial network. These algorithms are designed efficiently to detect deviations in the UAV GPS coordination such as alt, lat, long, and dynamic behavior such as velocity and acceleration. Therefore, these algorithms enhance security when integrated onboard UAV systems within 6G networks. In addition, this method ensures stable connections to devices connected to UAVs where the interruption affects the safety and quality of service provided to users. An interruption or deviation in the UAV is expected during the launching of the jamming and spoofing attack. Therefore, this technique is expected to enhance defense mechanisms such as changing the channels or flying away from malicious areas.

- 6G UAV system layer

  The 6G UAV network is expected to handle large amounts of data, so the algorithm in the first layer is essential as a primary detection algorithm to analyze the traffic data and identify the deviation patterns. Therefore, it ensures that the suspicious event is detected and identified correctly. Combining this layer with the second layer, which includes the OC-SVM and K-means, decreases the false positive rate.

  F1(z) is the algorithm used in the first layer at time z. R(z) represents traffic data in the transmission. T(z) is the feature of abnormality. FP(z) is the false positive rate.

$$F_1(z) = F_N(R(z), T(z))$$

The primary goal of F1(z) is to recognize and identify the changes in the normal transmission pattern to guarantee that abnormality is detected. Also, the reduction of the false positive rate is represented by combining both algorithm layers as

$$F(z) = F(F_1(z), F_2(z))$$

- It meets the adaptability in the dynamic environment of the 6G UAV networks. Label data collection is a critical process in ML training, but it is limited and expensive; therefore, an unsupervised learning technique is preferred. This model can be effectively updated with unseen data will address the issues of UAVs operating on 6G network. Furthermore, the unsupervised technique can detect a new type of attack.

  The adaptability of the suggested approach represented in $n(t)$ included the algorithms in the time duration T and $n(t)$.

  The state of the approach is represented in this equation as

$$A(t + \Delta t) = f(A(t), V(t), A_d(t), A_c(t))$$

  where is $A(t + \Delta t)$ the time that was selected to update the system at this time. f is the process function of the system $A(t)$, and $V(t)$ is the threat vector.

- leveraging 6G resources;

  Two features are expected in 6G, higher data rate volume and faster speeds. Therefore, by handling the anomaly detected in two levels of algorithms, computational resource issues are mitigated; hence, the system response is adequate.

- It is suited for UAV network security issues.

The main target of the jamming attack is to disrupt the communication between UAV and satellite or GC: so the RF is used to recognizes the sequential patterns. In addition, the main goal of the spoofing attack is to mislead the UAV from its location by manipulating the navigation system. Therefore, OC-SVM can effectively identify the deviation in the UAV path.

- It meets module adaptability

  The evolution of the threat vectors of the 6G UAV network is adaptable, with the suggested model able to add new algorithms to enhance the system's longevity.

  In this model, adaptability is represented as the ability to adjust the model based on the threat model $T(z)$ and expected vulnerabilities in the system described as $V(z)$ at the time. Therefore, it allows new algorithms $G(z)$ to be added to the model to enhance longevity and support the system's resilience to face emerging hackers. Therefore, the adaptability is represented as

$$D(t + \Delta t) = f(T(t), V(t), A(t))$$

- Ensures reliability in UAV operation.

  The system's two layers provide a robust validation process. This process guarantees the integrity of the mission, the safety of the UAV operation, and the reliability of the received data.

  Reliability in UAV performance is represented as $S(t)$ to ensure high safety and trustworthiness in the UAV mission. $G1(z)$ is the first algorithm to detect anomalies for validation. $G2(z)$ is the second algorithm used to ensure the integrity and safety of the mission in the second layer. The reliability in the two algorithms is achieved in this form

$$S(t) = f(G_1(z), G_2(z))$$

$f()$ represents reliability achieved through two detection and confirmation algorithms.

- Scalability in the 6G UAV network scenarios

  UAVs can be deployed in swarms and singles based on the mission plan. Jamming and spoofing have multiple impacts on the UAV mission, so the two layer algorithm detection techniques are comprehensive and meet scalability to address changes and various threats.

