# Characterizing CNAME Cloaking-Based Tracking on the Web

Ha Dao
The Graduate University for Advanced Studies (Sokendai)
Email: hadao@nii.ac.jp

Johan Mazel
ANSSI
Email: johan.mazel@ssi.gouv.fr

Kensuke Fukuda
NII/Sokendai
Email: kensuke@nii.ac.jp

***Abstract*—Third-party tracking on the web has been used for collecting and correlating user's browsing behavior. Many techniques have been developed to protect end-users from third-party tracking. However, third-party tracking has been getting more sophisticated in an arms race against these countermeasures. Due to the increasing use of ad-blocking and third-party tracking protections, tracking providers introduced a new technique called CNAME cloaking. It misleads web browsers into believing that a request for a subdomain of the visited website originates from this particular website, while this subdomain uses a CNAME to resolve to a tracking-related third-party domain. This technique thus circumvents the third-party targeting privacy protections.**

**The goal of this paper is to provide a thorough investigation of CNAME cloaking usage for web tracking. To the best of our knowledge, this is the first in-depth analysis of CNAME cloaking-based tracking. We crawl top pages of the Alexa Top 300,000 sites, and analyze the usage of CNAME cloaking. Our results show that 1,762 websites (0.59%) contain CNAME cloaking-based tracking as of January 2020, primarily on Business and Shopping websites in the United States. By evaluating four snapshots of Alexa 100,000 sites from 2016 to 2020, we find that the usage of CNAME cloaking-based tracking steadily increases. It demonstrates that CNAME cloaking-based tracking is not a new phenomenon, and it has been already deployed for at least four years. Finally, to provide a wider picture of current privacy protection techniques, we evaluate the effect of well-known filters, browsers, and extensions against CNAME cloaking-based tracking. We point out that browsers and privacy protection extensions are largely ineffective to deal with CNAME cloaking-based tracking except for Firefox with a developer's version of the uBlock Origin extension.**

## I. Introduction

Web tracking is becoming more and more ubiquitous, this thus results in an increase of privacy concerns from Internet users. In a TRUSTe study, 92% of British Internet users concern their online privacy [1]. A website (first-party domain) has many links to other resources (third-party domains). Some third-party domains are used for user tracking (third-party tracking) to provide functionalities, such as advertising, analytics, and social network integration [2]. For instance, with third-party tracking, advertisements on a website can be customized based on end-users' visits to other websites, which can be frightful for privacy-sensitive users.

There are some existing approaches to detect third-party tracking. Many privacy protections take blacklist approaches to detect third-party trackers [3], [4], [5], [6]. Some works identify tracking requests using cookies [7], or fingerprinting [8], [9]. Other studies intend to detect third-party tracker automatically using machine learning [10], [11], [12]. These are effective against third-party tracking.

However, web tracking is becoming more and more intricate. One of the emerging techniques is the use of Canonical Name Record or Alias (CNAME) record in Domain Name System (DNS) to hide usual tracking domains that are blocked by browser filter lists and extensions.

For instance, website *example.com* embeds a first-party *a.example.com*, which points to a tracking provider *ad.com* via the CNAME *x.ad.com*. Obviously, domain *a.example.com* is not blocked by filter lists. Ad blockers such as AdBlock [3], Adblock Plus [4] or uBlock Origin [6] and other extensions are blindsided by the CNAME Cloaking technique because browser extensions are usually not allowed access to the DNS layer of web requests. The only exception is Firefox and its DNS API that is actually used by the newest version of uBlock Origin [13].

In this paper, aiming at characterizing the CNAME cloaking-based tracking on the web, we crawl the Alexa Top 300,000 sites, apply wildcard matching based on well-known tracking filter lists, and analyze the usage of CNAME cloaking-based tracking. The main results of the paper are as follows. We detect 1,762 websites (0.59%) containing CNAME cloaking-based tracking in Alexa 300K sites as of January 2020 by matching with tracker lists (§ IV-B and § IV-C). Those websites are mainly business and Shopping websites in the United States. They use 56 tracking providers in total, and the most common one is Adobe [14]. Next, by analyzing longitudinal snapshot crawled data of Alexa Top 100K sites (§ IV-D), we show that the usage of CNAME cloaking-based tracking steadily increases from 2016 to 2020. Finally, we evaluate the detection ability of such tracking for major browsers and extensions (§ V).

## II. Background and term definitions

### A. Background

*1) Third-party tracking:* Privacy leakage occurs through communications with trackers. Third-party web tracking refers to the practice by which an entity (the tracker), other than the website directly visited by the user, identify and collect information about web users.

From the view of website's administrators, user tracking is useful for a variety of purposes such as behavioral advertising

978-3-903176-27-0 / ©2020 IFIP

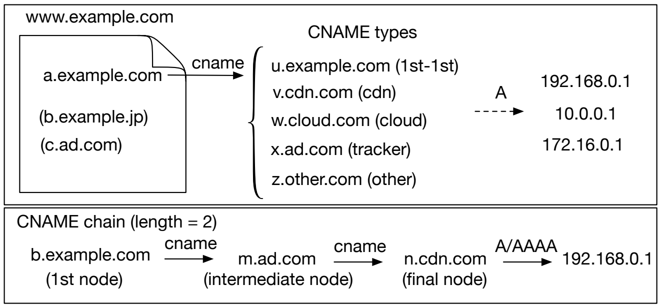

Fig. 1. Overview of CNAME cloaking-based tracking.

or website analytics. On the user's side, the larger number of browsing profiles, the greater loss of privacy.

*2) In-browser privacy protection techniques:* Several in-browser privacy protection techniques have been designed to protect end-users from third-party tracking, including extensions and browser itself.

Some anti-tracking extensions work effectively to detect third-party tracking, such as Ghostery [15], Disconnect [5], and uBlock Origin [6]. Some browsers also have built-in privacy protection features to protect end-users from third-party tracking, such as Firefox [16], Brave [17], and Tor Browser [18]. Firefox introduces Enhanced Tracking Protection (ETP) feature from Firefox version 69. It blocks user profile from browsing behavior observation across websites [19]. Brave has a feature called Shields which protects user's privacy by blocking ads and trackers, cookies, malicious code and malicious sites [17]. The Tor Browser is a browser based on the onion routing tool Tor and Mozilla's Extended Support Release (ESR) Firefox branch to enhance privacy and security. It includes both HTTPS-Everywhere and NoScript extensions which respectively enable HTTPS when possible, and allow users to block JavaScript [20].

## B. *Term definitions*

We first define some terms we use throughout this paper:

*1) Subdomain:* In the DNS hierarchy, a subdomain is any domain which is underneath a main domain. Subdomains are used to organize or divide contents of a website into specific sections. In this paper, we consider that a domain with prefix *www* is a subdomain. For example, *a.example.com* and *www.example.com* are subdomains of domain *example.com*.

*2) CNAME Cloaking-based tracking:* The usage of DNS CNAME records coupled with Content Delivery Network (CDN) is increasingly commonplace to improve website load times, reduce bandwidth costs and increase content availability and redundancy.

