Assignment 5: Anonymity

eSoc 488: Information Privacy with Applications

Due: 13 November 2018

Total Homework/Assignment Points: 100

---

1 Simple auditing

Even simple investigation of web traffic can be used to learn potentially-sensitive information. Here we set up a web server, and observe what information is visible to a few different perspectives. For this section, write your answers on their own line in a file called 'auditing' (no extension needed). Your mitmproxy dumps (described below) should be called 'dump_1' and 'dump_2.'

1. (10) Set up a simple web server using the following steps:
   (a) Set up a virtual private server (VPS) running Debian on Digitalocean.[1]
   (b) Disable password authentication, and use an SSH key to access the server with a new account (not root).
   (c) On your VPS, install ufw, and use it to ensure that your server drops all traffic that is not an SSH session from your current IP address, or a web request from 127.0.0.1 on port 8080.
   (d) Install apache2. Edit the file /var/www/html/index.html to contain exactly the string 'DOWN WITH CYBERCRUD'.
   (e) Add the IP address of your server to 'auditing'.
2. (5) Look at your logs
   (a) Use tail to view the access logs on your server *as they are changing* (you need a flag for tail). Visit your site. Answer the following in the file for this section:
      i. What information is visible about you?
      ii. How could this be used to learn more?
3. (5) Use mitmproxy to record a flow for a site you're visiting (i.e., filter for just the traffic for the site itself, not third parties). Save the dump as 'dump_1.' Now do the same for a third party advertiser or analytics company on a website. Create a text dump of the third party flow as well, and save it as 'dump_2.' Now search the text for something you find interesting, and record it in 'auditing.'

2 Tor[2]

You will submit these files for this section: your tor configuration file, 'tor_getinfo.py,' and 'geo.'

[1] Or a similar service. These instructions don't assume anything about the VPS provider other than that it allows you to run a Linux instance.

[2] Thanks to Maximilian Golla at RUB for sharing the questions in sections 3.2 and 3.3 below, which I have modified for our purposes.

2.1 Onion Proxy

In this exercise we learn to interact with the onion proxy (OP) running on your client, and get information from it.

1. (10) Open the control port of the Tor daemon using the Tor configuration file.[3] Then, write a python script called 'tor_getinfo.py' that uses Stem to get information about your Tor connections.[4] *You must include detailed comments in this script explaining what each significant part is doing.*

    - List the nodes used, with the IP-address and the fingerprint for each open Tor connection. Use the GeoIP Database[5] to find the locations of your Tor nodes. Record this in the 'geo' file, comma separated, with one line per node.

2.2 Specifying parts of the circuit

Tor allows to define the entry and exit nodes by modifying the config file torrc. It is your task to use the following entry and exit nodes, which are identified by fingerprints.[6]

Guard nodes:

- D529E870E7CCFCDA2CFEE9D317A8DC6E85497FDA
- C38286764201C7F0CDCC928ED59F2180F067C49D

Exit nodes:

- E4D1F25DFBE484208866BA4A1A958B73127CB0AD
- E6FAC9A7F33EE66F03C55C119770B2D45D3C576B

Answer the following questions in a file called 'tor_node_selection.'

1. (10) List the nodes used, with the IP-address and the fingerprint for each open Tor connection and use the GeoIP Database[7] to find the location of your nodes.
2. (5) How does a strict selection of the Tor exit and entry nodes influence the anonymity of the connection?
3. (5) Is it possible to exclude nodes? When should this be done?

2.3 Tor Bridges

Bridges are Tor relays that are not listed in the main Tor directory. Since there is no complete public list of them, even if your ISP (Internet Service Provider) is filtering connections to all the known Tor relays, they probably would not be able to block all the bridges. If you suspect your access to the Tor network is being blocked, you may want to use the bridge feature of Tor. It is your task to establish a connection to the Tor network via a bridge.[8]

1. (10) Get a bridge that you can use to connect to the Tor network and give its IP, port and fingerprint. How did you get these pieces of information? What are other methods to get a bridge (be sure to list them all)?

[3] The documentation can be found at `https://www.torproject.org/docs/tor-manual.html.en`.
[4] Stem documentation is at `https://stem.torproject.org/`.
[5] Found at `https://www.maxmind.com/en/geoip-demo`
[6] See footnote 3 above for the documentation.
[7] See the link in footnote 5.
[8] See `https://www.torproject.org/docs/bridges.html.en`

2. (5) In which situations is the application of bridges useful? Does the method to get a bridge that you chose before work in such situations?

2.4 Set up a Tor hidden service

Your next task is to set up a hidden service on the Tor network.

1. (20) Using your VPS, install tor, and set up a hidden service using the tor documention.[9] On your *local* computer, use the tor browser to visit your site (the address is found in $HOME/hidden_service/hostname). This site will need to remain available until you have received a grade for this assignment (not more than one week from the submission deadline). Turn in the onion address for your working hidden service in a file called 'hidden_service_address'.

Please answer the following questions about hidden services:

1. (5) How is the service hidden, i.e., why can users not reveal the location of the service (in the network) and still use it?
2. (5) Is there a way to still reveal information about the operators or the location of a hidden service? Monitor the connections to your hidden service (e.g., by watching the access log file).
3. (5) Can you distinguish between different users? Would this be different if your service was not a hidden service (and still accessed via Tor)?

[9] See `https://www.torproject.org/docs/tor-hidden-service.html.en`