# A Proactive Network Forensics Analysis Using Wireshark and Artificial Neural Network Model

## Mohammed AlQattan<sup>1</sup>

Department of Computer Networks and Communications, College of Computer Science and Information Technology, King Faisal University, P.O. Box 400 Al-Ahsa 31982, Saudi Arabia 217019040@student.kfu.edu.sa

Network forensics is a complex domain that demands a deep understanding of the available practical tools and properties to have an appropriate experience of how network forensics tools and methods can be employed to detect dangerous traffic. Wireshark is an effective open-source tool for sniffing packets and operates with many operating systems like Windows, macOS, and Linux. Many features help users sniff packets easier with its easy-to-use Graphical User Interface (GUI) or Command Line Interface (CLI), i.e., advanced filtering capabilities. This paper employed Wireshark to analyze the packets by simulating several attacks scenarios within protocols among some applications, i.e., Secure Shell Hash (SSH)/Telnet, Voice Over Internet Protocol (VoIP), WhatsApp, and Address Resolution Protocol (ARP) Spoofing. This study highlighted Wireshark's ability to analyze malicious behavior for information via network traffic. The data collection of 10000 benign and malicious instances was exported through Wireshark, which simulates education network traffic. The experiments employed an artificial neural network model to detect the network behavior, and the results showed significant accuracy in detecting potential network threats.

Keywords: NetworkForensics, Wireshark, Proactive detection, Artificial Neural NetworkAttacks Scenarios.

#### I. INTRODUCTION

Network forensics is a complicated field that requires deep insight into the available powerful tools and properties to have a reasonable understanding of how Network forensics tools and methods can be utilized to detect harmful behavior [1].

Wireshark is one of the network forensics softwarethat analyzes packets on the network traffic. It is also one of the tools that can show the live traffic of the network in actual real-time. Previously, it was known as Ethereal, developed by Gerald Combs in 1988 [2].It is considered one of the oldest tools used in network traffic, analyzer, troubleshooting, etc. Wireshark can run on many operating systems, such as Windows and Unix. Note that Unix can be available in different versions. Wireshark supports various protocols and uses Graphical User Interface (GUI) [3]. It supports not only

## Heider A. M. Wahsheh<sup>2\*</sup>

Department of Information Systems, College of Computer Science and Information Technology, King Faisal University, P.O. Box 400 Al-Ahsa 31982, Saudi Arabia \*Corresponding author:<u>hwahsheh@kfu.edu.sa</u>

GUI; that is why Wireshark is a powerful tool among many. It has unlimited access to the network traffic and provides many statistics, filters, protocols, etc. Later, we will discuss different protocols, functionalities, and attacks that could happen, like Address Resolution Protocol (ARP) spoofing or poisoning [4]. Capturing packets on Wireshark could be imported or exported as aPacket Capture(pcap) file, a tcpdump file, etc. Also, it can use one of the available interfaces on the Personal Computer (PC). Within a PC, it can have Network Interface Card (NIC), which is an interface that allows you to connect to the internet. For example, a wireless interface is widely used these several days to connect to the internet, as each NIC is unique from a different device [5]. Because each NIC contains a physical address, and the Media Access Control (MAC) address is uniquely identified for each device worldwide, MAC cannot be the same address for two devices. Note that the MAC address is 48 bits long. To sniff the network traffic, the user can select any NIC available within the PC, such as wireless, ethernet, and others. When the packets arrive at NIC and are accepted, we can say it is set into promiscuous mode. Promiscuous mode is one of the brilliant modes to allow packets to be copied to the memory. That is why Wireshark is so valuable for storing the packets in the memory [2-5].

This paper utilized Wireshark to investigate the packets by simulating several attacks scenarios within protocols among some applications, i.e., Secure Shell Hash (SSH)/Telnet, Voice Over Internet Protocol (VoIP), WhatsApp, and Address Resolution Protocol (ARP) Spoofing. Involving Wireshark in those scenarios implies the real practice behavior of each application and its protocols. This study highlighted Wireshark's ability to analyze malicious behavior for the information via network trafficand allow systems to countermeasure suspicious occurrences quickly. Moreover, a data collection of 10000 benign and malicious instances was exported through Wireshark, which simulates education network traffic. The experiments employed an artificial neural network model to detect the network behavior, and the results



showed significant accuracy in detecting potential network threats.

The rest of this paper is organized as follows: Section 2illustrates how Wireshark works. Section 3 simulates several attack scenarios within protocols among some popular applications. Section 4 explores the artificial neural network model and the performance evaluation. Section 5 draws the concluding remarks and discussion.

