Understanding the Privacy Implications of Adblock Plus's Acceptable Ads

Ahsan Zafar
azafar2@ncsu.edu
North Carolina State University
Raleigh, North Carolina, USA

Aafaq Sabir
asabir2@ncsu.edu
North Carolina State University
Raleigh, North Carolina, USA

Dilawer Ahmed
ahmed2@ncsu.edu
North Carolina State University
Raleigh, North Carolina, USA

Anupam Das
anupam.das@ncsu.edu
North Carolina State University
Raleigh, North Carolina, USA

ABSTRACT

Targeted advertisement is prevalent on the Web. Many privacy-enhancing tools have been developed to thwart targeted advertisement. Adblock Plus is one such popular tool, used by millions of users on a daily basis, to block unwanted ads and trackers. Adblock Plus uses EasyList and EasyPrivacy, the most prominent and widely used open-source filters, to block unwanted web contents. However, Adblock Plus, by default, also enables an exception list to unblock web requests that comply with specific guidelines defined by the Acceptable Ads Committee. Any publisher can enroll into the Acceptable Ads initiative to request the unblocking of web contents. Adblock Plus in return charges a licensing fee from large entities, who gain a significant amount of ad impressions per month due to participation in the Acceptable Ads initiative. However, the privacy implications of the default inclusion of the exception list has not been well studied, especially as it can unblock not only ads, but also trackers (e.g., unblocking contents otherwise blocked by EasyPrivacy).

In this paper, we take a data-driven approach, where we collect historical updates made to Adblock Plus's exception list and real-world web traffic by visiting the top 10k websites listed by Tranco. Using such data we analyze not only how the exception list has evolved over the years in terms of both contents unblocked and partners/entities enrolled into the Acceptable Ads initiative, but also the privacy implications of enabling the exception list by default. We found that Google not only unblocks the most number of unique domains, but is also unblocked by the most number of unique partners. From our traffic analysis, we see that of the 42,210 `Google` bound web requests, originally blocked by EasyPrivacy, around 80% of such requests are unblocked by the exception list. More worryingly, many of the requests enable 1-by-1 tracking pixel images. We, therefore, question exception rules that negate EasyPrivacy filtering rules by default and advocate for a better vetting process.

ASIA CCS '21, June 7–11,2021, Hong Kong, Hong Kong

ACM ISBN 978-1-4503-8287-8/21/06…$15.00
https://doi.org/10.1145/3433210.3437536

CCS CONCEPTS

• **Security and privacy → Privacy protections**; *Browser security*.

KEYWORDS

Online tracking; privacy-enhancing tools; adblock rules

ACM Reference Format:
Ahsan Zafar, Aafaq Sabir, Dilawer Ahmed, and Anupam Das. 2021. Understanding the Privacy Implications of Adblock Plus's Acceptable Ads. In *Proceedings of the 2021 ACM Asia Conference on Computer and Communications Security (ASIA CCS '21), June 7–11, 2021, Hong Kong, Hong Kong*. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3433210.3437536

1 INTRODUCTION

We spend a significant portion of our daily routine surfing the Web. Today, we interact with the Web in almost every facet of our lives: socializing, banking, health care, education, and entertainment. This makes the Web a gold mine for data brokers who quietly collect data about our lifestyles as we interact with different websites. For example, by simply looking at one's online activities, it is not difficult for a company to determine if she is pregnant or trying to lose weight; whether she is trying to switch jobs; what medications she is taking; where she is going on a trip and even where she is staying. In isolation, each of these facts may be innocent enough, but when such data is aggregated over a long period of time, it becomes a significant privacy invasion. Over time it becomes easy to determine a user's age, gender, race, social stature, political leanings, health condition, financial situation, taste in food, music, clothing, and so on. We have already witnessed several privacy breach incidents in the past [42, 54].

To counter such privacy-invasive tracking we have seen the rise of many privacy-enhancing web tools like ad- and tracker-blocking browser extensions, VPN services and anonymity networks. Among these tools ad- and tracker-blocking browser extensions have seen substantial adoption due to their ease of deployment and use. Adblock Plus [16] is one of the most widely used privacy-enhancing browser extensions, having more than 200 million active users [18]. Adblock Plus uses popular filter lists such as *EasyList* [29] and

EasyPrivacy [31] to block unwanted web resources. These filter lists both enhance our web experience [51, 71] and safeguard us against various security and privacy threats [52, 63, 67, 82]. However, Adblock Plus also maintains an exception list to unblock contents, and this exception list is endorsed by the Acceptable Ads initiative. Adblock Plus enables any ad providers to enroll into the Acceptable Ads initiative as long as their ads meet certain guidelines [19]. The purpose of the program is to allow ads that are generally unintrusive and do not interefere with the content. Interestingly, the Acceptable Ads feature is enabled *by default* and the privacy implications of such default inclusion has not been well studied.

In this paper, we take a data-driven approach to thoroughly analyze the evolution of the Acceptable Ads program and its impact on user's online privacy by answering the following research questions — **RQ1:** ***How has the exception list endorsed by the Acceptable Ads initiative evolved compared to EasyList and EasyPrivacy?*** Given that exception rules basically negate filtering rules, we study how the exception rules have evolved compared to filtering rules in terms of the various rule options used (e.g., third-party, image or domain-specific options). **RQ2:** ***What companies have been benefiting from Adblock Plus's Acceptable Ads initiative over time?*** We look at identifying top partners contributing to the exception list and also determine domains from which contents are prominently unblocked. **RQ3:** ***How are users impacted by Adblock Plus's default activation of Acceptable Ads?*** We collect real-world traffic and perform differential analysis to determine not only how many contents, but also what type of contents (e.g., tracker vs. ad) are unblocked in the presence of the exception list.

To answer these questions, we collect historical commits of EasyList, EasyPrivacy and the exception list endorsed by the Acceptable Ads program. We also crawl the Tranco [62] top 10k sites in Chrome and collect all web requests generated by each visit. Next, we perform a longitudinal analysis of how filtering rules have evolved over the years and how exception rules (ones that are used for the Acceptable Ads program) contrast to such evolution. We also go deep into analyzing the partners of the Acceptable Ads program and determine the dominant players, and the domains they unblock. Lastly, we analyze real-world web traffic to determine what filtering and exception rules are more commonly triggered and to what extent exception rules unblock contents from EasyList and EasyPrivacy. Our analysis provides much needed insight into the unwanted implications of default activation of Acceptable Ads.

In summary, we make the following contributions:

- We analyze how filtering and exception rules used by Adblock Plus have evolved over time and contrast their evolution in terms of the different rule options used. Such analysis helps us determine what type of contents are typically used to serve ads and what contents are being requested to be unblocked by the exception rules (§5).
- We also perform the first large-scale analysis of ad publishers partnering with Adblock Plus and highlight domains (i.e., domains from which contents are to be unblocked) they aggressively unblock. Our analysis identifies the top companies responsible for unblocking the largest amount of web contents (§6).
- Lastly, we utilize real-world web traffic to analyze the privacy implications of the exception list. We not only identify the most prevalent blocking and exception rules triggered, but also perform a differential analysis to determine the impact of the exception list on EasyList and EasyPrivacy. We show that the exception rules unblock many 1-by-1 pixel-based trackers — something that does not technically qualify as an ad in the first place, and thereby give users a false sense of protection against online trackers (§7).

The remainder of this paper is organized as follows. Section 2 provides a brief background on how Adblock Plus filtering and exception rules work. Section 3 describes related work. Section 4 describes our data collection process. We analyze the evolution of filters and exceptions in Section 5. Section 6 investigates dominant partners in the Acceptable Ads program. We analyze the impact of default inclusion of Acceptable Ads in Section 7. Section 8 discusses the implications of our findings and limitations. Finally, we conclude in Section 9.

2 BACKGROUND

Adblock Plus. In 2019, GlobalWebIndex reported that 47% of internet users globally use an ad blocker [66]. Adblock Plus [16] is one of the most widely used ad-blocking browser extensions with more than 200 million active users [18]. Adblock Plus blocks ads based on preexisting filters. By default Adblock Plus uses the EasyList,[1] but allows users to add custom filters. With these filters enabled, it can block specific web requests while loading a website; typically these web requests serve an ad. Furthermore, Adblock Plus also provides the option to enable the EasyPrivacy list to block online trackers.[2] However, Adblock Plus also enables, by default, an exception list in support of the 'Acceptable Ads' initiative, which unblocks non-intrusive ads. This exception list overturns existing blocking rules. Interestingly, any company can become a partner in the Acceptable Ads initiative and request specific web contents to be unblocked. Adblock Plus charges partners a licensing fee when they gain more than 10 million additional ad impressions per month due to participation in the Acceptable Ads initiative [15].

Blocking Filters. A blocking filter is essentially structured like a regular expression with additional options that adjust the scope of a filter to affect only specific contents or domains. There are mainly three types of filters, including [35]:

- Request filters: Applied on the network level to decide whether a request should be blocked (e.g., looking at domains).
- Content filters: Hide particular elements within a page (e.g., hiding elements by ID attribute or by name).
- Exception filters: Used to unblock certain requests or unhide certain elements on certain websites.

