Information Security and Privacy
Assignment 1 , Part 3

Lightbeam is a Firefox **add-on** that **monitors** third-party trackers. Lightbeam works by creating a **snapshot** of the websites you visit and **any** third-party trackers **running** on those pages. This allows users to see how many trackers are **following** them and which websites **are** sharing their data with third **parties.**

Third party trackers are cookies or other technologies placed on your computer by a website other than the **website** you are currently **viewing.** They are often used to track your **search** activity **on various** websites for advertising purposes.

Third-party trackers are **a set** of **codes found** on **many** websites. They collect **information about users' browsing history** and send **it** to other companies, often for advertising purposes. If the same third-party tracker is **available in multiple locations,** it can **create** a **full** profile of the user over time.

Below are the most common tracking mechanisms:

Cookies are the most widely known method of identifying users. They use small files (only 4 KB per chunk) stored in the browser by the web server. When the user visits the website for the first time, a cookie with a unique user ID (which can be generated by the user's computer) is stored on the user's computer.

You do not need to log in the next time you visit the Facebook page, because the browser will remember your information from the cookies stored when you first logged in.

Browser fingerprinting is a very effective method of identifying and tracking users as they browse the internet. The information is very well written and generally includes browser type and version, operating system and version, screen resolution, supported fonts, plug-ins, time zone, language and font preferences and is independent of hardware configuration.

These statistics may be created from general materials and are not fully personally identifiable. But usually one person in a million has the same characteristics as you.

Web beacons are small, often invisible objects on web pages or emails. A web beacon is also known as a "web bug"; also known as a "tag", "page tag", "tracker", "pixel tracker" or "pixel gif".

At their simplest, these are small, precise images, usually only one pixel in size. When a website loads or an email is opened, they are downloaded as images and called to the remote server to retrieve the images. The server calls companies to notify them that their email has just been opened or their website has just been visited

Unexpected Findings:

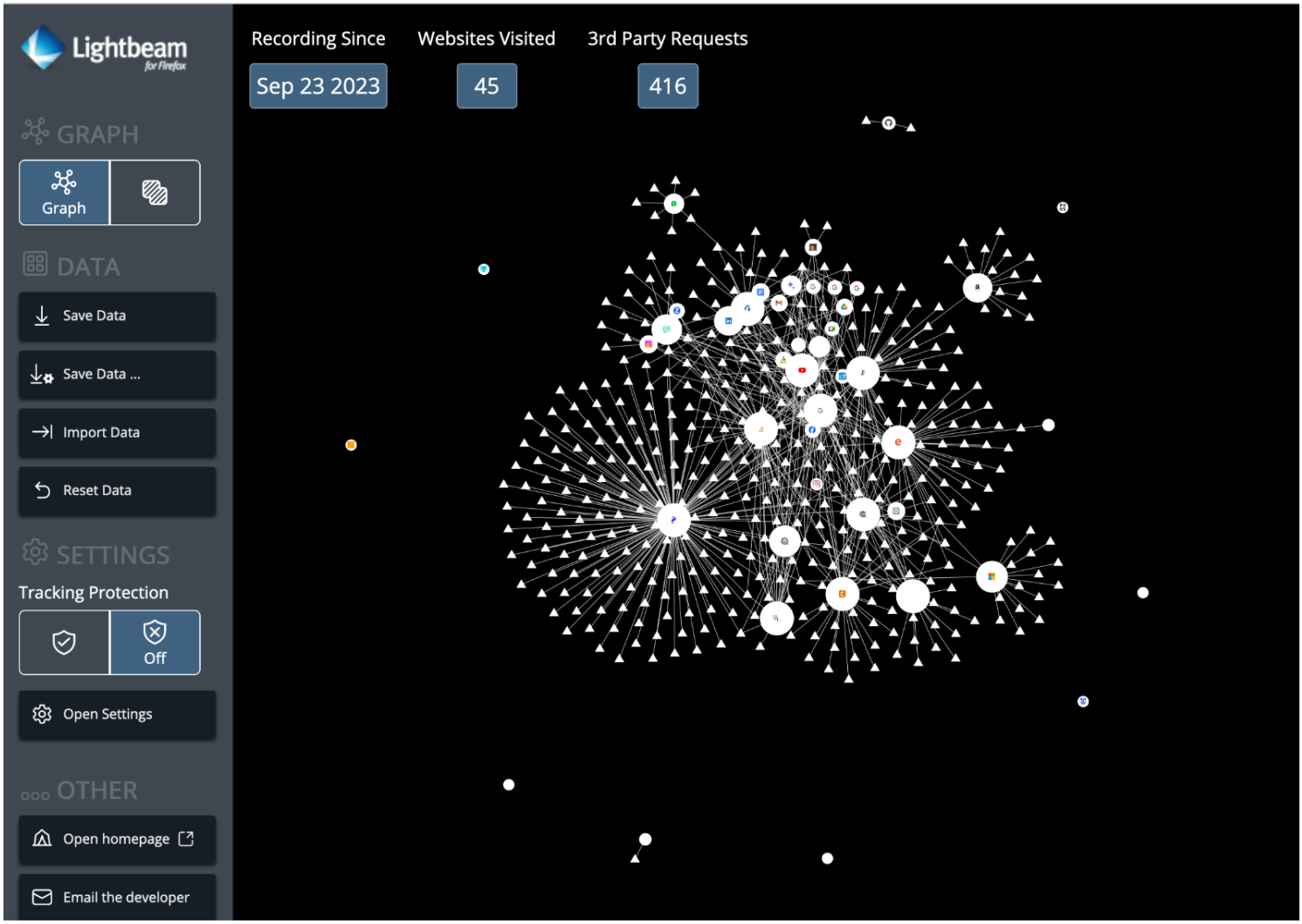
The analysis of the web trackers using lightbeam gave me lots of new findings which i did nor expect. One or the major finding was