  It is considered that U is the group of several UAVs forming a swarm. $J(s)$ is the jamming attack at time T, and (t) represents the spoofing attack action during the mission. A1(z) is the algorithm at the first level to detect anomalies, and A2(z) is the confirmation algorithm at the second level. The swarm that forms in

$$S(t) = f(A_1(z), A_2(z), J(t), T(t), U)$$

  where $f()$ is the function to represent the swarm, with the UAV number and the presence of the jamming and spoofing attack, and $A1$ and $A2$ are the two algorithms' detection. Therefore, scalability is achieved by keeping the swarm's performance stable and ensuring the detection of any changes in the overall swarm performance.

7

Future Research Directions and Conclusions

7.1 Future Research

The two layer algorithm detection technique in the 6G UAV network provides a significant step forward. This technique provides the ability to be extended to achieve new research goals and adapt to meet the rapid evolution of the upcoming technology. This technique is expected to be used in the future for security protection as follows.

Detecting anomalies is practical in critical systems. Therefore, measuring the severity of the detected anomaly increases the algorithms' efficiency. Using specific techniques by

incorporating them with the primary algorithm is effective. The algorithm can determine the potential impact and the severity level, either high or low. The statistical method is an efficient technique to assess the severity of threats. It can monitor the magnitude of the normal pattern and decide whether the severity is high or low. A security breach is another technique to determine the impact of severity by monitoring the fluctuation that could result from the low severity issue. In addition, temporal analysis can be included by using the time duration anomaly and frequency as metrics. Given the proposed technique, the primary algorithm autoencoder can be integrated with this technique to enhance an effective response strategy.

The K-means algorithm can be used to classify the jamming detected into more event types. This could improve countermeasures based on the classification of the detected anomaly.

7.2 Conclusions

The proposed methodology is composed of two-level algorithmic approach, including autoencoder as the first algorithm to detect the anomaly and the secound layer algorithms, OC-SVM to confirm the anomaly as a spoofing attack, and the K-means to confirm the anomaly as a jamming attack. This could provide significant advancement to enhance security in the UAV 6G network. The first algorithm, the autoencoder, is used as a filtration method to analyze the network data and convert it to be used to extract deviation in the data pattern. The second layer includes the OC-SVM and K-means algorithms used as detection system to validate the identified anomalies and categorize them as jamming or spoofing attack, respectively.

Using the unsupervised learning technique allows deployment even if the labeled data does not exist, making the suggested method scalable and practical. OC-SVM provides the ability to distinguish between input data that is normal and abnormal, especially in attempts of the spoofing attack to deviate the UAV from its planned path. On the other hand, the LSTM identifies the temporal and spatial patterns related to the jamming attack in the stream data.

In conclusion, the approach suggested for designing two-level algorithms represents an effective front-line defense methodology for the 6G UAV network. As the technology matures, this proposed approach can be key to enhancing UAV communication and assurance.

