https://youbroketheinternet.org/trackedanyway

So We Got Tracked Anyway

Did you install EFF's brilliant Privacy Badger or any other smart HTTP Cookie management tool? Or did you simply pick the privacy preference in your browser that *ignores all third-party cookies?* Did many websites you visit annoy you with permission-to-use-cookies pop-ups because of European legislation?

Guess what, it's all been useless.

Hamburg university researchers have examined closely how web browsers implement so-called TLS session resumption and how the top million popular websites make use of that feature. They found that 80% of websites make a correct use, unsuitable for tracking repeat visitors — just resuming an existing session within the last ten minutes.

Unfortunately though, Google is present on 80% of these websites in form of Analytics, Fonts or other third-party inclusions. And among 10% of sites that do not respect reasonable resumption times, Google sticks out as one of the most greedy ones — it allows for a web browser to stay offline for over a day, and still be recognized as the same web browser the next day. Considering that it is nearly impossible to surf the web without accessing some Google content, this means that Google can track all your surfing habits without any need for HTTP Cookies!

As Facebook isn't as pervasively present in all of the web, it went even further. It is enough for you to visit any website bearing a Like button every second day to allow Facebook to profile you, even if you never dreamt of logging into that service. Could it be our researchers just caught these companies with their hands deep in the cookie jar (pun intended)? For how long have they been collecting user data this way?

TLS stands for 'Transport Layer Security'. It is the protocol standard formerly known as SSL that powers the encryption in HTTPS. With the post-Snowden initiative of encrypting all web traffic, we inadvertedly introduced a new method of bulk surveillance.

The problem of TLS session tracking isn't news, actually. Back in 2010 fippo aka Philipp Hancke wrote a proof-of-concept implementation of such a tracking mechanism while he was refining the TLS implementation in psyced. We discussed the problem in the chatroom, but failed to make it public, thinking it was obvious. And back then only high security applications like banking and shopping were using HTTPS. Stuff you log into with all your data, anyway. Using encryption just to protect the privacy of regular websurfing was considered paranoid.

Eight years later, everyone in the business seems keen to point out that TLS version 1.3 will finally address this issue by encrypting session data, but that

would only protect us from passive observers in the network — it doesn't help if the server itself is trying to figure out who we are.

Only those of us who have systematically used Torbrowser or blocked Faceboogle domains on their firewalls or routers are exempted from yet another privacy failure. And those who happen to shutdown their computers completely each day, or otherwise maintain a habit of restarting their browser each day anew.

Even Tor users that have been surfing the web using any browser but the appropriate Torbrowser must now face the evidence that Google may not know where they were physically located, but it was granted data about most of the websites visited, possibly over the course of years. A scoundrel who thinks they would actually collect that data and use it according to their business model (like they have done before).

This time the breach is particularly painful, because it is affecting those people who thought they had taken measures against it. How many more times will we try our luck with band aids and hotfixes rather than demand an Internet that just cannot spy on us by design? We should demand this from technologists, but most of all from politicians. And we should go to the streets bearing banners. This is no small thing. We are losing our democracies.

Consider also how the Snowden revelations informed us that the Google cookie was NSA's way to identify targets, about a decade ago. Since Cookie filtering became commonplace, the NSA certainly needed new ways of identifying people on the net. To comply with the Freedom Act, the NSA still has PRISM running and companies such as Facebook and Google are simply obliged to collaborate. We may have no proof, but we know the current state of legislation. NSA just wouldn't be doing its job, if it wasn't exploiting this loophole.

Thanks to the Hamburg researchers for systematically looking into this problem and exposing the corporations that silently may have been poking fun at our anti-Cookie legislations and protection tools. If this breach is now uncovered, it only means that there are dozens more this broken Internet is capable of, that we haven't become aware of. Expect to be tracked anyway, in ways you never conceived of. Don't let this carry on. Speak up. Do something.

—lynX.

Appendix

Did you turn on that 'Block dangerous and deceptive content' feature in Firefox' security tab? It works by downloading a database from Google when you start the browser. Occasionally a website you visit can appear like it is contained in that database, so your browser will ask Google again to make sure the website isn't in there. In practice, your browser may consult Google fairly often behind your back. The issue was raised in 2006, that 'Safe Browsing' shouldn't send the regular Google "PREF" Cookie as if the user had intentionally wanted to visit Google, but odd argumentations made sure that the bug never got fixed,

possibly related to NSA's mission to keep entire humanity in check and depending on "PREF" for that purpose. Unfortunately even GDPR doesn't impede this business, as the user "intentionally" installed Firefox by "free will", as if they knew what they are getting themselves into. This whole ideology of people making choices while technology does whatever it wants behind their backs, as if the precondition of transparency and understanding could ever be met, is utter madness making dystopias come true. Thanks Ashkan, lixtra and others.