There are various options that can appear in blocking rules. These options modify the behavior of a filter and are separated with a comma (,) after a dollar sign ($) at the end of the filter. For example, in the following filter rule: `/ads/*$script,image`, the actual filter is `/ads/*`, and `script` and `image` are its options — signaling the

[1] Adblock Plus maintains their own copy of EasyList [21]
[2] Adblock Plus maintains their own copy of EasyPrivacy [22]

extension to block any script or image type content. Currently, the following options are supported by Adblock Plus [35]:

- script: scripts loaded via the HTML `script` tag
- image: images loaded via the HTML `img` tag
- stylesheet: external CSS stylesheet files
- object: content handled by browser plug-ins, e.g., Flash or Java
- xmlhttprequest: requests started using the XMLHttpRequest object
- document: the page itself, but only works for exception rules
- subdocument: embedded pages, usually included via HTML inline frames (iframes)
- elemhide: for exception rules only, similar to document but only turns off element hiding rules on the page rather than all filter rules
- popup: pages opened in a new tab or window
- font: external font files
- media: regular media files like music and video
- other: types of requests not covered in the list above

Furthermore, following restriction options can control how filters should be applied to outgoing web requests.

- domain: applied on pages loaded from the specific domains
- sitekey: applied on pages that provide a public key and a signature that can be verified by the very same public key contained in the filter
- third-party: applied to requests from a different origin than the currently viewed page

For example, in the following filter rule –

`||scdn.co/static/js/baba.js$script,domain=spotify.com`

`scdn.co` is the *blocking domain*, whereas `spotify.com` is the referring domain. We term the referring domain as the *surrogate* domain in the rest of the paper. Also, it is possible to apply inverse operation on certain types of options. For example, `~script` implies any content other than a script and `domain=~example.com` means that the filter should be applied to all pages from any domain, except `example.com`.

Exception Filters. Exception filters follow the same structure as blocking filters. They are regular expressions with one fundamental difference in that they start with an exception anchor *@@*. This anchor informs the extension that any web request that matches the filter should be unblocked. For example, the following exception rule `@@||maps.google.com/maps/$script,domain=171gifs.com` enables scripts from `maps.google.com` to be loaded, when such requests are launched from `171gifs.com`.

Exception rules, by default, override blocking rules. Thus, when a new web request is generated Adblock Plus first tries to match it against any existing blocking rule. If no match is found, the web request is allowed to go through (i.e., it is a *benign* request). On the other hand, if a matching filter is found, Adblock Plus will then check to see if it matches with any exception filter. If an exception filter is found, the web request will be *unblocked*, whereas the web request will be *blocked* when no matching exception rule is found. Figure 1 highlights this overall pipeline. We have confirmed the presence of exception filters in basic blocking lists, such as EasyList, in order to unblock important web resources that are essential to rendering a web page [7]. These exception filters are typically used to prevent site breakage and disruption caused by existing blocking filters.

Exception List for Acceptable Ads. Adblock Plus generates revenue mainly through the Acceptable Ads program [15]. This list contains exception rules for companies (domain owners) that have contacted Adblock Plus to exempt certain web resources from being blocked. We refer to these companies as *partners* (as termed in the exception list). These partners are interesting entities in the list as Adblock Plus charges a licence fee for companies that gain more than 10 million additional ad impressions through their participation in the Acceptable Ads program [15]. For example, Figure 2 represents information about a partnering entity in the Acceptable Ads initiative.

```
!:partner_token=Amazon Advertising
!:partner_id=ec725ef475df5236
!:type=partner
!:forum=https://adblockplus.org/forum/viewtopic.php?f=12&t=9791
! Amazon text ads
@@||adsensecustomsearchads.com^$elemhide,document,
subdocument,domain=d14qd3he45186l.cloudfront.net
```

Figure 2: Example of a partner in the exception list.

Here, we refer to `Amazon Advertising` as a *partner* to Adblock Plus's Acceptable Ads program, where `Amazon Advertising` is requesting Adblock Plus to unblock contents from `adsensecustomsearchads.com`. We refer to `adsensecustomsearchads.com` as an *unblocked* domain. The domains enlisted under the 'domain'

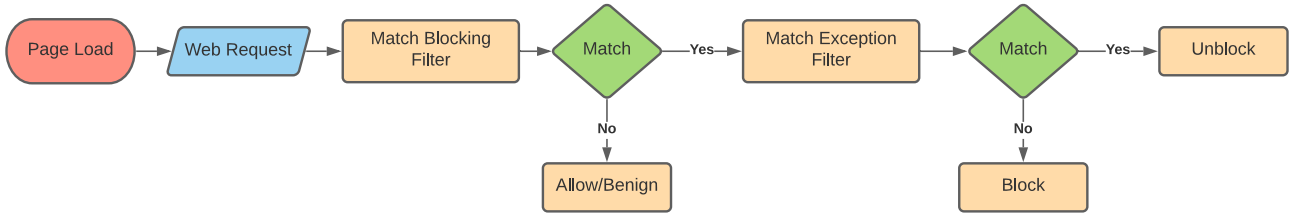


Figure 1: Adblock Plus filtering process. Each web request is either blocked, or unblocked or allowed to pass.

option (e.g., `d14qd3he45186l.cloudfront.net`) are the domains on which the exception rule is applied (i.e., they are *surrogate* domains). By domains we refer to the effective top-level-domains (eTLD+1). The Acceptable Ads initiative allows any vendors to register exception filters to unblock ads that follow certain criteria. These criteria provide specific instructions on the placement, size and clear distinction of ads. According to Adblock Plus, these standards are meant to ensure that the ads displayed are not intrusive in nature [19].

3 RELATED WORK

Online tracking. Online tracking refers to the process of companies collecting consumers' online activities across different websites. Typically, websites include various third-party resources, which then enable third-party companies to obtain information about our browsing activities. Websites include these resources primarily for *targeted advertisement* and analytics. Overtime online trackers have become smarter and stealthier. Online tracking can be classified into two broad groups: stateful and stateless tracking. Stateful tracking usually utilizes some form of local storage on the client's browser. For example, using cookies to store unique identifiers. However, as soon as browsers enabled users to clear cookies, trackers started using Flash Local Storage Object (LSO), HTML5 local and session storage, and ETags to store unique identifiers [60, 72, 75]. Alternatively, a tracker can completely avoid storing a local identifier, and instead recognize the host browser or device using a wide variety of software and hardware characteristics such as fonts [49], battery [70], canvas [68] or hardware-level features [43, 59]. Several studies have looked at the prevalence of such trackers in the wild [38, 39, 46, 47, 69]. Mobile platforms have also been shown to be vulnerable to fingerprinting [44, 45, 61]. Others have shown how hardware and software constraints on mobile platforms often lower the tracking precision for mobile browsers [46, 57, 76].

Countermeasures against online tracking. Over the years many privacy-enhancing web tools have emerged among which ad- and tracker-blocking browser extensions are the most widely used. These ad- and tracker-blocking extensions are different in terms of how they filter unwanted web contents. The main difference arises from how these tools derive the underlying filter rules. Typically, the rulesets can be categorized into three groups: community-driven, centralized, and algorithmic. Community-driven rulesets such as EasyList [29] and EasyPrivacy [31] are most popular, and are used by different browser extensions such as Adblock Plus [16], Adblock [17] and uBlock Origin [1]. Other blocking rulesets are more centrally managed and curated by specific third-party companies such as Ghostery [34] and Disconnect [25]. These rulesets are usually more compactly formed. The last category of blocking rules is derived algorithmically instead of relying on some regularly updated blacklists. These tools use heuristics to automatically detect trackers. For example, EFF's Privacy Badger extension [36] uses algorithmic methods to decide which third-party domain is a tracking domain by observing requests between first-party and third-party domains, and searching for the presence of the same high-entropy string (e.g., identifiers) across multiple first-party sites. Due to the difference in the underlying blocking techniques, the effectiveness of blocking unwanted web contents also varies across different tools. There have been many measurement studies that have looked at the effectiveness of privacy-enhancing browser extensions in blocking unwanted web contents [41, 48, 50–52, 55, 56, 64, 65, 67, 74, 77, 79, 80]. Moreover, privacy-geared browsers such as Brave [20], Cliqz [24] and Epic [32] are slowly becoming popular.

Evolution of filtering lists. In the last few years, we have seen researchers analyze crowd-sourced blocking lists such as EasyList. Alrizah et al. have looked at the errors and pitfalls within the crowd-sourcing process for maintaining EasyList [40]. Similarly, Vastel et al. have analyzed the trade-off between the growth and efficiency of EasyList [74]. Hashmi et al. have analyzed various open-sourced ad-blocking lists and have shown that most lists are updated by prioritizing ads and tracking domains found in popular websites [53]. On the contrary, Sjosten et al. propose new ways to enhance existing filter lists to cover parts of the web that have very small user-base [73]. Iqbal et al. have studied how anti-adblock filter lists have evolved over the years in their fight against ad blockers [58]. And lastly, Walls et al. were the first to look at how the acceptable ads program has changed since its inception in 2011 [78]. They show that the whitelist has been updated on average every 1.5 days, and grew from 9 filters in 2011 to over 5,900 filters in 2015.

Distinction from prior works. We also focus on the Acceptable Ads program, however, our work differs from the work done by Walls et al. [78] in three major directions. Firstly, we provide an in-depth analysis on how the distribution of various blocked contents has changed over time and how it correlates with the contents unblocked by the exception list. Secondly, we analyze how the partnership with various companies has evolved over time and who benefits from the exception rules. Lastly, to the best to our knowledge, we are the first to showcase how the exception list enabled by the Acceptable Ads program not only allows ads (something that is expected), but also trackers (such as 1-by-1 pixel-based trackers) commonly blocked by EasyPrivacy. Thus, automatically enrolling users into the Acceptable Ads program can potentially give users the false sense of protection against online trackers.

4 DATA COLLECTION AND PARSING

This section describes our data collection methodology to gather retrospective versions of blocking filters (EasyList and EasyPrivacy) and exception filters (Acceptable Ads). It also explains our experimental setup for generating traffic to model Adblock Plus's behavior in the presence of different filter lists. Figure 3 shows an overview of our data collection and analysis process.

