Assessing Censorship on Microblogs in China: Discriminatory Keyword Analysis and Impact Evaluation of the "Real Name Registration" Policy

KW Fu[1], CH Chan[2], M Chau[3]

[1] Journalism and Media Studies Centre, The University of Hong Kong, Hong Kong, China; Tel: (852) 3917-1643; Fax: (852) 2858-8736; Email: kwfu@hku.hk

[2] Journalism and Media Studies Centre, The University of Hong Kong, Hong Kong, China; Tel: (852) 3917-4114; Fax: (852) 2858-8736; Email: chainsawtiney@gmail.com

[3] School of Business, The University of Hong Kong, Hong Kong, China; (852) 3917-1014; (852) 2858-5614; Email: mchau@business.hku.hk

Citation:

King-wa Fu, Ch Chan, Michael Chau, "Assessing Censorship on Microblogs in China: Discriminatory Keyword Analysis and Impact Evaluation of the "Real Name Registration" Policy," IEEE Internet Computing, 11 Feb. 2013. IEEE computer Society Digital Library. IEEE Computer Society, <http://doi.ieeecomputersociety.org/10.1109/MIC.2013.28>

Abstract

Using 111 million microblogs collected between January 1 and June 30, 2012, we report our findings on investigating the use of microblogs in China, or *weibo* in Chinese, and the impact of censorship practices imposed on Chinese microbloggers. To better control for alternative explanations for censorship decisions that are attributable to an individual's characteristics and choices, we deployed a matched case-control study design to 1) determine a list of Chinese terms that discriminate the censored and uncensored posts when they are written by the same microbloggers, with the list including some homophones and puns created by Chinese microbloggers to circumvent the censors successfully, and 2) evaluate the impact of the Real Name Registration system on microbloggers' posting activities, and the findings suggest that the new policy might have stopped some microbloggers from writing about social and political subjects.

Keywords: Internet censorship; China; real name registration; microblog; Weibo

Introduction

China is known to be a society where news media are strictly regulated by the government and the public information is often censored. The country is routinely ranked by international bodies as a place where its citizens enjoy the least extent of freedom of speech and freedom of the press. But on the Internet platforms, for example online forums or blogs, Chinese people seem to have more autonomy to speak on public affairs and occasionally can set a social agenda that successfully draws the attention of traditional media and the public,[1] for example "My father is Li Gang" or "Guo Meimei," incidents which 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 to build the public agenda.[2, 3] Recently, the Twitter equivalent microblogging service in China, *weibo* in Chinese, has been enthusiastically depicted by Western media as a new "free speech platform."[4] This optimistic view is, however, challenged with the Chinese authorities' ubiquitous mechanisms to control the flow of public information, including the sophisticated filtering system known as the Great Firewall[5] and a recent new requirement that microbloggers must register with their real identities.[6]

According to the China Internet Network Information Center (CNNIC),[7] the total number of Chinese microbloggers reached 274 million by mid-2012, comprising 51% of the total Internet population. Currently, the two leading Chinese microblog platforms are Sina Weibo and Tencent Weibo, each claiming to possess over 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.

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

There have been a few scholarly examinations of the censorship practices in Chinese blogs[8, 9] and microblogs.[10] But as we will discuss in coming sections, these studies suffer from some limitations. This article presents a new approach to understanding the characteristics of censorship of Chinese microblogs. Moreover, we evaluate the impact of the enforcement of the Real Name Registration system in China on the microblogging, how often the microbloggers post, and what they write about.

