Assessing Censorship on Microblogs in China

Discriminatory Keyword Analysis and the Real-Name Registration Policy

The authors investigated the use of microblogs – or *weibos* – and related censorship practices using 111 million microblogs collected between 1 January and 30 June 2012. Using a matched case-control study design helped researchers determine a list of Chinese terms that discriminate censored and uncensored posts written by the same microbloggers. This list includes homophones and puns created by Chinese microbloggers to circumvent the censors successfully. The study's design also allowed for evaluating the new real-name registration system's impact on microbloggers' posting activities. Findings suggest that this policy might have stopped some microbloggers from writing about social and political subjects.

King-wa Fu, Chung-hong Chan, and Michael Chau
University of Hong Kong

China is known as a society in which the government strictly regulates news media and often censors public information. International bodies routinely rank the country's citizens as having the least freedom of speech and freedom of the press. However, on Internet platforms, such as online forums and blogs, the Chinese people seem to have more autonomy to speak on public affairs and can occasionally set a social agenda that successfully draws media and public attention.[1] Examples include the "My father is Li Gang" or "Guo Meimei" incidents that were widely covered in the Western media. These cases support a notion that the Internet in China has played an overarching role in constructing China's public sphere, empowering civil discourse, and helping build the public agenda.[2,3] Recently, China's Twitter equivalent microblogging service, *weibo*, has been enthusiastically depicted by Western media as a new "free speech platform."[4] This optimistic view is, however, challenged by the Chinese authorities' ubiquitous mechanisms for controlling the public information flow, including the sophisticated filtering system known as the *Great Firewall*[5] and a new requirement that microbloggers must register with their real identities.[6]

Published by the IEEE Computer Society
1089-7801/13/$31.00 © 2013 IEEE

According to the China Internet Network Information Center (CNNIC),[7] the total number of Chinese microbloggers reached 274 million by mid-2012, comprising 51 percent of the total Internet population. Currently, the two leading Chinese microblog platforms are Sina Weibo and Tencent Weibo, each of which claims more than 300 million registered accounts. To comply with the government's content regulations, all Chinese weibo service providers must establish an internal censorship department to filter sensitive posts.

Although a few scholars have examined censorship practices in Chinese blogs[8,9] and microblogs,[10] their studies suffer from limitations. Here, we present a new approach to understanding the censorship characteristics of Chinese microblogs and evaluate how the country's new real-name registration (RnR) system is affecting microblogging, how often microbloggers post, and what they write about.

China Internet and Censorship

China's Internet market has rapidly expanded over the past decade. Although still a growing industry, the Internet growth rate has stabilized. Total Chinese Internet users reached 538 million by mid-2012, penetrating 40 percent of the country's population.[7] The country's Internet penetration rate rose only 4 percent in 2012 compared to a 9 percent average yearly growth since 2007. One significant shift in China's online usage pattern is the extensive adoption of social media applications. In mid-2012, one of every two Chinese "netizens" claimed to be a weibo user, whereas usage of traditional online tools — including email and posting on online forums — has diminished.[7]

The Chinese government has built "the largest and most sophisticated filtering systems in the world" to censor Internet content.[5] These stringent measures make the Chinese Internet virtually an intranet, in which sensitive terms are filtered, and access to Western sites such as Facebook, Twitter, and YouTube is blocked. However, these practices are only one aspect of the country's censorship system. Another major part of content censorship is implemented by domestic Internet content providers.[9] To comply with their license conditions, Chinese ISPs must act as censors to screen customers' messages or disable accounts. For example, Sina reportedly set up a censorship department of a thousand people to monitor weibos. In addition to filtering, social media companies also develop new practices to counter the dissemination of sensitive information. According to a media report, Sina encourages weibo users to report each other for releasing "untrue information," and the company is required to delete "harmful" weibos within five minutes.

Another rigorous measure — the RnR system — requires microbloggers to disclose their real identities. After trials in major Chinese cities, the system was officially launched on 16 March 2012. Registered users must release their identity numbers or indirect identifiers — such as mobile phone numbers — to the service provider for government verification. Users can then choose to post either by nickname or real name; unregistered users can view posts but can't create content. Widespread concern exists that this true identity disclosure policy will have a chilling effect on online comments, especially on political criticism and other sensitive topics.

