Analyzing the Effectiveness of Internet Privacy Tools

Alish Akhmetzyanov
University of Minnesota
akhme004@umn.edu

Austin Altmann
University of Minnesota
altma064@umn.edu

Allen Hansen
University of Minnesota
hans5117@umn.edu

Josh Foschi
University of Minnesota
fosch006@umn.edu

Connor Van Drisse
University of Minnesota
vandr042@umn.edu

December 16, 2016

Abstract

Our paper presents a direct comparison between several web browsing extensions (Ghostery, Adblock Plus, Adblocker Ultimate, and Disconnect) and their effectiveness in blocking third-party cookies. Further, it examines the effectiveness of a Tor proxy in blocking third-party domains (note that this is not its goal). Data collection occurred over the top 5,000 websites, as ranked by Amazon Alexa, with each site being visited by 3 different instances of the Firefox browser for a total of 15,000 GET requests.

Our tests showed that Ghostery was the most effective privacy extension examined in terms of both the amount of third-party cookies it allowed as well as the percentage of first-party domains on which each third-party domain allowed by Ghostery appeared. Adblock Plus and AdBlocker Ultimate both performed very similarly to each other, while Disconnect performed slightly worse than Ghostery. The Tor proxy was not expected to be very effective, as cookie blocking is not its goal, but there was some improvement over a vanilla Firefox browser.

1 Introduction

Targeted advertising is far from uncommon in the current state of the Internet. In fact, it has been found to be of significant concern to 80% of modern day consumers[20]. Of course, there are many benefits to this method of advertising and the use of third-party cookies. For example: users receive a custom-tailored advertising experience, a significant portion of the Internet remains free to consumers, and research is able to be conducted into search trends and human events[13]. However, underlying problems persist in the perception and state of the third-party cookies that are being used. Users often feel that an invasion of privacy has occurred when information is recorded about them, sometimes involuntarily. And even if a user consents to the use of third-party cookies, his or her personal information can still be sold to an unknown entity[13]. Over time, profiles can be developed of a user that will contain crucial aspects of his or her life that are not limited to family, friends, hobbies, purchases, vacations, or browser history. This information can be used in background checks, determining insurance rates, evidence in legal investigations, politically targeted fundraising, etc.[13].

Privacy tools are some of the most popular browser extensions on platforms such as Firefox, Chrome, and Safari. In this paper, four browser extensions (Ghostery, Adblock Plus, AdBlocker Ultimate, and Disconnect) will be tested for their effectiveness at avoiding third-party tracking compared to each other and a default browser in order to examine how easily an average user can mitigate privacy concerns. This paper will also examine where these third-party trackers are coming from and which ones persist through these extensions. Further, the Tor network will also be tested as an alternative to extension-using on a vanilla Firefox browser for its ability to provide privacy against third-party tracking.

Throughout, this paper will extend upon the work in Steven Englehardt and Arvind Narayanan's "Online Tracking: A 1-million-site Measurement and Analysis" by using OpenWPM to analyze the top 5,000 websites of the Amazon Alexa top one million and directly comparing stateful browsing sessions with no extensions to Ghostery, Adblock Plus, AdBlocker Ultimate, and Disconnect, as well as a session of the top 5,000 sites with the Tor network. The session with no extensions is used as a control to analyze the effectiveness of an extension in maintaining privacy against third-party trackers, and the comparisons between extensions and the Tor network are to analyze the effectiveness of each utility's privacy implementations. Further detail on the methodology is provided in Section 4. We hope to highlight how simple additions to a browser can greatly improve privacy, as well as how third-party tracking is able to penetrate these tools and where this personal data is going to.

2 Motivation

The use of cookies has become a standard practice in most modern web applications. Cookies play a very important role in adding state to the stateless HTTP/S protocol and making internet transactions more efficient and individualized. Common interactions with cookies include saving page preferences, storing selections from previously visited pages, and enabling automatic authentication for logins. The cookies that provide such mechanism are associated with the web server specified by the page URL and are called first-party cookies.

In our paper, we are focusing on third-party cookies. Third-party cookies are loaded through the web page content that references servers in other domains such as links, images, JavaScript and HTML IFrames[21]. Although first-party, as well as many third-party cookies, are useful, there is a potential for third-party cookies to be used in ways that make users uncomfortable, typically stemming from profile building.

The integrity and confidentiality of third-party cookies, many used by advertising companies, is raising serious concerns. Cookie transactions, such as storing or sending cookies, are hidden from the users with default browser settings. Some of the ways advertising companies use cookies includes targeted advertising, behavioral analysis, and ad efficiency evaluation. The emergence of new technologies such as supercookies and flash cookies, which can be stored on a user's computer permanently, raise even more privacy concerns. Cookies can even be used in government tracking - the National Security Agency has been revealed to have the ability to pinpoint users based on identifiers that are contained in advertising and tracking cookies[18]. These identifiers are transmitted as plaintext between a user's browser and the servers that it communicates with, making cookie-harvesting attacks such as cache sniffing and XSS cookie sniffing possible[21].

