

HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

---□□□---



SOICT

CAPSTONE PROJECT REPORT

NETWORK STEGANOGRAPHY

Group 1:	Nguyen Duc Thang	20210778
	Nguyen Hoang Anh	20214945
	Nguyen Huu Duan	20214951
	Nguyen Huy Hoang	20214959

Lecturer:	Mr. Tran Quanh Duc
	Mr. Le Van Dong

HANOI, 01/2024

ABSTRACT

In today's world, keeping information safe and communications secure is extremely important. As information technology becomes more prevalent, transactions and communications via computer networks are also becoming more common. Normal communication channels can be vulnerable to information leakage, making it crucial to find more secure ways to communicate. This project explores covert communication by embedding an encrypted message into the destination port numbers of TCP packets. This covert channel is established between a sender and a receiver using a novel method that ensures the confidentiality and integrity of the transmitted message.

The sender encrypts the message using RSA encryption with the receiver's public key. The encrypted message is then divided into smaller chunks, with each bit embedded into the destination port number of a TCP packet. The length of the TCP packet payload is manipulated to indicate the position of each bit. The crafted packets are then transmitted to the receiver, who listens for incoming packets and extracts the message bits from the destination port numbers. Once the entire message is received, the receiver reassembles and decrypts it using their private key.

This approach not only illustrates a secure method of covert communication but also provides insights into how malware might exfiltrate data from a compromised system, thereby aiding in the development of better detection and prevention mechanisms.

Table of Contents

ABSTRACT	2
CHAPTER 1. INTRODUCTION.....	5
Introduction.....	5
Problem Statement	5
Background and Problems of Research	5
Contributions	6
CHAPTER 2. LITERATURE REVIEW	7
2.1. Scope of Research.....	7
2.2. Fields for Encoding	8
2.2.1. Type of Service (Tos)	8
2.2.2. IP Identification (IPID).....	8
2.2.3. IP Flags.....	8
2.2.4. IP Fragment Offset.....	8
2.2.5. IP Options.....	9
2.2.6. TCP Sequence Number.....	9
2.2.7. TCP Timestamp	9
2.2.8. Packet Order.....	9
CHAPTER 3. METHODOLOGY.....	10
3.1. Overview Methodology	10
3.2. Implementation in Python Code	11
3.2.1. Sender Implementation.....	11
3.2.2. Receiver Implementation	16
CHAPTER 4. RESULTS.....	20
4.1. Results.....	20
4.1.1. Evaluation Parameters	20
4.1.2. Simulation Method.....	20
4.1.3. Practical Execution.....	20
4.1.4 Observations.....	21
4.2. Conclusion.....	22

CHAPTER 1. INTRODUCTION

1.1. Introduction

1.1.1. Problem Statement

The problem we address concerns hiding information within network protocol header values to create steganographic covert channels. In this context, we examine a scenario involving three entities: Alice, Bob, and Walter. Alice can freely modify packets originating from a machine within Walter's network. Her objective is to leak a message to Bob, who can only monitor packets at the network's egress points. Alice's goal is to conceal the message from Walter, who can observe (but not modify) any outgoing packets from his network

In this scenario, the challenge lies in creating a covert communication channel that allows Alice to send information to Bob without raising suspicion from Walter, who has the ability to observe all outgoing traffic. Walter's position as a passive observer means he can scrutinize packet headers for anomalies that might indicate covert communication. Thus, the goal is to design a method that embeds information in such a way that it remains undetectable by Walter while still being retrievable by Bob

1.1.2. Background and Problems of Research

Embedding information into network protocol headers to create covert channels is a well-explored area in cybersecurity research. However, many existing techniques are susceptible to detection by passive observers due to the predictable patterns they introduce in network traffic. Such patterns can be detected using statistical analysis or machine learning algorithms, rendering the covert channel ineffective

Key challenges in this area include:

- **Detectability:** Ensuring that the covert communication channel is not easily detectable by anyone other than the intended recipient. This requires sophisticated techniques to avoid creating identifiable patterns in the network traffic.
- **Indistinguishability:** The covert channel must blend seamlessly with regular traffic, lacking any distinctive features that could alert a passive observer to its presence

- **Bandwidth:** Balancing the amount of data that can be hidden within each packet with the need to maintain the covert channel's stealthiness. Higher bandwidth covert channels are often more susceptible to detection.