Our Approach

We developed a data collection and visualization system called *Weiboscope*. To access the microblog data, we use the Sina Weibo Open API. We first constructed a list of popular microbloggers — those who have 1,000 or more followers — based on the User Search API. Systematically searching for these users in each region of China beginning in late 2010 and continuing throughout the study period gave us a list of about 350,000 microbloggers. We collected user posts through the User Timeline API function and then saved them in our PostgreSQL database. Our selection of high-follower-count samples can minimize the number of spam accounts in our data. Between 1 January 2012 and 30 June 2012, our system collected 111 million weibos.

Because the Chinese censor mechanism can remove posts swiftly, we must revisit the user timeline frequently enough to save as many copies of posts as possible before they're censored. However, because API use is subject to a per-hour rate limitation, we had to restrict the API calls to a list of prioritized users. We created three prioritized groups with various sampling frequencies:

- *Group 1* consisted of fewer than 10 China Media Project researchers at the University

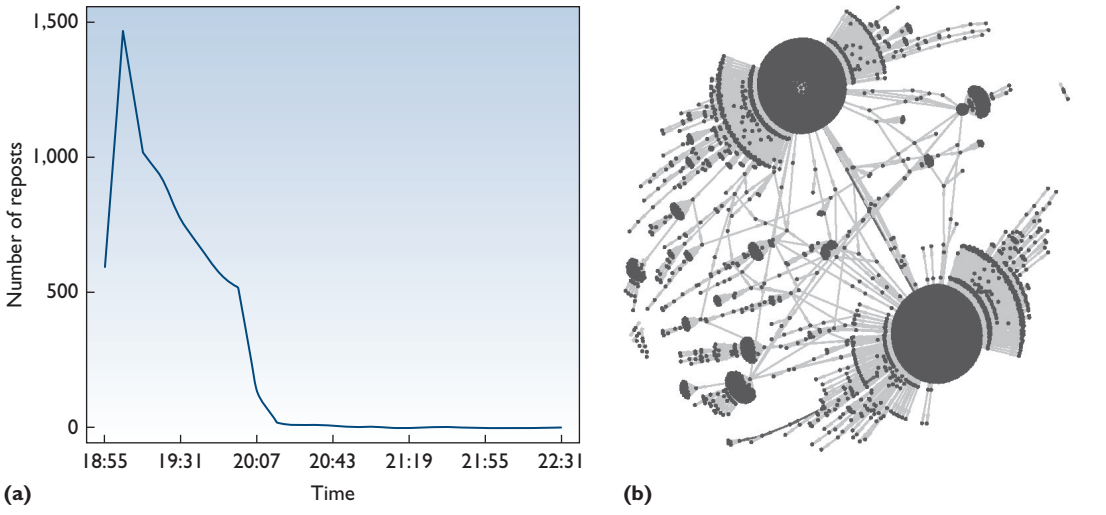

Figure 1. Example findings from the Weiboscope data collection and visualization system. (a) A graphic tracking Blogger H's original weibo post, diffusion, and censorship. (b) A repost graph shows that reposts were primarily diffused via the Internet space by a few microbloggers with many followers (node size is proportional to the microblogger's number of followers).

of Hong Kong who are scholars and active writers on the Chinese media industry. We checked and updated Group 1's timelines every three minutes.

- *Group 2* consisted of friends of Group 1 users; these are mostly Chinese dissident writers, journalists, and scholars. We updated this group's members automatically by adding users whose posts had been deleted. Group 2 consisted of about 5,000 users and was updated every six hours.
- *Group 3* consisted of users with an authenticated, or *VIP*, status and more than 10,000 followers. This group had approximately 38,000 users and was updated once a day.

