FLASH COOKIES AND PRIVACY II: NOW WITH HTML5 AND ETAG RESPAWNING

MIKA D. AYENSON[*], DIETRICH J. WAMBACH[†], ASHKAN SOLTANI[‡]
NATHANIEL GOOD[§] & CHRIS JAY HOOFNAGLE[**††]

IN AUGUST 2009, WE DEMONSTRATED THAT POPULAR WEBSITES WERE USING "FLASH COOKIES" TO TRACK USERS. SOME ADVERTISERS HAD ADOPTED THIS TECHNOLOGY BECAUSE IT ALLOWED PERSISTENT TRACKING EVEN WHERE USERS HAD TAKEN STEPS TO AVOID WEB PROFILING. WE ALSO DEMONSTRATED "RESPAWNING" ON TOP SITES WITH FLASH TECHNOLOGY. THIS ALLOWED SITES TO REINSTANTIATE HTTP COOKIES DELETED BY A USER, MAKING TRACKING MORE RESISTANT TO USERS' PRIVACY-SEEKING BEHAVIORS.

IN THIS FOLLOWUP STUDY, WE REASSESS THE FLASH COOKIES LANDSCAPE AND EXAMINE A NEW TRACKING VECTOR, HTML5 LOCAL STORAGE AND CACHE-COOKIES VIA ETAGS.

WE FOUND OVER 5,600 STANDARD HTTP COOKIES ON POPULAR SITES, OVER 4,900 WERE FROM THIRD PARTIES. GOOGLE-CONTROLLED COOKIES WERE PRESENT ON 97 OF THE TOP 100 SITES, INCLUDING POPULAR GOVERNMENT WEBSITES. SEVENTEEN SITES WERE USING HTML5, AND SEVEN OF THOSE SITES HAD HTML5 LOCAL STORAGE AND HTTP COOKIES WITH MATCHING VALUES. FLASH COOKIES WERE PRESENT ON 37 OF THE TOP 100 SITES.

WE FOUND TWO SITES THAT WERE RESPAWNING COOKIES, INCLUDING ONE SITE—HULU.COM—WHERE BOTH FLASH AND CACHE COOKIES WERE EMPLOYED TO MAKE IDENTIFIERS MORE PERSISTENT. THE CACHE COOKIE METHOD USED ETAGS, AND IS CAPABLE OF UNIQUE TRACKING EVEN WHERE ALL COOKIES ARE BLOCKED BY THE USER AND "PRIVATE BROWSING MODE" IS ENABLED.

---

[*] Mika D. Ayenson is a junior at Worcester Polytechnic Institute. Authors Ayenson & Wambach made equal contributions to this paper.
[†] Dietrich J. Wambach is a senior at the University of Wyoming.
[‡] MIMS (Berkeley 2009), independent researcher and consultant focused on privacy, security, and behavioral economics.
[§] Ph.D., Chief Scientist and Principal of Good Research.
[**] Lecturer in Residence, UC Berkeley Law.
[††] This work was supported exclusively by TRUST (Team for Research in Ubiquitous Secure Technology), which receives support from the National Science Foundation (NSF award number CCF-0424422) and the following organizations: AFOSR (#FA9550-06-1-0244), BT, Cisco, ESCHER, HP, IBM, iCAST, Intel, Microsoft, ORNL, Pirelli, Qualcomm, Sun, Symantec, Telecom Italia, and United Technologies. We are grateful for the opportunities offered by the TRUST Research Experiences for Undergraduates program (REU), and to its program leader, Dr. Kristen Gates.

INTRODUCTION

In a study of popular websites in 2009, we found widespread use of "Flash cookies."[1] Flash cookies, technically called "local shared objects," are files used by Adobe Flash developers to store data on users' computers. Our 2009 paper elucidated the advantages of Flash cookies from a developer perspective, and documented that some advertisers adopted Flash cookies because they were relatively unknown, more difficult for consumers to delete, and were more effective in tracking than HTTP cookies. We documented other tracking advantages of Flash cookies as well—they are more persistent than HTTP cookies, they can store 100KB of information by default (HTTP cookies only store 4KB), and they are stored such that all browsers on an computer can access them, meaning that even if a user switches browsers, Flash cookies enables the user to be tracked.[2]

RECENT RESEARCH

In recent years, there has been an explosion in research concerning user tracking online. In their ongoing investigations of web privacy issues, Bala Krishnamurthy, Konstantin Naryshkin, and Craig Wills studied how personal information flows from first to third party sites. They found that a majority of the popular sites they analyzed "directly leak sensitive and identifiable information to third-party aggregators."[3] This follows their multiple-year study of 1,200 websites, where they found increasing

---

[1] Ashkan Soltani, Shannon Canty, Quentin Mayo, Lauren Thomas, and Chris Jay Hoofnagle, *Flash Cookies and Privacy*, Aug. 10, 2009, available at: http://ssrn.com/abstract=1446862, *accepted for publication at* AAAI Spring Symposium on Intelligent Information Privacy Management 2010, CodeX: The Stanford Center of Computers and Law.

[2] For an in-depth discussion of the various advantages of different tracking vectors, *see* Sonal Mittal, *User Privacy and the Evolution of Third-party Tracking Mechanisms on the World Wide Web* (2010), available at http://www.stanford.edu/~sonalm/Mittal_Thesis.pdf.