CNAME has also been used for user tracking. Tracking providers ask their clients to delegate a subdomain for data collection and tracking and link it to an external server using a CNAME DNS record [21]. We call *CNAME cloaking-based tracking* the usage of CNAME to disguise requests to a third-party tracker as first-party ones. For example, when end-users access the website *example.com*, this website embeds a first-party tracker *a.example.com*, which points to a tracking provider *ad.com* via the CNAME *m.ad.com*. Tracking provider *ad.com* thus tracks activities of end-users on the website *example.com*.

TABLE I
SUMMARY OF CRAWLED DATA IN ALEXA TOP 300K SITES (JAN 2020).

| Metrics | | | Numbers | Percentage |
|---|---|---|---|---|
| 3rd party requests | | | 14,640,568 | 54,27% |
| 1st party requests | domain | | 5,919,965 | 21.94% |
| | subdomain | w/o CNAME | 3,245,361 | 12.03% |
| | | **w/ CNAME** | **3,172,304** | **11.76%** |
| **Total requests** | | | 26,978,198 | 100% |

## C. *CNAME chain*

CNAME chain corresponds to a series of CNAMEs from the initial first-party subdomain to all CNAME nodes before the resolution to an IP address (see Figure 1). We consider four CNAME types for a CNAME chain:

- First-party type: The domain of the final node in a CNAME chain is the same as the domain of the considered HTTP request, or the IP addresses of both the final node and the second-level domain are the same (*u.example.com*).
- CDN type: The domain of nodes in a CNAME chain is used for CDN service (*v.cdn.com*).
- Cloud and others type: The domain of nodes in a CNAME chain is used for other activities, such as cloud storage or firewall (*w.cloud.com*, *z.other.com*).
- Tracker type: The domain of nodes in a CNAME chain is used for tracking user activities (*x.ad.com*). We name this *CNAME cloaking-based tracking*.

# III. METHOD

In this section, we describe the website selection and data collections, then explain our methodology to detect, characterize, and analyze behavior of CNAME cloaking-based tracking.

## A. *Websites selection and Data collection*

The first step is the selection of websites that would be most appropriate for our work. We use the popularity index from Alexa [22] in all of our measurements, similar to past literature [9], [23], [24]. To characterize CNAME cloaking-based tracking, we use OpenWPM [9] to conduct large-scale automatic crawls on Alexa Top 300K sites. OpenWPM is based on Firefox version 52 and allows collecting all the HTTP/HTTPS requests emitted and their responses for each site. We performed the crawls with default settings on January 2020, with three IP addresses in Japan (Table I).

In addition, in order to track longitudinal behavior of CNAME cloaking-based tracking, we also rely on four other datasets (see Table II). We collected two datasets on Alexa Top 100K sites with OpenWPM in April 2018 and January 2020. The other two datasets are publicly available in Princeton Web Census Data [9]. They were collected in January 2016

TABLE II
LONGITUDINAL SNAPSHOT DATASETS.

| Time | Alexa | List gen. | Requests | Firefox version |
|---|---|---|---|---|
| Jan 2016 | 100K | 01/2016 | 9,487,367 | 41 |
| Feb 2017 | 100K | 11/2016 | 10,964,374 | 45 |
| Apr 2018 | 100K | 03/2018 | 9,926,080 | 52 |
| Jan 2020 | 100K | 12/2019 | 9,647,506 | 52 |

and February 2017 and targeted Alexa Top 100K sites. These datasets were also crawled with OpenWPM, so all the data sources are compatible and comparable. Note that the contents of Alexa lists are not the same among these four datasets because Alexa lists themselves are updated daily and change significantly from one day to the next [25]. The list used for each crawl is described in “List gen.” column of Table II.

Furthermore, the instability of the Alexa Top list drastically increased in January 2018 [25]. So, in order to make a fair comparison, we also use the intersection (26,162 sites) of the four Alexa Top 100k sites above.

### B. CNAME lookup

First of all, we separate the generic Top Level Domain (gTLD) and country code top-level domain (ccTLD) from the visited website for all HTTP requests using the Public Suffix List [26]. We only keep subdomain of an HTTP request if it is not null and its second-level domain is the same as the visited website domain. We look up and check CNAME records for each subdomain. We then resolve each CNAME answer set by DNS. We save all nodes in CNAME chain (see § II-C) to analyze the CNAME cloaking behind first party requests. We find that 45.73% of the HTTP requests are first-party requests in 2020 (Table I). We then only keep 11.76% of the HTTP requests that contain first-party CNAME.

Looking up CNAMEs for the longitudinal data, we additionally check historical forward DNS (FDNS) datasets provided by Rapid7 [27]. The coverage of the FDNS data in our CNAME data is not perfect. It missed 10% of CNAMEs in 2018 and 30% in 2016 and 2017. We intend to use DNSDB [28] in future research to improve this coverage.

### C. CNAME cloaking-based tracking detection and analysis

*1) CNAME cloaking-based tracking detection:* To detect CNAME cloaking-based tracking, we use an approach based on wildcards matching of tracking filter list.

First, we discard CNAME-related subdomains that are categorized as first-party type. We classify a CNAME chain as *first-party* if the domain of the final node in this chain is the same as the domain of the considered HTTP request, or if the IP addresses of both the final node and the second-level domain are the same.

We then intend to detect CNAME cloaking-based tracking inside remaining subdomains. We apply wildcard matching based on well-known tracking filter lists: Easy privacy list [29] and AdGuard tracking protection filter [30]. Easy Privacy list consists of nine sublists and the Adguard tracking filter list consists of eleven sublists. They contain many rules that remove all forms of tracking, including web bugs, tracking scripts and information collectors, thereby protecting user personal data. Focusing on tracking, we select *third-party tracking domains and international third-party tracking domains* sublists from Easy privacy list and *tracking servers list* sublist from AdGuard tracking protection filter as of Feb. 5 2020. These blacklists are partly overlapping. We build the union of the two filter lists above to make the final tracking filter list. Then, we build regular expressions from tracking domains to match with CNAME behind all remaining subdomain. For example: *eulerian.net∧third-party* is changed to *.eulerian.net.$*. This rule matches any CNAME ending with *.eulerian.net.*; We can thus detect any CNAME cloaking-based tracking from tracking provider like Eulerian [31]. Finally, we inspect individual CNAME nodes in all CNAME chains using this customized filter list. If any node in a CNAME chain is flagged by this list, we classify this CNAME chain as a *tracker*. In this case, CNAME cloaking is used for tracking.

To avoid false positives, we then group these CNAME chains by domain and inspect them carefully. Finally, we remove five third-party domains, as false positives: *cedexis.net*, *episerver.net*, *meteor.com*, *pp8.com* and *tanx.com*.

By doing this, we can make sure the remain CNAME chains are used for CNAME cloaking-based tracking.

We furthermore use CDN lists [32], [33] to check if remaining CNAME chains are *CDN*. If it is not the case, we consider them as *Others* (see Figure 1).