4.1 Collecting Different Filter Versions

EasyList and EasyPrivacy are community-driven filter lists that are publicly available through a Github repository [30]. Similarly, the exception list supported by the Acceptable Ads initiative is maintained on a Mercurial repository [10]. We scrape the latest version for each day a commit was made to any of these repositories.

Table 1 highlights the number of different versions we were able to retrieve. Blocking filters were being maintained almost on a daily basis. This is why we were able to find around one version for each day of the year. Our collection of both blocking lists starts from January 1, 2011 and ends on May 20, 2020. In total, we gathered

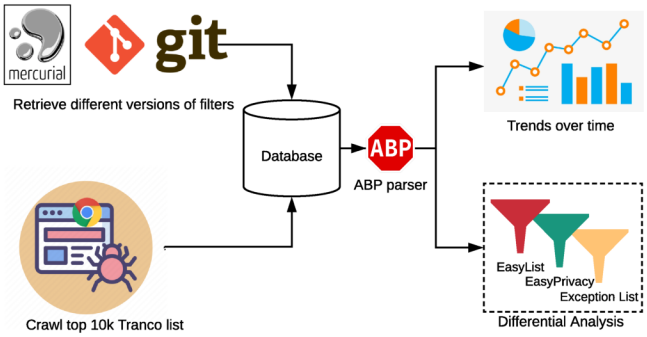


Figure 3: Data collection and analysis pipeline.

Table 1: Number of different versions of a list retrieved.

Year	EasyList	EasyPrivacy	Exception
2011	362	362	6
2012	366	366	38
2013	360	360	138
2014	358	358	165
2015	355	355	208
2016	358	358	252
2017	351	351	252
2018	356	356	243
2019	354	354	237
2020	141	141	155
Total	3371	3371	1694

3,371 versions for each list. The first version of the Exception list was logged on October 6, 2011. Our collection of the exception list spans from the first commit to the final commit made on August 11, 2020. Over this period, we were able to gather 1,694 versions. Unlike blocking filters, exception filters are not maintained on a daily basis. The early years of the exception filter was characterized by slow growth. In its first year, commits were made on only six different days. Over time, as companies started to join the Acceptable Ads initiative, the number of commits increased. In 2019, on average, almost 31.3 and 28.7 filters per day were added to and removed from the exception list, respectively.

Filters are basically regular expressions that describe some properties of a web request. These properties include the content types (script, image, font, etc.) or restriction types (specific surrogate domains on which to apply the rule). In order to extract this metadata from filters, we used an Adblock Plus filter parser [9], which is a Node.JS version of Brave's C++ Adblock Plus parser. This tool helped us parse each filter and provide information on its different properties such as the number of domains it targets, or the options it used, or whether it is an exception rule.

4.2 Collecting Web Traffic

We also collected real-world web traffic by visiting the top 10k sites listed by Tranco [62]. For this we ran a headless instance of Chrome and used Chrome's remote debugging interface to dump all web requests as Har files [23]. For each site visit, we spawn a fresh headless instance of Chrome and disabled caching. We waited for 15 seconds after the load event to allow any additional web contents to be loaded. The timeout period was set to 30 seconds and for any page load failure, we made two retry attempts with a delay of 5 seconds between consecutive attempts. We started our crawl on June 11, 2020 and it finished in around four days. We were able to parse Har files for 8,636 sites and for the remaining sites we obtained page not found, or server not found error messages.

Har files contain information about the web resources that were requested. In order to model the behavior of Adblock Plus extension and assess the implications of filters, we used the same aforementioned Adblock Plus filter parser [9] that matches a web request against filter lists that are fed to the parser. For each Har file, we first parsed its content to retrieve all the web resources requested. This information has to be complemented with the surrogate domain (i.e., the referring domain) and passed on to the parser. After instrumenting the parser, [3] we passed in a total of 1,333,925 web requests that were generated from the 8,636 page loads. The parser followed the filtration processed described in Figure 1 to match a web request with a filter. If the web request matched only a blocking filter, it got blocked. But if it matched an exception filter after matching with a blocking filter, it was unblocked. We output both the blocking status and triggered filter to derive a detailed analysis of real-world traffic in Section 7.

5 EVOLUTION OF FILTERS AND EXCEPTIONS

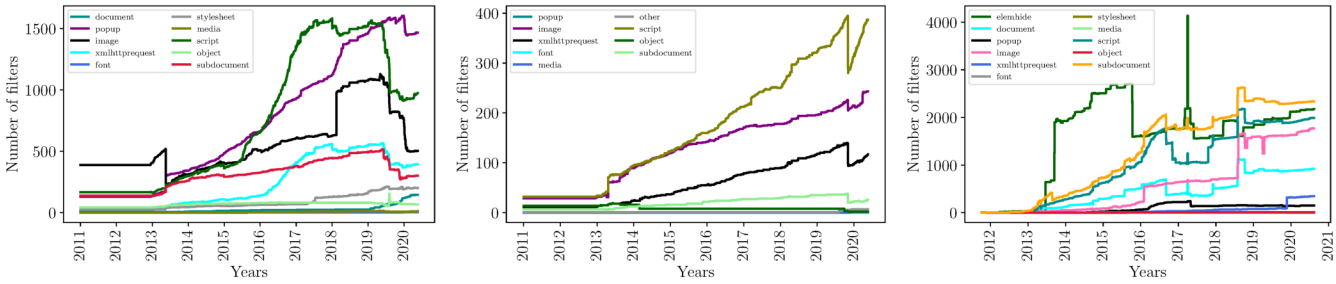
In order to understand the growth of Acceptable Ad's exception list, we must first understand the significance of blocking filters (i.e., EasyList and EasyPrivacy), since exception lists are triggered in response to the action of blocking filters. EasyList is the primary filter list that blocks most advertisements from web pages, including unwanted iframes, images and objects. On the other hand, EasyPrivacy is an optional supplementary filter list that filters tracking resources such as 1-by-1 tracking pixel images. In this section, we contrast the growth of the blocking lists and the exception list (used for the Acceptable Ads initiatives) in terms of both the raw number of rules used, and the use of various content and restriction options within the rules themselves. Such analysis will help us understand how the filters and exceptions have evolved over time.

5.1 Content Type

Content type rule options basically indicate the specific type of contents to block. For example, the following filtering rule: `||buzina.xyz^$script` will block any script originating from `buzina.xyz`. A full list of the different content types supported by Adblock Plus is described in Section 2. Over time filters have evolved to counter ways in which advertisers attempt to evade the blocking rules. Thus, understanding the evolution of the various rule options can help us gain insights into how advertisers have evolved over the years and how that correlates with contents that are being unblocked through the Acceptable Ads program.

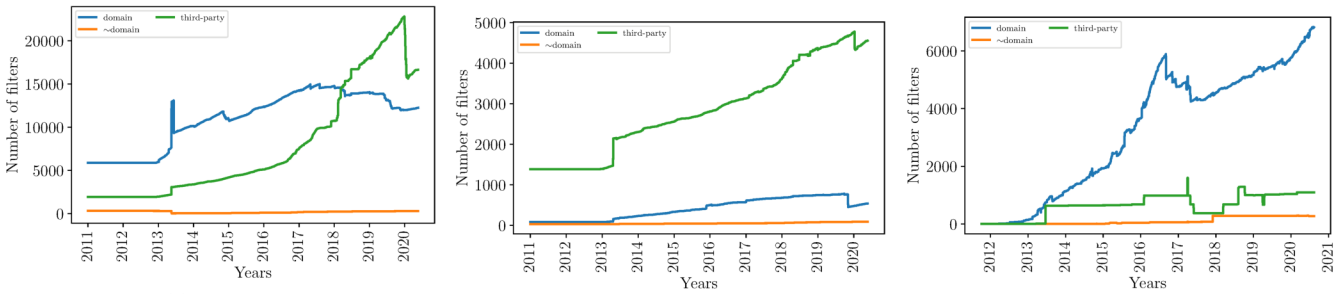
Figure 4 highlights how the distribution of various content types have evolved over the years. We can see from Figure 4a that 'popup', 'script' and 'image' are by far the most prevalent options appearing

[3]The parser we used relied on Bloom Filter matching to speed up the filter matching process. We removed this feature as it does not represent the filtration pipeline of ABP extension. Additionally, we changed the code base to allow the parser to output rules that were triggered.



(a) EasyList (b) EasyPrivacy (c) Acceptable Ads Exception list

Figure 4: Distribution of filter options based on content types for various filter and exception list. We see 'script' and 'image' options are popular across both blocking and exception rules.



(a) EasyList (b) EasyPrivacy (c) Acceptable Ads Exception list

Figure 5: Distribution of filter options based on restriction type. Blocking filters use 'third-party' restriction to block ads and trackers, whereas exception filters use 'domain' option the allow ads to be loaded from specific domains.

in EasyList, indicating ads are mostly served through such contents. When it comes to blocking trackers (i.e., using EasyPrivacy), we see that 'script' (e.g., fingerprinting scripts) and 'image' (e.g., 1x1 tracking pixel) options are most commonly used (as shown in Figure 4b). We also see many rules blocking XMLHttpRequests. Now, if we see the trend for the exception list endorsed by the Acceptable Ads initiatives, we see that 'subdocument' (e.g., iframes), 'elemhide' (ad banners), 'script' and 'image' are the popular options to unblock ads. While 'subdocument' and 'elemhide' are the top two content types unblocked by the exception list, we can see that 'script' and 'image' contents are also similarly unblocked.