[3] Krishnamurthy, B., Naryshkin, K., & Wills, C. E*., Privacy leakage vs. Protection measures: the growing disconnect*, *presented at* W2SP 2011: Web 2.0 Security and Privacy 2011 (2011), available at http://www.cs.wpi.edu/~cew/papers/w2sp11.pdf.

collection of information about users from an increasingly concentrated group of tracking companies.[4]

Researchers have also focused upon new vectors for tracking. As early as 2003, Dean Gaudet described unique user tracking through using "ETags," a feature of the cache in browsers.[5] Samy Kamkar has demonstrated the "Evercookie," a tracking mechanism that uses Flash storage, HTML, and a variety of other techniques (including ETags) in order to make it resistant to user attempts to delete cookies and other unique identifiers.[6] Peter Eckersley has demonstrated the privacy risks associated with browser fingerprinting, where server-side scripts can query a browser for enough information to identify a computer.[7]

In particular, recent research has focused upon the privacy implications of plugins such as Flash. As early as 2006, Corey Benninger noted that Flash cookies could be set without any visible sign to the user that Flash was running: "In fact, it would be difficult to reliably detect if an application were using flash cookies."[8] As Sipior, Ward, & Mendoza recently noted, addressing this risk by simply disabling Flash is unrealistic from a user perspective because an enormous amount of web content is delivered in

---

[4] Krishnamurthy, B., & Wills, C., *Privacy diffusion on the web: A longitudinal perspective*, Proceedings of the 18th ACM international conference on World wide web (2009)(p. 541-550), available at http://portal.acm.org/citation.cfm?id=1526782.

[5] Dean Gaudet, *Tracking Without Cookies*, Feb. 17, 2003, available at http://www.arctic.org/~dean/tracking-without-cookies.html ("other than cookies, there's typically only one other type of data a webserver can cause a browser to store on its local harddrive -- cacheable web content. this technique attempts to get the browser to store unique id information in its cache in a manner which will be communicated to the server at a later date. (the later communication will be via a GET If-Modified-Since, or If-None-Match.)")

[6] Samy Kamkar, *Evercookie* (2010) available at http://samy.pl/evercookie/.

[7] Peter Eckersley, *How unique is your web browser?,* Proceedings of the Privacy Enhancing Technologies Symposium (PETS 2010), Springer Lecture Notes in Computer Science (p. 1-18)(2010), available at http://www.springerlink.com/index/0J1M07443GU00H07.pdf.

[8] Corey Benninger, *AJAX Storage: A Look at Flash Cookies and Internet Explorer Persistence*, Foundstone Professional Services & Education, McAfee (2006), available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.128.2523.

formats requiring a plugin.[9] This is problematic from a privacy perspective, because once in place, the plugin infrastructure can be leveraged for unique user tracking and sharing of unique identifiers across domains.[10]

Important Flash security research related to our investigation concerns Flash's "cross domain" policies. According to Adobe, "A [cross-domain] policy file is a simple XML file that gives the Flash Player permission to access data from a given domain without displaying a security dialog. When placed on a server, it tells the Flash Player to allow direct access to data on that server, without prompting the user grant access."[11] This feature routes around the "same-origin policy" that underlies the security of the web, giving Flash applications the ability to read data on other domains and subdomains. In his 2008 analysis of websites with cross-domain policies, Jeremiah Grossman explained:

> *When a hostname is included in the circle of trust you allow them to read all data on the site that the user has access to, this includes any (authenticated) content and (session) cookies. So should a malicious attacker or website owner gain control of a website in the circle of trust (via a server hack or XSS), then they feasibly can compromise user data off that domain. This could easily leads to privacy violations, account takeovers, theft of sensitive data, and bypassing of CSRF protections (grabbing the key ahead of time).*[12]

In follow-up research, different teams have found websites with "wildcard" entries in their cross-domain policies, meaning that they have marked as trusted any other domain on the web. Sebastian Lekies and colleagues found such wildcard policies in 2.8 percent of a sample of almost 1.1 million

---

[9] Sipior, J., Ward, B., & Mendoza, R., *Online Privacy Concerns Associated with Cookies, Flash Cookies, and Web Beacons*, 10 Journal of Internet Commerce 1, 4 (2011), doi: 10.1080/15332861.2011.558454.

[10] *Id*. at 11.

[11] ADOBE, CROSS-DOMAIN POLICY FOR FLASH MOVIES, September 28, 2010, available at http://kb2.adobe.com/cps/142/tn_14213.html.

[12] Jeremiah Grossman, *Crossdomain.xml Invites Cross-site Mayhem*, May14, 2008, available at http://jeremiahgrossman.blogspot.com/2008/05/crossdomainxml-invites-cross-site.html.

domains.[13] Dongseok Jang and colleagues found wildcard policies in 6 percent of their sample of 50,000 domains.[14] (This apparent disparity is a result of a greater concentration of cross-domain policies among popular sites, with adoption falling off in less popular sites.) Focusing upon cross-domain problems and other security issues from Flash implementations, Kuzma and colleagues found minor vulnerabilities on almost all educational websites they sampled, and more serious vulnerabilities on 20% of their sample.[15]

The most important study related to our work was authored by Aleecia McDonald and Lorrie Faith Cranor of Carnegie Mellon.[16] Their 2011 investigation of Flash cookies found a dramatic decline in their use. For instance, McDonald et al. found that only 20% of top 100 websites used Flash cookies, and that only two sites respawned using Flash cookies. McDonald et al. were also careful to attempt to determine whether Flash cookie values were unique or not—six of the top 100 sites had Flash cookies that were not unique, and thus probably not used to track individuals.