#### II. HOW DOES WIRESHARK WORK?

Wireshark has a lot of protocols to capture within a single NIC. Note that Wireshark is not an Intrusion Detection System (IDS) [2]. That is because this software will not warn you if any attack happens. Wireshark is just a sniffer that views all traffic from the NIC that has been collected. In this section, we will discuss the features of the Wireshark GUI. The user should know the known protocols because they are the powerful understanding of the Wireshark environment, such as Transmission Control Protocol (TCP), User Datagram Protocol (UDP), Real-Time Protocol (RTP), etc.

All the information is valuable in Wireshark and can support fixing any issue on the network or understanding how the network traffic works [6]. International Organization for Standardization (ISO) has divided the network into seven layers called the Open System Interconnected (OSI) model. Those seven layers are the application, presentation, session, transport, network, data link, and physical, presentation, session, transport, network, network, data link, and physical layers. As shown in Figure 1, those are the seven layers starting with the physical layer from the bottom and ending with the application layer and vice versa. Those seven layers are essential when it comes to Wireshark. The Wireshark can decode the packets into those seven layers. Each layer is specified in the application based on the contents regarding of OSI model in the network traffic [7].

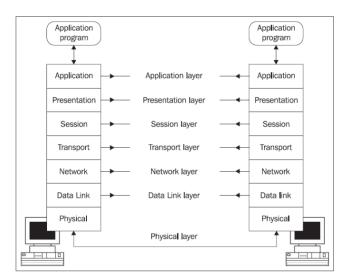


Figure 1: OSI reference model using different layers.

## III. TYPES OF POSSIBLE ATTACKS

This section will discuss several possible attacks that could happen in the educational networktraffic using Wireshark. We know that Wireshark is not an Intrusion Detection System (IDS) because it does not warn the user if a malicious attack ishappening to the user's PC [7]. Without further ado, let's discuss some of the possible attacks that could be seen in Wireshark. Those possible attacks use remote session protocols such as Telnet or Secure Shell Hash (SSH), Voice over Internet Protocol (VoIP), WhatsApp application chat, and Address Resolution Protocol (ARP) spoofing [8].

## A. Telnet vs. SSH

Telnet and Secure Shell Hash (SSH) are remote sessions to let us establish a connection from the remote site [9]. For example, a user located in Germany can access a router or any other device that is in France via a remote session like SSH or Telnet. Telnet does not provide any encryption for the data. However, SSH does provide encryption for the data to exchange between two computers in the network [10]. Both protocols can be used in the router, switch, or any remote connection to a device. In Figure 2, we can find that two routers are connected through an ethernet link using the GNS3 application. The first router on the left, called R1, has an IP address of "10.1.1.1" with a mask of "/30" or "255.255.255.252". For the second router has the IP address of "10.1.1.2" with a mask of "/30" or "255.255.255.252". The mask is "/30" or "255.255.255.252" is optimum because we need only 2 hosts to be connected within those devices. This means that R1 and R2 routers are the only hoststhat are connected, which means that it does eliminate the waste of IP addresses.

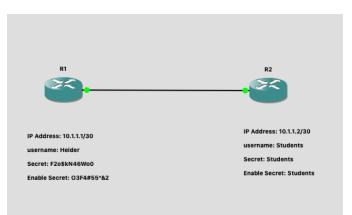


Figure2: Using Telnet and SSH through two routers via GN3 application.

SSH is a more secure remote connection, it uses Rivest-Shamir-Adelman (RSA) algorithm as public-key encryption [10]. Telnet does not use any kind of encryption. That is why



it is not appropriate to use. In practice, we use RSA as the 2048 bit key to establishing the SSH protocol. Figure 3 shows that the data is encrypted because SSH uses the RSA algorithm to encrypt the packet. However, Telnet does not use any type of algorithm to encrypt the packet. That is why it is recommended to use SSH protocol, as seenin the comparison between SSH and Telnet in Table 1. Telnet is not using any data encryption, making SSH stronger and better. So, it is hard to decrypt the packets if there is a man-in-the-middle attack. It also uses public-key cryptography, which uses public and private keys. The man-in-the-middle attack will find that the decryption is not easy to crack the SSH encryption [9-10].

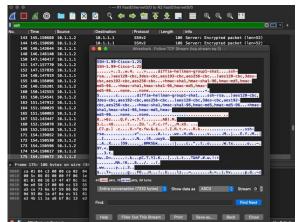


Figure 3: Encrypted data using remote SSH connection in Wireshark.

| Type of Protocol | SSH          | Telnet |  |  |  |
|------------------|--------------|--------|--|--|--|
| Encryption       | RSA          | None   |  |  |  |
| Version          | SSHv1, SSHv2 | Telnet |  |  |  |
| Cryptography     | Public key   | None   |  |  |  |

## Table 1: Comparison between SSH and Telnet protocols.

#### B. VoIP

Voice over Internet Protocol (VoIP) is a voice communication over IP addressescalled IP telephony [11]. This technology allows voice communication over the internet instead of a standard phone. People benefited from the cost of switching from regular phones to VoIP. VoIP costs less than a traditional phone because regular phones reserve a large amount of bandwidth. However, VoIP is used to share the bandwidth among all users connected to the internet. VoIP is the most effective way to deliver voice to other users [11-12].