There is not much that a user with a stock browser can do to prevent his or her identity from being revealed to the third-party content providers. One solution could be disabling cookies altogether; however, this method would also block first-party cookies, thus making many web-services unusable. In our research, we focus on anti-tracking browser extensions which are one of the simplest solutions to cookie-based privacy violations. We investigate several popular browser extensions and compare their effectiveness in blocking third-party cookies and tracking content. Finally, we analyze effectiveness of the Tor network in preventing tracking cookies.

3 Background and Related Work

Since its creation, OpenWPM has been widely used in several papers to automate the testing of web privacy tools and has proven to be adaptable and simple to use [14, 16, 22]. Researchers around the world have used it to track everything from cookies to various fingerprinting schemes. We will review some of these papers, and highlight both the similarities and differences to our research. We will also give an overview of the extensions we used in our testing, and provide a brief analysis of how they work.

3.1 Related Works

Instead of tracking clients and analyzing the effects of client tracking, ”Host FingerPrinting and Tracking on the Web”[22] uses OpenWPM to track web hosts over long periods of time to demonstrate the security and privacy implications that widespread host-tracking could bring. They spend some time explaining and analyzing tracking in the context of cookie-churn, which is related to one-time-only cookies which appear for certain websites and are supposed to promote online anonymity. They found that it is relatively easy to gather a robust profile on most hosts after a short period of tracking. They also give some tips on the client side to mitigate common tracking techniques.

”Online Tracking: A 1-million-site Measurement and Analysis”[15] is a research project undertaken by two of the original creators of OpenWPM. They study a number of different fingerprinting and tracking methods, and show how OpenWPM can be used to speed up this process, as well as to make the analysis of online tracking a more accessible undertaking for researchers. The goal of this paper was in part to illustrate how well OpenWPM hides the formidable engineering challenges that come with widespread analysis of the web. After giving valuable data on the effectiveness of not only tracking techniques but privacy add-ons that users employ, they conclude that OpenWPM has the potential to mitigate online privacy incursions, as well as bring the discussion about online privacy out of the lab and into the real world.

”The Web Never Forgets: Persistent Tracking Mechanisms in the Wild”[12] details the research done on three advanced web-tracking mechanisms - canvas figerprinting, evercookies, and the use of cookie syncing along with evercookies. Part of the struggle they faced was the automation of their data collection, as web privacy measurement was a young field at the time of this research (and still is at the time of this writing). As some of the researchers involved here are also cited on later OpenWPM papers, the automation process they developed here may have influenced the conception of OpenWPM. Their conclusion is that in the arms-race that is web privacy, the trackers currently have the upper hand. Evercookies and other fingerprinting techniques are incredibly and intentionally difficult for the everyday user to spot, and can be hard to defend against. They suggest that one goes beyond browser extensions and manually monitors their system regularly. They also recommend that “security focused” browsers should take a more aggressive stance when it comes to detecting and removing tracking techniques. Additionally, they recommend a series of back-end web crawls to detect unwanted trackers, and advise that website hosts take an in depth look at the trackers on their sites so that they have a complete knowledge of what data is being gathered and from whom.

”Cookies That Give You Away: The Surveillance Implications of Web Tracking”[17] is a research paper written by the creators of OpenWPM, which uses OpenWPM to track cookies and present a method in which widespread tracking can be accomplished using HTTP tracking cookies. They found that an adversary using proper techniques can reconstruct 62-73% of the web traffic of a given user. They also highlight how government agencies (both domestic and foreign) could (and sometimes do) use these techniques to conduct mass surveillance. The technique they describe consists of creating “clusters” of related cookies that essentially build a profile on a given user, and can provide insight into their habits and real-world identity. Throughout the paper they evaluate this attack and analyze its implications, using servers located inside and outside of the U.S. to conduct testing (this was to see the effect of tracking foreign users). Their advice to mitigate the risk of being tracked as they describe is to minimize the number of cookies accepted, to force HTTPS connections whenever possible, and to use different usernames between sites. They also recommend that cookie trackers only use unique identifiers if they are transmitted over HTTPS, as this will reduce piggybacking by other parties. The paper also mentions that the use of the Tor browser would likely defeat this attack, although Tor still has the potential to leak identity on popular websites (through the use of similar usernames across multiple sites).

3.2 Privacy Tools

Each privacy measure examined has a unique method for combating third-party tracking. However, an average user may also be interested in an unhindered browsing experience (many third-party cookies are helpful - their removal can lead to difficulties on certain websites) and the trustworthiness of the extension’s developers.

Ghostery works by analyzing each web server that is being called during a connection to a website and comparing them to known trackers. It will then inform the user of the tracker being deployed and, depending on configuration, potentially block the web server call from returning to the sender. Rather than explicitly blocking cookies, Ghostery will maintain web server communications but with a custom, more private response. Otherwise, it will simply disable communications with web servers[5].

Adblock Plus allows the user to enable a set of filters that block specified elements from loading in a website. Filters are entirely customizable and open source. By default, the extension uses filters called EasyList and Acceptable ads. Both are maintained by an open source community and will disable advertisements that are deemed to be invasive, as well as enable advertisements that are deemed to be acceptable. The criteria is mostly visual and requires the ad to be distinct yet not interrupt a website’s usability[1].