China Internet and Censorship

The Internet market in China has rapidly expanded since a decade ago. It is still a growing industry but the growth rate has stabilized. Total Chinese Internet users reached 538 million by mid-2012, penetrating 40% of the country's population.[7] The country's Internet penetration rate rose only 4 percent in 2012 compared to 9 percent average yearly growth since 2007. One significant shift in the pattern of online usage in China 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] The stringent measures make the Chinese Internet virtually an Intranet, in which sensitive terms are filtered and access to Western sites like Facebook, Twitter, and Youtube are blocked. But these practices are only part of the whole censorship system. A major process of censoring content is implemented by domestic Internet content providers.[9] To comply with their license conditions, Chinese service providers have to act as a censor to screen customer's messages or disable accounts. For example, Sina Corp. reportedly set up a censorship

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

department of a thousand people to monitor weibos. Besides filtering, social media companies also develop new practices to counter the dissemination of sensitive information on social media. According to media report, Sina Corp. encourages weibo users to report each other for releasing "untrue information" and is required to delete "harmful" weibos within five minutes.

Another rigorous measure involves new requirements for microbloggers to disclose real identities. This is known as the Real-Name Registration system. After trials in some major Chinese cities, the system was officially begun on March 16, 2012. Registered users release their identity numbers or indirect identifiers, say mobile phone numbers, to the service provider for verification by the government. The users then can choose to post either by nickname or real name, whereas unregistered users can view posts but are not allowed to create contents. There has been widespread concern that the true identity disclosure policy would have created a chilling effect on online comments, especially on political criticism and sensitive topics.

Related Works

Censorship practices in China

King, Pan, & Roberts deployed an undisclosed automatic data collection methodology to download 3,674,698 blog posts from 1,382 Chinese websites in 2011, and concluded that roughly 13% of all blog posts in China were censored.[8] Their findings made an interesting observation that the Chinese authorities can tolerate the blogs posting a wide range of criticism of the Chinese government and policies but tend to censor primarily the spread of posts that may lead to collective action.

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

Bamman, O'Connor, & Smith gathered 56 million posts using Sina's public timeline API in 2011. 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%.[10] Using Chinese posts on Twitter as comparison group, the authors constructed a list of politically sensitive keywords that had a higher probability of being censored.

Despite the efforts of the two research teams, a few limitations would pose a threat to the validity of the findings. The first issue is the omnipresence of spam and fake accounts on the Chinese microblog space. Sina Weibo has been discovered to carry huge amount of spam, which artificially inflates Internet traffic.[11] Even when spam removal procedures were deployed, the most frequent censored terms were still dominated by spam.[10]

The second issue is the influence of individual user characteristics that may possibly confound the results. Many Chinese Internet accounts are selectively monitored,[3] reportedly done manually by weibo providers' "editors." Censorship decisions can be made to target microbloggers who have specific attributes, that is to say in the Chinese context: dissidents, journalists, scholars, or rights-activists. Therefore, simply comparing terms between censored and uncensored sets may not disentangle the influence from individual characteristics. So instead of addressing the question of what topics or terms in general are more likely to be censored, we reframe the question into: when posts are made by same person, why are some posts censored but others not censored? What are the terms that characterize the censorship decision?

The third limitation is the definition of Chinese terms. Chinese text does not use space to separate terms. Rather than using tool to chop Chinese documents into terms for

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

analysis, the approach used in both studies was to sample the data corpus by searching pre-defined sets of dictionary terms, for example sourcing from open source Chinese-English dictionary and Chinese Wikipedia,[10] or self-devised lists of sensitive Chinese terms.[8] However, in the context of China, sensitive topics or names are commonly represented by puns, homophones, or special combinations of written Chinese characters which often are not usages in a dictionary. If a research study imposes a predetermined list of Chinese terms, it may fail to correctly identify the user-generated-terms and thus limit the scope of findings.

Effects of the Real-Name Registration system

Social science research has shown for long that people take into consideration conformity pressure or possible punishments when voicing their opinions.[12] Since Real-Name Registration has been in force, microbloggers have to 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 such as arrest and imprisonment, which are not uncommon in China.

