Flash Cookies and Privacy

Ashkan Soltani[A], Shannon Canty[B][1], Quentin Mayo[B][2], Lauren Thomas[B][3] & Chris Jay Hoofnagle[C]

School of Information[A]
Summer Undergraduate Program in Engineering Research at Berkeley (SUPERB) 2009[B]
UC Berkeley School of Law[C]
University of California, Berkeley
Berkeley, USA
correspondence to: choofnagle@law.berkeley.edu

Abstract

This is a pilot study of the use of "Flash cookies" by popular websites. We find that more than 50% of the sites in our sample are using Flash cookies to store information about the user. Some are using it to "respawn" or re-instantiate HTTP cookies deleted by the user. Flash cookies often share the same values as HTTP cookies, and are even used on government websites to assign unique values to users. Privacy policies rarely disclose the presence of Flash cookies, and user controls for effectuating privacy preferences are lacking.

Introduction

Advertisers are increasingly concerned about unique tracking of users online.[4] Several studies have found that over 30% of users delete first party HTTP cookies once a month, thus leading to overestimation of the number of true unique visitors to websites, and attendant overpayment for advertising impressions.[4]

Mindful of this problem, online advertising companies have attempted to increase the reliability of tracking methods. In 2005, United Virtualities (UV), an online advertising company, exclaimed, "All advertisers, websites and networks use [HTTP] cookies for targeted advertising, but cookies are under attack."[5] The company announced that it had, "developed a backup ID system for cookies set by web sites, ad networks and advertisers, but increasingly deleted by users. UV's 'Persistent Identification Element' (PIE) is tagged to the user's browser, providing each with a unique ID just like traditional cookie coding. However, PIEs cannot be deleted by any commercially available anti-spyware, mal-ware, or adware removal program. They will even function at the default security setting for Internet Explorer."[5] (Since 2005, a Firefox plugin called "BetterPrivacy", and more recently, a shareware program called "Glary Utilities Pro" can assist users in deleting Flash cookies.)

United Virtualities' PIE leveraged a feature in Adobe's Flash MX: the "local shared object,"[6] also known as the "flash cookie." Flash cookies offer several advantages that lead to more persistence than standard HTTP cookies. Flash cookies can contain up to 100KB of information by default (HTTP cookies only store 4KB).[7] Flash cookies do not have expiration dates by default, whereas HTTP cookies expire at the end of a session unless programmed to live longer by the domain setting the cookie. Flash cookies are stored in a different location than HTTP cookies,[7] thus users may not know what files to delete in order to eliminate them. Additionally, they are stored so that different browsers and stand-alone Flash widgets installed on a given computer access the same persistent Flash cookies. Flash cookies are not controlled by the browser. Thus erasing HTTP cookies, clearing history, erasing the cache, or choosing a delete private data option within the browser does not affect Flash cookies. Even the 'Private Browsing' mode recently added to most browsers such as Internet Explorer 8 and Firefox 3 still allows Flash cookies to operate fully and track the user. These differences make Flash cookies a more resilient technology for tracking than HTTP cookies, and creates an area for uncertainty for user privacy control.

It is important to differentiate between the varying uses of Flash cookies. These files (and any local storage in general) provides the benefit of allowing a given application to 'save state' on the users computer and provide better functionality to the user. Examples of such could be storing the volume level of a Flash video or caching a music file for better performance over an unreliable network connection. These uses are different than using Flash cookies as secondary, redundant unique identifiers that enable advertisers to circumvent user preferences and self-help.

With rising concern over "behavioral advertising," the US Congress and federal regulators are considering new rules to address online consumer privacy. A key focus surrounds users' ability to avoid tracking, but the privacy implications of Flash cookies has not entered the discourse.

Additionally, any consumer protection debate will include discourse on self-help. Thus, consumers' ability to be aware of and control unwanted tracking will be a key part of the legislative debate.

To inform this debate, we surveyed the top 100 websites to determine which were using Flash cookies, and explored the privacy implications. We examined these sites' privacy policies to see whether they discussed Flash cookies.

We also studied the privacy settings provided by Adobe for Flash cookies, in an effort to better understand the practical effects of using self-help to control Flash cookies. Because some sites rely so heavily on the use of Flash

content, users may encounter functionality difficulties as a result of enabling these privacy settings.