Finding 1: *'Script' and 'image' options are popular across both blocking and exception rules — suggesting ads are typically served through images and scripts.* In general, we find the use of both 'script' and 'image' options to increase over the years across all lists. We do, however, see some sudden fluctuations in the lists as evident from the spikes.

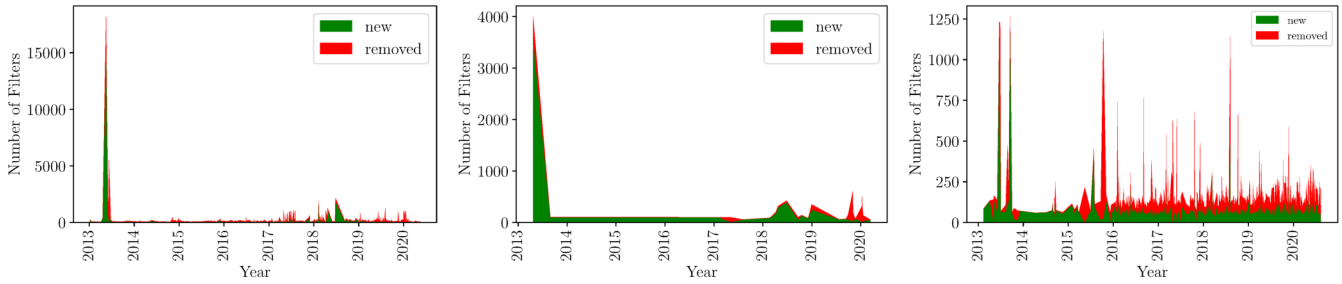
5.2 Restriction Types

Restriction types are used by filters to specify instances where they should be activated (or not). For example, the following filtering rule: `||adobedtm.com^$third-party,domain=adobe.com` restricts all third-party contents from `adobedtm.com` when the request is launched from `adobe.com`. The 'third-party' option specifies that the filter should only be applied to requests from a different origin than the currently viewed page.

Figure 5 highlights the distribution of restricting options over the years. We can see from Figure 5a that EasyList has incrementally used more of 'third-party' and 'domain' options to block advertisements, originating from third-party domains, to be loaded on specific sites. Figure 5b also shows an incremental usage of 'third-party' option to prohibit trackers originating from third-party domains while loading a web page. However, unlike EasyList, EasyPrivacy rules generally do not specify domains where rules should be activated. This allows anti-tracking rules to be applicable on all web pages.

Partners of acceptable ads only want to unblock content from specific surrogate domains. Figure 5c shows that exception filters extensively use 'domain' type restrictions to make sure that acceptable ads are displayed only on domains of their interest. However, unlike EasyList, not many exception filters specify 'third-party' restrictions. This allows acceptable ads to be unblocked from any origin.

Finding 2: *Blocking filters use 'third-party' restriction to block ads and trackers that do not follow same-origin policy. However, as the purpose of exception filters is to allow acceptable ads to be loaded from specific ad networks, we see more use of the 'domain' option.* This trend seems to be persistent, at least for the most recent three years.



(a) EasyList **(b) EasyPrivacy** **(c) Acceptable Ads Exception list**

Figure 6: Churn rate of different list. Exception rules are edited more frequently compared to blocking filters.

5.3 Churn Rate

In this section, we analyze the growth of blocking and exception filters based on the number of raw rules that have changed over time. It should be noted that exception filters are changed at the request of the partners who submit them in the first place [33]. The committee assesses whether the proposed changes violate any standard criteria for acceptable ads. In case it does not, the filter rules are added to the exception list. Since the number of committed versions for each day is very high, we only consider updates as significant if the total number of changes for that version is greater than 50 from the previous version. This allows us to identify discernible temporal changes made to the lists.

Figure 6a and Figure 6b show that both blocking filters (i.e., EasyList and EasyPrivacy) have grown through modest changes over the years. However, one key event in 2013 significantly increased the number of filters in both lists. On May 17, 2013, Fanboy's list merged with EasyList [2]. Similarly, Fanboy's tracking list merged with EasyPrivacy. We also see that in late 2019, both blocking filters underwent reduction in the number of rules (as evident by the red spikes). Inspection of Git log revealed that a lot of stale and problematic rules were being audited and removed accordingly [8].

In contrast, the exception list endorsed by the Acceptable Ads program witnessed a slow growth in early years. However, in 2013, two major spikes can be seen. The first surge occurred when Google joined the program [3]. For the second spike, we used mercurial logs and discovered that new partners and their corresponding filters were added to the list in one single push. In general, we see that after 2015, the rate at which the exception list updates is significantly higher compared to blocking filters. This can be attributed to the fact that after 2015, the management for Acceptable Ads program was handed over to an independent committee which eased the criteria for acceptable ads, thereby allowing more filters to be frequently updated [14].

Finding 3: *Exception rules are edited more frequently based on how partners want to benefit from them. Higher churn rate for such filters means on average an exception filter has shorter life span. In contrast, blocking filters are less frequently edited or removed, unless they cause websites to break or lose essential functionalities.* In the past five years, exception lists have been edited at an average rate of 29.2 filters per day. In contrast, the blocking filters have been edited at a rate of 9.16 filters per day.

6 ACCEPTABLE ADS PARTNERS

Acceptable Ads initiative allows many contributors to issue exception filters to unblock acceptable ads of their interest. Adblock Plus charges licensing fees to large entities who gain more than 10 million additional ad impressions per month due to participation. However, Adblock Plus claims that 90% of licences are issued to smaller platforms for free [15]. This section explores the major beneficiaries of the Acceptable Ads program over time.

6.1 Enrollment of Partners

We first study how the number of unique partners and unblocked domains have evolved over time. Retrieving partners was tricky since versions preceding November 2018 had no fixed template for specifying a partner in the list. In order to get the names, we used regular expressions.[4] Obtaining unique partners in recent versions was easier since the list had a fixed template that provided partner details (as depicted in Section 2), including name and token number. Similarly, for extracting *unblocked* and *surrogate* domains, we use string operations and regular expressions.

Figure 7 shows how the number of unique partners evolved over time. As mentioned in the previous section, we witness an increased growth of partners after an independent Acceptable Ads Committee took over the program in 2015 [14]. This accounts for a general upward trend. However, the most interesting change is the accelerated removal of partners between 2017 and 2019 as shown in Figure 7. Publishers such as `ThoughtCatalog` and `Uniregistry` left the Acceptable Ads program in 2017 [11, 12].

We also use mercurial logs and commit history to identify the partners and domains that were removed. We saw that on November 07, 2018, the exception list changed its template [13]. In the new template, a lot of stale partners were removed and their exception filters were re-issued under other existing partners. For example, `InfoSpace`, `Yidio` ads and `Nevada Enterprise` were all removed in the new template and their exception filters were registered under `System1`. In some cases, partnering companies had been bought out by a bigger organization. For instance, `InfluAds`, which was bought out by `BuySellAds` in 2013 [4], and as a result had its filters re-issued under its parent company in the new template.

[4]We tested the accuracy of our regular expression. Out of 329 partners in a given version, we were able to retrieve 322.

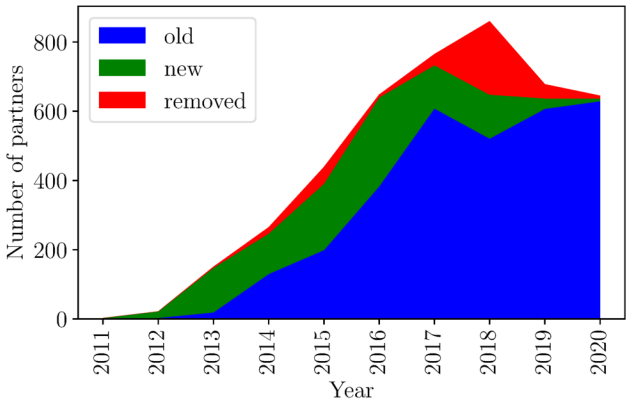


Figure 7: Evolution of partners joining the Acceptable Ads initiative. Exception list went through a major change in 2018 (i.e., removal of partners), but most of the exception filters were re-issued under large publisher.

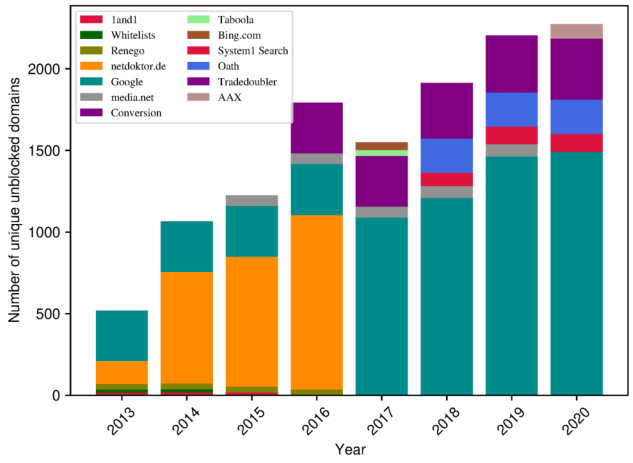
Finding 4: *The exception list had undergone a major revision in 2018 after which it removed zombie partners and stale domains. We see that many smaller partners ceased to exist as their rules were merged with larger entities.* Many of the exception rules issued by smaller entities were re-issued under large publishers like `AAX` and `Taboola`.