South Korea was the first country where regulations requiring use of real identities on the Internet were enacted.[13] An evaluation study aiming to assess the impact of that policy confirmed a significant reduction of uninhibited behaviors at the aggregate level with light users as the primary impacted group; heavy users did not seem to have changed behaviors.[13]

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

Our Approach

We developed a data collection and visualization system called Weiboscope. We make use of the Sina Weibo Open API to access the microblog data. Firstly, a list of popular microbloggers was constructed based on the User Search API. The inclusive criterion of our sample is those microbloggers who have 1,000 or more followers. We systematically searched these users in each region in China since late 2010. Through this process, we generated 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 January 1, 2012, and June 30, 2012, our system collected 111 million weibos.

As the Chinese censor can take down posts swiftly, we have to revisit the user timeline frequently enough to save as many copies as possible of posts before they are censored. However, since the API use is subject to a per-hour rate limitation, we restrict the API calls to a list of prioritized users on our list.

We created three prioritized groups with various sampling frequencies: Group 1, consisting of less than 10 China Media Project researchers at the University of Hong Kong, who are scholars and active writers on the Chinese media industry. We check and update Group 1's timelines every three minutes; Group 2, whose members were created by obtaining friends of Group 1 users. They are mostly Chinese dissident writers, journalists, and scholars. We updated the members of this group automatically by adding users whose posts have been discovered to be deleted. This list consists of about 5,000 users. An update for Group 2 is run every six hours; and Group 3, including those users

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

with an authenticated status (VIP status, hereafter) and more than 10,000 followers, has about 38,000 users. For Group 3, each update is run once a day.

While each user's timeline was being fetched, the user's recently modified timeline was compared to the immediately previous version. If there were posts 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 that is marked as missing comes with two potential messages from the API: "weibo does not exist" and "permission denied." Although Sina does not explicitly say what each message carried for the post they are marked with, we reverse engineered their empirical meaning. We tested the returned messages by known causes of deletion and verified by repeated trials. This approach yielded circumstantial evidence that posts missing for various reasons are marked differently. If the API response is "permission denied," the post is set by the censor to be inaccessible by other users, under the condition that Sina does not allow user setting to block outsider's access. On the other hand, posts that are voluntarily deleted by the user or removed entirely by the censor are marked "weibo does not exist." As there is no feasible method to determine who deleted a post, deleted posts were therefore excluded to avoid contamination of the truly censored weibos.

An Example

We illustrate how the system works in 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 expectedly often censored. He belongs to the Group 3, whose timeline was checked every day.

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

On May 2, 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 American 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 May 2. Our system updated his timeline and saved the copy of his post on 19:41. Next day, we detected "permission denied" from the API and confirmed its disappearance on 6:01 May 3. Figure 1a shows that the hourly repost count was drastically reduced between 20:07 and 20:16 May 2, which appeared to be the time when the original post was censored.

We then deployed Sina's Repost Timeline API to retrieve the whole set of retweets. The original weibo was totally reposted 6,800 times by 6,193 distinct microbloggers in one and half hours. To study its diffusion, we generated the repost graph in which each node represents a microblogger and an arrow denotes a retweet from one and another. Node size is proportional to the number of followers of a microblogger. 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 is not usually deleted by the user. Figure 1b presents the repost graph, which demonstrates that 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 higher likelihood of being censored, all "permission denied" posts captured from January 1, 2012, to June 30, 2012, were included in this part of the analysis.

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

To control for the influence of the characteristics of individual microbloggers, we deployed a case-control matching strategy, which is a common design used in observational study in epidemiology or social science, and is particularly useful when researching a topic when implementation of a randomized control experiment is not feasible. Each censored post was paired up 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.

The Stanford Word Segmenter was deployed to undertake Chinese word segmentation for each document in the corpus.[14] Subsequently, stopwords, punctuations, emoticons, URLs and "@usernames" (addressing a user) were removed from the text. Using R[15] and its supported libraries, each document in the corpus was then tokenized using the bigram tokenizer and a document-term matrix was generated with inverse-document frequency weighting. Terms with low frequency, less than five occurrences, were removed from the document-term matrix.

