Do-Not-Track Techniques for Browsers and Their Implications for Consumers

Martin Beck, Michael Marhöfer

Do-Not-Track techniques for browsers and their implications for consumers

Martin Beck[1] and Michael Marhöfer[2]

[1] Technische Universität Dresden, martin.beck1@tu-dresden.de
[2] Nokia Siemens Networks GmbH & Co. KG, michael.marhoefer@nsn.com

Abstract. Recent efforts to increase online privacy by offering the user a more general choice to opt out of online tracking were mainly pushed by the FTC in late 2010. As the FTC explicitly omitted technical details, browser developers started to implement what they thought might be appropriate to either limit user tracking directly, or let the advertiser know about the user's wish to not be tracked. This paper gives a short overview on the positions and arguments of stakeholders and evaluates the technical proposals and implementations aiming to support the consumer in keeping control over his personal data.

1 Introduction

Since the Internet Architecture Board (IAB) held its "Internet Privacy Workshop" in December 2010 [1] and the Federal Trade Commission (FTC) released its preliminary staff report "Protecting Consumer Privacy in an Era of Rapid Change" also in December 2010 [12], many discussions and comments from all parties involved led to a huge media interest about online privacy, resulting in the "Online Privacy: Towards Informational Self-Determination on the Internet" workshop at Dagstuhl [2], in a workshop at the world wide web consortium (W3C) [4], drafts for standardization of tracking protection features [20, 28] and even proposed bills [16, 19, 23].

Earlier on in November 25th, 2009 the European Parliament passed a directive to regulate storage of and access to information on the equipment of a user [24]. Commonly used HTTP-Cookies are only one example of such a storage mechanism. The consumer needs to be provided with "comprehensive" information about any activity that results in the described behaviour and the offer to refuse from engaging in this activity [24]. So a user must first give his consent[6], or the corresponding data processing would be unlawful. The time for transforming this Directive into national law was set to one and a half year, so on May 25th, 2011 all members of the European Union should have finished a corresponding local law. The only country that adopted the european directive in time is the UK [3], giving organisations and businesses one year to comply with the new law [22].

Throughout all these discussions the common synonym "Do Not Track" (DNT) is used with very different meanings. Development of unique definitions

for "tracking" and "privacy" is ongoing and undertaken, for instance, at the beforementioned W3C workshop. The expected results of "Do Not Track" and proper technical solutions are also to be discussed and defined. This paper gives a short overview on the positions and arguments of stakeholders and evaluates the technical proposals and implementations aiming to support the consumer in keeping control over his personal data.

Tracking is used in this document as identifying a consumer by a specific ID or some sort of fingerprinting and by getting to know his current action. If tracking is done by a third party not being the domain or party the user intentionally wanted to communicate with, this is called *third-party tracking*, as opposed to *first-party tracking* from the intentionally visited domain.

The paper will be structured as follows. In section 2 groups composed by similar interests will be observed regarding their preferences and arguments within the overall DNT debate. Section 3 covers proposed solutions by browser developers and finally Section 4 will conclude this paper.

2 Stakeholders

Restricting and regulating online tracking affects a wide range of parties, represented by stakeholders from the advertising industry, privacy advocates, publishers, browser developers and researchers. As a result, contrary goals need to be integrated to arrive at a commonly acceptable solution.

An often named application for online tracking is *behavioral targeted advertising*, which uses aggregated data collected through online tracking to show adverts for the extrapolated interests of a certain user. According to [5], behavioral targeted ads in 2009 had a 2.68 times higher average price compared to standard "Run of Network" ads, while in the same year the average conversion rate increased from 2.8% to 6.8% by a factor of 2.43. The total ad revenue for the participants of the study in 2009 was approximated to $3.3 billion, with 17.9% attributable to behavioral targeting.

Opposing the "Commercial Privacy Bill of Rights Act of 2011" [16], Google, Facebook, Yahoo, AOL and other companies sent a letter to Senator Alan Lowenthal [13], to vote against regulating privacy as proposed by the mentioned bill. Main arguments include that the regulation would be unnecessary, as current browsers already offer tracking protection and self-regulation programs "already addressing all areas of consumer privacy [...]" [13]. Internet economy and innovation would be harmed due to conflicting standards, "creating significant confusion and uncertainty for investors" [13].

From the consumer's perspective, represented by privacy advocates, pure technical solutions which are not backed up by legal rules are unlikely to reach their full potential, as advertising networks are not committed to respect the user's desired privacy policy. Companies could just ignore the implications of the deployed technical solutions without facing legal consequences. Also, without regulation and clear definitions of used terms and expected results, it remains

unclear to the consumer what the implications of using this privacy preserving technology are due to the lack of transparency. Making an educated choice would incorporate many—in the regulatory case unnecessary—constraints, such as reading lengthy privacy policies to extract the definitions of otherwise common terms like "tracking" or "privacy" on a per domain basis.

