Detection and Mitigation of De-authentication attacks in WIFI Networks

By

Anam Saud

Reg No: 00000274112

Supervisor

Dr. Waleed Bin Shahid

A thesis submitted in conformity with the requirements for the degree of Master of Science in Information Security

Department of Information Security

Military College of Signals (MCS)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

September 2022

Thesis Acceptance Certificate

Certified that final copy of MS Thesis written by **NS Anam Saud** Registration No. **00000274112** of **Military College of Signals** has been vetted by undersigned, found complete in all respects as per NUST Statutes/ Regulations/ MS Policy, is free of plagiarism, errors, and mistakes, and is accepted as partial fulfillment for the award of MS degree. It is further certified that necessary amendments as pointed out by GEC members and foreign/ local evaluators of the scholar have also been incorporated in the said thesis.

Signature : ______________________

Name of Supervisor : **Dr. Waleed Bin Shahid**

Date : ______________________

Signature(HOD) : ______________________

Date : ______________________

Signature(Dean/Principal) : ______________________

Date : ______________________

Plagiarism Undertaking

I solemnly declare that the research work presented in the thesis titled "Detection and Mitigation of De-authentication attacks in Wi-Fi Networks" is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me. I understand the zero-tolerance policy of the HEC and the Military College of

Signals, NUST towards plagiarism. Therefore, I as an author of the above-titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred to/cited. I undertake that if I am found guilty of any formal

plagiarism in the above-titled thesis even after awarding of MS degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized thesis.

Student/Author Signature : ____________________

Name : ______________________________________

Declaration

I, *Anam Saud* declare that this thesis titled "Detection and Mitigation of De-authentication attacks in WIFI Networks" and the work presented in it are my own and has been generated by me as a result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a Master of Science degree at NUST
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at NUST or any other institution, this has been clearly stated
3. Where I have consulted the published work of others, this is always clearly attributed
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work
5. I have acknowledged all main sources of help
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself

Anam Saud,

Reg No: 00000274112

Copyright Notice

This thesis is dedicated to my

beloved parents and dear husband and most importantly my son because of whom it took me this long

Abstract

One way to remove a client from the network is by de-authentication. The components of a wireless network, usually referred to as a Wi-Fi network, are an AP (Access Point) and a client. The de-authentication procedure may be initiated by either the AP or the client. To de-authenticate, a de-authentication frame is utilised. A de-authentication frame is a management frame. There are altogether three basic frame types in the IEEE 802.11 standard. Data frames that used to move information between stations. There are various distinct types of data frames, depending on the network. Performing area-clearing operations, channel acquisition and carrier-sensing maintenance tasks, and positive acknowledging of received data all need the employment of control frames in conjunction with data frames. The Management frames complete the process by performing supervision duties; they are used to enter and exit wireless networks. Management frames like de-authentication and disassociation result in the termination of a client's network connection. The transmission of management frames has always been done in clear text and without message authentication. Due to the fact that they are delivered in clear, de-authentication or disassociation frames can be readily spoofed on the part of a client or an AP. As a result, neither the client nor the AP will be in the 802.11 standard's authenticated state. Following then, all packets will be discarded until authentication is restored, which will result in the client's network services being cut off. This assault, a de-authentication attached, is comparable to the man-in-the-middle assault. This specific weakness in the 802.11 Management Frames involves very careful detection and mitigation of de-authentication attacks in Wi-Fi Networks. The goal of this study is to discover a de-authentication attached while it is occurring or has just begun.

Keywords: *De-authentication attack, detection, Machine Learning*

Acknowledgments

I would like to thank God Almighty, for letting me through all the difficulties. It was not easy especially after the birth of my son during the degree and a full time job but day by day Allah has given me strength to stand up and get going. You are the one who has let me finish my degree. I have always trusted you and will keep on trusting you. I would like to acknowledge my supervisor Dr. Waleed Bin Shahid for his continuous support of my MS study and research. It is he who made this work possible. I would also like to thank my committee members Dr.Waseem Iqbal and Dr. Hammad Afzal for their encouragement and cooperative attitude. Last but not the least, I would like to give special thanks to my dear husband for always standing by side and supporting me. Thanks to all my family members for their continuous support. Without you none of this would indeed be possible.

Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

Abbreviations

ML	Machine Learning
De-auth	De-authentication
IDS	Intrusion Detection system
PPV	Positive Predictive Value
SVM	Support Vector Machines
VM	Virtual Machine
TP	True Positive
FP	False Positive
FN	False Negative
TN	True Negative

Chapter 1

Introduction

1.1 Introduction

The world's communication networks with the quickest growth rate are wireless networks. Wireless networks enable information transmission between two points without the use of any physical connections, such as wires or cables. Radio Frequency (RF) is the medium they employ for communication. We live in a society where communication is ubiquitous, and wireless communication in particular is a crucial aspect of our daily life. Mobile phones, GPS receivers, remote controls, Bluetooth audio, and other wireless communication systems are some of the most often utilised wireless communication systems in our daily lives. An access point (AP) and clients, often known as nodes, are the two main components of a wireless network. These clients can take the form of a laptop, computer, or mobile device. As previously discussed, radio waves are used to communicate between these nodes. There are four primary types of wireless networks. [1] [2] [3]

- Internet access is made available inside a constrained region using wireless local area networks (WLAN). Initially employed in homes and workplaces, WLAN technology is now also present in shops and eateries because everyone today needs access to the internet. [4] [5].

- Metropolitan area wireless networks (WMAN) are the next topic. These are set up in urban areas around the world to give those who are in public places outside of a building, such an office or home network, access to the Internet. [6].

- Wide-ranging geographic areas are covered by wireless wide-area networks (WWAN), including nearby towns and cities. In order to enable access outside the coverage area of a wireless LAN or metropolitan network, wireless WANs use cellular technology.[7].
- Devices are connected by a wireless personal area network (WPAN) over a short distance, typically within a person's reach. They only permit communication within a 10 metre radius. WPAN uses technologies like Bluetooth. [8].

A wireless network technology known as Wi-Fi or WiFi is based on the IEEE 802.11 family of specifications [9] [10] [11]. Wi-Fi enables neighbouring digital devices, including routers and laptops/mobiles, to exchange data through radio waves. WLAN is used to access the internet. The most widely used wireless local area network (WLAN) protocol that uses 2.4 GHz UHF and 5 GHz ISM frequency channels is called Wi-Fi. Devices that are between 20 and 40 metres from the source of Wi-Fi can access the Internet. To make Wi-Fi more secure throughout time, the protocol has undergone security advancements and new security protocols have been added. Wi-Fi networks are still susceptible for a variety of reasons, though. De-authentication attacks are one of them, and they are a problem that many users encounter..

Devices like the wireless access point (AP), router, and clients on the wireless network should not be available to users outside of that network in order for the Wi-Fi network to be safe. Encryption is one of the primary methods for securing a Wi-Fi network. Data is encrypted using cryptographic keys by Wi-Fi security standards. Symmetrical encryption, employed by Wi-Fi systems, encrypts and decrypts data using the same key.

Currently, four wireless security protocols are offered:

- Wired Equivalent Privacy (WEP).
- Wi-Fi Protected Access (WPA).
- Wi-Fi Protected Access 2 (WPA 2).
- Wi-Fi Protected Access 3 (WPA 3).

All these above mentioned protocols provide protection to data frames. Which leads to the our main issue which is the protection of the management frame, more on this below.

However, the lack of protection of Management frames lead to a major vulnerability on Wi-Fi networks and that is the de-authentication attack.

In 802.11 protocol there are 3 types of frames. These are called Data Frames, Control Frames and Management Frames. The Management frames does probing, associating, roaming, and disconnecting clients from the Wi-Fi Network.

1.1.1 The Process of De-authentication and the De-Authentication Attack

Disconnecting a client or node from a wireless network is accomplished by de-authentication. Either device may send a de-authentication frame if it wants to de-authenticate from an Access Point (AP) or if it wants an AP to de-authenticate from a client. Since neither party can reject de-authentication, de-authentication occurs once a de-authentication packet is received and accepted by the destination. A management frame that is sent in clear text is the de-authentication packet or frame. The de-authentication packet's frame format is displayed below.

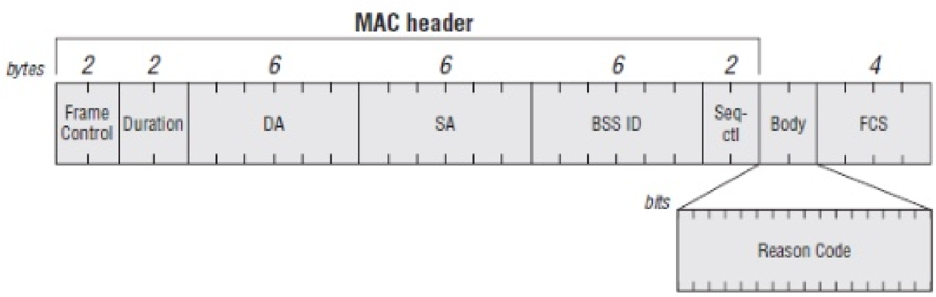

Figure 1.1: Frame format of De-authentication packet.

The de-authentication frame is vulnerable to attacks because it is not encrypted. The fact that management frames are not encrypted is a common target for Wi-Fi assaults. Using a de-authentication attack to impair wireless networks is disruptive. It is a member of the denial-of-service family and causes networks to go momentarily down. These strategies are typically low-key because they don't call for specialised knowledge or expensive tools. Deauthentication attacks target the exchange of information between a client and a router (AP). A client can be a computer, a laptop, or a mobile device. In a de-authentication attack, the attacker sends either the client or the access point

(AP) a phoney de-authentication packet, breaking their connection. [12] [13]. False requests that obstruct regular communication are a component of de-authentication attacks. Since 802.11-based wireless networks demand de-authentication frames whenever users end connections, they are vulnerable to attack by this tactic. The issue is that access points can fail to detect that the request came from an unauthorised source. Because arriving frames are not validated by networks, hackers can spoof them. Even when sessions use WEP, a lack of encryption feeds the flames. Additionally, WiFi networks lack a method of MAC address verification. [14]. Attackers could therefore use spoofing and de-authentication techniques. Connections are terminated using spoof frames.

The de-authentication request made by the attacker to disconnect the authorised user (client) from the network is shown in the figure below.

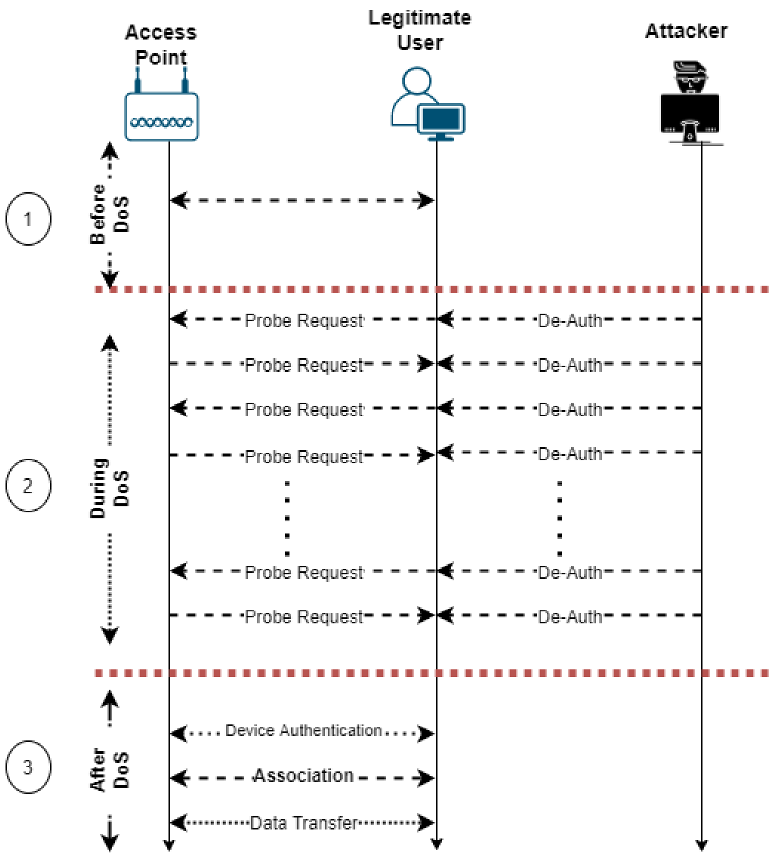

Figure 1.2: De-authentication attack flow.

There are numerous ways an attacker can conduct a de-authentication assault.

- The attacker can create a counterfeit de-authentication frame by configuring the source MAC address to be that of the client and the destination MAC address to be that of the access point. When the client receives the bogus de-authentication, the AP disconnects it from the network.

- The attacker then generates a fraudulent de-authentication frame with the client's MAC address as the destination and the AP's MAC address as the source in order to disconnect the user from the network. The first predicament is reversed in this instance.
- A packet created by the attacker might also have the AP's MAC address as its source and the destination as its broadcast MAC address (FF:FF:FF:FF:FF:FF). All users who are logged in to the AP are then disconnected.

1.2 Research Goals

The weakness in the 802.11 protocol to protect the Management frame is a critical problem in WIFI security. The problem has not been addressed properly hence the problem caused by unencrypted management frames de-authentication attacks continue to hamper WIFI communication. The aim of this thesis was therefore to gain an understanding of the de-authentication attack and to analyze WIFI network traffic so that we are able to differentiate between legitimate and illegitimate de-authentication packets and hence be able to detect de-authentication attacks. This detection will in turn help us in the mitigation of the attack.

1.3 Problem Statement and Motivation

The importance of network connectivity cannot be denied. A de-authentication attack is a form of Man in the Middle Attack. This attack can disrupt communication and dismantle network connectivity. Therefore the need to address this issue is very pressing. Work has been done to help overcome this weakness in the 802.11 standards but nothing has been done so far which can be adopted by all efficiently. Hence, a solution to this security issue needs to be addressed. De-authentication attack is considered one of the most powerful DoS attacks in the field of wireless communication, but it is also one of the most difficult to accurately identify. Therefore, the aim of the work is a practical study of the interaction between the client and the AP during the exchange of frames in normal conditions and during the DoS-attack. To solve the problem, the following tasks were set.

- Practical implementation of the de-authentication attack.
- Analysis of frames during the attack to identify anomalies.
- Development of an algorithm for detecting de-authentication attacks.

1.4 Scope of Work and research objectives

The scope of the research is limited to Wifi WLAN Networks only.