6.2 Top Partners and Unblocked Domains

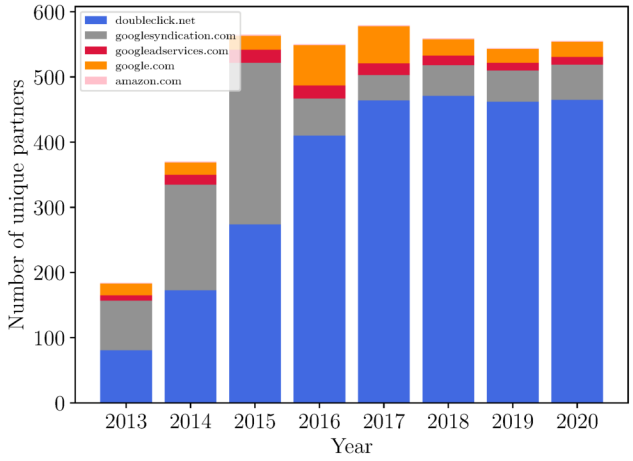
Partners that publish more exception filters have a greater chance of successful ad impressions. We identified top partners that benefit through the Acceptable Ads program by listing domains from which contents are unblocked. For this part of the analysis, we skip versions of the exception list from 2011 to 2012 because the number of partners was negligible.

To understand the evolution of the dominant partners of the Acceptable Ads program, we rank the partners by the number of unique domains they unblock. For our analysis, we only focus on the top five partners from each year's latest commit. Figure 8a highlights the most contributing partners from each year. In the early years, we see that `NetDoktor` remained the most dominant partner until 2016. Since 2016 it no longer appears as one of the top five contributing partners. After inspection, we found out that filters issued by `NetDoktor` were constantly violating the criteria set by the Acceptable Ad Committee. After issuing multiple warnings, the Acceptable Ad Committee started removing exception rules issued by `NetDoktor` [5]. At the same time, we see increased participation from `Google` as it has been securing more ad impressions from different domains. For the past four years, `Google` has been the top contributor to the exception list.

Figure 8b shows the evolution of the top five domains that are unblocked by the most number of unique partners. We see that `doubleclick.net`, `googlesyndication.com`, `googleadservices.com`, and `google.com` are present among the top domains from which partners unblock contents. All of these domains are owned by `Google`. Distribution of the unblocked domains for the most recent version of the exception list is given in Appendix A.



(a) Top partners



(b) Top unblocked domains

Figure 8: Top partners and unblocked domains in the exception list from 2013 to 2020. We see that `Google` not only unblocks the most number of unique domains, but is also unblocked by the most number of unique partners

Finding 5: *Google not only unblocks the most number of unique domains, but is also unblocked by the most number of unique partners. The domination of Google in both fronts means that Google and its ad exchange platforms have positioned themselves strongly before ad publishers.* This form of partnership translates into financial gains as `Google` can guarantee higher ad impressions even when users install ad-blocking tools like Adblock Plus.

6.3 Overlap among Partners

Partners in the Acceptable Ads program may unblock contents originating from similar domains. This can mean that they are unblocking contents from each other. We wanted to explore if there is a reciprocal relationship between partners.

The total number of partners in the latest version of the exception list is more than 600. However, a small number of partners are the major contributors to the list. Figure 9 shows a CDF of the unique domains unblocked by different partners. We can see that the top 20 partners [5] covered around 90% of unique unblocked domains. We, therefore, focus our analysis on these top 20 partners. Next, we look at the common domains they unblock.

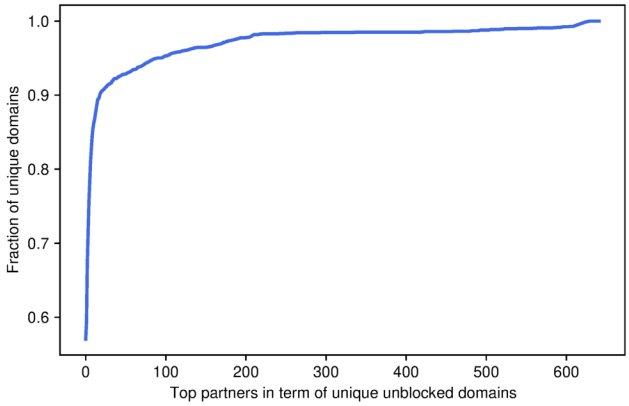


Figure 9: CDF of domain coverage by partners. Top 20 partners cover 90% of the unique domains.

Figure 10 shows a chord diagram for the top 20 contributors. Figure 10 enables us to visualize the overlap of domains unblocked by the top 20 partners. The arc length of each partner depicts the total number of unique domains it unblocks. Not surprisingly, `Google` has the largest arc length. The width of the connection between partners represents the volume of domains they share with each other. A connection that goes back to the node itself represents the domains that a given partner does not share with anyone (i.e., those domains are unique to a given partner).

Finding 6: *Apart from* `Google` *and* `TradeDoubler`*, the overlap of domains between partners is negligible. This implies that each partner is either unblocking content from its own domain(s) or another publisher that has not yet partnered with the Acceptable Ads program.*

7 IMPACT OF ACCEPTABLE ADS PROGRAM

Adblock Plus by default enables acceptable ads. This means that users are shown ads by default. While the Acceptable Ads Committee has provided specific criteria that ad publishers must follow to have their ads shown to end users, the overall vetting process is not fully transparent. As a result it is not surprising that we have seen instances where offended end users have harshly criticized some decisions made by the Acceptable Ads Committee [6].

In this section, we study the impact of the exception list endorsed by the Acceptable Ads program. For this purpose, we collect web traces by visiting the top 10k sites listed by Tranco [62]. We then analyze all web requests made while loading these web pages. In

[5]Top in terms of the number of domains unblocked

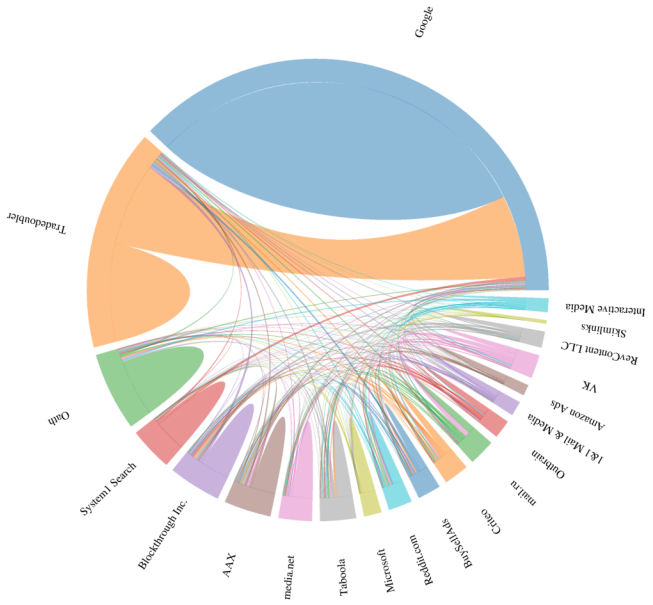


Figure 10: Overview of domains unblocked by the top 20 partners. The overlap of domains between most partners is negligible.

total, we analyzed 1,333,925 web requests. In order to model the behavior of Adblock Plus extension, we use Adblock Plus filter parser [9]. We instrumented the parser to return not only whether a web request should be blocked or not, but also the exact matching rule. The parser follows the pipeline depicted in Figure 1 to either block, unblock or allow a request to pass. The last case is classified as 'benign', since no blocking filter matched with the web request. We perform a *differential analysis*, where we apply different combinations of filters on the web requests and record the outcomes to provide insights into their impact on user privacy.

7.1 Triggered Filters

We analyze the impact of different combinations of filters and exceptions. Since there are three lists: EasyList, EasyPrivacy and Exception list. A total of six possible configurations are feasible.[6] Table 2 highlights the number of unique filtering and exception rules available under different combinations of lists and those that were triggered under the six configurations. One thing to note, even without the exception list endorsed by the Acceptable Ads initiatives there are exception rules in EasyList and EasyPrivacy. These exception rules exist as at times it is easier to block everything and allow a few exceptions instead of writing multiple blocking rules. We see that under the default setting when both EasyList and Exception lists are enabled, a very small percentage of rules are used/triggered. For instance, only 1.01% (156/14,159) of the exception filters and 1.79% (1,180/65,830) of the blocking filters are used under default setting. In contrast, when only EasyPrivacy is applied, 14.9% (2,475/16,515) of the blocking filters is activated. This outcome is largely consistent with our previous finding (Figure 5) in

[6]Out of the eight theoretical configurations two configurations are not feasible, one case where no lists are used and the other case is when only the exception list is used.

Table 2: Number of unique filters triggered and web requests blocked/unblocked under different filter configurations.

Configuration	Unique rules		Unique rules triggered		Number of web requests		
	Blocking	Exception	Blocking	Exception	Blocked	Unblocked	Allowed/benign
EasyList	65830	2393	1307	25	314649	255	1019021
EasyPrivacy	16515	666	2475	35	280097	895	1052933
EasyList + EasyPrivacy	82345	3059	3669	75	544400	2297	787228
EasyList + Exception *	65830	14159	1180	156	219373	95722	1018830
EasyPrivacy + Exception	16515	12432	2163	167	141647	78562	1113716
Combined	85005	14825	3380	283	406458	140270	787197

* default configuration in Adblock Plus extension

Section 5.2, where we show that unlike EasyPrivacy, both EasyList and Exception list use the 'domain' restrictions more extensively to limit the domains from which contents can be loaded.

Another major insight from Table 2 is the increase in the number of exception rules triggered (167-35=132) in the presence of the exception list while using EasyPrivacy. According to Table 2, these 132 exception filters were responsible for unblocking 77,667 (78,562-895) web resources, some of which are requests made by trackers. Exception list is meant to allow ads from publishers that are partners in the Acceptable Ads program. EasyPrivacy typically blocks tracking attempts like scripts that perform fingerprinting attempts or 1-by-1 pixel images that allow tracking. This is concerning because the exception list should unblock advertisements only. EasyPrivacy, on the other hand, is an anti-tracking list that has nothing to do with advertisements. Thus, this questions why some EasyPrivacy filters are negated by the exception list and what are their implications.

