Facebook Tracking Through Social Plug-ins

Technical report prepared for the Belgian Privacy Commission

Güneş Acar[1], Brendan Van Alsenoy[2], Frank Piessens[3], Claudia Diaz[1], Bart Preneel[1]

24 June 2015

Version 1.1

Outline

[1] COSIC, KU Leuven , iMinds
[2] ICRI/CIR KU Leuven, iMinds
[3] DistriNet, KU Leuven , iMinds

1 Introduction

This report provides a technical description of Facebook's online tracking capabilities enabled by its social plug-ins[4]. Social plug-ins are extremely popular, as website owners increase their audience if individuals share their content through online social networks. Facebook's Like Button, the most popular Facebook social plug-in, is present on 32% of the top 10.000 sites[5], covering almost all website categories including health and government websites[6].

The near-ubiquity of the social plug-ins also makes them the ideal tool for collecting the browsing activities of Web users, also known as tracking[7]. For the purposes of this report, “tracking” is defined as the collection of users' web browsing activities across different websites. The type of tracking facilitated by Facebook social plug-ins is commonly referred to as "third-party tracking", because the tracker (e.g. Facebook) is a different party from the (first-party) website visited by the user, as displayed in the user's browser address bar.

The way social plug-ins are commonly implemented forces the user's browser to fetch content (e.g., images or scripts) from social network servers, exposing information about user's visits to the social network operator. It is worth noting that Facebook is in a unique position, as it can easily link the browsing behavior of its users to their real world identities[8], social network interactions, offline purchases, and highly sensitive data such as medical information, religion, and sexual and political preferences. This renders the privacy implications of Facebook's tracking more invasive than any other third-party tracking setting, where, for example, advertisers or analytics companies may not have direct access to visitors' real world identities.

2 Scope

This report is limited to the analysis of cookie-based tracking enabled by the Facebook social plug-ins. Websites may also use “cookie-less” tracking mechanisms such as browser fingerprinting[9], Flash

[4] Facebook social plug-ins include Like Button, Share Button, Embedded Posts, Comments, Send Button, Follow Button, Activity Feed, Recommendations Feed, Like Box and Facepile. See, https://developers.facebook.com/docs/plug-ins

[5] According to Quantcast ranking, http://trends.builtwith.com/widgets/Facebook-Like

[6] Chaabane, A., Kaafar, M. A., Boreli, R., “Big friend is watching you: analyzing online social networks tracking capabilities”, Proceedings of the 2012 ACM Workshop on online social networks (WOSN), 2012.

[7] See, also the elaboration by Article 29 Data Protection Working Party on third-party cookies in the context of European Data Protection Directive: “Opinion 04/2012 on Cookie Consent Exemption”, WP 194, 7 June 2012.

[8] “What names are allowed on Facebook?”, https://www.facebook.com/help/112146705538576

[9] Eckersley, P. "How unique is your web browser?", in Proceedings of the 10th International Conference on Privacy Enhancing Technologies (PETS), Berlin, Germany, 2010.

cookies[10] or other types of evercookies[11] which are not covered in this report.

Our experiments were focused on long term, identifying cookies that can be used for third-party tracking. We did not assess the outcome of our experiments in terms of changes in the advertisements received by individuals, which would require a more extensive study and a different methodology[12].

Understanding the ultimate functionality and behavior of some Facebook cookies was not possible due to encryption and obscurity. Where possible, we referred to the explanations given by Facebook to the Irish Data Protection Commissioner (DPC) during its 2011 audit[13] and 2012 re-audit.

The findings we present in this report are based on experiments ran in March 2015. Facebook may change the behavior of its software and services anytime in the future.

3 Methodology

Our analysis is composed of a number of scenarios such as the tracking of Facebook users who are logged in or logged out, tracking of non-users and the functioning of the "opt-out" mechanism suggested by Facebook. We manually carried out possible user actions such as logging into Facebook or browsing to a web page that includes Facebook social plug-ins. Where necessary, we opened Facebook accounts to study the tracking of Facebook users.

Whenever possible, we followed a similar methodology to those documented in the Irish DPC Facebook audits. Yet, we updated the experimental setup to adapt to the changes introduced by Facebook since 2012[14].

3.1 Experimental Setup

We used a clean virtual machine to carry out each individual experiment. This helped us to isolate the effect of the browsing history of the machine used in the experiments. Also, the IP address of the test machine visible to websites was shared with thousands of other computers in the university NAT pool[15],

[10] Soltani, Ashkan, et al. "Flash Cookies and Privacy." AAAI Spring Symposium: Intelligent Information Privacy Management. 2010.

[11] http://samy.pl/evercookie/

[12] See, for example, Datta, A., Tschantz, M. C., Datta, A. "Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination", in *Proceedings of Privacy Enhancing Technologies Symposium*, July 2015 and Lécuyer, M. et al. "XRay: Enhancing the Web's Transparency with Differential Correlation", in Proceedings of the *23rd USENIX Security Symposium*. August 2014, San Diego, CA.

[13] O'Reilly, Dave. "Facebook Technical Analysis Report", 16th December 2011, available at https://dataprotection.ie/documents/facebook%20report/report.pdf/appendices.pdf

[14] This primarily includes Facebook's more extensive use of encrypted connection (HTTPS). See also "Network capture" part in Section 3.1 Experimental Setup

[15] University of Leuven, www.kuleuven.be

which renders the linking to the previous history by IP address alone technically infeasible.

We used the following software for virtualization and testing:

- **Host machine:** GNU/Linux Xubuntu 14.04 LTS.
- **Guest machine:** GNU/Linux Ubuntu 12.04 LTS.
  - We fully updated the system, installed the Gnome classic desktop environment and disabled error tracker submissions, automatic updates and upgrades to prevent the pollution of network captures[16].
- **Virtualization software:** Oracle Virtualbox 4.3.22 r98236
- **Browser:** Mozilla Firefox 36.0 with Adobe Flash Player 11.2.202.442.
  - We set the browser homepage to a blank page and disabled the following features to prevent automatic background connections[17]:
    - Firefox Health Report, Crash Reporter, link prefetching, add-on metadata updating, blocklist updating, auto-update checking, anti-phishing list updating, and anti-malware list updating.
- **Network capture:** We used Wireshark Network Analyzer[18] 1.6.7 and mitmproxy[19] 0.11.3 in regular proxy mode. Wireshark is used to capture all the network traffic on the default network interface. mitmproxy is used to intercept and decrypt the HTTPS traffic[20] which cannot be done straightforwardly with Wireshark[21]. We tested our setup on HTTPS enabled websites and made sure that the HTTP traffic is not disturbed by our setup and that it is captured properly. Both Wireshark and mitmproxy were run in the guest machine.

3.2 Data collection

We collected the following data for each experiment:

- Network and HTTP(S) capture: Wireshark captures (pcap) and mitmproxy dumps are stored and analyzed. This allowed us to find the cookies set by Facebook and assess other information transferred to Facebook by means of HTTP headers.

[16] https://help.ubuntu.com/community/AutomaticConnections
[17] https://support.mozilla.org/en-US/kb/how-stop-firefox-automatically-making-connections
[18] https://www.wireshark.org/
[19] https://mitmproxy.org/
[20] https://mitmproxy.org/doc/ssl.html
[21] http://wiki.wireshark.org/SSL

- **Firefox profile and cache folder:** After each experiment, we made a backup of Firefox's profile[22] and cache folder. The profile directory contains user data such as cookies, local storage and IndexedDB. We used SQLiteStudio[23] software to check the cookies and other databases. The cache directory is also retained for the record, since the browser cache can be used as an *evercookie*[24] mechanism to track users by storing unique identifiers in the cached content or the metadata (ETag).
- **Flash cookies (LSOs):** We took a copy of the ~/.macromedia/Flash_Player/#SharedObjects/ directory to inspect possible use of Flash cookies, otherwise known as local shared objects (LSO).

4 Tracking of Non-Facebook Users

We tested several scenarios involving tracking of non-Facebook users including the scenarios analyzed in the Irish DPC's 2011 audit.

4.1 Prior visit to a Facebook page

In this scenario, a non-Facebook user visits a page under the facebook.com domain and then visits other sites that include Facebook social plug-ins.

With a clean virtual machine, we visited Facebook's homepage (facebook.com). We found that, a cookie named “datr” with a 2-year lifetime was set. The “datr” cookie contained a 24-character random-looking alphanumeric string and was scoped to the domain *.facebook.com* and the path “/”, meaning the cookie will be sent when fetching resources from the domain facebook.com and all its subdomains. Moreover, three additional session cookies were set by Facebook, reg_fb_gate, reg_fb_ref, wd which keep track of the first and last Facebook page visited by the user and the inner dimensions of the browser window respectively.

