

**CMPE 209-03**

**Network Security**

**Spring 2017**

**Network Traffic Analyzer**

**Interim Project Report**

Under the guidance of

**Prof. Chao-Li Tarng**

**Project Team 12**

Kshama Shalini (010763792)

Rohit Kotalwar (010805080)

Shivang Singh (010939851)

Vinay Puttanna (010699858)

April 18, 2017

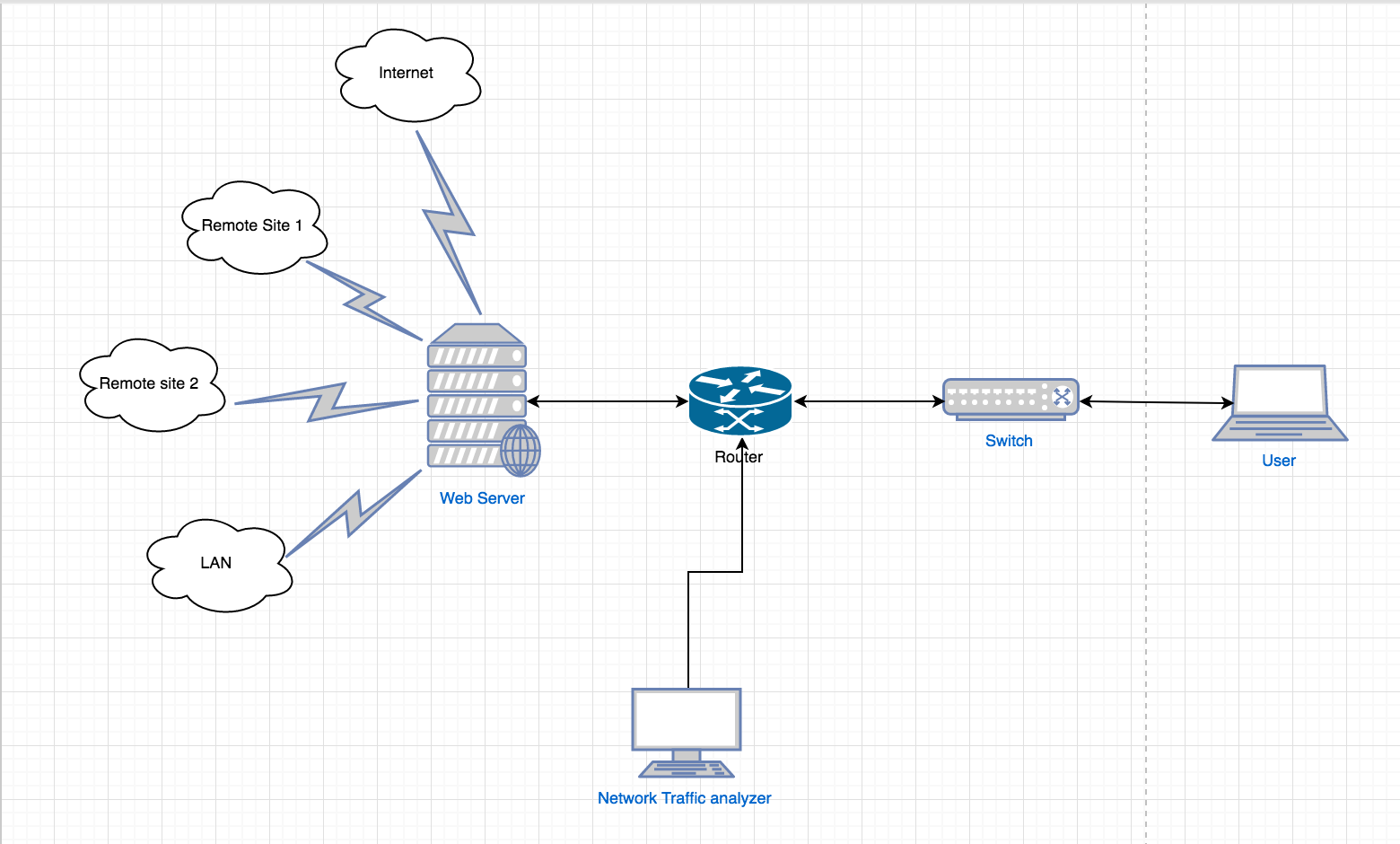
**Overview**

Cyberslacking is the big challenge nowadays for organizations and institutions across the globe. It may lead to legal issues that could cause severe damages to their reputation. So far, the measures adopted to handle the cyberslacking includes framing of various Internet policies and blocking websites that pose a threat to their network security. These measures failed to control cyberslacking completely because of various reason. The reason behind this is due to the creation of thousands of new websites, And thus making the process of monitoring the malicious website extremely complicated.

Although measure are being taken into consideration, but still even after several precautionary measures, people tend to access resources from the Internet that causes Copyright Violations or are inappropriate or prohibited content. Here at this point, By using Network Security tools we can understand the browsing pattern or behaviour of users worldwide. By using these tools we can easily locate the user on the geographical map who is accessing the malicious website and also we can locate the organization which host thce malicious website in the internet.

In this project, we are developing Network Traffic analyzer, to develop this project we will be using different types of tools such as Snort, Wireshark, Python Script, Python PyGeoIP Module, DPKT module and Google Maps. The aim behind developing this project is finding out the geographical location of the user who is accessing a malicious website and the location of the malicious website. The physical location may help investigators to find the location of malicious website to stop and prevent the attacks generated from such websites. We will be using Linux machine to execute the application.

