The Web is Watching You: A Comprehensive Review of Web Tracking Techniques and Countermeasures

Iskander Sanchez-Rola[1], Xabier Ugarte-Pedrero[2], Igor Santos[1], and Pablo G. Bringas[1]

[1]DeustoTech, University of Deusto
{iskander.sanchez,isantos,pablo.garcia.bringas}@deusto.es

[2]Cisco Talos Security and Research Intelligence Research Group
{xabipedr}@cisco.com

Abstract. Web tracking is a commonly-used practice on the Internet devoted to retrieve user information for activities such as personalization or advertisement. These techniques are said to drive the web economy, although they are commonly used to invade users' privacy. In the last years, a general concern raised about web tracking, looking forward to combat it in many ways like regulations, anti-tracking methods and even standardization. In this paper, we analyze and discuss the current techniques for web-tracking as well as techniques for its detection and analysis, and countermeasures to prevent web tracking.

Keywords: web privacy, web tracking, web security

1 Introduction

Web tracking is a widespread technique on the Internet that gathers user data to perform online advertisement, content personalization, or user authentication. In general, web tracking allows third-party or first-party websites to know the users' browsing history and browsing configuration to these ends. These techniques can be used to improve users' experience and to enhance their browsing on the Internet. However, since they sometimes involve retrieving user data, even without their consent, they can be considered a privacy violation by itself in some occasions. Web advertising companies refuse to acknowledge web tracking as a threat and they even publicly defend web tracking on the basis of its relevance in the Internet economy [1]. Web-tracking techniques can be of stateful or stateless nature, depending on whether or not they require data to be stored in user's computer to properly function. It is usually considered that stateless tracking methods are harder to limit and block because they easily bypass common countermeasures against tracking such as private browsing or removing cookies [2].

Recent work [2–6] has studied the prevalence of different types of web tracking and fingerprinting. These works have found that web tracking is a very usual

practice and that not only it is based on its less damaging forms as web analytics but also advanced fingerprinting techniques such as font probing or canvas fingerprinting.

Several countermeasures against massive web tracking have emerged in the literature. An important proposal to limit web tracking is the initiative *Do Not Track* [7] promoted by Mayer et al. that allows users to express their willingness to avoid being tracked. Another line of work is the use of custom browsers. For example, *Privaricator* [8] that subverts the linkability by adopting randomization policies or *TrackingFree* [9] that blocks the storage and communication stage of the tracking process.

In this paper, we comprehensively review the web-tracking research topic, extending our previous review [10]. We study this topic from a web-security research perspective, describing both web-tracking techniques and defenses. The goal is to bring understanding of the current state of the art in web tracking to allow a better understanding of its landscape and the proper discussion of its implications and future research trends. This review, is to the best of our knowledge the first one that reviews multiple web tracking techniques, countermeasures and legislation.

2 Web Tracking Applications

Albeit some web tracking techniques are privacy-invasive and raise a serious concern for users' privacy, they are an extended practice in the Internet. In this section, we review the different applications of tracking in the web and we detail non-privacy-invasive alternatives that can be used for these applications.

2.1 Types of Web Tracking

Roesner et al. [4] defined a taxonomy of third-party cookie-based web tracking that is currently the most common form of tracking. For this, they focused on two aspects: the functionality of the tracker and the type of behavior. Regarding the functionality, the following categories were defined: (i) third-party advertisement, (ii)third-party advertisement with pop ups, (iii)third-party advertisement networks, and (iv) third-party social widgets. With regards to the type of behaviour, the authors defined defined the following ones:

- **Behavior A (Analytics).** The tracker serves an analytics engine for each website. Only tracks users within that specific site
- **Behavior B (Vanilla)**.The tracker exhibits third-party storage that can be get and set only from a third-party position to track users across sites.
- **Behavior C (Forced)**. The cross-site tracker forces users to visit its domain directly (such as popups or redirections), placing it in a first-party position.
- **Behavior D (Referred).** The tracker relies on a type B, C, or E tracker to leak unique identifiers, rather than on its own client, to track users across sites.

– **Behavior E (Personal)**. The cross-site tracker is visited by the user directly in other contexts.