Moreover, VoIP transmits voice data in packets through the internet, making it a reliable and cost-effective protocol. Besides, the regular phone uses Public Switched Telephone Network (PSTN). PSTN is known as a traditional circuitswitched telephone network. However, VoIP is known as packet switching. Regular phones carry voice communication over analog signals, making it expensive for the Internet Service Provider (ISP). On the other hand, digital signals like VoIP make it cost-effective to use.

We will focus on the attack that could happen in VoIP using Wireshark.Line phone is an open source using messages, video, and VoIP. Line phone uses open standards like Session Initiation Protocol (SIP) and Real-Time Protocol (RTP). Toallow voice communication, SIP and RTP must work together. In addition, the Line phone is available in IOS, Android, macOS, Windows, and Linux. It provides a graphical user interface, enhanced instant messaging, audio, video, and secure communication. Although it says end-to-end encryption between users, the application provides no encryption by default. So, we have tested two users to figure that out through Wireshark. After the testing, we found thatthe voice communication was so evident between the two users, which indicates that any man-in-the-middle attack could collect the voice conversation from end to end. Thus, losing trust in this application. Encryption options exist in the Line phone application setting, such as SRTP, ZRTP, and DTLS. However, the choice is selected "None" by default. Not every user has experience with how encryption work. An application like Line phone or any other VoIP application needs to provide full encryption for end users by default and not without encryption. Data encryption is essential to prevent security information between end users.

## C. WhatsApp

WhatsApp is a phone chat application used by iOS and Android. WhatsApp is used by more than two billion people around the world. Many people use it as a secure and reliable application [13]. We will discuss the possible attack that could happen if there were a man-in-the-middle attack. We will sniff the traffic using Wireshark. Before we discuss it practically, WhatsApp provides end-to-end encryption between two or more end-users. Let us see what the practical tells us. In this example, we have tested two users to communicate with each other. In Figure 4, the communication between two end users shows encrypted data in Wireshark. Thus, it indicates that WhatsApp is a trusted application thatend-to-end users can use.



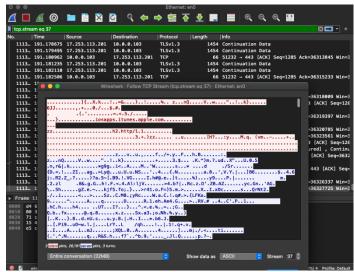


Figure 4: Encrypted messages between two end-user traffic from WhatsApp using Wireshark.

#### D. ARP Spoofing

Address Resolution Protocol (ARP) spoofing is a protocol that is vulnerable to attack by hackers to see the user's data information by sniffing the packet using Wireshark or any other sniffing tool [14]. ARP contains IP address and Media Access Control (MAC) address together. When a user wants to initiate a communication with another user within the local network for the first time, the user must send an ARP request containing the IP address of the end device with a broadcast MAC address "ff:ff:ff:ff:ff:ff". In Figure 5, when the end device has match of the IP address that the user had requested, the end device will reply, "this is my IP address, and I will give you my MAC address". Then, the user will put the end device's IP address and MAC address to its ARP cache to be saved and not apply ARP every time to the end device. The ARP spoofing is taking advantage of sending a reply message called gratuitous ARP to its corresponding users within the same network. Moreover, it takes advantage of the ARP cache to change the router's MAC address to its MAC address for the user's ARP cache. Without further ado, let's get into the practical work.



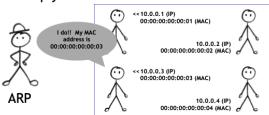


Figure 5: ARP reply to get the MAC address of a user.