Through this research I aim to to achieve following goals:

- Detection of de-authentication attacks when it occurs
- Predicting, by using the behavior of the Network traffic that a de-authentication attacks is about to hit a Wi-Fi Network
- Using those predictions to prevent a de-authentication attack from happening
- Ways to mitigate the de-authentication attacks ones they have taken place

1.5 Relevance to National Needs

In order to develop a vibrant Cyber Security Ecosystem within Pakistan cyber security issues need to be identified, analyzed, researched and eventually solved. Addressing the cyber security issue related to de-authentication attacks will be an addition to this secure ecosystem. This will also help establish a front line of defense against today's immediate threats that we as a nation face. This will strengthen the future cyber security environment within the country.

Some areas of application which are relevant to our national needs are as following:

- Data leakage prevention
- Securing Networks
- Prevention of Data ex filtration from within Public and Private Organizations

- Military setups working will sensitive national data

1.6 Main Contributions

The main contributions of the thesis are as follows:

- **Review of existing literature:** We have performed an extensive review of solutions proposed to mitigate de-authentication attack. The literature has been analyzed and weaknesses have been discussed. Solutions are compared with each other and using the existing study the conclusions and directions of future work have been carefully carved.
- **Differentiating legitimate and illegitimate de-authentication packets:**Wireshark has been used extensively to analyse network traffic so that we can distinguish between real de-authentication packets and fraudulent ones. This is a significant milestone in our study since it is utilised to provide a defence against de-authentication attacks.
- **Proposed a Machine Learning based solution:** A machine learning-based intrusion detection system (IDS) has been proposed, which will be more reliable and accurate in spotting a de-authentication assault and, consequently, at mitigating it.
- **Feature Selection for the Machine Learning based IDS:**So far most researches have analyze only a few frame exchange characteristics. In this research I have tried to cover most of these characteristics to make sure our results are more accurate.
- **Experimental validation of proposed methods:** The proposed methodology has been validated using a larger dataset. The experiments have shown very good results of identifier a de-authentication attack.

1.7 Thesis Organization

The subsequent section provides the outlines of the given chapters:

- Chapter 2: This chapter contains the background and brief description of existing literature for de-authentication attack and its detection and mitigation Additionally, it also briefly explains the basic concepts used in this work e.g. machine learning, and the feature extraction for this purpose.
- Chapter 3: This chapter represents the overall methodology and the techniques used to detect and mitigate the de-authentication attacks. This chapter consists of experimental settings and the results obtained by it.
- Chapter 4: Lastly, this chapter includes the conclusion of this study and discusses possible future directions.

Chapter 2

Background and Literature Review

2.1 Background

Wi-Fi vulnerabilities and IEEE 802.11 security methods have been under research and scrutiny for a long time [15] [16] [17]. As a result, continued investigation into the IEEE 802.11 standard's flaws is required to stop the resurgence of new crimes in this field. The issue of security is now the most serious. [18] [19]. Wi-Fi networks are prone to attacks due to the shortcomings and limitations of the IEEE 802.11 protocols [20] [21]. One of the major shortcoming in the protocol is the lack of encryption or authentication of the Management frame. This means that the Management frame which is responsible for de-authentication moves across a Wi-Fi network in plain text. This then leads to the de-authentication attack which is the scope of our research [22]. Management frames are a very important data packets that control the communication between an Access Point and other nodes [23] [24].

Three different kinds of frames exist:

1. Management frames (type 00)
2. Control frames (type 01)
3. Data frames (type 10)

Data communication between the wireless clients and the access point is regulated by control frames.

The data frames are secured by security protocols like WEP, WPA, or WPA2 as discussed in Chapter 1 1. and contain the real data that is received from the network layer.

Lastly, the Management frame which are used to control and monitor the communication between an access point and a wireless client/node. It is responsible for authentication, association and de-authentication between the nodes within the Wi-Fi network. However, unfortunately management frames are not encrypted, like data frames which are transmitted over the network in encrypted form. Due to the lack of encryption, 802.11 management frames are vulnerable de-authentication attacks [25].

2.2 Existing Research Work

A lot of work and research has been done to detect de-authentication attack and methods have been suggested to mitigate this attack. Below is a summary of all the existing research related to de-authentication in Wi-Fi network.

2.2.1 Protocol Modification

Solutions have been proposed to modify the current protocol. These solutions suggest modifying the current authentication framework by authenticating the de-authentication frames.

Lets have a look at the de-authentication frame to understand their structure and functionality. A de-authentication frame is basically used to terminate or end a Wi-Fi connection. It can be send by either a client or an AP. It is a notification and has to be accepted by either party.

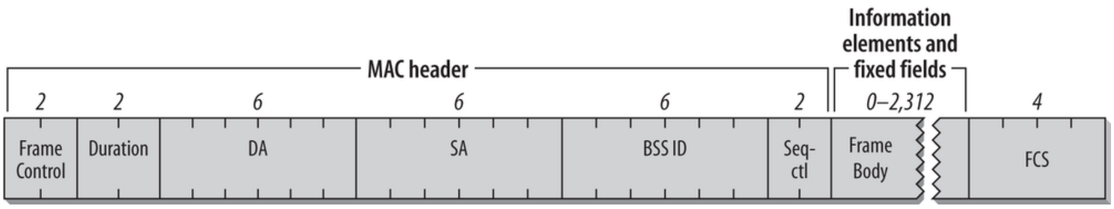

Figure 2.1: De-authentication Frame Structure.

The following is contained in the frame body of the de-authentication frame.

1. Reason Code (2 byte)
2. Vendor Specific Information (one or more)
3. 802.11w (MFP) info

The de-authentication frame is both unauthenticated and unencrypted. There is no source authentication in the 802.11 Wi-Fi standard since there is no way or procedure for ensuring that a packet truly originates from the source it claims to. This means that if an attacker manages to "spoof" the MAC address of a valid network user, they may simply "spoof" another node and request several MAC-layer services. As a result, this underlying issue gives rise to de-authentication attacks.

Bellardo [26] offered an encryption-based approach. This alteration, however, is not based on a cryptographic calculation. This method advises authenticating all management frames so that if the de-authentication packet was actually a fake, it wouldn't be able to pass authentication. By using frame authentication, this strategy can thwart the de-authentication attack, but it necessitates updating the firmware on both the client and the AP. Each management frame would incur additional expenses if authentication were added, which would affect both clients and AP. Additionally, since authentication requires a lot of processing, validating every management communication would quickly deplete the batteries of portable electronics like cellphones and PDAs, etc.

Arora [27], has suggested a comparable technique for confirming management frameworks. However, cryptography was used in this approach. This approach uses a one-way hashing mechanism, which makes it computationally impossible to circumvent. The differences in how the protocol appeared before and after the adjustment are seen in Figures "ref"fig:demo figure 2.2" and "ref"fig:demo figure 2.3" below. This technique may be applied as a straightforward firmware update because it doesn't call for sophisticated cryptographic calculations. The application of cryptographic algorithms and the subsequent generation of distinctive tokens to create a secure communication protocol might, however, be time-consuming and ineffective for roaming devices like mobile phones.

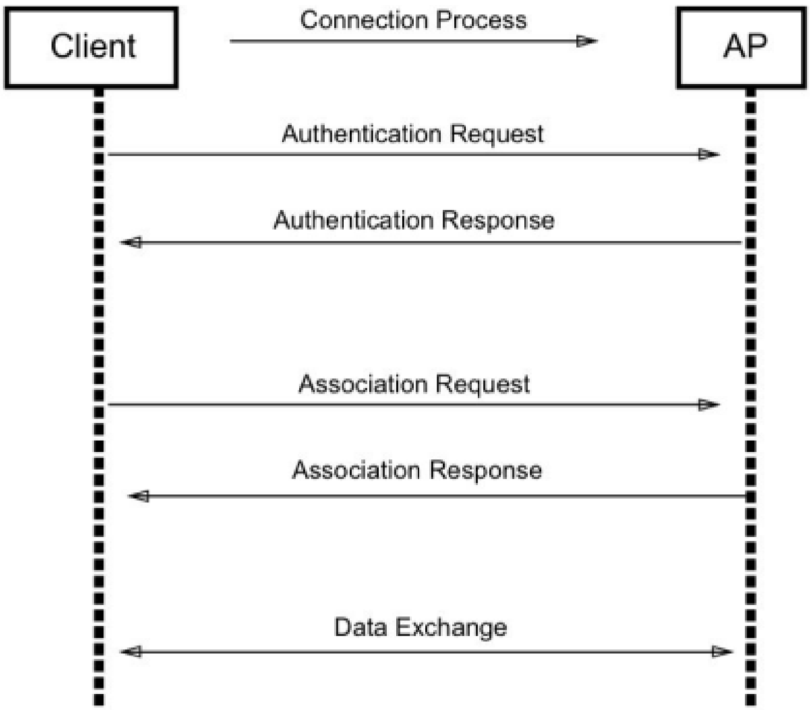

Figure 2.2: Original Association Process.

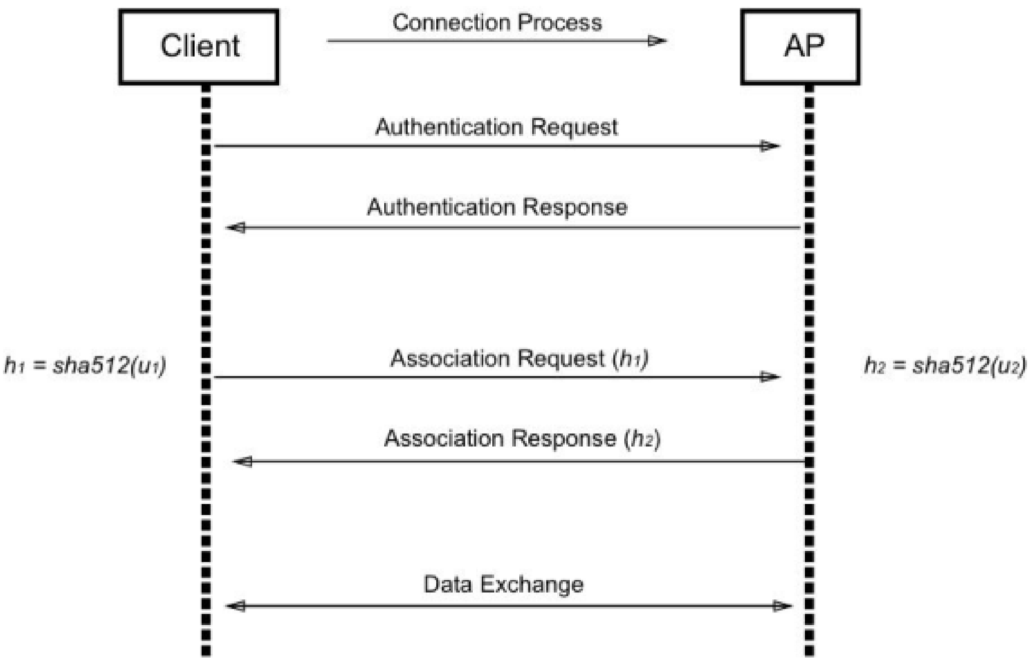

Figure 2.3: Modified Association Process.

2.2.2 Delay the Processing of Management Frames

Bellardo [26] also recommended an alternative tactic, delaying processing of management frames while delaying processing of de-authentication frames. Here is an example to demonstrate this. If an AP receives a data frame from the same client after receiving a de-authentication frame from the same client that has already been authenticated, the previous de-authentication frame(s) will be disregarded and erased. This is because by delaying the execution of de-authentication and queuing such requests for 5–10 seconds, an AP may monitor these client packets. The request will be queued if a de-authentication packet shows up when a client is already authorised and delivering data to the AP. If a data packet appears again, the de-authentication request is then disregarded because a legitimate client wouldn't create packets in such sequence. The same approach may be used in reverse to decrease fraudulent or counterfeit de-authentication packets provided to the client on behalf of the AP. This approach has the benefit that it may be implemented with a simple firmware update to existing NICs and access points without the need for a new management structure. However, delaying the processing of all management frames will result in problems with hand-off and association for clients who are roaming. [26]

2.2.3 Using Reverse Address Resolution Protocol (R-ARP)

Edgar D. Cardenas [28] suggested to use RARP (Reverse Address Resolution Protocol) to avoid de-authentication attacks. By using RARP (Reverse Address Resolution Protocol) we can detect spoofed frames. However, an intelligent and seasoned attacker can manipulate the IP address of the client to bypass the RARP technique.

Reverse Address Resolution Protocol is based on computer networking which is used by a client computer to request it's IP address. It does this by sending the it's physical address (MAC) to a special RARP server that is on the same Local area network. The RARP request is a broadcast message telling all nodes in the LAN that its MAC address is this hence please tell his IP address. The response however is uni-cast. See figure 2.4 shows the RARP protocol. However, attacks on RAPR can result in the failure of this technique [29]

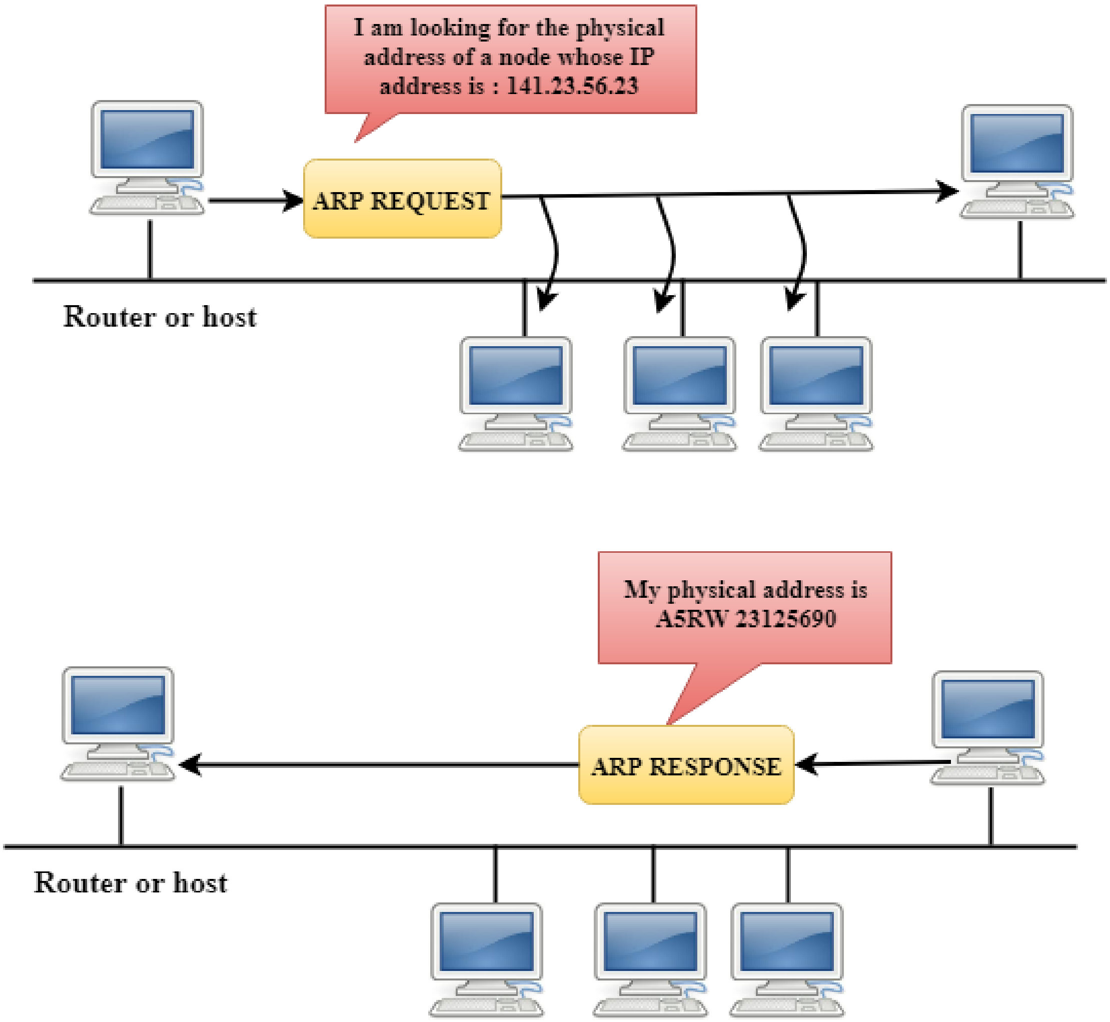

Figure 2.4: Reverse Address Resolution Protocol.