*2) CNAME cloaking-based tracking analysis:* Having gathered the CNAME chains using CNAME cloaking-based tracking, the next step concentrates on their analysis. First of all, we consider the ranking of website containing CNAME cloaking-based tracking to assess how a real user would be affected in the real world by this type of tracking. Then, we consider the website category and country using CNAME cloaking-based tracking. We look up the IP address of these websites, then using Free IP Geolocation API [34] to determine the website location. We use FortiGuard Web Filtering [35] dataset from January 2020 for the website category classification. Note that Free IP Geolocation and FortiGuard Web Filtering datasets are not perfect, and we intend to address it in future work.

Finally, we consider tracking providers behind CNAME cloaking-based tracking. Tracker domains with the same second level domain are merged into a single tracker domain. For example, *cnn.122.2o7.net.* and *ikea.com.122.2o7.net.* are grouped into *2o7.net.* We then link domains to tracking providers using Disconnect’s blocklist [36].

## IV. CHARACTERIZING CNAME CLOAKING-BASED TRACKING

### A. CNAME chains structure

In this section, we focus on the characteristics of CNAME chains for first-party subdomain in Alexa Top 300K sites. Firstly, we present the CNAME usage of first party request by subdomain in Table III. The most common CNAME type is requests referring to resources of the first party (57.99%).

TABLE III
CNAME TYPES OF FIRST PARTY REQUEST BY SUBDOMAIN (ALEXA 300K SITES IN 2020).

| **Metric** | **1st-1st** | **Tracker** | **CDN, Cloud and others** |
|---|---|---|---|
| HTTP requests | 1,839,728/57.99% | **4,421**/0.14% | 1,328,155/41.87% |
| Subdomains | 48,365/39.47% | **1,899**/1.55% | 72,276/58.98% |

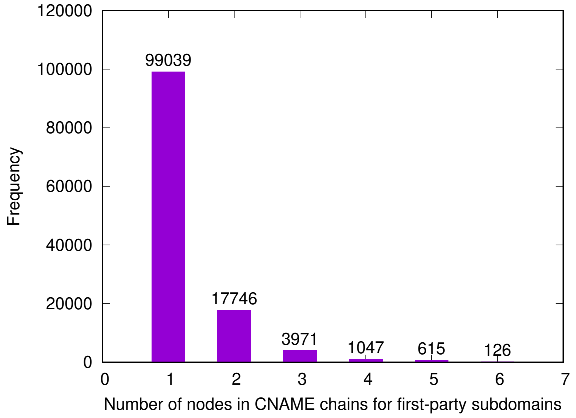

Fig. 2. The number of nodes in CNAME chains for first-party subdomains (Alexa Top 300K sites in 2020).

CDN and cloud also represent a large proportion of CNAME type (41.87%). Overall, we detect 4,421 CNAME cloaking-based tracking URLs. Furthermore, we find that these URLs belong to 1,762 websites (0.59%) on Alexa Top 300K sites.

Then, we breakdown the number of nodes in CNAME chains for first-party subdomains in our latest dataset (Alexa Top 300K sites in 2020) in Figure 2. We observe that about 80% of CNAME chains are very simple, just consisting of one CNAME. However, we also observe longer chains whose maximum length is six. These longest chains are mainly used by Microsoft likely for load balancing. This result suggests that checking only the first CNAME might be not enough for detecting CNAME cloaking-based tracking, because tracker websites may appear in intermediate nodes in the chain.

Finally, we show the breakdown of CNAME types regarding their position in CNAME chains in Figure 3. We note that the position represents the location of a CNAME in a CNAME chain. For example, the second node in the CNAME chain in Figure 1 is *m.ad.com* and its position is 1.

In Alexa Top 300K sites, the tracking-related domain inside a CNAME chain is mainly located at the first position. We however also observe tracking domains in the second position.

### B. Websites using CNAME cloaking-based tracking

Next, we focus on the characteristics of websites containing CNAME cloaking-based tracking.

Figure 4 presents the Empirical Cumulative Distribution Function (ECDF) of the Alexa ranking of websites containing CNAME cloaking-based tracking. Websites containing CNAME cloaking-based tracking are spread in the Alexa ranking. It illustrates that 30% of the CNAME cloaking-based tracking belongs to the top 20K website. Popular websites use more CNAME cloaking-based tracking.

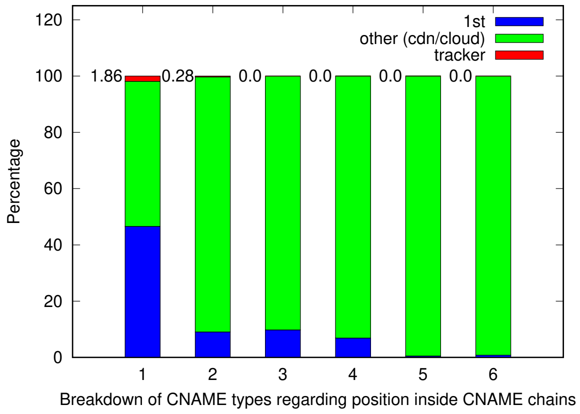

Fig. 3. Breakdown of CNAME types regarding position inside CNAME chains (Alexa Top 300K sites in 2020).

Then, we discuss the website category of websites containing CNAME cloaking-based tracking. For 1,762 websites containing CNAME cloaking-based tracking, the percentages of websites in Business, Shopping, Information Technology, Finance, Travel, Personal Vehicles categories are 15.8%, 15.1%, 14.9%, 10.2%, 8.9%, and 2.9%, respectively. In contrast, for the breakdown of these websites inside each website category in Figure 5, the percentages of these websites account for 0.4%, 1.5%, 0.5%, 2.6%, 2.4%, and 1.7%, respectively. In general, Finance, Personal Vehicles, and Travel use CNAME cloaking-based tracking more than other website categories.

Next, we analyze the website country of websites containing CNAME cloaking-based tracking. 66.0% of websites are located in the United States, 8.1% are located in Germany, and other countries have significantly lower percentages. On the other hand, for the breakdown of these websites inside each website country in Figure 6, the percentage of the United States, Germany, Ireland, Belgium, Czechia, Sweden, and the United Kingdom are 0.7%, 0.9%, 2.2%, 1.9%, 1.8%, 1.7%, and 1.5% respectively. In general, the number of websites using CNAME cloaking-based tracking in Alexa Top 300K sites from Ireland and Belgium is larger than other website countries.

### C. Tracking providers using CNAME cloaking-based tracking

We provide the breakdown of tracking providers behind CNAME cloaking-based tracking in Figure 7. We confirm 56 tracking providers using this techniques. The major player in Alexa Top 300K sites is Adobe (61%). Besides Adobe, we see some well-known tracking providers, such as Act-on [37], Eulerian [31], and Intent Media [38] (6.1%, 2.2%, and 2.1%, respectively).

Moreover, Table IV shows the breakdown of the tracking providers by category. We observe some specific pairs between the trackers and the category; Act-on for Business (34.8%), sp-prod.net for Media (32.8%), Extole and Eulerian for Shopping

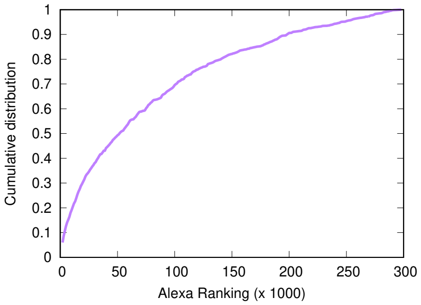

Fig. 4. ECDF of the Alexa ranking of websites containing CNAME cloaking-based tracking (Alexa Top 300K sites in 2020).