Finding 7: *The exception list endorsed by Acceptable Ads program not only unblocks advertisements, but also trackers generally filtered by EasyPrivacy, thereby potentially exposing end users to online tracking.* This potentially violates the purpose of Acceptable Ads of allowing only non-intrusive ads and not trackers.

7.2 Impact on Trackers

Adblock Plus claims that some acceptable ads comply with Do Not Track policy, and/or ads which are served by the domain which is wholly owned by the same company [27]. However, these rules may not necessarily apply to all the unblocked ads. In this section, we try to analyze the benefits that top publishing companies gain through the Acceptable Ads program and understand what type of privacy implications that has on end users. The idea is to compare how the content distributed by top ad publishers are affected under different filter configurations. The inferences drawn from our results can also help readers make an informed decision about the best configuration that suits their privacy needs.

We start by mapping our web requests to the owners of the domains. For this purpose, we merged the tracker list provided by WhoTracksMe [37] and Disconnect [26]. Both of these lists map domains owned by different trackers. We then first extract eTLD+1 from web requests and use this merged tracker list to map domains to companies. We were able to map the domains of web requests to 843 unique companies. For this analysis, we focused on the top 10 companies that owned most of the requested resources. Out of the 1,333,925 web requests, 32% were owned by the top 10 companies (i.e., they were most prevalent). Using the Adblock Plus filter parser we compute what proportion of web requests, owned by these top 10 companies, would be *blocked*, *unblocked* or *allowed*.

Figure 11 highlights the resulting distribution of blocked, unblocked or allowed web requests. Figure 11a to Figure 11c show the portion of traffic allowed and blocked without the presence of the exception list endorsed by the Acceptable Ads program, and thus provide us with baseline numbers. Figure 11d to Figure 11f show the impact of Acceptable Ads program on the proportion of web requests that are unblocked (in orange color). We see that except for `Facebook` and `Verizon`, all top companies benefit from exception rules. For example, `Google` owned 232,780 of the web requests. In Figure 11b, almost 18% of `Google's` web requests are blocked by EasyPrivacy. However, when the exception list is enabled in Figure 11e, 14.6% of web requests are unblocked, giving `Google` almost 33,959 more ad impressions. Thus, out of the 42,210 Google owned web requests that were blocked in the EasyPrivacy-only configuration, around 80% of such requests were unblocked by the exception list. Similarly, we see in Figure 11a that all web requests owned by `PubMatic` (25,896 in total) are blocked due to EasyList. However, with the exception list enabled (Figure 11d), almost 56% of these web resources become unblocked. However, since none of `PubMatic`'s web resources are blocked by EasyPrivacy (as seen from Figure 11b), this implies that the company generally operates as an ad publisher.

Furthermore, Figure 11e also shows exception filters unblocking contents from `AppNexus`, `Adobe` and `Google` that were originally blocked by EasyPrivacy (as evident from Figure 11b). Ideally, Acceptable Ads program should not interfere with contents blocked by EasyPrivacy — an anti-tracking list and not an ad-blocking list.

We also determine the exact filtering and exception rules that are triggered the most. Table 3 lists the top five triggered filter and exception rules. We are interested in determining the exception rules that are overpowering EasyPrivacy's blocking filters. From Table 3 we can see that the most triggered rule was:
`@@||cm.g.doubleclick.net/pixel?$image,third-party`
which was activated 17,867 times. This is a 1-by-1 pixel-based tracker. Similarly, we found several instances of other 1-by-1 pixel-based trackers as shown below –

(a) EasyList (b) EasyPrivacy (c) EasyList + EasyPrivacy

(d) EasyList + Exception (e) EasyPrivacy + Exception (f) EasyList + EasyPrivacy + Exception

Figure 11: Impact of exception list on traffic bounded for top 10 destinations. We see that many publishers benefit from the Acceptable Ads program. Worryingly, exception rules can also unblock trackers that would otherwise be blocked by EasyPrivacy.

```
@@||adnxs.com/pixie^$image,third-party,
@@||adroll.com/pixel^$script,image,third-party,
@@||pixel.mathtag.com/sync/img?$image,third-party,
@@||beacon.krxd.net/pixel.gif$image,third-party
```

All of these domains are owned by companies that distribute tracking resources for advertisement purposes. In fact, according to WhoTracksMe, 20.2% of web traffic is tracked by `doubleclick.net` [28].

Finding 8: *The default inclusion of Acceptable Ads creates privacy risks as the exception rules can unblock trackers that would otherwise be blocked by EasyPrivacy.* Furthermore, the term 'Acceptable Ads' can be misleading as it creates the impression that the platform facilitates only non-intrusive visible ads, which in this case is not fully true.

8 DISCUSSION

Summary. We analyze the evolution of the exception list endorsed by the Acceptable Ads program and contrast it with the growth of anti-ad (EasyList) and anti-tracking (EasyPrivacy) filters to understand the key web resources that are being unblocked by the exception list. We also collect real-world web traffic and model Adblock Plus's behavior under different configurations to understand the impact of enabling exception filters on the different types of web resources fetched.

Our work provides the following insights: 1) the exception list endorsed by the Acceptable Ads program poses privacy risks for end users as it can unblock tracking resources on the browser. We model the extension to show how exception rules unblocked 1-by-1 pixel images requested from different third-party tracking domains; 2) the default inclusion of Acceptable Ads needs to be revisited and possibly users should be prompted for consent or at least be made aware of the implications; 3) the Acceptable Ads Committee should review exception lists for rules that negate filters issued by EasyPrivacy. More, importantly the committee should properly vet what type of blocking filters the exception rules modify, possibly through some automated auditing process. We have open sourced our parsing tools to the research community [81].

Limitations. There are a few limitations in our work. First, in explaining sudden fluctuations in trends we relied on official reports and dedicated forum pages to understand the reasons for such abrupt changes. We, therefore, may miss some undisclosed information. Second, we only crawl the homepage of the top 10k popular websites to model the filtering behavior of Adblock Plus under different settings. Crawling sub-pages may result in somewhat different results. However, even with simply visiting homepages we show the privacy risks imposed by the default inclusion of Acceptable Ads. Our results can therefore be thought of as lower bounds. Third, our traffic analysis involved mapping domain to organizations. We used tracker information from WhoTracksMe [37] and Disconnect [26]. However, we were still unable to account for many of the requests belonging to lesser known entities. Automatically mapping domains to organizations is still an open problem. However, we believe our mapping is sufficient to cover the dominant players in the exception list. Moreover, our analysis on the overlap between partners was based on how the unblocked domains are being shared between the partners. This analysis could further benefit from looking deeper into the frequency of the unblocked

Table 3: Top 5 rules triggered under different configurations.

Configuration	Blocking Filters	Exception Filters
EasyList	||pubmatic.com^$third-party	@@||adfox.ru^$ third-party
	||cm.g.doubleclick.net^$popup,third-party	@@||npttech.com/advertising.js$script
	||rubiconproject.com^$third-party	@@||cheatsheet.com^$generichide
	/\.com\/[0-9]{2,9}\/$/$script,stylesheet,third-party,xmlhttprequest	@@||infowars.com^$generichide
	/\.com\/[0-9]{2,9}\/$/$script,stylesheet,third-party,xmlhttprequest	@@||streamtheworld.com^$media,third-party
EasyPrivacy	||doubleclick.net^$image,third-party	@@||munchkin.marketo.net/munchkin.js
	||cm.g.doubleclick.net^	@@||adobedtm.com/launch-$script
	||facebook.com/tr$third-party	@@||omtrdc.net/rest/$xmlhttprequest
	||bidswitch.net^$third-party	@@||amplitude.com/libs/amplitude-$script,third-party
	||everesttech.net^$third-party	@@||marketo.com/js/forms2/$script,stylesheet
EasyList+ EasyPrivacy	||pubmatic.com^$third-party	@@||adfox.ru^$ third-party
	||cm.g.doubleclick.net^$popup,third-party	@@||munchkin.marketo.net/munchkin.js
	||rubiconproject.com^$third-party	@@||adobedtm.com/launch-$script
	||openx.net^$third-party	@@||omtrdc.net/rest/$xmlhttprequest
	/\:\/\/[a-z0-9]{5,}\.com\/[A-Za-z0-9]{3,}\/$/$script,third-party,xmlhttprequest	@@||amplitude.com/libs/amplitude-$script,third-party
EasyList + Exception	/\.com\/[0-9]{2,9}\/$/$script,stylesheet,third-party,xmlhttprequest	@@||cm.g.doubleclick.net/pixel?$image,third-party
	||pubmatic.com^$third-party	@@||pubmatic.com/AdServer/Pug?$image,third-party
	/\:\/\/[a-z0-9]{5,}\.com\/[A-Za-z0-9]{3,}\/$/$script,third-party,xmlhttprequest	@@||match.adsrvr.org/track/cmf/$image,third-party
	||33across.com^$third-party	@@||openx.net/w/1.0/sd?$image,third-party
	||gumgum.com^$third-party	@@||stats.g.doubleclick.net^$script,image
EasyPrivacy + Exception	||google-analytics.com/analytics.js	@@||cm.g.doubleclick.net/pixel?$image,third-party
	/r/collect?	@@||stats.g.doubleclick.net^$script,image
	||districtm.io^$third-party	@@||x.bidswitch.net/sync^$image,third-party
	||33across.com^$third-party	@@||pagead2.googlesyndication.com/pagead/gen_204?$image
	||bidr.io^$third-party	@@||sync-tm.everesttech.net/upi/pid^$image,third-party
Combined	/\:\/\/[a-z0-9]{5,}\.com\/[A-Za-z0-9]{3,}\/$/$script,third-party,xmlhttprequest	@@||cm.g.doubleclick.net/pixel?$image,third-party
	/\.com\/[0-9]{2,9}\/$/$script,stylesheet,third-party,xmlhttprequest	@@||pubmatic.com/AdServer/Pug?$image,third-party
	||google-analytics.com/analytics.js	@@|http://$popup,sitekey=MFww..Q
	/r/collect?	@@||match.adsrvr.org/track/cmf/$image,third-party
	||pubmatic.com^$third-party	@@||stats.g.doubleclick.net^$script,image

content that is shared between the partners. Lastly, in determining partners, we were hindered by the unstructured formatting of the exception list in versions that preceded November 2018. To obtain the names of the partners, we used regular expressions that sometimes failed to return valid names for the partners. However, we manually verified the accuracy of our approach. Out of 329 partners in a given version, we were able to retrieve 322 (i.e., our approach had 98% accuracy).

