KARABUK UNIVERSITY
FACULTY OF ENGINEERING
COMPUTER ENGINEERING DEPARTMENT


# CME 422 PROJECT REPORT

## NETWORK TRAFFIC MONITORING AND ANOMALY DETECTION

**ÖZKAN DERNEK**

**2014010217019**

**YÜKSEL ÇELİK**

**DATE (JUNE, 2020)**

I certify that in my opinion the thesis submitted by ÖZKAN DERNEK titled “Network traffic monitoring and anomaly detection ” is fully adequate in scope and in quality as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Yüksel Çelik ……

Thesis Advisor, Department of Computer Engineering

….. / ….. / 20..

The degree of Master of Science by the thesis submitted is approved by the Administrative Board of the Institute of Graduate Programs, Karabuk University.

Assist. Prof. Dr. Hakan Kutucu ……

Head of Institute of Graduate Programs

*“I declare that all the information within this thesis has been gathered and presented in accordance with academic regulations and ethical principles and I have according to the requirements of these regulations and principles cited all those which do not originate in this work as well.”*

ÖZKAN DERNEK

## Executive summary

Computer communications firstly developed in 1969 by Project name Advanced Research Projects Agency Network (ARPANET). The first computers were connected in 1969, the main goal is to communicate other computers and data sharing with computers ,data divided chunks and sent to accross computer . Behind this theory control data and manage easily .

A protocol was developed ( TCP/IP protocol suite) by which multiple separate networks could be joined into a network of networks. TCP/IP protocol suite has many different protocol inside to manage and send data correctly across the networks .

After this development all world changed. Today's Internet works

with TCP/IP protocol suite . We send and take data almost whole day. Although most people use the Internet , They do not know answers 'How computer systems communicates each other , How to works Internet ' these questions .

I am going to explain this questions , and will focus TCP/IP protocol suite , Firstly we need to understand 'What is the network ?' , The definition of computer network from wikipedia "A computer network is a group of computers that use a set of common communication protocols over digital interconnections for the purpose of sharing resources located on or provided by the network nodes . The interconnections between nodes are formed from a broad spectrum of telecommunication network technologies, based on physically wired, optical, and wireless radio-frequency methods that may be arranged in a variety of network topologies " , Basically understanding of network at least two computers commuciate each other via a medium , but if there are not at least two computers, there is no need for a network.

Based on my research data monitoring is very important to see what's going on your network , Probably almost every day we have risk on the Internet .

An attackers or hackers can attack your systems . If we can not detect attacks, hackers or attackers can damage your system . In this situation , I researched how can we basically monitor our network's data and detecttion basic attacks . There are a lot of attacks outside but I will a give basic understanting data throughput and understand packets .

I explained basic understanding of networks and data communications of computers , we entered TCP/IP protocol suite , I will explain other protocols and details of TCP/IP and OSI later.

## ACKNOWLEDGMENT

First of all, I would like to give thanks to my advisor, Assist. Prof. Dr. Yüksel Çelik, for his great interest and assistance in preparation of this thesis.

## TABLE OF CONTENTS



**List of figures and tables**

# INTRODUCTION

Welcome to introduction section , we will focus fundamentals of TCP/IP protocol suite , our main goal is solid understanding of TCP/IP suite. Before we start understanding of TCP/IP suite , we need to explain what is the Internet , How is it works .?

Let's understand behind theory of real Internet communication . we use the ınternet everyday and every hour . But almost %90  of people can not explain how it is work. Basically it is a only electrical signal of data. Computers use 1's and 0's to represent data .  For example 'A' character represent to 11110000 like that in todays computer . When you type a message to your friend computer converts your data binary system . Engineers and researchers developed TCP (tranmission control protocol ) and IP (Internet Protocol) to delivery your data correctly . Engineers and researchers  thought  computers needs uniquely address to transmit data correctly . They developed IP logical address to make uniquely each computers on the world .There are billions devices works with internet today's World .How is it possible incredible data send or recieve correctly . To control data we use TCP protocol , Transmission control protocol controls data to send and recieve correctly .I will explain this part Understanting Tcp/ip section .

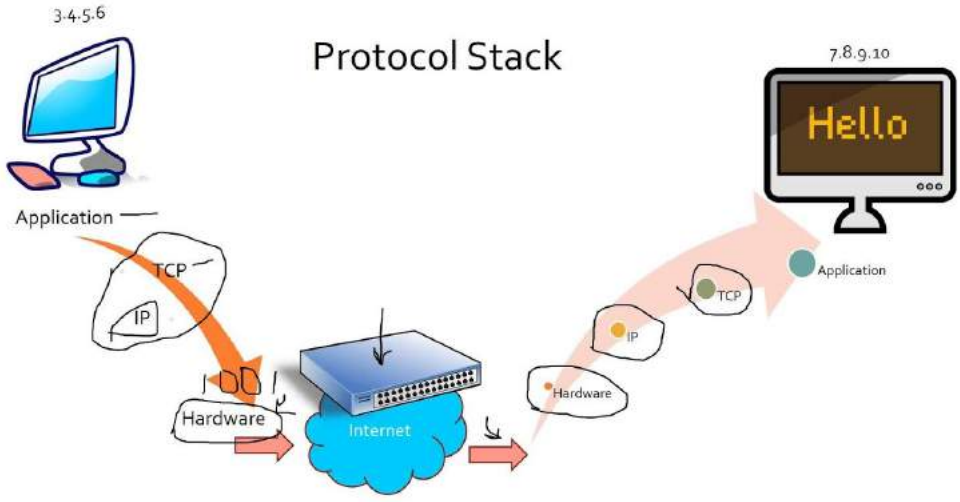



Figure 1

When we look at the Picture left computer has a ip 3.4.5.6 and right computer has a diffrent ip address 7.8.9.10 they can communicate each othet with TCP/IP protocol suite .

