The Center for Internet and Society at Stanford Law School is a leader in the study of the law and policy around the Internet and other emerging technologies.

TRACKING THE TRACKERS: SELF-HELP TOOLS

By **Jonathan Mayer** on September 13, 2011 at 4:35 am

A number of technologies have been touted to offer consumers control over third-party web tracking. This post reviews the tools that are available and presents empirical evidence on their effectiveness. Here are the key takeaways:

1. Most desktop browsers currently do not support effective self-help tools. Mobile users are almost completely out of luck.
2. Self-help tools vary substantially in performance.
3. The most effective self-help tools block third-party advertising.

Following the usage model in the FTC staff's 2010 preliminary online privacy report, this post is oriented towards the user who wants a simple, persistent, comprehensive solution such that with high confidence no third party collects her browsing history. We assume that some third-party trackers will use non-cookie tracking methods including supercookies and fingerprinting (e.g. Microsoft, KISSmetrics, Epic Marketplace, BlueCava, Interclick, Quantcast).

Thanks to Jovanni Hernandez and Akshay Jagadeesh for assisting with data collection, and to Arvind Narayanan and Peter Eckersley for input on drafts.
{C}

Self-Regulatory "Opt-Out" Cookies

For over a decade a minority of third-party trackers, most prominently the members of the self-regulatory Network Advertising Initiative (NAI) and Digital Advertising Alliance (DAA), have offered users the ability to set an "opt-out" cookie.

As a technological matter, cookies are a poor mechanism for storing persistent and comprehensive user preferences. Users often delete their cookies, wiping out "opt outs." Cookies can expire (e.g. Chitika's "opt-out" cookie). And because each "opt-out" cookie is scoped to specific domain, a user has to periodically install cookies for companies that have newly begun to offer an "opt out." The user experience for setting "opt-out" cookies is unnecessarily arduous (see the DAA "opt-out" page), and in some cases "opt-out" cookies aren't even set correctly. That said, there are several browser extensions that significantly improve the usability and persistence of "opt-out" cookies (e.g. Abine Taco, Beef Taco, Google Keep My Opt Outs, NAI Consumer Opt Out Protector, and PrivacyChoice Keep More Opt Outs).

You may have noticed that "opt out" is scrupulously placed in quotes throughout this discussion. That's because, setting aside technical issues, **"opt-out" cookies don't actually opt users out of tracking**. As we explained in an earlier post, "opt-out" cookies only opt users out of seeing ads based on tracking—not tracking itself. And as we showed in a later post on the DAA's self-regulatory icon initiative, both the NAI and DAA use slippery, deceptive language in describing their "opt-out" programs.[1]

Do Not Track

Do Not Track uses an HTTP header to signal a user's preference to opt out of third-party tracking. Browsers have been quick to adopt the proposal, user adoption is skyrocketing (1, 2), and tools are under development for detecting violations. But, for the moment, most tracking companies steadfastly refuse to comply. We believe Do Not Track is the right way to provide consumer choice on third-party tracking (learn more at DoNotTrack.Us), and we recommend users enable the feature to send a signal to regulators, legislators, and tracking companies. While we are pleased with Do Not Track's progress, convincing stakeholders to adopt the proposal is a lengthy process. In the interim, users must look elsewhere for effective protection against third-party tracking.

Browser Profile Clearing

Users are often advised to regularly clear their cookies, cache, history, and other browser profile settings to prevent third-party tracking. There are several reasons this approach does not adequately protect users.

First, many third-party tracking methods continue to work. Tracking techniques that do not require storing state in the browser are wholly unaffected. As for stateful tracking, the user must play Whac-A-Mole with third parties. To remove ETag cookies, the user must clear the browser's cache. To remove Flash cookies, she has to independently clear her Flash plugin data. In short: the user has to scrub *anyplace* the browser or a plugin can store state.

Second, clearing the browser profile only provides periodic protection. In the intervals between when a user clears his settings, every tracking method works.

Third, clearing the browser profile undermines beneficial functionality. Many of the lost features result in significant annoyances (e.g. stored logins and browsing history). Some even introduce security vulnerabilities (e.g. stored authentication tokens and HTTP Strict Transport Security).