We direct the reader to our methods section, as it highlights key differences of our investigation. The McDonald team visited the landing page of the top 100 sites, plus a selection of random sites. Our current and 2009 studies are different in that we visit the top 100 sites and make 10 clicks on the same domain, to simulate a user session. As a result of this difference, McDonald et al. acknowledged that their scan represented a "lower bound" in counting of Flash cookies.

---

[13] Lekies, S., Johns, M., & Tighzert, W., *The State of the Cross-domain Nation*, W2SP 2011: Web 2.0 Security and Privacy 2011 (2011), available at http://w2spconf.com/2011/papers/cross_domain_Nation.pdf.

[14] Jang, D., Venkataraman, A., Sawka, G. M., & Shacham, H., *Analyzing the Crossdomain Policies of Flash Applications*, W2SP 2011: Web 2.0 Security and Privacy 2011 (2011) available at http://www.w2spconf.com/2011/papers/crossDomainFlash.pdf.

[15] Kuzma, J., Price, C., & Henson, R., *Flash vulnerabilities analysis of US educational websites*, International Journal of Electronic Security and Digital Forensics, 3(2), 95-107 (2011), available at http://inderscience.metapress.com/index/9W7J37484G5Q848L.pdf.

[16] McDonald, A. M., & Cranor, L. F., *A Survey of the Use of Adobe Flash Local Shared Objects to Respawn HTTP Cookies*, CMU-CyLab-11-001 (2011), available at http://www.casos.cs.cmu.edu/publications/papers/CMUCyLab11001.pdf.

McDonald et al. also emphasizes the normative implications of Flash use for user tracking where sites are not using respawning. The use of Flash cookies for unique user tracking is problematic, because it is functionally equivalent to respawning. This is because users are generally not aware of Flash cookies and until very recently, browser controls did not address them. Whether or not a website respawns, if it uses Flash cookies, it can uniquely and persistently track individuals even in situations where the user has taken reasonable steps to avoid online profiling.

With our focus on respawning, we did not adequately articulate this problem in 2009. In fact, we referred to local shared objects as "Flash cookies" in order to make the issue more accessible to policymakers and others. But this caused many to speciously argue that Flash cookies are really no different than HTTP cookies. Local shared objects are not just like HTTP cookies—they are far more flexible than HTTP cookies, and the infrastructure that gave rise to them enabled an obscure and persistent tracking mechanism that largely is still in place today. Table 1 below sets forth the basic differences among the cookies analyzed in this paper.

HTML5 Web Storage

Flash cookies may be just a bridge technology for online trackers. HTML5 storage offers many advantages over ordinary cookies, and since it does not involve using a plugin (like Flash), HTML5 may become a more universal tracking mechanism. Like Flash cookies, HTML5 storage is more persistent than HTTP cookies. HTTP cookies expire by default, and in order to make them persistent, developers must use a complex syntax and constantly update the expiration date. HTML5 data are persistent until affirmatively deleted by a web site or user. Storage size is important too. While Flash cookies have a default limit of 100KB, HTTP cookies store just 4KB, compared to 5Mb for HTML5 storage.[17]

[17] Bruce Lawson & Remy Sharp, Introducing HTML5 142-3 (New Riders 2011).

Table 1: Key Characteristics of HTTP Cookies, Flash Cookies, and HTML5 Storage

	HTTP Cookies	Flash cookies	HTML5 storage
Storage	4KB	100KB by default	5Mb by default
Expiration	Session by default	Permanent by default	Permanent by default
Location	In SQL file (Firefox)	Stored outside the browser	In SQL file (Firefox)
Access	Only by browser	By multiple browsers on same machine	Only by browser

Several commentators have highlighted the privacy risks presented by HTML5. Others have argued that HTML5 has a great potential to enable more privacy-preserving advertising models.[18]

However, to our knowledge, no one has performed a survey of HTML5 privacy practices. Thus, as part of our update to our original Flash cookies investigation, we also captured and analyzed HTML5 data.

TECHNICAL AND POLICY DEVELOPMENTS SURROUNDING FLASH COOKIES

Our 2009 paper concluded:

> *Flash cookies are a popular mechanism for storing data on top 100 websites. Some top 100 websites are circumventing user deletion of HTTP cookies by respawning them using Flash cookies with identical values. Even when a user obtains a NAI opt-out cookie, Flash cookies are employed for unique user tracking. These experiences are not consonant with user expectations of private browsing and deleting cookies. Users are limited in self-help, because anti-tracking tools effective against this technique are not widespread, and presence of Flash cookies is rarely disclosed in privacy policies.*

[18] *See generally*, Arvind Narayanan and Jonathan Mayer, *DoNotTrack: an approach to tracking protection*, Workshop on Internet Tracking, Advertising, and Privacy (WiTap), July 22, 2011 (presentation at workshop); Arvind Narayanan and Jonathan Mayer, *The Do Not Track Cookbook* (n.d.), available at http://donottrack.us/cookbook.