AdBlocker Ultimate is very similar to Adblock Plus in its use of filters, implementing one called English Filter by default rather than EasyList and Acceptable ads. However, it was created as a response to the business decisions of other adblockers, alluding heavily toward Adblock Plus. Adblock Plus allows consenting users to view whitelisted advertisements that are deemed to be unobtrusive[19]. This feature is profitable for Eyeo, the creators of Adblock Plus, because they earn a percentage of the ad revenue and charge fees to large advertisers[1]. AdBlocker Ultimate is against this practice and will not display any advertisements unless a website is explicitly whitelisted by the user or the extension is disabled. And rather than being owned by a for-profit company, AdBlocker Ultimate is nonprofit and open source[2]. It should be noted that this paper does not endorse one extension over the other, nor does it support a particular business model. But these business models are not insignificant when it comes to analyzing the cookies that are present in a browsing session.

Disconnect works by categorizing network requests (requests from anything other than the current site that a user is browsing) into several groups: Google, Facebook, Twitter, Advertising, Analytics, Social, and Content, and blocking anything outside of the Content categorization[4]. The inclusion of Content network requests is due to their usefulness in website behavior - if they are blocked, a website may become unusable or behave in an unintended way. Disconnect will not block first-party ads but will block third-party ads due to their potential to implement tracking. Tracking is defined by Disconnect as a behavior that both collects data and retains it[4]. Further, Disconnect does not block trackers that conform to Do Not Track standards defined by the Electronic Frontier Foundation, trackers that have an explicit opt-in policy, and trackers that are aimed at providing a non-malicious, improved user experience[4].

Tor uses onion routing, a distributed overlay network, to encrypt messages sent over the network in an attempt to preserve user anonymity[11]. Because blocking third-party tracking cookies is not an explicit goal of Tor, we were not expecting a great reduction in the number of tracking cookies during this trial. The results in Section 5 support this hypothesis, as the small reduction can be explained by an increased number of crashes compared to the other measurements.

4 Measurement Platform

When performing web privacy measurements, replicating an average user is important in order to simulate normal browsing behavior. This can be accomplished in one of two ways - automating the browsing process or utilizing crowdsourcing. Crowdsourcing requires incentives and privacy protection for the participants. Taking this into account, we sought to automate the browsing process. To date, there are numerous web privacy measurement platforms, each varying in implementations. Our selection process involved choosing a platform that had modular design combined with a high-level abstraction. In addition, the platform's implementation was taken into account for the sake of creating a user-like experience while collecting data from a browser's session. OpenWPM satisfies these requirements and was our web privacy measurement platform of choice.

4.1 OpenWPM

OpenWPM is an open source web privacy measurement platform created by researchers at Princeton University. To our knowledge, OpenWPM is the only platform to date that allows for stateful measurements. Stateful measurements are necessary in order to properly simulate a user profile. If stateless, the browser will always appear as a new user, resulting in unrealistic results[14]. Utilizing Selenium, OpenWPM supports consumer browsers such as Firefox, Chrome, and Internet Explorer. For the purposes of this paper, Firefox was main browser of choice. Using a lightweight browser does not provide support for new web technologies such as Adobe Flash, WebGL, and features provided by HTML5[14]. Without access to these technologies during testing, the results would be an inaccurate representation of typical modern web browsing. In addition, utilizing a consumer browser allows for the inclusion of API's. In turn, this has allowed us to measure the effectiveness of several existing privacy tools.

Selenium was originally created for developers to test their web applications. Due to its blocking API, Selenium alone does not support large scale testing. OpenWPM's integration and encapsulation of Selenium has resolved page timeouts, browser crashes, and other related issues during testing. This allows for the testing of large data sets with no data loss or corruption[14].

Lastly, OpenWPM utilizes a data aggregator which provides standardized results among repeated trials. Data is collected with an SQLite aggregator which collects data at the network level by utilizing Selenium's API. This information is in the form of HTTP requests and profile cookies. The aggregator tracks the changes in these cookies after each page visit during testing[14]. OpenWPM provides access to other content, such as JavaScript; however, for the purposes of this paper, an analysis of the profile cookies collected was necessary.

4.2 Measurement Configuration

Our tests were run on a Lenovo ThinkCentre M90. This desktop was equipped with an Intel Core i5-650 Processor 3.2GHz, 4GB of RAM, and Ubuntu 14.04. These configurations allowed for three browser instances. Each was setup to load the same URL and collect HTTP requests and profile cookies. The same URL was loaded in an attempt to prevent crashes, as in some instances a site would not load on one browser, yet load on another, due to a timeout. Each measurement utilized the top 5,000 URLs of the Alexa top 1 million site list (http://www.alexa.com). During testing, each browser instance was granted a 60 second timeout to fully load the URL's homepage. If the load was successful, a collection of the data occurs with no further interaction (e.g. scrolling, clicking internal links, logging in) with the site[14]. In the event of a timeout, the browser process is killed and restarted in order to continue with testing.