To address these challenges, our research builds upon the criteria outlined by D. Lamas and Miller for an ideal covert channel:

- **Detectability:** The covert channel should be measurable only by the intended recipient, making it invisible to others
- **Indistinguishability:** The covert channel should lack any identifiable characteristics that differentiate it from normal traffic
- **Bandwidth:** The covert channel should provide a sufficient number of data hiding bits per channel used without compromising its stealthiness

1.1.3. Contributions

In this thesis, we contribute the following:

- **Development of a Novel Method:** We propose a new method for embedding data into the destination port numbers of TCP packets. This method leverages the variability in TCP header values to hide information in a manner that is difficult to detect.
- **Detailed Implementation Methodology:** We provide a comprehensive implementation methodology, including encryption of the message using RSA, embedding the encrypted message bits into TCP packets, and transmitting these packets in a way that minimizes detectability.
- **Demonstration of Effectiveness:** We demonstrate the effectiveness of our method through practical experiments. These experiments show how the method can be used to transmit covert messages without detection by passive observers. We also discuss the potential applications and limitations of our approach.
- **Real-World Scenario Analysis:** We analyze a real-world scenario involving Alice, Bob, and Walter to illustrate how our method can be applied in practice. This analysis highlights the robustness of our approach against passive observation.

CHAPTER 2. LITERATURE REVIEW

2.1. Scope of Research

This research focuses on developing a covert communication method that embeds encrypted messages into the destination port numbers of TCP packets. Unlike traditional steganographic techniques that often fail to produce output conforming to the distribution of unmodified TCP/IP implementations, this approach leverages the inherent properties of TCP/IP protocols to create a covert channel.

The Internet Protocol (IP) allows the transportation of datagrams between fixed length addresses and supports fragmentation and reassembly of long datagrams, though it does not provide reliability guarantees. In contrast, the Transmission Control Protocol (TCP) provides a reliable, stream-oriented communication channel that maintains reliability even in the presence of packet loss, reordering, and duplication.

Certain fields in TCP/IP headers, especially the destination port number in TCP packets, offer potential as carriers for steganographic covert channels. These fields can be manipulated in a way that appears plausible to a passive observer, thereby maintaining the stealth of the communication. This section examines the relevant TCP/IP header fields and evaluates their suitability for embedding covert messages, focusing on the available entropy and the likelihood of these modifications going undetected.

TCP/IP steganography exploits the fact that most header fields remain unchanged during transit. While some fields, like the time-to-live (TTL) field in IP packets, provide limited opportunities for steganography due to their predictable initial values, the destination port number in TCP packets offers a more viable medium. This is because it can be manipulated to encode message bits while still appearing legitimate to network monitoring tools and passive observers.

By embedding encrypted messages within the destination port numbers, this research aims to create a robust and undetectable steganographic communication channel. This approach leverages the reliable and widespread nature of TCP communications, ensuring that the covert messages can be transmitted and reassembled accurately without arousing suspicion.