> *A tighter integration between browser tools and Flash cookies could empower users to engage in privacy self-help, by blocking Flash cookies. But, to make browser tools effective, users need some warning that Flash cookies are present. Disclosures about their presence, the types of uses employed, and information about controls, are necessary first steps to addressing the privacy implications of Flash cookies.*

Much has happened since we published the original Flash cookies work. The realization that advertisers were working around expressed privacy preferences with technology led to attention from the Federal Trade Commission, European regulators, and the plaintiff bar. The Federal Trade Commission recognized the problem in its staff report on privacy:

> *…consumers are not likely to be aware of the technical limitations of existing control mechanisms. For example, they may believe they have opted out of tracking if they block third-party cookies on their browsers; yet they may still be tracked through Flash cookies or other mechanisms.*[19]

Additionally, former Commissioner Pamela Jones Harbour warned companies about tracking that evaded users' intent:

> *Even where consumers have the ability to opt-out, the effects are limited. If consumer data are unavailable from one source, often they can be obtained from another. Flash cookies and other technology largely circumvent cookie controls. We may soon long for the day when all we worried about were cookies. For every company crafting a response that addresses notice, choice, or transparency, there are several more firms trying to parse and evade the intent of*

[19] FEDERAL TRADE COMMISSION, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: A PROPOSED FRAMEWORK FOR BUSINESSES AND POLICYMAKERS 65-66, Dec. 2010.

> *Commission guidance. We have entered a digital arms race, and the current outlook is troubling.*[20]

In January 2010, Adobe released an update to Flash Player that made it compatible with "Private Browsing Modes" in major web browsers.[21] In March 2011, a new update to Flash Player enabled users to delete cookies within the browser, and added a control panel for users to make privacy and other settings.[22] Meanwhile others have created browser extensions to enumerate and manage Flash cookies and other tracking vectors.[23]

Adobe officials have condemned the practice of respawning. In a letter to the Federal Trade Commission, MeMe Rasmussen, Adobe's CPO, wrote:

> *Applications and Web sites built for use with Adobe® Flash® Player are enjoyed by the vast majority of computer users today. Many of these applications depend on Local Storage1 to store data necessary to make the applications easy to use in a way that is consistent with the user's expectations.*
>
> *However, we are aware of one use of Flash Local Storage that is inconsistent with the user's expectations. This is the practice of using Local Storage to back up browser cookies for the purpose of restoring them after they have been deleted by the user. This restoration happens without the user's knowledge and express consent.*

---

[20] Commissioner Pamela Jones Harbour, Remarks Before FTC Exploring Privacy Roundtable, Dec. 7, 2009 (Washington, DC), available at http://www.ftc.gov/speeches/harbour/091207privacyroundtable.pdf.

[21] Jimson Xu & Tom Nguyen, *Private browsing in Flash Player 10.1*, June 30, 2010, available at http://www.adobe.com/devnet/flashplayer/articles/privacy_mode_fp10_1.html

[22] Martin Brinkmann, *A Close Look At Adobe Flash Player 10.3 Beta*, Mar. 8, 2011, available at http://www.ghacks.net/2011/03/08/a-close-look-at-adobe-flash-player-10-3-beta/

[23] See e.g. Abine's Privacy Suite: http://www.abine.com/apps.php; Sonal Mittal's FoxTracks: http://www.cdt.org/foxtracks/webbugs; and and BetterAdvertising's Ghostery: http://www.ghostery.com/.

> ***Adobe condemns this type of misuse of Local Storage.*** *We encourage developers to use technology responsibly, and certainly not in ways that circumvents the user's intentions or reasonable expectations.*[24]

The Network Advertising Initiative (NAI), a US-based self-regulatory project hosted by a public relations firm, said that its members should not use Flash cookies for online behavioral advertising purposes until, "…such time as web browser tools allow for the same level of transparency and control as is available today for standard HTTP cookies."[25] We note that the developments of this year arguably greenlight the use of Flash cookies under the NAI rule, and in any case, the NAI ban only pertains to OBA-related uses of Flash cookies.

METHODS

We largely followed the methods of our 2009 paper with some improvements to ensure a clean state between sessions. We crawled the top 100 U.S. websites based upon QuantCast.com's ranking of July 13, 2011. The data collection occurred on July 21, 2011. We used two PCs with virtualized Linux/Ubuntu OSes, being careful to restore the virtual machine after each website visited, in order to avoid contamination. Using Firefox version 5, we called the site URL and then made 10 arbitrary clicks on each website, being careful to remain on the same top-level-domain. We collected HTTP, HTML5, and Flash cookies from these crawling sessions. We never "signed in" to a website in this process.

Because of the dynamic nature of websites and online advertising, any given survey may produce different advertisements and correspondingly different HTTP, HTML5, and Flash cookies. Thus, our snapshot may differ

---

[24] ADOBE SYSTEMS INC., COMMENTS FROM ADOBE SYSTEMS INCORPORATED – PRIVACY ROUNDTABLES PROJECT NO. P095416, Jan. 27, 2010, available at http://www.ftc.gov/os/comments/privacyroundtable/544506-00085.pdf (emphasis in original).