For testing purposes, we performed six stateful trials. The first trial was a control in which no privacy tools were utilized. A trial was then conducted for each of the privacy tools mentioned in Subsection 3.2. On average, each trial took 48-50 hours to complete.

5 Results

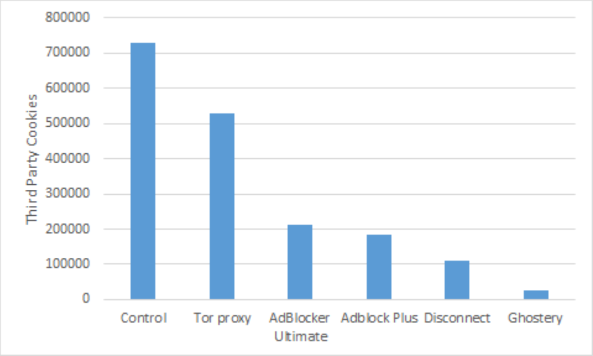

Figure 1: Third-party cookies gathered from the top 5,000 site measurement for each of the privacy tools as well as the control trial (Firefox without any extensions).

Figure 1 shows the number of cookies gathered by each of the measurements over the top 5,000 sites. Each site was visited in three different browsers in an attempt to reduce the number of sites which were unable to be accessed due to timeouts or failed requests. Based on this chart, Ghostery was the most pronounced in terms of the reduction in third-party cookies, reducing the number of third-party cookies seen in the control measurement by approximately 97%. Steven EngleHardt and Arvind Narayanan point out in their 1 million site analysis that measuring the total number of third-party cookies may be an ineffective method of analysis, as not all third-party cookies are in fact tracking cookies[14]. However, as we will see, tracking cookies make up a majority of the third-party cookies. Thus, a reduction as pronounced as that of Ghostery's will inevitably greatly reduce the number of tracking cookies.

5.1 Unsuccessful requests

Before making any inferences based on the reduction in third-party cookies due to the Tor proxy, it is necessary to take into account the number of failed GET requests during each trial. Failed GET requests are due to either a timeout while making a request to the server or when the response status received from the requested domain is not ”OK”.

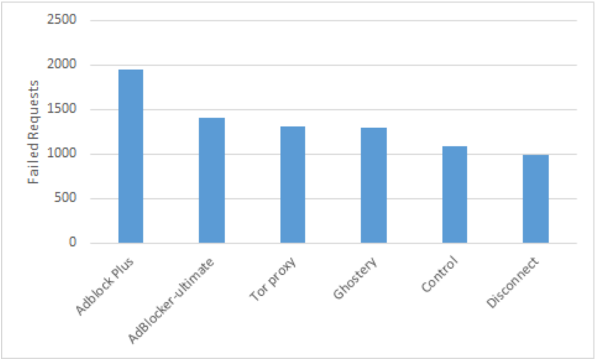

Figure 2: Failed GET requests during each measurement. There were a total of 15,000 GET requests during each (5,000 sites multiplied by 3 browsers).

Figure 2 shows that Adblock Plus had the greatest number of failed GET requests with 1,948 of the 15,000 requests made, failing. We suspect that this is due to the popularity of Adblock Plus (with over 19 million users on Firefox and over 14 times the number of users as Ghostery, the second most popular tool that was analyzed, Adblock Plus is by far the most popular tool for ad blocking), as sites have learned to detect the presence of Ablock Plus and sometimes request that a user disables it[7, 3]. While this metric is informative, it is also important to note how many sites were completely inaccessible during a trial, that is, how many sites a tool a measurement trial failed to access on all three browser instances.

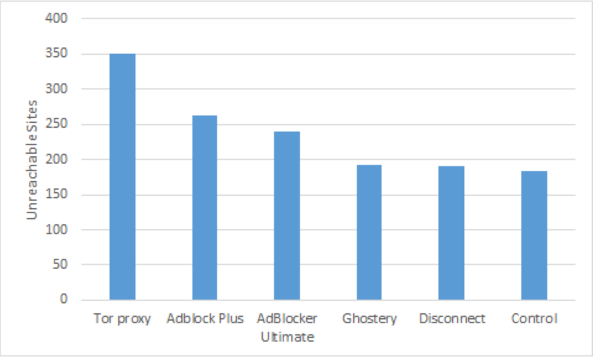

Figure 3: The number of unreachable sites during each measurement. An unreachable site is a site in the top 5,000 that could not be loaded due to a timeout or a failed GET request.

Figure 3 shows the number of sites which were inaccessible during a given trial. While Adblock Plus had the greatest number of failed GET requests among all of the privacy tools, the Tor proxy failed to access the greatest number of sites entirely; that is, none of the three browser instances were able to load the site. This could be due to the location of the exit node (websites can match ads based on geographic locations[6]) target or because of an increased number of timeouts as a result of the overhead to Onion Routing. It is also possible to identify Tor users using a public DNS-based list of Tor exit nodes and block connections based on this, but it is not clear whether this is occurring or not[9].