The relative frequency of each keyword occurrence in two distinct sets of posts was compared by χ^2 feature selection algorithm.[16] Keywords with higher χ^2 value represent a greater discriminatory ability among all others. We further deployed a relative risk (RR) measure to determine which terms are able to characterize censored posts, that is to say RR>1.

$$RR = \frac{\text{Frequency of occurrence in censored posts} + 0.5}{\text{Frequency of occurrence in uncensored posts} + 0.5}$$

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

Totally 17,594 censored posts, which were submitted by 4,667 distinct microbloggers, were captured during the study period. These posts were individually matched with 35,184 uncensored posts. Two posts were not matched because one microblogger did not have enough uncensored posts for one-to-two matching. The top 30 keywords with the highest χ^2 value and RR > 1 are listed in Table 1a. (The full list is available online or upon request). These keywords were mostly related to the Bo Xilai scandal, Chen Guangcheng diplomatic incident, United States 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 a number of terms that were created by Chinese microbloggers to circumvent the censors. For example Pingxi Wang (平西王, literally "King who pacifies the west," referring to Bo Xilai), CGC (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).

Going down the list, we uncovered a number of "sensitive" terms which have lower discriminatory power, representing that they had higher survival rates for censorship circumvention. They included "tomato" (西红柿, "western red city" which referred to Chongqing) and "head nurse" (护士长, referring to Wang Lijun, a key figure in the Bo Xilai scandal).

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

Evaluating the Impact of Real-Name Registration

We defined the period between December 08, 2011, and March 15, 2012, as T1 and the one between March 16, 2012, and June 22, 2012, as T2. T1 covers the 99 days before the enforcement of the Real-Name Registration (hereafter, RnR in short) scheme in China while T2 is the same-length period after the enforcement. Totally 166,725 microbloggers who posted at least one weibo during T1 were included.

The total number of posts submitted in T1 and T2 were calculated for each sample. Those who made no posts during T2 were defined as potentially affected microbloggers (hereafter, PAMs in short). We hypothesized that their reduced user activity was largely attributed to the chilling effect of RnR.

We first conducted aggregate-level analysis. Daily frequency of posts is plotted in Figure 1c, in which we noticed three sharp drops. One was before the Lunar New Year, another immediately followed March 16, and the last was on April 20 (linked to an Internet crackdown). The volume of posts appeared to decrease gradually after March 16 for roughly 30 days and then bounced up. The pattern may be linked to the enforcement of the RnR. But on March 15, Bo Xilai was coincidentally ousted, which would stimulate more censorship. Nevertheless, we undertook ARIMA time-series analysis and found the drop between March 16 to April 21 was indeed not statistically significant.

We reexamined the question using individual level analysis, employing logistic regression. Finding predictors for PAM, including user characteristics and keywords of posts from T1, 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 or not the user allowed others to comment, self-

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

reported gender, and self-reported geographic location. Microbloggers who reported as originating from "Taiwan," "Hong Kong," "Macao" and "others" were categorized as from non-RnR regions. Independent variables with adjusted odds ratio (AOR in short hereafter) significantly larger or smaller than 1 at 5% level were declared as significant independent predictors for PAM status.

We identified 57,155 (34.3% of 166,725) PAM who did not post during T2. The statistically significant predictors are listed and sorted by effect size in descending order: claiming to have originated from RnR regions (AOR: 2.124, 95% CI: 2.051 to 2.201), non-VIP (1.725, 95% CI: 1.772 to 1.820), allowing all to comment on timeline (AOR: 1.730, 95%CI: 1.645 to 1.819), female (AOR: 1.051, 95% CI: 1.027 to 1.075), every reduction of 100 friends (1.010, 95% CI: 1.008 to 1.012), every reduction of 100 followers (AOR: 1.002, 95% CI: 1.001 to 1.003).