2.2.4 Letter Envelop Protocol

Nguyen [30] presented a Letter-envelope protocol that established a secret key between the client and the AP and then made use of it to verify every time a de-authentication frame was received.

This protocol relies on a one-way hash function f(.). Computing x in this situation where y = f(x) is computationally impossible; However, it is simple to calculate N given x. Following is the letter-envelope protocol:

- Initially, client randomly generates x1, then computes y1 = f(x1). Similarly, AP generates x2 and computes y2 = f(x2).
- The client transmits a "envelope" containing y1 to the AP during the authentica-

tion process, and the AP sends a "envelope" containing y2 to the client. So they used an envelope to exchange their x1 and y1.

- When the client wants to sever its connection with the AP, it sends the de-authentication or disassociation frame along with x1 to the AP. The "letter" is the name given to this variable. The frame is verified and will be treated correctly if this "letter" matches the previously transmitted "envelope," i.e. f(x1) = y1. If nothing else, the frame is put in the trash.
- In a similar manner, the AP transmits the disassociation/de-authentication frame along with the x2 value when it wants to cut off communication with the client. If f(x2)=y2, the client disconnects from the AP.

This strategy is effective in preventing de-authentication attacks, but both the client side and the AP require firmware updates.

2.2.5 Detection of spoofed packets based on Sequence Number

Guo [31], Xia [32], and Anjum [33] Several techniques for spotting spoofing attacks have been put forth, all based on sequence number analysis. The sequence number rises by 1 in each frame. The following frame would be sent with the sequence number x+1, x+2, and so on if the client supplied the previous frame with the sequence number x. In order to predict or estimate the sequence number beforehand and escape detection, an experienced attacker can send a frame with the sequence number "x+1". The approach is based on the assumption that when there are numerous frames to send, it might be difficult to transmit a frame with the proper sequence number at the precise timing.

2.2.6 Setting up a Threshold Number

Agarwal [34] By placing a limit on the quantity of de-authentication frames a client receives, the de-authentication attack was discovered. A client is alerted to the possibility of a de-authentication attack if it receives more de-authentication frames than the threshold number. The technique is susceptible to human mistake and judgement since the administrator sets a static threshold that is a quantitative value. There are a lot of false positives with this method because it just considers one parameter and ignores other wireless network-related parameters. Second, an attacker may send de-authentication frames at irregular intervals, making it impossible to identify the attack. The authors set a threshold value for the number of de-authentication or disassociation frames that they considered to be typical. An assault has been recognised if the intensity—that is, the number of frames received at a single moment—rises over this threshold. This suggests that the network is overloaded, resulting in congestion and a denial of service (DoS).

2.2.7 Using Machine Learning Approach

Here machine learning based approach is used to detect the attack [35]. This solution is closest to ours but the features they are using are limited however, in this research we have shortlisted more features.

2.2.8 802.11w : MFP or Management Frame Protection

To address the aforementioned problem, the 802.11w standard was released in 2008. This comprises technologies that provide data integrity, authenticity of the packet's origin, and replay protection. 802.11w was designed to offer data secrecy of management frames. WPA3 has this security feature built in. Both the client device and the AP must support WPA3 for this to function. However, today's high-end gadgets and pricier router versions are the only ones that support these standards. WPA3 is not supported by older devices. This needs to be able to support the standard on both the AP and the client end for it to work. The problem still exists because it will take more than ten years for devices to switch from WPA2 to WPA3. WPA3 use will slow down the processing performance overall.[36].

Management frames like de-authentication and disassociation, which were covered in the preceding sections, are always unauthenticated and unencrypted. The AP adds a Message Integrity Check Information Element (MIC IE) to each network management packet when 802.11w/WPA3 is enabled. [37]. Utilizing Integrity Group Temporal Key, this is accomplished (IGTK). During the four-way key handshake, this IGTK is created. This key cannot be duplicated, changed, or replayed since doing so would render the MIC useless. Additionally, the management frame contains certain information that is encrypted.

Shared below is a screenshot of a slide that was used by Jameson Blandford [38], Technical Marketing Engineer Cisco so explain Management Frame Protection (MFP)

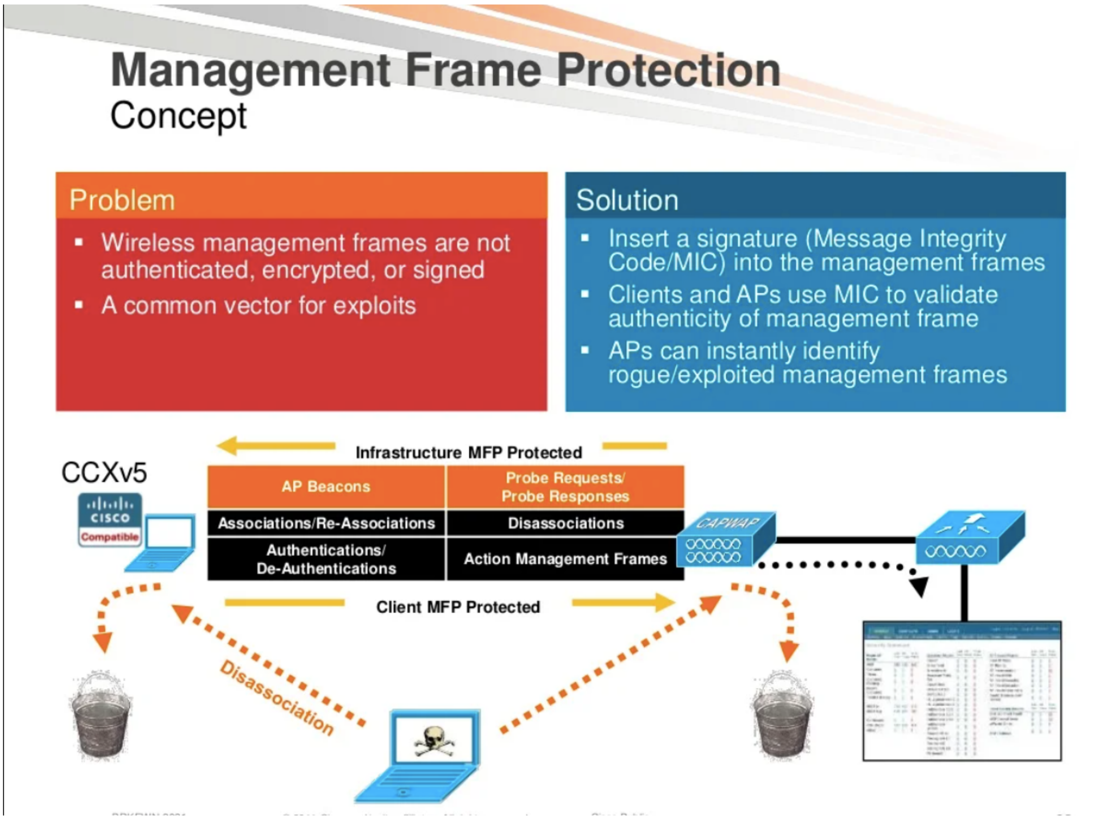

Figure 2.5: Concept of Management Frame Protection .

Chapter 3

Methodology Applied

3.1 Overview

This section explains the steps and the approach that was taken in order to detect when a Wi-Fi network is under a de-auth attack [39] [40]. Our research is using the Machine Learning approach for this. We will use an IDS Based on Machine Learning to detect a de-authentication attacks that takes place in a Wi-Fi network and generate alarm if the attack is detected. Figure 3.1 shows the approach that is taken. In the figure you can see that the ML-IDS is placed near the AP. This ML - IDS will monitor all the data coming and going from this particular AP only and not others.

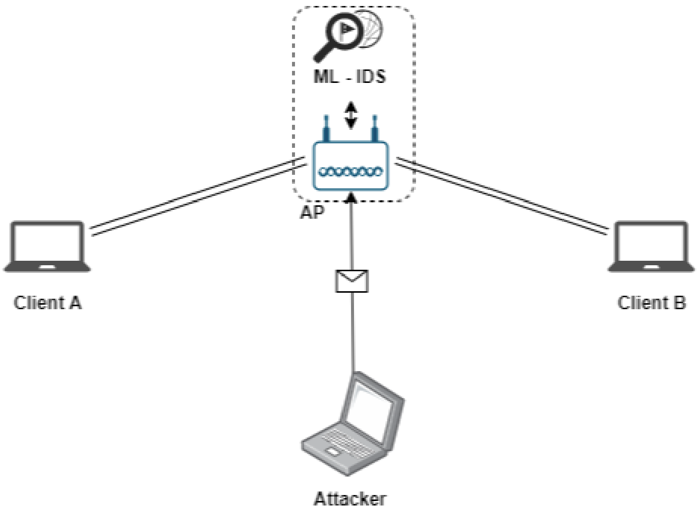

Figure 3.1: Proposed Approach.

In order for us to train our IDS, we made use of the training data. In order for us to have that training data we had to generate and collect our our data set. We could have used an already available data set for this purpose but as per our knowledge there are no such data sets available that can be used for the training of this IDS. Based on the training data, the ML predictor is trained so that it can then identify attack packets from benign packets and once this training is done it will be deployed on a live network. The ML-IDS will then analyse the live network traffic send across the Wi-Fi network. If the attack has occurred then an alarm will be generated.

3.2 Data Collection

The data collection we generated is being used to train the ML IDS. Since there is no publicly available data set for de-authentication attacks in Wi-Fi networks, the data set was created internally. We conducted a data de-authentication attack for the goal of data collection on a Wi-Fi network specifically set up for this. Wireshark was used to gather traffic before, during, and after the assault.

3.2.1 Performing a De-authentication Attack

In order to perform the de-authentication a Wi-Fi network was setup. This experimental setup consists of an AP and two clients (Laptops) and an attacker machine.

- **Access Point (AP):** The AP here serves as a connection point for two legitimate users in a network. In our experimental testbed, we are using a laptop as an access point. The operating system either Windows or Linux can be used. The reason for using a laptop as an AP is that we have to configure and install our ML-IDS at the same point where all communication traffic passed through AP could be analyzed.

- **Client A and B:** Client A and B are legitimate users who are connected to the AP. Both are communicating with each other and browsing the internet via AP. The implementation feasibility of both users is that both can be configured on the same laptop by using two different Virtual Machines installed. Windows operating system is fine for both VMs. While the network is connected with the above AP.

The MAC address table of AP will have two machines registered with two different MAC addresses.

- **Attacker Machine:** The attacker machine can also be on the same laptop using a third VM. However, the operating system will be Kali Linux because the tool aircrack-ng is pre-installed in Kali Linux.
- **Machine learning-based IDS:** The IDS will be installed on the same laptop which is the AP. As all the communication will be forwarded from the AP if IDS will be placed where it can detect and mitigate de-authentication attacks.

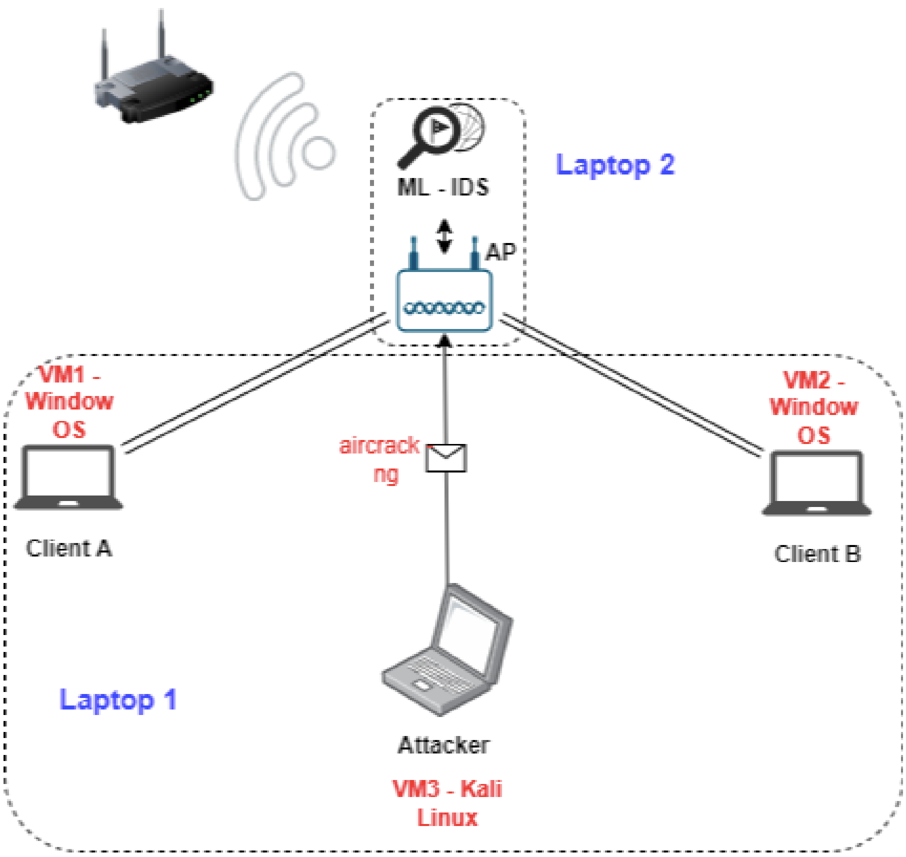

Figure 3.2: Experimental Setup for carrying out a De-authentication Attack

In order to collect our dataset we performed the spoofing and send malicious attack packets to the AP. Here are the steps we took to do that:

1. An experimental setup was arranged as shown above in figure 3.2
2. The attacker sends the de-authentication attack packet with a spoofed address.

