Blocking Privacy

A Final Project Report for CS 4501 - Privacy in the Internet Age

By

Ans Shaukat (hgn9pv) and Heru Avila (qns7rq)

Project Overview

For our project, we created a web app through Django which is deployed on Heroku that explores the information a web server can glean from the client through various features:

- Client Information from IP Address
    - Based on the IP address of the client, we use an API to get all available information including: Country, Region Name/State, City, Zip code, Lat/Long Coordinates, Timezone, ISP, and of course IP
- Browser Information
    - By reading the HTTP header, we can get info such as what Browser the client is using, browser version, OS type and version, and what kind of device the client is using
- Cookies
    - We used cookies for various purposes, we store the number of times someone visited a website, the date and time they last visited the website, the last location they visited the website from and also their “home location” which is the location used for VPN detection. All the cookies had an expiration of 90 days, but “home location” cookies reset every 12 hours, just in case someone is traveling.

- Tor Detection
    - We use a simple method to detect Tor connections which is by retrieving the list of exit nodes that the Tor service publishes and checking if any on the list match the Client's IP. This can be circumvented if they use something like Tor Guard, which is a VPN service, but it frequent uses of it would be caught by our VPN detection
- VPN Detection
    - For VPN checking, we are using location. Whenever a user visits the website for the first time we save their location into "home location". Then every time the user visits the website again, we check their location with the saved "home location" which is saved in cookies. If the user is in a different city, but same region and country then we treat this as VPN not # being used. If the user is in a different region and same country or just entirely different country we treat it as VPN being used. We are currently allowing the home location to be reset every 12 hours, just in case someone is traveling. For example, if you connect from Charlottesville, Virginia, the app will set that as your "home location" for 12 hours and if we detect that the client connects from outside the state in that time period, we recognize it as a VPN. While this is a somewhat naive approach, it still prohibits common or frequent uses of VPNs.

- Browser Version Status

    - From the Browser information, we can make an API call to find the latest version of their browser and compare it to their version and inform the client if their browser is up to date or not.
- Geolocation specific content/Tor Specific Content
    - Based on previous features, we can give specific content based on the location of the client or if they are connecting via Tor, currently we are just detecting if they are within or outside the US for geo-specific content. We have three tabs on our website for U.S Only, Non U.S and Non Tor, if a user visits the U.S Only tab while being outside the U.S then they are taken to an access denied page. If a user outside the U.S tries to visit the U.S only tab then they are taken to the access denied page and finally the same is true for Tor users trying to access tab for Non Tor users. We did this to show how websites can use your personal information to show you specific content and block you from other content.

Discoveries

By creating this web app we were hoping to discover how easy or difficult it would be for companies/websites to collect information from their clients and to detect if they were trying to protect their privacy through the use of services such as Tor or VPNs. There were 3 main areas where we learned the most but we were also surprised just how easy it was to get information such as location, device type, and browser type just from the client connection. The 3 main areas that we learned about were Tor connections, VPN connections, and cookies.

For Tor connections, we were surprised at how easy it was to block or detect if clients were using it as all we had to do was retrieve the list of exit nodes. While it is hardly a

bulletproof method to block all Tor users, it allows us to block or redirect traffic for a majority of Tor users who won't go beyond just using the basic service.

For VPN connections, we learned that implementing any method to detect VPNs was complicated and there always seemed to be some way to circumvent detection. We didn't opt to try and compare IP lists as we already did that for Tor and we wanted to explore other methods. Our method is a bit naive but we believe it can stop a good amount of VPN traffic or disallow clients from spoofing multiple different locations in a short amount of time. Overall, our implementation gave us insight to just how difficult it is to block/detect VPNs and how much privacy it can afford its users.

Finally, for setting cookies we were pleasantly surprised about how easy it was to set cookies, we only had to follow this format: `response.set_cookie('key',value,expiration date)` . We learned that cookies can be really helpful for functionality purposes, for example for us we are using the "home_location" cookie to detect VPN usage, but another thing that we learned was that cookies are not user proof, if a user decides to delete their cookies every time they visit our website then they will be able to bypass the VPN detection system. Overall, setting cookies showed us how easily any website can save information about us and then use it to profile us later on.