1 : Big Tech Companies usually blamed for stealing and selling data are just tip of an iceberg :

Below is the screenshot of Lightbeam visualization of 3 days worth of browsing the web for various purposes. Majority of my search consisted of

- browsing on google
- opening facebook
- Watching youtube
- Scrolling instagram
- Reading tech articles and so on.

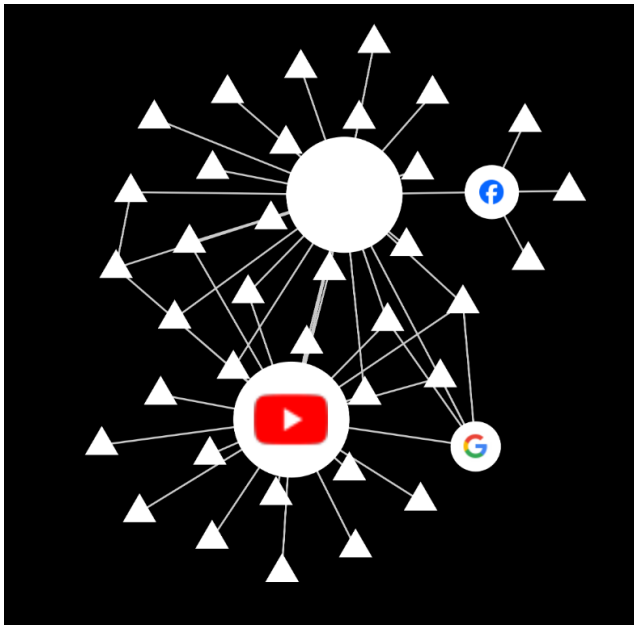
By a single look at the below graph one can infer that our daily activity is being watched by hundreds of trackers and multiple companies are benefiting from it. Whats surprising is when the number of trackers watching us when we visit tech sites like Programiz , StackOverflow and Eventbrite is much higher than the trackers you find on amazon, instagram and facebook even.
Below image clearly shows that Programiz has the highest trackers out of all.