These last categories are not mutually exclusive with the exception of B and E, because the user can only visit the web tracker's domain. In their study in 2012, they found out that in the top visited 500 webpages in the well-known Alexa ranking, the most common tracker type was the B (Vanilla) one. With these categories, we can understand the different levels of danger that web trackers pose depending on their use of data. In addition to these categories, advanced fingerprinting is an stateless tracking technique that bypasses user's browsing configuration as we will be detail later.

2.2 Advertising and Analytics Services

The most well-known and extended use of web tracking and arguably the main application of it, are web advertisement and analytics. Web tracking is used in analytics to track the visitors of a particular website, its articles and sub-pages. In addition other information and demographics such as browser version or operating system can also be obtained [11]. These user profiles are stored and updated as the user interacts with websites. In advertisement, the browsing history of a particular user is retrieved by gathering the user identifier from the websites that host the particular advertisement network, knowing which websites she has visited. Some of the aforementioned techniques can be used to enhance user identification and allow to share the browsing history among different advertisement companies. Based upon the user profiling, the advertisement network will also update its advertisement placement specially targeting to the user visiting the website in its network.

Although these implementations may variate, nearly every vendor has adopted one of two typical models. Some offer analytics as a paid service and they cannot utilize the client's analytics information, protecting the obtained information. Other vendors offer a free analytics service, but they use the obtained data for ad targeting, market understanding, and so on. Advertising companies do not always depend on the data sold by the analytics services. They use their own techniques to categorize the user. Information transference between banners is one of the most employed techniques [4]. In addition to pre-packaged solutions, some websites use their own implementations [12]. They are sometimes even obfuscated to evade detection systems. As these methods do not follow any specific information flow, they are much more difficult to detect and stop.

2.3 Alternatives to Privacy-Invasive Tracking

Advertising companies have affirmed that web tracking is required for maintaining web economy [1]. However, other techniques have been proposed for analytics and targeting preserving users' privacy. *Adnostic* [13] was proposed by Toubiana et al. to overcome the privacy issues of online behavioral advertisement. They proposed a practical architecture to performing the user targeting directly in the

user browser instead of sending any information to third parties. *PrivAd* [14] implements a very similar approach but utilizes a trusted party to anonymize the client. *RePriv* [15] uses the browser to allow interest profiling and generically enables personalization. Bilenko et al. presented a technique [16] to store the user profile and recent browsing history in a cookie without the intervention of any third-party. *ObliviAd* [17] is an approach for privacy-preserving online behavioral advertisement that employs secure hardware-based private information gathering for distribution of advertisements. All these approaches can provide a continuation to web advertisement without violating users' privacy.

3 Web Tracking Techniques

3.1 Stateful Tracking

This type of web tracking techniques use the available methods to store information within the client's computer. To this end, third-party websites access different aspects of the websites to retrieve user data and to store it.

Lou Montulli proposed *cookies* mainly to allow users' stateful web navigation [18]. Cookie usage permits a website to store data on the users' computers that is retrieved when the user returns to the site. In this way, websites can maintain the navigation state that otherwise would be lost and thus, increase the usability.

Although they were not designed for it, the abuse of their stateful nature started shortly after their first appearance. Since websites are composed of different resources that may be allocated in the web hosting, the main page and also in third-party servers, these external resource providers have the capability of including cookies along with their provided resource. If the third-party server provides resources to a large number of other websites that this particular user visits, it can certainly gather user's browsing history and profile her browsing habits. For example, a website `site.com` includes an image from a third-party domain called `advertisement.com`. The server `advertisement.com` sends the image along with a `HTTP Set-Cookie` header, that will be stored on her machine. The user visits another website `site-two.com` that also utilizes images from `advertisement.com`, when asking for the image `site-two.com`. It will send user's previously set cookie to `advertisement.com`. It will recognize the user and start tracking her. This behavior is called *third-party cookies* and it is arguably the most extended technique for web tracking.