We found that Flash cookies are a popular mechanism for storing data on the top 100 sites. From a privacy perspective, this is problematic, because in addition to storing user settings, many sites stored the same values in both HTTP and Flash cookies, usually with telling variable names indicating they were user ids or computer guids (globally unique identifiers). We found that top 100 websites are using Flash cookies to "respawn,"[1] or recreate deleted HTTP cookies. This means that privacy-sensitive consumers who "toss" their H`TTP cookies to prevent tracking or remain anonymous are still being uniquely identified online by advertising companies. Few websites disclose their use of Flash in privacy policies, and many companies using Flash are privacy certified by TRUSTe.

Flash Cookies

Some exposition on Adobe's system for managing Flash cookies is necessary here.

Flash data is stored in a different folder on different computing platforms. For instance, on an Apple, Flash local shared objects (labeled .sol) are stored at:

/users/[username]/Library/Preferences/Macromedia/Flash Player/

On a Windows computer, they are stored at:

\Documents and Settings\[username]\Application Data\Macromedia\Flash Player

Several subdirectories may reside at that location: "#SharedObjects" contains the actual Flash cookies and subdirectories under "Macromedia.com" contains persistent global and domain-specific settings for how the Flash player operates. As such, there will be a subdirectory for each Flash-enabled domain a user visits under the "Macromedia.com" settings folder. This has privacy implications that will be visited in section IV(F) below.

A Flash cookie can be set when a websites embeds first party or third party Flash content on a page. For instance, a website may include animated Flash banner advertisements served by a company that leases the advertising space or they may embed a hidden SWF used solely to provide metrics on the user. Thus, merely visiting some websites (without actually clicking on an advertisement or video) can cause Flash data from a third party advertiser to be stored on the user's computer, often unbeknownst to the user.

Methods

We analyzed HTTP and Flash cookies from the top 100 domains ranked by QuantCast results of July 1, 2009. The data for this survey were captured on July 27, 2009.

[1] We use the popular gamer word "respawn" to describe the recreation of a HTTP cookie after its affirmative removal by the user.

We also analyzed six additional government websites: CDC.gov, DATA.gov, DHS.gov, IRS.gov, NASA.gov, and Whitehouse.gov. We took care not to leave the top-level domain when analyzing these sites. That is, the URL always displayed the domain to be analyzed during our browsing session.

Potential for Tracking

We used Mozilla Firefox 3.5 (release June 30, 2009) and Windows XP Professional Version 2002 Service Pack 3 for capturing data from the top 100 websites. To avoid contamination from different domains visited, we created a small program to handle the process of deleting all data stored between sessions since Firefox's "Clear Private Data" tool does not remove stored Flash objects.

Each session consisted of starting on a Firefox about:blank page with clean data directories. We then navigated directly to the site in question (by entering the domain name into the browser's navigation bar) and mimicked a 'typical' users session on that site for approximately 10 pages. For example, on a video site, we would search for content and browse videos. On a shopping site, we would add items to our shopping cart. We did not create accounts or login for any of the sites tested. As a result, we had to 'deep link' directly into specific user pages for sites such as Facebook.com or Myspace.com since typically these sites do not easily allow unauthenticated browsing.

We used SoThink SWF Catcher, a Firefox plugin which identifies all SWF files present on a webpage, to capture the Flash content encountered throughout the user session. We also quit the browser after each session and ran a program to capture the resulting persistent data such as HTTP cookies, Flash objects, and the Firefox cache.

Because of the dynamic nature of websites and online advertising, any given survey may produce different advertisements and correspondingly different Flash data from varied advertising networks. Thus, our snapshot of HTTP and Flash cookies may differ from another user's experience. However we feel that this provides reasonable sample for an initial study.

Respawning Deleted HTTP cookies

To manually test for HTTP cookie respawning, we used Safari 4.0.1 in a clean state (no HTTP or Flash cookies as well as no items in the browser cache) to visit a top 100 site. After browsing on the site and HTTP and Flash cookies are acquired, we deleted all HTTP cookies, cleared the cache, and restarted the browser, but did not modify the Flash cookies. We then visited the same site and noted the values of HTTP cookies set and whether they matched the Flash cookies set in the previous session.