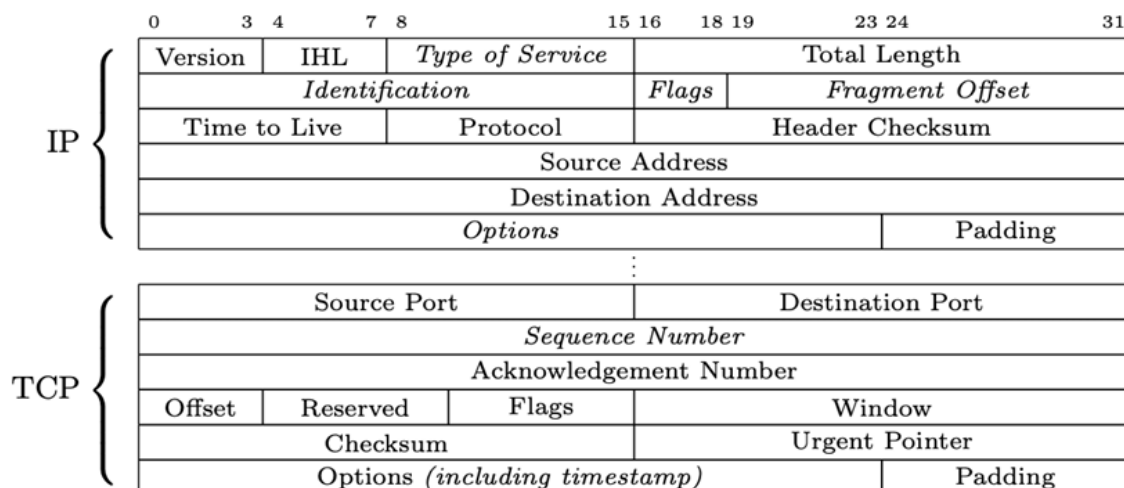


Figure 2.1 TCP/IP Headers

2.2. Fields for Encoding

2.2.1. Type of Service (Tos)

The eight Type of Service (ToS) bits in the IP header are now rarely used with their original semantics but have been repurposed, for example, in the implementation of DiffServ. While there is potential for using these bits as a steganographic carrier, it's easily detected as the field is often set to zero in default operating system configurations.

2.2.2. IP Identification (IPID)

The IP Identification field (IPID) aids in assembling datagram fragments and is allocated 16 bits in the IP header. Unique and unpredictable values are necessary to prevent fragment mix-ups and idle scanning. However, schemes for embedding data in this field are detectable due to the non-random nature of IP IDs.

2.2.3. IP Flags

IP packets include two flags: Do Not Fragment (DF) and More Fragments (MF). The use of the DF bit for steganographic signaling is proposed, but its predictability makes it detectable by a warden.

2.2.4. IP Fragment Offset

IP packets contain an offset field when fragmented, allowing for covert transmission by modulating fragment sizes. However, this method is easily detected, especially in environments where path MTU discovery is used.

2.2.5. IP Options

IP packets rarely contain options, limiting their steganographic potential. The use of the IP Timestamp option is described but is easily detectable and of limited use due to hop constraints

2.2.6. TCP Sequence Number

TCP sequence numbers support reliability and flow control. While a prototype implementation of steganography using TCP ISNs exists, it's detectable due to constraints on sequence number generation.

2.2.7. TCP Timestamp

The TCPtimestamp option allows accurate measurement of round trip time and mitigates sequence number wrap-around issues. Covert channels based on modulating TCP timestamps exist but are detectable due to deviations from expected timestamp distributions.

2.2.8. Packet Order

Packet ordering can carry information, especially in IPSec networks, but it's limited in applicability and easily noticeable due to uncommon packet reordering.

CHAPTER 3. METHODOLOGY

3.1. Overview Methodology

The methodology of this project focuses on covert communication by embedding an encrypted message into the destination port numbers of TCP packets. This covert channel is established between a sender and a receiver using the following detailed steps:

Senders:

- Encrypt the message:
 - The sender starts by using RSA encryption to securely encrypt the message with the receiver's public key. RSA is an asymmetric cryptographic algorithm that ensures only the receiver, who possesses the corresponding private key, can decrypt the message.
 - The message is converted into ciphertext using the receiver's public key, creating a secure, encrypted version of the message.
- Embed Encrypted Message into Packets:
 - The encrypted message is divided into smaller chunks or bits. Each bit of the encrypted message is then embedded in the destination port number of TCP packets.
 - The length of the payload in a TCP packet is manipulated to indicate the position of each bit within the destination port number. This ensures that each TCP packet carries a specific bit of the encrypted message.
- Transmit Crafted Packets:
 - The sender transmits these crafted TCP packets to the receiver. Each packet contains a part of the encrypted message embedded within its destination port number
 - The transmission continues until all parts of the encrypted message have been sent.
- Send End-of-Message Indicator
 - After transmitting the entire message, the sender sends a special packet to the receiver to indicate the end of the message transmission. This packet helps the receiver identify that no further message bits are to be expected.