References

1. Gawas, A. U. An overview on evolution of mobile wireless communication networks: 1G-6G. *International Journal on Recent and Innovation Trends in Computing and Communication* **3,** 3130–3133 (2015).
2. Santhi, K., Srivastava, V., SenthilKumaran, G. & Butare, A. *Goals of true broad band's wireless next wave (4G-5G)* in *2003 IEEE 58th Vehicular Technology Conference. VTC 2003-Fall (IEEE Cat. No. 03CH37484)* **4** (2003), 2317–2321.
3. Wang, M. *et al.* Security and privacy in 6G networks: New areas and new challenges. *Digital Communications and Networks* **6,** 281–291 (2020).
4. Halonen, T., Romero, J. & Melero, J. *GSM, GPRS and EDGE performance: evolution towards 3G/UMTS* (John Wiley & Sons, 2004).
5. Mshvidobadze, T. *Evolution mobile wireless communication and LTE networks* in *2012 6th International Conference on Application of Information and Communication Technologies (AICT)* (2012), 1–7. doi:10.1109/ICAICT.2012.6398495.
6. Chen, H. *et al.* Ultra-Reliable Low Latency Cellular Networks: Use Cases, Challenges and Approaches. *IEEE Communications Magazine* **56,** 119–125. doi:10.1109/MCOM.2018.1701178 (2018).
7. Lauridsen, M., Gimenez, L. C., Rodriguez, I., Sorensen, T. B. & Mogensen, P. From LTE to 5G for Connected Mobility. *IEEE Communications Magazine* **55,** 156–162. doi:10.1109/MCOM.2017.1600778CM (2017).
8. Yang, T., Zhao, J., Hong, T., Chen, W. & Fu, X. *Automatic Identification Technology of Rotor UAVs Based on 5G Network Architecture* in *2018 IEEE International Conference on Networking, Architecture and Storage (NAS)* (2018), 1–9. doi:10.1109/NAS.2018.8515719.
9. Wang, Y., Zhang, Z., Zhang, P., Ma, Z. & Liu, G. *A new cloud-based network framework for 5G massive Internet of Things connections* in *2017 IEEE 17th International Conference on Communication Technology (ICCT)* (2017), 412–416. doi:10.1109/ICCT.2017.8359672.
10. Chowdhury, M. Z., Shahjalal, M., Ahmed, S. & Jang, Y. M. 6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions. *IEEE Open Journal of the Communications Society* **1,** 957–975 (2020).
11. Guo, H., Li, J., Liu, J., Tian, N. & Kato, N. A survey on space-air-ground-sea integrated network security in 6G. *IEEE Communications Surveys & Tutorials* **24,** 53–87 (2021).
12. HrISToV, G. V., ZAHArIEV, P. Z. & BELoEV, I. H. A review of the characteristics of modern unmanned aerial vehicles. *Acta technologica agriculturae* **19,** 33–38 (2016).
13. Basan, E. *et al.* GPS-spoofing attack detection technology for UAVs based on Kullback–Leibler divergence. *Drones* **6,** 8 (2021).
14. Liu, J., Shi, Y., Fadlullah, Z. M. & Kato, N. Space-air-ground integrated network: A survey. *IEEE Communications Surveys & Tutorials* **20,** 2714–2741 (2018).
15. Mozaffari, M., Lin, X. & Hayes, S. Toward 6G with connected sky: UAVs and beyond. *IEEE Communications Magazine* **59,** 74–80 (2021).

16. Xiao, Z., Dong, H., Bai, L., Wu, D. O. & Xia, X.-G. Unmanned aerial vehicle base station (UAV-BS) deployment with millimeter-wave beamforming. *IEEE Internet of Things Journal* **7,** 1336–1349 (2019).

17. Wang, S., Wang, J., Su, C. & Ma, X. *Intelligent Detection Algorithm Against UAVs' GPS Spoofing Attack* in *2020 IEEE 26th International Conference on Parallel and Distributed Systems (ICPADS)* (2020), 382–389.

18. Greco, C., Pace, P., Basagni, S. & Fortino, G. Jamming detection at the edge of drone networks using Multi-layer Perceptrons and Decision Trees. *Applied Soft Computing* **111,** 107806 (2021).

19. Jasim, K. S., Ali Alheeti, K. M. & Najem Alaloosy, A. K. A. in *Advances in Cybersecurity, Cybercrimes, and Smart Emerging Technologies* 97–110 (Springer, 2023).

20. Anantvalee, T. & Wu, J. A survey on intrusion detection in mobile ad hoc networks. *Wireless network security,* 159–180 (2007).

21. Axelsson, S. Intrusion detection systems: A survey and taxonomy (2000).

22. Tsai, C.-F., Hsu, Y.-F., Lin, C.-Y. & Lin, W.-Y. Intrusion detection by machine learning: A review. *expert systems with applications* **36,** 11994–12000 (2009).

23. Ramadan, R. A., Emara, A.-H., Al-Sarem, M. & Elhamahmy, M. Internet of Drones Intrusion Detection Using Deep Learning. *Electronics* **10,** 2633 (2021).

24. Omar, S., Ngadi, A. & Jebur, H. H. Machine learning techniques for anomaly detection: an overview. *International Journal of Computer Applications* **79** (2013).