Implications of Manipulating User Controls

We tested usability to explore how a hypothetical privacy-sensitive user's experience would differ if his/her settings were changed to restrict Flash cookies. The test was

performed using Mozilla Firefox with the BetterPrivacy 1.29 add-on installed. BetterPrivacy provides an easy-to-use interface to review, protect or delete Flash cookies. Flash player settings are controlled via a webpage on Adobe.com's website called the Adobe Flash Player: Settings Manager[8].

The user navigated to each of the top 100 websites and took notes of any pop-ups, broken content, or any other abnormalities experienced while browsing the site. Each session began with clearing all non-Adobe Flash Player shared object files (i.e. those not under the Macromedia.com folder), navigating to the site in question, and then mimicking a 'typical' user's session. Caution was taken not to navigate away from the domain of the site being tested. After each session, BetterPrivacy was checked for the appearance of any Flash cookies that may have been accumulated while browsing the site.

We attempted to identify changes in user-experience after restricting the ability for third party Flash objects from being stored on a user's computer (first party objects were still allowed). This option is enabled by: navigating to the Adobe Flash Player Settings Manager, locating the 'Global Storage Settings' option panel, then deselecting the option that reads, "Allow third party flash content to store data on your computer."

Results and Discussion

Presence of Flash and HTTP Cookies

We encountered Flash cookies on 54 of the top 100 sites. These 54 sites set a total of 157 Flash shared objects files yielding a total of 281 individual Flash cookies.

Ninety-eight of the top 100 sites set HTTP cookies (only wikipedia and wikimedia.org lacked HTTP cookies in our tests). These 98 sites set a total of 3,602 HTTP cookies.

Thirty-one of these sites carried a TRUSTe Privacy Seal. Of these 31, 14 were employing Flash cookies.

Thus, both HTTP and Flash cookies are a popular mechanism on top 100 websites.

Common Flash Cookie Variable Names

We attempted to infer the potential use of Flash cookies via examining the actual variable names for each cookie. Often, developers will use the term 'uid' or 'userid' to refer to a unique identifier whereas 'volume' could suggest volume settings for a music or video player. Below is a table of the most frequently occurring names in our sample.

Cookie Name	Frequency
volume	21
userid	20
user	14
id	8
lts	6
_tpf	6
_fpf	6
uid	5
perf	5
computerguid	5

The most frequently occurring Flash cookie outside of those used in the Flash Player system directory was 'volume'. Given the dominance of Flash video on the web, it is reasonable to expect that volume settings would be a commonly occurring use of Flash cookies. However, it is surprising with which the prominence of Flash cookies such as 'userid, user, and id', which were found to store unique identifiers which could be used to track the user, were found. It's also worth mentioning that '_tpf' and '_fpf' were found to also contain unique identifiers which were also found to contain overlapping values as the ones found in HTML cookies for 'uid' or 'userid'.

Shared Values Between HTTP and Flash Cookies

Of the top 100 websites, 31 had at least one overlap between a HTTP and Flash cookie. For instance, a website might have an HTTP cookie labeled "uid" with a long value such as 4a7082eb-775d6-d440f-dbf25. There were 41 such matches on these 31 sites.

Most Flash cookies with matching values were served by third-party advertising networks. That is, upon a visit to a top 100 website, a third party advertising network would set both a third party HTTP cookie and a third party Flash cookie. Our tests revealed 37 matching HTTP and Flash values from the following advertisers: ClearSpring (8), Iesnare (1), InterClick (4), ScanScout (2), SpecificClick (14), QuantCast (6), VideoEgg (1), and Vizu (1).

In 4 cases, the following first-party domains HTTP cookies matched Flash cookie values: Sears, Lowe's, AOL, and Hulu.

Flash Cookie Respawning

Shared values between HTTP and Flash cookies raises the issue of whether websites and tracking networks are using Flash cookies to accomplish redundant unique user tracking. That is, storing the same values in both the Flash and HTTP cookie would give a site the opportunity to backup HTTP cookies if the user deleted them.

We found that taking the privacy-conscious step of deleting HTTP cookies to prevent unique tracking could be circumvented through "respawning" (See Figures 1-3). The Flash cookie value would be rewritten in the standard HTTP cookie value, thus subverting the user's attempt to prevent tracking.

We found HTTP cookie respawning on several sites.

On About.com, a SpecificClick Flash cookie respawned a deleted SpecificClick HTTP cookie. Similarly, on Hulu.com, a QuantCast Flash cookie respawned a deleted QuantCast HTTP cookie.

