Table of Contents

ECE 8930 Lab 3 Guide

Network background traffic is very important in DDoS attack detection studies. The amount of traffic generated in a network depends on the time of the day, the allowed network services (e.g., email, cloud services, VoIP, downloading/uploading large files), and the average number of users. It constitutes a ground truth during a detection. While signature-based detection approaches the search for known patterns in background traffic, anomaly-based detection approaches use the sudden deviation in the observed feature (such as the number of packets received in a second) of the background traffic for attack detection. Therefore it is crucial to use operational network background traffic in these studies. However, it is not always possible to have access to an operational network for testing. Researchers overcome this issue by using simulated network background traffic, generating traffic in a computer cluster or replaying packet traces collected from an operational network in their studies.

In this assignment you will:

- Learn how to replay captured network traffic.
- Observe and understand the difference between operational network background traffic and a simulated background traffic

About the Time Series:

Time series refers to the number of packets captured per second. Time series analysis is more practice than packet analysis to study the DDoS traffic. It takes much less resource to perform during heavy load of traffic.

In this lab, the time series is used to compare the difference between real traffic and simulated traffic.

Thus, the so-called simulated traffic is simply a series of generated time stamps indicating packets arriving times, instead of a pcap file.

Lab Setup:

Campus traffic is broadcasting on the Clemson network. For students at Charleston, please ssh to any Clemson host machine to capture the campus traffic. Detailed instructions will be provided in the next section.

In this lab, you will need:

- One host machine on the Clemson campus to capture campus traffic.
- A pre-configured VM to
  - replay the captured traffic
  - execute a python script to compare the real background traffic versus simulated background traffic.

Credential to the pre-configured VM:
username: student
password: private123

Clemson host machine available to capture campus traffic:

192.168.10.9

Lab Guide:

- ## Important!

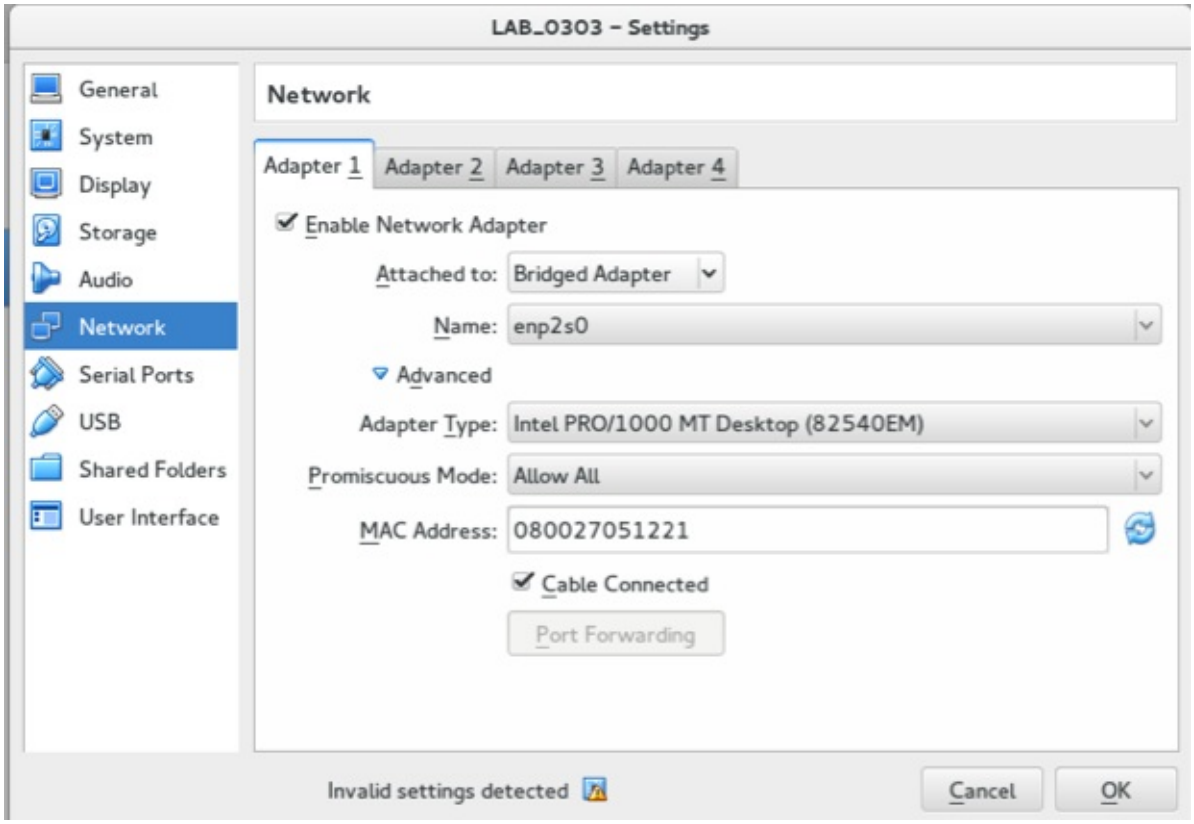
- When copying any code, don't copy the following mentioned symbols at the beginning of each line.
- "$" precedes Linux commands that are typed at a regular user's shell prompt, usually your host machine
- "#" precedes Linux commands that are typed at a root shell prompt, usually in a VM
- ">>>" precedes Python or Scapy commands