[25] Network Advertising Initiative, FAQs (n.d.), available at http://www.networkadvertising.org/managing/faqs.asp#question_19.

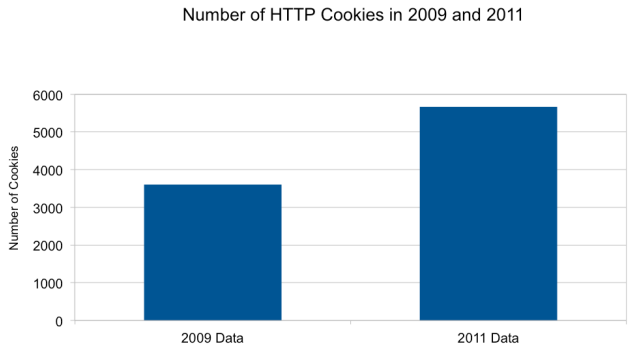
from another user's experience. However we feel that this provides reasonable sample for study.

We used several methods to detect and confirm respawning cookies, including manually deleting HTML cookies to see whether they reappeared. We also injected arbitrary values into objects to see whether those same values would later appear in HTTP and HTML5 cookies.

RESULTS & DISCUSSION

HTTP COOKIES

We detected cookies on all top 100 websites. In total, we detected 5,675 HTTP cookies. This is dramatically higher than the 3,602 we detected in 2009. Twenty sites placed 100 or more cookies, including seven that placed more than 150 (wikia.com, 242; legacy.com, 230; foxnews.com, 185; bizrate.com, 175; drudgereport.com, 168; myspace.com, 151; time.com 151).



The most frequently appearing cookie keys were: uid, id, PREF, __utmz, __utma, __utmb, and UID. Many of these keys are commonly associated with unique user tracking. For instance, __utma is used by Google for identifying unique visitors.

Most cookies—4915 of them—were placed by a third party host.

We detected over 600 third party hosts among the 4915 third party cookies. Google had cookies on 89 of the top 100 sites; the company's ad tracking

network, doubleclick.net, had cookies on 77. Combined, Google has a presence on 97 of the top 100 websites. This includes popular government websites such as usps.com, irs.gov, and nih.gov. Only microsoft.com, ups.com, and wikipedia.org lacked some type of Google cookie.

Other third party trackers with a strong presence in the top 100 included scorecardresearch.com (61), and atdmt.com (56). Among top 100 sites, wikia.com, legacy.com, foxnews.com, drudgereport.com, and bizrate.com hosted the most cookies from third party domains.

Flash Cookies – Local Shared Objects

We found 100 Flash cookies on the top 100 sites, down from the 281 we found in 2009. These Flash cookies appeared on 37 sites, down from the 54 sites we found in 2009.

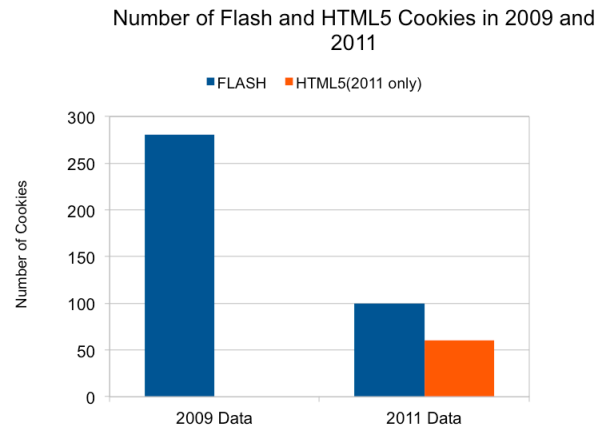
Flash cookies can store many keys and values. MTV.com had 8 flash cookies, one of which stored over 140 values. We found 454 key/value pairs in 100 Flash cookies detected. The most common keys used were: expiration, volume, creation, domainHash, campaignTracking, id, and time.

Two sites had shared values between Flash cookies and HTTP cookies: hulu.com and foxnews.com. In the case of foxnews.com, the value was shared in HTML5 local storage as well.

HTML5 Storage

Seventeen of the top 100 sites were using HTML5 local storage. These 17 sites had a total of 60 key/value pairs.

We found matching values among HTML5 local storage and HTTP cookies in several cases. Twitter.com, tmz.com, squidoo.com, nytimes.com, hulu.com, foxnews.com, and cnn.com had such matching values. In most of these cases, the matching value was with a third party service, such as meebo.com, kissanalytics.com, and polldaddy.com.



RESPAWNING

We found three respawning behaviors on two sites: hulu.com and foxnews.com.

In 2009, we reported that a QuantCast cookie was respawned on hulu.com. After our 2009 paper, QuantCast executives contacted authors Hoofnagle and Soltani almost immediately, and quickly acted to change the behavior of their service in order to prevent respawning.[26]

Nevertheless, hulu.com, QuantCast, and other companies were sued for the practice, and the case settled this year. In a summary of Flash cookies filed with the court, it was claimed that websites such as Hulu did not know that third party services provided by QuantCast and Clearspring tracked users through Flash.[27] This assertion effectively shifted the blame from

[26] Ryan Singel, *Online Tracking Firm Settles Suit Over Undeletable Cookies*, Wired Epicenter, Dec. 5, 2009, available at http://www.wired.com/epicenter/2010/12/zombie-cookie-settlement/.