Your application data (Maybe skype ,Google chrome so on .) converted 1's and 0's . Tcp/ıp protocol suite added additional information to your packet . A physical devices (routers ,switch ,hub etc) takes data and sent via medium target host.

This explanation is the basic understanding of network .Our main goal is the observing data 1's and 0's on networks . I developed a basic tool to observing data with python language.This tool made for sniffing data on your host machine . Incoming packets IP address , port , protocol number , and anomaly detection based on İp address , to make this I used different databases on the World . Virustotal , Abusedİp .

## Technologies

Python Programming Language is the main technology behind our tool . I choosed python language because python language is a all in one language to backend and frontend development . There a lots of diffirent gui options in python language and libraries is most reliable with Linux based systems .

Let's quick introduction about python " **Python** is an interpreted, high-level,general-purpose programming language Created by Guido van Rossum and first released in 1991, Python's design philosophy emphasizes code readability with its notable use of significant Whitespace . Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects.

Python has very easy syntax to understand . In Linux operating systems most of distributions on the ınternet has pyhton language pre-installed . If you use any kind of Linux operating system you can develop your Project with python language . Most reliable with Linux operating systems . Python language accross platform language independent environment . This programming language Works Windows , Macos , Unix/Linux based operating systems .

To download python and install your operating system , it is easy just you need to go www.python.org official website .

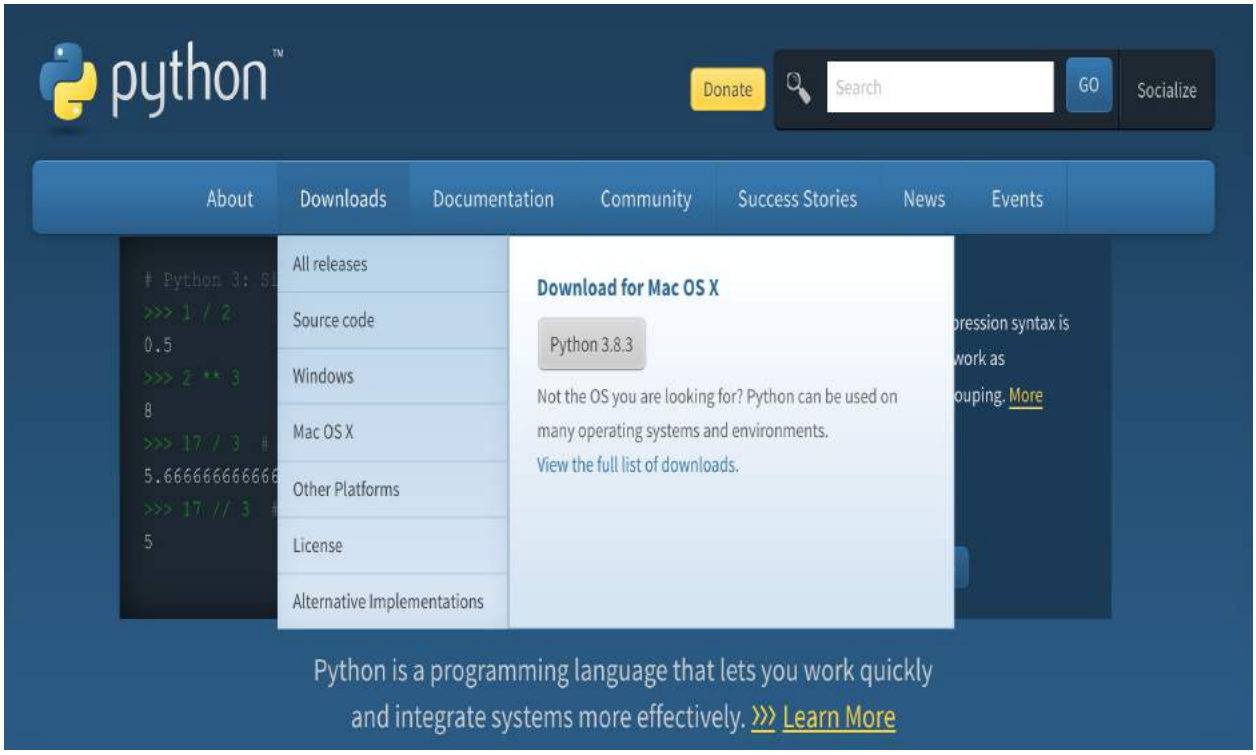



Figure 2

You can download python , Downloads section on the menu bar . Choose your operating system what you use and download python your system .

## Environments

Devolopment place and Environments place is Linux operating systems . While I developing this tool , I used different distros of Linux . We need the understand Linux operating system .

There are diffirent operating systems on the world . Most common operating system is Windows operating system . But there is a another option to develop your code and programs most secure and stable operating system Linux . İt is open source operating system you can customize or develop this system uniquely .

Linux is a member of open source Unix-like operating systems based on the Linux-kernel, an operating system kernel first released on september 1991 by Linus Torvalds.

A view of Linux operating system .