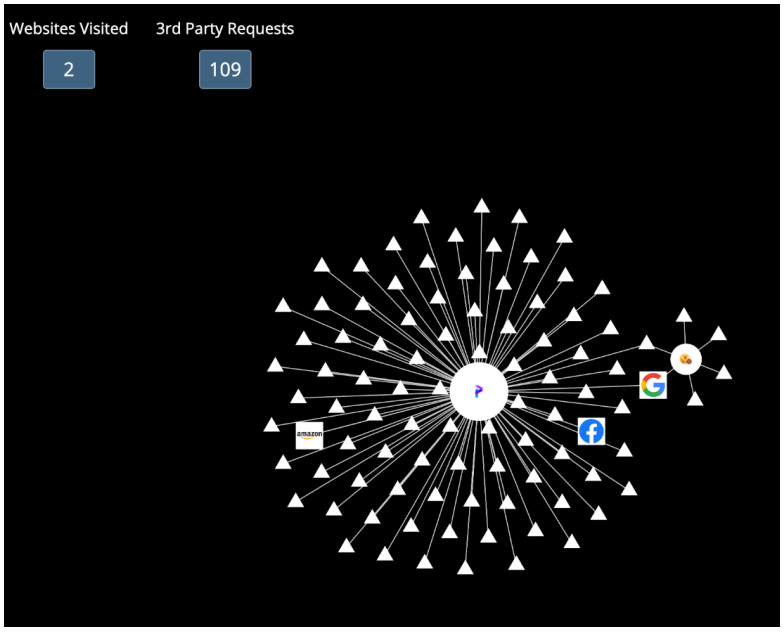


Inorder to get clear picture of the amount of tracker, I cleared the browser cache and visited each of the website individually through google search engine.

Below are the screenshots of the same:

- Programiz has 109 trackers tracking it.
- In the case of Facebook visualization, Even though I haven't visited youtube, a request has been made to youtube and that intern has few other trackers getting info from it.

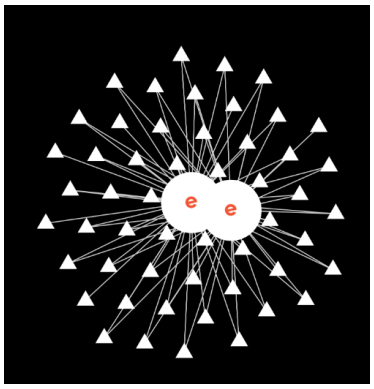


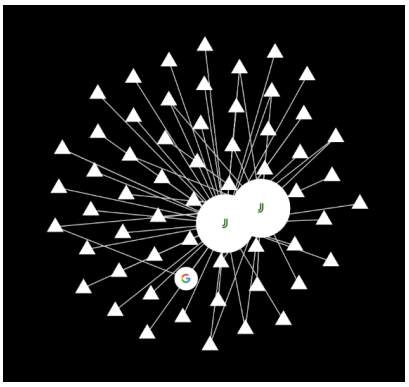


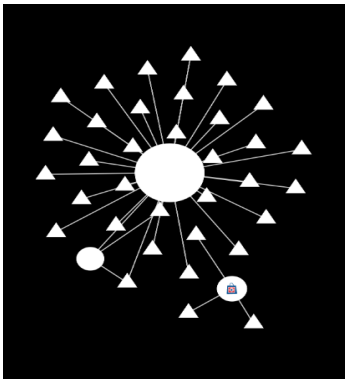
Facebook Trackers Visualisation

Programiz Tracker Visualisation

Few other websites with numerous trackers are as follows:







Eventbrite

Juniper Networks

HDFC Bank

2 : Companies you don't use might have more data about you than you think.

When i was looking at the third party trackers and their sources i found the that there are many other ad agency apart from Google and facebook ads. Few of the ones i found were

- Privacymanager.io -> the name is deceiving but its an ad agency
- Clarity.ms -> which is an analytic service run by microsoft.
- Amazon analytics -> analytics run by Amazon.
- Yahoo.com -> i do not have an account but still being tracked
- Snapshot.com -> i do not have an account but still being tracked.