5.2 Did using a Tor proxy reduce the number of third-party cookies?

The answer isn’t clear. Regardless of why so many sites were inaccessible while using the Tor network, it is important to consider this while analyzing the results. Recall that Figure 1 showed the number of cookies gathered by each of the measurements over the top 5,000 sites. Use of a Tor proxy resulted in a comparatively small, but not insignificant reduction in the number of third-party cookies. With the potential caveats in the preceding paragraph accounted for, its possible that this reduction would disappear. If this experiment were repeated, the timeout duration could be increased in order to rule out timeouts due to the overhead of encryption. Recall, however, that blocking third-party cookies is not a goal of the Tor network, and running the same tests using the Tor Browser Bundle would be far more intriguing as it provides certain defenses against anonymity leaks in the application layer.

5.3 The most prevalent third-party domains

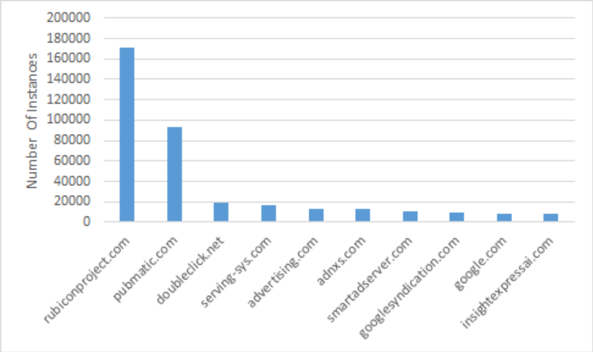

Figure 4: Third-party cookies from the the top 5,000 measurement for the control trial, each accessed through 3 different Firefox instances, categorized by the top ten domains.

The above chart shows the number of instances of each third-party domain without any privacy tools enabled. The most prevalent third-party domain was rubiconproject.com followed by pubmatic.com and doubleclick.net, respectively. If instead of looking at the total number of instances of each domain, we look at the percentage of first-party domains (one of the 5,000 visited), it turns out that rubiconproject.com no longer holds the top spot and pubmatic.com disappears from the top 10 entirely (See Figure 5). This suggests that while the most third-party cookies came from these two domains, they were present on a smaller subset of the top 5,000 websites. In their 1 million site measurement, Steven EngleHardt and Arvind Narayanan found that news sites had the largest number of third-party tracking cookies, which may lend some insight into this phenomena[14]; however, further research is needed to determine what sites these third-party domains appeared on.

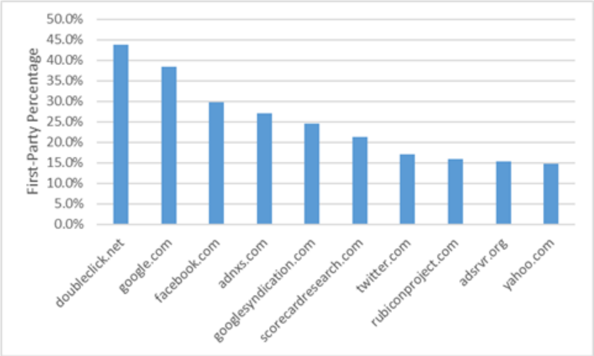

Figure 5: Percentage of the top 5,000 first-party domains on which a given third-party domain appears without any privacy tools enabled.

5.4 A closer look at Ghostery

Interestingly, Figure 3 showed that Ghostery was on par with Disconnect for the least number of inaccessible sites despite Ghostery also blocking the most third-party cookies. The main reason for Ghostery out-performing other privacy tools tested is likely due to the design of its API. Upon loading a URL, Ghostery will check the tags contained in the HTML file being requested for the classification of data collectors by comparing them to known trackers. If trackers are found, Ghostery will either modify its response or block all communication with the web server that the tracker originated from. In contrast, other privacy tools, such as Adblock Plus, AdBlocker Ultimate, and Disconnect, detect tracking instances by utilizing a blacklist. Once trackers are found, a modified response is sent with these web servers[1, 2, 4]. In addition, Adblock Plus is known to let non-intrusive ads through and Disconnect will not block third-party cookies that conform to Do Not Track standards, as mentioned in Subsection 3.2. With these design implementations in mind, the results for Ghostery are justified.

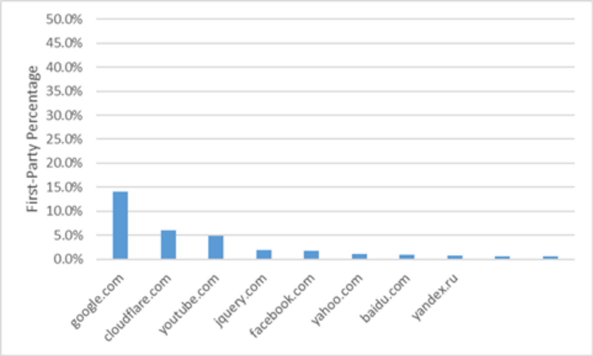

Figure 6: Percentage of first-party sites that a third-party domain appears on with Ghostery enabled.