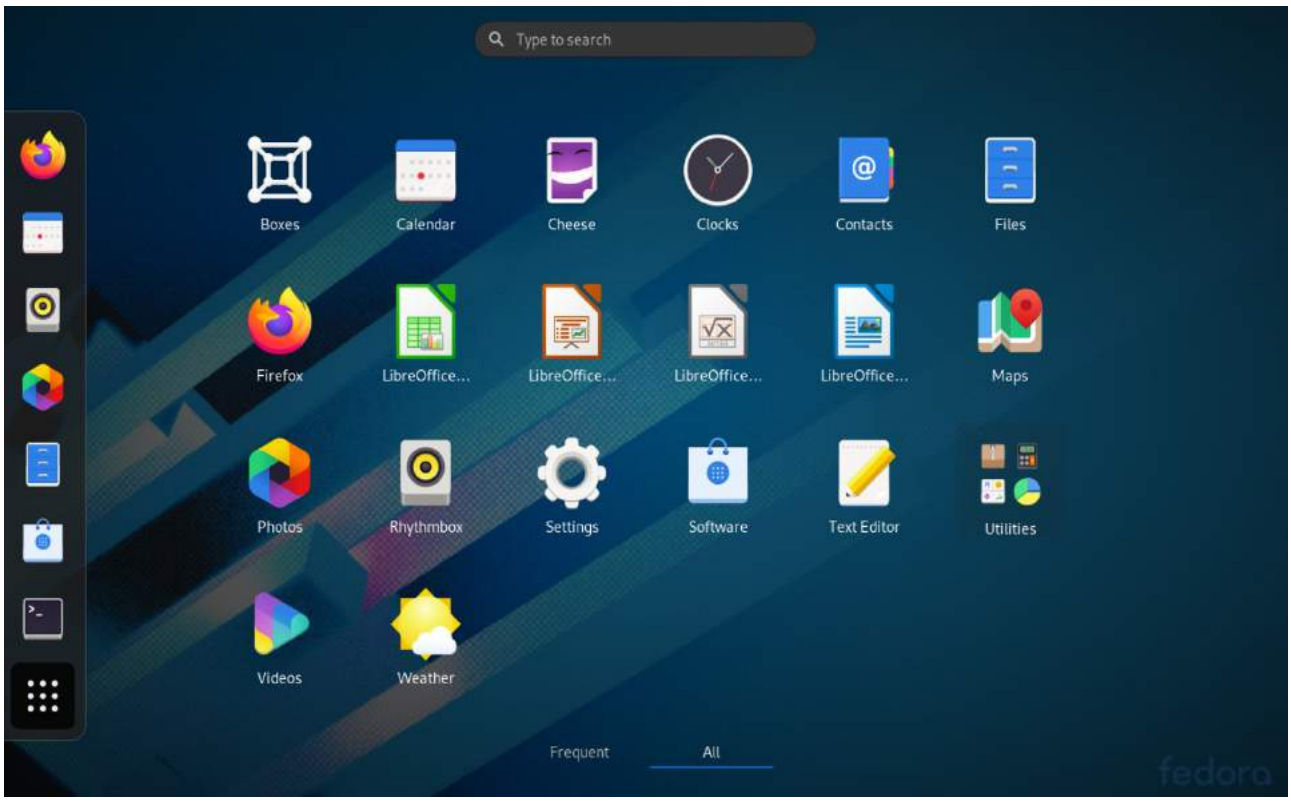
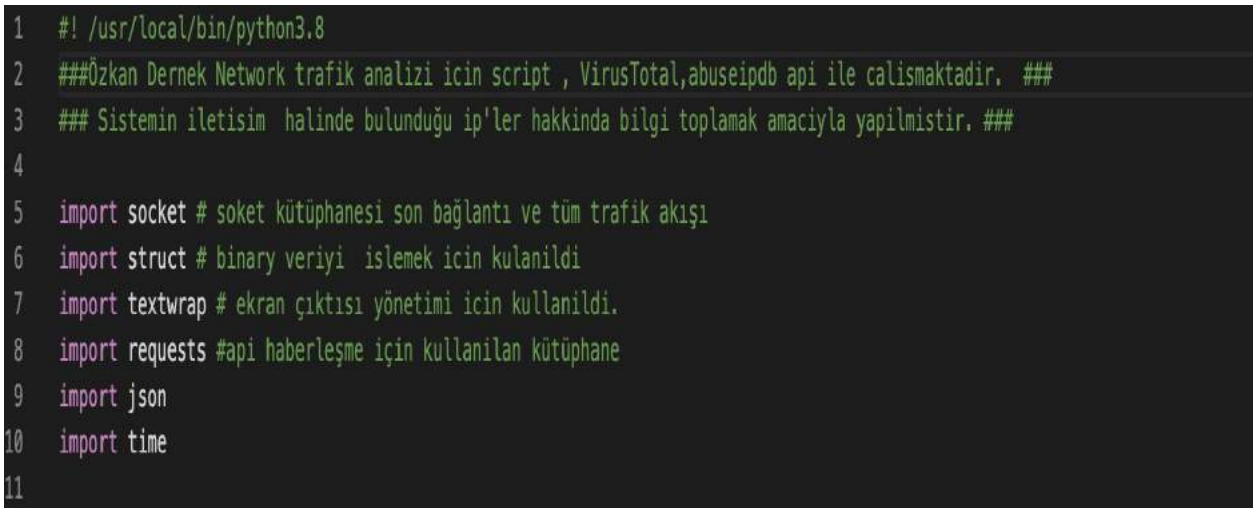



Figure 3

## Libraries Of Python



```
1  #! /usr/local/bin/python3.8
2  ###Özkan Dernek Network trafik analizi icin script , VirusTotal,abuseipdb api ile calismaktadir.  ###
3  ### Sistemin iletisim  halinde bulunduğu ip'ler hakkinda bilgi toplamak amaciyla yapilmistir. ###
4
5  import socket # soket kütüphanesi son bağlantı ve tüm trafik akışı
6  import struct # binary veriyi  islemek icin kulanildi
7  import textwrap # ekran çıktısı yönetimi icin kullanildi.
8  import requests #api haberleşme için kullanilan kütüphane
9  import json
10 import time
11
```

Figure 4

**Socket library** is to create a socket in your software . This library provides access to the BSD *socket* interface. It is possible on all modern Unix systems, Windows, MacOS, and probably additional platforms.

What is the socket ?

Sockets allow communication between two different processes on the same or different computers. To be more precise, it's a way to talk to other computers using standard Unix file descriptors. In Unix, every I/O action is done by writing or reading a file descriptor. A file descriptor is just an integer associated with an open file and it can be a network connection, a text file, a terminal, or something else.