We then visited a web page on *gayworld.be*, a website that includes a Facebook social plug-in. The inspection of the network traffic revealed that the “datr” cookie is sent to facebook.com domain in the *Cookie* header of the HTTP requests. The *Referer* [sic] header of the same request includes the URL of the currently visited page. In addition, the URL of the page to be liked is included in the “*href*” parameter.

[22] https://support.mozilla.org/en-US/kb/profiles-where-firefox-stores-user-data
[23] http://sqlitestudio.pl/
[24] Ayenson, M. et al. “Flash Cookies and Privacy II: Now with HTML5 and ETag respawning.” World Wide Web Internet and Web Information Systems, 2011. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1898390

Table 1: The cookies placed when a non-Facebook user visits Facebook page.

Name	Sample Value	Contains*	Expires	Secure
datr	`jicEVbqr3GxEtizEbP6XEG_c`	Browser ID	2 years	No
reg_fb_gate	`https%3A%2F%2Fwww.facebook.com%2Fpolicy.php`	URL of the first visited Facebook page¶	Session	No
reg_fb_ref	`https%3A%2F%2Fresearch.facebook.com%2F`	URL of the last visited Facebook page¶	Session	No
reg_ext_ref†	`https%3A%2F%2Fwww.google.be%2F[...]`	URL of the external referrer¶	Session	No
wd	`1280x653`	Browser window dimensions	Session	No

**: The behavior of these cookies seems to be unchanged since the Irish DPC's report.*
¶: URL is stored in percent encoded form.
†: reg_ext_ref cookie is only placed when the user is referred from an external site such as a search engine.

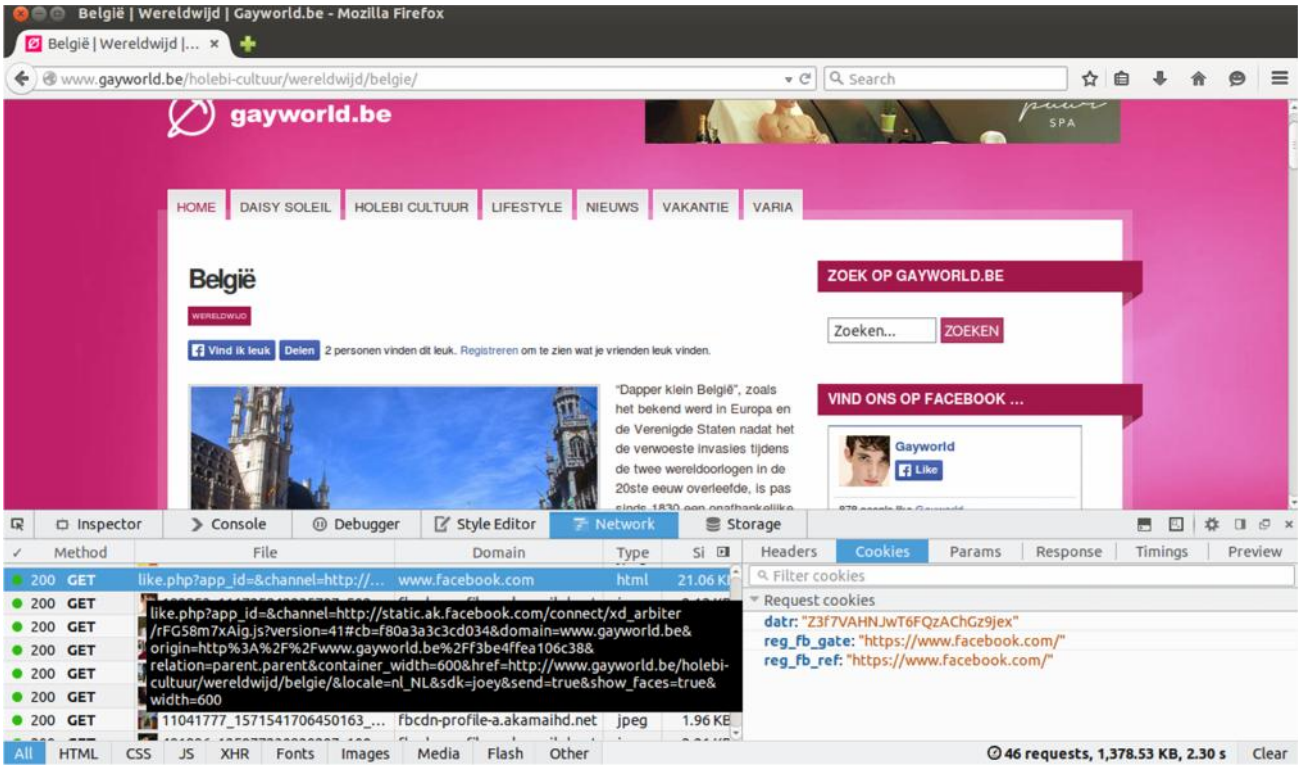

Figure 2 *Facebook receives the cookies previously set on the Facebook page while loading the Like Button.*

```
2015-03-07 23:11:14 GET https://www.facebook.com/plugins/like.php?app_id=&channel=http%3A%2F%2Fstatic.ak.facebook.com%2Fconnect%2Fxd_arbiter%2FrFG58m7xAig.js%3Fversion%3D41%23cb%3Df80a3a3c3cd034%26domain%3Dwww.gayworld.be%26origin%3Dhttp%253A%252F%252Fwww.gayworld.be%252Ff3be4ffea106c38%26relation%3Dparent.parent&container_width=600&href=http%3A%2F%2Fwww.gayworld.be%2Fholebi-cultuur%2Fwereldwijd%2Fbelgie%2F&locale=nl_NL&sdk=joey&send=true&show_faces=true&width=600
                       ← 200 text/html 7.31kB 637.94kB/s
Request                                                Response
Host:              www.facebook.com
User-Agent:        Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:36.0) Gecko/20100101 Firefox/36.0
Accept:            text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language:   en-US,en;q=0.5
Accept-Encoding:   gzip, deflate
Referer:           http://www.gayworld.be/holebi-cultuur/wereldwijd/belgie/
Cookie:            datr=Z3f7VAHNJwT6FQzAChGz9jex; reg_fb_ref=https%3A%2F%2Fwww.facebook.com%2F;
                   reg_fb_gate=https%3A%2F%2Fwww.facebook.com%2F
Connection:        keep-alive
```

Figure 1 *A more detailed look at the information received by Facebook while visiting a page that includes a Facebook Like Button. The "datr" cookie, which identifies the browser, is sent with other information about browser, operating system and language preferences.*

According to the Audit Report of the Irish DPC (2011), the "datr" cookie identifies a browser used to connect to Facebook[25]. The "datr" cookie is not flagged as secure, hence it may be sent over the unencrypted connections allowing the tracking of non-Facebook users by adversaries who can monitor

[25] http://www.dataprotection.ie/documents/facebook%20report/final%20report/report.pdf

the network[26]. In our experiments we have witnessed several cases where the "datr" cookie is placed or sent in the clear, without encryption.

In a related experiment, we started with a clean virtual machine and **made a Google search with the terms "*facebook data policy*"**. We visited the first search result, which happened to be the Facebook Data Policy page[27]. While loading the policy page, Facebook placed the "datr" cookie with a 2-year lifetime. There was no formal notice regarding any cookie being stored. We then visited a Belgian website related to prostate cancer treatment[28] which includes a Facebook social plug-in. By inspecting the captured network traffic, we found that the "datr" cookie is sent to Facebook while requesting the Facebook Like Button.

Finally, we confirmed the finding that Facebook sets a long-term, identifying "datr" cookie on other pages belonging to the facebook.com domain, by visiting a Facebook event page and a fan page following the same methodology.