Receiver:

- Listen for Incoming Packets:
 - The receiver continuously listens for incoming TCP packets. It specifically monitors the destination port numbers of these packets to extract the embedded message bits.
 - Each bit extracted from the destination port number is stored temporarily until the entire encrypted message is received.
- Reassemble and Decrypt the Message:
 - The receiver reassembles the extracted bits to form the full encrypted message. The reassembled message is a binary representation of the encrypted message
 - Upon receiving the special end-of-message packet, the receiver knows that all parts of the message have been received. The receiver then decrypts the reassembled message using their private key to retrieve the original plaintext message.
 - Decryption is performed using RSA decryption with the receiver's private key, reversing the encryption process done by the sender.

3.2. Implementation in Python Code

3.2.1. Sender Implementation

Sender Code (sender.py)

Explanation of the Code:

- Initialize and read the public key: The sender starts by reading the public key from a file and initializing an RSA encryption object to encrypt the message.

```
1  from Crypto.PublicKey import RSA
2  from Crypto.Cipher import PKCS1_v1_5
3
4  with open('public_key.pem', 'rb') as f:
5      bin_key = f.read()
6      key = RSA.import_key(bin_key)
7      cipher = PKCS1_v1_5.new(key)
8
```

- Convert number to list of bits: The 'num_to_bits' function converts a number into a list of bits for easier manipulation.

```
1 def num_to_bits(num):  
2     return [int(i) for i in bin(num)[2:].zfill(8)]
```

- Convert list of bits to positions of 0s and 1s: The 'bits_to_pos' function converts a list of bits into positions of 0s and 1s, facilitating the embedding of bits into the port number.

```
1 def bits_to_pos(bits):  
2     list_0s = []  
3     list_1s = []  
4     for i in range(1, len(bits)):  
5         if bits[i] == 0:  
6             list_0s.append(i)  
7         else:  
8             list_1s.append(i)  
9     return list_0s, list_1s  
10
```

- Craft TCP packet: The 'craft' function creates a TCP packet with payload data from a file or random payload and other pre-specified attributes including destination port number, arbitrarily derived source port number, ip address obtained taken from user input, and whose flag is the "E" flag

```

1  from scapy.all import *
2  from scapy.layers.inet import IP, TCP
3  import random
4
5  dest = ""
6  d_port = random.randint(0, 65535)
7
8  def craft(file, start_pos, length, flag="E", payload="
file"):
9      if payload == "file":
10         file.seek(start_pos)
11         character = file.read(length)
12         pkt = IP(dst=dest) / TCP(sport=123, dport=d_port,
flags=flag) / character
13     else:
14         # create random payload
15         character = ''.join(chr(random.randint(0, 255))
for _ in range(length))
16         pkt = IP(dst=dest) / TCP(sport=123, dport=d_port,
flags=flag) / character
17
18     return pkt
19