**Struct library** is used to handling binary data stored from network connections.

**Textwrap library** ,In python the textwrap library is used to format and wrap  texts. There are some options to format the texts by adjusting the line breaks in the input paragraph.

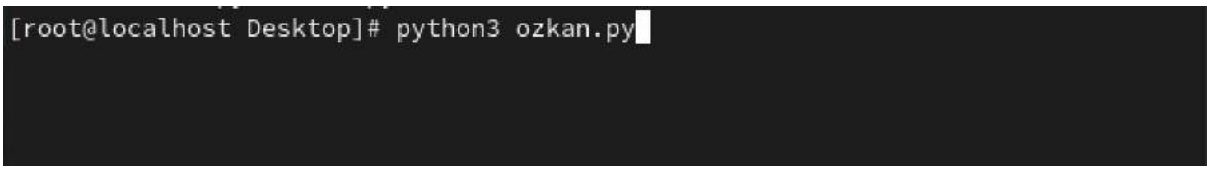
**Requests library ,** this library used to communicate other platforms with api . Our tool queries ip addresses on Virustotal and Abuseipdb .

# USAGE

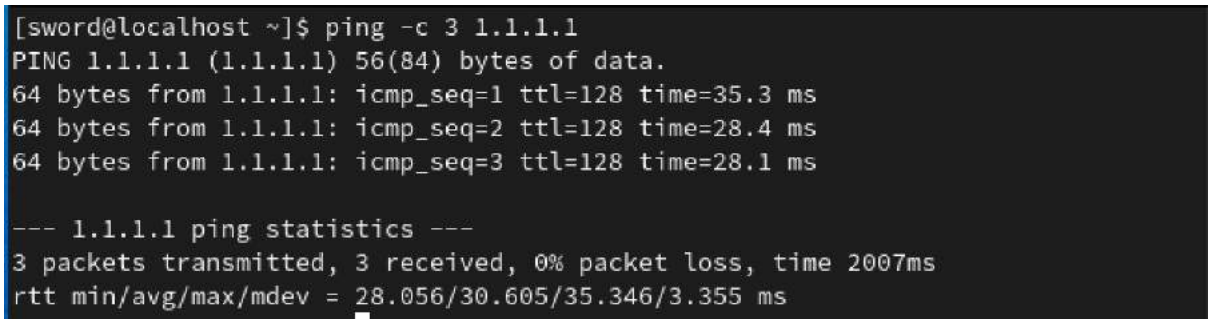
Before you use our tool , there are some requirements . I mentioned libraries previous pages . You need to install these libraries . İt is easy use tool for end users . There is no complicated things. I will represent information how to use it any Linux based operating systems .


Figure 5

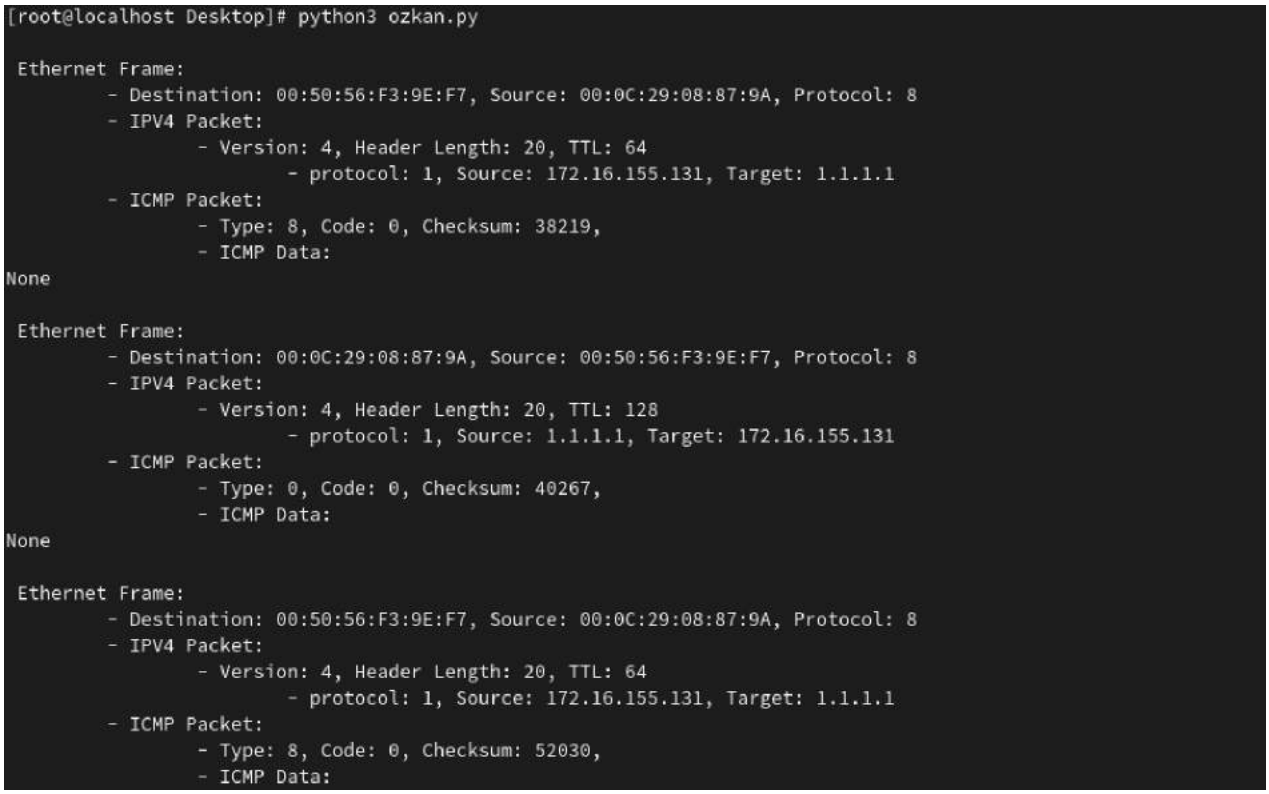
First of all , we need to run this tool root privilege because of our tool wants to access socket information .After this step we can run our tool easily .