25. Bhattacharyya, D. K. & Kalita, J. K. *Network anomaly detection: A machine learning perspective* (Crc Press, 2013).

26. Xiao, L., Wan, X., Lu, X., Zhang, Y. & Wu, D. IoT security techniques based on machine learning: How do IoT devices use AI to enhance security? *IEEE Signal Processing Magazine* **35,** 41–49 (2018).

27. Banerjee, N., Giannetsos, T., Panaousis, E. & Took, C. C. *Unsupervised learning for trustworthy IoT* in *2018 IEEE international conference on fuzzy systems (FUZZ-IEEE)* (2018), 1–8.

28. Altawy, R. & Youssef, A. M. Security, privacy, and safety aspects of civilian drones: A survey. *ACM Transactions on Cyber-Physical Systems* **1,** 1–25 (2016).

29. Mukherjee, A., Fakoorian, S. A. A., Huang, J. & Swindlehurst, A. L. Principles of physical layer security in multiuser wireless networks: A survey. *IEEE Communications Surveys & Tutorials* **16,** 1550–1573 (2014).

30. Pinto, P. C., Barros, J. & Win, M. Z. *Physical-layer security in stochastic wireless networks* in *2008 11th IEEE Singapore International Conference on Communication Systems* (2008), 974–979.

31. Poor, H. V. & Schaefer, R. F. Wireless physical layer security. *Proceedings of the National Academy of Sciences* **114,** 19–26 (2017).

32. Pan, F., Pang, Z., Luvisotto, M., Xiao, M. & Wen, H. Physical-layer security for industrial wireless control systems: Basics and future directions. *IEEE Industrial Electronics Magazine* **12,** 18–27 (2018).

33. Zhou, X., Song, L. & Zhang, Y. *Physical layer security in wireless communications* (Crc Press, 2013).

34. Wang, D., Bai, B., Zhao, W. & Han, Z. A survey of optimization approaches for wireless physical layer security. *IEEE Communications Surveys & Tutorials* **21,** 1878–1911 (2018).

35. Niu, J. *et al.* Defending jamming attack in wide-area monitoring system for smart grid. *Telecommunication Systems* **60,** 159–167 (2015).

36. Wang, L. & Wyglinski, A. M. *A combined approach for distinguishing different types of jamming attacks against wireless networks* in *Proceedings of 2011 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing* (2011), 809–814.

37. Šimon, O., Götthans, T. & Popela, M. *Commercial uav jamming possibilities* in *2022 32nd International Conference Radioelektronika (RADIOELEKTRONIKA)* (2022), 1–6.

38. Duan, B. *et al. Anti-jamming path planning for unmanned aerial vehicles with imperfect jammer information* in *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (2018), 729–735.

39. Zou, Y., Zhu, J., Wang, X. & Hanzo, L. A survey on wireless security: Technical challenges, recent advances, and future trends. *Proceedings of the IEEE* **104,** 1727–1765 (2016).

40. Pirayesh, H. & Zeng, H. Jamming attacks and anti-jamming strategies in wireless networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials* (2022).

41. Slimane, H. O., Benouadah, S., Khoei, T. T. & Kaabouch, N. *A Light Boosting-based ML Model for Detecting Deceptive Jamming Attacks on UAVs* in *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)* (2022), 0328–0333.

42. Darsena, D., Gelli, G., Iudice, I. & Verde, F. Detection and blind channel estimation for UAV-aided wireless sensor networks in smart cities under mobile jamming attack. *IEEE Internet of Things Journal* **9,** 11932–11950 (2021).

43. Maksutov, A. A., Valter, D. A., Borisenko, G. V. & Ovchinnikov, K. A. *Real-time simulation of the GLONASS system signals using SDR* in *2019 IEEE conference of russian young researchers in electrical and electronic engineering (EIConRus)* (2019), 26–28.

44. Larcom, J. A. & Liu, H. *Modeling and characterization of GPS spoofing* in *2013 IEEE international conference on technologies for Homeland Security (HST)* (2013), 729–734.

45. Oligeri, G., Sciancalepore, S., Ibrahim, O. A. & Di Pietro, R. *Drive me not: GPS spoofing detection via cellular network: (architectures, models, and experiments)* in *Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks* (2019), 12–22.