```

- Send the message: In the 'client' function, the user enters the ip, the system will check the validity of the ip, if not valid, it will report an error and exit. The user will then enter a message they want to send, which is then encoded and divided into bits. These bits are embedded in the destination port number of the TCP packet, which is then sent. The user can specify a file to be the packet's payload or the system will generate a random payload for the packet. Finally, a special packet is sent to signal the end of the message

```

1     def client():
2         global dest
3         destination = input('Enter the destination IP address
4         : ')
5         # check if the IP address is valid with regex
6         ^((25[0-5]|(2[0-4]|1\d|[1-9])\d)\.?\b){4}$
7         if destination != '':
8             if not re.match(r'^((25[0-5]|(2[0-4]|1\d|[1-9])\
9         d)\.?\b){4}$', destination):
10                print('Invalid IP address')
11                return
12            dest = destination
13
14        while True:
15            message = input('Enter your message: ')
16            if message == 'exit':
17                return
18
19            ciphertext = cipher.encrypt(message.encode())
20            # ciphertext = message.encode()
21
22            cipher_len = len(ciphertext) + 1
23
24            file_name = input('Enter the file name: ')
25            if file_name != '':
26                # check if the file exists
27                try:
28                    open(file_name, 'rb')
29                except FileNotFoundError:
30                    print('File not found')
31                    return
32                with open(file_name, 'rb') as file:
33                    file_len = len(file.read())
34                    start_pos = 0
35                    packet_len = (((file_len // (cipher_len *
36            8)) - 16) // 1000) * 16

```

```

33         print(ciphertext)
34         for char in ciphertext:
35             list_bits = num_to_bits(char)
36             list_0s, list_1s = bits_to_pos(
num_to_bits(d_port))
37             for i in list_bits:
38                 if i == 0:
39                     ind = list_0s[random.randint
(0, len(list_0s) - 1)]
40                 else:
41                     ind = list_1s[random.randint
(0, len(list_1s) - 1)]
42                 pkt = craft(file, start_pos,
packet_len + ind)
43                 start_pos += packet_len + ind
44                 send(pkt, verbose=False)
45
46             # send special packet to indicate the end
of the message
47             pkt = craft(file, start_pos, packet_len)
48             send(pkt, verbose=False)
49             # close the file
50             file.close()
51         else:
52             # create random packet length that is a
multiple of 16
53             packet_len = random.randint(1, 50) * 16
54             for char in ciphertext:
55                 list_bits = num_to_bits(char)
56                 list_0s, list_1s = bits_to_pos(
num_to_bits(d_port))
57                 for i in list_bits:
58                     if i == 0:
59                         ind = list_0s[random.randint(0,
len(list_0s) - 1)]
60                     else:
61                         ind = list_1s[random.randint(0,
len(list_1s) - 1)]
62                 pkt = craft(None, 0, packet_len + ind
, payload="random")
63                 send(pkt, verbose=False)
64             # send special packet to indicate the end of
the message
65             pkt = craft(None, 0, packet_len, payload="
random")

```

```

66         send(pkt, verbose=False)
67     if __name__ == "__main__":
68         client()
69

```

3.2.2. Receiver Implementation

Receiver Code (receiver.py)

Explanation of the Code:

- Initialize and read the private key: The receiver starts by generating an RSA key pair if they do not already exist and reads the private key from a file.

```

1     from Crypto.PublicKey import RSA
2     from Crypto.Cipher import PKCS1_v1_5
3     import os
4
5     def generate_rsa_keys():
6         key_gen = RSA.generate(1024)
7         public_key = key_gen.publickey().export_key()
8         private_key = key_gen.export_key()
9         with open('public_key.pem', 'wb') as fi:
10             fi.write(public_key)
11         with open('private_key.pem', 'wb') as fi:
12             fi.write(private_key)
13
14     if not (os.path.isfile('public_key.pem') and os.path.
15 isfile('private_key.pem')):
16         generate_rsa_keys()
17
18     with open('private_key.pem', 'rb') as f:
19         bin_private_key = f.read()
20
21     key = RSA.import_key(bin_private_key)
22     cipher = PKCS1_v1_5.new(key)

```


- Convert number to list of bits: The 'num_to_bits' function converts a number into a list of bits for easier manipulation

```
1 def num_to_bits(num):  
2     return [int(i) for i in bin(num)[2:].zfill(8)]  
3
```

- Convert list of bits to number: The 'bits_to_num' function converts a list of bits back into a number to reassemble the original message

```
1 def bits_to_num(bits):  
2     return int("".join([str(i) for i in bits]), 2)  
3
```