What is perhaps most impressive about Figure 3 is that it shows Ghostery all but eliminated the most prominent third-party tracking cookies listed in "Online tracking: A 1-million-site Measurement and Analysis"[14]. The paper classifies third-party cookies as tracking cookies when they are present in the context of a fingerprinting JavaScript call (which are also defined in the paper). The third-party domain doubleclick.net was present in at least 25% of all first-party sites in all other simulations (See Appendix for charts from the other measurements). In this trial, however, it was blocked by Ghostery on the majority of visits. Further, not only was doubleclick.net present on the greatest number of first-parties (See Figure 5), it has also been identified as tracking in nearly all instances[14]. Other top 20 third-party tracking domains that were blocked significantly more often by Ghostery than any other privacy tools included facebook.com, googlesyndication.com, and twitter.com.

5.5 Adblock Plus versus AdBlocker Ultimate

Figure 1 showed the total number of third-party cookies accumulated during each measurement. Despite being created in response to Adblock Plus' business model (See Subsection 3.2), AdBlocker Ultimate appears to perform slightly worse than Adblock Plus.

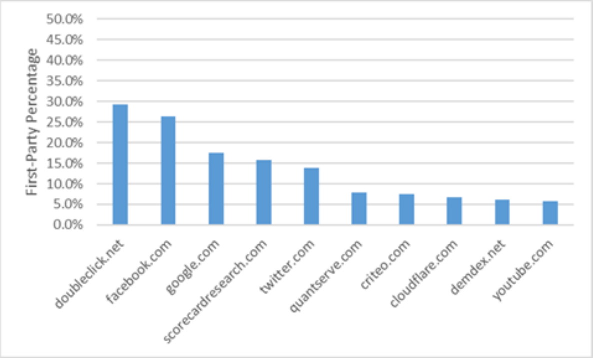

Figure 7: Percentage of first-party sites that a third-party domain appears on with Adblock Plus enabled.

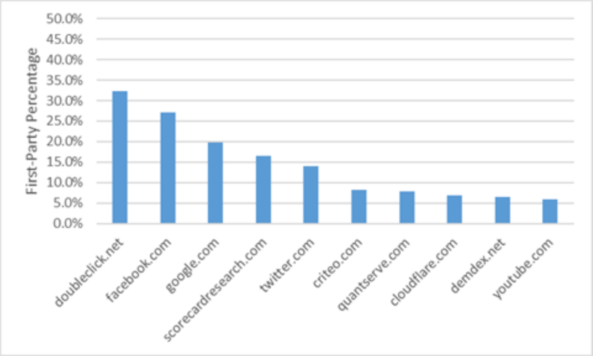

Figure 8: Percentage of first-party sites that a third-party domain appears on with AdBlocker Ultimate enabled.

Figure 7 shows the percentage of first parties that a given third-party domain appears on with Adblock Plus enabled; Figure 8 displays the same statistics, but with AdBlocker Ultimate enabled instead. One can see that Adblock Plus was more effective in terms of the prevalence of third-party domains across all 5,000 first-party domains, and in fact it blocked a greater number of third parties in total. Statistics on the number of instances of each third-party domain can be found in the appendix.

6 Conclusion and Future Work

For further research, it would be interesting to adapt OpenWPM to utilize a dynamic privacy extension such as Privacy Badger. The extensions examined in our paper have static filters and lists used to block predetermined cookies. Privacy Badger, on the other hand, examines gathered cookies while browsing the web for signs of tracking. If a cookie appears to be tracking, it will then be added to a blacklist.[8] Such a method could potentially gather cookies that the static filters and lists failed to include. However, a browser using only Privacy Badger would be open to tracking until the extension learns a filter. In order for OpenWPM to utilize this extension, changes would need to be made to the way OpenWPM handles crashes, as when a timeout occurs the browser is restarted and this resets PrivacyBadger's filter. Alternatively, a method of loading filters could be added to PrivacyBadger.

This experiment was performed on a relatively small sample size of 5,000. With more powerful equipment, a larger sample size could be used. Further, it would be interesting enable multiple extensions during a trial. Privacy tools such as Disconnect require specific implementations, such as being enabled first, but using a dynamic filter such as Privacy Badger alongside a static filter such as Adblock Plus could eliminate third-party trackers bypass the filter. It would be a way of combining the unique qualities of each extension and analyzing how beneficial, or detrimental, the effect would be.

Our analysis was unable to use context to classify each cookie as a tracker or non-tracker, as this requires the researcher to save JavaScript calls (something we did not do, but is supported by OpenWPM)[14]. Instead, we examine the top cookies seen in our extensions against the top cookies seen in the OpenWPM 1 million site measurement with regard to whether they are predominantly classified as trackers. While we believe this is effective, it would be beneficial if classifications were made more precise.

Overall, Ghostery performed the best out of all of the extensions. However, there is more to consider than simply the effectiveness of an extension when choosing a privacy tool. As mentioned in section 3.2, each extension has a different business model, and more than one has come under criticism. Ghostery, while effective, will send back anonymous data to the companies that own the third-party domains that users are aiming to avoid. Although this, like Adblock Plus's program, is opt-in[5].

This was not a comprehensive analysis. While we made an effort to examine some of the most common privacy tools, there are more available with varying degrees of effectiveness and complexity. Some of these options were mentioned before, such as Privacy Badger and Tor Browser, but an even more privacy-conscious user may opt for something more robust such as Tails OS[10].

