Your web surfing history accessible via JavaScript: researchers

December 3 2010

"JavaScript is a great thing, it allows things like Gmail and Google Maps and a whole bunch of Web 2.0 applications; but it also opens up a lot of security vulnerabilities. We want to let the broad public know that history sniffing is possible, it actually happens out there, and that there are a lot of people vulnerable to this attack," said UC San Diego computer science professor Sorin Lerner.

(PhysOrg.com) -- The Web surfing history saved in your Web browser can be accessed without your permission. JavaScript code deployed by real websites and online advertising providers use browser vulnerabilities to determine which sites you have and have not visited, according to new research from computer scientists at the University of California, San Diego.

The researchers documented JavaScript code secretly collecting browsing histories of Web users through "history sniffing" and sending

that information across the network. While history sniffing and its potential implications for privacy violation have been discussed and demonstrated, the new work provides the first empirical analysis of history sniffing on the real Web.

“Nobody knew if anyone on the Internet was using history sniffing to get at users’ private browsing history. What we were able to show is that the answer is yes,” said UC San Diego computer science professor Hovav Shacham.
The computer scientists from the UC San Diego Jacobs School of Engineering presented this work in October at the 2010 ACM Conference on Computer and Communications Security (CCS 2010) in a paper entitled, “An Empirical Study of Privacy-Violating Information Flows in JavaScript Web Applications”.

History Sniffing

History sniffing takes place without your knowledge or permission and relies on the fact that browsers display links to sites you’ve visited differently than ones you haven’t: by default, visited links are purple, unvisited links blue. History sniffing JavaScript code running on a Web page checks to see if your browser displays links to specific URLs as blue or purple.

History sniffing can be used by website owners to learn which competitor sites visitors have or have not been to. History sniffing can also be deployed by advertising companies looking to build user profiles, or by online criminals collecting information for future phishing attacks. Learning what banking site you visit, for example, suggests which fake banking page to serve up during a phishing attack aimed at collecting your bank account login information.

“JavaScript is a great thing, it allows things like Gmail and Google Maps

and a whole bunch of Web 2.0 applications; but it also opens up a lot of security vulnerabilities. We want to let the broad public know that history sniffing is possible, it actually happens out there, and that there are a lot of people vulnerable to this attack," said UC San Diego computer science professor Sorin Lerner.

The latest versions of Firefox, Chrome, and Safari now block the history sniffing attacks the computer scientists monitored. Internet Explorer, however, does not currently defend against history sniffing. In addition, anyone using anything but the latest versions of the patched browsers is also vulnerable.

Sniffing out History Sniffing

"We built a dynamic data flow engine for JavaScript to track history sniffing in the wild. I don't know of any other practical tool that can be used to do this kind of extensive study," said Dongseok Jang, the UC San Diego computer science Ph.D. student who developed the JavaScript monitoring technology. The researchers plan to broaden their work and study what information is being leaked by applications on social media and other Web 2.0 sites.

The computer scientists looked for history sniffing on the front pages of the top 50,000 websites, according to Alexa global website rankings. They found that 485 of the top 50,000 sites inspect style properties that can be used to infer the browser's history. Out of 485 sites, 63 transferred the browser's history to the network. "We confirmed that 46 of them are actually doing history sniffing, one of these sites being in the Alexa global top 100," the UC San Diego computer scientists write in the CCS 2010 paper.

Table 1 in the paper outlines the websites the computer scientists found that performed history sniffing during the data collection period. In

some cases, the websites created their own history sniffing systems. In other cases, advertisements served by outside companies contained JavaScript code performing the history sniffing.

History Sniffing in Perspective

The computer scientists say that history sniffing does not pose as great a risk to your privacy or identity as malicious software programs (malware) that can steal your banking information or your entire Facebook profile. But, according to Shacham, “history sniffing is unusual in effectively allowing any site you visit to learn about your browsing habits on any other site, regardless if the two sites have any business relationship.”

To see history sniffing in action, visit: www.whattheinternetknowsaboutyou.com .

“I think people who have updated or switched browsers should now worry about things other than history sniffing, like keeping their Flash plug-in up to date so they don’t get exploited. But that doesn’t mean that the companies that have engaged in history sniffing for the currently 60 percent of the user population that is vulnerable to it should get a free pass,” said Shacham.

Tracking History Sniffing

The UC San Diego history-sniffing detection tool analyzes the JavaScript running on the page to identify and tag all instances where the browser history is being checked. The way the system tags each of these potential history tracking events can be compared to the ink or paint packets that banks add to bags of money being stolen.

“As soon as a JavaScript tries to look at the color of a link, we immediately put ‘paint’ on that. Some sites collected that information but never sent it over the network, so there was all this ‘paint’ inside the browser. But in other cases, we observed ‘paint’ being sent over the network, indicating that history sniffing is going on,” explained Lerner. The computer scientists only considered it history sniffing when the browser history information was sent over the network to a server. “We detected when browser history is looked at, collected on the browser and sent on the network from the browser to their servers. What servers then do with that information is speculation,” said Lerner.

The “paint” tracking approach to monitoring JavaScript could be useful for more than just history sniffing, Lerner explained. “It could be useful for understanding what information is being leaked by applications on Web 2.0 sites. Many of these apps use a lot of JavaScript.”

More information: Dongseok Jang, Ranjit Jhala, Sorlin Lerner, and Hovav Shacham. “An Empirical Study of Privacy-Violating Information Flows in JavaScript Web Applications.” In A. Keromytis and V. Shmatikov, eds., Proceedings of CCS 2010, pages 270–83. ACM Press, Oct. 2010.

Provided by University of California - San Diego