- Parse received packets: The "parse" function listens for incoming TCP Packets, checks the flags and source port to identify the relevant packets, extracts bits from the destination port number, and reassembles these bits into the encrypted message. Once enough bits are gathered to form a byte, it adds the byte to the encrypted message.

```

1  from scapy.all import *
2  from scapy.layers.inet import IP, TCP
3
4  lst = {}
5  cipher_text = {}
6  payloads = {}
7
8  def parse(pkt):
9      global cipher_text
10     global lst
11     if not pkt.haslayer(TCP) or not pkt.haslayer(IP):
12         return
13
14     flag = pkt[TCP].flags
15     sport = pkt[TCP].sport
16     sip = pkt[IP].src
17
18     if flag == 0x40 and sport == 123:
19         dst_port = pkt[TCP].dport
20         list_bits = num_to_bits(dst_port)
21         payload = pkt[TCP].payload
22         bytes_of_payload = bytes.fromhex(bytes_hex(
23             payload).decode())
24
25         if len(bytes_of_payload) % 16 == 0:
26             try:
27                 msg = cipher.decrypt(cipher_text[sip], b'
28 ')
29                 print(sip, msg)
30                 cipher_text[sip] = b''
31                 return
32             except ValueError as e:
33                 print(e)
34                 print(cipher_text[sip])

```

```

33
34         if sip in lst:
35             lst[sip].append(list_bits[len(
bytes_of_payload) % 16])
36             payloads[sip] += bytes_of_payload
37         else:
38             lst[sip] = [list_bits[len(bytes_of_payload) %
16]]
39             payloads[sip] = bytes_of_payload
40
41         if len(lst[sip]) == 8:
42             num = bits_to_num(lst[sip])
43             if sip not in cipher_text:
44                 cipher_text[sip] = bytes([num])
45             else:
46                 cipher_text[sip] += bytes([num])
47             lst[sip] = []
48

```

- Listen for incoming packets: The 'server' function uses 'sniff' to listen for incoming packets on the specified network interface and calls the 'parse' function to process each packet.

```

1     def server():
2         sniff(iface="eth0", prn=parse)
3
4     if __name__ == "__main__":
5         server()
6

```

CHAPTER 4. RESULTS

4.1. Results

In this section, we present the results of implementing the covert communication system described in the methodology. The practical experiments were conducted to demonstrate the effectiveness of the proposed method. The following subsections provide details of the experimental setup, execution, and observations

4.1.1. Evaluation Parameters

To evaluate the covert communication method, we focused on the following parameters:

- Stealthiness: The ability to remain undetected by network monitoring tools.
- Reliability: Successful transmission and correct decryption of the message.
- Bandwidth: The amount of data that can be transmitted covertly within the TCP packets

4.1.2. Simulation Method

We implemented the covert communication system using Pythonscripts for both the sender and receiver. The sender script encrypts the message using RSA and embeds the encrypted bits into the destination port numbers of TCP packets. The receiver script listens for incoming TCP packets, extracts the embedded bits, and reassembles and decrypts the message

Set up:

- Sender and receiver machines were set up in a controlled network environment
- Network traffic was monitored using Wireshark to capture and analyze the packets

4.1.3. Practical Execution

Sender:

- The sender script was executed, and the destination IP address and message were provided as inputs
- The script encrypted the message, embedded it into the TCP packet headers, and transmitted the packets

Receiver:

- The receiver script was executed and monitored the incoming network traffic.

- It successfully extracted the embedded message bits from the destination port numbers and reassembled the encrypted message.
- The message was then decrypted to retrieve the original plaintext message.