The functionality of HTTP cookies was later adopted by other components of the browser like Flash Cookies [19] (also named Local Shared Objects or LSO), HTML5 [20], or JavaScript. Flash cookies store more data, they lack an expiration date, and they are not controlled by the browser. HTML5 allows websites and third-party trackers to store information in the browser without setting a cookie. The `Window.name` is a non-persistent property of JavaScript that can be used to share data between different websites.

Third-party cookies raised the concern of the research community [4, 5, 21–23] that was also extended to popular media and the general public. As a result,

the community responded in several ways. For example, comScore (a popular online advertisement company) presented a study [24] that showed that 1 out of 3 users removed first and third party cookies within a month. Extensions also appeared to block third-party tracking [25, 26] and to visualize cookie sharing [27]. Finally, the currently widespread *Private Browsing mode* was born to allow users to navigate the web without leaving any trace in their local storage.

However, trackers responded with a set of techniques derivatively created to circumvent these solutions. Cookie syncing is a practice of tracking companies to bypass the Same-Origin Policy and allow different trackers to use the same user identifiers and share them. This technique synchronizes cookies among trackers by passing user identifications. Since each website cannot read other cookies, cookie syncing or synchronization provides a method to, according to Google, facilitate targeting and real-time bidding [4]. *Cookie respawning* and *Evercookies* intentionally abuse the storage methods of the browsers to restore the previously removed cookies. Soltani et al. [28] discovered the use of Flash cookies to regenerate removed HTTP cookies. In a later study [19], they discovered several websites using *ETags* and HTML `localStorage` API to respawn cookies. Evercookie was created by Kamkar [29] as a resistant tracking method with different storage mechanisms such as Flash cookies, `localStorage`, `sessionStorage` and *Etags*. Evercookie also employed various novel methods, working altogether to regenerate cookies.

3.2 Stateless Tracking

Stateless tracking does not require to store any information on the users' computer. This stateless device fingerprinting has become an increasingly common practice performed by advertisement and anti-fraud enterprises. This stateless techniques permit these companies to bypass the private browsing mode and also the current cookie-related regulations in Europe and the United States. In addition, these techniques allow advertisers to increase the previous gathered user data and an easier sharing of the user identifications between different tracking services. Currently, the are several features to uniquely fingerprint a device that can be gathered from different aspects of the browser such as *Javascript* or *plugins*.

JavaScript-based device fingerprinting is performed by inspecting its accessible browser resources. For example, NAVIGATOR contains data about the browser vendor and version, supported plugins and MIME types, and information about the operating system and architecture. Another object commonly used is SCREEN that contains information about the user monitor resolutions as well as the color and pixel depth. Mayer [30] experimented fingerprinting 1,328 users by hashing the contents of the JavaScript accessible browser features `navigator`, `screen`, `navigator.plugins`, and `navigator.MimeTypes`, allowing to uniquely fingerprint more than 96% of the users. Eckersley [31] extended that browser features by adding a list of the installed fonts, timezones and a browser's ACCEPT headers, combining them to create an unique device-specific identifier. To this end, *Panopticlick* [32] was developed and evaluated on around half million users,

identifying correctly the 94.2% of the half million users, demonstrating that it was possible to be identified and tracked without the need of any stateful client-storage mechanism. Eckersley also demonstrated that the list of installed fonts was the most accurate feature for device fingerprinting. In addition, the browsing history can also be gathered by abusing the JavaScripts visited-link color feature [33]. Other researchers have proposed the usage of performance benchmarks to identify the different JavaScript engines [34], raising errors of the standard tests [35] or computing the differences by elements created with the `canvas` HTML element [36].

This last technique is known as *canvas fingerprinting* and by the utilization of the Canvas API of modern web browsers, subtle differences in rendering the same text or WebGL scenes can be abused to extract a unique fingerprint. The main property that allows this technique to work so well is the fact that the same text can be rendered in different manners depending on the operating system, due to the differences in the rasterization such as anti-aliasing, hinting, sub-pixel smoothing, system fonts, API implementations or the display. The technique draws as many different letters as possible to the canvas with the intention of maximizing the diversity.