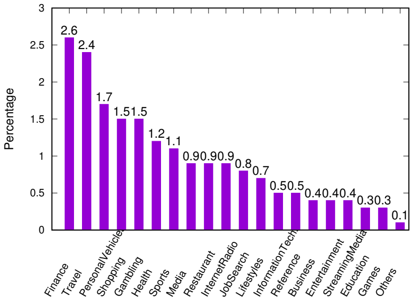

Fig. 5. Breakdown of websites containing CNAME cloaking-based tracking inside each website category (Alexa Top 300K sites in 2020).

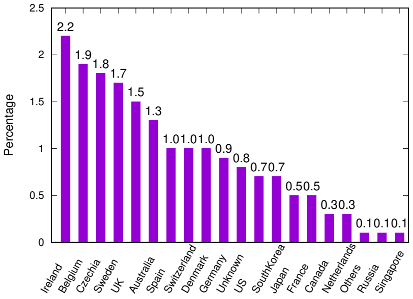

Fig. 6. Breakdown of websites containing CNAME cloaking-based tracking inside each website country (Alexa Top 300K sites in 2020).

TABLE IV

Breakdown of tracking providers inclusion in website by website category in Alexa Top 300K sites. The values have the following meaning: raw/percentage for category/percentage for tracking provider.

| **Category** | **Adobe** | **sp-prod.net** | **Act-On** | **Extole** | **Oracle** | **Eulerian** | **IntentMedia** | **wt-eu02** | **segment.com** | **Other** |
|---|---|---|---|---|---|---|---|---|---|---|
| Finance | 132/11.9/**71.7** | 0/0/0 | 7/6.3/3.8 | 3/4.8/1.6 | 6/10.5/3.3 | 3/7.3/1.6 | 1/2.6/0.5 | 3/7.7/1.6 | 1/3.3/0.5 | 28/14.4/15.2 |
| Travel | 96/8.6/**58.2** | 0/0/0 | 6/5.4/3.6 | 1/1.6/0.6 | 0/0/0 | 7/17.1/4.2 | 36/**92.3/21.8** | 6/15.4/3.6 | 0/0/0 | 13/6.7/7.9 |
| PersonalVehicles | 41/3.7/**78.8** | 5/3.6/9.6 | 0/0/0 | 0/0/0 | 2/3.5/3.8 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 4/2.1/7.7 |
| Shopping | 171/15.4/**61.3** | 0/0/0 | 1/0.9/0.4 | 32/**50.8**/11.5 | 4/7/1.4 | 10/**24.4**/3.6 | 0/0/0 | 8/**20.5**/2.9 | 4/13.3/1.4 | 49/**25.1**/17.6 |
| Gambling | 5/0.4/**29.4** | 0/0/0 | 0/0/0 | 0/0/0 | 3/5.3/17.6 | 2/4.9/11.8 | 0/0/0 | 1/2.6/5.9 | 0/0/0 | 6/3.1/35.3 |
| Health | 49/4.4/**75.4** | 1/0.7/1.5 | 3/2.7/4.6 | 6/9.5/9.2 | 2/3.5/3.1 | 0/0/0 | 0/0/0 | 1/2.6/1.5 | 1/3.3/1.5 | 2/1/3.1 |
| Sports | 33/3/**78.6** | 1/0.7/2.4 | 1/0.9/2.4 | 1/1.6/2.4 | 1/1.8/2.4 | 1/2.4/2.4 | 0/0/0 | 0/0/0 | 2/6.7/4.8 | 2/1/4.8 |
| Media | 97/8.7/**58.1** | 45/**32.8**/26.9 | 2/1.8/1.2 | 0/0/0 | 0/0/0 | 3/7.3/1.8 | 1/2.6/0.6 | 3/7.7/1.8 | 1/3.3/0.6 | 15/7.7/9 |
| Restaurant | 19/1.7/**82.6** | 1/0.7/4.3 | 0/0/0 | 1/1.6/4.3 | 0/0/0 | 1/2.4/4.3 | 0/0/0 | 0/0/0 | 0/0/0 | 1/0.5/4.3 |
| InternetRadio | 3/0.3/30.0 | 4/2.9/**40.0** | 1/0.9/10 | 0/0/0 | 0/0/0 | 1/2.4/10 | 0/0/0 | 0/0/0 | 0/0/0 | 1/0.5/10 |
| JobSearch | 11/1/**78.6** | 0/0/0 | 2/1.8/14.3 | 1/1.6/7.1 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 |
| Lifestyles | 13/1.2/**48.1** | 8/5.8/**29.6** | 1/0.9/3.7 | 3/4.8/11.1 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 1/3.3/3.7 | 1/0.5/3.7 |
| IT | 129/11.6/**48.0** | 40/**29.2**/14.9 | 26/**23.2**/9.7 | 3/4.8/1.1 | 18/**31.6**/6.7 | 7/17.1/2.6 | 1/2.6/0.4 | 4/10.3/1.5 | 9/**30.0**/3.3 | 32/16.4/11.9 |
| Reference | 8/0.7/**50.0** | 2/1.5/12.5 | 4/3.6/**25.0** | 1/1.6/6.3 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 1/0.5/6.3 |
| Business | 176/15.8/**60.7** | 7/5.1/2.4 | 39/**34.8**/13.4 | 3/4.8/1 | 14/**24.6**/4.8 | 5/12.2/1.7 | 0/0/0 | 10/25.6/3.4 | 7/23.3/2.4 | 29/14.9/10 |
| Entertainment | 34/3.1/**77.3** | 7/5.1/15.9 | 2/1.8/4.5 | 0/0/0 | 0/0/0 | 1/2.4/2.3 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 |
| StreamingMedia | 10/0.9/**83.3** | 0/0/0 | 1/0.9/8.3 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 1/0.5/8.3 |
| Education | 20/1.8/**50.0** | 1/0.7/2.5 | 8/7.1/20 | 3/4.8/7.5 | 4/7.0/10.0 | 0/0/0 | 0/0/0 | 0/0/0 | 4/13.3/10.0 | 0/0/0 |
| Games | 4/0.4/26.7 | 10/7.3/**66.7** | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 1/0.5/6.7 |
| Others | 62/5.6/**65.3** | 5/3.6/5.3 | 8/7.1/8.4 | 5/7.9/5.3 | 3/5.3/3.2 | 0/0/0 | 0/0/0 | 3/7.7/3.2 | 0/0/0 | 9/4.6/9.5 |

TABLE V

Breakdown of tracking providers inclusion in website by website country in Alexa Top 300K sites. The values have the following meaning: raw/percentage by country/percentage by tracking provider.