We then identified the Chinese terms in the microblogs created during T1 that were associated with PAM status. In order to adjust for the confounders for the PAM status, such as personal characteristics or reasons, a case-control matching strategy was used. Each PAM was paired with a matching non-PAM, with the same status in terms of gender, province, VIP, and allowing microbloggers' comments. Two additional criteria for matching are set to select those whose followers count and friends count are within +/- 30% ranges and whose frequency of microblog posts counted in T1 fell within the 50% to 200% range. If multiple matched non-PAM were identified, the one with maximum ratio of sustained user activity, i.e. post count in T2 over post count in T1, was selected. PAM without matching were discarded. We eventually assembled a group with equal number of PAM and non-PAM with similar group-wise characteristics.

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

We randomly selected 3,000 PAM from all microbloggers who claimed to originate from Beijing, Shanghai, Guangdong, and Tianjin, which were the officially announced trial cities to enforce the RnR. We then generated 3,000 matched non-PAM. All their weibos posted in T1, totaling 437,153 posts, were retrieved for analyzing terms associated with PAM status. The text of the corpus was processed as previously described. A list of keywords with high discriminatory power for PAM status, which is the top 30 terms as ranked by χ^2 value and RR > 1, was presented at Table 1b.

Among all top keywords with RR > 1, most were found to be connotatively referred to political scandals, international affairs, social events or figures, for instance "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), or Article 73 (73 条, new legislation allowing authorities to detain any parties suspected of national security threats) and corruption (腐败).

In summary, despite an absence of evidence of significant changes in overall activity after the RnR, when comparing those microbloggers who did not post after the RnR to those who posted as usual by scrutinizing their contents 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 if the RnR were not enforced, some PAM microbloggers could have posted weibos related to politics as usual after the RnR. While the impact of the RnR can not be observed ecologically, it might exert a chilling effect selectively on *some* microbloggers who are characterized as: stationed inside

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

Mainland China, with a non-authenticated identity, welcoming other microbloggers to comment (stimulating public debate), and less online networked (likely grassroots citizens with fewer followers and friends). It might cause them to stop writing on sensitive issues via the same account, whereas we can not completely rule out that they would change accounts and continue to post. The finding should be considered as preliminary. Further study in a longer term is needed before drawing conclusion, especially when China government announced additional requirements for the RnR.[6]

Readers should cautiously interpret the findings. Our samples were collected from 350,000 high-followers-count microbloggers, which constitute a small percentage of the overall microblogger population. Despite their ability to diffuse information, the findings can not be overly generalized to all microbloggers. Nevertheless, we do not expect a significant change in results if users with fewer followers are sampled. Another study has found a small number of high-follower-count Chinese microbloggers contributes a majority portion of posts and draws most attention.[17] Thus high-follower-count users should be deemed representative. Our system also can not detect the complete set of deleted posts, especially those rapidly censored one. Moreover, the abandonment of microblogging might be attributable to factors other than chilling effect. There is no previous study on the reasons for abandonment of microblogging under censorship. We suggest future studies should interview users about their practices before and after the policy change.

Conclusion

We report our method for researching Internet censorship in China by deploying computational approaches as well as matched case-control designs, demonstrating a

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

novel method to investigate online censorship in China. We strive to address several limitations of the previous studies, especially controlling for explanations for a censorship decision according to individual attributes. We successfully identify a list of Chinese terms that discriminate between censored and uncensored posts, including some homophones and puns that were created by Chinese microbloggers to circumvent the censors successfully. Moreover, in the impact evaluation study, we find evidence to support that Real-Name Registration may stop some specific microbloggers from writing about social and political subjects. Despite the fact that China is an authoritarian state and the contexts are totally different, this result should have global implications as recently more international social media companies, including Youtube and Google, have expressed interest in forcing their users to disclose real names to minimize trolling.