The third party next to advertisers and privacy advocates, which is directly involved, is composed of publishers, service providers and general content providers, who used targeted advertising to increase their revenue. According to the argumentation of the advertising industry, publishers highly depend on behavioral targeting to be able to keep their services running as we know them today [27]. Often this inferes that behavioral targeting is equivalent to or at least necessary for digital advertising. However, behavioural targeting only counts for one part of all available user targeted advertising methods. Targeted adverts using contextual, search, placement or social network data are not engaging in any user tracking and are thus not subject to this regulation. As behavioral advertising only had a share of about 4% of overal internet advertising revenue [26] advertising will, with high probability, continue with only a small impact, caused by the Do-Not-Track regulation [21].

Contrary to the position of advertisers and privacy advocates, browser developers did not define a common language for communicating consumer privacy whishes in the first place. Instead most large browsers got equipped with various privacy enhancing technoloies. Implications from that will follow in section 3.

Within the W3C, a tracking protection working group was formed[3], which tries to define methods for expression of user preferences and thus to improve user control and privacy. The final recommendation compiled by this working group should be available in June, 2012. The current Draft [8] already proposes many technical details regarding the DNT protocol and answers questions and issues regarding the underlying model.

3 Proposed Technical Solutions

When looking at the technical aspects of the - advertising-industry-preferred - self-regulation, which has been evolved over the past years, the same technology that mainly started tracking is used to limit it—HTTP Cookies. These are now called *Opt-Out Cookies* [7]. Consumers may go to specific web sites[4], which offer a list of supported advertising companies allowing to opt-out of behavioral advertising for these.

Several problems arise with such a system.

[3] http://www.w3.org/2011/tracking-protection/

[4] Like: http://www.aboutads.info/choices/, http://www.networkadvertising.org/managing/opt_out.asp

Pros

- Can be used directly, without any client software updates
- Is already available and supported by some companies
- The user can check if new Cookies are being installed and existing ones still used
- Advertiser gets to know what the user wants and what his tracking preferences are

Cons

- User must visit special web site, giving a list of networks to opt-out from
- "Opt-Out" is not equally defined across all advertisers, data collection may continue or not [17]
- Opt-Out Cookies will be deleted together with normal Cookies
- User needs to maintain the list of Opt-Out Cookies through these websites and check for new companies
- Most tracking information is still sent to the advertiser, also handing over control for this data

So, in its current design, self-regulation does not offer an appropriate privacy solution to customers. One of the downside aspects, namely Opt-Out Cookies being deleted together with all other Cookies, can be prevented by using a browser plugin for making such Cookies permanent. Google released such an extension called "Keep my Opt-Outs"(KMOO) [15] for "Google Chrome", which takes a copy of the NAI[5] consumer opt out registry[6] to reinstall all Cookies described within that list each time they got deleted. The user cannot choose to allow specific companies, add Cookies or update the registry. Instead updates to the KMOO extension are required to reflect changes made to the original NAI registry, or the consumer would have to manually edit this file, respectively replace it by a more current version. A similar plugin for Mozilla Firefox is "Taco Beef", which also makes Opt-Out Cookies permanent.

Another proposed solution—*DNT-Header*—comes from Mozilla and describes an extension to the sent HTTP header, which will carry a new entry called "DNT" [20]. Upon reception of this flag, carrying the value 1, the receiving server gets to know that the user does not want to be tracked. What this means is still to be defined. Next to the Mozilla Firefox browser[7], the Microsoft Internet Explorer also added support for this header since version 9. Apple's Safari browser will be equipped with a compatible header in an upcoming version [9]. Again we shall look at the arguments surrounding this technology.

[5] Network Advertising Initiative, http://www.networkadvertising.org/
[6] located at: http://networkadvertising.org/optoutprotector/registry.json
[7] since Mozilla Firefox version 4.0

Pros

- Generic broadcast solution that can be set once within the browser and forbids tracking on all supported networks
- Good usability, as the consumer does not need to maintain a list or keep track of updates
- Tracking preference is conveyed to advertising company

Cons

- Updated browser needed to enable header support
- Adoption by companies just started, having nearly no practical usage right now
- Again, most tracking information is still sent to the advertiser, also handing over control for this data

Once the used terms are commonly defined, DNT support is included within the major browsers and web servers, and advertising companies respect the setting, this technical solution could provide a very convenient way to disable online tracking. As far as browser support is concerned, as Apple Safari also ships with DNT support, 84% of the browser market is ready to protect the user's privacy through this technology [9]. Also support from some companies was announced [10], giving initial usage to the DNT-Header approach.