| **Category** | **Adobe** | **sp-prod.net** | **Act-On** | **Extole** | **Oracle** | **Eulerian** | **IntentMedia** | **wt-eu02** | **segment.com** | **Other** |
|---|---|---|---|---|---|---|---|---|---|---|
| Ireland | 20/1.8/**35.1** | 23/16.8/**40.4** | 0/0/0 | 1/1.6/1.8 | 0/0/0 | 6/14.6/10.5 | 0/0/0 | 1/2.6/1.8 | 1/3.3/1.8 | 5/2.6/8.8 |
| Belgium | 4/0.4/**80.0** | 0/0/0 | 0/0/0 | 0/0/0 | 1/1.8/**20.0** | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 |
| Czechia | 5/0.4/**27.8** | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 13/6.7/72.2 |
| Sweden | 7/0.6/**58.3** | 0/0/0 | 0/0/0 | 0/0/0 | 5/8.8/**41.7** | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 |
| United Kingdom | 37/3.3/**45.7** | 31/**22.6/38.3** | 1/0.9/1.2 | 0/0/0 | 1/1.8/1.2 | 0/0/0 | 0/0/0 | 1/2.6/1.2 | 0/0/0 | 10/5.1/12.3 |
| Australia | 20/1.8/**76.9** | 0/0/0 | 1/0.9/3.8 | 0/0/0 | 1/1.8/3.8 | 0/0/0 | 0/0/0 | 0/0/0 | 2/6.7/7.7 | 2/1.0/7.7 |
| Spain | 9/0.8/**60.0** | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 4/9.8/**26.7** | 0/0/0 | 0/0/0 | 0/0/0 | 2/1.0/13.3 |
| Switzerland | 6/0.5/**75.0** | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 2/1.0/**25.0** |
| Denmark | 6/0.5/**100.0** | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 |
| Germany | 48/4.3/**30.0** | 12/8.8/7.5 | 2/1.8/1.3 | 0/0/0 | 1/1.8/0.6 | 0/0/0 | 15/**38.5**/9.4 | 32/**82.1**/20 | 0/0/0 | 50/**25.6/31.3** |
| Unknown | 11/1.0/**84.6** | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 1/2.4/7.7 | 0/0/0 | 0/0/0 | 0/0/0 | 1/0.5/7.7 |
| United States | 805/**72.3/67.2** | 66/**48.2**/5.5 | 101/**90.2**/8.4 | 56/**88.9**/4.7 | 44/**77.2**/3.7 | 12/**29.3**/1.0 | 13/**33.3**/1.1 | 3/7.7/0.3 | 26/**86.7**/2.2 | 72/**36.9**/6 |
| South Korea | 2/0.2/**16.7** | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 10/5.1/**83.3** |
| Japan | 36/3.2/**92.3** | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 3/1.5/7.7 |
| France | 5/0.4/13.2 | 3/2.2/7.9 | 0/0/0 | 0/0/0 | 2/3.5/5.3 | 16/**39.0/42.1** | 5/12.8/13.2 | 0/0/0 | 0/0/0 | 7/3.6/18.4 |
| Canada | 25/2.2/**67.6** | 0/0/0 | 3/2.7/8.1 | 4/6.3/10.8 | 1/1.8/2.7 | 2/4.9/5.4 | 0/0/0 | 0/0/0 | 0/0/0 | 2/1.0/5.4 |
| Netherlands | 13/1.2/**68.4** | 0/0/0 | 2/1.8/10.5 | 0/0/0 | 1/1.8/5.3 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 3/1.5/15.8 |
| Russia | 6/0.5/**46.2** | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 5/12.8/**38.5** | 0/0/0 | 0/0/0 | 2/1.0/15.4 |
| Singapore | 6/0.5/**66.7** | 0/0/0 | 1/0.9/11.1 | 1/1.6/11.1 | 0/0/0 | 0/0/0 | 0/0/0 | 0/0/0 | 1/3.3/11.1 | 0/0/0 |
| Others | 42/3.8/**70.0** | 2/1.5/3.3 | 1/0.9/1.7 | 1/1.6/1.7 | 0/0/0 | 0/0/0 | 1/2.6/1.7 | 2/5.1/3.3 | 0/0/0 | 11/5.6/18.3 |

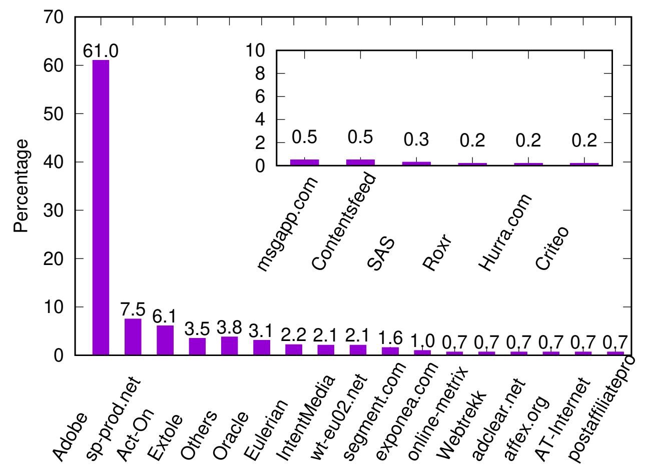

Fig. 7. Tracking provider providing CNAME cloaking-based tracking (Alexa Top 300K sites in 2020).

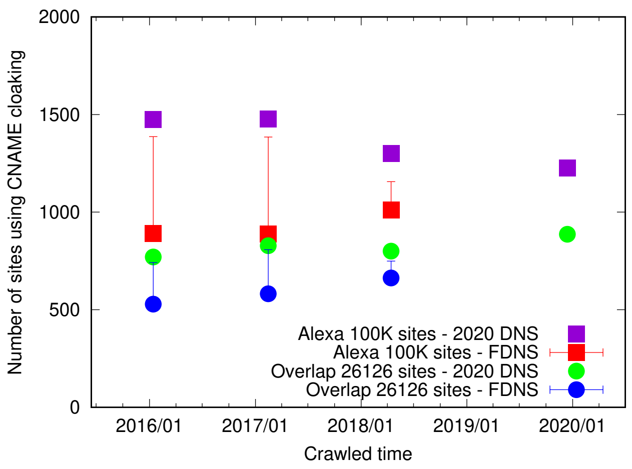

Fig. 8. The number of websites using CNAME cloaking-based tracking along time.

(50.8% and 24.4%), and Intent Media for Travel (92.3%), Oracle for Information Technology (31.6%). Other tracking providers were distributed in different types of websites. Except Internet Radio and Game categories (40% and 66.7% for sp-prod.net), Adobe is the most popular tracking provider in almost all categories. Furthermore, Table V shows the breakdown of the tracking providers by country. Tracking providers cooperating with websites such as Adobe, sp-prod.net, Act-on, Extole, Oracle, and segment.com are mainly located in the United States (72.3%, 48.2%,90.2%, 88.9%, 77.2%, and 86.7%, respectively). We also observe that some tracking providers are mainly located in specific countries, e.g., Eulerian in France (39.0%), Intent Media and Wt-eu02 in Germany (38.5% and 82.1%). Again, Adobe is the most popular tracking provider in almost all countries, except France (Eulerian with 42.1%) and Ireland (sp-prod.net with 40.4%).