Last, as a practical matter, clearing the browser profile is an unworkable solution. The average user cannot, and should not, reasonably be expected to diligently vacuum her browser on a monthly basis—let alone every week or every day.[2]

Private Browsing Mode

While implementation specifics vary by browser, private browsing modes share a common goal: eliminate evidence of browsing that resides on the computer. To a first approximation private browsing modes function the same as clearing the browser profile, except the user proactively declares a session to be private (automatically clearing profile changes when the session ends) instead of retroactively clearing the profile. Private browsing has the very same shortcomings as clearing the browser profile: it does not stop all tracking methods, it provides only periodic protection (the user can be tracked within a private browsing session), beneficial web functionality breaks, and as a practical matter a user will not adjust a setting every time his browser launches.

Third-Party Cookie Blocking

All the major web browsers include an option to prevent third-party domains from setting cookies. Because cookies are just one of many ways third parties track users, third-party cookie blocking provides limited protection. And unless a browser blocks third-party cookies from being read,[3] clicking a tracker's ad or visiting a tracker's website (e.g. Facebook or Google) once is enough to set an indefinite tracking cookie.[4]

Targeted Cookie Blocking

Internet Explorer, Firefox, Chrome, and several browser extensions offer the ability to prevent cookies from certain domains from being read or set. Just like third-party cookie blocking, this approach does not mitigate non-cookie tracking technologies. It also largely eliminates interactive functionality on websites that are both a first party and a third-party tracker (e.g. Facebook or Google).

Execution Blocking

A number of tools are available for preventing the execution of JavaScript (e.g. NoScript), Flash (e.g. Flashblock), and other script content that could be used for tracking. While there are many other reasons to use these tools (including security, speed, and power consumption), they only mitigate a subset of tracking mechanisms.

Content Blocking

[Updated 9/14 to include a note on Request Policy. Thanks to Joe Hall for the suggestion.]

Because of the myriad methods for tracking, many privacy tools focus on preventing the browser from even requesting certain third-party content. While content blocking can effectively prevent third-party tracking, a content blocking tool is only as effective as its list of rules on what to block (often called a "blocklist"). Most content blocking tools consist of nothing more than a regularly updated blocklist (or family of blocklists), in either Adblock Plus or Tracking Protection List format. Request Policy, a Firefox extension, takes the opposite approach: all requests to third-party domains are blocked, save those the user explicitly allows. While Request Policy offers nearly comprehensive protection from third-party tracking, properly configuring it requires substantially greater patience and expertise than the average user can reasonably be expected to possess.

Please note: Chrome, Safari, Mobile Safari, and the Android browser DO NOT presently support content blocking.[5] Firefox extensions are able to block content, and users can install blocklists in Internet Explorer 9.

Effectiveness Measurement

We conducted a study of the effectiveness of twelve web privacy tools at mitigating third-party web tracking. **Please note: several of the blocklists we studied in Adblock Plus format are also available in the less expressive Tracking Protection List format. The change in formats may impact performance.**

> Abine's Tracking Protection List blocks many online advertising and marketing technologies that can track and profile you as you browse the Web. This list is updated weekly to keep you safer and more private.
>
> EasyList is the primary subscription that removes adverts from English webpages, including unwanted frames, images and objects. It is the most popular list for Adblock Plus, with over 7 million daily users, and forms the basis of over a dozen combination and supplementary subscriptions.
>
> EasyPrivacy is an optional supplementary subscription that completely removes all forms of tracking from the internet, including web bugs, tracking scripts and information collectors, thereby protecting your personal data.
>
> Ghostery allows you to block scripts from companies that you don't trust, delete local shared objects, and even block images and iframes. Ghostery puts your web privacy back in your hands.
>
> PrivacyChoice maintains a comprehensive database of tracking companies, including domains used by nearly 300 ad networks and platforms, tracking methods, summaries of key policies, oversight, and opt-out and opt-in processes. PrivacyChoice has created Tracking Protection Lists based on this data. You have the option of installing two lists. The first list blocks companies that are not subject to oversight by the NAI and the second list blocks all tracking company domains in the PrivacyChoice database. These lists will be automatically updated with new tracking domains discovered through continuous website scanning and user panels.
>
> Complete control over online tracking using multiple methods, including cookie blocking, persistent opt-out cookies, Flash and HTML5 control, and Do Not Track signals.
>
> TRUSTe is the leading online privacy certification and services provider. TRUSTe's TRUSTed Tracking Protection List enables relevant and targeted ads from companies that demonstrate

respectful consumer privacy practices and comply with TRUSTe's high standards and direct oversight. TRUSTe helps users get good ads, without compromising personal privacy.