46. Tamazin, M., Karaim, M., Noureldin, A. & Rustamov, R. GNSSs, signals, and receivers. *Multifunctional operation and application of GPS,* 119–139 (2018).

47. Van den Bergh, B. & Pollin, S. Keeping UAVs under control during GPS jamming. *IEEE Systems Journal* **13,** 2010–2021 (2018).

48. Wei, X. & Sikdar, B. *Impact of GPS time spoofing attacks on cyber physical systems* in *2019 IEEE international conference on industrial technology (ICIT)* (2019), 1155–1160.

49. Zheng, X.-C. & Sun, H.-M. Hijacking Unmanned Aerial Vehicle by Exploiting Civil GPS Vulnerabilities Using Software-defined Radio. *Sensors & Materials* **32** (2020).

50. Arthur, M. P. *Detecting signal spoofing and jamming attacks in UAV networks using a lightweight IDS* in *2019 international conference on computer, information and telecommunication systems (CITS)* (2019), 1–5.

51. Aissou, G., Slimane, H. O., Benouadah, S. & Kaabouch, N. *Tree-based supervised machine learning models for detecting GPS spoofing attacks on UAS* in *2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)* (2021), 0649–0653.

52. Du, X., Cao, Y., Wen, L. & Yang, Z. A Review of Intrusion Detection in FANETs. *Secure and Digitalized Future Mobility: Shaping the Ground and Air Vehicles Cooperation,* 99 (2022).

53. Aboelfottoh, A. A. & Azer, M. A. *Intrusion Detection in VANETs and ACVs using Deep Learning* in *2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)* (2022), 241–245.

54. Whelan, J., Sangarapillai, T., Minawi, O., Almehmadi, A. & El-Khatib, K. *Novelty-based intrusion detection of sensor attacks on unmanned aerial vehicles* in *Proceedings of the 16th ACM symposium on QoS and security for wireless and mobile networks* (2020), 23–28.

55. Boucetta, C., Nour, B., Hammami, S. E., Moungla, H. & Afifi, H. *Adaptive range-based anomaly detection in drone-assisted cellular networks* in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)* (2019), 1239–1244.

56. Zhang, R., Condomines, J.-P., Chemali, R. & Larrieu, N. *Network intrusion detection system for drone fleet using both spectral analysis and robust controller/observer* PhD thesis (ENAC, 2018).

57. Asif, R., Hu, Y.-F., Ali, M., Li, J.-P. & Abdo, K. *Signal Classification for Safety Critical Aeronautical Communications for Anti-Jamming using Artificial Intelligence* in *2021 IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)* (2021), 1–6.

58. Moustafa, N. & Jolfaei, A. *Autonomous detection of malicious events using machine learning models in drone networks* in *Proceedings of the 2nd ACM MobiCom Workshop on Drone Assisted Wireless Communications for 5G and beyond* (2020), 61–66.

59. Ouiazzane, S., Addou, M. & Barramou, F. A multiagent and machine learning based denial of service intrusion detection system for drone networks. *Geospatial Intelligence: Applications and Future Trends,* 51–65 (2022).

60. Li, Y. *et al.* Jamming Detection and Classification in OFDM-Based UAVs via Feature-and Spectrogram-Tailored Machine Learning. *IEEE Access* **10,** 16859–16870 (2022).

61. Viana, J. *et al.* Two methods for Jamming Identification in UAVs Networks using New Synthetic Dataset. *arXiv preprint arXiv:2203.11373* (2022).

62. Dang, Y., Benzaıd, C., Yang, B. & Taleb, T. *Deep Learning for GPS Spoofing Detection in Cellular-Enabled UAV Systems* in *2021 International Conference on Networking and Network Applications (NaNA)* (2021), 501–506.

63. Shafique, A., Mehmood, A. & Elhadef, M. Detecting signal spoofing attack in uavs using machine learning models. *IEEE Access* **9,** 93803–93815 (2021).

64. Fraser, B., Al-Rubaye, S., Aslam, S. & Tsourdos, A. *Enhancing the security of unmanned aerial systems using digital-twin technology and intrusion detection* in *2021 IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)* (2021), 1–10.