Finally, we further investigate the number of tracking providers in each website. Most websites (1,695) deploy only one tracking provider, as expected. However, we also find 66 websites using two providers, and 1 website using three providers. Typical pairs of the providers are (Adobe, sp-prod.net), (Adobe, online-metrix.net), and (adclear.net, wt-eu02.net). We do not identify any plausible reasons of deploying multiple providers, but they might be used for different purposes (e.g., analytics and advertisement).

We conclude that, besides the biggest player Adobe, CNAME cloaking tracking providers operate on many website categories and countries.

### D. Longitudinal analysis of CNAME cloaking-based tracking

In this section, we analyze the longitudinal evolution of the number of websites using CNAME cloaking-based tracking. Figure 8 indicates the number of websites using CNAME cloaking-based tracking in Alexa 100K sites. We combine two crawled datasets and two DNS lookup datasets: (1) for the crawled data, the number of websites in each Alexa 100K and those in the overlap among all Alexa 100K datasets (26,126 sites); (2) for two DNS lookup datasets, DNS lookup in 2020 and lookup with the FDNS data (collected in February 2017, the oldest available snapshot, and June 2018). We then plot four combinations: the number of websites in each Alexa 100K sites with 2020 DNS (purple rectangles) and with FDNS (red rectangles). Those in the overlap among all Alexa 100K datasets with DNS in 2020 (green circles) and with FDNS (blue circles). The error bars in the figure show the number of unsolved CNAMEs due to the coverage of the FDNS data.

We show the growth of websites introducing CNAME cloaking-based tracking over the years. At a glance, the number of websites containing CNAME cloaking-based tracking is slightly decreasing in Alexa Top 100K websites with the latest DNS (purple rectangles). However, this decrease is due to biases of DNS lookup. Considering the historical DNS data (red rectangles), we confirm a slight increasing trend even though the error bars are large in 2016 and 2017. We see the similar trends in the overlap data (green and blue circles). Thus, not only websites newly ranked-in Alexa 100K sites, but also the commonly appeared websites in Alexa 100K sites introduced more CNAME cloaking-based tracking in four years.

Furthermore, 550 websites in 2016 still use this technique to track users in 2020. This is worrying for the end-users. Because up to our knowledge, there is no widely-used privacy protection technique that counters CNAME cloaking-based tracking (except the developer feature recently added to uBlock Origin [13], see also § V).

## V. Protection techniques effectiveness against CNAME Cloaking?

We analyze and compare browsers and extensions regarding privacy protection against CNAME cloaking-based tracking.

### A. Filter list

In order to block CNAME cloaking-based tracking, EasyPrivacy [29] and AdGuard tracking protection [30] require the

TABLE VI

DETECTION PERFORMANCE: EASY PRIVACY LIST AND ADGUARD TRACKING PROTECTION FILTER (ALEXA TOP 300K SITES IN 2020).

| Metric | AdGuard Tracking | Easy Privacy | All (combined) |
|---|---|---|---|
| HTTP requests | 1,733/39.20% | 3,104/70.21% | 3,114/70.44% |
| Subdomains | 1,068/56.24% | 1,475/77.67% | 1,483/78.09% |
| Sites | 1,051/59.65% | 1,426/80.93% | 1,431/81.21% |

identification of first party subdomains which are fronts for CNAME cloaking. For example, EasyPrivacy has a rule to block tracking provider Eulerian: *f7ds.liberation.fr*^. So, when website *liberation.fr* makes a request to the third-party tracker Eulerian through *f7ds.liberation.fr*, the request is blocked.

We assess the efficiency of the filter lists as counter-measures. We use adblockparser [39] to directly match blocking list rules with all HTTP requests in the Alexa 1,762 sites that contain CNAME cloaking-based tracking in Table III. We inspect individual CNAME cloaking-based tracking URLs using these well-known filter lists in January 2020. The results of this experiment is shown in Table VI. We find that 3,114 CNAME cloaking-based tracking URLs have been flagged by these filters. This represents 70.44% of all CNAME cloaking-based tracking URLs in Alexa Top 300K sites. Beside that, the Easy Privacy list detects almost as much CNAME cloaking-based tracking as combined lists. This means that CNAME cloaking domains detected by Adguard tracking filter list are almost always detected by Easy Privacy. Overall, tracker blocking lists thus do not effectively deal with CNAME cloaking-based tracking. Subdomains being used for CNAME cloaking may change often, which makes day-to-day filter lists updating tedious and time-consuming, and thus explain filter list poor performances.

### B. Browsers and extensions

Some browsers focus on security and privacy by blocking trackers. Browser extensions also use several techniques (such as blacklisting, or traffic monitoring) to block third-party tracking. We evaluate the ability of common browsers and extensions to block CNAME cloaking-based tracking.

We investigate five major browsers and six popular privacy protecting extensions which supports these browsers. We choose following popular browsers [40]: Chrome 80.0 [41], Opera 66.0 [42], Brave 1.4.92 [17], Firefox 73.0 [16] and Tor Browser 9.0.2 [18]. Regarding extensions, we use two criteria: blocking trackers and supporting multiple browsers. The privacy extensions that meet our criteria are Adblock 4.5.0 [3], Adblock Plus 3.7 [4], Privacy Badger 2020.1.13 [43], Disconnect 5.19.3 [5], Ghostery 8.4.6 [15], uBlock Origin 1.24.4 [6] and 1.24.5rc1 (developer's version) [13]. Ublock Origin 1.24.5rc1 has an anti CNAME cloaking-based tracking feature [13]. We include this version to provide an up-to-date picture of CNAME cloaking-based tracking counter-measures. We then collect all the HTTP requests and responses on the 1,762 websites containing CNAME cloaking-based tracking in Table III. We use Atrica[1] [44], a multi-browser crawling library, to gather data on websites with CNAME cloaking-based tracking. To conduct a general comparison of browsers and privacy protection techniques, we crawl 1,762 websites using 40 different profile configurations (five browsers × eight extensions including the vanilla/bare setting). All the measurements were performed in March 2020 with three IP addresses in Japan. One crawling took approximately 4 to 6 hours on commodity hardware.

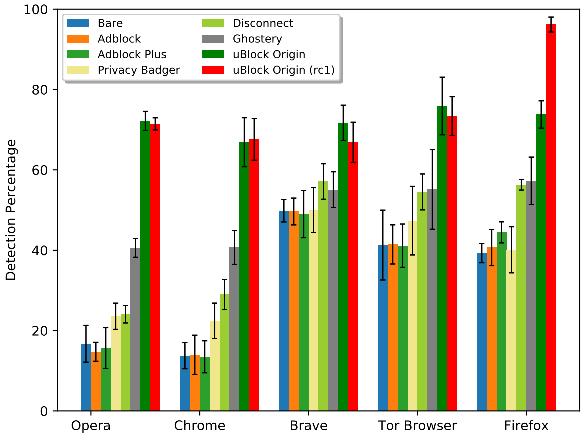

Fig. 9. Detection performance of browsers and extensions regarding websites containing CNAME cloaking-based tracking. The mean and standard deviation are computed on three crawls.