[27] "The Customer Defendants, on their own behalf and on behalf of their corporate parents and affiliates, have represented to Quantcast and Clearspring that the Customer Defendants were unaware that LSOs were being used to store information regarding consumers who accessed their websites and web content. Quantcast and Clearspring do not dispute that representation and, to the extent of their knowledge, information, and belief, adopt and incorporate it here." *In Re Quantcast Advertising Cookie Litigation*, 2:10-cv-05484-GW–JCG, (Cal. C.D. 2011)(Joint Submission of Supplemental Information Regarding Plaintiffs' Motion for Preliminary Approval of Class

consumer-facing websites to the third party tracking companies involved. In the settlement flowing from the suit, QuantCast and Clearspring explicitly promised to not respawn cookies using Flash, or to use Flash as an alternative to HTTP cookies for tracking purposes.[28] These obligations did not apply to consumer-facing websites, such as hulu.com.

We found two different methods of cookie respawning on hulu.com.

First, hulu.com used standard Flash respawning to reinstantiate a HTTP cookie with the key "guid," mirroring a stored object with the key "computerguid." There are two important points to raise about this: unlike the situation in 2009, where a third party respawned the cookies, this use of Flash is in-house at hulu.com. And, while Adobe points out that local storage enables the delivery of rich content, hulu.com's use of Flash appears to fall into the category of unique user tracking condemned by Adobe. Adobe argues that such uses of Flash should be subject to express user consent.[29]

Second, we found first party HTTP and HTML5 cookies respawned on hulu.com through a service hosted at kissmetrics.com. This respawning employed the cache to mirror values, specifically ETags. To our knowledge, this is the first demonstration of this ETag tracking "in the wild."

ETag tracking and respawning is particularly problematic because the technique generates unique tracking values even where the consumer blocks HTTP, Flash, and HTML5 cookies. In order to block this tracking, the user would have to clear the cache between each website visit. Even in private browsing mode, ETags can track the user during a browser session. Additionally, the ETag respawning we observed set a first party cookie on hulu.com. This means that other sites subscribing to the kissmetrics.com service could synchronize these identifiers across their domains.

Action Settlement).

[28] *In Re Quantcast Advertising Cookie Litigation*, 2:10-cv-05484-GW–JCG, (Cal. C.D. 2011)(Settlement Agreement at §4.19).

[29] ADOBE SYSTEMS INC., COMMENTS FROM ADOBE SYSTEMS INCORPORATED – PRIVACY ROUNDTABLES PROJECT NO. P095416, Jan. 27, 2010, available at http://www.ftc.gov/os/comments/privacyroundtable/544506-00085.pdf.

The script for this function, hosted at http://doug1izaerwt3.cloudfront.net, includes other code that indicates its author is aware of tracking and the risk of data collection about the user. For instance, it includes a function to detect the collection of information that credit card companies require websites to control more carefully.

On June 30, 2011, Hulu.com updated its privacy policy to include disclosures surrounding Flash cookies.[30] This update appears to be driven by obligations in the Flash cookies settlement, which requires consumer-facing websites to include, "in its online Privacy Policy, a disclosure of its use of LSOs and a link to at least one website or utility offering users the ability to manage LSOs, if such website or utility is available."[31]

In this updated policy, Hulu.com includes a link to Adobe's Flash cookie manager, discloses that it uses Flash cookies, but then downplays their potential for tracking: "Local Shared Objects are similar to browser cookies, but can store data more complex than simple text. By themselves, they cannot do anything to or with the data on your computer."

We object to this last sentence in particular. While it is technically true that *by themselves* Flash cookies cannot do anything to the data on a user's computer, in reality, Flash cookies never are used *by themselves*. It is the code accompanying Flash cookies that enables them to mirror other data, and can be used to reinstantate that data when deleted by the user.

The hulu.com privacy policy does not mention respawning of any kind, and even claims that "You can configure your Internet browser to warn you each time a cookie is being sent or to refuse cookies completely. However, unless you accept cookies, you will not have access to certain Hulu Services."

---

[30] HULU.COM, PRIVACY POLICY, June 30, 2011, available at http://www.hulu.com/privacy (We have updated our Privacy Policy to provide more details about our information practices, including…Our use of "Local Shared Objects" in connection with Adobe's Flash Player.")

[31] *In Re Quantcast Advertising Cookie Litigation*, 2:10-cv-05484-GW–JCG, (Cal. C.D. 2011)(Settlement Agreement at §4.20.4).

Hulu.com's updated policy also describes "Web beacons." It is unclear whether this section of the policy describes kissmetrics.com cache respawning. The description would not lead an average user to understand that the cache was being used to undo cookie deletion.

We find it surprising that months after settling a suit involving unique user tracking through third parties, hulu.com has moved Flash tracking and respawning in-house. Furthermore, the use of kissmetrics cache cookie respawning is very similar to the respawning we found in 2009—hulu.com used a third party to engage in tracking that users do not know about, cannot detect, and effectively cannot block.