Step 1. Import the pre-configured VM:

- The command in this step should be executed on a host machine.
- Replace the 'username' with your username in the following command.

Import the pre-configured VM: Kali_DDoS_Lab6

Change the network configuration of the newly created VM. Generate a new mac address by click the icon to the right of the MAC address input box. Then start the VM:



Step 2. Capture and replay the campus background traffic:

- The first ssh command is only for Charleston students.

Execute the following command on your host machine and log in to the machine with campus traffic:
192.168.10.9

```
$ ssh 192.168.10.9
```

This command captures 1000 packets and saves it in file campus_traffic.pcap in pcap format.

```
$ tshark -i enp3s1 -c 1000 -w campus_traffic.pcap -F libpcap
```

This command captures packets arriving in 100 seconds:

```
$ tshark -a "duration:120" -w campus traffic1.pcap -F libpcap
```

The following command should be **executed on the preconfigured VM**.

This command will get scrips from NAS:

```
root@localhost~# sftp username@192.168.10.5
root@localhost~# cd VMs/
root@localhost~# get ECE8930_lab3_spring_2018_scripts.tgz
root@localhost~# exit
root@localhost~# tar zxvf ECE8930_lab3_spring_2018_scripts.tgz
```

This command will copy the captured file back to your VM. Replace username with your username and the IP to the host IP you just logged in to capture campus traffic:

```
root@localhost~# scp username@192.168.10.x:~/campus_traffic.pcap .
```

This command will replay the campus_traffic.pcap file, open wireshark and listen to channel "lo" before run it:

```
root@localhost~# sudo tcpreplay --intf1=lo campus_traffic.pcap
```

Save the replay traffic and compare to the real one.