4 Web Tracking Analysis and Detection

The first techniques that were adopted to block web tracking were based on blacklists such as *EasyPrivacy* [37] and *Ghostery* [25]. These solutions contain a list of names of scripts known to perform some sort of web tracking and domains known to host web tracking. In this way, when a website and the URL of its first-party or third-party scripts is available, they are searched in the blacklists. If a script in the website is named as one in the blacklists or it is hosted in a blacklisted domain then its functionality is blocked by the blacklisting solution. Other solutions exists like *Privacy Badger* [26] or *Disconnect* [38] that utilize heuristics to determine that a third-party domain is loading and sending content to the website.

In addition to these solutions, several researchers have performed studies of web tracking. In this way, one the first studies analyzed the prevalence of HTML cookies on the Internet [39]. Mayer & Mitchell [3] analyzed the diverse web tracking methods, developing a framework to evaluate websites' privacy. Roesner et al. [4] developed a taxonomy to categorize the different web tracking methods, measuring their prevalence. Nikiforakis et al. [5] focused their study in three well-known fingerprinting companies, discovering that 40 websites out of 10,000 sites used advanced fingerprinting techniques such as font probing. Acar et al. [2] presented *FPDetective* that based on different manually derived rules, was capable of detecting fingerprinting and advanced web tracking techniques. Later, Acar et al. [6] focused on the recent canvas fingerprinting method and, by the means of manually-set rules, measured its prevalence in the Internet, finding that 5% of the websites in the top 100,000 of the Alexa ranking. They also

measured the prevalence of cookie syncing and rewspaning techniques, showing that they are extended practices.

5 Countermeasures to Web Tracking

5.1 Anti-tracking Techniques

Several browser configurations can be used in order to avoid some specific types of web tracking. Even though the most effective countermeasure is to just disable JavaScript since most of the attacks described use or depend somehow in it, it is not desirable since it affects common user browsing. A more viable option is to use the temporary modes such as private or guest mode that most current browsers implement. In this way, the browser will not save or cache any visited website or downloaded file [40] and will avoid classic types of cookies. Since most fingerprinting techniques retrieve specific configurations of the browsers in order to compute a unique identification for the user, another possibility for avoiding being tracked and fingerprinted is to disable certain specific font sets, disabling cookies and using domain and script blacklisting techniques. Anyhow, these countermeasures will not avoid every type of web tracking technique and advanced techniques may bypass these blocking methods. Unfortunately, disabling JavaScript would prevent some methods for web tracking but it also causes many websites to render incorrectly. Disabling other secondary features used in web tracking is a more promising approach because the number of websites that rely on them is smaller.

Spoofing a user profile can also be considered as a countermeasure against web tracking. However, the best possible spoofing configuration to avoid web tracking will require every user to share the same user profile, which renders inviable nowadays. Spoofing data and using different random profiles would help to circumvent web tracking because it hinders the uniqueness of the user's browser. Some of the properties to spoof are browser, platform, time zone or screen resolution. Nevertheless, spoofing could be counterproductive because these attempts to hide the identity can also be used for device fingerprinting [5].

There also exist fully functional anti-tracking web browsers (e.g., FlowFox [41], *TrackingFree* [9], or Privaricator [8]) that implement a precise and general information analysis and control, devoted to protect the user against web tracking. In order to limit their effect in user browsing, these privacy-aware browsers seek a very low performance overhead. In addition, there is also some work focused in the analysis of user's browsing [42]. In this way, all the accessed websites could be analyzed without exception, taking into account that the user is the weakest link in the security chain. By applying taint analysis or dynamic controls, and using several policies, it is possible to detect web tracking [43, 44]. The main problem of these methods is that they only take into account certain fields and privacy attacks. Understanding and controlling every type of privacy attack would enormously improve web-browsers. Nevertheless, the biggest disadvantage of a general control method as taint analysis, is its computational complexity.

5.2 Standardization

One possible solution to the problem of web tracking is to standardize the control of the information that is being transmitted. Two main projects have been advanced for giving users control over their personal data: *Do Not Track* (DNT) [7] and *Platform for Privacy Preferences* (P3P) [45].