We have tested two computers in this example; one computer is acting as a user using Windows operating system. The other computer works as a hacker using the Linux Kali operating system, intending to sniff the packets from the user. We have used an application called Ettercap [15]. This application is used in Unix operating systems to be able to use the ARP spoofing attack on the user. First, the user will have a stable connection to the router and its corresponding MAC address when the hacker uses the Ettercap application to target the user via sending a gratuitous ARP. The user Windows machine ARP cache of the router will change its MAC address to the hacker's MAC address. Of course, the user will have no idea that the packets will reach the hacker. The user will think that the packets are reaching the router because, in the ARP cache, it appears the router's IP address. Although the user sees the router's IP address in the ARP cache of the Windows operating system, the MAC address is not the router's MAC address. More likely, it is the hacker's computer MAC address. Using Wireshark, we can see that in Figure 6,all the packets from the user with an IP address of "192.168.1.3" hits the hacker device traffic first before going to the router with an IP address of "192.168.1.1". Even though ARP spoofing looks s fantastic way to sniff the packets, it is not the optimum technique to attack and look for the packets. ARP spoofing solution is easy to solve it. Instead of making the ARP cache of the user set to dynamic to get the MAC address of the router or any other device within the Local Area Network (LAN).The user needs to manually type the MAC address of the devices in the LANon its cache to change from dynamic to static mode. With this solution, the hacker's ARP spoofing attackwill be useless to the user.