After each user's timeline was fetched, the user's recently modified timeline was compared to the immediately previous version. If posts were missing in the new copy compared with the old copy, we found the missing posts and verified their absence by another API call before labeling them as missing. A weibo marked as missing comes with two potential messages from the API: "weibo does not exist" or "permission denied." Although Sina doesn't explicitly say what each message carried for the post they are marked with, we reverse engineered their empirical meaning. We tested the returned messages based on known causes of deletion and verified them via repeated trials. This approach yielded circumstantial evidence that posts missing for various reasons are marked differently. If the API response is "permission denied," the censor sets the post to be inaccessible to other users; we assume this because Sina doesn't allow a user setting to block outsider access. Posts that the user voluntarily deletes or the censor removes entirely are marked "weibo does not exist." Because no feasible method exists for determining who deleted a post, we excluded deleted posts to avoid contamination of the truly censored weibos.

An Example

To show how the system works, we'll use a real example. Blogger H is a famous scholar and writer in China. Followed by 181,000 readers, he writes regularly on social topics in China, and his microblog posts are — as expected — often censored. In our study, Blogger H belongs to Group 3, so we check his timeline every day.

On 2 May 2012, the blind Chinese activist Chen Guangcheng was confirmed to be seeking protection from the US Embassy in Beijing. Chinese authorities demanded an apology from US officials for the incident, and Blogger H posted a comment.

Figure 1a shows how this weibo was posted, diffused, and finally censored. The post was created by Blogger H at 18:46 on 2 May. Our system updated his timeline and saved the copy

of his post at 19:41. The next day, we detected "permission denied" from the API and confirmed the post's disappearance at 6:01 on 3 May. As Figure 1a shows, the hourly repost count was drastically reduced between 20:07 and 20:16 on 2 May, which appears to be the time that the original post was censored.

We then deployed Sina's Repost Timeline API to retrieve the whole retweeted set. The original weibo was reposted a total of 6,800 times by 6,193 distinct microbloggers in 1.5 hours. To study its diffusion, we generated the repost graph (Figure 1b) in which each node represents a microblogger, and an arrow denotes a retweet. Node size is proportional to the number of a microblogger's followers. We extracted the user mentions contained in the repost contents, in the format of "//@username," which showed the immediate source of repost, assuming that this default text isn't usually deleted by the user. As Figure 1b shows, the reposts were primarily diffused via the Internet space by a few microbloggers with many followers.

Analyzing Censored Keywords

To study the Chinese terms that lead to a higher likelihood of censorship, we included all "permission denied" posts captured from 1 January 2012 to 30 June 2012 in this part of the analysis.

To control for the influence of individual microbloggers' characteristics, we deployed a case-control matching strategy, which is a common observational study design used in epidemiology and social science, and is particularly useful when researching a topic for which implementing a randomized controlled experiment is infeasible. We paired each censored post with two randomly selected uncensored posts published by the same microblogger during the study period to construct a corpus of matching censored and uncensored microblog posts.

We used the Stanford Word Segmenter to undertake Chinese word segmentation for each document in the corpus.[11] We then removed stopwords, punctuation, emoticons, URLs, and "@usernames" (addressing a user) from the text. Using R[12] and its supported libraries, we tokenized each document in the corpus using the bigram tokenizer and generated a document-term matrix with inverse-document frequency weighting. We removed terms with low frequency (fewer than five occurrences) from the document-term matrix.

We used the χ^2 feature selection algorithm[13] to compare the relative frequency of each keyword occurrence in two distinct sets of posts. Keywords with higher χ^2 value represent a greater discriminatory ability than all others. We further deployed a relative risk (RR) measure to determine which terms can characterize censored posts – that is to say, $RR > 1$.

$$RR = \frac{\text{Frequency of occurrence in censored posts} + .5}{\text{Frequency of occurrence in uncensored posts} + .5}$$

We captured a total of 17,594 censored posts, submitted by 4,667 distinct microbloggers, during the study period. These posts were individually matched with 35,184 uncensored posts. Two posts weren't matched because one microblogger didn't have enough uncensored posts for one-to-two matching. Table 1 shows the top 30 keywords with the highest χ^2 value and $RR > 1$. (The full list is available online at http://tinyurl.com/hkuweibokeywords). These keywords were mostly related to the Bo Xilai scandal, the Chen Guangcheng diplomatic incident, the US Ambassador to China Gary Locke's finance disclosure, the one-child policy, housing policy, and the pension system. Other major keywords included political terms: "two meetings" (两会, the two annual meetings that make national-level political decisions), National People's Congress (人大代表), leaders in the Communist Party (书记), officials (官), refuting rumors (辟谣), content deletion (删), and profanity.

