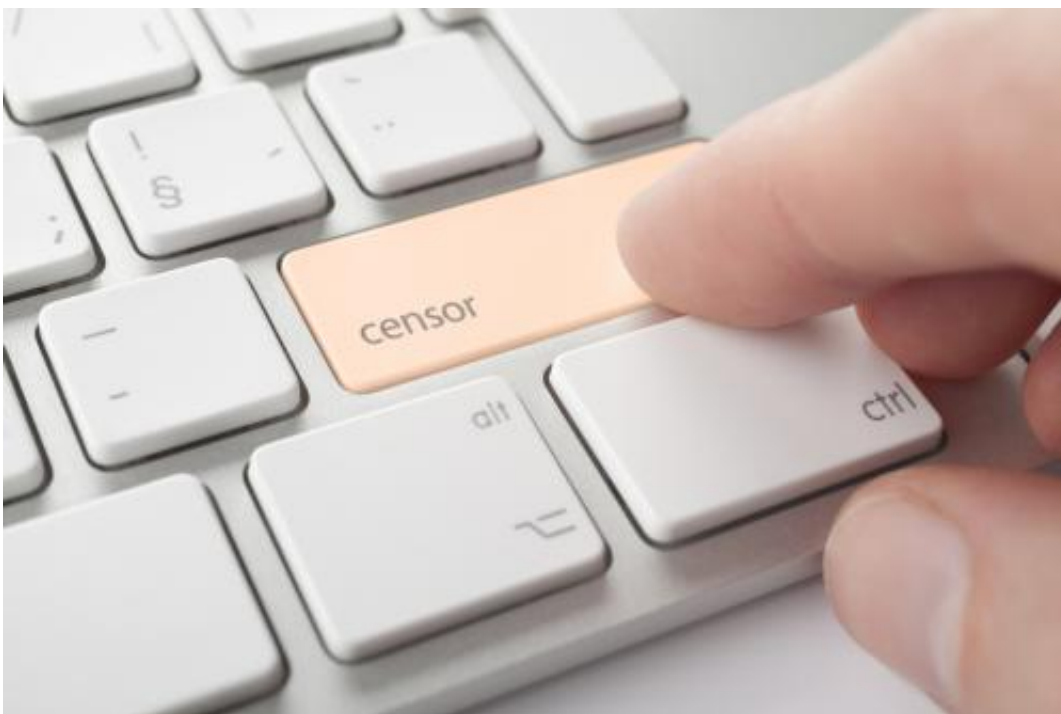# Study shows just how fast censorship can occur in social media

March 14 2013, by Jade Boyd

A new study analyzes how controversial posts are deleted in near real time from the Twitter-like Chinese social media service called Weibo, which hosts about 100 million messages per day. Credit: Photos.com/Rice University

(Phys.org) —An analysis of censorship patterns on the Twitter-like Chinese social media service called Weibo gives the clearest picture yet of how the site's operator, Sina Weibo, finds and deletes controversial posts in near real time, despite a daily volume of 100 million messages. The study, which was conducted by an independent researcher and

collaborators at Rice University and the University of New Mexico (UNM), is available online and undergoing peer review.

"Other people have explored censorship on Weibo, but this work is focused on the speed at which censorship happens," said lead researcher Dan Wallach, professor of computer science at Rice and co-author of a forthcoming study that was recently posted on the pre-print site arxiv.org.

A team led by Wallach and UNM's Jed Crandall worked with the study's lead author, an independent researcher named Tao Zhu. Their analysis indicates that Sina Weibo uses a combination of keyword-matching software and human censors to monitor and delete potentially controversial posts on Weibo. By closely monitoring individuals who frequently post controversial messages, Sina Weibo is able to delete many objectionable posts in less than five minutes, the study found.

Launched three years ago, Weibo, like Twitter, allows users to post 140-character messages with usernames and hashtags. About 300 million people use Weibo, which is China's most popular microblogging service. Users post 100 million messages each day on Weibo.

For the study, researchers began by following 25 "sensitive" users that they had discovered by doing a search for people who had used words previously banned by Weibo. To broaden their search, the researchers added more than 3,000 users who had reposted one of the 25 sensitive users more than five times. They then followed this expanded group for a period of time and measured how often and how quickly their posts were deleted. Any user with more than five deleted posts was added to the pool of sensitive users.

After 15 days, the sensitive group included 3,567 users. The researchers found that on average, about 4,500 posts by the sensitive users were

deleted each day, including about 1,500 that were deleted at the network level by Sina Weibo. The team's censorship-tracking software was able to track, within one minute, the amount of time a post remained online before it was deleted.

The researchers found that deletions happened most heavily in the first hour after an original post had been made, and nearly 90 percent of deletions occurred within 24 hours. The analysis also revealed a sophisticated mechanism to remove all reposts of deleted posts, often within five minutes of the original post's deletion. Deletion times were found to be significantly shorter for a subset of users who tended to post deleted content most often, an indication that Sina Weibo actively monitors the activity of some users.

"Roughly 12 percent of the total posts from our sensitive users were eventually deleted," Wallach said. "We have enough of these posts to be able to run topical analysis algorithms that let us extract the main subjects that Weibo's censors seemed concerned with on any given day."

To date, researchers have collected 470 million posts from the Weibo public timeline and 2.38 million posts from a user timeline.

"Measuring how censorship is practiced by Weibo allows us to examine industry practices that may become more widespread," said Crandall, assistant professor of computer science at UNM. "There has been considerable debate in the U.S. recently about extending copyright law enforcement to include various kinds of filtering online. China already has laws in place for companies within China to filter online content."

Wallach and Crandall said that Weibo, as one of China's largest social networking companies, faces the dual challenge of keeping its users engaged—and thereby watching advertisements and making money for Weibo—while keeping the content it hosts compliant with local laws.

"Weibo gives us a window into the future for what Internet censorship of [social media](#) around the world may look like," Wallach said. "Former Supreme Court Justice Louis Brandeis championed transparency a century ago when he wrote, 'Sunlight is said to be the best of disinfectants.' We hope that our research shines a light on how laws created by governments and implemented by the private sector can affect free speech everywhere, including here in the U.S."

More information: Research available for download:

The Velocity of Censorship: High-Fidelity Detection of Microblog Post Deletions
arxiv.org/abs/1303.0597
Tracking and Quantifying Censorship on a Chinese Microblogging Site
arxiv.org/pdf/1211.6166v1.pdf
An Analysis of Chinese Search Engine Filtering
arxiv.org/pdf/1107.3794.pdf

Provided by Rice University