Eve Tanios

Privacy Lab

A screenshot of a computer

Description automatically generated**Unencrypted DNS Data Analysis**

Many of these entries are queries from the DNS recursive resolver until it reaches the authoritative DNS nameserver to bring the user to the correct website. However, some of these entries (not all shown in this image) are data collection services. According to CloudFare, the DNS recursive resolvers themselves often contain data centers, and this data is later taken for targeted ads, or sold for other uses.1

I chose the NYTimes for my test case, and after navigating to the website, noticed many DNS queries that were not for the resolver, and instead for data collection. For instance, Google, though *googlesyndication.com*, *beacons.gcp.gvt2*, and *doubleclick.net*, can see what websites a user has visited based on the DNS queries. This data is then used for targeted ad loading across the web. Another notable observer is *adsrvr.org* (The Trade Desk), which is a bidding ad exchange that collects “non-personally identifiable information that may include, but is not limited to, your IP host address, and pages viewed…age, gender, income, education, interests, and user activity from third parties”.2 Though this information is listed as “non personally identifiable, it can easily be connected to individuals when combined with other company databases, especially considering the wide range of data collected. This can create ethical issues relating to the deanonymization of data and the formation of huge databases that essentially create entire profiles for individuals, complete with increasingly personal information, in order to target ads for increased profit.

|  |  |  |
| --- | --- | --- |
| **Company** | **Domain Names** | **Concerns** |
| Google | Googlesyndication.com, beacons.gcp.gvt2, doubleclick.net, googletagservices.com | Many data collection domains, large possibility of the formation of a “master database” with combined information that creates a de-anonymized profile of user |
| The Trade Desk | Adsrvr.org | Collects “non-personally identifiable” information, but also mentions sharing data on interests, education, and third-party user activities. This is done largely without informed consent, and collecting information on “interests” is vague and broad. However, it is of great interests for other companies that wish to make profit on targeted ads. |
| Amazon | Awsdns-35.com | Amazon now hosts many websites through AWS, which means that they likely store much information about which websites are visited by the user. In general, this is also a concern because the user isn’t informed that through visiting the NYTimes, they are also informing Amazon about their potential interests and news sources. |

These companies gain access to the user’s website visiting data through the DNS queries. When the recursive resolver queries a DNS root nameserver, the nameserver responds with the DNS server, which holds information about its domains. These servers can store information about the user in this querying process, and they can also be connected to these large companies (such as AWS, which hosts websites, and Google).

**Encrypted DNS Data AnalysisA screenshot of a computer

Description automatically generated**

My data shows that the ad tracking entries are much less prevalent when over encrypted DNS. Instead, the DNS queries picked up by Wireshark show only the threads used by Apple (which may or may not be collecting data for ads). Without encryption, the user’s ISP, transit providers, and theoretically other people on the WiFi network could spy on the websites that are being visited and the modify DNS queries and responses in malicious ways. With encryption, the query details are less visible on Wireshark, since the DNS message is embedded into a secure TLS channel. This means that the ISPs and resolvers are hindered in collecting data on the user.3

A potential issue regarding this encryption is that the user must trust that the company, such as Nord VPN, is not selling the user’s data or using it for other purposes as well.

Interestingly, I found that Cloudfare is the only provider that satisfies Mozilla’s strict resolver policy.3, 4 This policy entails the following:

* The resolver may only retain user data for the purpose of operating the service and only for 24 hours
* The resolver must not retain, sell, or transfer to any third party any personal information, IP addresses or other user identifiers, or user query patterns from the DNS queries
* The resolver must not combine the data that it collects from queries with any other data in a way that can identify individuals

Of course, there is still a layer of trust to this system, and with any aspect of DNS queries over the internet. It appears, however, that Cloudfare is working on ensuring that there is privacy and security for free for users on the internet.

**Sources**

1. <https://www.cloudflare.com/learning/dns/what-is-dns/>
2. <https://www.thetradedesk.com/us/privacy>
3. <https://blog.cloudflare.com/dns-encryption-explained/>
4. <https://wiki.mozilla.org/Security/DOH-resolver-policy>