Furthermore, we discovered many terms that Chinese microbloggers created to circumvent the censors. Examples include

- "Pingxi Wang" (平西王), literally the "King who pacifies the west" in reference to Bo Xilai;
- "CGC" (the initials of Chen Guangcheng);
- "crown prince" (储君), referring to Xi Jinping, China's new leader; and
- "grass" (艹), an obscure alternative writing of 草, a homophone of a vulgar word.

We also uncovered "sensitive" terms that have lower discriminatory power, meaning they had higher survival rates for censorship circumvention. Examples here included "tomato" (西红柿) for "western red city," referring to

*Table 1. Top 30 keywords for censorship and potentially affected microblogger (PAM) status according to χ^2 value.**

	Predictors for Censorship			Predictors for PAMs		
Rank	Terms	χ^2	RR	Terms	χ^2	RR
1	重庆 (Chungking)	302.38	3.25	两会 (two meetings)	75.88	5.67
2	光诚 (Guangcheng)	248.62	32.82	雷 锋 (Lei Feng)	56.66	3.41
3	陈 光诚 (Chen Guangcheng)	237.42	46.72	白色 情人节 (White Day)	37.07	5.86
4	两会 (two meetings)	232.29	5.39	人大代表 (National People's Congress)	33.48	3.14
5	骆家辉 (Gary Locke)	212.82	12.38	情人节 (Valentine's Day)	30.66	1.42
6	辟谣 (refuting rumors)	203.14	5.16	王 立军 (Wang Lijun)	30.57	4.91
7	代表 (representative)	200.1	2.75	立军 (Lijun)	30.57	4.91
8	薄 (Bo, a family name)	187.26	5.07	锋 (Feng)	27.7	1.87
9	日报 (daily newspaper)	179.82	4.21	妇女节 (Women's Day)	24.97	3.61
10	公布 财产 (announced assets)	178.93	24.59	Bed (the English word)	22.26	2.01
11	北京 日报 (Beijing Daily)	172.05	7.91	民主 (democracy)	22.05	1.92
12	薄熙来 (Bo Xilai)	159.18	7.01	三八 (March 8)	22.04	2.56
13	人大代表 (National People's Congress)	152.5	4.23	Bed 凌乱 (messy)	21.96	3.74
14	骆家辉公布 (Gary Locke, announced)	152.2	305.96	学 雷 (learn, Lei)	21.68	3.36
15	财产 (assets)	150.95	3.86	吴英 (Wu Ying)	21.52	5.13
16	转发 (retweet)	146.52	1.34	叙利亚 (Syria)	20.65	5.46
17	转 (retweet)	144.09	1.81	方舟子 (Fang Zhozi)	19.92	2.35
18	王立军 (Wang Lijun)	134.24	4.18	吴 英案 (Wu Ying's case)	19.68	10.2
19	求证 (seeking confirmation)	127.34	4.56	英案 (Ying's case)	19.68	10.2
20	转发 微博 (retweet, weibo)	126.49	1.34	315	19.51	3.11
21	请 骆家辉 (asking, Gary Locke)	126.14	253.97	市长 (mayor)	19.33	2.09
22	书记 (secretary)	124.51	2.87	的意料 (expectation of)	19.24	2.95
23	转 么 (retweet)	122.13	245.97	代表委员 (representative member)	19.04	6.56
24	鬼子 转 (little devil, retweet)	122.13	245.97	出乎 的 (exceed someone)	19.03	2.81
25	公布 (announced)	120.01	2.86	文革 (Cultural Revolution)	18.48	2.9
26	求 辟谣 (seeking refutation)	110.06	7.33	三八节 (March 8 Festival)	18.29	3.7
27	删 (deleted)	108.53	2.97	公知 (public intellectual)	18.29	3.7
28	陈 (Chen, a family name)	106.82	2.2	两会 期间 (two meetings, duration)	18.05	37
29	微博 (Weibo)	104.57	1.26	党 (party)	17.97	1.63
30	养老不 (social security for the elderly and "don't")	100.08	201.97	临时工 (temporary job)	17.95	4.29

*Keywords in red are in both lists.