9 CONCLUSION

In this paper, we perform an in-depth analysis of the Adblock Plus's Acceptable Ads program. We perform a longitudinal analysis of filtering and exception rules, and show that most ads are served through scripts and images. We also show that exception rules have shorter life span compared to blocking rules. When it comes to partners contributing most to the exception list, we see that Google not only unblocks the most number of unique domains, but is also unblocked by the most number of partners. Lastly, we find that the exception list endorsed by the Acceptable Ads program unblocks many pixel-based trackers, originally blocked by EasyPrivacy. Thus, The Acceptable Ads initiative poses privacy risks to end users as it negates EasyPrivacy filtering rules, by default.

ACKNOWLEDGEMENT

We thank our anonymous reviewers for their feedback. This material is based upon work supported in parts by the National Science Foundation under grant number CNS-1849997. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] 2007. *uBlock Origin.* https://github.com/gorhill/uBlock/
[2] 2013. *Fanboy merges with easylist.* https://easylist.to/2013/05/17/easylist-merges-with-fanboy-s-list.html.
[3] 2013. *Google search ads added.* https://hg.adblockplus.org/exceptionrules/rev/8bdf815a5291.
[4] 2013. *InfluAds bought by BuySellAds.* http://www.finsmes.com/2013/09/influads-shuts.html.
[5] 2013. *NetDoktor removed.* https://adblockplus.org/forum/viewtopic.php?f=12&t=15361.
[6] 2014. *Taboola ads.* https://adblockplus.org/forum/viewtopic.php?f=12&t=25991.
[7] Adblock Plus 2016. *Help Create An Exception To Easy List Filter.* Adblock Plus. https://adblockplus.org/forum/viewtopic.php?f=1&t=43814.
[8] 2016. *Problematic filters.* https://forums.lanik.us/viewtopic.php?f=64&t=44687.
[9] 2017. *ABP Filter Parser.* https://github.com/bbondy/abp-filter-parser
[10] 2017. *Mercurial repository for Adblock Plus's Exception list.* https://hg.adblockplus.org/exceptionrules/file.
[11] 2017. *Removed whitelists for Text ads on parked domains from Uniregistry.* https://hg.adblockplus.org/exceptionrules/rev/2e08652dbcf7.
[12] 2017. *Removed whitelists for thoughtcatalog.com ads.* https://hg.adblockplus.org/exceptionrules/rev/2f597226c09.
[13] 2018. *Restructured partner comments and sections.* https://hg.adblockplus.org/exceptionrules/rev/f65613789b62.
[14] 2020. *AAC history.* https://eyeo.com/acceptable-ads-committee-history/.
[15] Adblock Plus 2020. *Acceptable Ads.* Adblock Plus. https://adblockplus.org/en/about#monetization
[16] 2020. *Adblcok Plus.* https://adblockplus.org/
[17] 2020. *AdBlock: Block Ads.* https://getadblock.com/
[18] 2020. *Adblock Plus passes 200 million active users.* https://eyeo.com/acceptable-ads-by-the-numbers/
[19] Adblock Plus 2020. *Allowing acceptable ads in Adblock Plus.* Adblock Plus. https://adblockplus.org/acceptable-ads.
[20] 2020. *Brave Browser.* https://brave.com/

[21] Adblock Plus 2020. *Canonical repository for EasyList.* Adblock Plus. https://easylist-downloads.adblockplus.org/easylist.txt

[22] Adblock Plus 2020. *Canonical repository for EasyPrivacy.* Adblock Plus. https://easylist-downloads.adblockplus.org/easyprivacy.txt

[23] 2020. *Chrome Har capturer.* https://github.com/cyrus-and/chrome-har-capturer.

[24] 2020. *Cliqz.* https://cliqz.com/en/desktop

[25] 2020. *Disconnect.* https://disconnect.me/

[26] 2020. *Disconnect tracker list.* https://github.com/disconnectme/disconnect-tracking-protection

[27] 2020. *Do not track ads.* https://adblockplus.org/acceptable-ads#privacy-friendly-acceptable-ads.

[28] 2020. *Double Click tracker stats.* https://whotracks.me/trackers/doubleclick.html.

[29] 2020. *EasyList.* https://easylist.to/

[30] 2020. *EasyList GitHub.* https://github.com/easylist

[31] 2020. *EasyPrivacy.* https://easylist.to/easylist/easyprivacy.txt

[32] 2020. *Epic Privacy Browser.* https://www.epicbrowser.com/

[33] 2020. *Get whitelisted.* https://adblockplus.org/acceptable-ads#get-whitelisted.

[34] Cliqz 2020. *Ghostery.* Cliqz. https://www.ghostery.com/

[35] Adblock Plus 2020. *How to write filters.* Adblock Plus. https://help.eyeo.com/en/adblockplus/how-to-write-filters

[36] 2020. *Privacy Badger.* https://www.eff.org/privacybadger

[37] 2020. *Tracker List provided by whotracksme.* https://whotracks.me/trackers.html.

[38] Gunes Acar, Christian Eubank, Steven Englehardt, Marc Juarez, Arvind Narayanan, and Claudia Diaz. 2014. The Web Never Forgets: Persistent Tracking Mechanisms in the Wild. In *Proceedings of the 21st ACM SIGSAC Conference on Computer and Communications Security (CCS)*. 674–689.

[39] Gunes Acar, Marc Juarez, Nick Nikiforakis, Claudia Diaz, Seda Gürses, Frank Piessens, and Bart Preneel. 2013. FPDetective: Dusting the Web for Fingerprinters. In *Proceedings of the 20th ACM SIGSAC Conference on Computer and Communications Security (CCS)*. 1129–1140.

[40] Mshabab Alrizah, Sencun Zhu, Xinyu Xing, and Gang Wang. 2019. Errors, Misunderstandings, and Attacks: Analyzing the Crowdsourcing Process of Ad-blocking Systems. In *Proceedings of the 19th Internet Measurement Conference (IMC)*. 230–244.

[41] Rebecca Balebako, Pedro G. Leon, Richard Shay, Blase Ur, Yang Wang, and Lorrie Faith Cranor. 2012. Measuring the effectiveness of privacy tools for limiting behavioral advertising. In *Web 2.0 Workshop on Security and Privacy (W2SP)*.

[42] Michael Barbaro and Tom Zeller Jr. 2006. *A Face Is Exposed for AOL Searcher No. 4417749.* The New York Times. https://www.nytimes.com/2006/08/09/technology/09aol.html

[43] Yinzhi Cao, Song Li, and Erik Wijmans. 2017. (Cross-) Browser Fingerprinting via OS and Hardware Level Features.. In *Proceedings of the 24th Annual Network and Distributed System Security Symposium (NDSS)*.

[44] Anupam Das, Nikita Borisov, and Matthew Caesar. 2016. Tracking mobile web users through motion sensors: Attacks and defenses. In *Proceeding of the 23rd Annual Network and Distributed System Security Symposium (NDSS)*.

[45] Anupam Das, Nikita Borisov, and Edward Chou. 2018. Every move you make: Exploring practical issues in smartphone motion sensor fingerprinting and countermeasures. *Proceedings on the 18th Privacy Enhancing Technologies (PoPETs)* 1 (2018), 88–108.

[46] Peter Eckersley. 2010. How Unique is Your Web Browser?. In *Proceedings of the 10th Privacy Enhancing Technologies Symposium (PETS)*. 1–18.

[47] Steven Englehardt and Arvind Narayanan. 2016. Online Tracking: A 1-million-site Measurement and Analysis. In *Proceedings of the 23rd ACM SIGSAC Conference on Computer and Communications Security (CCS)*. 1388–1401.

[48] Steven Englehardt, Dillon Reisman, Christian Eubank, Peter Zimmerman, Jonathan Mayer, Arvind Narayanan, and Edward W. Felten. 2015. Cookies That Give You Away: The Surveillance Implications of Web Tracking. In *Proceedings of the 24th International Conference on World Wide Web (WWW)*. 289–299.

[49] David Fifield and Serge Egelman. 2015. Fingerprinting Web Users Through Font Metrics. In *Proceedings of the 19th International Conference on Financial Cryptography and Data Security (FC)*. 107–124.

[50] Gertjan Franken, Tom Van Goethem, and Wouter Joosen. 2018. Who Left Open the Cookie Jar? A Comprehensive Evaluation of Third-Party Cookie Policies. In *Proceedings of the 27th USENIX Security Symposium (USENIX Security)*. 151–168.