**Project Architecture**

****

***Fig. 1- System Architecture***

Two approaches to capture network traffic:

1) Using wireshark and DPKT Module.

In this approach we will be using wireshark tool to capture the network traffic. This captured traffic in PCAP file will parsed using DPKT module. The output of this parse file can be used to find the source IP address and destination IP address. By using these IP addresses we will find the source IP address which are accessing the the malicious websites.

2) Using SNORT IDS and application script.

The another approach to capture traffic is by using SNORT Intrusion Detection System. We will be installing SNORT on local machine. SNORT will listen the traffic on the basis of rules configured.

Example for Rules:

**alert icmp any any -> $HOME\_NET any (msg:"ICMP test detected"; GID:1; sid:10000001; rev:001; classtype:icmp-event;)**

Explanation of above rule: here the alert will be given by SNORT IDS when an ICMP packet from any port and any IP address comes to any port of Home network

Starting Snort on interface of the machine:

**$ sudo /usr/local/bin/snort -A console -q -u snort -g snort -c /etc/snort/snort.conf -i eth0**

We will capture the logs from SNORT and we will store the captured logs into the text file.

These captured logs within the text file will be parsed using application script. The output of the parsed application script contains source IP address and Destination IP address.

After getting the IP address of source and destination we will check which source IP address is accessing malicious website.

After finding the source IP address which is accessing the malicious website, we will locate the source IP address and malicious website IP address.

To locate the map the IP address to the geographical location we will be using PygeoIP python library.

**Approaches**

* **pyGeoIP (IPtables)**

We will be using pyGeoIP tool for getting the Geolocation of the IP address. Python GeoIP or pyGeoIP is a python library and a python tool used to provide the geolocation information of the particular IP address. The library was developed on the top of PHP GeoIP tool. The library takes the path of filename as parameter for the creation of GeoIP class. This filename is the path to GeoIP database. GeoIP class which is created takes the IP address as parameter and returns the Geolocation information of the IP address. The Geolocation information such as country, region, city, organization, Internet Service Provider and ASN can be found out using pyGeoIP. pyGeoIP uses the following databases to access the Geolocation of the IP address.

* **Wireshark**

Wireshark, an open-source network traffic analyzer is used to sniff the packets in the network. Also, it will be used for network troubleshooting and for the analysis of packets.

* **Dpkt Python Module**

The Python module which is used for decoding the PCAP file. By using the Dpkt module we can parse PCAP files from the Wireshark and then we can inspect each packet in the PCAP file.

* **Python Scripting Language**

We will be using Python scripting language to develop the application script.

* **SNORT IDS**

We will be using the SNORT IDS to collect the traffic. The traffic will be captured according to the rules specified in the configuration.

**Project Plan and Schedule**

Following tables shows the project flow of our project.

|  |  |
| --- | --- |
| **Month/Week** | **Task** |
| April 1 | Installing Python, PyGeoIP, DPKT Module and Wireshark. |
| April 10 | Collecting the traffic through Wireshark. |
| April 15 | Parsing PCAP file using DPKT module. |
| April 25 | Developing Python script for analyzing parsed PCAP file. |
| April 30 | Converting the analyzed data from PCAP file to .KML file format. |
| May 2 | Displaying Data from .KML file to Google Maps |

**Accomplished and planned activities**

We have successfully set up the environment needed to develop our project. We used virtualBox to create Ubuntu Linux virtual machine. We installed required tools like wireshark on our virtual machine. We also installed Python and its libraries and modules like PyGeoIP and DPKT Module. Along with that, we have collected the traffic and captured the packets using wireshark and stored those packets as pcap file.

During the course of our project, we would aim at finding the physical location of the blacklisted sites. We will use the already captured pcap files for this purpose. We will use a malicious database, that will contain the list of all blacklisted sites that are malicious and can harm the users in any way. We will use scripts to parse the packets to find the source and destination IP addresses. And then we will use PyGeoIP to map IP addresses to their corresponding physical addresses. We will further map the physical locations to Google maps for better visualization. We will also setup a proxy server to sniff usernames and passwords. Then we will also identify users who downloaded any content from such blacklisted sites caught in the above procedures.

**Issues and Risks**

Python language is being used to develop application script for this project. It is new to all the team members and there was a substantial learning curve involved initially. The final output of our project is a web application that serves the objective of finding and viewing the geolocation of IP addresses. For this, we had to decide among the available Python web frameworks. The three most popular Python web frameworks available were Django, Pyramid and Flask. We have thoroughly studied the pros and cons of each framework before deciding upon the Pyramid framework. Though it provides less options than Django, it provides more flexibility for choosing external libraries, database. The learning curve is more as it is a new framework to work with for all team members.

Per our application algorithm, the Python script generates the KML output in the plain text format in the shell. This KML output has to be exported to a KML file which acts as an input to the Google Maps API to display geolocations of the IP addresses. Currently, this step has posed a challenge for us to achieve in our project.

We have also confirmed that geolocation tracking of an IP address does not pose any legal issues because it does not cause any privacy and security risks to the user. Location tracking does not provide the exact location of the IP address. Instead it provides only the city and zip code information.