Do Not Track is a proposal that combines technology and policies in order to declare user's preferences regarding web tracking. This information is sent via an HTTP header: `DNT`. All modern browsers (e.g., Chrome, Firefox, Opera, Safari, and Internet Explorer) support a Do Not Track opt-out preference (i.e., `DNT: 1 header`). This policy also indicates that websites must stop tracking the user for whatever reason when they receive a `DNT` header.

Platform for Privacy Preferences is devoted to facilitate to websites the task of communicating their privacy habits in a standard format that can be automatically obtained and understood by user agents. Users have the possibility of coming to a decision based on the privacy practices indicated by the website [46]. Thanks to that, users do not need to read the privacy policies of all the webpages they access, they just need to read it's practices. Websites implementing these policies have to make their habits public. Browsers can help the user to interpret those privacy habits with user-friendly interfaces.

Although many stakeholders (policy makers, consumer advocates and researchers) think that Do Not Track could decidedly reduce tracking and data collection on the web, as the final decision of taking it into account only resides in websites, it is not followed as expected [47]. The case of P3P is similar, due to the lack of support from current browsers for the implementation, the P3P Specification Working Group suspended the project.

5.3 Regulations

After understanding the magnitude of the problem, we should understand the existing regulations in the United States and European Union [3]. It was not until recently that these regulations were introduced in order to restrict large-scale collection of personal data [48].

In the *United States*, one of the missions of the Federal Trade Commission (FTC) is the promotion of consumer protection. They can only prevent practices of businesses that are either unfair or deceptive under 15 U.S.C. § 45. First violations will incur on a small payment, but subsequent violations get big monetary penalties. On 2012 the FTC issued its final report [49] establishing four best practices for companies to protect the privacy of all American consumers and give them the possibility to have more control of tracking options and personal information collection. The report expands on a preliminary report released in 2010 [50] which proposed a framework for consumer privacy control because of the new technologies that allow information collection that is often not perceivable by consumers. The objective is to balance the personal data of consumers with innovation.

Regarding the *European Union*, the Directive 2002/58/EC on Privacy and Electronic Communications, also known as E-Privacy Directive, indicates that the use of electronic communications networks to store information or to gain access to information stored in the terminal equipment of a user is only allowed on condition that the user concerned is provided with clear and comprehensive information about the purposes of the processing, and is offered the right to refuse such processing [51]. If the above indications are not met, penalties could be up to 2% of the revenue. The Article 29 Working Party (WP29) addresses the topic of device fingerprinting in the Opinion 9/2014 , which extends over the previous Opinion on Cookie Consent Exemption [52], and indicates that websites cannot process device fingerprints which are generated through the gaining of access to or the storing of information on the users terminal device if there is not a explicit consent of the user (unless some specific exemptions) [53].

There have also been attemps of *self-regulation*. In 2009, many of the largest advertising and marketing companies and associations, supported by the Council of Better Business Bureaus, created a self-regulatory program with the principal objective of giving total control over the collection and use of private data to the users [54]. Websites should have clear options regarding to the data collection and use, letting the user decide if they want that collection or not. There should also be a limit on the specific data type obtained if it is sensitive information. Until that moment, all the different actors worked interdependently in this area. Nevertheless, it is only indicated for the data collection used to predict user interests to deliver online advertising. These principles do not apply to websites that collect that information for its own uses. 2 years later, the Digital Advertising Alliance (DAA) announce an expansion of the program in order to include the non-advertising businesses to the self-regulation [55]. These new principles prohibit third parties to collect, use or transfer any multi-site information. However, these data was mostly covered in the areas of insurance, credit, employment or health.

Despite the fact that many regulations exist, there is not a continuous control of the websites to check if they are actually following them. Creating a organization responsible for this would secure the compliance of regulations and therefore improve the privacy control of the users.

6 Discussion and Concluding Remarks

Since the introduction of cookies, web tracking has evolved along with the techniques to evade it. New advanced stateless device fingerprinting techniques have been proposed and adopted by trackers, including well-known advertisement and analytics companies, as recent work discovered [2, 6]. These advanced methods not only avoid the necessity of storing on the client's machine but also, since they are not contemplated in current regulations against web tracking, they are able to operate with no regulation even though its purpose is the same as cookies.