3. The attacker spoofs a legitimate MAC Address of the client and runs periodic frames of the de-authentication [41]. Because de-authentication requests cannot be ignored, the access point therefore responds immediately to these requests

4. The AP considers this as a legitimate packet from a legitimate user thus disconnecting the victim (client) from the network. The figure 3.3 shows the spoofed address of the victim by the attack.

5. Once the attack is successful, the client disconnects from the Wi-Fi network and cannot connect to the network back again until the attacker stops the attack. [42]

6. We used the Aircrack-ng suite in Kali Linux is to perform this attack.

7. Used Wireshark in monitor mode to collect the network traffic before, during and after the attack.

8. The data collect via this experiment was then used as our dataset.

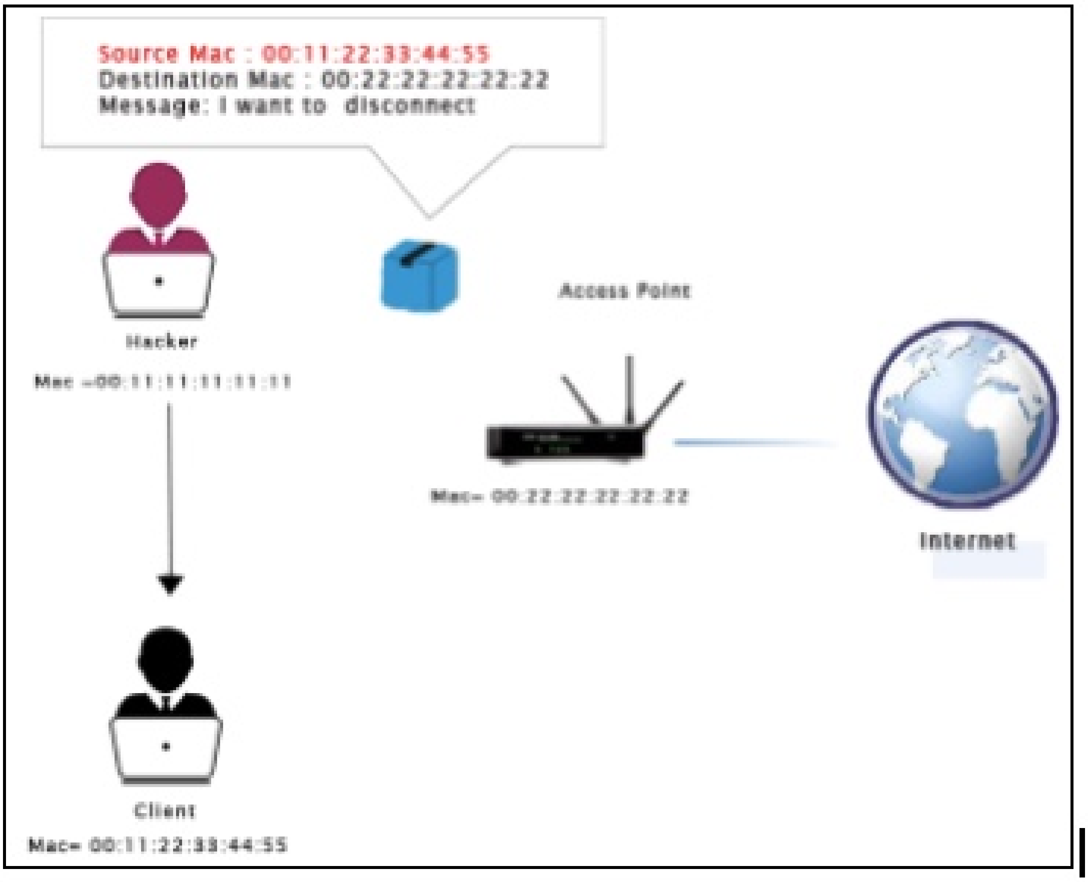

Figure 3.3: Spoofing done by the Attacker.

In this research the Kali Linux OS is used and the Aircrack-ng tool to run a de-authentication attack. This tool has powerful utilities that can be used to put various wireless network cards in monitoring modes and used for packet injection [26] The attack comprises of few steps to successfully de-authenticate someone:

- Run ifconfig and iwconfig
- Set wireless adapter on monitor mode by running airmon-ng start wlan0 (wlan1 or wlan2)
- Search the target machine or victim which you want to de-authenticate by running airodump-ng wlan0mon
- Once we choose our target from step 3, now we want more information about the target machine by running airodump-ng -d "target's BSSID" -c "target's channel number" "wireless adapter monitor mode name" e.g. airodump-ng -d 40:D7:BF:DC:4C:E8 -c 6 wlan0mon. You will get the mac address of the target machine.
- Now the final command to de-authenticate is aireplay-ng -0 0 -a “target's BSSID” -c “target's mac address” wlan0mon

3.2.2 Wi-fi Network Traffic Collection

The experimental setup explained in Section 3.2.1 and the de-authentication attacks that was performed generates network traffic. We collected network traffic under normal conditions meaning before launching a de-authentication attack and during the de-authentication attack.

Our aim was to collect a large set of data. Wireshark tool was used to collect the network traffic. Wireshark was used to collect network traffic before, during and after attack.The client machine were performing daily tasks including browsing the internet and using social media etc. Meanwhile the attacker was made to select any random time to target a chosen client and launch the de-auth attack on them. The data set is collected over a period of 5 hours so that enough data is gathered to do the training and then testing. For purposes of training the Machine Learning predictor we use 60 percent of the data set generated and kept the rest of 40 percent for testing purposes.

3.2.3 Feature Selection for the Machine Learning based IDS

Feature selection is the most important and crucial part of this research [43] [44] [45]. So far different researches have used a set of features to work with [25] [46] [47]. However, we aimed to shortlist all possible features that impact the behaviour of the network traffic and after that as per their significant and impact features are shortlisted. Wireshark's ability to collect de-auth frames in both normal and attack scenarios was examined.

Below are the features that have been shortlisted along with the reason of why those features will be used in the ML [48] [49]. These features will then be added to the model training [50].

1. Time Difference(TD) between de-auth and auth packets: In normal situations a client that is connected with AP sends in a disconnection request this means that now the client wants to end the connection. If after sending a disconnection requests, the client immediately sends a request to reconnect, this is suspicious behaviour. The interval between the client being disconnected and when it is re-authenticated with the same AP is what is meant by the time difference feature in this case. The client's request is really the difference in time between these two requests. Time Difference, for instance, equals t2 - t1 if the client disconnects at time t1 and reconnects at time t2. t2 - t1 = TD This value is quite modest in the de-authentication assault scenario. This occurs because when a client is abruptly unplugged, it instantly tries to re-authenticate with the same AP. As a result, the Time Difference has a relatively low value.

2. Number of de-authentication Frames send by client: Under normal situations when a user wants to disconnect, they will send a de-auth packet and get disconnected. However, under attack situations this is a bit tricky. The attacks aim is to make sure the attack is effective and for that to happen he sends multiple de-auth packets to make sure that the connection drops. Hence, the number of de-auth packets that are coming from a client or AP is an important factor. The more packets send, the more chances of the attack to be successful. This is called de-authentication frame flooding [25].

3. Exchange rate of the Numbers of Frames in a session: The number of frames sent and received by the client and the AP during a session is the total number of frames

sent and received by them. The term "session" refers to the period of time starting with the client's authentication and ending when the client disconnects from the network. The frame exchange counter is reset to 0 when the same client reconnects. The client's Frame Exchange value will be low if the attacker is often sending faked de-authentication packets on the client's behalf. The length or duration of a single session will be quite short since de-authentication attacks frequently cause disconnections. During normal traffic the number of frame exchange is significant and not so small as during an attack condition.

4. Number of Authentication Frames: This feature is the opposite of, Amount of de-authentication Frames feature. When a client is disconnected from a network due to malicious and spoofed de-auth packet send by the attacker on behalf of the client, the client immediately tries to reconnect and hence sends a authentication packet. Therefore, this feature serves as a counter for the quantity of authentication frames that are exchanged following a client disconnect and attempt to rejoin. When a client disconnects during regular traffic, it barely ever tries to rejoin. The client attempts to reconnect to the same AP during a de-authentication assault, which raises the number of this feature.

5. Number of TCP Frames: This function maintains track of the total number of TCP frames sent and received by each client. Under typical circumstances, quite a few TCP frames are exchanged. The client is automatically removed from the AP during a de-authentication attempt, signaling suspected attack activity, hence this number significantly drops.

6. The Reason Code of the de-auth packet:Research indicates that reason code 7 is used by the majority of de-authentication attack tools in all de-authentication frames. The valid de-authentication frame may also use the same reason code, but using the same reason code repeatedly is abnormal. Consequently, take into account this variable.

As in previous techniques, mostly reason code and MAC timestamp feature have been used to detect the de-authentication attack. If the reason code is unspecified and also with this MAC timestamp is not set according to the legitimate user, then we can say the particular packet is from the attacker's side. But we can enhance our technique by considering other features as explained above. Using more features for predictions improves accuracy and reduces false negatives.

3.2.4 Reason code in De-authentication frame

The de-authentication frame contains the reason code, which explains why the connection is interrupted. Here are some of the common reason codes Table 3.1

Code	Reason
0	Reserved
1	Unspecified reason
2	Previous authentication no longer valid
3	Station is leaving (or has left) IBSS or ESS
4	Disassociated due to inactivity
5	Disassociated because AP is unable to handle all currently associated stations
6	Class 2 frame received from non-authenticated station
7	Class 3 frame received from non-associated station
8	Disassociated because sending station is leaving (or has left) BSS
9	Station requesting (re)association is not authenticated with responding station
10	Disassociated because the information in the Power Capability element is unacceptable

Table 3.1: Reason codes for authentication cancellation

3.3 Attribute Comparison of a malicious de-authentication packet and a benign de-authentication packet

In order to understand the difference between a regular de-authentication and a de-authentication attack, a comparison is done between the attributes of the two in this section.

The following attributes were selected to see how they differ in a normal de-authentication and a de-authentication attack.

- Number of Headers in Wireshark
- Interface id:4 wlan0mon (Frame Header)
- Encapsulation Type (Frame header)
- Time Difference (Frame Header)
- Protocol in Frame (Frame Header)
- Radiotapheader v0, Length 18
- Data Rate
- Channel Frequency
- Antenna Signal
- Header 802.11 radio Information
- Frame Control Field
- Reason Code

Attribute	**Normal De-auth**	**De-auth Attack**
Number of Headers in Wireshark	5	3
Interface id:4 wlan0mon (Frame Header)	1	0
Encapsulation Type (Frame Header)	Encapsulation type: IEEE 802.11 plus radio-tap radio header (23)	Encapsulation type: IEEE 802.11 Wireless LAN (20)
Time difference (frame header)	[Time delta from previous captured frame: 0.003065695 seconds]	Time delta from previous captured frame: 0.004930000 seconds]
Frame length (frame header)	44 bytes	26 bytes
Protocol in frame (frame header)	[Protocols in frame:radiotap:wlan$_{r}adio:wlan$]	[Protocols in frame: wlan]
Radiotapheader v0,length 18	1	0
Data Rate	1.0 Mb/s	N/A
Channel frequency	2412[BG 1]	N/A
Antenna signal	-77 dbm	N/A
Header 802.11 radio information	1	0
Frame control field	0xc000	0xc000
Reason code	Reason code: Previous authentication no longer valid (0x0002)	Reason code: Class 3 frame received from non-associated STA (0x0007)

Table 3.2: Attribute Comparison between a real de-authentication and a de-authentication Attack

- **Number of Headers in Wireshark:** The header can include up to 40 bytes of options and has a minimum header size of 5 words and a maximum header size of 15 words, giving it a minimum size of 20 bytes and a maximum size of 60 bytes. [51]

- **Interface id:4 wlan0mon (Frame Header):** The wireless card should be configured to monitor mode in order to gather communications. As opposed to already processed 802.11 frames, in this mode you may view the actual frames that were broadcast and received in the air. Additionally, you can see every packet in the air, not just the ones that are sent to your machine.

- **Encapsulation Type :** Encapsulation is a general term for the process by which a lower-layer protocol inserts data from a higher-layer protocol into the data component of its frame. Encapsulation is the act of employing another form of packet to surround a certain type of packet. A typical de-authentication packet's encapsulation time differs slightly from an attack de-authentication packet's encapsulation time.

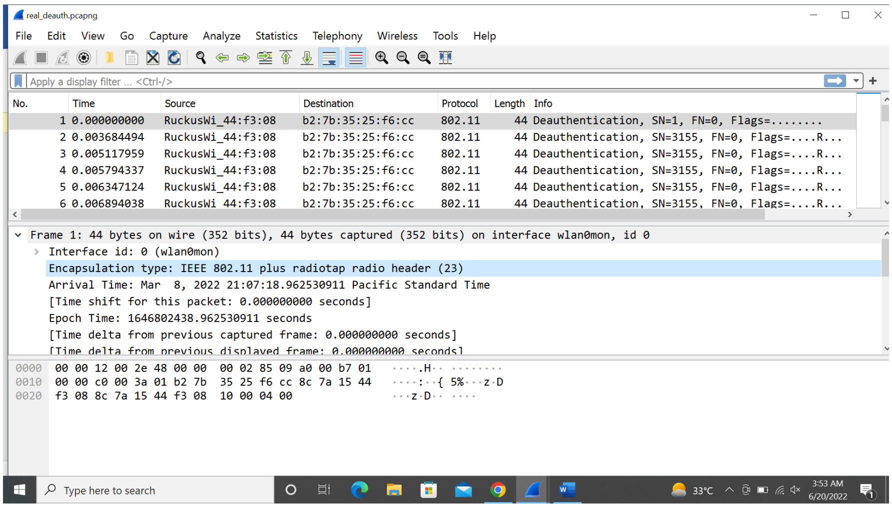

Figure 3.4: Encapsulation time of a real de-authentication packet.