- Abine Tracking Protection List

  In our initial testing, the Abine list performed very poorly; manually inspecting the list we identified several typos. We called our findings to Abine's attention, and the company responded with an updated list. We present below our findings on both the original and updated Abine lists.

- Adversity Ads + Privacy + Antisocial Adblock Plus lists
- EasyList Adblock Plus list
- EasyPrivacy Adblock Plus list
- EasyList + EasyPrivacy Adblock Plus lists
- Fanboy's List Ads + Tracking + Annoyance Adblock Plus lists
- Ghostery browser extension (configured to block all trackers, "experimental" cookie blocking not enabled)
- PrivacyChoice 1 Tracking Protection List
- PrivacyChoice 2 Tracking Protection List
- PrivacyChoice TrackerBlock browser extension (configured to block all trackers, opt-out cookies not enabled)
- TRUSTe Tracking Protection List

Effectiveness Measurement - Methodology

For each blocking tool we conducted a crawl of the Alexa U.S. top five hundred websites using the FourthParty web measurement platform. To ensure broad coverage of third parties we crawled the list three times in series, and to provide fresh browser state for each page load we cycled private browsing mode off and on. We also conducted a baseline crawl for comparison. Our crawl data is available on request.

We compiled three measurements with each blocking tool:

HTTP Requests. The number of crawled pages on which each domain (public suffix + 1) receives at least one HTTP request. Almost all third-party web content is served using HTTP, so there likely few if any false negatives. But this measurement includes false positives: some resources are served from a third party that does not track. For example, the Google Libraries API (googleapis.com) serves static content and instructs the browser to cache it for a year.

HTTP Set-Cookie Responses. The number of crawled pages on which each domain (public suffix + 1) sends at least one HTTP response that includes a Set-Cookie header. This metric has some false negatives since it includes neither trackers that do not set cookies over HTTP nor trackers that set their cookies in a first-party context (e.g. Twitter). There are few false positives since in almost all cases cases if a web service wants to preserve state across multiple sites it will just use a unique identifier.

Cookies Added - Cookies Deleted. The number of cookies added less the number of cookies deleted by each domain (fully qualified domain name). Measuring the difference between cookies added and deleted neglects trackers that do not use cookies or set cookies only as a first party, and is overinclusive of first-party sites that set a large number of cookies. Scripts can behave erratically when a browser blocks content, introducing significant noise into this measurement. We include it as a rough benchmark for cookie blocking tools.

Effectiveness Measurement - Results

The following graph reports the average across all tracking domains of the relative difference in each measurement.[6] We encourage interested users to examine the complete spreadsheet of measurements.

Some observations from inspecting the tools and analyzing the crawl data:

- Self-help tools vary significantly in their effectiveness. Some (especially the Tracking Protection Lists from Abine before 9/6 and from TRUSTe) offer very little protection.[7] No tool is comprehensive.
- Installing multiple self-help tools can *decrease* user privacy. Despite receiving widespread negative press for the practice, TRUSTe continues to serve a Tracking Protection List that overrides other lists to *allow* tracking by BlueKai, comScore, Scorecard Research, and others.
- Some websites depend on the presence of certain third-party scripts (e.g. the Google Analytics ga.js or urchin.js). Ghostery cleverly circumvents this issue by replacing several popular scripts with dummy stand-ins. (See also NoScript surrogate scripts.) It may be worthwhile to add support for dummy scripts to blocklist formats.
- The top performers (EasyList + EasyPrivacy and Fanboy's List Ads + Tracking + Annoyance) are community-maintained blocklists.
- Both top performers require installing more than one blocklist.
- All the top and near-top performers (EasyList + EasyPrivacy, Fanboy's List Ads + Tracking + Annoyance, Ghostery, and Adversity Ads + Privacy + Antisocial) block third-party advertising. This result should come as little surprise: third-party tracking is often inextricably commingled with third-party advertising.
- Most self-help tools do a poor job of blocking social plugins, even from the most popular social networks and sharing platforms.

Policy Implications

In the debates surrounding online privacy, many tracking companies have assumed that if they can hold out against Do Not Track, their business practices will continue. That's not necessarily the case. Some users will turn to the next-best alternative, and we now know what that is: ad blocking. Internet Explorer 9 already supports ad blocking with two clicks. Representatives from Mozilla have repeatedly delivered the ultimatum that if effective regulation or self-regulation does not occur, Firefox will provide users with self-help tools. W3C is working to standardize a blocklist format. The extent to which users adopt ad blocking will, of course, depend on usability, advocacy, and much more. But it likely won't take much persuading: users dislike advertising, and ad blockers are already the most popular extensions for Firefox, Chrome, and Safari. Third parties should not be so hasty to play Russian roulette with the Internet economy. And publishers should not be so willing to let them.