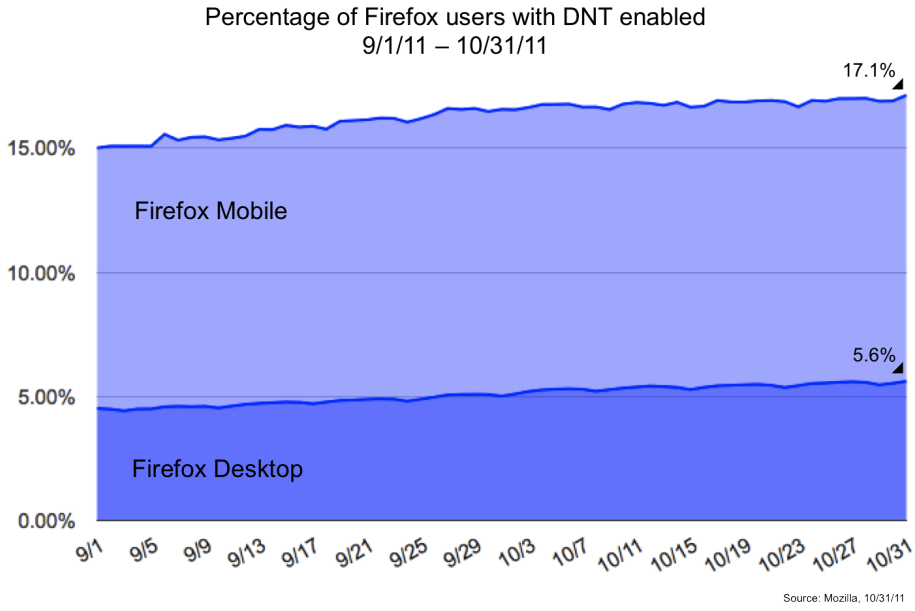

Fig. 1. DNT adoption on desktop and mobile devices

Figure 1 shows the adoption of the DNT option in Mozilla Firefox on desktop and mobile devices throughout September and October in 2011.

As a third proposed way to protect the privacy of customers, Microsoft implemented black and white lists within the latest version of its "Internet Explorer" browser[28]. The scope for checking URLs against the lists is always on third party domains, implying that the originally visited domain (first party) and all its sub-domains cannot be subject to tracking protection. These lists, called "*Tracking Protection Lists*" (TPL), are implemented by three complementing mechanisms, which are:

- A personalized, semi-automatic list
- External lists by third party providers
- A tracking protection exception list

The personalized list gets filled automatically based on a user-defined threshold, specifying how often a single third party element must at least be present across different domains to get included onto the list. Once an element made it onto the personalized list, the user can choose to manually select if this item should be black or white listed with allowance being the default, or automatically blacklist all entries. The only possibility a user has to influence what gets included is through the beforementioned threshold. The list cannot be edited easily to add or remove elements.

External lists are provided by third parties to allow users to benefit from already approved list entries without the need to build up lists on their own. Microsoft officially linked[8] to four such list providers upon the release of their Internet Explorer v9. Two of these lists are the EasyPrivacy list, known from the AdBlock Plus project, that also compiles the EasyList for blocking adverts, and TRUSTe tracking protection list. While "tracking protection" would imply that the user is to be protected against tracking and thus actually less to no tracking would take place after such a list would be installed, the TRUSTe list actually whitelists companies marked as trustworthy to TRUSTe.

Another problem can also be described based on the mentioned white list: Allow rules always have higher priority than Block rules. These rules are applied to each requested element that is referenced within the visited website and resides in a third party domain. So once a blocking rule is selected, the request is dropped and no data is transferred to the third party host. Those are the main arguments in favor and against blocking lists and TPL especially:

Pros

- All data and thus control over personal data remains at the user side
- Established and known lists exist which can be used immediately (EasyPrivacy from AdBlock Plus)
- As long as a user relies on external lists, no manual maintenance is needed

Cons

- Blocking lists are not tailored to tracking, but may block everything, including adverts or functionality
- May break functionality if wrong scripts are blocked
- Advertiser does not know that a user wants to opt out of tracking

[8] http://ie.microsoft.com/testdrive/Browser/TrackingProtectionLists/Default.html

External lists may include a number for specifying a time-to-live interval, after which the list should be updated automatically, which is typically in the range of three to five days. During this period, updates by the list providers are not pushed to the consumers.