Chongqing; and "head nurse" (护士长), referring to Wang Lijun, a key figure in the Bo Xilai scandal.

Evaluating the Impact of Real-Name Registration

To evaluate RnR's impact, we defined

- T1 as the period between 8 December 2011 and 15 March 2012 — that is, the 99 days before the enforcement of the RnR scheme in China; and
- T2 as the period between 16 March 2012 and 22 June 2012 — that is, the 99 days after RnR enforcement.

In all, we included 166,725 microbloggers who posted at least one weibo during T1. We calculated the total number of posts submitted in T1 and T2 for each sample. Those who made no posts during T2 were defined as potentially affected microbloggers (PAMs). We hypothesized that their reduced user activity was largely attributable to RnR's chilling effect.

First, we conducted aggregate-level analysis. Figure 2 plots the daily frequency of posts and features three sharp drops: one before the Lunar New Year, another immediately following 16 March, and the third on 20 April (linked to an Internet crackdown). The volume of posts appeared to decrease gradually after 16 March for roughly 30 days and then bounced up. The pattern might be linked to RnR enforcement. But, on 15 March, Bo Xilai was ousted, which would stimulate more censorship. Nevertheless, we undertook an autoregressive integrated moving average (ARIMA) time-series analysis and found the drop between 16 March and 21 April was statistically insignificant.

We reexamined the question using logistic regression for individual-level analysis. Finding predictors for PAM — including user characteristics and T1 post keywords — can shed light on the mechanism for the reduced activity. The dependent variable was the PAM status. The independent variables included number of followers, number of friends, VIP status, whether the user let others comment, self-reported gender, and self-reported geographic location. Microbloggers who reportedly originated from Taiwan, Hong Kong, Macao, or "other" were categorized as being from non-RnR regions. Independent variables with an adjusted odds ratio (AOR) significantly larger or smaller than 1 at a 5-percent level were declared significant independent predictors of PAM status.

We identified 57,155 PAMs (34.3 percent of 166,725) who didn't post during T2. The statistically significant PAM predictors — sorted by effect size, in descending order — follow here; "CI" stands for "confidence interval."

- Claiming to have originated from RnR regions — AOR: 2.124; 95 percent CI: 2.051 to 2.201.
- Non-VIP — AOR: 1.725; 95 percent CI: 1.772 to 1.820.
- Allowing anyone to comment on timeline — AOR: 1.730; 95 percent CI: 1.645 to 1.819.
- Female — AOR: 1.051; 95 percent CI: 1.027 to 1.075.
- Every reduction of 100 friends — AOR: 1.010; 95 percent CI: 1.008 to 1.012.
- Every reduction of 100 followers — AOR: 1.002; 95 percent CI: 1.001 to 1.003.

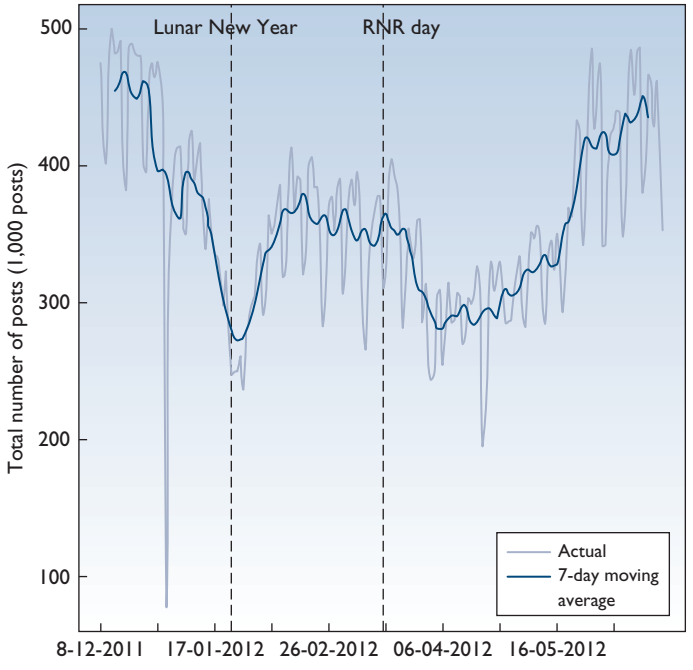

Figure 2. The daily posting frequency of 166,725 microbloggers indicates a gradual decrease in number of posts after 16 March for roughly 30 days and then a rebound. However, a time series analysis reveals that the rise and drop were statistically insignificant.