|       | Capturing from eth0 _ 🗖 🗙 |       |      |            |      |   |       |     |      |     |      |       |      |      |       |        |         |          |             |      |      |       |      |              |      |      |              |      |
|-------|---------------------------|-------|------|------------|------|---|-------|-----|------|-----|------|-------|------|------|-------|--------|---------|----------|-------------|------|------|-------|------|--------------|------|------|--------------|------|
| File  | Edit                      | V     | ew   | <u>G</u> o | C    | aptu  | re    | Ana | lyze | 2   | stat | istic | 5    | Tele | ph    | ony    | Wire    | eles     | is <u>T</u> | ools | Help | C     |      |              |      |      |              |      |
|       |                           | ł     |      |            |      |   |       | X   | Ć    |     | ٩    | <     |      | =    |       | $\geq$ | 1       | <u>.</u> |             |      |      | ŧ     |      | <b>a</b> , ( | 2    |      | Ē            |      |
| A A   | oply a                    | a dis | olay | filte      | r    | <ct< th=""><th>r -/&gt;</th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th>-</th><th>• +</th></ct<> | r -/> |     |      |     |      |       |      |      |       |        |         |          |             |      |      |       |      |              |      |      | -            | • +  |
| No.   |                           | Time  |      |            |      | Sou   | rce   |     |      |     |      | De    | esti | nat  | ion   |        |         |          | Proto       | col  | Leng | th Ir | nfo  |              |      |      |              | 4    |
|       | 28                        | 18.5  | 114  | 5067       | 3    | 192   | .168  | .1. | 3    |     |      | 19    | 2.   | 168  | .1.   | 255    |         |          | NBNS        |      | 9    | 2 N   | ame  | quer         | y I  | VB W | ORKG         | ROUF |
|       | 29                        | 18.5  | 239  | 3264       | 8    | 172   | .16.  | 148 | .1   |     |      | 17    | 2.:  | 16.  | 148   | . 255  |         |          | NBNS        |      | 9    | 2 N   | ame  | quer         | ÿΙ   | VB W | ORKG         | ROUF |
|       |                           | 19.2  |      |            |      |   | .16.  |     |      |     |      |       |      |      |       | .255   |         |          | NBNS        |      |      |       |      | quer         |      |      |              |      |
|       | 31                        | 19.2  | 478  | 6065       | i1   | 172   | .16.  | 148 | .1   |     |      | 17    | 2.:  | 16.  | 148   | .255   | ;       |          | NBNS        |      | 9    | 2 N   | ame  | quer         | y I  | VB W | <b>ORKG</b>  | ROUF |
|       |                           | 19.2  |      |            |      |   | .168  |     |      |     |      |       |      |      |       | 255    |         |          | NBNS        |      | 9    | 2 N   | ame  | quer         | ÿ I  | VB W | <b>IORKG</b> | ROUF |
|       | 33                        | 20.4  | 082  | 5486       | 8    | 192   | .168  | .1. | 3    |     |      | 19    | 2.   | 168  | .1.   | 1      |         |          | ICMP        |      | 7    | 4 E   | cho  | (pir         | g)   | req  | uest         | ic   |
|       | 34                        | 20.4  | 150  | 9003       | 12   | 192   | . 168 | .1. | 3    |     |      | 19    | 12.  | 168  | .1.   | 1      |         |          | ICMP        |      | 7    | 4 E   | cho  | (pir         | ng)  | red  | uest         | ic   |
|       |                           |       |      |            |      |   | 47:4  | 3:1 |      |     |      |       |      |      |       | ree-   | (for    |          |             |      |      |       |      | Root         |      |      |              |      |
|       | 36                        | 22.0  | 073  | 1891       | .5   | 192   | .168  | .1. | 3    |     |      | 19    | 2.:  | 168  | .1.   | 255    |         |          | BROW        | SER  | 21   | 6 G   | et I | Backu        | ip l | List | Reg          | uest |
|       | 37                        | 22.0  | 126  | 9039       | 4    | 192   | .168  | .1. | 3    |     |      | 19    | 2.   | 168  | .1.   | 255    |         |          | NBNS        |      | 9    | 2 N   | ame  | auer         | νı   | VB W | /ORKG        | ROUF |
|       |                           |       |      |            | 8    |   |       | 3:f |      |     |      |       |      |      | a-t   | ree-   | (for    |          |             |      |      | 0 R   |      | Root         | í =  |      |              |      |
|       | 39                        | 22.7  | 567  | 3021       | 2    | 192   | .168  | .1. | 3    |     |      |       |      |      |       | 255    |         |          | NBNS        |      |      |       |      | quer         |      |      |              |      |
|       | 40                        | 23.5  | 018  | 5446       | )1   | 192   | .168  | .1. | 3    |     |      | 19    | 2.   | 168  | .1.   | 255    |         |          | NBNS        |      |      |       |      | quer         |      |      |              |      |
|       | 41                        |       |      |            |      |   |       | 3:f |      |     |      |       |      |      | a - t |        | (for    |          |             |      |      |       |      | Root         |      |      |              |      |
|       | 42                        | 25.9  | 056  | 2253       | 15   | 192   | .168  | .1. | 3    |     |      |       |      | 168  |       |        |         |          | ICMP        |      |      |       |      | (pir         |      |      |              |      |
|       | 43                        | 25.9  | 111  | 8084       | 19   | 192   | .168  | .1. | 3    |     |      |       |      | 168  |       |        |         |          | ICMP        |      |      |       |      | (pir         |      |      |              |      |
|       |                           | 26.2  |      |            |      | 192   | .168  | .1. | 3    |     |      |       |      |      |       | 255    |         |          | BROW        | ISER |      |       |      | Backu        |      |      |              |      |
|       | 45                        | 26.2  | 625  | 3127       | 4    | 192   | .168  | 1   | 3    |     |      | 19    | 2.   | 168  | 1     | 255    |         |          | NBNS        |      |      |       |      | quer         |      |      |              |      |
|       |                           | 26.7  |      |            |      |   | 47:4  |     |      |     |      |       |      |      |       |        | (for    |          |             |      |      |       |      | Root         |      |      |              |      |
|       |                           | 27.6  |      |            |      |   | .168  |     |      |     |      |       |      |      |       | 255    |         |          | NBNS        |      |      |       |      | quer         |      |      |              |      |
| 4     |                           |       |      |            |      |   |       |     | -    |     |      |       | _    |      |       |        |         |          |             |      | _    |       |      |              | , .  |      |              | Þ    |
| )⊧ Fr | ame                       | 1: 6  | 0 b  | ytes       | on   | Wi  | re (  | 480 | bi   | ts) | , 6  | 9 by  | tes  | s c  | apt   | ured   | (48     | θb       | its)        | on   | inte | rfa   | ce e | eth0,        | ic   | i O  |              |      |
| 0000  | 01                        | 80    | c2 6 | 0 0        | 9 00 | 0c  | 47    | 43  | f8   | 3b  | 01   | 00    | 27   | 42   | 42    |        |         |          | C · ; ·     |      |      |       |      |              |      |      |              |      |
| 0010  | 03                        | 00    | 90 G | 2 0        | 2 30 | 80  | 01    |     |      |     |      | 3b I  |      |      |       |        | <       |          | GC          | 1    |      |       |      |              |      |      |              |      |
|       | 00                        | 00    | 30 G | 1 0        | : 47 | 43  | f8    |     |      |     |      | 00    |      |      |       |        | · · · G | с·       |             |      |      |       |      |              |      |      |              |      |
|       | 02                        | 00    | of 6 | 0 0        | 00   | 00  | 00    |     | 00   |     |      |       |      |      |       |        |         |          | 1           |      |      |       |      |              |      |      |              |      |
|       |                           |       |      |            |      |   |       |     |      |     |      |       |      |      |       |        |         |          |             |      |      |       |      |              |      |      |              |      |
|       |                           |       |      |            |      |   |       |     |      |     |      |       |      |      |       |        |         |          |             |      |      |       |      |              |      |      |              |      |
|       |                           |       |      |            |      |   |       |     |      |     |      |       |      |      |       |        |         |          |             |      |      |       |      |              |      |      |              |      |

Figure 6: Linux Kali runs the Wireshark application to sniff the packets from the user tothe router.

sckots: 47 . Displayed: 47 (100.0%) Profile: Default

## IV. ARTIFICIAL NEURAL NETWORK MODEL AND PERFORMANCE EVALUATION

In this work, we analyze particular education network traffic by Wireshark and export the data into a CSV file to be used within the artificial neural network to develop the proactive detection of malicious behavior on network traffic.