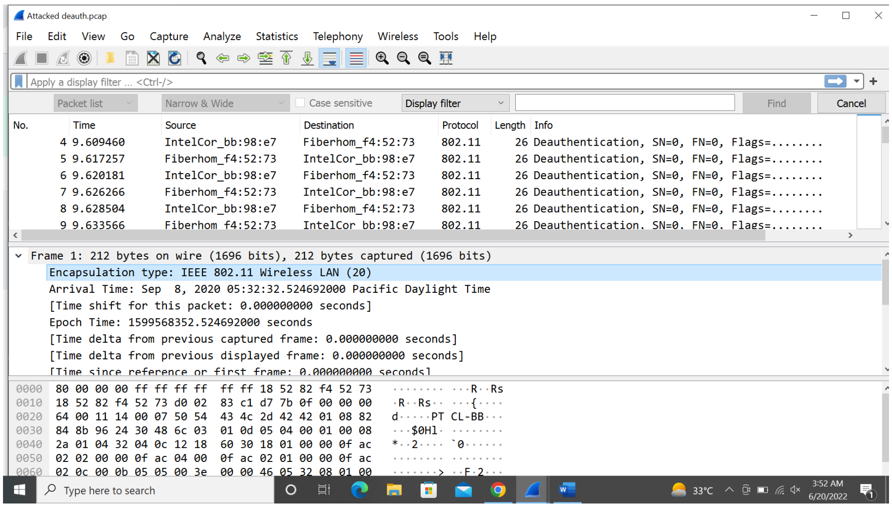

Figure 3.5: Encapsulation time of an attack de-authentication packet.

- **Time Difference :**

  The "time delta from previous displayed frame" is the difference in time stamps between the packet in question and the packet before it in the packet list.

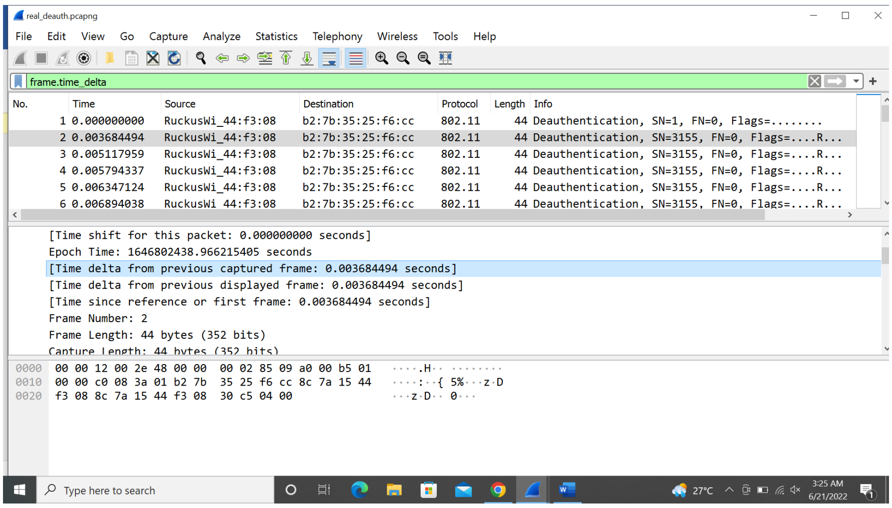

Figure 3.6: Time Difference between a normal de-authentication packet.

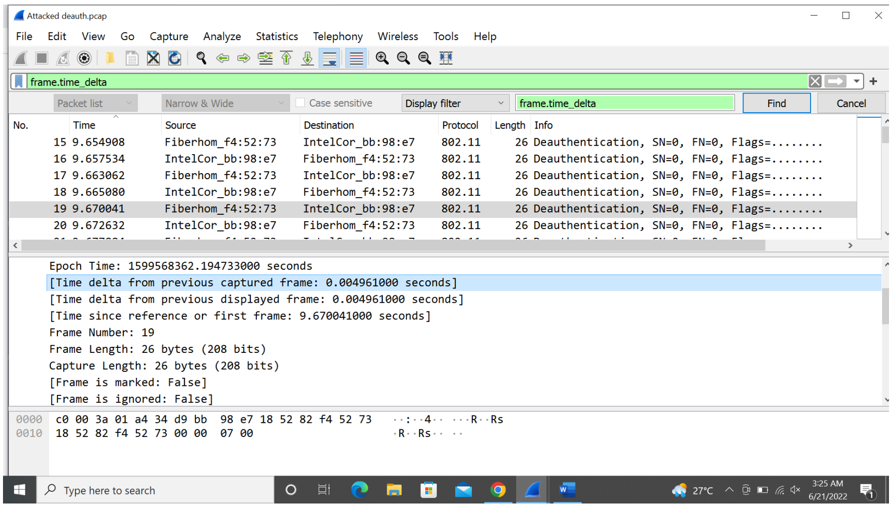

Figure 3.7: Time Difference between an attack de-authentication packet.

- **Frame Length :**

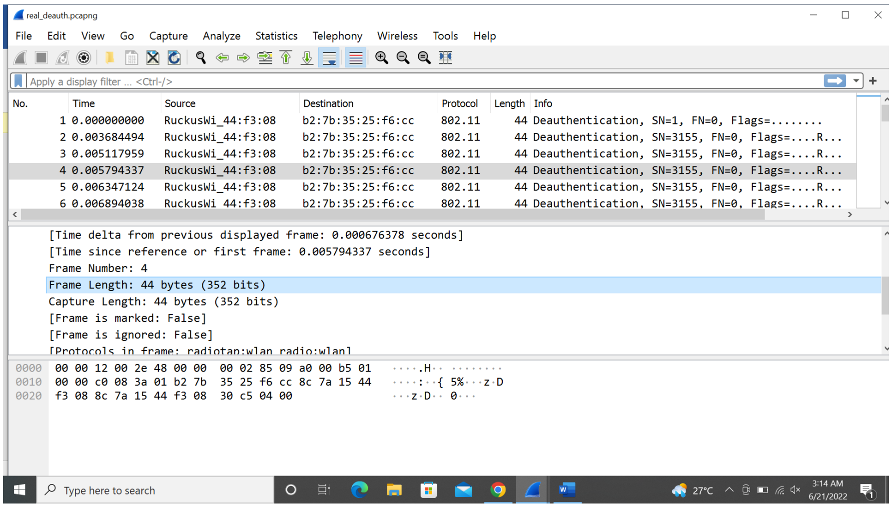

Figure 3.8: Frame length of a real de-authentication packet.

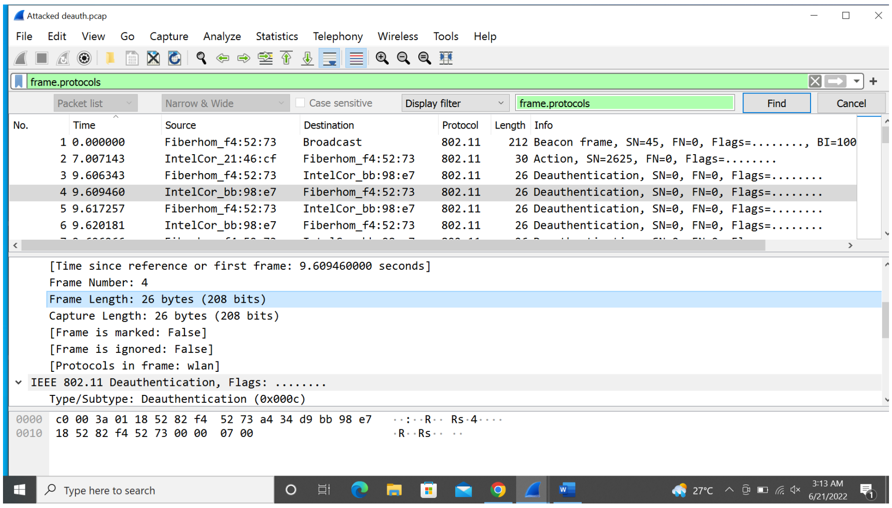

Figure 3.9: Frame Length of an attack de-authentication packet.

- **Protocol in Frame :** Although not a true protocol in and of itself, Wireshark uses the frame protocol as the foundation for all protocols built upon it. It displays data from the capture process, such as the precise moment a certain frame was taken. It might be considered a counterfeit dissector.

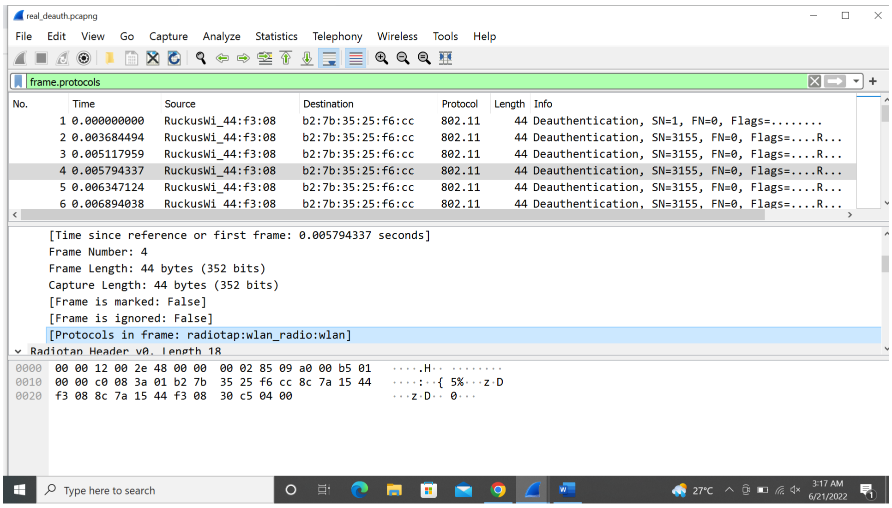

Figure 3.10: Protocol frame of a real de-authentication packet.

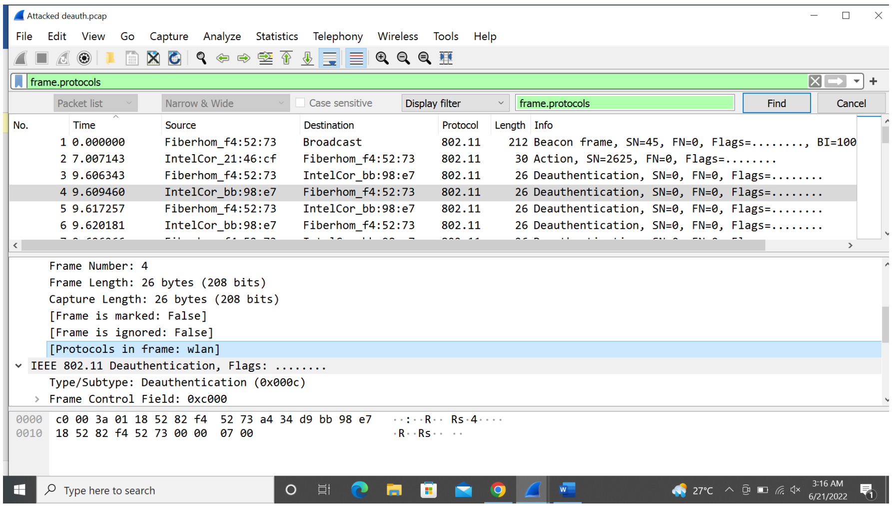

Figure 3.11: Protocol frame of an attack de-authentication packet.

- **Radiotapheader v0,length 18 :**

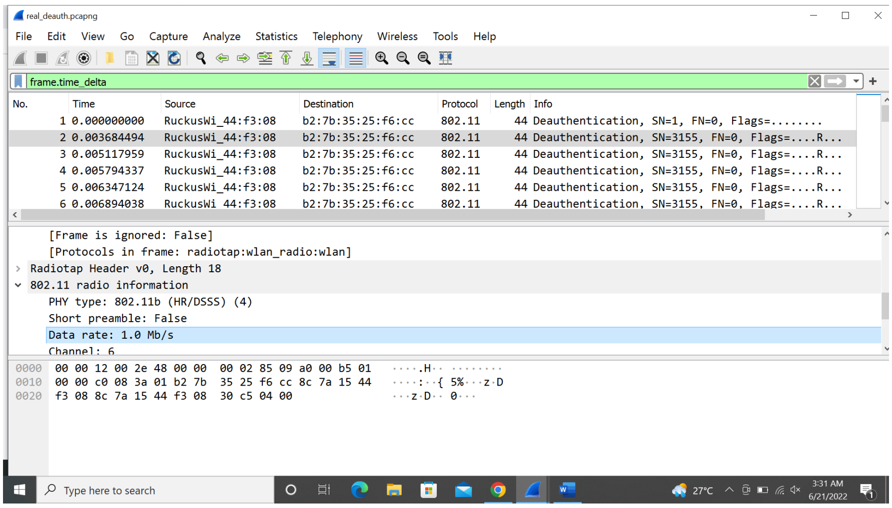

Figure 3.12: Real de-authentication packet

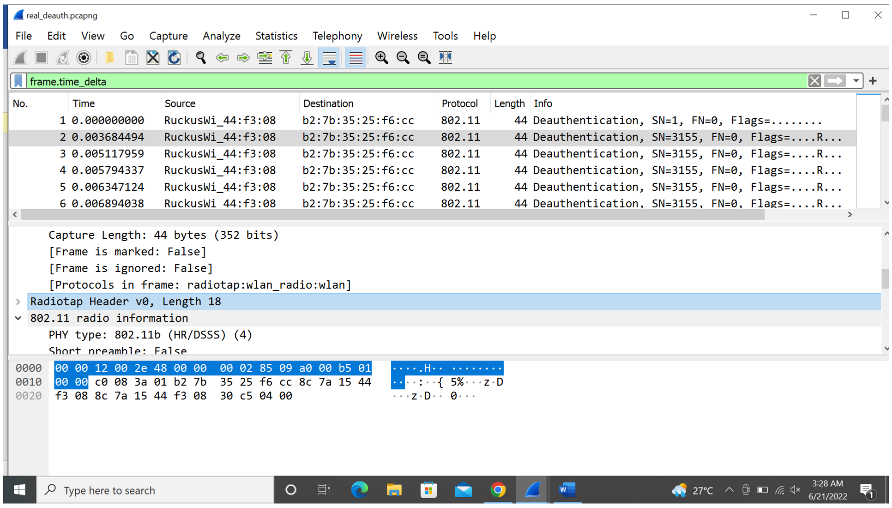

Figure 3.13: Attack de-authentication packet.