In 1996 I was worried about HTTP ETags becoming a tool for tracking, They still are today. And legislators are once again completely unaware of them. Thanks, Chris Morgan.

HTTPS session identifiers can be disabled in Mozilla products manually by setting '**security.ssl.disable_session_identifiers**' in about:config. Thanks, gruez. Torbrowser simply has this and dozens of other settings configured to put privacy and security first.

Thanks toast0 for details on TLS 1.3.

Also breaking news: activists and local population have requisited the Google Campus in Berlin. No wait, police already intervened. Google and Mastercard Cut a Secret Ad Deal to Track Retail Sales. Google found the perfect way to link online ads to store purchases: credit card data. "People don't expect what they buy physically in a store to be linked to what they are buying online."

Questions & Answers

"Those session resumption tokens save you redownloading 1- 10kb of certificates every fresh connection and the multiple round trips for the TLS handshake. Its a bandwidth and latency optimization." says tlarkworthy.

So should we always trade in our privacy and civil rights for performance optimizations? And did anyone even ask us? Is it enough if some of us opt out, if in the end result most of humanity becomes individually predictable and easily subject of micro-invasive mass manipulation? Understand how these performance optimizations nibble at the foundations of what's left of democracy?

_jal suggests, "This is why you block the surveillance shops' IP space. They are professional panty-sniffers, dependent on doing so for their dinner. Like trolls and narcissists, your only value to them is instrumental. Assuming you care about these things, communicating with them in any way is against your interests. The internet is a much nicer, safer place when you blackhole the commercial-Stasi-wannabes."

Would be nifty to have ready to use tools and configs for average people to block all webservers that abuse HTTPS sessions for surveillance. Oh wait, the folks from 'Occupy Google' have a list of recommendations. Then again, it doesn't help humanity if only a bunch of aware intellectuals opt out.

stubish says, "Google needs to be perceived as the best place to spend your advertising budget. What business case is there for Google to work hard at tracking people, defeating anti-tracking measures?"

Faceboogle have become the best places to spend your advertising budget *because* despite the high price they can target just the people you need. And most of all you can't afford not to do this type of unethical surveillance targeting, because all your competitors are doing it. This creates a massive market distortion in which Faceboogle are the two mind-reading monopolists. Since this approach is entirely unethical and of no value for society, it would be reasonable to bring back equal conditions for advertising by making mass surveillance impossible again, as it was before 1995.

yuhong asks, "I'm under the impression Firefox intends to do DNS over HTTPS with Cloudflare shortly. If X and Y are Cloudflare fronted domains, can they now pair sessions? I'm guessing a DOH session queries for domain X and immediately an HTTPS connection appears for X, then queries for Y and another appears at Y. Then the whole DOH session becomes identifiable once Cloudflare fronts any service with email sign-in."

Sounds very realistic to me and worth a closer examination, maybe in form of an upcoming university study?

More trivia on the issue of so-called 'Safe Browsing':

manicdee illustrates: "You issue quest on hash A then hash B. Google guesses that because of other activity it has seen today, visits to a site marked by hash A followed by visits to site marked hash B means you are following a link from Alex Jones' blog to a flat earth holocaust denial web site, and thus prepares to serve your IP address ads for tin foil hats and prepper magazines. The chances of your traffic pattern of hash A then hash B colliding with, say, my browsing of the MLP fan club and following a link to cosplay photos from Dragon Con are pretty slim, even though the MLP fan club URL hash collided with the Alex Jones blog hash. Google aren't just looking at the one thing you viewed, they are following you everywhere."

How likely is it, that visiting two partitions of the worldwide web subsequently is giving clues on which specific websites you probably were visiting? And even if in a certain percentage of cases this assertion is wrong, does it matter commercially? Or is a certain percentage of algorithmic failure irrelevant to Google, just not irrelevant to you? Will it have political consequences in your life, if you are stored as an Alex Jones visitor when you are actually just an MLP fan? Sounds like another fine case for a scientific research paper to be written — to find out if 'Safe Browsing' is a threat to liberty even if cookies and TLS session identificators were eliminated.

Last Change: 2021-06-30

Further discussion on Ycombinator (We are not endorsing that site, it's just that the *attention economy* chooses its own data agglomeration centers).

No advertising, no tracking, no profiling, no data mining. Viewable as ybti.psyciumunsqarzsehz5xlgsi2mg4dkvntwf5bwj5kwbcbazwiuhna2ad.onion as much as youbroketheinternet.org or via freenet.