---

[1] Sometimes even the NAI and DAA member companies misunderstand what the self-regulatory programs require. Here are two examples from Google's Keep My Opt Outs tool (1, 2):

> Today we're making available Keep My Opt-Outs, which enables you to opt out permanently from ad tracking cookies.
>
> **Will this persistently opt me out of every cookie on the web?**
> No, this will not opt you out of cookies that are not related to personalized online ads.

[2] Some browsers offer options for clearing components of the browser profile on exit. These options may somewhat mitigate usability issues with regularly cleaning the profile.

[3] Third-party cookie blocking in Internet Explorer, Chrome, and Safari only prevents cookies from being set, not read. Chrome does provide a separate "experimental" option in about:flags that prevents third parties from reading cookies. Firefox's third-party cookie blocking prevents both setting and reading cookies.

[4] There may also be trivial ways to circumvent third-party cookie blocking. In Safari, for example, a redirect through a domain or a POST to a domain will allow setting cookies.

[5] The blocking API in WebKit (used in Chrome and Safari extensions) has a number of shortcomings. First, it doesn't prevent network requests, just loading content into the DOM. Second, the API doesn't allow blocking for all HTTP requests. Last, to support even modestly comprehensive blocking without a significant performance impact, it requires a synchronous message passing feature that Chrome lacks. A more

comprehensive blocking API for Chrome is currently under development with no set release date. A previous effort towards a Chrome blocklist feature was cancelled after six months of development.

Android users can block third-party web (but not app) content by running Firefox with Adblock Plus.

[6] We treated a domain as a third-party tracking domain if its metric value was greater than six in the baseline crawl. In other words, we considered a domain to be a third-party tracker if it, to a first approximation, consistently appeared on more than two sites. We found our results quite robust against changing the threshold value for considering a domain a third-party tracker.

To conserve space, the graph above does not show values below zero.

[7] We did not conduct a crawl with the Enhanced Privacy Tracking Protection List, though a cursory inspection of the list revealed that it blocks very few third-party trackers.

Focus Areas: Privacy

Related Topics: Do Not Track

Comments

Dan February 17, 2012 at 2:28 pm permalink

Yes, we must do all that we can to protect the consumer. It has been shown that too many companies are in it for the money and apparently do not care much for the consumer and see them only as an object of profit. I think a lot falls on the consumer to demand more from their companies and have more products made in the States as well as not tracking the consumer. In addition, we need to take care in protecting ourselves because remember as a consumer you must learn and be responsible for your own actions. I simply prefer my privacy over get this because it is free and in turn you support the advertiser. I for one would even prefer a Facebook that had no ads and a minimally subscription price per year to help support itself.

reply

Dimitris October 10, 2011 at 5:51 pm permalink

I am using the lists for AdBlock that you mention, EasyList and Fanboy's lists, Ghostery and Priv3.
Is that too much?
Am I doing any good or is it meaningless to have more than one tool?
Also, if at some website AdBlock is blocking some element I want to see and I disable it, does that allow tracking cookies to pass or do the other tools "catch" them?
Thanks for any help!
D

reply

Adam from Ghostery September 26, 2011 at 12:27 pm permalink

Greetings. Do you work for Mozilla? We're in regular contact with them and I wasn't aware that anyone who worked with/for them had this opinion.
Ghostery is anything but spyware. Better Advertising is now named Evidon, and it operates as a neutral privacy technology and services company. We do not serve ads, and any data that is collected through Ghostery is done so on a completely volunteer basis through our opt-in, anonymous GhostRank panel.
If you (or anyone else) would like more information, we've written extensively on this: http://mygho.st/D

reply

The guy from Mozilla September 14, 2011 at 4:24 pm permalink

Did you ever notice that Ghostery belongs to Better Advertising company?
http://forums.mozillazine.org/viewtopic.php?f=7&t=1905935

reply

Add new comment

E-mail

Your name

The content of this field is kept private and will not be shown publicly.

Comment *

Disable rich-text

Notify me when new comments are posted

Once you hit Save, your comment will be held for moderation before being published. You will not see a confirmation message once you hit the Save button but please be assured your comment has been submitted and we will review it.

What code is in the image? *

Image CAPTCHAEnter the characters shown in the image.

SAVE PREVIEW