Due to the proliferation and the awareness that recent work have raised in the community, new techniques for detection of web tracking behavior have been

proposed. Besides common blacklists, that can be subverted easily, heuristics [26] and rule-based systems [4, 5, 2, 6] have also been proposed by the community. These methods provide a way to detect these web tracking techniques and block them if necessary. These methods have been primarily utilized to perform studies about the current ecosystem of web tracking in the Internet, raising a general concern about the use of these methods. However, they can be adopted in the future in order to be used as complement of current blacklisting methods to ensure reactive protection of users against web tracking methods.

More proactive methods have also been proposed such as custom browsers, standardization, or regulation. In web browsers, even if randomizating the parameters that these techniques usually employ for generating the unique identifiers is a proposal, if detected, may also be used by web trackers as a parameter for generating an unique fingerprint. Therefore, other methods should be explored for a complete user protection against web tracking. Standardization and regulation policies are an important effort to fight against the generalization of web tracking and also the users' lack of knowledge about being tracked and fingerprinted. Even though they are important steps against tracking — and online surveillance in general — it should be complemented because new and more advanced tracking techniques may be developed, bypassing the restrictions imposed by laws or standards.

In general, web tracking is a widespread technique that despite it can be used to enhance user experience and even its security, it violates user's privacy, sometimes without her explicit knowledge and consent. New techniques have been developed that make fingerprinting and tracking easier and more difficult to combat. In addition, its emerging creation of targeted malware campaigns (as the cases reported by Symantec [56] and FireEye [57] this last year) link directly this technique to the creation of targeted and personalized malicious software, raising a relevant problem not only for user privacy but also in the field of malware analysis that has to cope with more complex samples.

Acknowledgments This research was partially supported by the Basque Government under the pre-doctoral grant given to Iskander Sanchez-Rola.

References

1. Singer, N.: Do not track? advertisers say dont tread on us. The New York Times (October 13, 2012) `http://www.nytimes.com/2012/10/14/technology/do-not-track-movement-is-drawing-advertisers-fire.html`.
2. Acar, G., Juarez, M., Nikiforakis, N., Diaz, C., Gürses, S., Piessens, F., Preneel, B.: FPDetective: dusting the web for fingerprinters. In: Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS). (2013)
3. Mayer, J.R., Mitchell, J.C.: Third-party web tracking: Policy and technology. In: Proceedings of the International Symposium on Security and Privacy (Oakland). (2012)
4. Roesner, F., Kohno, T., Wetherall, D.: Detecting and defending against third-party tracking on the web. In: Proceedings of the USENIX conference on Networked Systems Design and Implementation (NDSI). (2012)