We also found HTTP cookie respawning across domains. For instance, a third-party ClearSpring Flash cookie

respawned a matching Answers.com HTTP cookie. ClearSpring also respawned HTTP cookies served directly by Aol.com and Mapquest.com. InterClick respawned a HTTP cookie served by Reference.com

Interaction with NAI Opt-Out

"The NAI (Network Advertising Initiative) is a cooperative of online marketing and analytics companies committed to building consumer awareness and establishing responsible business and data management practices and standards."[9] Since some of the sites using Flash cookies also belong to the NAI, we tested the interaction of Flash cookies with the NAI opt-out cookie.

We found that persistent Flash cookies were still used when the NAI opt-out cookie for QuantCast was set. Upon deletion of cookies, the Flash cookie still allowed a respawn of the QuantCast HTML cookie. It did not respawn the opt-out cookie. Thus, user tracking is still present after individuals opt out.

Presence of Flash Settings Files

Adobe Flash settings files (those in the Macromedia.com folder) were set by Flash player in visits to 89 of the top 100 sites. A total of 201 settings files were present among these 89 sites. This is relevant, because each settings file is stored in its own directory, labeled by domain. This creates a type of history file parallel to the one created by the browser. However, the Flash history is not deleted when browser controls are used to erase information about sites previously visited. This means that users may falsely believe that they have fully cleared their history when using the standard browser tools.

Privacy Policies

We searched the privacy policies of the top 100 sites, looking for terms such as "Flash," "PIE," or "LSO." Only 4 mentioned the use of Flash as a tracking mechanism.

Given the different storage characteristics of Flash cookies, without disclosure of Flash cookies in a privacy policy, it is unclear how the average user would even know of the technology. This would make privacy self-help impossible except for sophisticated users.

Government Sites

The Obama Administration is considering whether to change policy concerning the use of HTTP cookies on government websites. Currently, government officials require a "compelling need" to use persistent HTTP cookies, and must disclose their use in a privacy policy.

In light of this we arbitrarily chose six government websites to determine whether Flash was being used to assign unique values to visitors. Of the 6 government sites we tested, 3 had Flash cookies. Three were set by whitehouse.gov, one of which was labeled, "userId." Five of these sites used persistent HTTP cookies.

Whitehouse.gov disclosed the presence of a tracking technology in its privacy policy, but the policy does not specify that Flash cookies are present, nor does it provide any information on how to disable Flash cookies.[10]

User Experience

Since users generally do not know about Flash cookies, it stands to reason that users lack knowledge to properly manage them. In comments to the New York Times, Emmy Huang of Adobe said, "It is accurate to say that the privacy settings people make with regards to their browser activities are not immediately reflected in Flash Player. Still, privacy choices people make for their browsers aren't more difficult to do in Flash Player, and deleting cookies recorded by Flash Player isn't a more difficult process than deleting browser cookies. However, it is a different process and people may not know it is available."[11]

A separate issue arises with user controls: if a privacy sensitive individual knows about them and employs them, will they still be able to use the internet normally?

When disabling third party content, we found that 84 of the sites had no functionality issues after third-party Flash content was disabled. Sixteen sites stored some type of Flash data.

Ten sites did not function optimally with third party context storage disabled. Nine of these 10 sites would not display Flash content. One site displayed an advertisement intermittently that never stabilized.

Conclusion

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.

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.

Acknowledgments

This work was supported in 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 Summer Undergraduate Program in Engineering Research at Berkeley (SUPERB), and to its program leaders, Dr. Kristen Gates, Dr. Sheila M. Humphreys, and Beatriz Lopez-Flores.

References

[1] Shannon Canty is a senior at Clemson University majoring in bioengineering.

[2] Quentin Mayo is a senior at Jacksonville State University majoring in computer science.

[3] Lauren Thomas is a senior at Louisiana State University majoring in industrial engineering.

[4] M. Abraham, C. Meierhoefer, and & A. Lipsman, "The Impact of Cookie Deletion on the Accuracy of Site-Server and Ad-Server Metrics: An Empirical Comscore Study," 2007, available at http://www.comscore.com/Press_Events/Presentations_Whitepapers/2007/Cookie_Deletion_Whitepaper.