To reduce measurement error, we conducted three crawls and computed the relative standard error of the mean percentage of websites using CNAME cloaking-based tracking. We notice that there are also several possible sources of noise in our data. Some of these are internal and known, such as failure to connect to a website on a special time, or may also be external factors, such as network unreliability. To make a fair comparison, we set the website crawl timeout to 60 seconds. After this duration, if any website does not finish loading, we remove it.

Finally, we apply the same method (§ III-B and § III-C) to detect CNAME cloaking-based tracking among these profiles. Figure 9 shows the detection percentage of the CNAME cloaking-based tracking among browsers and their extensions. Overall, all browsers and extensions have a different impact on CNAME cloaking-based tracking. The most aggressive browser is Brave. It has the best performance among five browsers without any extension and blocks around 50% of websites that use CNAME cloaking-based tracking. We speculate that Shields feature (§ II-A2) is effective at detecting CNAME cloaking-based tracking. We also manually confirm that Shields blocks some CNAME cloaking-related subdomains, such as smetrics.10daily.com.au (Adobe), f7ds.liberation.fr (Eulerian), and 5ijo.01net.com (Eulerian).

Another group of browsers provides average performances (Firefox and Tor Browser). The remaining browsers (Opera

[1] Atrica currently supports chromium-based and Firefox-based browsers.

TABLE VII
TRACKING PROVIDERS USING CNAME CLOAKING-BASED TRACKING MISSED BY FIREFOX WITH UBLOCK ORIGIN 1.24.5RC1.

| Tracking provider | Firefox and uBlock Origin rc1 |
|---|---|
| intentmedia.net(Intent Media) | 12 |
| omtrdc.net (Adobe) | 12 |
| actonsoftware.com (Act-On) | 11 |
| exactag.com | 5 |
| sas.com (SAS) | 3 |
| at-o.net(AT Internet) | 3 |
| others | 12 |

and Chrome) provide low protection abilities.

For all browsers, the most effective extension is uBlock Origin that reduces around 70% of the websites containing CNAME cloaking. Adblock and Adblock Plus provide low protection abilities for all browsers. This result is not surprising because these extensions target ad-blocking. Another notable point is that uBlock Origin version 1.24.5rc1 with anti-CNAME cloaking-based tracking technique is better than uBlock Origin version 1.24.4. It however only impacts to Firefox browser because other browsers do not provide an API that allows an extension to perform DNS lookups [45].

Table VII shows the list of 18 tracking providers missed by Firefox with uBlock Origin developer build 1.24.5rc1, though they were initially detected by the Easy privacy list or the AdGuard tracking protection filter. This result shows that the unreleased feature from uBlock Origin is not able to completely block CNAME cloaking-based tracking and there is thus room for protection improvement.

## VI. RELATED WORK

### A. Third-party tracking detection and privacy protection technique comparison

The privacy hazards of online web tracking have been studied extensively with various approaches proposed to detect third-party tracking. Schelter and Kunegis [46] extracted third-party embeddings from more than 3.5 billion web pages to study global online tracking and found that 9 out of the 20 predominant third-party domains belong to trackers. Roesner et al. [47] developed a client-side method for detecting and classifying five kinds of third-party trackers over 500 unique trackers on the 500 most popular and 500 less popular websites according to the Alexa rankings. Wu et al. [10] developed DMTrackerDetector which detects third-party trackers automatically using structural hole theory and supervised machine learning. Harass et al. [11] presented an unsupervised method that leverages application-level traffic logs to automatically detect services running some tracking activities, thus enabling the generation of curated blacklists.

Besides, several works analyze privacy protection techniques. Ruiz-Martinez et al. [48] presented a survey of the theoretical comparison of the solutions and the main tools for privacy concern when users and surfing on the Internet. Mayer et al. [2] surveyed the current policy debate surrounding third-party web tracking and explains the relevant technology and uses a fourth-party web measurement platform to collect HTTP requests, responses and cookies. In previous studies, Mazel et al. [23] proposed a reliable methodology for privacy protection techniques comparison and compared a wide range of privacy protection techniques.

Our work not only focuses on privacy extensions but also extensively compares a wide set of browsers and privacy protection techniques against CNAME cloaking-based tracking.

### B. DNS analysis and CNAME cloaking-based tracking

DNS analysis is useful in detecting malicious activities on the web. Khalil et al. [49] proposed a method to discover malicious domains by analyzing passive DNS data by taking advantage of the dynamic nature of malicious domains to identify strong connections among malicious domains from a set of existing known malicious ones. Peng et al. [50] suggested a method to detect the malicious domains via domains that are not resolved to IP addresses directly but only appear in DNS CNAME records.

Krishnamurthy et al. [51] looked up CNAMEs to analyze third-parties across first-party sites. They did not focus on CNAME cloaking-based tracking, but they analyzed the third-party tracking across a large set of popular websites. In the uBlock Origin's GitHub issues page, a user presented a website loading first-party request, that pointed to a tracking provider [45]. This issue was then addressed in several discussions [52], [53], [54].

## VII. CONCLUSION

In this paper, we characterized the CNAME cloaking-based tracking on the web. We conducted experiments to assess the occurrence and evolution of CNAME cloaking-based tracking. The results show that 1,762 websites in the Alexa Top 300K sites in January 2020 contain CNAME cloaking-based tracking, primarily on website in the United States, within the Business and Shopping websites [2].

We also characterized a longitudinal analysis of CNAME cloaking-based tracking from 2016 to 2020. We found a significant evidence that the top websites have injected more CNAME cloaking-based tracking in the last four years. We believe that our method can help privacy experts understand third party tracking based on CNAME cloaking, and thus, improve existing counter-measures.

As future work, we intend to use Wayback Machine's archive (as in [55]) to add longitudinal measurements to improve our data collection process. We also want to improve existing counter-measures to deal with third-party tracking, especially CNAME cloaking-based tracking to protect end-users from tracking.

[2] The lists of websites and subdomains containing CNAME cloaking-based tracking in Alexa Top 300K sites (Jan 2020) are available at https://github.com/fukuda-lab/cname_cloaking

# References

[1] (2016) 2016 truste/ncsa consumer privacy infographic - gb edition. [Online]. Available: https://www.trustarc.com/resources/privacy-research/ncsa-consumer-privacy-index-gb/

[2] J. R. Mayer and J. C. Mitchell, "Third-party web tracking: Policy and technology," in *IEEE S&P*, 2012, pp. 413–427.

[3] Adblock. [Online]. Available: https://getadblock.com/

[4] Adblock plus. [Online]. Available: https://adblockplus.org/

[5] Disconnect. [Online]. Available: https://disconnect.me/

[6] R. Hill. ublock origin - an efficient blocker for chromium and firefox. fast and lean. [Online]. Available: https://github.com/gorhill/uBlock

[7] V. Dudykevych and V. Nechypor, "Detecting third-party user trackers with cookie files," in *2016 Third International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T)*. IEEE, 2016, pp. 78–80.

[8] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz, "The web never forgets: Persistent tracking mechanisms in the wild," in *Proceedings of ACM SAC*, 2014, pp. 674–689.