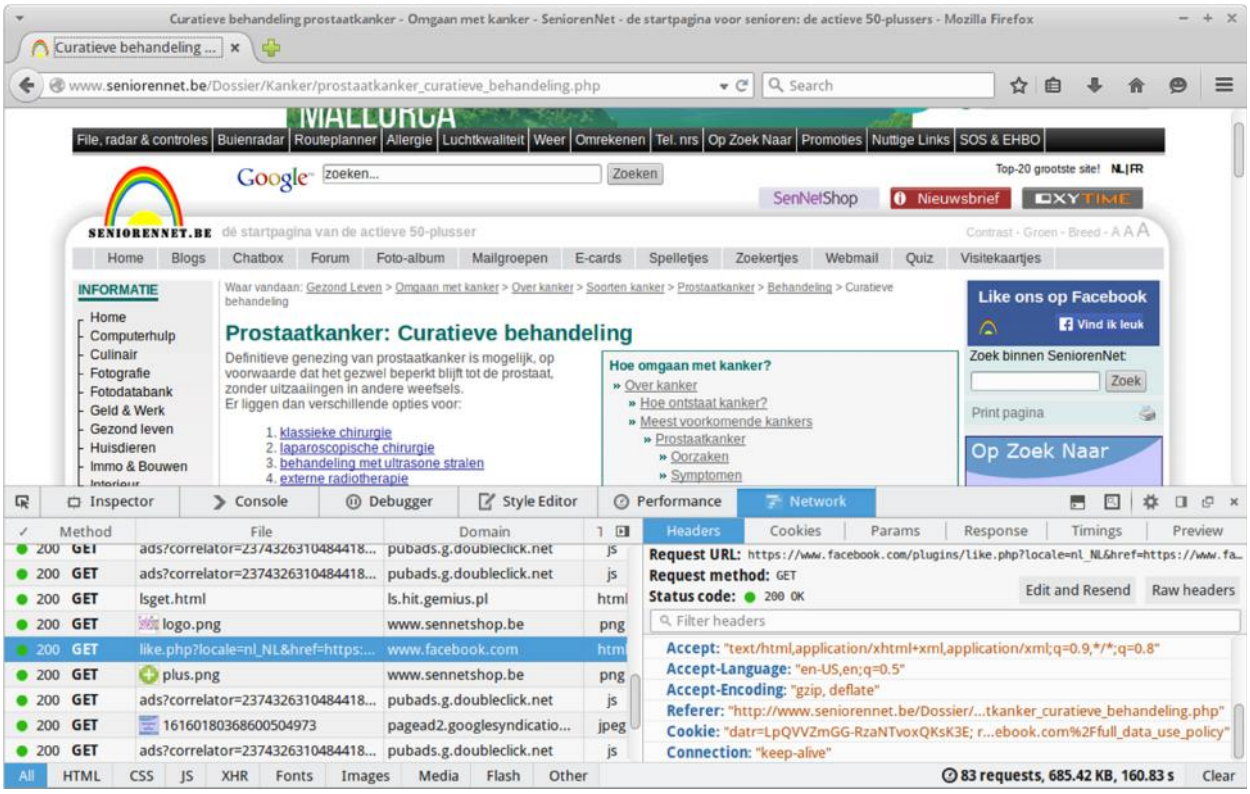

Figure 3 *Cookies placed on the Facebook policy page sent to Facebook while visiting a cancer related website*

[26] See, e.g., "How the NSA & FBI made Facebook the perfect mass surveillance tool", http://venturebeat.com/2014/05/15/how-the-nsa-fbi-made-facebook-the-perfect-mass-surveillance-tool/ and S. Englehardt, "How cookies can be used for global surveillance" (https://freedom-to-tinker.com/blog/englehardt/how-cookies-can-be-used-for-global-surveillance/)

[27] https://www.facebook.com/policy.php

[28] http://www.seniorennet.be/Dossier/Kanker/prostaatkanker_curatieve_behandeling.php

4.2 No prior visit to Facebook page

In this scenario *a non-Facebook user never visits a page from the domain facebook.com, but visits sites that include Facebook social plug-ins.*

With a clean virtual machine, we visited the home pages of imdb.com, hln.be, rtbf.be sites[29]. All of these sites include Facebook Like Buttons. By inspecting the cookies transmitted and retained after the visits, we found that the **Facebook social plug-ins did not set a cookie** in this scenario.

In order to test different social plug-ins provided by Facebook[30], we set up simple test pages and added a Facebook social plug-in to each page. We had different pages for different embedding options of each plug-in[31]. Overall, we set up more than 25 pages, all of which include a Facebook plug-in. We then visited the test pages on a clean virtual machine and confirmed that **no cookie was set by Facebook**.

4.2.1 Cookies set by Facebook on non-Facebook pages

Although the finding of the previous section might suggest that one can avoid tracking by Facebook social plug-ins by not visiting Facebook, the following cases show that Facebook sometimes sets cookies when it's in the third-party position, i.e. on pages outside facebook.com. We first look into the case of cookies set by the Facebook social plug-ins and then to the case of certain websites that include Facebook's authentication library, called "Facebook Connect."

Cookies set by Facebook social plug-ins

We turned to publicly available data from the HTTP Archive[32] to search for cookies set by Facebook social plug-ins on third party domains. We queried the HTTP Archive database for the data collected in March 2015[33] using Google BigQuery[34]. The queries revealed that, although Facebook never sets a cookie when the browser fetches the social plug-in, in some cases, social plug-ins initiate a request to pixel.facebook.com domain, which then sets a "datr" cookie[35]. We confirmed this behavior on several

[29] imdb.com is used in the 2011 audit by the Irish DPC. The latter two websites belong to Belgian news media.
[30] There are 11 different types of social plug-ins provided by Facebook, see footnote 1.
[31] HTML5, XFBML, Iframe and URL. (As of March 22, these integration options are removed. Now, Facebook offers only one way of integrating social plug-ins.)
[32] HTTP Archive is a publicly available archive of HTTP requests and responses from 500,000 websites.
[33]
```
SELECT pages.rank, pages.self, pages.url, req_referer, respCookieLen, respBodySize, req.url, status
  FROM [httparchive:runs.2015_03_01_requests] AS req JOIN (
  SELECT DOMAIN(url) self, url, pageid, rank FROM [httparchive:runs.2015_03_01_pages])
  AS pages ON pages.pageid = req.pageid
  WHERE req.url CONTAINS "pixel.facebook.com" AND respCookieLen > 0 AND reqCookieLen = 0 AND
  DOMAIN(req_referer) = "facebook.com" AND req_referer CONTAINS "/plugins/" ORDER BY pages.rank;
```
[34] https://bigquery.cloud.google.com
[35] This can be verified by searching "datr" on http://httparchive.webpagetest.org/export.php?test=150301_0_BC4&run=2&cached=0&pretty=1 and checking the

websites[36] using our experimental setup and found that after the page finishes loading, a request was made to a URL that starts with *https://pixel.facebook.com/si/kappa/*[37].

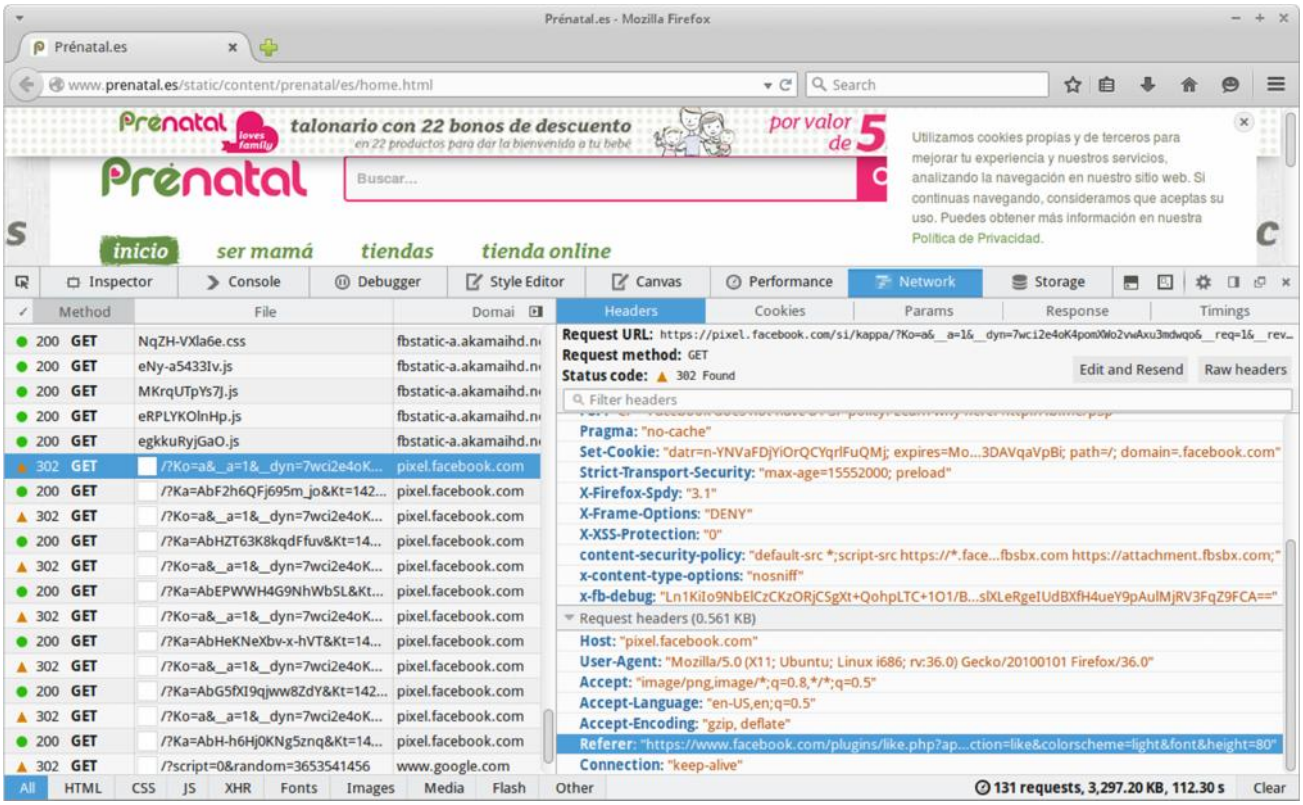

Figure 4 *Facebook sets "datr" cookie in response to the request made by Like Button source code.*

The HTTP Referer [sic] header of these requests was always a Facebook social plug-in URL[38], meaning that it was initiated by the social plug-in frame[39]. We also found that the pixel.facebook.com domain was mentioned in the plug-in source code[40].