Kissmetrics.com has a privacy policy as well, but it is targeted to commercial buyers of the kissmetrics.com service, rather than average web users.

We also found respawning on foxnews.com associated with voting in web polls. A third party polldaddy.com Flash cookie (hosted at i0.poll.fm) respawned an HTML5 cookie on foxnews.com. This cookie's key corresponded to the number assigned to the poll that our researcher engaged in. It appears to prevent the user from voting in the same poll twice.

Foxnews.com's privacy policy does disclose it "may use" Flash and other cookies.[32] It does not mention respawning.

CONCLUSION

In 2009, we surveyed the most popular websites to determine how they were using Flash cookies. In this followup study, we found that fewer websites are using Flash cookies. Fewer are also respawning cookies using Flash. However, one popular site is using both Flash and the

[32] FOXNEWS.COM, PRIVACY POLICY, Feb. 22, 2011, available at http://www.foxnews.com/about/privacy-policy/.

user's cache to respawn HTTP and HTML5 cookies in a way that cannot be blocked currently by the browser.

We also found many HTTP cookies on top sites, most of which originate from third parties. Google in particular has the ability to track user behavior across nearly all top sites—97 of them.

Although there is much potential for privacy-enhancing applications of HTML5 local storage, it nevertheless may emerge as a new tracking vector. Seventeen of the sites we surveyed employed HTML5 local storage, several did so in order to mirror a tracking identifier from a third party.

Table 2: Key Results and Comparison with Other Studies

	Soltani 2009	McDonald 2011	Ayenson Wambach et al. 2011
Number of sites with Flash cookies (top 100 sites)	54	20	37
Total number of Flash cookies (top 100 sites)	281	Not reported	100
Sites with respawning (top 100 sites)	6	2	2
Number of websites with HTTP Cookies (top 100 sites)	98	98	100
Total HTTP Cookies set (top 100 sites)	3602	Not reported	5,675
Sites with shared Flash/HTTP values on top 100	31	Not reported	2
Total shared Flash/HTTP values on top 100	41	8	2
Sample	Top 100 websites and six government sites	Top 100 websites and 600 random sites	Top 100 websites
Method	Visited homepage and then made 10 clicks on the same domain	Visited homepage multiple times	Visited homepage and then made 10 clicks on the same domain

Figure 1: Upon visiting hulu.com, we receive a tracking identifier through an ETag generated by Kissmetrics. This identifier persists when visiting spotify.com, which also uses Kissmetrics.