- **Data Rate :**

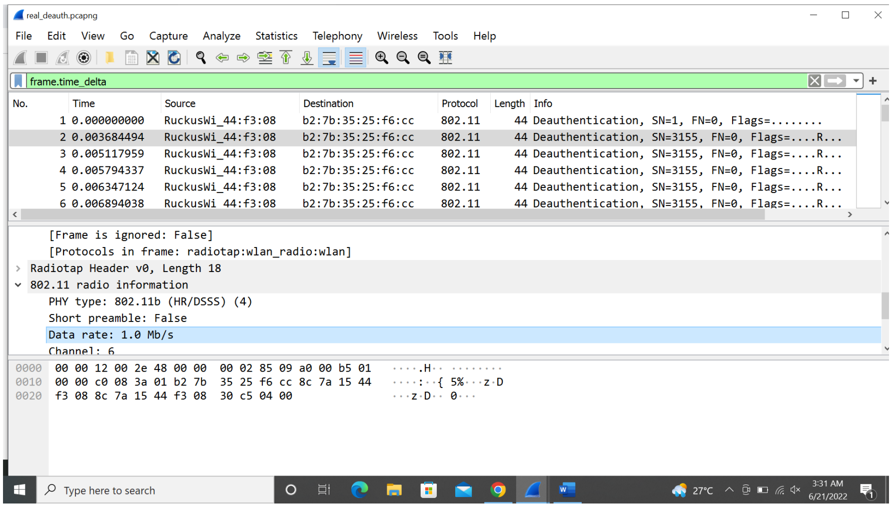

Figure 3.14: Data rate of a real de-authentication packet.

- **Channel Frequency :**

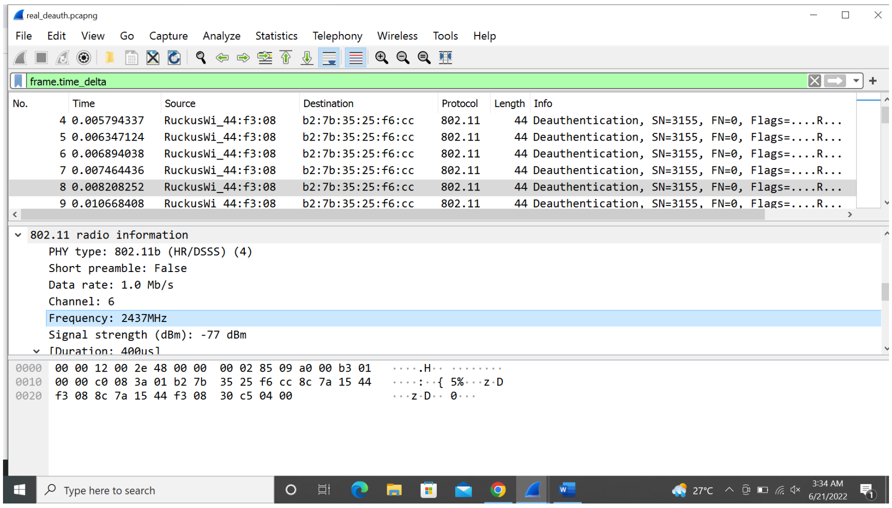

Figure 3.15: Channel Frequency of a real de-authentication packet.

- **Antenna Signal :**

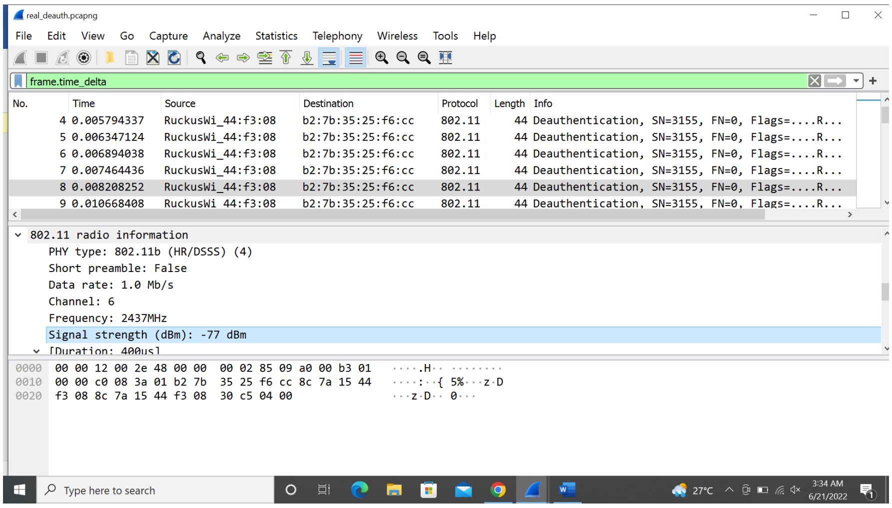

Figure 3.16: Antenna Signals of a real de-authentication packet.

- **Header 802.11 radio information :**

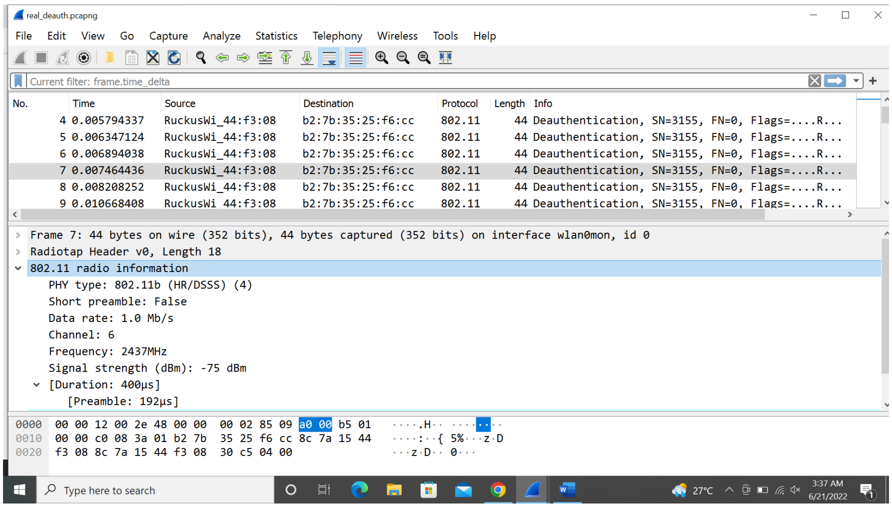

Figure 3.17: Header 802.11 radio information of a real de-authentication packet.

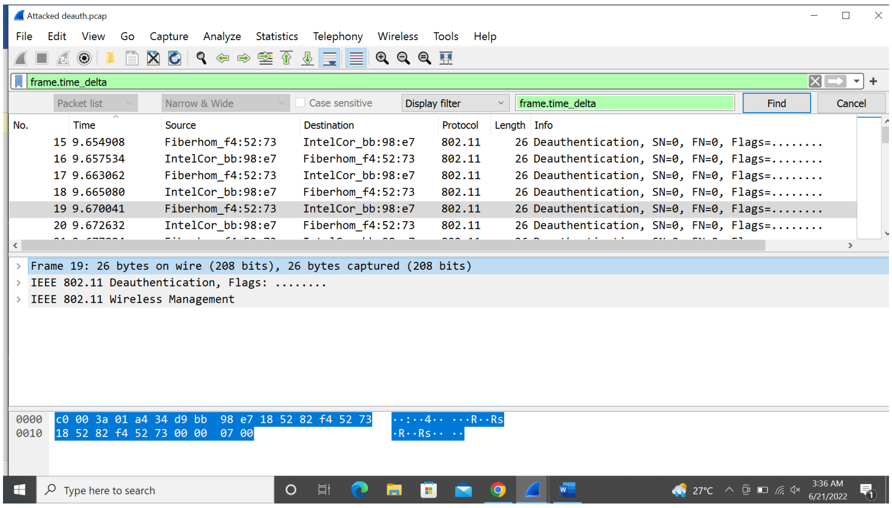

Figure 3.18: Header 802.11 radio information of an attack de-authentication packet.

- **Control field in the frame :**

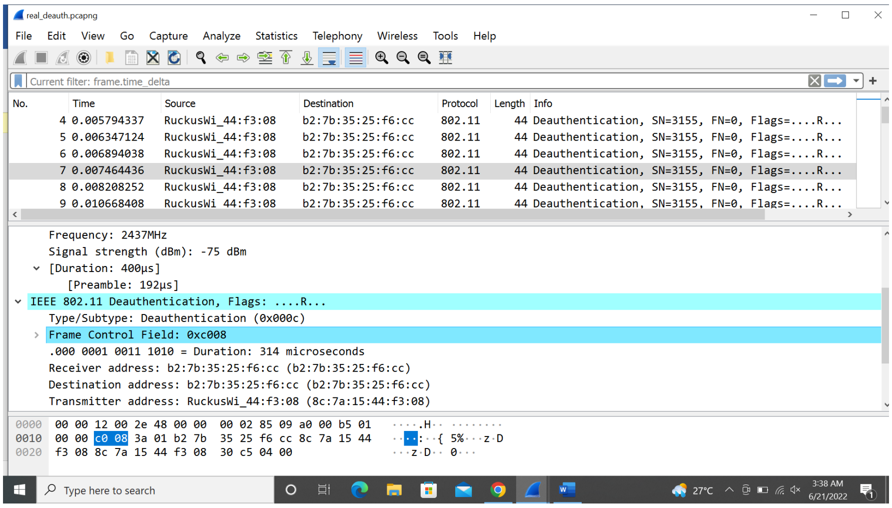

Figure 3.19: Control Field of a real de-authentication packet.

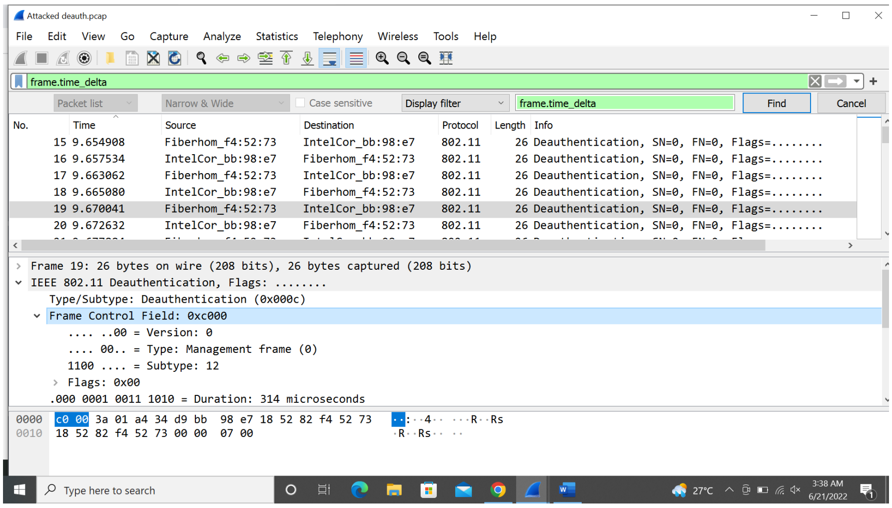

Figure 3.20: Control Field of an attack de-authentication packet.

- **Reason Code :** Reason Code is already explained in section 3.2.4

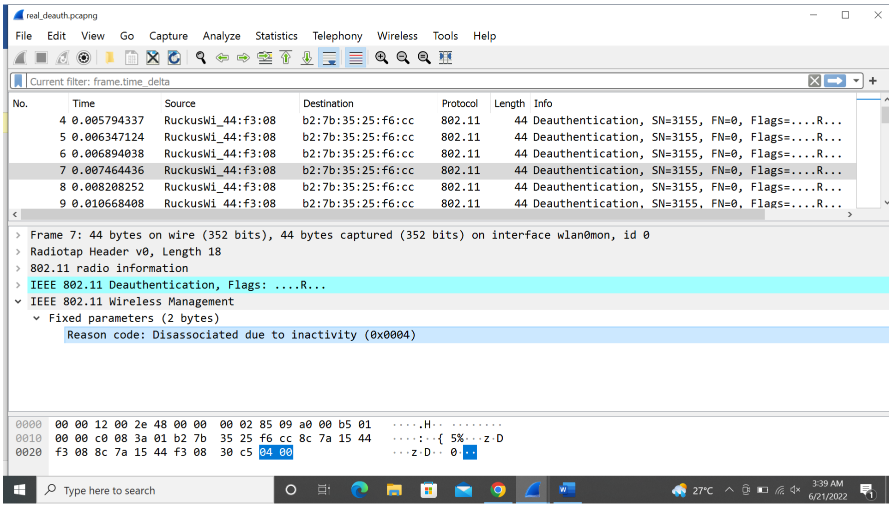

Figure 3.21: Reason Code of a real de-authentication packet.

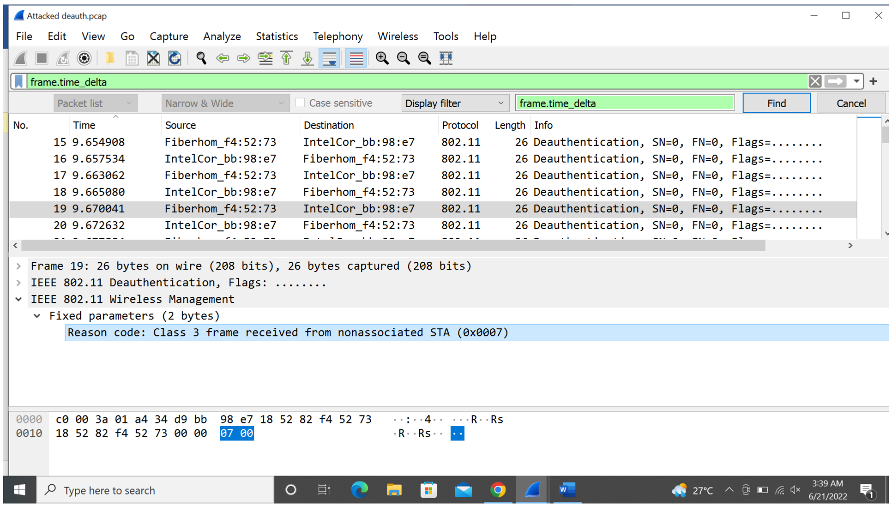

Figure 3.22: Reason Code of an attack de-authentication packet.

3.4 Training Data and Test Data

In machine learning the data set is divided into two subsets [52]. One is the training data - it's a part of our actual data set it is fed into the machine learning model to learn patterns. In this way, it trains the machine learning model. The other subset is called the testing data. Testing data is the portion of data we use to test the machine learning model.

Our training data set consists of 1069 packets. Out of these 1069 packets, 549 packets are malicious de-authentication packets and 520 packets are benign de-authentication packets.

Our test data set consists of 1021 packets. Out of these 1021 packets, 540 packets are malicious de-authentication packets and 481 packets are benign de-authentication packets.

3.5 Selection of the Machine Learning Classifier

The classifier that is chosen will largely determine how successful a machine learning-based IDS is.[53] [54]. An algorithm known as a classifier places input data into one of several categories or groupings. In this example, a classifier's task is to differentiate between malicious and innocent attack packets. We initially discussed our categorization methods in this section, along with the outcomes they produced. Each classifier has its own set of benefits and drawbacks in terms of parameters such as accuracy, sensitivity, specificity, and positive predictive value (PPV). These values are also used to select the classifier. [45].