We observed that, when a request to pixel.facebook.com was made, five additional requests follow the first one, separated by intervals that of five seconds or more. Moreover, by experimenting with the websites on which this behavior was observed, we found that the requests to pixel.facebook.com were not made on all visits. Also note that, Facebook offers "Conversion Pixel"[41] and "Custom Audience

referrer of the request.

[36] Including *prenatal.es, digitalnest.in, kateleong.com, endlesssimmer.com.*

[37] An example URL observed in our experiments was https://pixel.facebook.com/si/kappa/?Ko=a&__a=1&__dyn=7wci2e4oK4pomXWo2vwAxu3mdwqo&__req=1&__rev=1645171&__user=0&asyncSignal=4201&locale=en_US&lsd=AVrFxPzr.

[38] e.g. *https://www.facebook.com/plugins/like.php/[...]*

[39] Facebook social plug-ins are rendered in an IFrame element. See, https://en.wikipedia.org/wiki/IFrame.

[40] Exact snippet was as follows: *["TrackingConfig", [], {"domain": "https:\/\/pixel.facebook.com"}, 325].*

[41] https://www.facebook.com/help/1563508590530683

pixel”[42] to allow website owners to add their visitors to custom segments and retarget them on Facebook with Facebook ads. But the URL used for these pixels are different than pixel.facebook.com[43].

HTTP Archive contains data from crawls that were run every two weeks since November 2010. We searched the archive to find the first time this behavior was observed. We identified 1 August 2014 as the earliest date a Facebook social plug-in set a cookie by using the pixel.facebook.com domain.

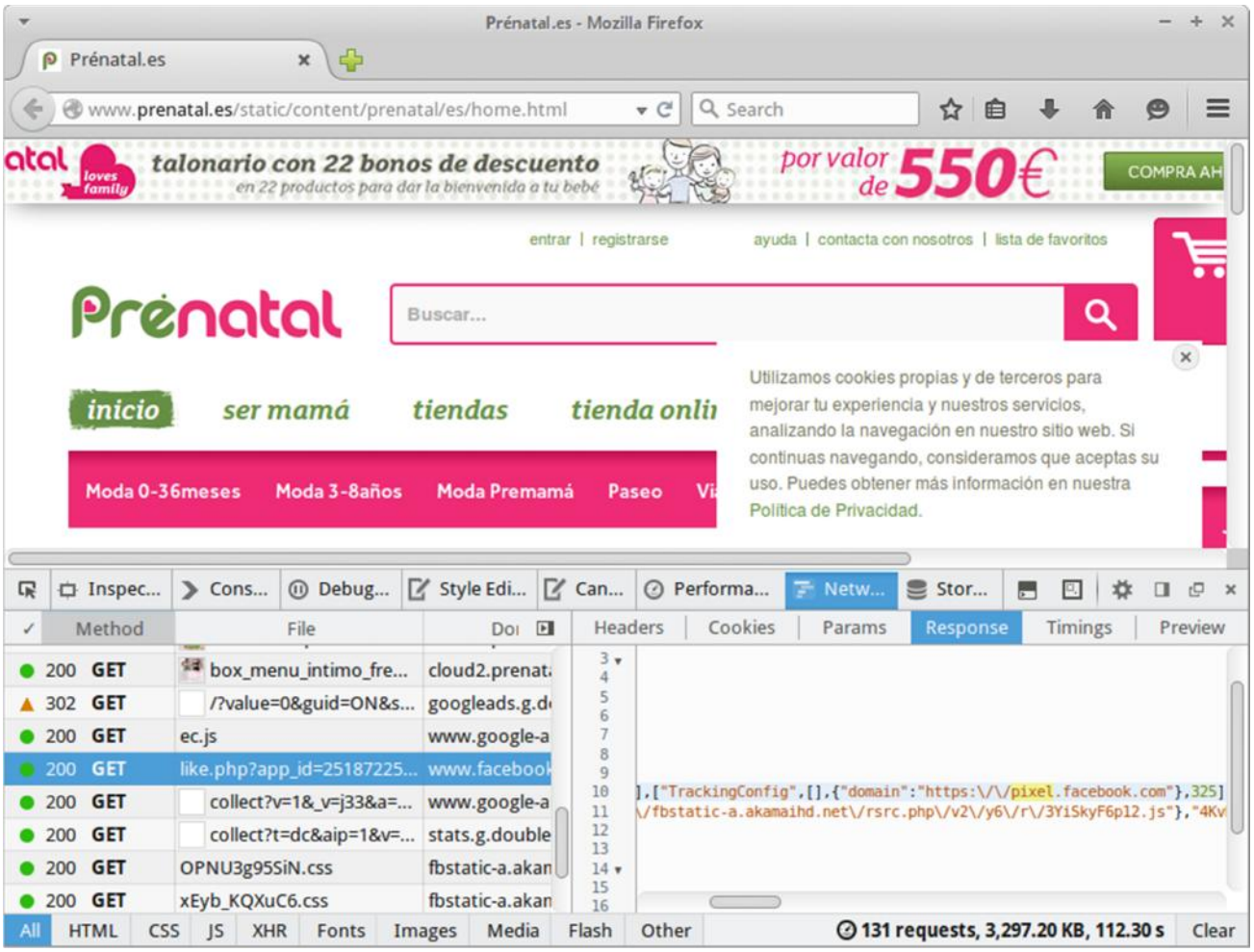

Figure 5 *The pixel.facebook.com domain mentioned in the source of the Facebook Like button*

[42] https://developers.facebook.com/docs/marketing-api/custom-audience-website/faq/v2.3#fbpixel
[43] These pixels use a URL starting with the following: *https://www.facebook.com/tr?id=*

Sites with Facebook Connect

By querying the HTTP Archive[44], we found that, on certain websites, Facebook sets a cookie when it's in the third-party position, while fetching a script from the *connect.facebook.com* subdomain. We then studied these sites more closely using our experimental setup. By visiting these candidate sites with a clean virtual machine, we found that Facebook sets a "datr" cookie on websites including *myspace.com*, *okcupid.com* and *mtv.com*[45] while fetching a script (sdk.js or all.js) from the *connect.facebook.com* subdomain[46]. We did not interact with the page such as logging in or clicking links. Visiting the homepage of these three sites was enough for the placement of the "datr" cookie and there was no visible presence of any Facebook plug-in.

The findings suggest that, **Facebook sets a "datr" cookie on certain non-Facebook pages, thus enabling the tracking by social plug-ins even if the user never visits a Facebook page.**

[44] We ran the following query against HTTP Archive using Google BigQuery:

```
SELECT pages.rank, pages.self, req_referer, respCookieLen, respBodySize, req.url, status
    FROM [httparchive:runs.2015_03_01_requests] AS req JOIN (SELECT DOMAIN(url) self, pageid, rank
    FROM [httparchive:runs.2015_03_01_pages]) AS pages ON pages.pageid = req.pageid
    WHERE (domain(req.url) = "facebook.com") AND req.url contains "connect.facebook.com"
    AND (NOT self = "facebook.com") AND (NOT self = "fb.me") AND (NOT self = "fbsbx.com")
    AND (NOT self = "fbcdn.net")  AND respCookieLen > 0 AND reqCookieLen = 0 AND
    NOT req_referer contains "plugin" ORDER BY pages.rank;
```

[45] The following publicly available pages on HTTP Archive can be used to verify our finding that Facebook sets a "datr" cookie on *okcupid.com* and *mtv.com* websites:
http://httparchive.webpagetest.org/export.php?test=150215_0_4TE&run=1&cached=0&pretty=1
http://httparchive.webpagetest.org/export.php?test=150222_0_168&run=2&cached=0&pretty=1

[46] Note that the cookies set by Facebook Connect have been analyzed in Roosendaal, A., "We Are All Connected to Facebook … by Facebook!", in S. Gutwirth et al. (eds), European Data Protection: In Good Health?, Springer, 2012, p. 3-19. An earlier version of this paper is available on SSRN as "Facebook tracks and traces everyone: Like this!", Tilburg Law School Legal Studies Research Paper Series, No. 03/2011, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1717563