Acknowledgements

This study is funded by the University of Hong Kong, Seed Funding Programme for Basic Research. The authors wish to acknowledge the contribution made by Cedric Sam in developing the programs for data collection.

References

[1] J. Hassid, “Safety valve or pressure cooker? Blogs in Chinese political life,” *Journal of Communication,* vol. 62, no. 2, pp. 212-230, April 2012, 2012.
[2] G. Qian, and D. Bandurski, "China's emerging public sphere," *Changing media, changing China*, S. L. Shirk, ed., pp. 38-76, New York: Oxford University Press, 2011.
[3] G. Yang, *The power of the Internet in China: citizen activism online*, New York: Columbia University Press, 2009.
[4] K. B. Richburg, "In China, microblogging sites become free-speech platform," *The Washington Post*, 2011.
[5] OpenNet Initiative, "China," http://opennet.net/research/profiles/china, [28 August 2012, 2009].

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

[6] K. Bradsher, "China Toughens Its Restrictions on Use of the Internet," *The New York Times*, http://www.nytimes.com/2012/12/29/world/asia/china-toughens-restrictions-on-internet-use.html, [December 28, 2012].

[7] China Internet Network Information Center, *30th statistical report on Internet development in China*, China Internet Network Information Center, Beijing, China, 2012.

[8] G. King, J. Pan, and M. Roberts, "How censorship in China allows government criticism but silences collective expression. Working Paper, 2012. copy at http://j.mp/LdVXqN," 2012.

[9] R. MacKinnon, "China's Censorship 2.0: How companies censor bloggers," *First Monday [Online],* 14, http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2378/2089 [16 August, 2012, 2009].

[10] D. Bamman, B. O'Connor, and N. Smith, "Censorship and deletion practices in Chinese social media," *First Monday,* vol. 17, no. 3-5, 2012.

[11] L. Yu, S. Asur, and B. A. Huberman, "Artificial Inflation: The True Story of Trends in Sina Weibo," *Arxiv preprint arXiv:1202.0327*, 2012.

[12] 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, pp. 689-703, Jun, 1990.

[13] D. Cho, S. Kim, and A. Acquisti, "Empirical Analysis of Online Anonymity and User Behaviors: The Impact of Real Name Policy," in International Conference on System Sciences, Hawaii, 2012, pp. 3041-3050.

[14] H. Tseng *et al.*, "A conditional random field word segmenter," *Fourth SIGHAN Workshop on Chinese Language Processing*.

[15] R Development Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing Vienna, 2011.

[16] Y. Yang, and J. O. Pedersen, "A comparative study on feature selection in text categorization," in 14th International Conference on Machine Learning, 1997, pp. 412-420.

[17] K. W. Fu, and M. Chau, "Reality Check for the Chinese Microblog Space: a random sampling approach," under review.

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

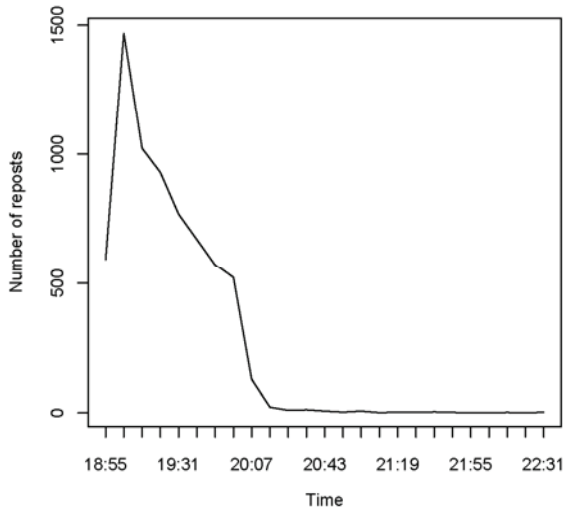
Table 1a/b: Top 30 keywords for censorship and PAM status according to χ^2 value. Keywords in red exist in both lists.