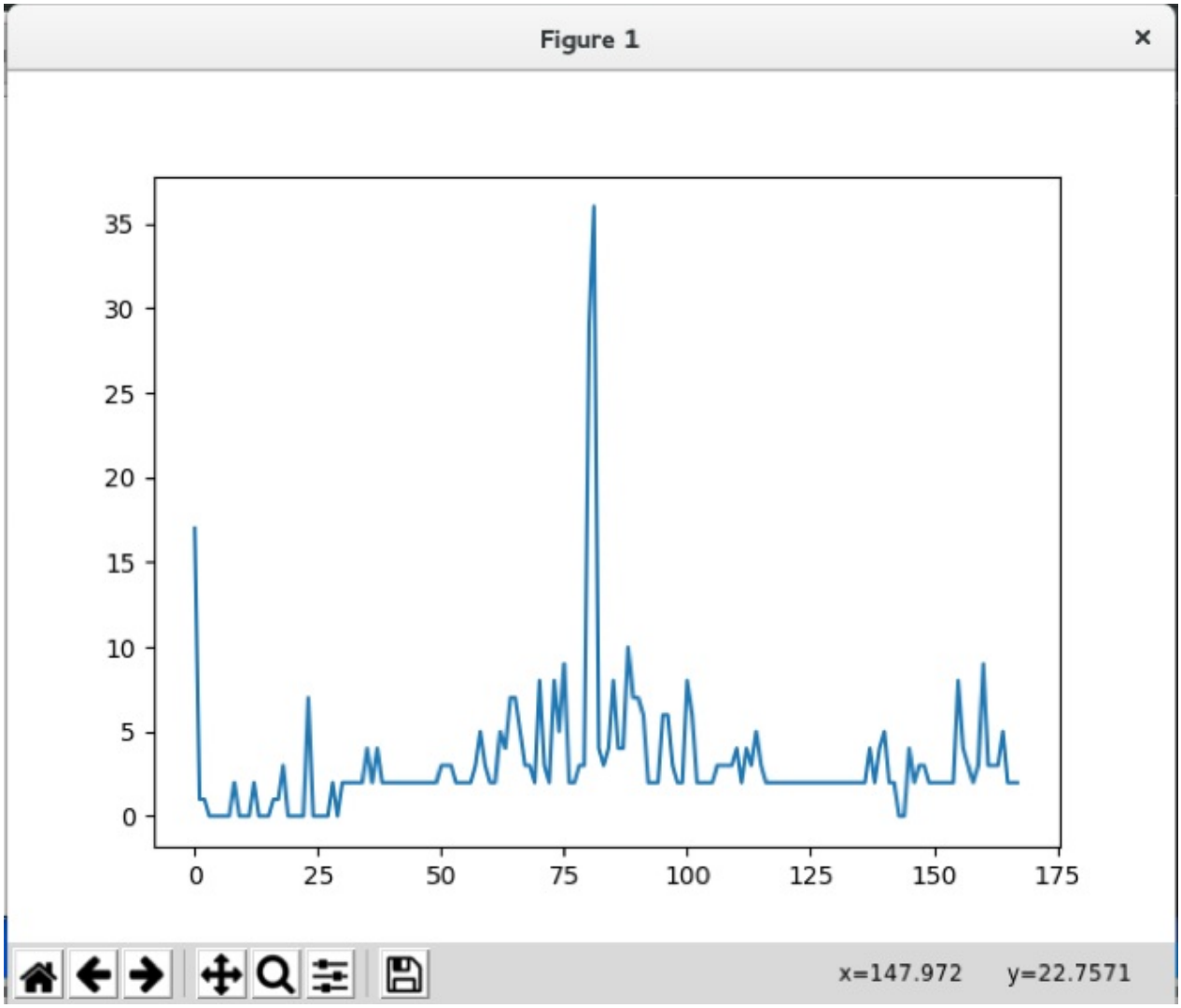
Step 3. Compare the real time series versus the generated time series:

- The command in this step should be executed on the preconfigured VM.