65. Khoei, T. T. *et al. A Comparative Analysis of Supervised and Unsupervised Models for Detecting GPS Spoofing Attack on UAVs* in *2022 IEEE International Conference on Electro Information Technology (eIT)* (2022), 279–284.

66. Wang, B., Wang, Z., Liu, L., Liu, D. & Peng, X. *Data-driven anomaly detection for UAV sensor data based on deep learning prediction model* in *2019 Prognostics and System Health Management Conference (PHM-Paris)* (2019), 286–290.

67. Galvan, J., Raja, A., Li, Y. & Yuan, J. *Sensor Data-Driven UAV Anomaly Detection using Deep Learning Approach* in *MILCOM 2021-2021 IEEE Military Communications Conference (MILCOM)* (2021), 589–594.

68. Baig, Z., Syed, N. & Mohammad, N. Securing the Smart City Airspace: Drone Cyber Attack Detection through Machine Learning. *Future Internet* **14,** 205 (2022).

69. Whelan, J., Almehmadi, A. & El-Khatib, K. Artificial intelligence for intrusion detection systems in unmanned aerial vehicles. *Computers and Electrical Engineering* **99,** 107784 (2022).

70. Shrestha, R., Omidkar, A., Roudi, S. A., Abbas, R. & Kim, S. Machine-learning-enabled intrusion detection system for cellular connected UAV networks. *Electronics* **10,** 1549 (2021).

71. Da Silva, L. M., Ferrão, I. G. & Branco, K. R. *A systematic mapping study in intrusion detection system for unmanned aerial vehicles security* in *2022 Latin American Robotics Symposium (LARS), 2022 Brazilian Symposium on Robotics (SBR), and 2022 Workshop on Robotics in Education (WRE)* (2022), 43–48.

72. Ahmed, M., Cox, D., Simpson, B. & Aloufi, A. Ecu-ioft: A dataset for analysing cyber-attacks on internet of flying things. *Applied Sciences* **12,** 1990 (2022).

73. Ahn, H. *Deep learning based anomaly detection for a vehicle in swarm drone system* in *2020 International Conference on Unmanned Aircraft Systems (ICUAS)* (2020), 557–561.

74. Pardhasaradhi, B. & Srihari, P. *Stealthy GPS spoofer design by incorporating processing time and clock offsets* in *2021 IEEE 18th India Council International Conference (INDICON)* (2021), 1–6.

75. Borhani-Darian, P., Li, H., Wu, P. & Closas, P. *Deep neural network approach to detect GNSS spoofing attacks* in *Proceedings of the 33rd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2020)* (2020), 3241–3252.

76. Whelan, J., Sangarapillai, T., Minawi, O., Almehmadi, A. & El-Khatib, K. *UAV Attack Dataset* 2020. doi:10.21227/00dg-0d12. https://dx.doi.org/10.21227/00dg-0d12.

77. Kosmanos, D. *et al. RF Jamming Dataset for Vehicular Wireless Networks* 2023. doi:10.21227/4zwk-yw78. https://dx.doi.org/10.21227/4zwk-yw78.

78. Mozaffari, M., Saad, W., Bennis, M., Nam, Y.-H. & Debbah, M. A tutorial on UAVs for wireless networks: Applications, challenges, and open problems. *IEEE communications surveys & tutorials* **21,** 2334–2360 (2019).

79. Gadal, S. *et al.* Machine Learning-Based Anomaly Detection Using K-Mean Array and Sequential Minimal Optimization. *Electronics* **11,** 2158 (2022).

80. Radha, B. & Sakthivel, D. Network Anomaly Detection using Data Mining Algorithms. *NVEO-NATURAL VOLATILES & ESSENTIAL OILS Journal— NVEO,* 5047–5056 (2021).

81. Alexandridis, T. K. *et al.* Novelty detection classifiers in weed mapping: Silybum marianum detection on UAV multispectral images. *Sensors* **17,** 2007 (2017).

82. Hoang, T. M., Nguyen, N. M. & Duong, T. Q. Detection of eavesdropping attack in UAV-aided wireless systems: Unsupervised learning with one-class SVM and k-means clustering. *IEEE Wireless Communications Letters* **9,** 139–142 (2019).