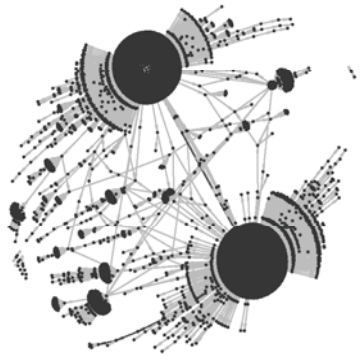
	Predictors for censorship			Predictors for PAM		
Rank	Terms	χ^2	RR	Terms	χ^2	RR
1	重庆 (Chungking)	302.38	3.25	两会 (Two meetings)	75.88	5.67
2	光诚 (Guangcheng)	248.62	32.82	雷 锋 (Lei Feng)	56.66	3.41
3	陈 光诚 (Chen Guangcheng)	237.42	46.72	白色 情人节 (White Day)	37.07	5.86
4	两会 (Two meetings)	232.29	5.39	人大代表 (National 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 rumours)	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	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 (Mar-8)	22.04	2.56
13	人大代表 (National People's Congress)	152.5	4.23	bed 凌乱 (bed, mess)	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 ('s expectation)	19.24	2.95
23	转 么 (retweet)	122.13	245.97	代表 委员 (representive, committee)	19.04	6.56

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

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 (Mar-8 Festival)	18.29	3.7
27	删 (deleted)	108.53	2.97	公知 (Public Intellectual)	18.29	3.7
28	陈 (Chen (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, not)	100.08	201.97	临时工 (Temporary job)	17.95	4.29

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE





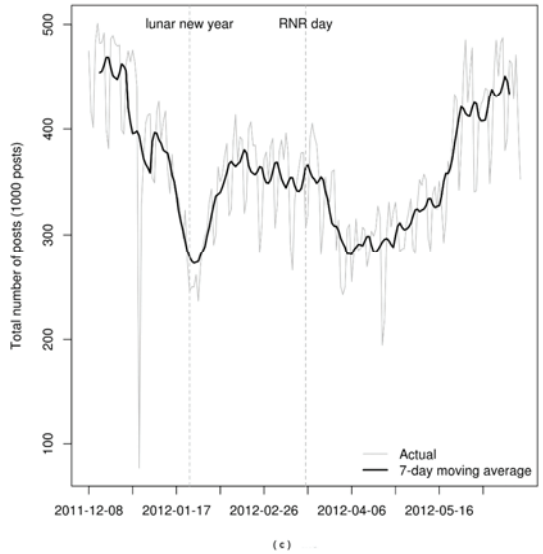


Figure 1a) How Blogger H's weibo was posted, diffused and finally censored

Figure 1b) A repost graph showing that the reposts were primarily diffused via the Internet space by a few microbloggers with a large amount of followers (node size is proportional to the number of followers of a microblogger)

Figure 1c) Daily frequency of posts made by 166,725 microbloggers

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE

About Authors

King-wa Fu is an assistant professor at the Journalism and Media Studies Centre, The University of Hong Kong. His research interests include political participation and online media use, health communication, and computational and statistical methods in media studies. He received a bachelor's and an MPhil's degree in Electrical and Electronics Engineering and hold a PhD degree in Journalism and Media Studies Centre, The University of Hong Kong. He is a member of the IEEE. Contact him at kwfu@hku.hk

Chung-hong Chan is a research assistant at the Journalism and Media Studies Centre, the University of Hong Kong. He is interested in the application of statistical methods to detect content censorship in social media. He 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 at the University of Hong Kong. His current research interests include Web mining, business intelligence and analytics, knowledge management, and IT education. Chau received a PhD in management information systems from the University of Arizona. He is a senior member of the IEEE, the ACM, and the Association for Information Systems. Contact him at mchau@business.hku.hk.

Digital Object Indentifier 10.1109/MIC.2013.28 1089-7801/$26.00 2013 IEEE