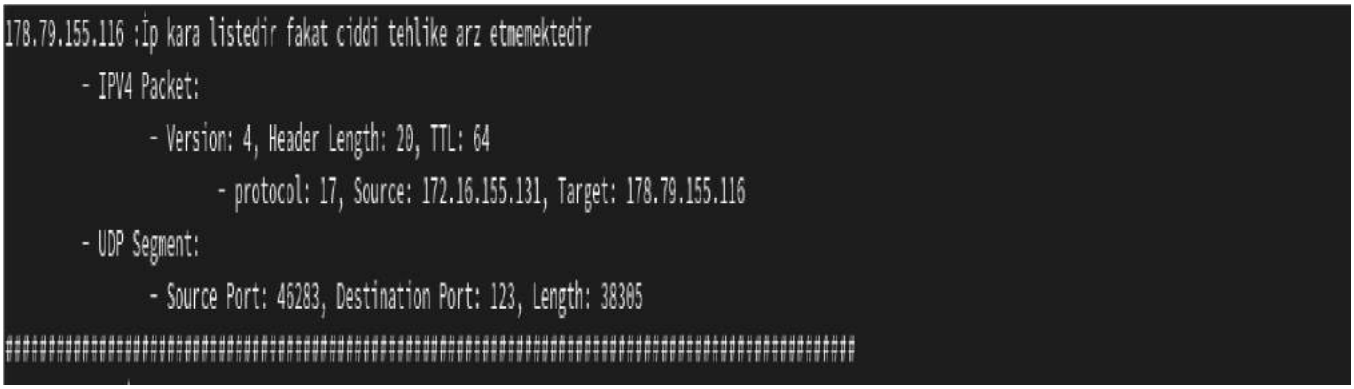

I developed this tool python3 3.8.2 version . This does not work python older version , Please be careful and check your python programming language version. Then I run my tool on system. Sniffer is working , I sent ICMP packet to Dns server to create traffic on network. Tool is sniffing data and converts binary data to string , TCP/IP suite protocols will demuxing in this step.



Let’s check our tool outputs . What we should see ? We sent ping to Cloudflare DNS ,our tool understand packet. This is icmp packet and destination ip ,source ip will be output on the screen.




As we can see there is information about  packets . This packets are icmp packets . I use VMware virtulization platform for running my Linux operating system thus my source ip address private address . But target machine address is public ip address .And our tool has a feature to check ip address is bad or good source ? How is that possible , There are lots of platform on the internet , They serving information about ip address reports , I used Virustotal and Abuseipdb because of free queries .




Our tool used databases for query this ip address , This ip address reported black list but it is not dangerous ip address . We will use this information anomaly detection on networks .

Tool can detect ping flood attack , ping flood attack is a type of a Dos (Denial Of service attack) . Our tool can detect this type of attack and blocks incoming packets from source. I sent many icmp packets to our host machine and our tool detect ping flooding attack to our host machine after that tool dropped packets .

Basic theory of ping flooding , I will give a Picture about ping flooding .