We then identified the Chinese terms in the microblogs created during T1 that were associated with PAM status. To adjust for the confounders of PAM status — such as personal characteristics or personal reasons — we used a case-control matching strategy. We paired each PAM with a matching non-PAM, using the same status in terms of gender, province, VIP, and allowing microbloggers' comments. Two additional criteria for matching were set to select those whose followers count and friends count are within plus or minus 30 percent ranges and whose frequency of microblog posts counted in T1 fell within the 50 to 200 percent range. If multiple matched non-PAMs were identified, we selected the one with the maximum ratio of sustained user activity — that is, a post count in T2 divided by the post count in T1. We discarded PAMs without matches. We eventually assembled a group with an equal number of PAMs and non-PAMs with similar group-wise characteristics.

We randomly selected 3,000 PAMs from all microbloggers who claimed to originate from Beijing, Shanghai, Guangdong, and Tianjin, which were the officially announced trial cities

for enforcing the RnR. We then generated 3,000 matched non-PAMs. We retrieved all of their weibos posted in T1 (totaling 437,153 posts) to analyze terms associated with PAM status. We processed the text of the corpus as previously described. Table 1 shows a list of keywords with high discriminatory power for PAM status — that is, the top 30 terms as ranked by χ^2 value and $RR > 1$.

Among all top keywords with $RR > 1$, most connotatively referred to political scandals, international affairs, social events, or figures. Examples include "two meetings," Lei Feng (雷锋), Wang Lijun (王立军), Wu Ying (吴英, a businesswoman convicted of financial fraud and sentenced to death), Syria (叙利亚), Cultural Revolution (文革), Wukan (乌坎, a village where an anti-corruption protest took place), Article 73 (73 条, new legislation allowing authorities to detain any parties suspected of national security threats), and corruption (腐败).

So, despite an absence of evidence of significant changes in overall activity following RnR, when we compared microbloggers who didn't post after the RnR to those who posted as usual by scrutinizing the contents they published before the RnR, we discovered that the best discriminatory terms between two groups were mostly related to political and social issues. We therefore suspected that, if the RnR weren't enforced, some PAMs would have posted weibos related to politics as usual after the RnR. Although the RnR's impact can't be observed ecologically, it might exert a chilling effect selectively on some microbloggers who are characterized as

- stationed inside mainland China, with a non-authenticated identity;
- welcoming other microbloggers to comment (stimulating public debate); and
- having a smaller online social network (likely grassroots citizens with fewer followers and friends).

RnR might cause these microbloggers to stop writing on sensitive issues using the same account, though we can't completely rule out their changing accounts and continuing to post. These findings should be considered preliminary; further study over a longer term is needed before we can draw conclusions, especially given that the Chinese government has recently announced additional requirements for the RnR.[6]

Readers should cautiously interpret our findings. Our samples were collected from 350,000 high-follower-count microbloggers, which constitute a small percentage of the overall microblogger population. Despite their ability to diffuse information, our findings can't be overly generalized to all microbloggers. Nevertheless, we don't expect a significant change in results when sampling users with fewer followers. Another study found that a small number of high-follower-count Chinese microbloggers contribute a majority of the posts and draw the most attention.[14] Thus high-follower-count users should be deemed representative.

Our system also can't detect the complete set of deleted posts, especially the rapidly censored ones. Moreover, the abandonment of microblogging might be attributable to factors other than the RnR chilling effect. Given that no study exists on the reasons for abandoning microblogging under censorship, we suggest future studies to interview users about their practices before and after the policy change.

Despite the fact that China is an authoritarian state and the contexts are totally different, our results have global implications: recently, some international social media companies — including YouTube and Google — have expressed interest in forcing their users to disclose real names to minimize trolling.

