All Your Browsing History Are Belong to Us

by Justin Brookman [1]
March 23, 2010

For several years, it has been a poorly kept secret that any Web site you went to could secretly search your browser's history file to see what sites you had previously visited. All the site owner had to do was ask. And while browser history "sniffing" has been around for a long time, companies are finally starting to actively take advantage of it. The time to act to prevent this clear threat to personal privacy is now.

The History of Browser History Sniffing

Browser history sniffing exploits the functionality of all Web browsers that displays hyperlinks of visited and non-visited sites in different colors. That is, when you visit a Web site that contains links to a number of other urls, the links to sites you have not previously visited will be shown in blue, while the links to sites that you had previously visited will be shown in purple. The links appear this way because the Web page is allowed to query to user's browser history in order to know what color to render the links on the Web page. Web sites can game this functionality by listing hundreds of Web addresses (often hidden to the user, who doesn't see the links at all, blue or purple) to get answers from the user's browser about what color to display the links. In this way, Web sites can effectively play "go fish" with a user's browser history file, asking if the visitor has visited www.facebook.com [2], or www.nytimes.com [3], or, perhaps more personally, www.viagra.com [4] or www.gamblersanonymous.org [5]. If you're curious to see how it works, the site www.whatthelnternetknowsaboutyou.com [6] provides several useful demonstrations.

The existence of this trick to query whether site visitors have visited a predetermined list of urls has been known for a long time. It has been identified as "Bug 147777 [7]" in Mozilla's development forum for nearly eight years. And for years, researchers [8] and privacy advocates [9] have ask that the issue be addressed. To date, nothing has been done, and unscrupulous Web site owners still maintain the capacity to determine whether site visitors have previously visited any other Web site.

Quite apart from the fact that Web sites don't have the right to see where you've been on the Web, there are real dangers to unrestricted access to browser history files. Identity thieves could find out what bank and credit card sites you use for better targeted phishing attacks [10]. Furthermore, recent research suggests that sites could use data about visited urls to accurately determine the identity of site visitors. One study [11] released last month shows how a site could correlate browser history queries with publicly available information about group membership on popular social networking sites to reliably identify a large percentage of visitors to a particular Web site. Thus, if you're active on any number of popular social networking sites, any Web site you try to visit anonymously could very likely figure out who you are.

Browser History Sniffing Goes Mainstream

The case to address this vulnerability has taken on greater urgency in recent weeks with news that Web analytics companies are starting to market products that directly take advantage of this hack. Last month, Eric Peterson reported [12] on an Israeli firm named Beencounter that openly sells a tool to Web site developers to query whether site visitors had previously visited up to 50 specific urls.

Another company called Tealium has already been marketing a product taking advantage of this exploit for nearly two years. Tealium's "Social Media" service runs daily searches of a customer's name for news and blog postings mentioning the customers, and then runs a JavaScript application on the customer's site to determine whether visitors had previously read any of those stories. The service allows Tealium customers a unique insight into what sites visitors had previously read about the company that may have driven them to the company's Web site. On the other hand, Tealium achieves this insight by rifling through a site visitor's browser history file without disclosure to or

permission from consumers.

When CDT recently visited Tealium's Web site, we determined they were using their own Social Media product to search our Web browser's history file. Tealium checked whether we had visited over 100 news or blog posts that had mentioned the company (because we were interested in learning about the company, we had, in fact, visited many of those sites). This search of our browsing history was conducted in the background with no visibility to the user --- we only knew about the search because we were using a JavaScript-monitoring Firefox plug-in to watch for it.

None of the websites we detected running Tealium's script provided upfront, clear and conspicuous disclosure about the browser history sniffing, nor did any ask for permission before searching through visitors' history files. Privacy policies for the companies running the scripts offered either misleading descriptions of the browser history sniffing, or none at all. Tealium's own privacy policy describing the Social Media product was written in vague, confusing terms, though Tealium has told CDT it intends to make its privacy policy disclosure more clear. Tealium does offer Web consumers the opportunity to download an opt-out cookie [13] to disable browser history searching on sites utilizing the Tealium Social Media service, but absent any meaningful disclosure to users whose files are being searched, Internet consumers would have no reason to seek out the Tealium opt-out cookie to avoid being snooped upon.