83. Fawcett, T. An introduction to ROC analysis. *Pattern recognition letters* **27,** 861–874 (2006).

84. Li, B., Fei, Z. & Zhang, Y. UAV communications for 5G and beyond: Recent advances and future trends. *IEEE Internet of Things Journal* **6,** 2241–2263 (2018).

85. Khan, M. A. *et al.* Swarm of UAVs for network management in 6G: A technical review. *IEEE Transactions on Network and Service Management* (2022).

86. Tsao, K.-Y., Girdler, T. & Vassilakis, V. G. A survey of cyber security threats and solutions for UAV communications and flying ad-hoc networks. *Ad Hoc Networks,* 102894 (2022).

87. Yang, L., Li, S., Li, C., Zhang, A. & Zhang, X. A survey of unmanned aerial vehicle flight data anomaly detection: Technologies, applications, and future directions. *Science China Technological Sciences,* 1–19 (2023).

88. Lu, X., Xiao, L., Dai, C. & Dai, H. UAV-aided cellular communications with deep reinforcement learning against jamming. *IEEE Wireless Communications* **27,** 48–53 (2020).

89. Krayani, A., Alam, A. S., Marcenaro, L., Nallanathan, A. & Regazzoni, C. Automatic Jamming Signal Classification in Cognitive UAV Radios. *IEEE Transactions on Vehicular Technology* (2022).

90. Kim, K. *et al.* Security analysis against spoofing attacks for distributed UAVs. *Decentralized IoT Systems and Security* (2020).

91. Tlili, F., Ayed, S., Chaari, L. & Ouni, B. *Artificial Intelligence Based Approach for Fault and Anomaly Detection Within UAVs* in *International Conference on Advanced Information Networking and Applications* (2022), 297–308.

92. Šimon, O. & Götthans, T. A Survey on the Use of Deep Learning Techniques for UAV Jamming and Deception. *Electronics* **11,** 3025 (2022).

93. Titouna, C., Naıt-Abdesselam, F. & Moungla, H. *An online anomaly detection approach for unmanned aerial vehicles* in *2020 International Wireless Communications and Mobile Computing (IWCMC)* (2020), 469–474.

94. Price, J. *et al. Real-time Classification of Jamming Attacks against UAVs via on-board Software-defined Radio and Machine Learning-based Receiver Module* in *2022 IEEE International Conference on Electro Information Technology (eIT)* (2022), 1–5.

95. Bae, G. & Joe, I. in *Advanced Multimedia and Ubiquitous Engineering* 305–310 (Springer, 2019).

96. Ajakwe, S. O., Ihekoronye, V. U., Kim, D.-S. & Lee, J. M. Pervasive Intrusion Detection Scheme to Mitigate Sensor Attacks on UAV Networks., 1267–1268 (2022).

97. Park, K. H., Park, E. & Kim, H. K. *Unsupervised intrusion detection system for unmanned aerial vehicle with less labeling effort* in *International Conference on Information Security Applications* (2020), 45–58.

98. Khan, S., Liew, C. F., Yairi, T. & McWilliam, R. Unsupervised anomaly detection in unmanned aerial vehicles. *Applied Soft Computing* **83,** 105650 (2019).

99. Park, K. H., Park, E. & Kim, H. K. Unsupervised fault detection on unmanned aerial vehicles: Encoding and thresholding approach. *Sensors* **21,** 2208 (2021).

100. Viana, J. *et al. A Convolutional Attention Based Deep Learning Solution for 5G UAV Network Attack Recognition over Fading Channels and Interference* in *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)* (2022), 1–5.

101. Bae, G. & Joe, I. *UAV anomaly detection with distributed artificial intelligence based on LSTM-AE and AE* in *Advanced Multimedia and Ubiquitous Engineering: MUE/FutureTech 2019 13* (2020), 305–310.

102. Moustafa, N. & Slay, J. *UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)* in *2015 Military Communications and Information Systems Conference (MilCIS)* (2015), 1–6. doi:10.1109/MilCIS.2015.7348942.