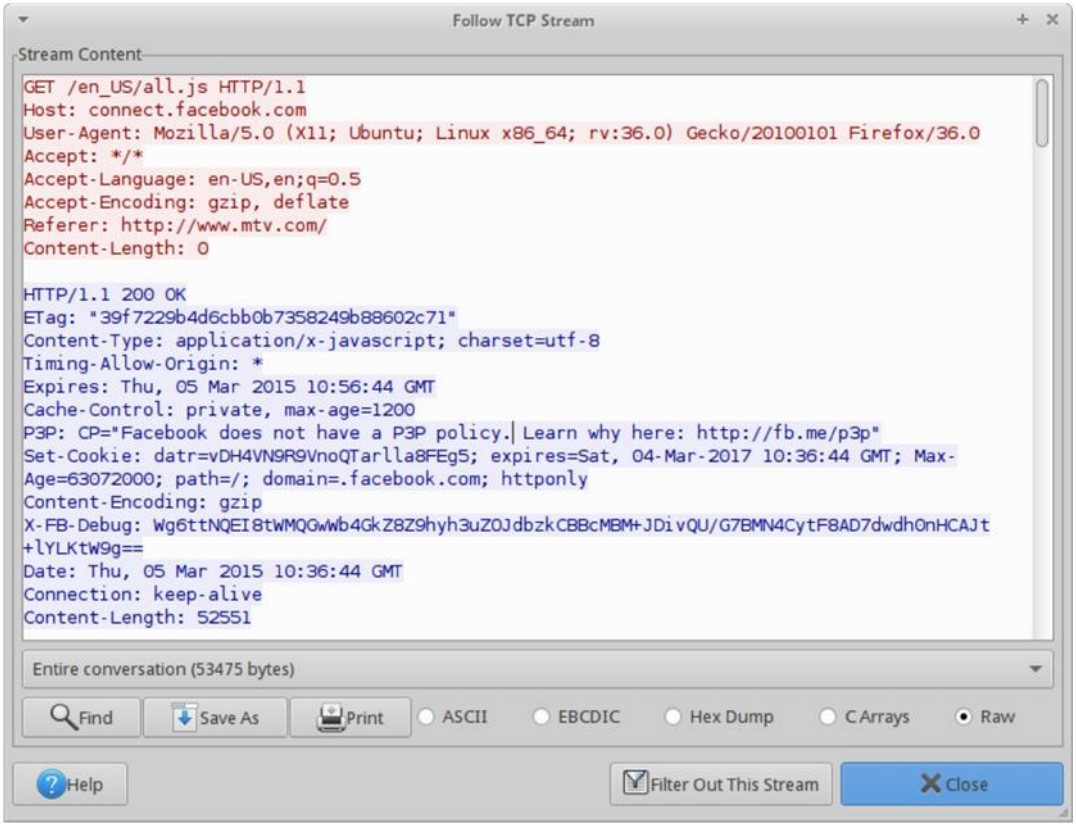

Figure 6 *Facebook sets "datr" cookie on mtv.com as a third-party.*

5 Tracking of Facebook Users

5.1 Logged in Facebook Users

In order to analyze the tracking of Facebook users, we logged in to the Facebook account we created for the experiments. We visited some Facebook and non-Facebook pages. We then checked the cookies sent to Facebook when visiting a site with social plug-ins; Table 2 lists the cookies we have identified.

If the Facebook user is logged in when visiting a site that include Facebook social plug-ins, Facebook received a total of 11 cookies in addition to the URL of the page being visited. The cookies include the Facebook ID (c_user), the browser ID (datr), the encrypted Facebook ID and browser ID (fr). Even if the user closes the Facebook tab (but not the browser), the cookies will be sent to Facebook as they are retained until the browser is closed. Depending on the status of the "Keep me logged in" checkbox, the lifetime of some cookies may vary.

Table 2: The list of cookies sent to Facebook when a logged in user visits a page with social plug-ins.

Name	Sample Value	Contains	Expires	Secure†
c_user	`100004223456398`	Facebook ID	Session/ 1 Month¶	Yes
datr	`S3fJVgeTh7_ikK5frtHsHPmE`	Browser ID	2 Years	No
fr	`0goRJJKaszKOLdKz8.AWXGHlRrxSLM3PHeHxfrORvl0H8.BCVChV.Sj.FUJ.0.AWWSuv8a`	Encrypted Facebook ID and Browser ID*	1 Month	No
lu	`wfKm8ItfbXqRklNoERo10HlH`	Encrypted ID of the last user*	2 Years	Yes
p	`-2`	User's channel partition*	Session	No
presence	`EM426705095EuserFA21B09211298286A2EstateFDutF1426705095426Et2F`	Chat state*	Session	Yes
s	`Aa67DZudqH2wPHl9`	?	Session/ 1 Month¶	Yes
xs	`244%3AjIZKp45fK9ceMA%3A2%3A1426705088%3A3455`	Session number and secret*	Session/ 1 Month¶	Yes
csm	`2`	Insecure indicator[47]	Session/ 1 Month¶	No
act	`1426704200575%2F14`	Timestamp and counter of user actions[48]	Session	No
wd	`1280x653`	Browser window dimensions	Session	No

***:*The descriptions are taken from the Irish DPC Audit Report[49] and the follow-up Review Report[50]. ¶: the cookie's lifetime depends on the "Keep me logged in" checkbox. If the box is checked, the cookie will expire in 1 month, otherwise it will be removed at the end of the session. †: If the secure attribute of the cookie is set (Yes), then the cookie will always be sent over the secured (HTTPS) connections.*

[47] https://www.facebook.com/notes/facebook-engineering/secure-browsing-by-default/10151590414803920
[48] https://www.nikcub.com/posts/facebook-fixes-logout-issue-explains-cookies/
[49] O'Reilly, Dave. "Facebook Technical Analysis Report", 16th December 2011, available at https://dataprotection.ie/documents/facebook%20report/report.pdf/appendices.pdf
[50] "Facebook Ireland Audit Review Report", 21 September 2012, available at http://www.dataprotection.ie/docs/21-09-12-Facebook-Ireland-Audit-Review-Report/1232.htm

When a logged in Facebook user visits a site with Facebook social plug-ins, Facebook receives the Facebook ID and browser ID, along with the URL of the page being visited.

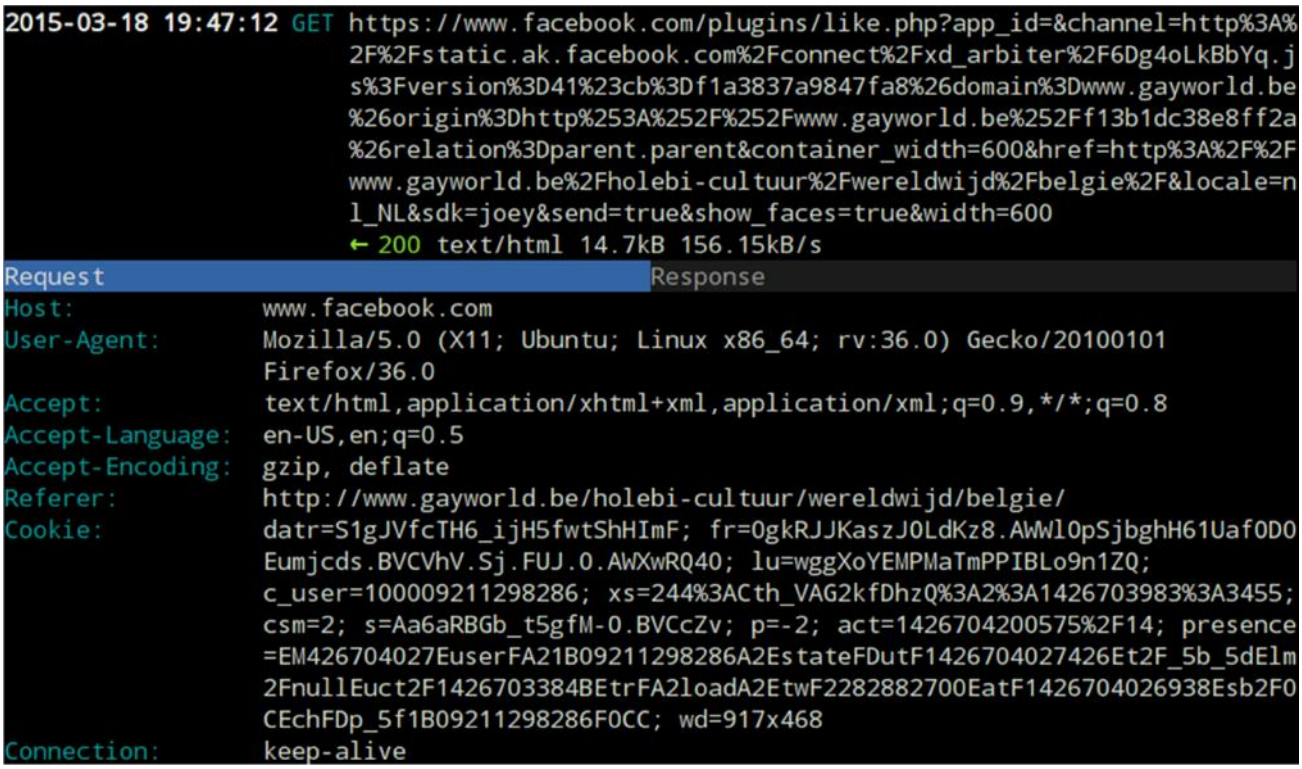

Figure 7 *If the Facebook users is logged in when she visit a site with a social plug-in, a total of 11 cookies (including "c_user" which contains the Facebook ID) are sent to Facebook, along with the URL of the page being visited.*