5. Nikiforakis, N., Kapravelos, A., Joosen, W., Kruegel, C., Piessens, F., Vigna, G.: Cookieless monster: Exploring the ecosystem of web-based device fingerprinting. In: Proceedings of IEEE Symposium on Security and Privacy (Oakland). (2013)
6. Acar, G., Eubank, C., Englehardt, S., Juarez, M., Narayanan, A., Diaz, C.: The web never forgets: Persistent tracking mechanisms in the wild. In: Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS). (2014)
7. Mayer, J., Narayanan, A., Stamm, S.: Do not track: A universal third-party web tracking opt out. IETF Request for Comments (2011)
8. Nikiforakis, N., Joosen, W., Livshits, B.: Privaricator: Deceiving fingerprinters with little white lies. In: Proceedings of the International Conference on World Wide Web (WWW). (2015)
9. Pan, X., Cao, Y., Chen, Y.: I Do Not Know What You Visited Last Summer: Protecting Users from Third-party Web Tracking with TrackingFree Browser. In: Proceedings of the Annual Network and Distributed System Security Symposium (NDSS). (2015)
10. Sánchez-Rola, I., Ugarte-Pedrero, X., Santos, I., Bringas, P.G.: Tracking users like there is no tomorrow: Privacy on the current internet. In: Proceedings of the International Conference on Computational Intelligence in Security for Information Systems (CISIS), Springer (2015)
11. Altaweel, I., Cabrera, J., Choi, H.S., Ho, K., Good, N., Hoofnagle, C.: Web privacy census: Html5 storage takes the spotlight as flash returns
12. Jang, D., Jhala, R., Lerner, S., Shacham, H.: An empirical study of privacy-violating information flows in javascript web applications. In: Proceedings of the 17th ACM conference on Computer and communications security, ACM (2010) 270–283
13. Toubiana, V., Narayanan, A., Boneh, D., Nissenbaum, H., Barocas, S.: Adnostic: Privacy preserving targeted advertising. In: Proceedings of the Network and Distributed System Symposium (NDSS). (2010)
14. Guha, S., Cheng, B., Francis, P.: Privad: practical privacy in online advertising. In: Proceedings of the USENIX conference on Networked Systems Design and Implementation (NDSI). (2011)
15. Fredrikson, M., Livshits, B.: Repriv: Re-imagining content personalization and in-browser privacy. In: Proceedings of the IEEE Symposium on Security and Privacy (Oakland). (2011)
16. Bilenko, M., Richardson, M., Tsai, J.: Targeted, not tracked: Client-side solutions for privacy-friendly behavioral advertising. In: Proceedings of the Privacy Enhancing Technologies (PETS). (2011)
17. Backes, M., Kate, A., Maffei, M., Pecina, K.: Obliviad: Provably secure and practical online behavioral advertising. In: Proceedings of the IEEE Symposium on Security and Privacy (Oakland),. (2012)
18. Schwartz, J.: Giving the web a memory cost its users privacy. `http://www.nytimes.com/2001/09/04/technology/04COOK.html` (2001)
19. Ayenson, M., Wambach, D., Soltani, A., Good, N., Hoofnagle, C.: Flash cookies and privacy II: Now with HTML5 and Etags respawning (2011). Social Science Research Network Working Paper Series (2011)
20. West, W., Pulimood, S.M.: Analysis of privacy and security in html5 web storage. Journal of Computing Sciences in Colleges **27**(3) (2012) 80–87
21. Krishnamurthy, B., Wills, C.E.: Generating a privacy footprint on the internet. In: Proceedings of the 6th ACM SIGCOMM conference on Internet measurement, ACM (2006) 65–70

22. Krishnamurthy, B.: Privacy leakage on the internet (2010)
23. Englehardt, S., Reisman, D., Eubank, C., Zimmerman, P., Mayer, J., Narayanan, A., Felten, E.W.: Cookies that give you away: The surveillance implications of web tracking. In: Proceedings of the International Conference on World Wide Web. (2015)
24. comScore: The Impact of Cookie Deletion on Site-Server and Ad-Server Metrics in Australia (2011)
25. Ghostery. `http:wwww.ghostery.com`
26. EFF: Privacy Badger. `https://www.eff.org/es/node/73969`
27. Atul Varma: Lightbeam: Discover whos tracking you online. `http://www.mozilla.org/en-US/lightbeam/`
28. Soltani, A., Canty, S., Mayo, Q., Thomas, L., Hoofnagle, C.J.: Flash cookies and privacy. In: AAAI Spring Symposium: Intelligent Information Privacy Management. Volume 2010. (2010) 158–163
29. Kamkar, S.: Evercookie-virtually irrevocable persistent cookies. `http://samy.pl/evercookie/` (2010)
30. Mayer, J.R.: Any person... a pamphleteer: Internet anonymity in the age of web 2.0 (2009) Senior Thesis, Stanford University.
31. Eckersley, P.: How unique is your web browser? In: Proceedings of the Privacy Enhancing Technologies (PETS), Springer (2010)
32. Electronic Frontier Foundation: Panopticlick: Is your browser safe against tracking? `https://panopticlick.eff.org/`,
33. Mowery, K., Bogenreif, D., Yilek, S., Shacham, H.: Fingerprinting information in javascript implementations. In: Proceedings of the Web 2.0 Workshop on Security and Privacy (W2SP). (2011)
34. Miyazaki, A.D.: Online privacy and the disclosure of cookie use: Effects on consumer trust and anticipated patronage. Journal of Public Policy & Marketing **27**(1) (2008) 19–33
35. Mulazzani, M., Reschl, P., Huber, M., Leithner, M., Schrittwieser, S., Weippl, E., Wien, F.: Fast and reliable browser identification with javascript engine fingerprinting. In: Proceedings of the Web 2.0 Workshop on Security and Privacy (W2SP). (2013)
36. Mowery, K., Shacham, H.: Pixel perfect: Fingerprinting canvas in HTML5. Proceedings of the Web 2.0 Workshop on Security and Privacy (W2SP) (2012)
37. EasyPrivacy. `https://easylist.adblockplus.org/`
38. Disconnet. `https://disconnect.me//`
39. Krishnamurthy, B., Wills, C.: Privacy diffusion on the web: a longitudinal perspective. In: Proceedings of the 18th international conference on World Wide Web (WWW). (2009)
40. Aggarwal, G., Bursztein, E., Jackson, C., Boneh, D.: An analysis of private browsing modes in modern browsers. In: USENIX Security Symposium. (2010) 79–94
41. De Groef, W., Devriese, D., Nikiforakis, N., Piessens, F.: Flowfox: a web browser with flexible and precise information flow control. In: Proceedings of the 2012 ACM conference on Computer and communications security, ACM (2012) 748–759
42. Hedin, D., Birgisson, A., Bello, L., Sabelfeld, A.: Jsflow: Tracking information flow in javascript and its apis. In: Proc. 29th ACM Symposium on Applied Computing. (2014)
43. Sen, K., Kalasapur, S., Brutch, T., Gibbs, S.: Jalangi: A selective record-replay and dynamic analysis framework for javascript. In: Proceedings of the 2013 9th Joint Meeting on Foundations of Software Engineering, ACM (2013) 488–498