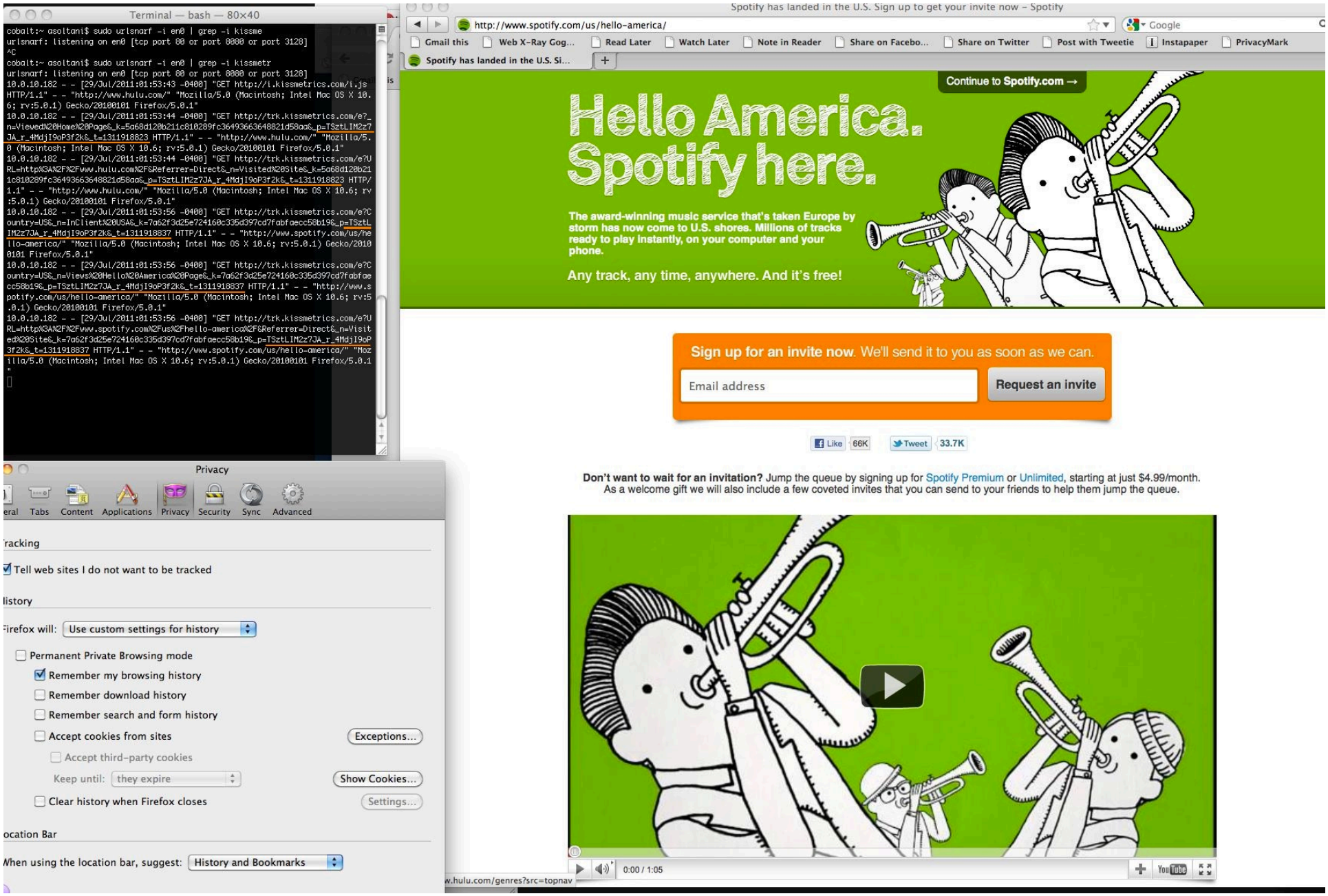


Figure 2: The Kissmetrics ETag is highlighted.

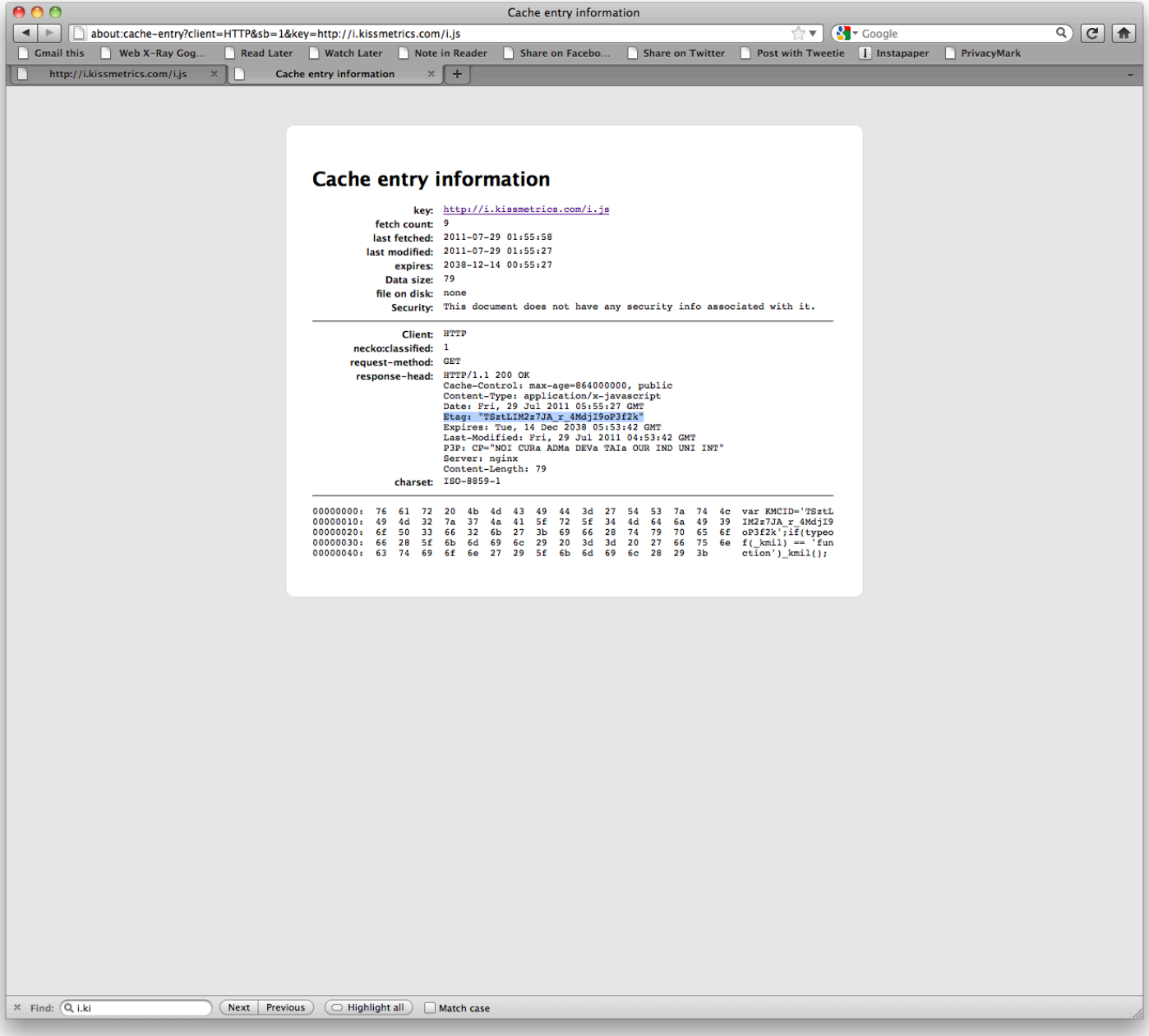


Figure 3: A Kissmetrics identifer based on ETags persists across the domains gigaom.com and spotify.com, with private browsing mode and do not track enabled.