The proposed standard [28] submitted by Microsoft also includes the DNT-Header introduced by Mozilla together with a new DOM element "doNotTrack" under "document.navigator". This element can be accessed by Javascript to offer scripts the possibility to detect the user's privacy preferences for the current site. Checking this element should return 1 in case tracking protection is set for this domain.

However, since the given priority calculation in the W3C proposal [28] differs from the implementation of the "Internet Explorer 9.0", which again differs from the "Internet Explorer 9.0 Release Candidate 1" implementation, complex blocklists may produce unexpected results.

4 Conclusion

As very different technical solutions are developed, many different kinds of advantages and drawbacks were recognized. Overall, the solutions can be categorized into either letting the advertiser know that tracking is not desired, or blocking requests and thereby limiting information flow to the advertisers.

A combination of both techniques enhances the privacy protection even further, as non-trustworthy companies which are not responding to DNT requests could be blocked, while other networks which respect the privacy setting can be allowed at the same time. Another aspect of online tracking is largely ignored: passive fingerprinting for (re-)identification. The EFF[9] hosts a web site which tries to measure the amount of information a user's browser gives away[10]. In order to at least disturb the companies who use unique combinations of transmitted information by the browser to build profiles on top of these fingerprint IDs, they could be confronted with a less unique appearance. More generic values within the HTTP header, sorted font lists and alike can help to increase the anonymization set. Such measures can be implemented by using a specific Firefox profile like "JonDoFox"[11] or equivalent.

As far as current usability for web-based opt-out solution, built in browser DNT options, blocking lists or browser plug-ins goes [18], non-technical consumers face a rather long list of issues. The selection of companies within opt-out lists is not based on a meaningful per-company decision by the user, but rather by excluding a few common company names or just disabling traffic to all available companies. Default options are always set to explizitly opt-out of tracking, even after downloading and installing tracking protection plugins like TACO and Ghostery, which implies that a user most probably wants to stop companies

[9] Electronic Frontier Foundation, https://www.eff.org/
[10] https://panopticlick.eff.org/
[11] http://anonymous-proxy-servers.net/en/jondofox.html

tracking his online behaviour. The same is true for the IE9, where a user is not supported in adding third party lists after enabling tracking protection.

The implemented interfaces for configuring the opt-out are either too simplistic, so the user is not well informed (IE9 slider), or contain too many technical terms to be understood (Ghostery). As such, these tools are ineffective at communicating their purpose and guiding users [18]. Interfaces are further confusing to the user, as it wasn't clear if tracking protection, for those companies intended, was already enabled, or even that third party TPL lists could be added to the IE9 protection feature. This is partly influenced by a lack of feedback for the tested tools, which leaves the user in the unknown state, whether tracking protection works or not [18].

The average user at the end will most probably not notice any difference at all. The block list feature of the Microsoft Internet Explorer would have the highest impact, which could block third party adverts all together. The notification mechanisms of DNT-Header, DOM elements or Opt-Out Cookies will still let advertisers deliver ads, but hopefully not tailored to that user.

As long as targeted advertising is of main interest, several solutions for performing privacy-preserving targeting are proposed and developed [25, 14, 11], focusing around stopping the unwanted flow of personal information to companies within the online advertising ecosystem. These solutions have local information extraction in common, which allows data mining against much of the local available information to generate highly probable interests for the consumer. Some of these systems even allow privacy-preserving gathering of correct adverts. Another point for discussion within such systems is whether the user would allow advertisers to get relevant personal information in case the consumer is able to give and actually gave informed consent.

Bibliography

[1] Internet privacy workshop - how can technology help to improve privacy on the internet?, December 2010. URL http://www.iab.org/about/workshops/privacy/.

[2] Perspectives workshop: Online privacy: Towards informational self-determination on the internet, February 2011. URL http://www.dagstuhl.de/11061.

[3] The privacy and electronic communications (ec directive) (amendment) regulations 2011, May 2011. URL http://www.legislation.gov.uk/uksi/2011/1208/made/data.pdf.

[4] W3c workshop on web tracking and user privacy, April 2011. URL http://www.w3.org/2011/track-privacy/.

[5] Howard Beales. The value of behavioral targeting. Technical report, Network Advertising Initiative, March 2010.

[6] Data Protection Working Party. Opinion 15/2011 on the definition of consent, July 2011. URL http://www.statewatch.org/news/2011/jul/eu-art-29-wp187-consent.pdf.

[7] Pam Dixon. THE NETWORK ADVERTISING INITIATIVE: Failing at Consumer Protection and at Self-Regulation. Technical report, World Privacy Forum, November 2007. URL http://www.worldprivacyforum.org/pdf/WPF_NAI_report_Nov2_2007fs.pdf.