At the moment, Tealium's product only offers aggregate data to customers for Web analytics purposes, so companies cannot use Tealium to determine whether particular Web visitors have visited certain sites or not. However, there are obviously no inherent limitations on the use of the browser history sniffing. Indeed, Tealium has filed a pending patent application [14] asserting a wide range of other uses of its product, including behavioral advertising and employee monitoring. If companies are allowed to freely search site visitors' browser history files, it is only a matter of time before we start to see even more invasive applications of this exploit.

The Need for Redress

For years, Web site developers have had the ability to search an incoming visitor's history file. Now it appears that some Internet companies are taking real advantage of this exploit. The time to act is long overdue.

- Browsers
  Browser makers have been aware of this problem for years, but, to date, have failed to act to protect their user's privacy. They have a responsibility to act with all due haste to fix it now. Researchers have previously released browser add-ons [15] (no longer supported) to address the privacy hole, but no browser has incorporated functionality to protect history files from snooping.

  Fortunately, Mozilla, at least appears to be working actively [7] to solve the problem for Firefox after several years of inaction. We call on the makers of the other Web browsers to commit resources to plug the holes in their respective products as well.

- Web site operators
  Web site operators (and the marketing and analytics companies they hire) must stop taking advantage of this security hole in direct opposition to consumers' reasonable expectations about the privacy of their previous Internet communications. Querying site visitor browser history is a line that should not be crossed.

  Certainly, companies may seek to legitimately gain user consent before scanning history files, but such consent should be given on an opt-in basis, and only after clear and conspicuous disclosure about what information about the consumer the company intends to access and what the company intends to do with such information. Companies cannot hide notice of browser history inquiries deep in terms of service or privacy policies where consumers are unlikely to notice the disclosures.

- Law enforcement
  Finally, if browsers continue to fail to protect their users, and Internet marketing companies continue to gain access to Web visitors' history files, then law enforcement should step in and bring affirmative cases against companies that gain access to such data without user permission. Taking advantage of a security hole by embedding hidden urls in a Web site to snoop through a visitor's history file has all the markings of a Section 5 claim --- either as a deceptive practice or under the FTC's unfairness authority. Subsequent aggregation or anonymization of user data does not absolve companies from searching through consumer's hard drives without authorization, nor does boilerplate and inconspicuous disclosure in a lengthy, legalistic privacy policy. The Federal Trade Commission has done excellent work in recent years holding companies responsible for unauthorized tracking of consumers' Internet behavior. [16] If companies continue to exploit browser security holes to snoop on site visitors' past Web surfing, the FTC (and state Attorneys General) have a responsibility to act to protect consumers' reasonable expectations of privacy.

-
- sniffing
- javascript
- History
- browser sniffing

Links:
[1] https://www.cdt.org/personnel/justin-brookman
[2] http://www.facebook.com
[3] http://www.nytimes.com
[4] http://www.viagra.com
[5] http://www.gamblersanonymous.org
[6] http://www.whattheInternetknowsaboutyou.com
[7] https://bugzilla.mozilla.org/show_bug.cgi?id=147777
[8] http://crypto.stanford.edu/safecache/sameorigin.pdf
[9] http://www.cdt.org/files/pdfs/20091221_ftc_comments_privacy_design.pdf
[10] http://www.ravenwhite.com/files/invasivesniff052.pdf
[11] http://www.darkreading.com/insiderthreat/security/client/showArticle.jhtml?articleID=223100436&cid=RSSfeed
[12] http://tech.Web analyticsdemystified.com/2010/02/know-where-your-visitors-have-been-beencounter.html
[13] http://www.tealium.com/privacy.html
[14] http://www.faqs.org/patents/app/20090287713
[15] https://addons.mozilla.org/en-US/firefox/addon/1502
[16] http://www.ftc.gov/opa/2009/06/sears.shtm