5.2 Logged out Facebook Users

In order to analyze the tracking of users who are logged out from Facebook, we ran the following experiment. We first logged in to Facebook without checking the "*Keep me logged in*" check box. We then logged out and restarted the browser to get rid of session cookies. We found four cookies retained in the browser (Table 3), all of which were scoped to .facebook.com domain and "/" path, meaning they will be sent to Facebook while fetching resources from facebook.com and its subdomains. We then visited a page that includes a Facebook social plug-in[51] and verified that all four cookies (datr, fr, lu, locale) are sent to Facebook while requesting the social plug-in.

[51] http://www.seniorennet.be/Dossier/Kanker/prostaatkanker_curatieve_behandeling.php

Table 3: Facebook retains the encrypted Facebook ID and browser ID even when the user logs out.

Name	Sample Value	Contains*	Expires	Secure
datr	`jicDVaqr2GxErizEbP6XEG_c`	Browser ID	2 years	No
fr	`0ZuGN96ZBkLEA1JM3.AWUNZHOO08Z1ODyL5rtIr3wSPWI.BVAyeV.An.AAA.0.AWVap1JO`	Encrypted Facebook user ID and browser ID	3 months	No
lu	`RANYg9GZTworKrnDvBE5m6aQ`	Auto-login state†	2 years	Yes
locale¶	`en_US`	Locale of the last user	1 week	No

**: According to Facebook's response to the 2011 Irish DPC Audit Report[52] and the 2012 Audit Review Report[53]*. ¶: "locale" cookie is set when a user logs out from Facebook. †: Part of the "lu" cookie holds the user ID of the previously logged in user, but this is set to zero when the user explicitly logs out.

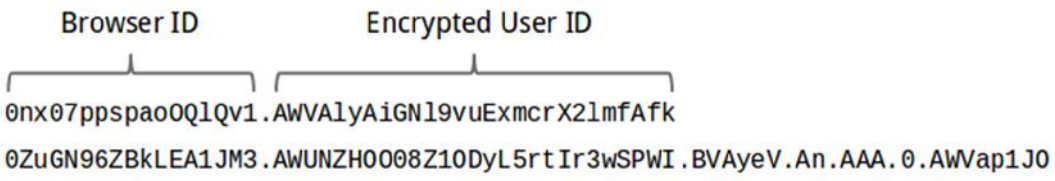

Figure 8 *"fr" cookie content as explained in Irish DPC's 2012 Audit Review Report (above) and as it is observed in our experiments (below). Despite the addition of new parts (to the right), browser ID and encrypted user ID parts seem to have remained the same.*

The cookies listed in Table 3 were studied in the Irish DPC's Facebook Audit Report and the Audit Review Report. According to Facebook's explanation to the Irish DPC, the "fr" cookie is used for advertising and contains the encrypted Facebook user ID and the browser ID. The lifespan of the "fr" cookie was noted as 1 month in the audit report, which was the exact lifespan we observed during our experiments in early March 2015. However, during our experiments we noted that the lifespan of the cookie was extended to 3 months somewhere in March 2015.

[52] O'Reilly, Dave. "Facebook Technical Analysis Report", 16th December 2011, available at https://dataprotection.ie/documents/facebook%20report/report.pdf/appendices.pdf

[53] "Facebook Ireland Audit Review Report", 21 September 2012, available at http://www.dataprotection.ie/docs/21-09-12-Facebook-Ireland-Audit-Review-Report/1232.htm

The finding suggests that **when a Facebook user explicitly logs out, Facebook keeps uniquely identifying "fr" and "datr" cookies in the browser, which are then used to track logged-out users across the web using its social plug-ins.**

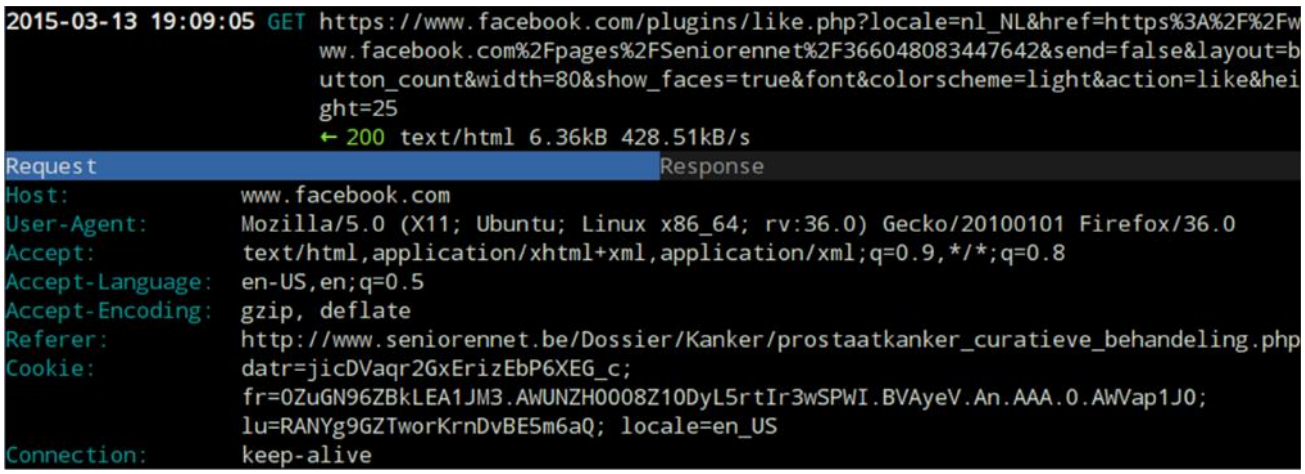

Figure 9 *When a logged-out Facebook user visits a page with a Facebook social plug-in, uniquely identifying "fr" and "datr" cookies are sent to Facebook along with the visited page.*

5.3 Deactivated Facebook Users

Facebook allows its users to deactivate their accounts. In order to assess the effect of deactivation to Facebook's tracking by social plug-ins, we deactivated our Facebook account and analyzed the cookies sent to Facebook while visiting pages with social plug-ins. Specifically, using a clean virtual machine, we logged in to our Facebook account and clicked "*Deactivate your account*" link on the Security Settings page. That took us to the deactivation page where Facebook requires users to provide a "*Reason for leaving*". We chose "*I have a privacy concern*" and clicked the "*Deactivate*" button. After confirming our password, Facebook displayed the message "*Your account has been deactivated*" and logged our user out.

We then restarted the browser and checked the cookies retained after the deactivation. We found that the cookies named "fr", "datr", "lu" and "locale" have not been deleted during the deactivation. We then inspected the network traffic while visiting two websites that include Facebook social plug-ins[54] *and confirmed that* "fr", "datr", "lu" and "locale" cookies are sent to Facebook while loading the social plug-ins.

As noted in the previous section on logged out users, according to Facebook, the "fr" cookie is used for advertising purposes and contains the encrypted Facebook ID and the browser ID. In addition, the "datr"

[54] http://www.seniorennet.be/Dossier/Kanker/prostaatkanker_curatieve_behandeling.php, http://www.hln.be/

cookie contains the browser ID. The cookies retained after deactivation were the same as the ones retained after the log-out.