[5] United Virtualities, "United Virtualities develops ID backup to cookies, Browser-Based 'Persistent Identification Element' will also restore erased cookie, Mar. 31, 2005, available at http://www.unitedvirtualities.com/UV-Pressrelease03-31-05.htm.

[6] Antone Gonslaves, "Company bypasses cookie-deleting consumers, March 31, 2005, available at http://www.informationweek.com/news/security/privacy/showArticle.jhtml?articleID=160400801.

[7] J. Lott, D. Schall, and K. Peters, Actionscript 3.0 Cookbook, O'Reilly, 2006, p. 410.

[8] Adobe, Flash Settings Manager, available at http://www.macromedia.com/support/documentation/en/flashplayer/help/settings_manager.html

[9] NAI, About the NAI, available at http://www.networkadvertising.org/about/.

[10] The White House, Our Online Privacy Policy, n.d., available at http://www.whitehouse.gov/privacy/

[11] Stone, Brad. "Adobe's Flash and Apple's Safari Fail a Privacy Test." Technology - Bits Blog – NYTimes.com. 30 Dec. 2008. 23 June 2009 http://bits.blogs.nytimes.com/2008/12/30/adobes-flash-and-apples-safari-fail-a-privacy-test

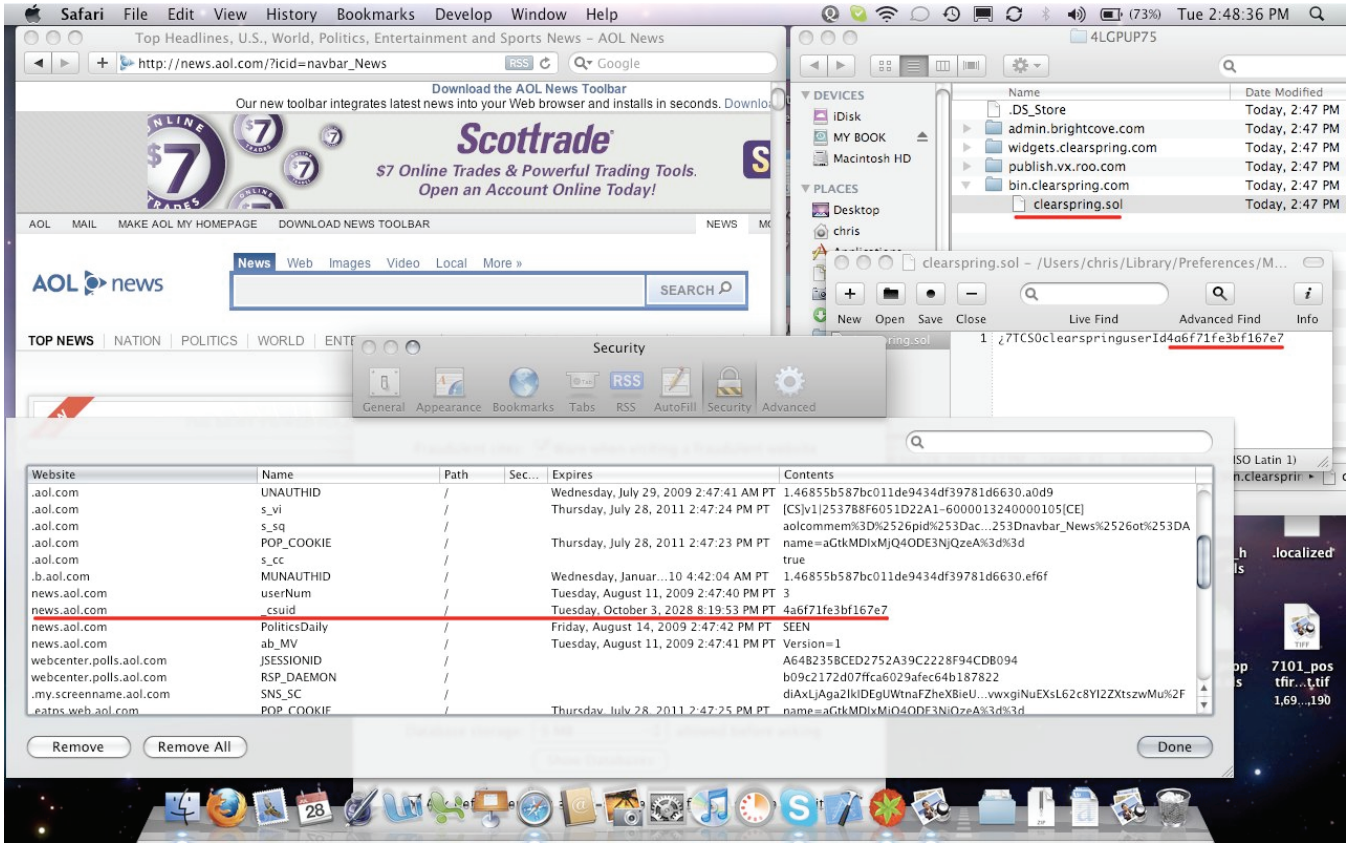

Figure 1: A matching Flash and HTTP cookie is set by AOL.com and ClearSpring.

Figure 2: The researcher deleted HTTP cookies and cleared the cache, leaving the Flash cookies unaltered

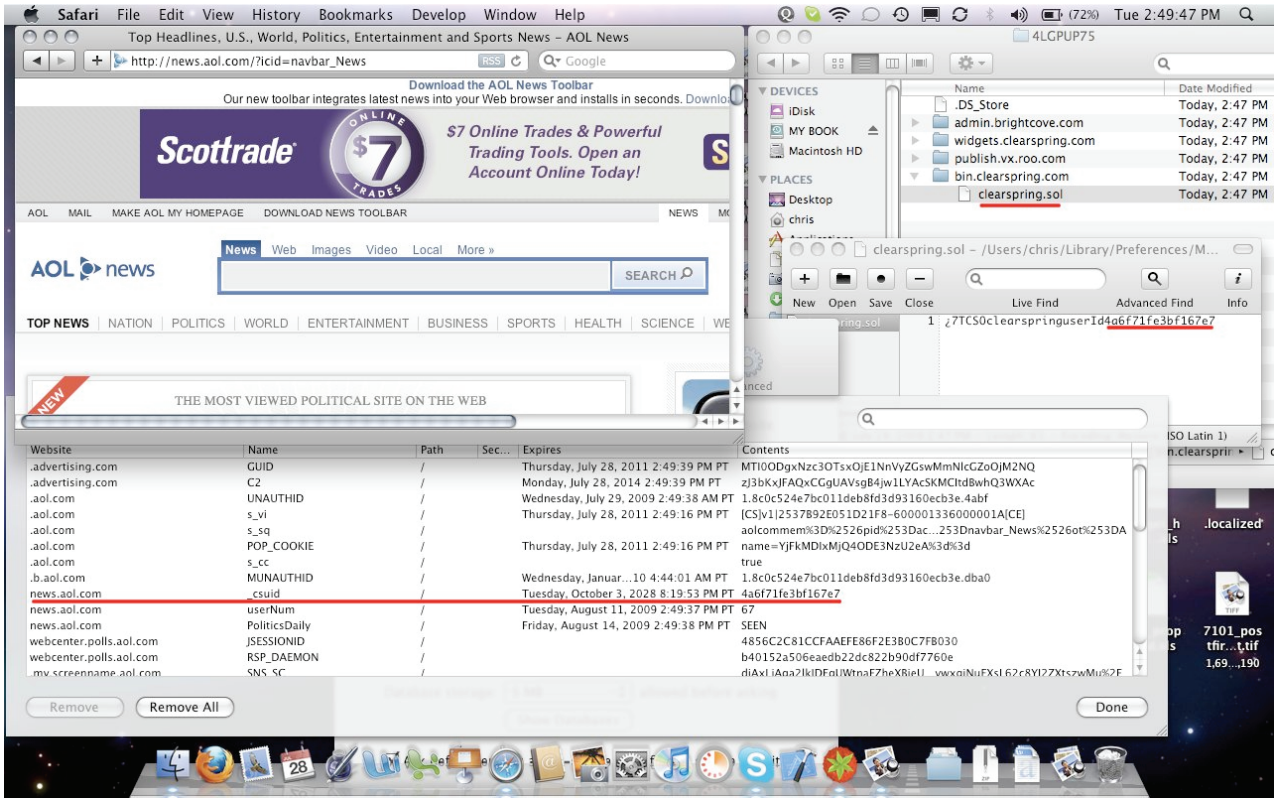

Figure 3: Upon revisiting AOL.com, a new HTTP cookie is set with the same value before HTTP cookies were deleted