Every one of the examined extensions successful in blocking third-party cookies to some degree, and many offer additional settings to fine-tune the private browsing experience. For example, Adblock Plus and AdBlocker Ultimate have the option of including more robust or specific filters on top of the defaults. The only tool that cannot be recommended for a role in blocking third-party cookies is the Tor proxy. However, this is expected, as that is not its goal. If a user wishes to further anonymize themselves with the Tor proxy it is necessary to use additional tools that provide application layer anonymity. The Tor Browser Bundle implements many of these features in an attempt to preserve anonymity while interacting with the Tor network, and thus it poses an opportunity for further research involving web privacy measurement over OpenWPM.

7 References

[1] 2016. Adblock Plus. (2016). `https://adblockplus.org/`

[2] 2016. AdBlocker Ultimate. (2016). `https://adblockultimate.net/`

[3] 2016. Detect Adblock. (2016). http://www.detectadblock.com/

[4] 2016. Disconnect. (2016). https://disconnect.me/

[5] 2016. Ghostery. (2016). https://www.ghostery.com/

[6] 2016. Google AdWords. (2016). https://adwords.google.com/home/how-it-works

[7] 2016. Mozilla Add-Ons. (2016). https://addons.mozilla.org/en-US/firefox/extensions/?sort=users

[8] 2016. Privacy Badger. (2016). https://www.eff.org/privacybadger

[9] 2016. The public TorDNSEL service. (2016). https://www.torproject.org/projects/tordnsel.html.en

[10] 2016. Tails. (2016). https://tails.boum.org/

[11] 2016. Tor. (2016). https://www.torproject.org/index.html.en

[12] Gunes Acar, Christian Eubank, Steven Englehardt, Marc Juarez, Arvind, and Claudia Diaz. 2014. The web Never Forgets: Persistent Tracking Mechanisms in the Wild. In *ACM Conference on Computer and Communications Security (CCS) 2014*.

[13] Linda Christiansen. 2011. Personal privacy and Internet marketing: An impossible conflict or a marriage made in heaven? *Business Horizons* 54 (2011), 509 – 514.

[14] Steven Englehardt and Arvind Narayanan. 2016a. Online tracking: A 1-million-site measurement and analysis. In *Proceedings of ACM CCS 2016*.

[15] Steven Englehardt and Arvind Narayanan. 2016b. Online tracking: A 1-million-site measurement and analysis. In *Proceedings of ACM CCS 2016*.

[16] Steven Englehardt, Dillon Reisman, Christian Eubank, Peter Zimmerman, Jonathan Mayer, Arvind Narayanan, and Edward W Felten. 2015a. Cookies that give you away: The surveillance implications of web tracking. In *Proceedings of the 24th International Conference on World Wide Web*. ACM, 289–299.

[17] Steven Englehardt, Dillon Reisman, Christian Eubank, Peter Zimmerman, Jonathan Mayer, Arvind Narayanan, and Edward W. Felten. 2015b. Cookes That Give You Away: The Surveillance Implications of Web Tracking. In *World Wide Web Conference 2015*.

[18] Andrew Hilts and Christopher Parsons. 2015. *Half Baked: The Opportunity to Secure Cookie-Based Identifiers from Passive Surveillance.* Technical Report. University of Toronto.

[19] Rich Mullikin. 2016. *Adblock Plus Enters Ad-Tech With Launch of SSP/Ad Platform.* Technical Report. Business Wire: A Berkshire Hathaway Company.

[20] Kurt Schiller. 2010. *Companies Reacting to Consumers' Views on Targeted Ads.* Technical Report. EContent.

[21] Rodica Tirlea, Claude Castelluccia, and Demosthenes Ikonomou. 2011. Bittersweet cookies. Some security and privacy considerations. European Network and Information Security Agency. (2011).

[22] Ting-Fang Yen, Yinglian Xie, Fang Yu, Roger Peng Yu, and Abadi Martin. 2012. Host Fingerprinting and Tracking on the Web:Privacy and Security Implications. In *The 19th Annual Network and Distributed System Security Symposium (NDSS) 2012.*

8 Appendix

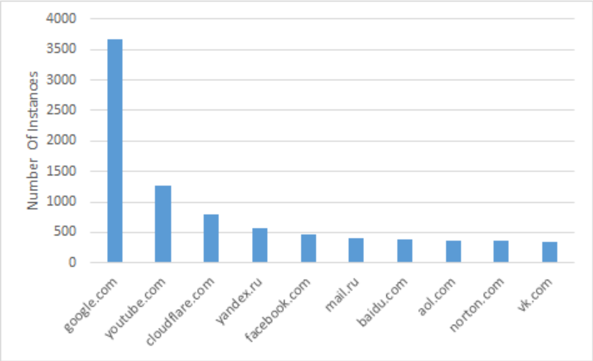

Figure 9: Third-party cookies from the the top 5,000 measurement for the Ghostery trial, each accessed through 3 different Firefox instances, categorized by the top ten domains.