[51] Kiran Garimella, Orestis Kostakis, and Michael Mathioudakis. 2017. Ad-blocking: A study on performance, privacy and counter-measures. In *Proceedings of the 2017 ACM on Web Science Conference*. 259–262.

[52] Arthur Gervais, Alexandros Filios, Vincent Lenders, and Srdjan Capkun. 2017. Quantifying web adblocker privacy. In *Proceedings of the 22nd European Symposium on Research in Computer Security (ESORICS)*. 21–42.

[53] Saad Sajid Hashmi, Muhammad Ikram, and Mohamed Ali Kâafar. 2019. A Longitudinal Analysis of Online Ad-Blocking Blacklists. *CoRR* abs/1906.00166 (2019). http://arxiv.org/abs/1906.00166

[54] Kashmir Hill. 2012. *How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did.* Forbes. https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#46438eb96668

[55] Raymond Hill. 2014. *Comparative benchmarks against widely used blockers: Top 15 Most Popular News Websites.* https://github.com/gorhill/httpswitchboard/wiki/Comparative-benchmarks-against-widely-used-blockers:-Top-15-Most-Popular-News-Websites

[56] Raymond Hill. 2015. *uBlock and others: Blocking ads, trackers, malwares.* https://github.com/gorhill/uBlock/wiki/uBlock-and-others%3A-Blocking-ads%2C-trackers%2C-malwares

[57] Thomas Hupperich, Davide Maiorca, Marc Kührer, Thorsten Holz, and Giorgio Giacinto. 2015. On the robustness of mobile device fingerprinting: Can mobile users escape modern web-tracking mechanisms?. In *Proceedings of the 31st Annual Computer Security Applications Conference (ACSAC)*. ACM, 191–200.

[58] Umar Iqbal, Zubair Shafiq, and Zhiyun Qian. 2017. The Ad Wars: Retrospective Measurement and Analysis of Anti-adblock Filter Lists. In *Proceedings of the 17th Internet Measurement Conference (IMC)*. 171–183.

[59] T. Kohno, A. Broido, and K. C. Claffy. 2005. Remote physical device fingerprinting. *IEEE Transactions on Dependable and Secure Computing* 2, 2 (April 2005), 93–108.

[60] Pierre Laperdrix, Nataliia Bielova, Benoit Baudry, and Gildas Avoine. 2019. Browser Fingerprinting: A survey. *CoRR* abs/1905.01051 (2019). http://arxiv.org/abs/1905.01051

[61] Pierre Laperdrix, Walter Rudametkin, and Benoit Baudry. 2016. Beauty and the beast: Diverting modern web browsers to build unique browser fingerprints. In *Proceedings of the 37th IEEE Symposium on Security and Privacy (S&P)*. 878–894.

[62] Victor Le Pochat, Tom Van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczyński, and Wouter Joosen. 2019. Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation. In *Proceedings of the 26th Annual Network and Distributed System Security Symposium (NDSS)*.

[63] Zhou Li, Kehuan Zhang, Yinglian Xie, Fang Yu, and XiaoFeng Wang. 2012. Knowing your enemy: Understanding and detecting malicious web advertising. In *Proceedings of the 19th ACM conference on Computer and Communications Security (CCS)*. 674–686.

[64] Jonathan Mayer. 2011. *Tracking the Trackers: Self-Help Tools.* http://cyberlaw.stanford.edu/blog/2011/09/tracking-trackers-self-help-tools

[65] J. R. Mayer and J. C. Mitchell. 2012. Third-Party Web Tracking: Policy and Technology. In *Proceedings of the 33rd IEEE Symposium on Security and Privacy (S&P)*. 413–427.

[66] TJ McCue. 2019. *47 Percent Of Consumers Are Blocking Ads.* Forbes. https://www.forbes.com/sites/tjmccue/2019/03/19/47-percent-of-consumers-are-blocking-ads/#5d9d4dd42037

[67] Georg Merzdovnik, Markus Huber, Damjan Buhov, Nick Nikiforakis, Sebastian Neuner, Martin Schmiedecker, and Edgar Weippl. 2017. Block me if you can: A large-scale study of tracker-blocking tools. In *Proceedings of the 2017 IEEE European Symposium on Security and Privacy (EuroS&P)*. 319–333.

[68] Keaton Mowery and Hovav Shacham. 2012. Pixel Perfect: Fingerprinting Canvas in HTML5. In *Web 2.0 Workshop on Security and Privacy (W2SP)*.

[69] N. Nikiforakis, A. Kapravelos, W. Joosen, C. Kruegel, F. Piessens, and G. Vigna. 2013. Cookieless Monster: Exploring the Ecosystem of Web-Based Device Fingerprinting. In *Proceedings of the 33rd IEEE Symposium on Security and Privacy (S&P)*. 541–555.

[70] Łukasz Olejnik, Gunes Acar, Claude Castelluccia, and Claudia Diaz. 2015. The leaking battery. In *Proceedings of the 10th International Workshop Data Privacy Management, and Security Assurance*. 254–263.

[71] Enric Pujol, Oliver Hohlfeld, and Anja Feldmann. 2015. Annoyed users: Ads and ad-block usage in the wild. In *Proceedings of the 15th Internet Measurement Conference (IMC)*. 93–106.

[72] Franziska Roesner, Tadayoshi Kohno, and David Wetherall. 2012. Detecting and Defending Against Third-Party Tracking on the Web. In *Proceedings of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*. 155–168.

[73] Alexander Sjosten, Peter Snyder, Antonio Pastor, Panagiotis Papadopoulos, and Benjamin Livshits. 2019. *Generation of Filter Lists for Regions that are Underserved.*

[74] Peter Snyder, Antoine Vastel, and Ben Livshits. 2020. Who Filters the Filters: Understanding the Growth, Usefulness and Efficiency of Crowdsourced Ad Blocking. In *Proceedings of the 2020 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*. 75–76.

[75] Ashkan Soltani, Shannon Canty, Quentin Mayo, Lauren Thomas, and Chris Jay Hoofnagle. 2009. Flash Cookies and Privacy. *Available at SSRN 1446862* (2009).

[76] Jan Spooren, Davy Preuveneers, and Wouter Joosen. 2015. Mobile device fingerprinting considered harmful for risk-based authentication. In *Proceedings of the 8th European Workshop on System Security (EuroSec)*. ACM, 1–6.

[77] Stefano Traverso, Martino Trevisan, Leonardo Giannantoni, Marco Mellia, and Hassan Metwalley. 2017. Benchmark and comparison of tracker-blockers: Should you trust them?. In *Proceedings of the 1st Network Traffic Measurement and Analysis Conference (TMA)*. 1–9.

[78] Robert J. Walls, Eric D. Kilmer, Nathaniel Lageman, and Patrick D. McDaniel. 2015. Measuring the Impact and Perception of Acceptable Advertisements. In *Proceedings of the 15th Internet Measurement Conference (IMC)*. 107–120.

[79] Craig E. Wills and Doruk C. Uzunoglu. 2016. What ad blockers are (and are not) doing. In *Fourth IEEE Workshop on Hot Topics in Web Systems and Technologies*

(HotWeb). 72–77.

[80] Zhonghao Yu, Sam Macbeth, Konark Modi, and Josep M. Pujol. 2016. Tracking the Trackers. In *Proceedings of the 25th International Conference on World Wide Web (WWW)*. 121–132.

[81] Ahsan Zafar. 2020. *Analysis of Acceptable Ads.* https://github.ncsu.edu/wspr/acceptable-ads

[82] Apostolis Zarras, Alexandros Kapravelos, Gianluca Stringhini, Thorsten Holz, Christopher Kruegel, and Giovanni Vigna. 2014. The dark alleys of madison avenue: Understanding malicious advertisements. In *Proceedings of the 14th Conference on Internet Measurement Conference (IMC)*. 373–380.

APPENDIX

A TOP 20 SURROGATE AND UNBLOCKED DOMAINS

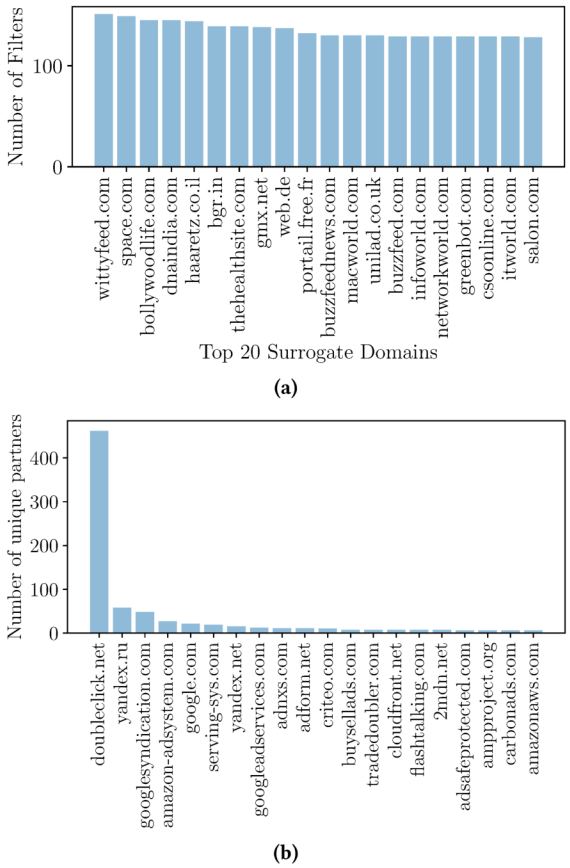


Figure 12: Top 20 surrogate and unblocked domains in the latest version of the exception list. (a) surrogate domains that featured the highest number of times; (b) domains that are unblocked by the most number of unique partners.