1. Decision Tree: The most common applications of decision trees in machine learning are for categorization problems. The model is trained to detect whether or not the data corresponds to a known object class in this supervised machine learning task. Models are trained to assign class labels to processed data. We have two different kinds of de-auth packets in this instance: malicious and benign. Please refer to the Naive Bayesian Classifier Code, Appendix A.3 for more details.

2. Naive Bayesian Classifier: A straightforward and effective classifier is the Naive Bayesian method. The Naive Bayes technique is advised for working with a data collection that has millions of records with specific qualities. Naive Bayes makes use of the Bayes Theorem. It determines the chance that a given record or piece of data belongs to a certain class by calculating membership probabilities for each class. We have two different kinds of de-auth packets in this instance: malicious and benign. Please refer to the Naive Bayesian Classifier Code, Appendix A.1 for more details.

3. Linear Regression:A machine learning approach called linear regression is based on supervised learning. A regression test is performed. Regression develops a goal prediction value based on independent variables. It is mostly employed to ascertain how variables and forecasts relate to one another. The sort of link that different regression models take into account between the dependent and independent variables, as well as the number of independent variables utilised, varies. We have two different kinds of de-auth packets in this instance: malicious and benign. Please refer to the Naive Bayesian Classifier Code, Appendix A.2 for more details.

Each classifier, as was already established, has benefits and disadvantages of its own. Based on the classifier's precision, sensitivity (Detection Rate), specificity, and positive predictive value, the classifier is chosen (PPV).

- Accuracy: It provides you with the model's overall accuracy, or the percentage of all samples that the classifier properly identified. To calculate accuracy, use the following formula: (TP+TN)/(TP+TN+FP+FN).

Classifier	TP	TN	FP	FN	Accuracy
Decision Tree	479	529	2	10	0.988235294
Naive Bayesian	461	345	194	20	0.790196078
Linear Regression	461	509	30	10	0.96039604

Table 3.3: Accuracy of the Classifiers

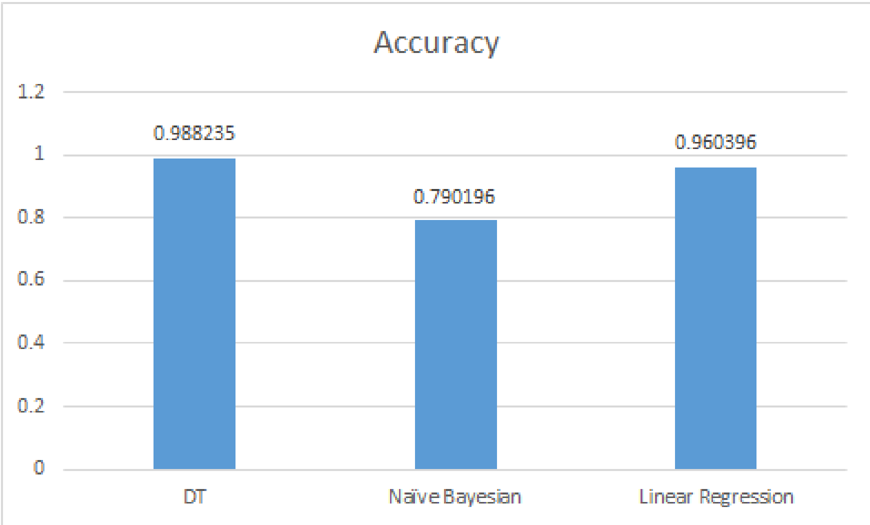

Figure 3.23: Accuracy of the Classifiers.

- Sensitivity (Detection Rate) : It reveals what percentage of all positive samples the classifier properly identified as positive. True Positive Rate (TPR), Sensitivity, and Probability of Detection are other names for it. To calculate Recall, use the following formula: TP/(TP+FN).

Classifier	TP	FN	Sensitivity
Decision Tree	479	10	0.979550102
Naive Bayesian	461	20	0.958419958
Linear Regression	461	10	0.978768577

Table 3.4: Sensitivity of the Classifiers

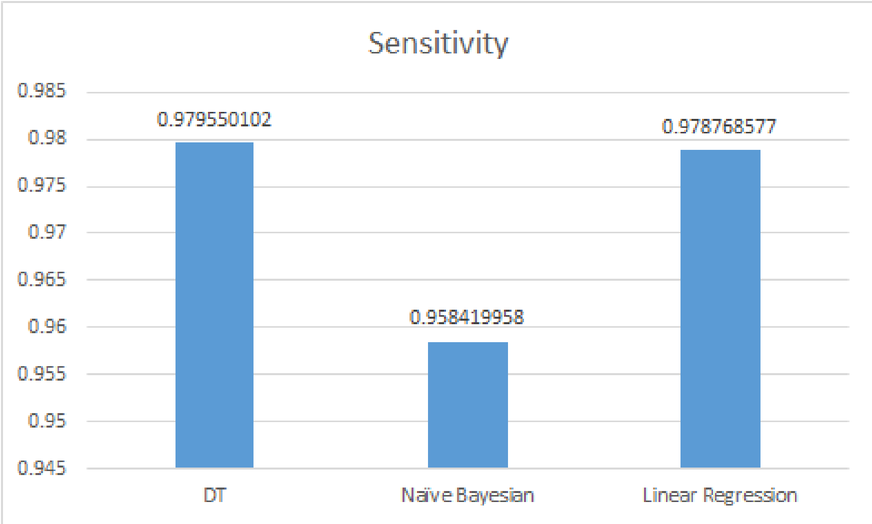

Figure 3.24: Sensitivity of the Classifiers.

- Specificity: It reveals what percentage of all negative samples the classifier properly identified as negative. Another name for it is True Negative Rate (TNR). To calculate specificity, use the following formula: TN/(TN+FP).

Classifier	**TN**	**FP**	**Specificity**
Decision Tree	529	2	0.996233522
Naive Bayesian	345	194	0.640074212
Linear Regression	509	30	0.944341373

Table 3.5: Specificity of the Classifiers

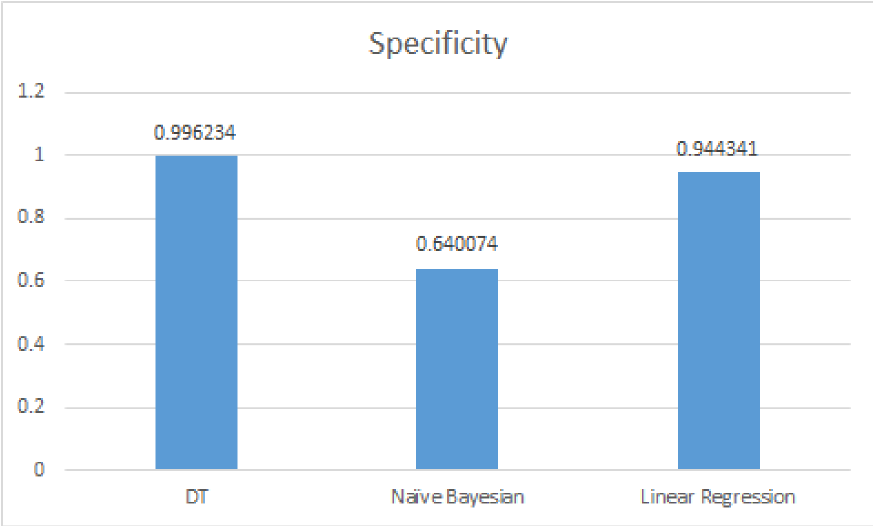

Figure 3.25: Specificity of the Classifiers.

- PPV: The percentage of favourably categorised instances that were actually positive is known as a binary classifier's positive predictive value (PPV). To calculate specificity, use the following formula: TP/(TP+FP).

Classifier	**TP**	**FP**	**PPV**
Decision Tree	479	2	0.995841996
Naive Bayesian	461	194	0.703816794
Linear Regression	461	30	0.938900204

Table 3.6: Positive Predictive Value (PPV) of the Classifiers

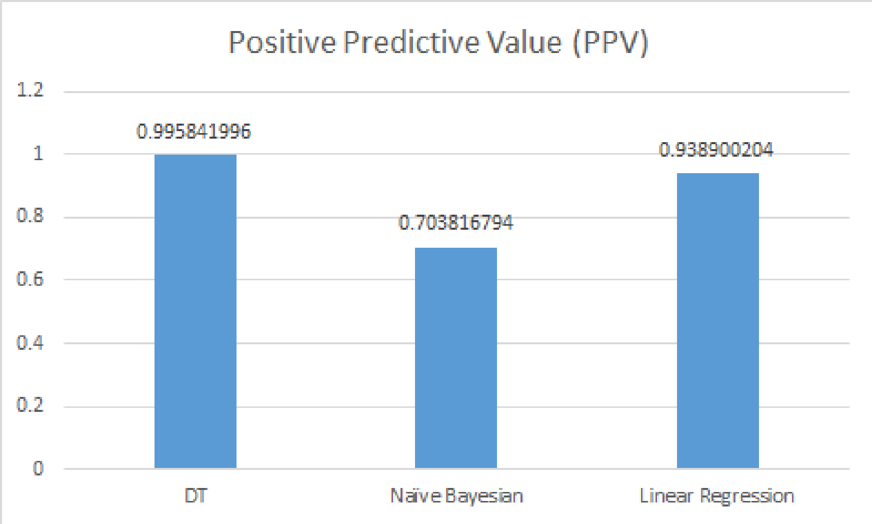

Figure 3.26: PPV of the Classifiers.

- NPV: The likelihood that a network is not actually under an attack after receiving a negative test result is known as the negative predictive value. To calculate specificity, use the following formula: TN/(TN+FN).

Classifier	**TN**	**FN**	**NPV**
Decision Tree	529	10	0.981447124
Naive Bayesian	345	20	0.945205479
Linear Regression	509	10	0.980732177

Table 3.7: Negative Predictive Value (NPV) of the Classifiers

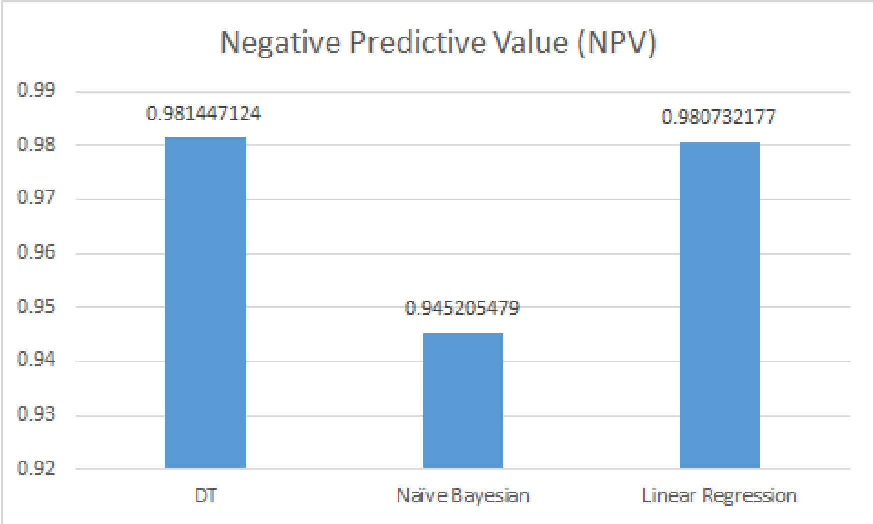

Figure 3.27: NPV of the Classifiers.

- Balanced Accuracy: Balanced accuracy is a metric we can use to assess the performance of a classification model. It is calculated as: Balanced accuracy = (Sensitivity + Specificity) / 2

Classifier	**Sensitivity**	**Specificity**	**Balanced Accuracy**
Decision Tree	0.979550102	0.996233522	0.987891812
Naive Bayesian	0.958419958	0.640074212	0.799247085
Linear Regression	0.978768577	0.944341373	0.961554975

Table 3.8: Balanced Accuracy of the Classifiers

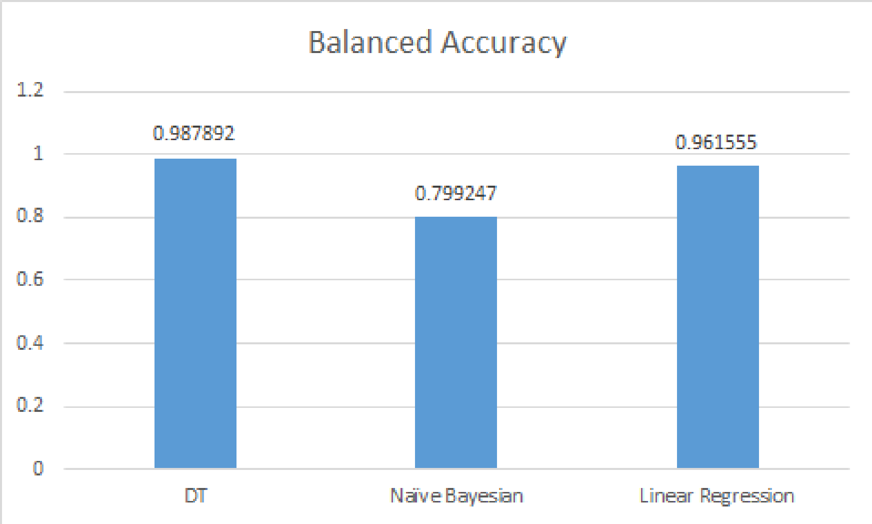

Figure 3.28: Balanced Accuracy of the Classifiers.

For the above formulas , TP means True Positive, FP means False Positive and FN stands for False Negative. When an actual assault occurs and is recognised as such by the ML IDS, the result is a True Positive (TP) result. When IDS considers a normal or non-attack action as an attack activity, an FP results. When the IDS views a harmful action as harmless, then it is False Negative (FN). When the ML IDS classifies a benign action as harmless, a TN happens.

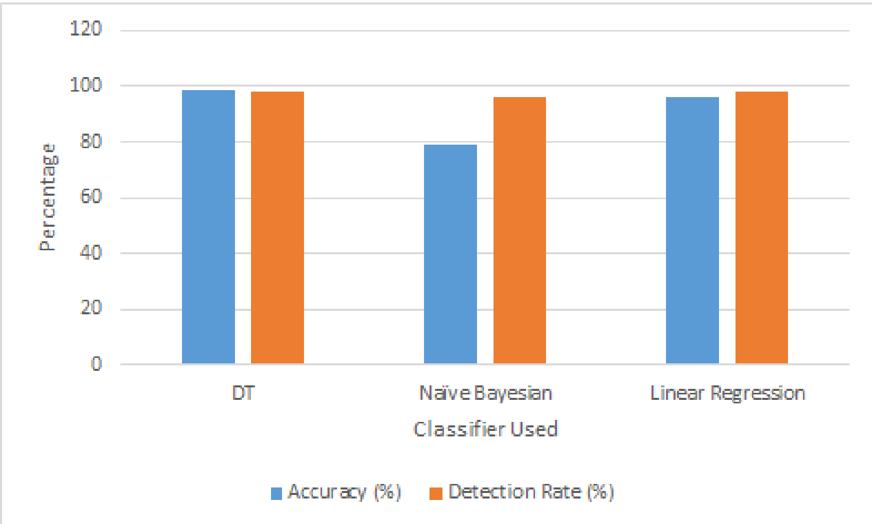

Figure 3.29: Classifiers Performance based on Accuracy and Detection rate.