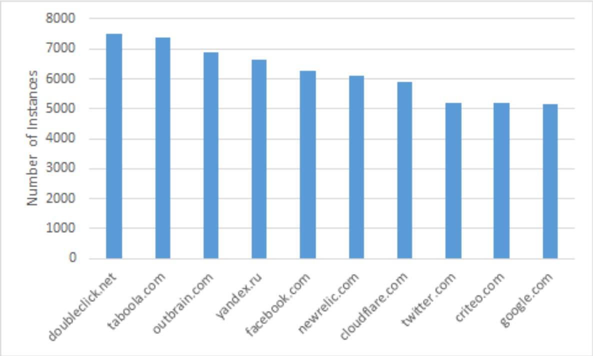

Figure 10: Third-party cookies from the the top 5,000 measurement for the Adblock Plus trial, each accessed through 3 different Firefox instances, categorized by the top ten domains.

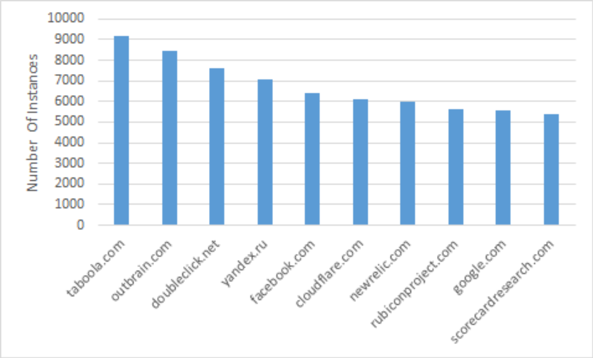

Figure 11: Third-party cookies from the the top 5,000 measurement for the AdBlocker Ultimate trial, each accessed through 3 different Firefox instances, categorized by the top ten domains.

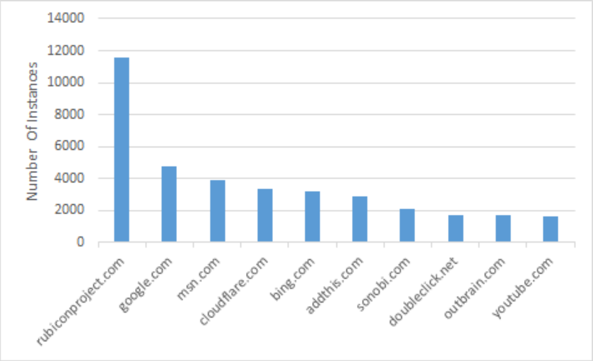

Figure 12: Third-party cookies from the the top 5,000 measurement for the Disconnect trial, each accessed through 3 different Firefox instances, categorized by the top ten domains.

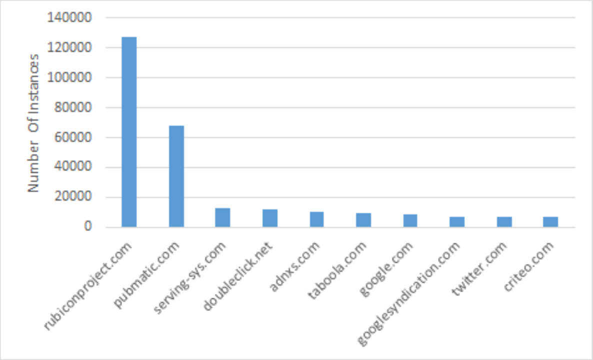

Figure 13: Total third-party cookies from the the top 5,000 measurement for Tor proxy categorized by the top ten domains

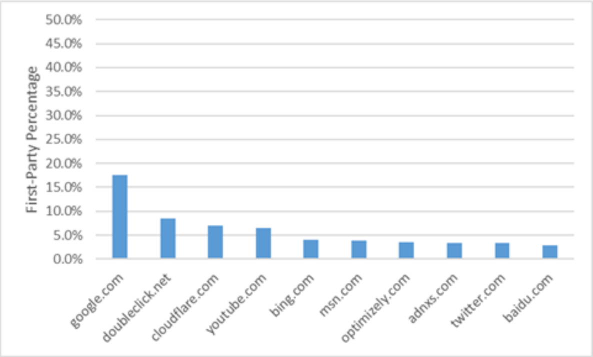

Figure 14: Percentage of first-party sites that a third-party domain appears on with Disconnect enabled.

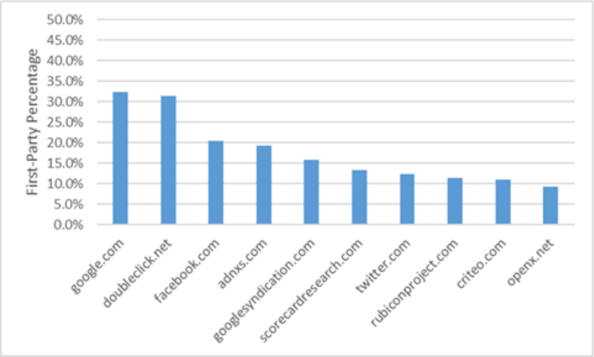

Figure 15: Percentage of first-party sites that a third-party domain appears on with Tor proxy enabled.