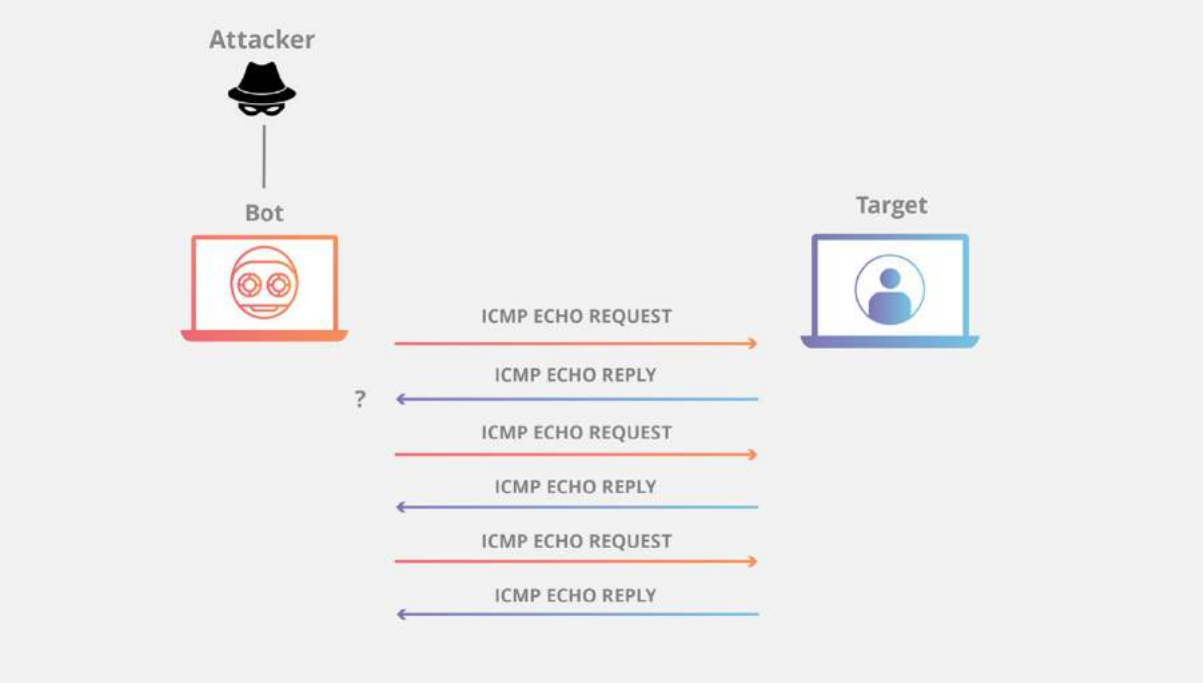
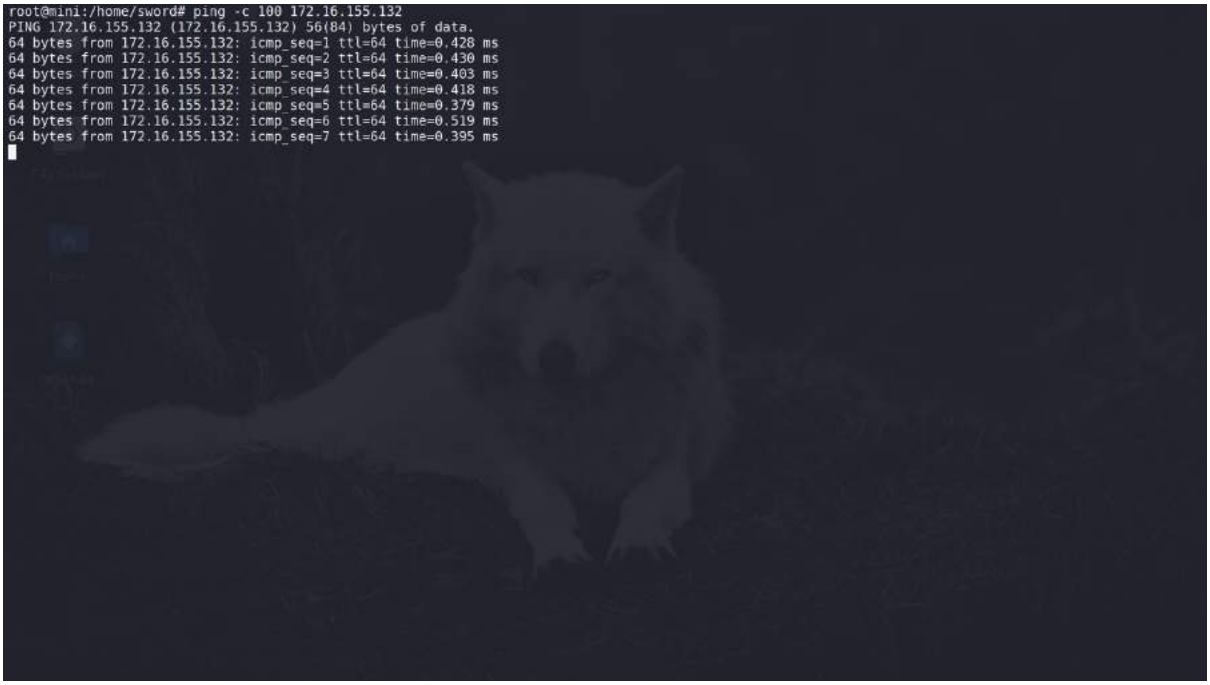
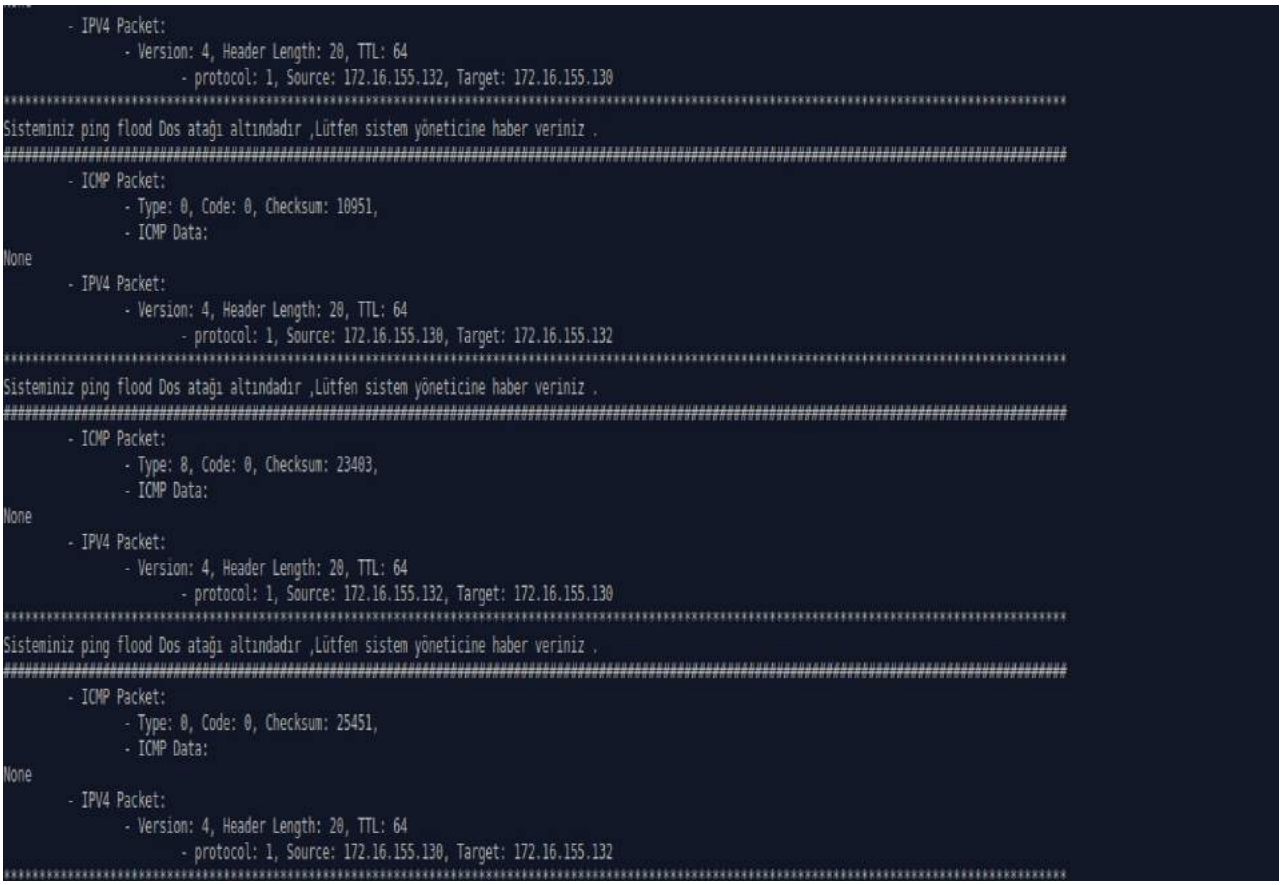



Figure 6

Attacker sent to many Icmp request our systems , if we can’t detect this attack , multi-thread attack can consume our network resources .