The state of th

Multi-layer perceptron (MLP) is a supplement of a feedforward artificial neural network, consisting of three layers; the input layer, output layer, and hidden layer. The input layer contains the input signal. The output layer executes essential tasks such as prediction, recognition, and classification. Several hidden layers can be placed between the input and output layers, which are the actual computational engine of the MLP. MLPs are developed to approximate any continuous function and deal with not linear problems. MLPs are formed of neurons called perceptions, as shown in Figure 7. a uses n numeric attributes perceptron as input (a = a1, a2, ..., an), and each of these attributes has a weight [16].

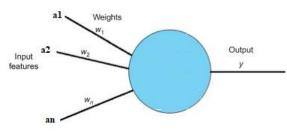


Figure 7: MLP Perceptron with n Input Features.

Figure 8 presents the main structure of MLP with three layers, three input features, four hidden neurons, and one output.

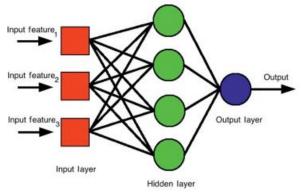


Figure 8: The main structure of the MLP Model.

The data collection is balanced and consists of 10000 instances; 5000 as malicious and 5000 as benign. The extracted features from Wireshark are summarized in Table 2, which corresponds to [17-24].

| Table 2 | . The list of select | ed features |
|---------|----------------------|-------------|
|---------|----------------------|-------------|

| # | FeatureName    | Description                                       |
|---|----------------|---|
| 1 | time           | timeoftrafficcapturing                            |
| 2 | app            | honeypotcapturedthetraffic                        |
| 3 | dest           | destip  |
| 4 | dest_port      | destport  |
| 5 | dionaea_action | eitherDionaeahoneypotacceptorrejecttheconnec tion |

| 6  | direction            | thedirectionofthecapturedtrafficeitherinorout                            |
|----|----------------------|--|
| 7  | eth_dst              | thedest macaddress   |
| 8  | eth_src              | thesourcemacaddress  |
| 9  | host                 | Splunkserveriporhostname   |
| 10 | ids_type             | thetypeoftheusedids  |
| 11 | ip_id                | thepacketid  |
| 12 | ip_len               | packetlength   |
| 13 | ip_tos               | packettypeofservice  |
| 14 | ip_ttl               | packettimetolive   |
| 15 | linecount            | thenumberoflinesofthecaptured traffic                                    |
| 16 | p0f_app              | protocolusedbyP0fforfingerprinting                                       |
| 17 | p0f_link             | theconnectiontypeattheattackersidelikemodemo<br>rdsl                     |
| 18 | p0f_os               | the operating system of the machinegeneratingtheattack                   |
| 19 | p0f_uptime           | howlongsincetheattackingmachineisup                                      |
| 20 | protocol             | tcporudp   |
| 21 | sensor               | idassignedbyMHNperhoneypot   |
| 22 | severity             | severityrankoftheattack  |
| 23 | signature            | thesignatureoftheattackasmatchedbysnort                                  |
| 24 | snort_classification | anumbergiven bysnorttoclassifythetraffic                                 |
| 25 | snort_header         | theruleheader  |
| 26 | snort_priority       | assignsaseverityleveltorules   |
| 27 | source               | inputdatasource(neededbySplunk)  |
| 28 | sourcetype           | inputdatatype(neededbySplunk)  |
| 29 | splunk_server        | Splunkiporhostname   |
| 30 | src                  | attacksrcip  |
| 31 | src_port             | attacksourceport   |
| 32 | ssh_password         | passwordusedbytheattackertryingtogetsshacc ess                           |
| 33 | ssh_username         | usernameusedbytheattackertryingtogetsshacc ess                           |
| 34 | ssh_version          | attackersshclientversion   |
| 35 | tcp_flags            | indicate a particular connection state<br>orprovideadditionalinformation |
| 36 | tcp_len              | packetlength   |
| 37 | timeendpos           | atwhichbyteintotheeventthetimestampends                                  |
| 38 | timestartpos         | atwhichbytethetimestampstarts  |
| 39 | transport            | transportprotocoltypetcporudp  |
| 40 | type                 | honeypoteventtype  |
| 41 | udp_len              | packetlength   |
| 42 | vendor_product       | nameofthehoneypotthatcapturesthetraffic                                  |
| 43 | raw                  | raw(not parsed)event   |

We adopted a 10-fold cross-validation in which the set of samples of the data collection is divided into ten elements. Nine are used for training, and one for testing. A crossvalidation procedure is duplicated ten times to test all ten elements.