3: Our data might not be secure

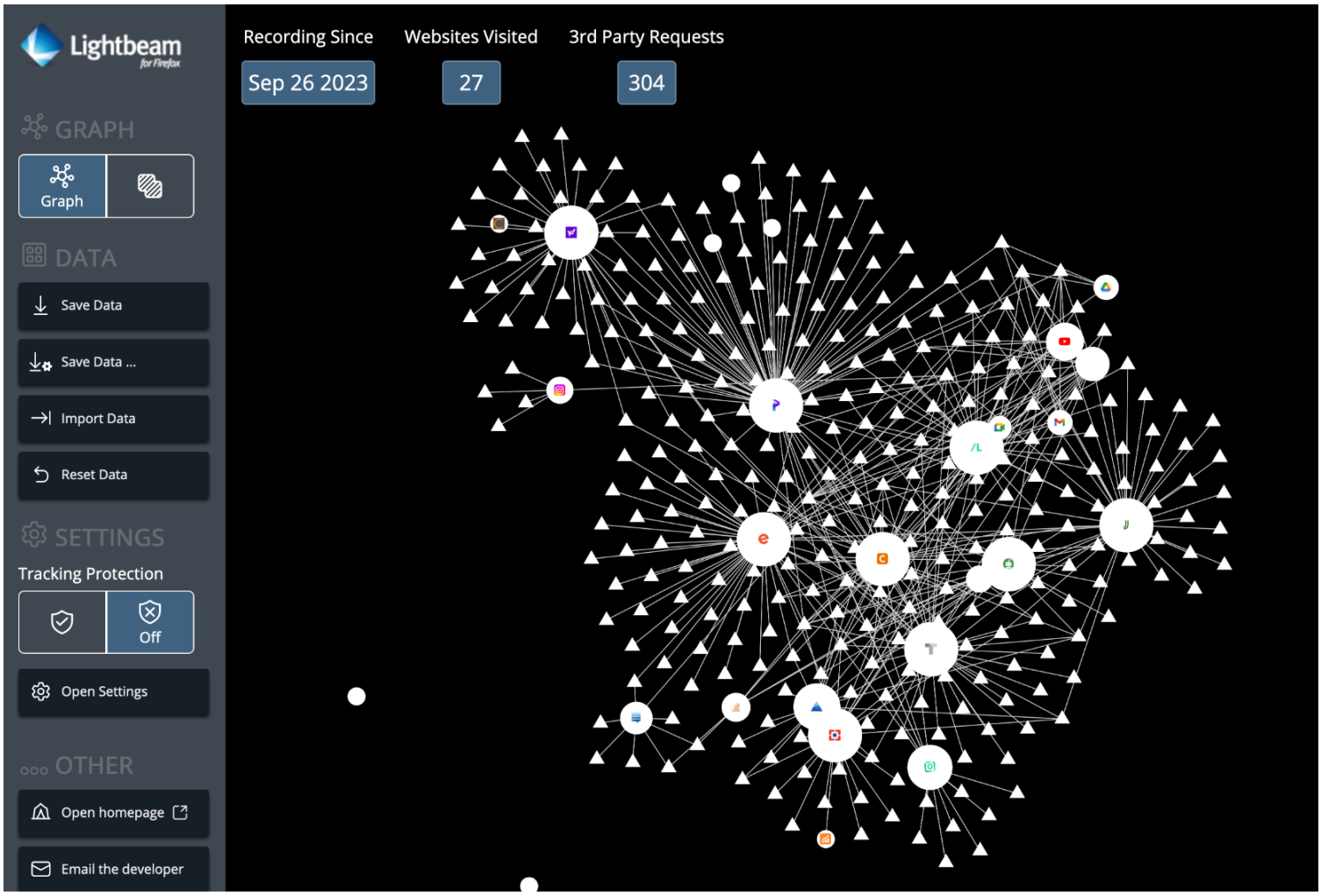
While trying to find the source of such trackers I found many unsecure websites which was blocked by browsers.
Yet they are given free access to our behavior. Few such websites i found was:

- Dis.criteo.com
- ads.pubmatic.com

4: Exclamation marks are not always what they seem.

One more interesting thing i found was Exclamation marks. When we see exclamation marks in a website out thought process is that its just another key in the typewriter.
But that is not the case. These seemingly lame characters are used to blend in **web beacons** inside the website which then notifies its source that the website was open and might also carry location and system info.

Snapshot of Lightbeam after clearing history and revisiting the website



If you compare the snapshots of 3 days of search and less than 5 minutes of search, we cannot find any drastic difference except the fact that the number trackers are a bit less in the second snapshot. But still 300 trackers are a lot to exploit and exploit the data that they are gonna get.

These trackers are so hungry for data that they start tracking the second you come online.