First large-scale analysis of 'soft' censorship of social media in China

March 7 2012

Researchers in Carnegie Mellon University's School of Computer Science analyzed millions of Chinese microblogs, or "weibos," to uncover a set of politically sensitive terms that draw the attention of Chinese censors. Individual messages containing the terms were often deleted at rates that could vary based on current events or geography.

The study is the first large-scale analysis of political content censorship in social media, a topic that drew attention and controversy earlier this year when Twitter announced a country-by-country policy for removing tweets that don't comply with local laws.

In China, where online censorship is highly developed, the researchers found that oft-censored terms included well-known hot buttons, such as Falun Gong, a spiritual movement banned by the Chinese government, and human rights activists Ai Weiwei and Liu Xiaobo. Others varied based on events; Lianghui, a term that normally refers to a joint meeting of China's parliament and its political advisory body, became subject to censorship when it emerged as a code word for "planned protest" during pro-democracy unrest that began in February 2011.

The CMU study also showed high rates of weibo censorship in certain provinces. The phenomenon was particularly notable in Tibet, a hotbed of political unrest, where up to 53 percent of locally generated microblogs were deleted.

The study by Noah Smith, associate professor in the Language

Technologies Institute (LTI); David Bamman, a Ph.D. student in LTI; and Brendan O'Connor, a Ph.D. student in the Machine Learning Department, appears in the March issue of *First Monday*, a peer-reviewed, online journal.

"A lot of studies have focused on censorship that blocks access to Internet sites, but the practice of deleting individual messages is not yet well understood," Smith said. "The rise of domestic Chinese microblogging sites has provided a unique opportunity to systematically study content censorship in detail."

The so-called Great Firewall of China, which prevents Chinese residents from accessing foreign websites such as Google and Facebook, is China's best known censorship tool. Other countries also are known to block Web access, such as when Egypt shut down Twitter and other social media sites during last year's Arab Spring protests.

But blocking access to all sites and services is impossible if China or any other country is to harness the Web's commercial and educational potential, Bamman said. An alternative is to allow access to sites, but police the content, eliminating messages deemed objectionable. Automated methods may be used to eliminate some messages, while others are deleted manually, he noted. Seldom are all weibos with a sensitive term deleted, but anecdotal evidence is overwhelming that certain messages are targeted.

"You even see some weibos where the writer asks, 'Is this going to be deleted?'" O'Connor said. In late 2010, New York Times columnist Nicholas Kristof opened an account on a Chinese microblog site; within an hour of sending a message about Falun Gong, his account was shut down.

To study this "soft" censorship, the CMU team analyzed almost 57

million messages posted on Sina Weibo, a domestic Chinese microblog site similar to Twitter that has more than 200 million users. They collected samples of weibos from June 27 to Sept. 30, 2011, using an application programming interface (API) that Sina Weibo provides to developers so they can build related services.

Using the same API, they later checked a random subset of weibos to see if they still existed and another subset that included terms known to be politically sensitive. If a weibo was deleted, Sina would return what the researchers came to regard as an ominous message: "target weibo does not exist."

In late June and early July, for instance, rumors began circulating of the death of Jiang Zemin, a former general secretary of the Communist Party of China who came to power during the Tiananmen Square protests of 1989. On July 6, at the height of the rumor, 64 of 83 messages containing his name were deleted; on July 7, 29 of 31 such messages were deleted.

As another check, the researchers compared the frequency of such messages on Sina Weibo with those on the Chinese language version of Twitter, which officially is blocked by China but can still be accessed by net-savvy people. On July 6, Jiang's name appeared in one out of every 75 tweets, but just one out of every 5,666 messages on Sina Weibo – another indication that the Jiang conversations on Sina Weibo are suppressed.

Many weibos with high deletion rates included terms and names known to be politically sensitive, such as Fang Binxing, the architect of the Great Firewall of China, and references to state propaganda. Others reflect sensitivity to events; a term meaning "to ask someone to resign" became subject to deletion following the high-speed rail crash that killed 40 people in Wenzhou last July and apparently referenced the minister

of railways.

Censored terms are not always political. Following the March 2011 Fukushima nuclear disaster in Japan, weibos containing such politically innocuous terms as iodized salt and radioactive iodine had high deletion rates. The researchers believe these deletions were the result of government efforts to quash false rumors about the nuclear accident causing salt contamination.

Not all deletions are necessarily state-instigated censorship, the researchers noted. Spam and pornographic messages also are subject to deletion, just as they are in the United States.

By establishing a methodology for studying soft censorship in China, the researchers say they now have a means for actively monitoring social media censorship as it changes over time. They also may have the means to probe deeper, identifying code words and metaphors used to sidestep censors.

Provided by Carnegie Mellon University

Citation: First large-scale analysis of 'soft' censorship of social media in China (2012, March 7) retrieved 23 April 2024 from https://phys.org/news/2012-03-large-scale-analysis-soft-censorship-social.html

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.