The finding suggests that **when a Facebook user deactivates her account, Facebook does not remove the uniquely identifying "fr" and "datr" cookies. These cookies are subsequently used to track deactivated users across the web using Facebook's social plug-ins.**

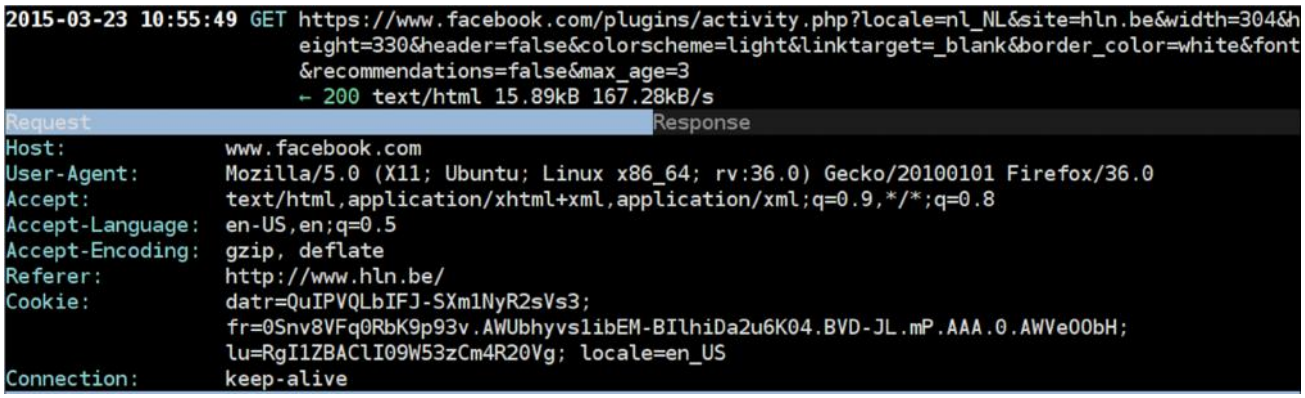

Figure 10 *When a deactivated Facebook user visits a page with a Facebook social plug-in, uniquely identifying "fr" and "datr" cookies are sent to Facebook.*

6 The "opt out" mechanism proposed by Facebook

Facebook's "About Facebook Ads" page[55] points out websites that users can visit and opt-out from interest-based advertising:

> "If you don't want Facebook or other companies to collect or use information based on your activity on websites, devices or apps off Facebook for the purpose of showing you ads, you can opt out from all participating companies through the Digital Advertising Alliance in the USA, the Digital Advertising Alliance of Canada in Canada, or the European Interactive Digital Advertising Alliance in Europe."

As noted in Section 2, we describe the effect of opt-out in terms of cookie-based tracking. We do not study the use of other tracking mechanisms, nor do we assess the claimed effect of opt-out on the advertisements targeted to users.

6.1 Opting-out with a clean profile

In the following, we analyze the case of individuals who don't have a cookie from Facebook at the

[55] https://www.facebook.com/about/ads/

moment of opt-out. Non-users who have never visited a Facebook page, or Facebook users who clear their cookies after logging out from Facebook would fall into this category.

6.1.1 European Opt-out Site

With a clean virtual machine, we visited the European Interactive Digital Advertising Alliance (EDAA) opt-out website (www.youronlinechoices.eu). We clicked the Belgium/Flemish link and "Je advertentievoorkeuren" (*Your Ad Choices*) button and waited for the website to populate the status of the participating companies which included Facebook. **After the status check was complete, we found that Facebook placed the cookie named "datr"**[56] along with three other session cookies "reg_fb_gate", "reg_fb_ref" and "reg_ext_ref." The "datr" cookie was set over an unencrypted connection and contained a unique identifier.

We then clicked the "Alle bedrijven uitzetten" (*Turn off all companies*) button to opt-out from the listed companies. During the opt-out, Facebook placed a cookie named "oo" with the value "1". The cookie name "oo" presumably stands for "opt-out". The "datr" cookie which was set on the status check page was not removed by Facebook during or after the opt-out.

Using the same virtual machine and the browser, we then visited a site that includes a Facebook social plug-in. By inspecting the network traffic, we confirmed that both the "oo" and "datr" cookies were sent to Facebook while loading resources from the domain facebook.com.

EDAA offers localized versions of their website for different countries and languages. In addition to Belgium – Flemish version, we confirmed our finding on the UK[57] version of the opt-out site by following the same methodology.

Note that, Facebook is not the only company that sets a long-term identifying cookie on the EDAA opt-out page. But we observed that some companies follow a better practice, for example, by removing the identifiers in the cookies[58].

The finding suggests that **Facebook places a long-term, uniquely identifying cookie on the website suggested by Facebook to European users for opting out from interest-based advertising. All the**

[56] We would like to thank Steven Englehardt from Princeton University for confirming this finding.
[57] http://www.youronlinechoices.com/uk/
[58] For instance, during the status check, Google's third-party advertising domain doubleclick.net placed a uniquely identifying cookie named "id." But, after we opted-out, the unique identifier in the cookie was replaced with "OPT_OUT" (i.e. the unique identifier was removed). On the other hand, we found that Google placed two new identifying cookies (NID and PREF) for its first party domain (google.com) after we clicked "*Turn off all the companies*" to opt-out.

later visits to pages that include Facebook social plug-ins can be linked by Facebook using this cookie which has a lifespan of two years.

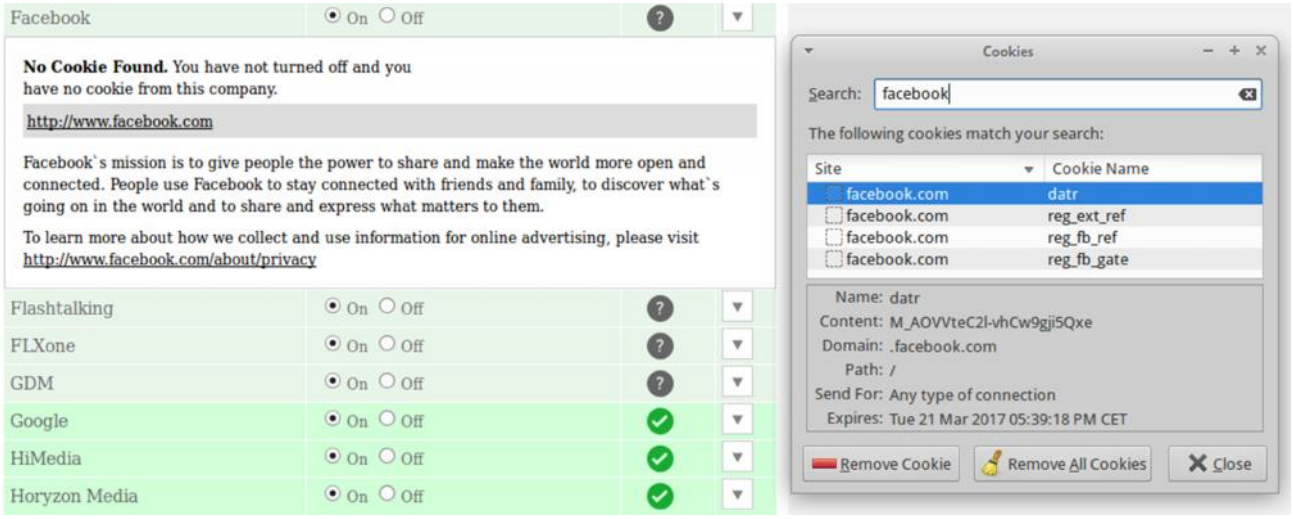

Figure 11 *Facebook sets four cookies during the status check on the EDAA opt-out site but the site reports "No Cookie Found" for Facebook. The cookie status was not corrected after we reloaded the page.*