We calculated the accuracy, precision (P), recall (R), and Fmeasure (F-M) using the prediction quality measurements; True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN), as shown in formulas 1-4 [25].

$$Accuracy_{i} = \frac{TP + TN}{TP + FP + TN + FN}_{(1)}$$
$$Recall_{i} = \frac{TP}{TP + FN}$$

(2)

$$Precision_{i} = \frac{TP}{TP + FP}$$

$$F - measure = \frac{(2 \times TP)}{(2 \times TP) + FP + FN}$$
(3)
(3)
(3)
(4)

When evaluating the model, we explore the highest values of TP, precision, recall, and F-measure. On the other hand, the FP rates should be lowest toward the optimal value, which is zero [25].

MLP model yielded an accuracy result of 92.56% with an error rate of 7.4%; Table 3 explores the detailed results.

| Class         | TP Rate | FP Rate | Precision | Recall | F-M   |
|---------------|---------|---------|-----------|--------|-------|
| Malicious     | 0.861   | 0.010   | 0.988     | 0.861  | 0.920 |
| Benign        | 0.990   | 0.139   | 0.877     | 0.990  | 0.930 |
| Weighted Avg. | 0.926   | 0.074   | 0.933     | 0.926  | 0.925 |

Table 3. Detailed Accuracy Results for SVM Model.

As shown in Table 3, the MLP yielded optimal prediction results for benign behavior with an accuracy of 99%. The accuracy result for malicious detection was 86.1%, which shows the considered detection accuracy result.

A confusion matrix is a table design that visualizes an algorithm's performance, typically in supervised learning models. Every row presents the actual instances class label in the confusion matrix, while every column shows the instances' predicted class label.

Among the 10000 instances, the MLP model correctly classified 9256cases and failed to classify 744cases to their correct class label, as shown in Table 4.

Table 4. Confusion Matrix of MLP Model.

|              | Predicted Class |           |  |  |  |  |  |  |
|--------------|-----------------|-----------|--|--|--|--|--|--|
| Actual Class | Benign          | Malicious |  |  |  |  |  |  |
| Benign       | 4949            | 51        |  |  |  |  |  |  |
| Malicious    | 693             | 4307      |  |  |  |  |  |  |

The results showed that the MLP model could be utilized as an early warning proactive tool for predicting malicious behavior on network traffic.

#### V. CONCLUSION AND DISCUSSION

1015

In conclusion, we find that Wireshark distinguishes between several packets through the network traffic, whether it is encrypted or unencrypted. Wireshark considers one of the most powerful tools used to sniff network traffic. The Wireshark is still the dominant sniffing tool in Linux, macOS, and Windows operating systems. The features of GUI helped us a lot with specific possible attacks. The possible attacks we have implemented in the practical work are SSH/Telnet, VoIP, WhatsApp, and ARP Spoofing. Each possible attack has its mechanism work and how we found the data encrypted through the Wireshark. In SSH/Telnet, we saw that SSH uses encryption to encrypt the data. However, Telnet does not have any encryption. Thus, the data is obvious to see. VoIP is one of the brilliant network features that has voice calls over the network. We had the test on the application that supported VoIP called Line phone. We found that by default phone does not provide any encryption, which makes it untrusted. In Wireshark, the voice is evident through the network traffic. Then, we tried to sniff the packets from the WhatsApp application.

We found that the data is encrypted end-to-end using the Wireshark sniffing tool. We also showed the ARP spoofing using Linux Kali as a hacker and a Windows user. We found that the Linux Kali had changed the ARP cache of Windows users to allow all the traffic from Windows to the Linux Kali instead of going to the router directly. The solution was to use a static rather than dynamic ARP cache. Moreover, this work employed Wireshark to analyze network packets among various attack scenarios. A dataset of 10,000 several packets was collected and labeled each instance as malicious or benign. An MLP model was used to detect network behavior proactively, and the results were promising to allow network administrators to countermeasure the suspicious events in the network traffic.

#### ACKNOWLEDGMENT

This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. GRANT980]'.

#### REFERENCES

- [1] A. Bijalwan."Network Forensics: The Privacy and Security". Chapman and Hall/CRC, 2021.
- [2] V. Jain, "Wireshark Fundamentals: A Network Engineer's Handbook to Analyzing Network Traffic", Apress; 1st ed. edition (March 4, 2022), pp. 1-276
- [3] S. Qureshi, S. Tunio, F. Akhtar, A. Wajahat, A. Nazir, F. Ullah. "Network Forensics: A Comprehensive Review of Tools and

Techniques", International Journal of Advanced Computer Science and Applications, Vol. 12, No. 5, 2021.