44. Chugh, R., Meister, J.A., Jhala, R., Lerner, S.: Staged information flow for javascript. In: ACM Sigplan Notices. Volume 44., ACM (2009) 50–62
45. World Wide Web Consortium: Platform for privacy preferences (p3p) project. http://www.w3.org/P3P
46. Byers, S., Cranor, L.F., Kormann, D., McDaniel, P.: Searching for privacy: Design and implementation of a p3p-enabled search engine. In: Privacy Enhancing Technologies, Springer (2005) 314–328
47. Mayer, J.: Tracking the trackers: early results. http://cyberlaw.stanford.edu/blog/2011/07/tracking-trackers-early-results (2011)
48. Goldfarb, A., Tucker, C.E.: Privacy regulation and online advertising. Management Science **57**(1) (2011) 57–71
49. Federal Trade Commission: Protecting consumer privacy in an era of rapid change: Recommendations for businesses and policymakers". https://www.ftc.gov/reports/protecting-consumer-privacy-era-rapid-change-recommendations-businesses-policymakers (2012)
50. Federal Trade Commission: Protecting consumer privacy in an era of rapid change, a proposed framework for businesses and policymakers". https://www.ftc.gov/reports/preliminary-ftc-staff-report-protecting-consumer-privacy-era-rapid-change-proposed-framework (2010)
51. European Parliament: Directive 2002/58/ec. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:en:HTML (2002)
52. Article 29 Data Protection Working Party: Opinion 04/2012 on cookie consent exemption. http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp194_en.pdf (2012)
53. Article 29 Data Protection Working Party: Opinion 9/2014 on the application of directive 2002/58/ec to device fingerprinting. http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp224_en.pdf (2014)
54. Digital Advertising Alliance: Self-regulatory principles for online behavioral advertising, behavioral advertising. http://www.aboutads.info/resource/download/seven-principles-07-01-09.pdf (2009)
55. Digital Advertising Alliance: Self-regulatory principles for multi-site data. http://www.aboutads.info/resource/download/Multi-Site-Data-Principles.pdf (2011)
56. Symantec – Security Response team: The Waterbug attack group. http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/waterbug-attack-group.pdf (2015)
57. Threat Intelligence, FireEye: Pinpointing Targets: Exploiting Web Analytics to Ensnare Victims. https://www2.fireeye.com/rs/848-DID-242/images/rpt-witchcoven.pdf (2015)