Figure 3.29 The table above displays the characteristics of classifiers utilised for the proposed IDS's. These characteristics include accuracy and detection rate. It can be see that the three classifier types that are used are able to give encouraging outcomes and results. When compared to other classifiers, the Naive Bayes classifier has a low accuracy of 79.02 percent and a high recall rate of 95.84 percent. When compared to Naive Bayes, Linear Regression which is the other classifier used, it performs better. For a linear regression, the accuracy and recall are 96.04 percent and 97.88 percent, respectively. Finally, there is Decision Tree, which has the best accuracy (98.82 percent) and detection rate (97.96 percent). Based on these findings, we have determined that a decision tree is the best type of classifier.

Chapter 4

Conclusion

WLANs are vulnerable to de-auth attacks due to unencrypted management frames and a lack of authentication mechanisms, which can completely disconnect users from the network. There is a lack of a robust and efficient solution in this field of research. All previous work using machine learning to detect a de-auth attack has chosen a small number of features for machine learning classifiers. As a result, the results are imprecise. n this study, a de-authentication attack on a WLAN network is carried out, and data is collected to generate training and test data for a machine learning classifier. Decision Tree, Naive Bayesian Classifier, and Linear Regression are the three classifiers tested. For De-auth attacks in 802.11 WiFi networks, we proposed a Machine Learning-based Intrusion Detection System. The proposed intrusion detection system detects the De-auth attack with a high detection rate and a low false positive rate. Because both precision and recall exceed 97 percent, the proposed IDS employs the Decision Tree classifier. One significant advantage of the Machine Learning-based IDS is that it does not require any changes to the current protocol, encryption algorithms, or firmware upgrades. Aside from that, the proposed work can be applied to both legacy and modern systems.

References

[1] Saja Mamoori. Wifi networks architecture. 01 2009.

[2] Amal Maamari and Nadir Salih. Study of attacks on wireless networks. pages 21–26, 08 2022.

[3] Patrick-Benjamin Bök, Andreas Noack, Marcel Müller, and Daniel Behnke. *Wireless Networks*, pages 91–115. 07 2020. ISBN 978-3-658-29408-3.

[4] Martin Sauter. Wireless local area network (wlan). pages 465–532, 01 2021.

[5] Kurt Hauser. Wireless local area networks (wlan). 08 2022.

[6] Raj, Jain, and Raj Jain. Wireless metropolitan area networks (wmans). 08 2022.

[7] Anthony Smith and Raymond Hansen. *Wireless Wide Area Networks*, pages 1191–1199. 11 2011. ISBN 9780471784593.

[8] Todor Cooklev. *Standards for Wireless Personal Area Networking (WPAN)*, pages 133–223. 08 2011. ISBN 9780738140667.

[9] Dong Chen. A survey of ieee 802.11 protocols: Comparison and prospective. 01 2017.

[10] Dong Chen. A survey of ieee 802.11 protocols: Comparison and prospective. 01 2017.

[11] Jian Wu. Performance analysis and optimization of ieee 802.11 protocol. *Applied Mechanics and Materials*, 602-605:3515–3517, 08 2014.

[12] Tushar Rakhra, Arjit Kaushal, Sarvesh Tanwar, Priyanka Datta, and Ajay Rana. De authentication attack: A review. pages 1–6, 12 2020.

[13] Raghava K, G. Chandra, and P.N.S.B.S.V V. Optimized de-authentication attack in ieee 802.11 networks. *International Journal of Recent Technology and Engineering*, 7:788–793, 07 2021.

[14] Mohammad Alaa Al-Hamami and Asso.Prof.Dr.Ghossoon Alsadoon. Development of a network-based: Intrusion prevention system using a data mining approach. pages 641–644, 01 2013.

[15] Smriti Pandya Smita Parte. A comprehensive study of wi-fi security – challenges and solutions. *INTERNATIONAL JOURNAL OF SCIENTIFIC ENGINEERING RESEARCH*, 3, 2012.

[16] Hassana Ganame Lusekelo Kibona. Wireless network security: Challenges, threats and solutions. a critical review. *International Journal of Academic Multidisciplinary Research (IJAMR)*, 2(4):19–27, 2018.

[17] Deepika Dhiman. Wlan security issues and solutions. *IOSR Journal of Computer Engineering*, 16:67–75, 01 2014.

[18] Saurabh Malgaonkar, Rohan Patil, Aishwarya Rai, and Aastha Singh. Research on wi-fi security protocols. *International Journal of Computer Applications*, 164: 30–36, 04 2017.

[19] Abdallah, Shukor Razak, and Coulibaly Yahaya. Detection and prevention of denial of service attacks (dos) in wlans infrastructure. *Journal of Theoretical and Applied Information Technology*, 71:417–423, 01 2015.

[20] Angelos Stavrou Constantinos Kolias, Georgios Kambourakis and Stefanos Gritzalis. Intrusion detection in 802.11 networks: Empirical evaluation of threats and a public dataset. *IEEE Communications Surveys and Tutorials*, 18(1):184–208, 2016.

[21] Suroto Suroto. Wlan security: Threats and countermeasures. *JOIV : International Journal on Informatics Visualization*, 2, 2018.

[22] Suroto Suroto. Wlan security: Threats and countermeasures. *JOIV : International Journal on Informatics Visualization*, 2, 06 2018.

[23] M. Chan Aung and K. Thant. Detection and mitigation of wireless link layer attacks. *2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA)*, pages 173–178, 2017.

[24] Sudeshna Chakraborty, Maliha Khan, Amrita, Preeti Kaushik, and Zia Nasseri. *An Extensive Review of Wireless Local Area Network Security Standards*, pages 591–604. 01 2021. ISBN 978-981-16-3066-8.

[25] V. Voskoboinyk R. Korolkov, S. Kutsak. Analysis of deauthentication attack in ieee 802.11 networks and a proposal for its detection. *Mathematical modeling. Information technology. Automated control systems*, (50), 2021.

[26] John Bellardo and Stefan Savage. 802.11 denial-of-service attacks: Real vulnerabilities and practical solutions. *Proceedings of 12 USENIX Security Symposium*, pages 2–2, 08 2003.

[27] Ananay Arora. Preventing wireless deauthentication attacks over 802.11 networks. *Creative Commons, CC BY-NC-SA 4.0*, 2013.

[28] Edgar D Cardenas. Mac spoofing – an introduction. *GIAC Security Essentials Certification (GSEC), GIAC Security Essentials Certification (GSEC)*, 2003.

[29] Petr Stepanov, Galina Nikonova, Tatyana Pavlychenko, and Anatoly Gil. Attack on the address resolution protocol. pages 1–3, 11 2020.

[30] Thuc Nguyen, Duc Nguyen, Bao Tran, Hai Vu, and Neeraj Mittal. A lightweight solution for defending against deauthentication/disassociation attacks on 802.11 networks. pages 185–190, 08 2008.

[31] Fanglu Guo and tzi-cker Chiueh. Sequence number based mac address spoof detection. pages 309–329, 09 2005. ISBN 978-3-540-31778-4.

[32] Haidong Xia and José Brustoloni. Detecting and blocking unauthorized access in wi-fi networks. volume 3042, pages 795–806, 05 2004. ISBN 978-3-540-21959-0.

[33] Farhan Anjum, S. Das, Praveen Gopalakrishnan, L. Kant, and Byungsuk Kim. Security in an insecure wlan network. pages 292 – 297 vol.1, 07 2005.

[34] Mayank Agarwal, Santosh Biswas, and Sukumar Nandi. Detection of deauthentication denial of service attack in 802.11 networks. pages 1–6, 12 2013.

[35] Sukumar Nandi Mayank Agarwal, Santosh Biswas. Detection of de-authentication dos attacks in wi-fi networks: A machine learning approach. *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 2015.

[36] Weijia Wang and Haihang Wang. Weakness in 802.11w and an improved mechanism on protection of management frame. *2011 International Conference on Wireless Communications and Signal Processing (WCSP)*, 2011.

[37] Cisco. Frequently asked questions about management frame protection (mfp). *Document ID:SMB5442 (https://www.cisco.com/c/en/us/support/docs/smb/wireless/cisco-small-business-wireless-access-points/smb5442-frequently-asked-questions-about-management-frame-protection.htmlq1)*, 2017.

[38] Jameson Blandfort. Wireless lan security, policy, and deployment best practices. *Cisco, CCIE 27687*, 2011.

[39] Shweta Sharma and Meenakshi Mittal. Detection and prevention of de-authentication attack in real-time scenario. *International Journal of Innovative Technology and Exploring Engineering*, 8:3324–3330, 08 2019.

[40] Jaspreet Kaur. Mac layer management frame denial of service attacks. pages 155–160, 09 2016.

[41] K.M.Pattani Deep Joshi, Dr. Ved Vyas Dwivedi. De-authentication attack on wireless network 802.11i using kali linux. *International Research Journal of Engineering and Technology (IRJET)*, (Vol. 4):pp. 1666–1669, 2017.

[42] Korolkov R.Y. and Kutsak S.V. The features of a deauthentication attack implementation in networks 802.11. *Ukrainian Information Security Research Journal*, (vol. 21, no. 3):pp. 175–181,, 2019.

[43] J. M. Patel A. Kumar, J. Naughton and X. Zhu. To join or not to join?: Thinking twice about joins before feature selection. *SIGMOD*, page pp .19–34, 2016.

[44] Dewi C. Huang SW. et al Chen, RC. Selecting critical features for data classification based on machine learning methods. *J Big Data 7*, 52 (2020).

[45] M. Anila and Pradeepini Gera. Study of prediction algorithms for selecting appropriate classifier in machine learning. *Journal of Advanced Research in Dynamical and Control Systems*, 9:257–268, 2017.

[46] Andreas Janecek, Wilfried Gansterer, Michael Demel, and Gerhard Ecker. On the relationship between feature selection and classification accuracy. *Proceedings of Machine Learning Research*, 4:90–105, 2008.

[47] H. Liu M. Dash. Feature selection for classification. *Intelligent Data Analysis*, 1, Issue:131–156, 1997.

[48] Rehman SU et al. Asim S, Muhammad H. A comparative study of feature selection approaches: 2016–2020. *Int J Sci Eng Res*, 1, Issue:11:469–78, 2020.

[49] Md Asaduzzaman, Mohammad Majib, and Md. Mahbubur Rahman. Wi-fi frame classification and feature selection analysis in detecting evil twin attack. 2020.

[50] Kulothungan K. Muthurajkumar S. et al Ganapathy, S. Intelligent feature selection and classification techniques for intrusion detection in networks: a survey. *J Wireless Com Network*, 2013.

[51] Brent Salisbury. What are ethernet, ip and tcp headers in wireshark captures. http://networkstatic.net/what-are-ethernet-ip-and-tcp-headers-in-wireshark-captures/#:~:text=The%20minimum%20size%20header%20is,of%20options%20in%20the%20header., 2008.

[52] Steven Euijong Whang Yuji Roh, Geon Heo. A survey on data collection for machine learning. 2019.

[53] Alceu Jr and Luiz Soares de Oliveira. Dynamic selection of classifiers—a comprehensive review. *Pattern Recognition*, 47:3665–3680, 2014.

[54] A. Elisseeff I. Guyon. An introduction to variable and feature selection. *J. Mach. Learn. Res.*, 3:pp. 1157–1182, 2003.

Appendix A

Appendix

A.1 Naive Bayesian Classifier Code

```
import pandas as pd
import os
from sklearn.naive_bayes import GaussianNB

#Create a Gaussian Classifier
df = pd.read_csv("E:\\train-deauth.csv")
X = df.iloc[:,:15]
y = df.iloc[:,15]
model = GaussianNB()
model.fit(X, y)
print(model)


import numpy as np
import csv
with open('E:\\test-deauth.csv') as csv_file:
    csv_reader = csv.reader(csv_file, delimiter=',')
    counter=0
    for row in csv_reader:
        z=row
```

```
        z = np.dtype('float64').type(z)
        x=model.predict([z[:15]])
        print (x , "+" , z[15])
        if ((x == [1] and z[15] == 0)) :
            counter = counter + 1
    print(counter)
```

A.2 Regression Code

```
import pandas as pd
from sklearn.linear_model import LinearRegression

import os
from sklearn import svm

#Create a Gaussian Classifier
df = pd.read_csv("E:\\train-deauth.csv")
X = df.iloc[:,:15]
y = df.iloc[:,15]
model = LinearRegression()
model.fit(X, y)

import numpy as np
import csv
with open('E:\\test-deauth.csv') as csv_file:
    csv_reader = csv.reader(csv_file, delimiter=',')
    counter=0
    for row in csv_reader:
        z=row
        z = np.dtype('float64').type(z)
        x=model.predict([z[:15]])
        print (x , "+" , z[15])
        if ((x >= 0.5 and z[15] == 0)) :
```

```
            counter = counter + 1
    print(counter)
```

A.3 Decision Tree Code

```
import pandas as pd
import seaborn as sn
import matplotlib.pyplot as plt
df = pd.read_csv("E:\\test-deauth.csv")

X = df.iloc[:,:15].astype(int)
y = df.iloc[:,15].astype(int)

corrMatrix = df.corr()
sn.heatmap(corrMatrix, annot=True)
plt.show()

#Training
from sklearn import tree
clf = tree.DecisionTreeClassifier()
clf = clf.fit(X, y)

#Visualizer
import os
import graphviz
os.environ["PATH"] += os.pathsep + "C:\\Program Files (x86)\\Graphviz2.38\\bin"
tree.plot_tree(clf)
dot_data = tree.export_graphviz(clf, out_file=None)
graph = graphviz.Source(dot_data)
graph.render(filename="E:\\DT.dot")

# Prediction
import numpy as np
```

```
import csv
from sklearn.metrics import classification_report, confusion_matrix
import pandas as pd

df = pd.read_csv("E:\\train-deauth.csv")

X_test = df.iloc[:, :15].astype(int)
y_test = df.iloc[:, 15].astype(int)
y_pred = clf.predict(X_test)
from sklearn.metrics import classification_report, confusion_matrix

print(confusion_matrix(y_test, y_pred))
print(classification_report(y_test, y_pred))
```