Packet Capture Analysis:

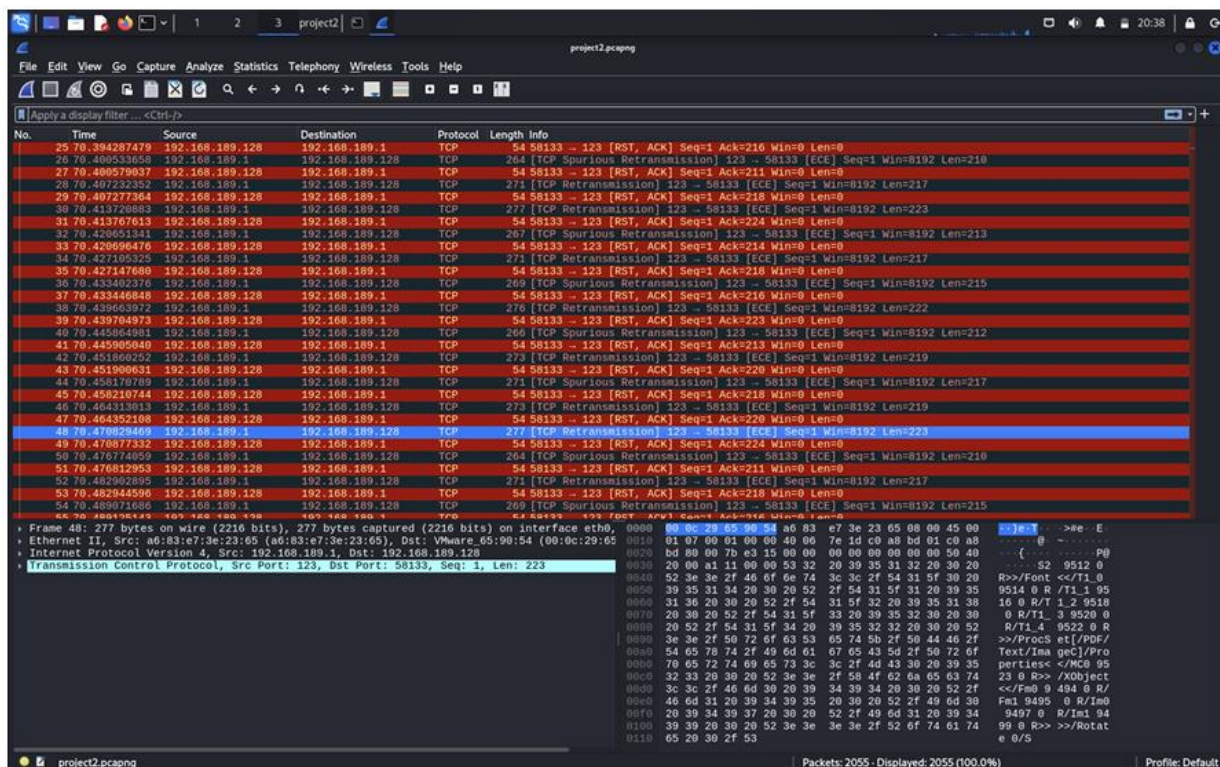
- Wireshark was used to capture and analyze the network traffic during the transmission.
- The captured packets were examined to verify the embedding of message bits into the TCP headers.

4.1.4 Observations

The result of the practical experiments are presented below:

Wireshark Packet Capture:

- The captured packets show the transmission of TCP packets from the sender to the receiver.
- The destination port numbers contain the embedded message bits, as illustrated in the provided screenshots.



Detailed Analysis:

- Stealthiness: The embedding of message bits into the destination port numbers was not detected by standard network monitoring tools, indicating a high level of stealthiness.
- Reliability: The receiver successfully reassembled and decrypted the message, demonstrating the reliability of the method
- Bandwidth: The method allows for the covert transmission of small to moderate sized messages without compromising stealthiness.

4.2. Conclusion

The practical experiments validate the effectiveness of the proposed covert communication method. The embedding of encrypted message bits into the destination port numbers of TCP packets provides a secure and stealthy way to transmit information. The results indicate that this method is reliable and remains undetectable under standard network monitoring conditions. Further research could explore optimizing the bandwidth and extending the method to other protocol fields for enhanced covert communication capabilities