- [4] I. J. Okonkwo, I. D. Emmanuel. "Comparative study of EIGRP and OSPF protocols based on network convergence", International Journal of Advanced Computer Science and Applications, Vol; 11, No. 6, 2020.
- [5] S. Qureshi, G. Das, S. Tunio, F. Ullah, A. Nazir, A. Wajahat. "Performance Analysis of Open Source Solution" ntop" for Active and Passive Packet Analysis Relating to Application and Transport Layer", International Journal of Advanced Computer Science and Applications. Vol. 10, No. 3, 2019.
- [6] L. Bock."Learn Wireshark: Confidently navigate the Wireshark interface and solve real-world networking problems". Packt Publishing Ltd, 2019.
- [7] M. G. M.Santos, P. A. A.Marcillo."Security in the data link layer of the OSI model on LANs wired Cisco". Journal of Science and Research, Vol. 3, (CITT2017), pp. 106-112, 2018.
- [8] T. Reisinger, I. Wagner, E. Boiten. "Security and privacy in unified communication". ACM Computing Surveys (CSUR), Vol. 55, No. 3, pp. 1-36, 2022.
- [9] G. Fahrnberger. "Realtime Risk Monitoring of SSH Brute Force Attacks". In International Conference on Innovations for Community Services (pp. 75-95). Springer, Cham. 2022.
- [10] J. Lee, H. Lee. "An SSH predictive model using machine learning with web proxy session logs". International Journal of Information Security, Vol. 21, No. 2, pp. 311-322, 2022.
- [11] V. Mahor, R. Padmavathy, S. Chatterjee. "Chebyshev chaotic mapbased efficient authentication scheme for secure access of VoIP services through SIP". International Journal of Security and Networks, Vol. 17, No. 1, pp. 39-47, 2022.
- [12] G. Bella, P. Biondi, S. Bognanni. "Attacking and Protecting Network Printers and VoIP Phones alike". arXiv preprint arXiv:2202.10832, 2022.
- [13] F. Karpisek, I. Baggili, F. Breitinger."WhatsApp network forensics: Decrypting and understanding the WhatsApp call signaling messages". Digital Investigation, Vol. 15, pp. 110-118, 2015.
- [14] N. Ahuja, G. Singal, D. Mukhopadhyay, A. Nehra."Ascertain the efficient machine learning approach to detect different ARP attacks". Computers and Electrical Engineering, Vol. 99, 107757, 2022.

- [15] Z. Trabelsi, Z. "IoT based Smart Home Security Education using a Hands-on Approach". In 2021 IEEE Global Engineering Education Conference (EDUCON), pp. 294-301,202.
- [16] M. Jassim, G. Coskuner, M. Zontul. "Comparative performance analysis of support vector regression and artificial neural network for prediction of municipal solid waste generation". Waste Management & Research, Vol. 40, No. 2, pp. 195-204, 202.
- [17] N. Moustafa, J. Slay. "The significant features of the UNSW-NB15 and the KDD99 data sets for network intrusion detection systems". In 2015
   4th international workshop on building analysis datasets and gathering experience returns for security (BADGERS) (pp. 25-31). IEEE, 2015.
- [18] A. Alshaibi, M. Al-Ani, A., Al-Azzawi, A., Konev, A., Shelupanov. "The Comparison of Cybersecurity Datasets". Data, 7(2), 22, 2022.
- [19] J. Fuhr, F. Wang, Y. Tang. MOCA."A Network Intrusion Monitoring and Classification System". Journal of Cybersecurity and Privacy, 2(3), 629-639., 2022.
- [20] H. Alshalah, H. Wahsheh. "Artificial Intelligence Model for Network Security Analysis". NeuroQuantology, Vol. 20, No. 13, pp. 2036-2044, 2022.
- [21] H. Ahmed, A. Hameed, N. Bawany. "Network intrusion detection using oversampling technique and machine learning algorithms". PeerJ Computer Science 8:e820, 2022.
- [22] M., Alani. "Implementation-Oriented Feature Selection in UNSW-NB15 Intrusion Detection Dataset". In International Conference on Intelligent Systems Design and Applications (pp. 548-558). Springer, Cham., 2022
- [23] M. Sarhan, S. Layeghy, M. Portmann. "Feature Analysis for Machine Learning-based IoT Intrusion Detection".2022,
- [24] A. Mahfouz, A. Abuhussein, F. Alsubaei, S. Shiva. "Toward A Holistic, Efficient, Stacking Ensemble Intrusion Detection System using a Real Cloud-based Dataset", International Journal of Advanced Computer Science and Applications(IJACSA), 13(9), 2022.
- [25] M. Al-Zahrani, H. Wahsheh, F. Alsaade. "Secure Real-Time Artificial Intelligence System against Malicious QR Code Links". Security and Communication Networks, 2021.