The social media industry in China is dynamic, and new technologies emerge every few years. WeChat, for example, is a popular mobile messaging application that has become a new superstar of social media in China; it currently claims to have more than 300-million accounts. Future work is crucially needed to study new social media applications in China, investigating their development and social impact as well as the Chinese government's evolving censorship policy.

Acknowledgments

This study is funded by the University of Hong Kong Seed Funding Program for Basic Research. We also acknowledge Cedric Sam's contribution in developing the programs for data collection.

References

1. J. Hassid, "Safety Valve or Pressure Cooker? Blogs in Chinese Political Life," *J. Comm.*, vol. 62, no. 2, 2012, pp. 212–230.

Related Work in Chinese Censorship and Censorship Policy

Researchers have conducted a few studies on the Chinese censorship policy in relation to social media.

Censorship Practices in China

In 2011, Gary King, Jennifer Pan, and Margaret Roberts deployed an undisclosed automatic data collection methodology to download 3,674,698 forum and blog posts from 1,382 Chinese websites and concluded that roughly 13 percent of all posts in China were censored.[1] In their findings, they observed that Chinese authorities can tolerate posts that write on a wide range of criticism of the Chinese government and its policies, but tend to be more sensitive to censoring the spread of posts that might lead to collective action. However, their study didn't cover weibos.

Also in 2011, David Bamman, Brendan O'Connor, and Noah A. Smith gathered 56 million posts using Sina's public timeline API. They then checked a random sample of 1.3 million weibos, of which 212,583 were deleted, suggesting an overall message deletion rate of 16.25 percent.[2] Using Chinese posts on Twitter as a comparison group, the authors constructed a list of politically sensitive keywords that had a higher probability of being censored.

However, a few limitations in these studies could threaten the findings' validity. The first issue is the omnipresence of spam and fake accounts in the Chinese microblog space. Sina Weibo carries a huge amount of spam, which artificially inflates Internet traffic.[3] Even when researchers deployed spam removal procedures, the most frequent censored terms were still dominated by spam.[2]

The second issue is the influence of individual user characteristics that might confound the results. Many Chinese Internet accounts are selectively monitored[4] – reportedly, this is done manually by weibo providers' "editors." Censorship decisions can target microbloggers who have specific attributes; in the Chinese context, the typical targets are dissidents, journalists, scholars, and rights-activists. Therefore, simply comparing terms between censored and uncensored sets doesn't necessarily disentangle the influence from individual characteristics. So, rather than address the question of what topics or terms in general are more likely to be censored, we reframe the question: When posts are made by the same person, why are some posts censored but others not censored? What are the terms that characterize the censorship decision?

The third limitation here is the definition of Chinese terms. Chinese text doesn't use space to separate terms. Rather than using a tool to chop Chinese documents into terms for analysis, both studies sampled the data corpus by searching predefined sets of dictionary terms, – sourced, for example, from an open source Chinese-English dictionary and Chinese Wikipedia[2] – or self-devised lists of sensitive Chinese terms.[1] However, in the context of China, sensitive topics or names are commonly represented by puns, homophones, or special combinations of written Chinese characters that don't appear in dictionaries. Research studies that impose a predetermined list of Chinese terms might fail to correctly identify the user-generated terms and thus limit the scope of their findings.

True-Identity Policies

Social science research has long shown that people consider conformity pressures or possible punishments when voicing their opinions.[5] Since real-name registration has been in force, microbloggers must register their true identities and thus no longer have the anonymity that they once enjoyed. This could discourage them from making critical posts due to fear of punishments – including arrest and imprisonment – which aren't uncommon in China.

South Korea was the first country to enact regulations requiring the use of real identities on the Internet.[6] An evaluation study aimed at assessing that policy's impact confirmed a significant reduction in uninhibited behaviors at the aggregate level. In that study, the primary impact was seen among light users, while the behavior of heavy users seemed unchanged.[6]