[9] S. Englehardt and A. Narayanan, "Online tracking: A 1-million-site measurement and analysis," in *Proceedings of ACM CCS*, 2016, pp. 1388–1401.

[10] Q. Wu, Q. Liu, Y. Zhang, and G. Wen, "Trackerdetector: A system to detect third-party trackers through machine learning," *Computer Networks*, vol. 91, pp. 164–173, 2015.

[11] H. Metwalley, S. Traverso, and M. Mellia, "Unsupervised detection of web trackers," in *Proceedings of IEEE GLOBECOM*, 2015, pp. 1–6.

[12] N. Kushmerick, "Learning to remove internet advertisements," in *Proceedings of AGENTS'99*, 1999, pp. 175–181.

[13] R. Hill. ublock origin - developer build 1.24.5rc0. [Online]. Available: https://github.com/gorhill/uBlock/releases/tag/1.24.5rc0

[14] Adobe: Creative, marketing and document management solutions. [Online]. Available: https://www.adobe.com/

[15] Ghostery makes the web cleaner, faster and safer! [Online]. Available: https://www.ghostery.com/

[16] Firefox browser. [Online]. Available: https://www.mozilla.org/en-US/exp/firefox/

[17] Brave browser. [Online]. Available: https://brave.com/

[18] Tor browser. [Online]. Available: https://www.torproject.org/

[19] (2019) Today's firefox blocks third-party tracking cookies and cryptomining by default. [Online]. Available: https://blog.mozilla.org/blog/2019/09/03/todays-firefox-blocks-third-party-tracking-cookies-and-cryptomining-by-default/

[20] (2008) The design and implementation of the tor browser [draft]. [Online]. Available: https://2019.www.torproject.org/projects/torbrowser/design/

[21] (2019) Data collection cnames and cross-domain tracking. [Online]. Available: https://docs.adobe.com/content/help/en/id-service/using/reference/analytics-reference/cname.html

[22] The top 500 sites on the web. [Online]. Available: https://www.alexa.com/topsites

[23] J. Mazel, R. Garnier, and K. Fukuda, "A comparison of web privacy protection techniques," *Computer Communications*, vol. 144, pp. 162–174, 2019.

[24] G. Merzdovnik, M. Huber, D. Buhov, N. Nikiforakis, S. Neuner, M. Schmiedecker, and E. Weippl, "Block me if you can: A large-scale study of tracker-blocking tools," in *IEEE Euro S&P*, 2017, pp. 319–333.

[25] Q. Scheitle, O. Hohlfeld, J. Gamba, J. Jelten, T. Zimmermann, S. D. Strowes, and N. Vallina-Rodriguez, "A long way to the top: significance, structure, and stability of internet top lists," in *Proceedings of ACM IMC*, 2018, pp. 478–493.

[26] Public suffix list. [Online]. Available: https://publicsuffix.org/list/

[27] Rapid7 open data, forward dns (fdns). [Online]. Available: https://opendata.rapid7.com/sonar.fdns_v2/

[28] Dnsdb database. [Online]. Available: https://www.dnsdb.info/

[29] Easyprivacy. [Online]. Available: https://easylist.to/easylist/easyprivacy.txt

[30] Adguard tracking protection filter. [Online]. Available: https://filters.adtidy.org/extension/chromium/filters/3.txt

[31] Marketing attribution and data management - eulerian. [Online]. Available: http://www.eulerian.com/

[32] S. Kayan. cdnfinder. [Online]. Available: https://github.com/turbobytes/cdnfinder/blob/master/assets/cnamechain.json

[33] W. Ma. china-cdn-domain-whitelist. [Online]. Available: https://github.com/mawenjian/china-cdn-domain-whitelist/blob/master/china-cdn-domain-whitelist.txt/

[34] Free ip geolocation api. [Online]. Available: https://freegeoip.app/

[35] Fortiguard web filtering. [Online]. Available: https://fortiguard.com/webfilter

[36] Tracking protection lists - trackers we block. [Online]. Available: https://github.com/mozilla-services/shavar-prod-lists/blob/master/disconnect-blacklist.json

[37] Act-on — exceptional marketing automation software. [Online]. Available: http://www.act-on.com/

[38] Intent media. [Online]. Available: http://www.intentmedia.com/

[39] Python parser for adblock plus filters. [Online]. Available: https://github.com/scrapinghub/adblockparser

[40] Browser statistics. [Online]. Available: https://www.w3schools.com/browsers/

[41] (2008) Chrome browser. [Online]. Available: https://www.google.com/chrome/

[42] Opera browser. [Online]. Available: https://www.opera.com/

[43] Privacy badger — electronic frontier foundation. [Online]. Available: https://www.eff.org/privacybadger

[44] V. Gerest. Atrica. [Online]. Available: https://github.com/fukuda-lab/atrica

[45] (2019) Address 1st-party tracker blocking #780. [Online]. Available: https://github.com/uBlockOrigin/uBlock-issues/issues/780#issuecomment-566845764

[46] S. Schelter and J. Kunegis, "Tracking the trackers: A large-scale analysis of embedded web trackers," in *Proceedigns of AAAI Conference on Web and Social Media*, 2016.

[47] F. Roesner, T. Kohno, and D. Wetherall, "Detecting and defending against third-party tracking on the web," in *Proceedings of USENIX NSDI*, 2012, pp. 12–12.

[48] A. Ruiz-Martínez, "A survey on solutions and main free tools for privacy enhancing web communications," *Journal of network and computer applications*, vol. 35, no. 5, pp. 1473–1492, 2012.

[49] I. Khalil, T. Yu, and B. Guan, "Discovering malicious domains through passive dns data graph analysis," in *Proceedings of ACM ASIACCS*, 2016, pp. 663–674.

[50] C. Peng, X. Yun, Y. Zhang, S. Li, and J. Xiao, "Discovering malicious domains through alias-canonical graph," in *2017 IEEE Trustcom/BigDataSE/ICESS*. IEEE, 2017, pp. 225–232.

[51] B. Krishnamurthy and C. Wills, "Privacy diffusion on the web: a longitudinal perspective," in *Proceedings of WWW*, 2009, pp. 541–550.

[52] O. Poitrey. (2019) Nextdns first to support blocking of all third-party trackers disguised as first-party. [Online]. Available: https://medium.com/nextdns/nextdns-added-cname-uncloaking-support-becomes-the-first-cross-platform-solution-to-the-problem-e3f437f84342

[53] R. Cointepas. (2019) Cname cloaking, the dangerous disguise of third-party trackers. [Online]. Available: https://medium.com/nextdns/cname-cloaking-the-dangerous-disguise-of-third-party-trackers-195205dc522a

[54] J. Leyden. (2019) Web trackers using cname cloaking to bypass browsers' ad blockers. [Online]. Available: https://portswigger.net/daily-swig/web-trackers-using-cname-cloaking-to-bypass-browsers-ad-blockers

[55] A. Lerner, A. K. Simpson, T. Kohno, and F. Roesner, "Excavating web trackers using web archaeology." *; login:*, vol. 41, no. 4, 2016.