An output our tool , Detecting ping flooding attack .



# Development

In development step , I used V model for software development . Most powerful development model in software engineering , almost all steps is best .

- Planning
- Requirements
- Architecture
- Detailed Design
- Implementation
- Unit Testing
- Integration testing
- System and Acceptance Test
- Maintance

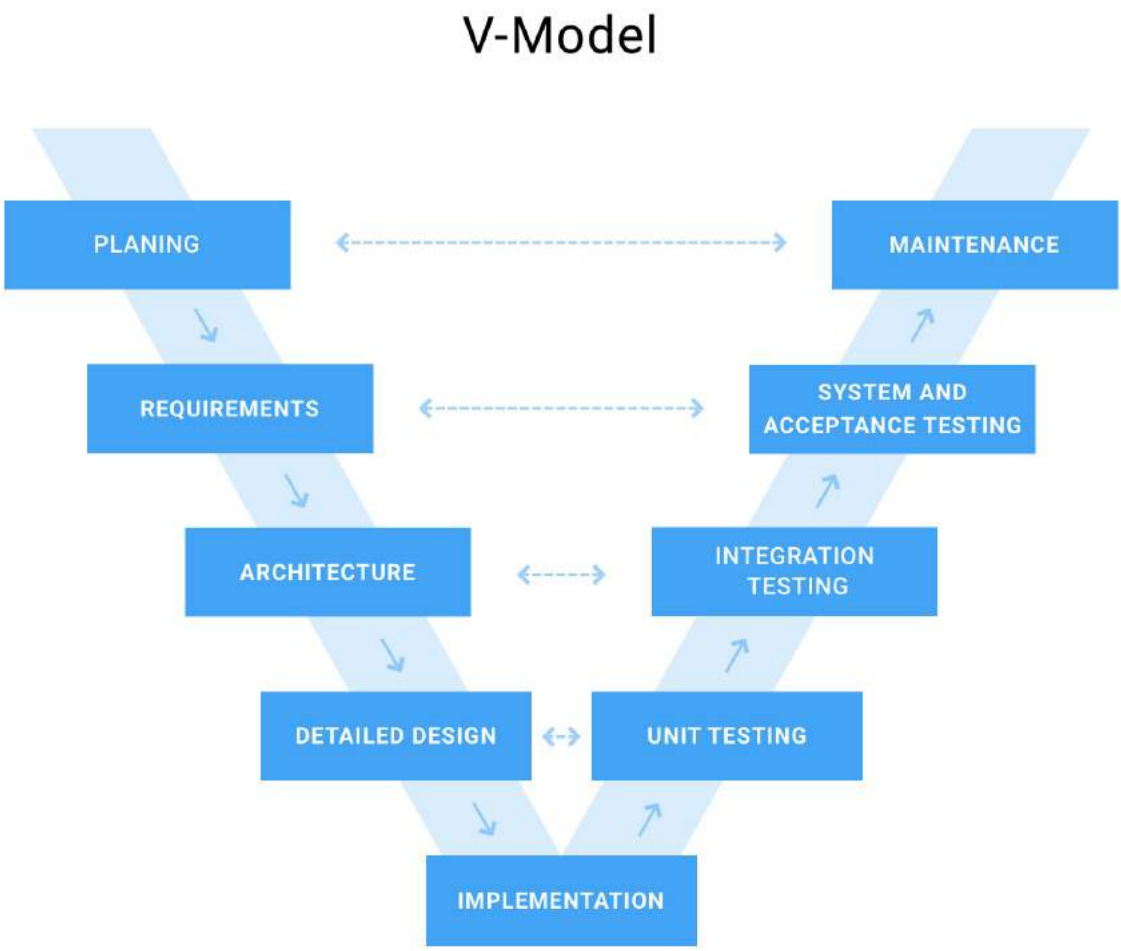
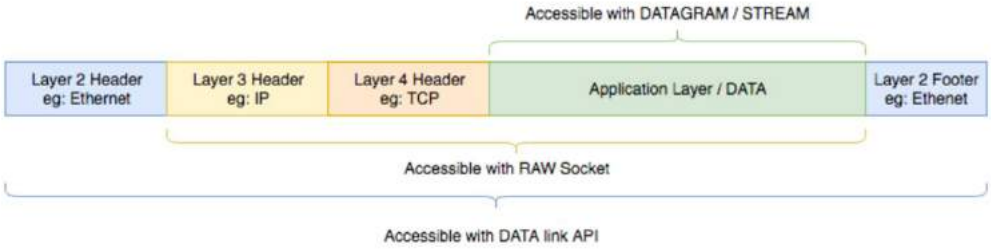



Figure 7

I developed simple network sniffer and anomaly detection with socket module in python , Basically I takes raw data from socket after this step we will observe data which protocol number has packet ,then we will fetching data from packets.