References

1. G. King, J. Pan, and M. Roberts, "How Censorship in China Allows Government Criticism but Silences Collective Expression," *Am. Political Science Rev.*, forthcoming, 2013.
2. D. Bamman, B. O'Connor, and N. Smith, "Censorship and Deletion Practices in Chinese Social Media," *First Monday*, vol. 17, nos. 3–5, 2012; http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3943.
3. L. Yu, S. Asur, and B.A. Huberman, "Artificial Inflation: The True Story of Trends in Sina Weibo," forthcoming; preprint available at http://arxiv.org/pdf/1202.0327v1.
4. G. Yang, *The Power of the Internet in China: Citizen Activism Online*, Columbia Univ. Press, 2009.
5. T. Connolly, L.M. Jessup, and J.S. Valacich, "Effects of Anonymity and Evaluative Tone on Idea Generation in Computer-Mediated Groups," *Management Science*, vol. 36, no. 6, 1990, pp. 689–703.
6. D. Cho, S. Kim, and A. Acquisti, "Empirical Analysis of Online Anonymity and User Behaviors: The Impact of Real Name Policy," *Proc. Int'l Conf. System Sciences*, IEEE, 2012, pp. 3041–3050.

2. G. Qian and D. Bandurski, "China's Emerging Public Sphere," *Changing Media, Changing China*, S.L. Shirk, ed., Oxford Univ. Press, 2011, pp. 38–76.
3. G. Yang, *The Power of the Internet in China: Citizen Activism Online*, Columbia Univ. Press, 2009.
4. K.B. Richburg, "In China, Microblogging Sites Become Free-Speech Platform," *Washington Post*, 27 Mar. 2011; http://articles.washingtonpost.com/2011-03-27/world/35260490_1_weibo-free-speech-posts.

5. OpenNet Initiative, *Internet Filtering in China 2004–2005: A Country Study*, 9 Aug. 2012; http://opennet.net/research/profiles/china.
6. K. Bradsher, "China Toughens Its Restrictions on Use of the Internet," *New York Times*, 28 Dec. 2012; www.nytimes.com/2012/12/29/world/asia/china-toughens-restrictions-on-internet-use.html.
7. *30th Statistical Report on Internet Development in China*, China Internet Network Information Center, Beijing, 2012.
8. G. King, J. Pan, and M. Roberts, "How Censorship in China Allows Government Criticism but Silences Collective Expression," *Am. Political Science Rev.*, forthcoming, 2013.
9. R. MacKinnon, "China's Censorship 2.0: How Companies Censor Bloggers," *First Monday*, vol. 14, no. 2, 2 Feb. 2009; www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2378/2089.
10. D. Bamman, B. O'Connor, and N. Smith, "Censorship and Deletion Practices in Chinese Social Media," *First Monday*, vol. 17, nos. 3–5, 2012; http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3943.
11. H. Tseng et al., "A Conditional Random Field Word Segmenter," *Proc. 4th SIGHAN Workshop on Chinese Language Processing*, Assoc. for Computational Linguistics, 2005; http://research.microsoft.com/pubs/78905/tseng05crf.pdf.
12. R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, 2011; www.r-project.org.
13. Y. Yang and J.O. Pedersen, "A Comparative Study on Feature Selection in Text Categorization," *Proc. 14th Int'l Conf. Machine Learning*, Morgan Kauffman, 1997, pp. 412–420.
14. K.W. Fu and M. Chau, "Reality Check for the Chinese Microblog Space: A Random Sampling Approach," *PLoS One;* http://dx.plos.org/10.1371/journal.pone.0058356.

King-wa Fu is an assistant professor at the Journalism and Media Studies Center, University of Hong Kong. His research interests include political participation and online media use, health communication, and computational and statistical methods in media studies. Fu has a PhD in journalism and media studies from the University of Hong Kong. He's a member of IEEE. Contact him at kwfu@hku.hk.

Chung-hong Chan is a research assistant at the Journalism and Media Studies Center, University of Hong Kong. He's interested in the application of statistical methods to detect content censorship in social media. Chan has an MSc in epidemiology and biostatistics from the Chinese University of Hong Kong. Contact him at chainsawtiney@gmail.com.

Michael Chau is an associate professor in the School of Business, University of Hong Kong. His research interests include Web mining, business intelligence and analytics, knowledge management, and IT education. Chau has a PhD in management information systems from the University of Arizona. He's a senior member of IEEE, ACM, and the Association for Information Systems. Contact him at mchau@business.hku.hk.