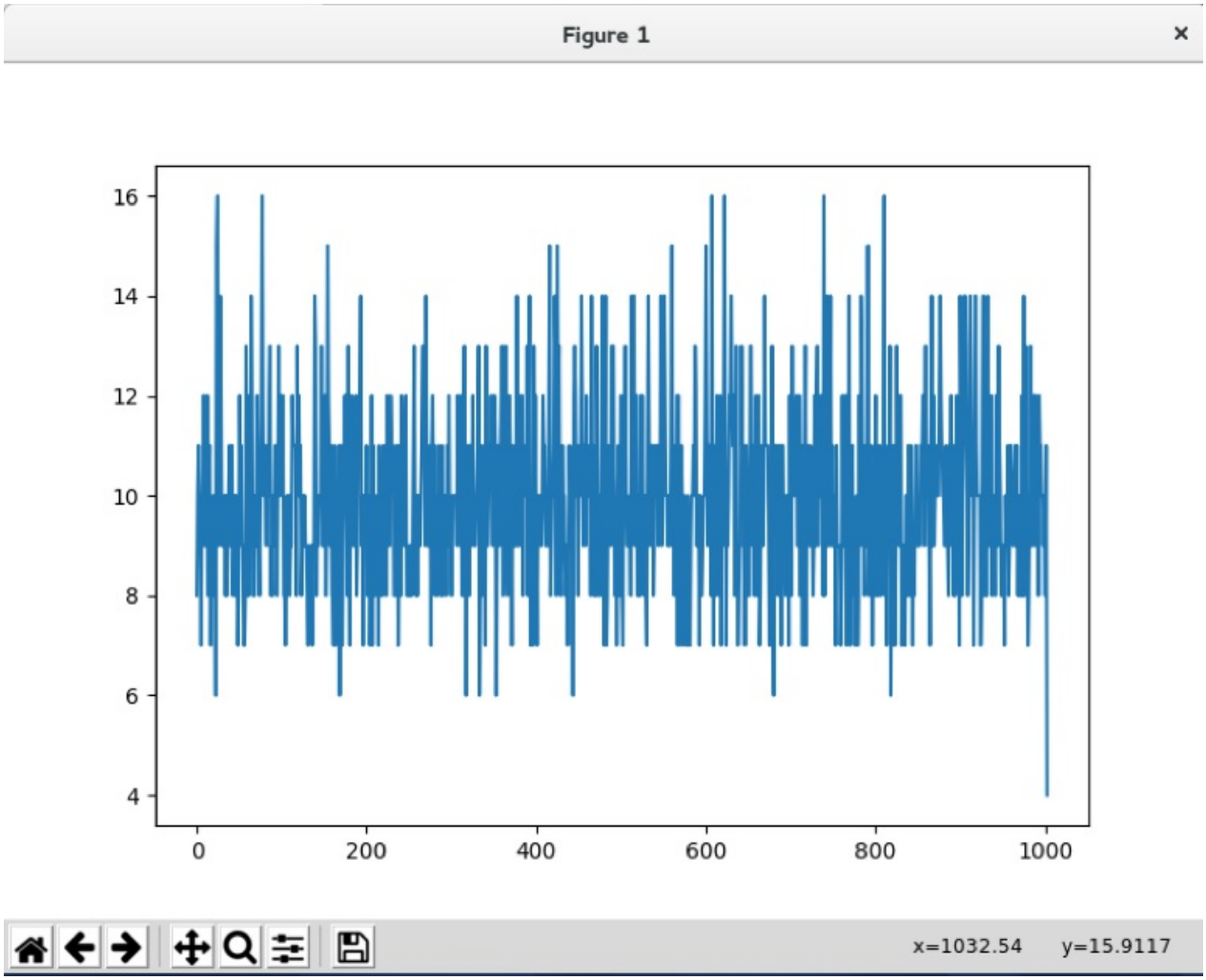
Execute the following command will generate two figures. (the captured pcap file should be in the same folder with scripts, and remove other pcap files).

```
student@localhost~$ python plot_time_series_example.py
```

The first one is the time series from the real campus background traffic, may need zoom in the get the figure below. **If doesn't work, use wireshark to plot time series graph for captured traffic.**



After close the first one, the second one will appear, which is the time series of simulated background traffic:



Step 3.1. Modify the perimeters of the simulated arriving times:

Use your favorite text editor such as VIM or NANO to edit **plot_time_series_example.py**. The line 28 of the plot_time_series_example.p_y have the **option**s (whose details can be seen in rand_arr_time.py) to change different arriving time algorithms. The available values are integers between 0 to 6. **The number_of_packets** refers to the total number of time stamps and **expected_duration** modify the total duration from the first time stamp to the last time stamp.
Modify those options and see if you get a figure that closer to the real background traffic.