[8] Roy T. Fielding and Adobe. Tracking preference expression (dnt), November 2011. URL http://www.w3.org/2011/tracking-protection/drafts/tracking-dnt.html.

[9] Chris Foresman. Safari to join "do not track" crowd, leaving google behind, April 2011. URL http://arstechnica.com/apple/news/2011/04/safari-to-gain-do-not-track-support-in-lion.ars.

[10] Alex Fowler. Advertisers and publishers adopt and implement do not track, March 2011. URL http://blog.mozilla.com/blog/2011/03/30/advertisers-and-publishers-adopt-and-implement-do-not-track/.

[11] Matthew Fredrikson and Benjamin Livshits. RePriv: Re-Imagining Content Personalization and In-Browser Privacy. In *IEEE Symposium on Security and Privacy*, May 2011.

[12] FTC. Protecting consumer privacy in an era of rapid change. Technical report, Federal Trade Commission, December 2010. URL http://www.ftc.gov/os/2010/12/101201privacyreport.pdf.

[13] Google, Facebook, Yahoo, and AOL. Sb 761 (lowenthal) — opposition, April 2011. URL http://static.arstechnica.com/oppositionletter.pdf.

[14] Saikat Guha, Bin Cheng, and Paul Francis. Privad: practical privacy in online advertising. In *Proceedings of the 8th USENIX conference on Networked systems design and implementation*, NSDI'11, pages 13–13, Berkeley, CA, USA, March 2011. USENIX Association. URL http://portal.acm.org/citation.cfm?id=1972457.1972475.

[15] Sean Harvey and Rajas Moonka. Keep your opt-outs, January 2011. URL http://googlepublicpolicy.blogspot.com/2011/01/keep-your-opt-outs.html.

[16] John Kerry and John McCain. Commercial privacy bill of rights act of 2011, April 2011.

[17] Saranga Komanduri, Richard Shay, Greg Norcie, and Lorrie Faith Cranor. Adchoices? compliance with online behavioral advertising notice and choice requirements. Technical Report 82, Carnegie Mellon CyLab, March 2011.

[18] Pedro G. Leon, Blase Ur, Rebecca Balebako, Lorrie Faith Cranor, Richard Shay, and Yang Wang. Why johnny can't opt out: A usability evaluation of tools to limit online behavioral advertising. Technical report, Carnegie Mellon University - CyLab, October 2011.

[19] Edward J. Markey and Joe Barton. Do not track kids act of 2011, May 2011. URL http://online.wsj.com/public/resources/documents/billdraft0506.pdf.

[20] J. Mayer, A. Narayanan, and Sid Stamm. Do not track: A universal third-party web tracking opt out draft-mayer-do-not-track-00. Technical report, Internet Engineering Task Force - Network Working Group, March 2011. URL http://datatracker.ietf.org/doc/draft-mayer-do-not-track/.

[21] Jonathan Mayer. Do not track is no threat to ad-supported businesses, January 2011. URL http://cyberlaw.stanford.edu/node/6592.

[22] Information Commisioner's Office. Ico gives website owners one year to comply with cookies law, May 2011. URL http://www.ico.gov.uk/~/media/documents/pressreleases/2011/enforcement_cookies_rules_news_release_20110525.pdf.

[23] Jay Rockefeller. Do-not-track online act of 2011, May 2011. URL http://commerce.senate.gov/public/?a=Files.Serve&File_id=85b45cce-63b3-4241-99f1-0bc57c5c1cff.

[24] The European Parliament and the Council of the EU. DIRECTIVE 2009/136/EC. *Official Journal of the European Union*, November 2009.

[25] Vincent Toubiana, Helen Nissenbaum, Arvind Narayanan, Solon Barocas, and Dan Boneh. Adnostic: Privacy preserving targeted advertising. *Network and Distributed System Security Symposium*, 2010.

[26] Henry A. Waxman and Joe Barton. Memorandum. Technical report, Congress of the United States, House of Representatives, Committee on Energy and Commerce, Washington, DC, USA, November 2010. URL http://democrats.energycommerce.house.gov/documents/20101201/Briefing.Memo.12.01.2010.pdf.

[27] Michael Zaneis. 'do not track' rules would put a stop to the internet as we know it, January 2011. URL http://www.usnews.com/opinion/articles/2011/01/03/do-not-track-rules-would-put-a-stop-to-the-internet-as-we-know-it.

[28] Andy Zeigler, Adrian Bateman, and Eliot Graff. Web tracking protection, February 2011. URL http://www.w3.org/Submission/2011/SUBM-web-tracking-protection-20110224/.