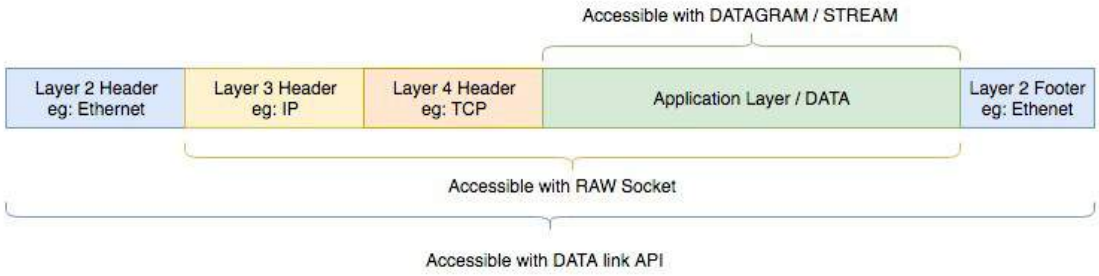

1. Import the modules:

**import socket**

2. we can create an INET raw socket:

3. **s=socket.socket(socket.AF_INET, socket.SOCK_RAW, socket.IPPROTO_TCP)**

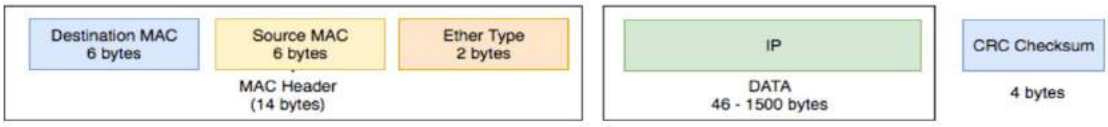
- When we take data from socket there are some constant to developing unique systems . Socket.af_inet is a address family . Second parameter is to raw data , third parameter is protocol number , There is some conventions for protocol numbers , IANA (Internet Assigned Numbers Authority) ,



## Fetching information into the packet

We can fetching the data that we sniffed, and unpack. the headers. To parse a packet, we need to have an idea of the Ethernet frame and the packet headers of the IP.

The Ethernet frame structure is as follows:



# CONCLUSIONS

This report describes basic understanding of computer communication , and answers about most popular questions of network communications . Our main goal is the understand how to communicate computers on real world.I observed TCP/IP protocol suite in this report . TCP/IP protocol suite has a layered architecture to encapsulation data . I developed a python script for sniffing data on socket. After that we parsed data with socket library in python .

We have raw data about traffic on your network , each packet demuxed special fuction into our tool . ICMP data , TCP data etc has a unique function to parsing data .After we took data I made meaningful data to understand some systems . IP address incoming packets , Flags and so on . I did not mention TCP/IP deeply but I researched too much theory and protocols , while I developing this tool . We answered some popular questions .

- How data communiction Works on real network ?
- Basic Understanding of TCP/IP , Ethernet Frame ?
- What is the network anomaly ?
- Ping flooding and How to detect our system .
- Ssh brute-force detection and prevent this attack

Finally I finished my project , In this project main goal is understand of theory . My purpose is not develop best product for marketing . My aim is the real understan low-level network communications and attack types on networks.

## Pictures of some codes in Software

```
# AF (Address Family AF_PACKET kullanildi Windows ve Linux dağıtımlarında çalışıyor . Macos Uyumsuz.)
#socket.ntohs(3) buradaki kullanilan parametre verininin nasil işlendiği ile ilgili IPV4 çaliştigimiz icin kullanildi.
conn = socket.socket(socket.AF_PACKET, socket.SOCK_RAW, socket.ntohs(3))
print("Trafik Dinleme basladi.\n")
```

```
if target :
    url = 'https://api.abuseipdb.com/api/v2/check'
    querystring = {
        'ipAddress': target,
        'maxAgeInDays': '90'
                    }

    headers = {
        'Accept': 'application/json',
        'Key': '24650a75122305b55858c94f7ac828b40595963eb1498197fd4f989545932dc84bdb1c5bebab50db'
           }

    response = requests.request(method='GET', url=url, headers=headers, params=querystring)

# Formatted output
    decodedResponse = json.loads(response.text)
    cevap = decodedResponse['data']['isWhitelisted']
    rapor_durumu = decodedResponse['data']['abuseConfidenceScore']
```

```
    if rapor_durumu <30:

            print(target+ " :" + "İp kara listedir fakat ciddi tehlike arz etmemektedir")

    elif rapor_durumu >=30 and rapor_durumu <=50:
            print(target+" :"+ "İp kara listedir , trafik hareketleri izlenmelidir.")

    elif rapor_durumu>50 and rapor_durumu <= 70:
           print(target+" :"+"İp kara listedir , ip haberleşme kısıtlanmalıdır.")

    elif rapor_durumu >70:
            print(target+" :"+"İp kara listedir.Gelen paket istekleri droplanmalıdır.")

    else:
            print(target+" :"+"İp kara listedir ,fakat rapor bulunamamistir.Diger referans kaynaklara göz
```

```
if src not in white_list or target not in white_list:
        if counter > 10:
            print("*"*150)
            print ("Sisteminiz ping flood Dos atağı altındadır ,Lütfen sistem yöneticine haber veriniz .")
            print("#"*150)
            komut = "iptables -I INPUT -s " +src+" -j DROP"
            os.system(komut)
```

## Sources


- www.python.org
- https://en.wikipedia.org/
- TCP/IP Protocol Suite  Fourth Edition  (Behrouz A.Forouzan)
- TCP / IP Illustrated, W. Richard Stevens

**AUTOBIOGRAPHY**

Özkan DERNEK was born in 1995 in Zonguldak. Özkan has completed his first, middle and high school education in Zonguldak. In 2014, he got into Karabuk University - Computer Engineering of Engineering Faculty. In 2020 graduated Karabuk University as a Computer Engineer.

**CONTACT:**

Address : Karapınar Beldesi Çaycuma Zonguldak

GSM : (536) 071 41 95

E-mail : ozkandernek@gmail.com