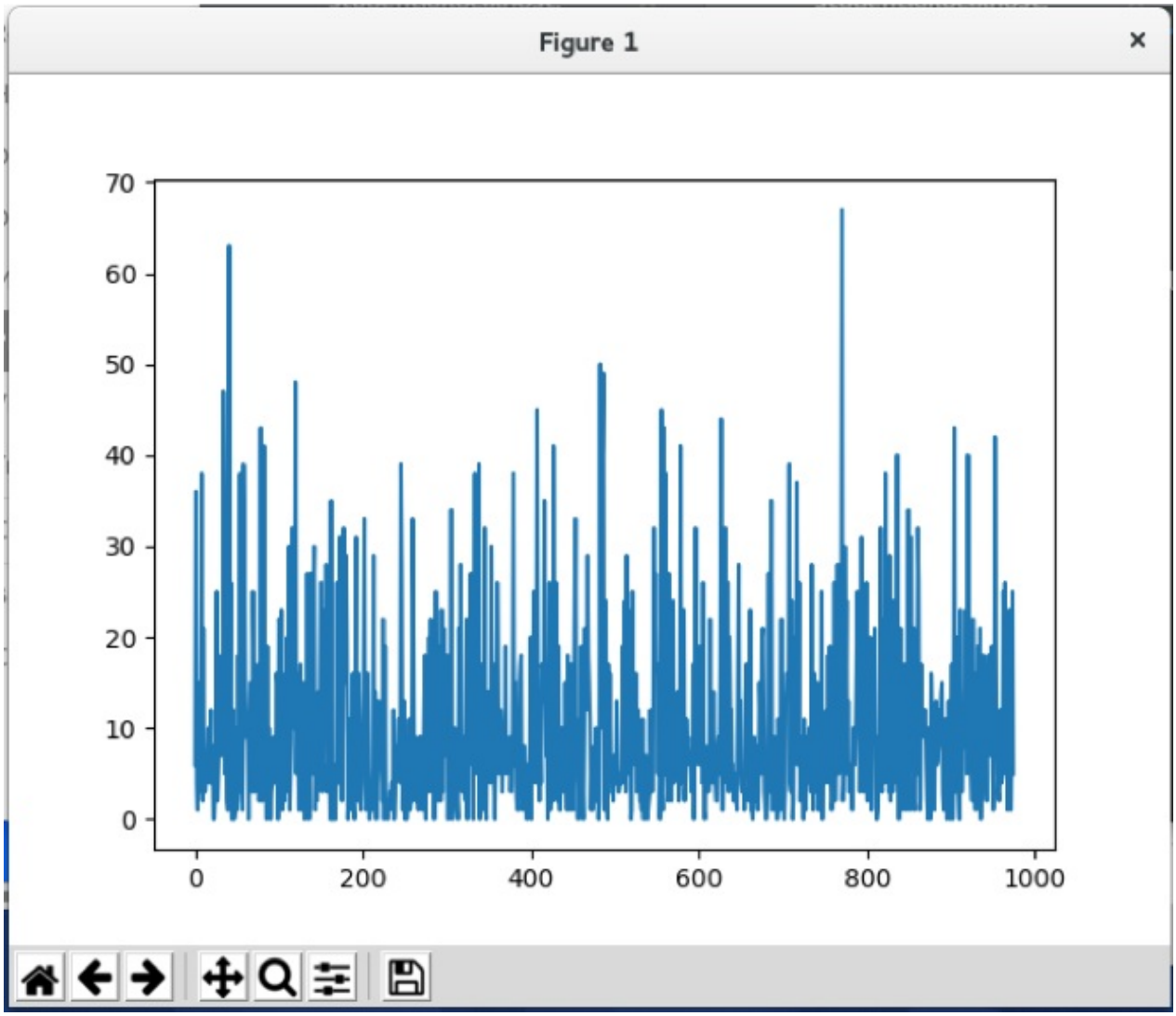
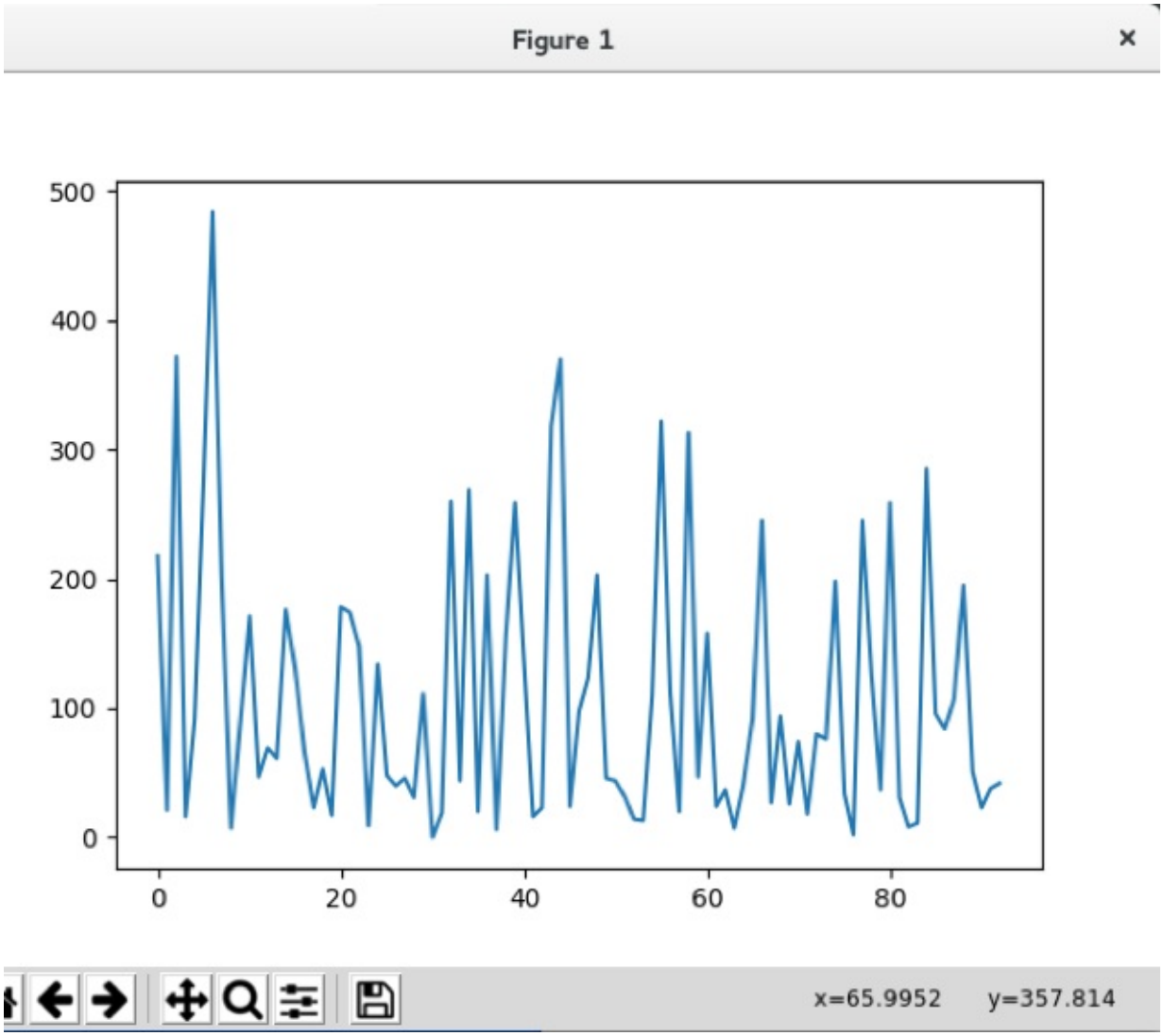


Figure 1
70
60
50
40
30
20
10
0
0
200
400
600
800
1000



Questions:

Q1 Plot the background traffic datasets generated/used in three scenarios where x-axis shows time, and the y-axis shows the number of packets received. Compare the figures and discuss their difference.

Q2 Network background traffic statistic generation script (plot_time_series_example.py) generates datasets using different probability distribution functions (PDF). Compare datasets generated using different PDFs. Discuss the difference between datasets and the data-set converted from operational network data.

Q3 What kind of traffic did you run between hosts in the second scenario (Replay a pcap file with tcpreplay). Justify why you think it can represent operational network traffic.

Q4 You can replay and forward captured pcap files on a link in a controlled network environment to use as background traffic. If you perform a DDoS attack on this link, you can observe the effects of the DDoS attack without jeopardizing the operational network.

However, some of the effects cannot be observed with replayed/forwarded background traffic. List some of these effects and explain the reason why they can not be observed.

Q5 Number of packets received by a node on the network is one of the popular metric used in DDoS detection applications. In this assignment, we focused on packet count statistic. Discuss what other metrics that can be used for DDoS detection.