Computer scientists identify Yelp security leak

November 4 2011

Computer scientists at Harvard, Boston University, and Yale stumbled upon a privacy leak in the mobile version of the popular Yelp social networking review site (m.yelp.com) in late October.

In the course of their ongoing research, which studies the interplay between social networks and Internet commerce, the team—Michael Mitzenmacher, Gordon McKay Professor of Computer Science at the Harvard School of Engineering and Applied Sciences; John Byers, Associate Professor of Computer Science at Boston University; and Giorgos Zervas, Simons Postdoctoral Fellow at Yale University and an Affiliate at the Center for Research on Computation and Society at Harvard—inadvertently found a servlet on m.yelp.com that could reveal some user information that was intended to be private.

Data at risk included certain user-specific fields such as email addresses, birth dates, gender, and full names. Even though no financial information was leaked, the team felt that the exposure of personally identifiable information presented a major threat. After double-checking the finding they alerted Yelp.

The group then worked with the company's engineers to help them gain a fuller understanding of the problem, which was then resolved with a workaround the very same day.

"Yelp's team responded in an exemplary fashion," says Mitzenmacher. "After we contacted them, Yelp's Michael Stoppelman and members of

the engineering staff listened to our presentation and description of the vulnerability seriously, and, as they describe in their blog post, took immediate action to correct the problem."

The researchers also noted Yelp's willingness to make the issue public to help alert users and to prevent any possible related problems on similar websites.

Mitzenmacher and Byers give full credit to Zervas for identifying the privacy risk. He came across the vulnerability in the course of a case study on Yelp as a site that provides economic information in the form of user-generated reviews.

"As part of our research and data collection, Giorgos [Zervas] was looking at Yelp's various interfaces, including the mobile web site," explains Mitzenmacher. "To be clear, he was not 'hacking' the site in any way—just interacting with it via a standard browser and normal HTTP requests."

Zervas, using an HTTP logger (a standard browser tool that allows a user to watch the exchange of data between the browser and the web servers), discovered that when he checked a particular restaurant for reviews and then clicked on the button asking for more reviews, entire reviewer records were leaked in JSON (JavaScript Object Notation) format. Those records contained non-encrypted information such as email addresses, gender, birth dates, and full names.

Ordinary users accessing the site from a mobile device would not have seen such sensitive information, as client-side JavaScript displayed only the non-sensitive information (such as the review text, date, and the reviewer's handle).

In the blog posting, Yelp's Stoppelman writes that the company

engineers “analyzed the servlet’s access logs to see if anyone exploited the hole...[and] did not find any evidence that user information had actually been collected.”

“This example shows the importance of having multiple redundant layers of security when handling personally identifiable information,” says Mitzenmacher. “In the Yelp post, they describe the redundancies they have added to prevent such leakage in the future.”

Provided by Harvard School of Engineering and Applied Sciences