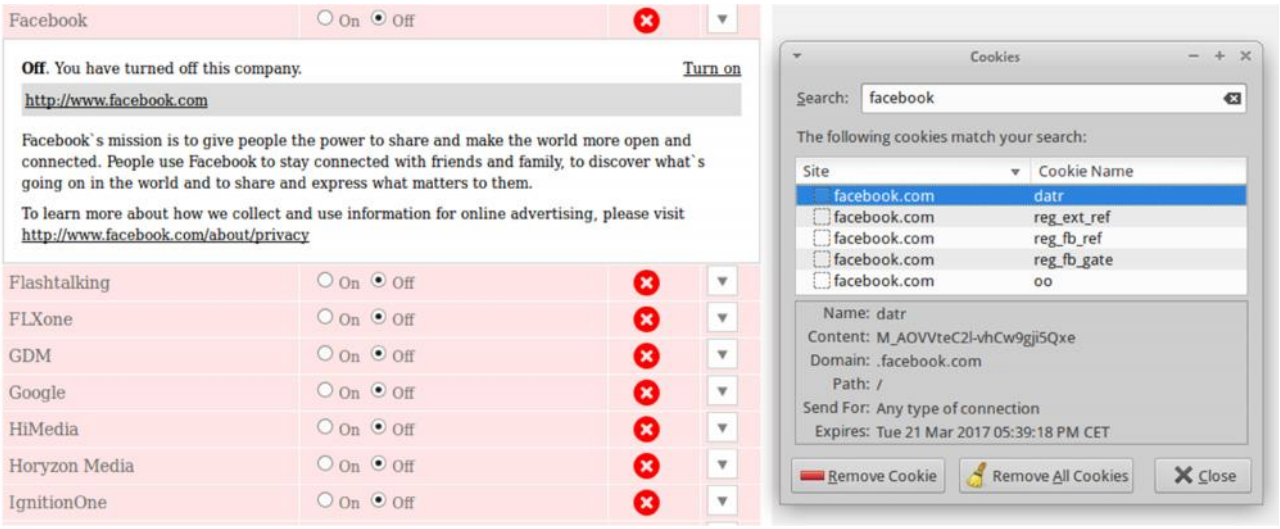

Figure 12 *Facebook retains the "datr" cookie after the opt-out.*

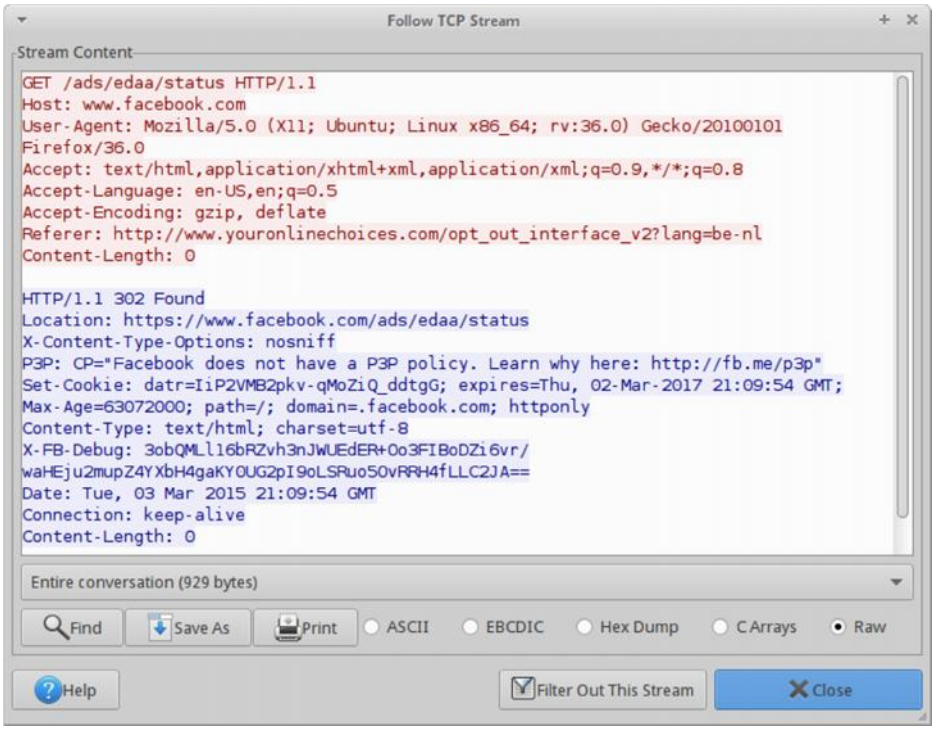

Figure 13 *Facebook sets a tracking cookie ("datr") on the EDAA opt-out site. EDAA opt-out site is suggested by Facebook to European users to control interest-based advertising.*

6.1.2 US and Canadian Opt-out Sites

We compared Facebook's cookie setting behavior on EDAA opt-out website to USA[59] and Canadian[60] Digital Advertising Alliance (DAA) opt-out sites suggested by Facebook[61]. By visiting and opting out on these sites with clean virtual machines, we found that **Facebook did not place "datr" or any other long term identifying cookie on the US and Canadian opt-out sites.** On these two sites, Facebook only placed a (non-identifying) cookie named "oo" with the value "1", which had a lifespan of 5 years.

6.2 Opting-out as a Facebook user

In the following, we describe the effects of the opt-out in relation to tracking of Facebook users.

6.2.1 European opt-out site

We logged in to our Facebook account and visited the opt-out site recommended by Facebook to European users (www.youronlinechoices.eu). We clicked the Belgium/Flemish link and "Je

[59] http://www.aboutads.info/choices/
[60] http://youradchoices.ca/
[61] https://www.facebook.com/about/ads/

advertentievoorkeuren" (*Your Ad Choices*) button and waited for the website to populate the status of the participating companies. We then clicked the "Alle bedrijven uitzetten" (*Turn off all companies*) button to opt-out from the listed companies. **During the opt-out**, **Facebook placed a cookie named "oo" with the value "1" but did not remove any of the cookies stored in the browser, including the "fr" cookie, which, according to Facebook's 2012 statements[62], is used for advertisement purposes**. Visiting two sites that contain Facebook social plug-ins, **we confirmed that Facebook still receives the uniquely identifying cookies such as "c_user", "datr", "lu" and "fr" after the user opts out.**

We then **logged out from** our **Facebook** account and analyzed the cookies received by Facebook when an opted-out user also logs out from Facebook. Visiting a site that includes a Facebook social plug-in, we found that Facebook still received the "fr", "datr", "lu" and "locale" cookies in addition to the "oo" cookie placed on the opt-out site. Thus, **even if a Facebook users opts-out from interest-based advertising and logs out from her account, Facebook still tracks her browsing activity through social plug-ins. One of the cookies collected by Facebook is, according to Facebook's 2012 statements, used for advertisement purposes.**

6.2.2 US and Canadian Opt-out Sites

Using a clean virtual machine, we logged in to our Facebook account and visited the USA[63] Digital Advertising Alliance (DAA) opt-out site suggested by Facebook. We opted-out from all the companies and then visited a website that includes a social plug-in. We confirmed that Facebook still receives the uniquely identifying cookies such as "c_user", "datr", "lu" and "fr" after the opt-out. We then logged out from our Facebook account and analyzed the cookies received by Facebook when an opted-out user also logs out from Facebook. Analyzing the network traffic we found that Facebook still received the "fr", "datr", "lu" and "locale" cookies in addition to the "oo" cookie placed on the opt-out site. Thus, **even if a Facebook users opts-out from interest-based advertising on the US opt-out site and logs out from her account, Facebook still tracks her browsing activity using social plug-ins**.

We confirmed this finding on the Canadian Digital Advertising Alliance (DAA)[64] opt-out site using the same methodology.

The findings suggest that **there is no difference between North American and European opt-out sites in terms of their effects on tracking of Facebook users**. **The sites recommended by Facebook to its**

[62] O'Reilly, Dave. "Report on Facebook Ireland (FB-I) Audit 2-3 May & 10-13 July 2012", 21 September 2012, p. 34, https://dataprotection.ie/documents/press/Facebook_Ireland_Audit_Review_Report_21_Sept_2012.pdf
[63] http://www.aboutads.info/choices/
[64] http://youradchoices.ca/

users for opting-out from interest-based advertising does not stop tracking by Facebook social plug-ins. Facebook still collects the browsing information of its users, even if they log out from Facebook after opting-out on the recommended sites. Facebook still receives uniquely identifying cookies, including one (“fr”) that is used for advertising and contains the encrypted Facebook ID and browser ID.

Acknowledgements

We would like to thank Steven Englehardt for confirming our finding about the opt-out websites, and Marc Juarez and Iraklis Symeonidis for their helpful discussions.

Appendix - Social Plug-ins on Popular Websites in Belgium

The following is a list of top sites visited by Belgian web users that contain Facebook social plug-ins. The list is compiled in cooperation with the *College Bescherming Persoonsgegevens* (Dutch Data Protection Agency) based on a crawl that was run between 12 and 16 March 2015. To find the websites containing social plugins, 100 most visited sites by Belgian web users were crawled and 5 random links within the same domain were clicked.

The crawl data was captured using *mitmproxy* and contained all the HTTP headers and bodies. The browser (Firefox) was automated using a Python script based on *Selenium*[65] browser automation framework. The data was collected on a Ubuntu 14.10 (64-bit) computer and analyzed using the *libmproxy*[66] Python library. The crawling computer was connected to the Internet using a standard consumer ISP with no proxies. Of the 100 popular sites, only 95 could be visited without a problem, the remaining 5 sites was either not accessible or a software error has occurred.

Only in one case, we included a site (microsoft.com) that was not originally in the top-100 list but redirected from a top-100 site (live.com) which belongs to the same company (Microsoft Corporation). This is due a link clicked by the crawler which redirected the browser from live.com to microsoft.com domain.

The 34 websites that contained Facebook social plug-ins are listed below:

1. 2dehands.be
2. 2ememain.be
3. 7sur7.be
4. 9gag.com
5. aliexpress.com
6. allocine.fr
7. amazon.fr
8. commentcamarche.net
9. deredactie.be
10. dhnet.be
11. diply.com
12. gva.be
13. hln.be
14. ikea.com
15. imdb.com
16. imgur.com
17. knack.be
18. kuleuven.be
19. lalibre.be
20. lameuse.be
21. lavenir.net
22. lesoir.be
23. levif.be
24. microsoft.com *(linked from live.com)*
25. msn.com
26. nieuwsblad.be
27. pinterest.com
28. reimageplus.com
29. rtbf.be
30. rtl.be
31. skynet.be
32. standaard.be
33. what-character-are-you.com
34. zalando.be

[65] http://www.seleniumhq.org/
[66] https://mitmproxy.org/doc/scripting/libmproxy.html