

Limitations of Chinese Social Media Censorship and Their Relationships with Government Policies

Zhizhuang Chen^{1,*}

¹Shanghai Shangde Experimental School, Shanghai, China

*Corresponding author: chenzhizhuang@shangdejy.com

Abstract. In an increasingly anxious global context, the Chinese Government is putting more pressure on media censorship because of the Ideological dichotomy and political disputes. Social media, as a more influential platform than conventional media, make serval changes to new policies made by the Chinese Government. The first part of the essay investigates two typical examples of these social media to conclude the limitations of censorship, this includes 1. The Community Specification Contains Many Ambiguous Statements, 2. No Banning Details Given, 3. Differentiated Censorship Intensity, and 4. Unwritten Rules Besides Community Specification. These limitations in social media censorship are generated by the influence of government policies. The second part explains the factors that cause social media companies to choose to sacrifice user experiences to fit into regulations under the pressure. In addition, the third part evaluates the harm of these limitations and censorship trends, also the possibility of future problem-solving. Overall, the essay focuses on the relations between policies and platform regulation, emphasizing the impact of government involvement.

Keywords: Government Policies, Social Media, Censorship, Campaign-style enforcement.

1. Introduction

Since China's reform of the market economy in 1978, the Chinese government loosened control of the media to motivate the commercialization process [1]. But as China's economic power and social media market has grown rapidly, speech censorship has gradually tightened. Especially in the recent two years (2020~2022), the Covid-19 pandemic and the increasingly anxious international situation have made this problem more serious. Around this issue, many articles were already discussing how severe censorship helps the Chinese government consolidate power and guide public opinion from a macroscopic perspective [2]. But there were fewer articles discussing how emerging social media companies react exactly to the policy changes, simply put, the relationship between government policies and social media censorship. Meanwhile, most of them are out of date, because it has changed rapidly during the pandemic period. The reason that makes it worth studying is that Chinese public social media platforms are more influential among the general public and more capable of determining the direction of social opinion nowadays, compared with conventional media such as paper and television. Also, studying its reaction mechanism is helpful to understand the complete image of Chinese censorship.

The essay body can be divided into three parts, elaboration of limitations, the possible factors of limitations, and the possibility of future problem-solving. About the research methodology, the essay chooses two typical social media platforms that have different audience basis and community philosophies as research subjects, their code names are A and B respectively. The summary of limitations mainly relies on the collection of user experience and the detailed provisions of platform censorship, subject to the non-disclosure of detailed data. The investigation of the relationship between government policies and censorship based on the government regulatory records, through the timeline, it could show how the platform makes changes after certain regulations or events. In addition, evaluating the possibility of future problem-solving on two aspects of phenomena and factors.

2. The Four Main Limitations of Chinese Social Media Censorship

Chinese Social Media Censorship is accused of being too strict and impersonal [3], for this phenomenon, the author believes that it can be concluded into four main limitations by reviewing the collection of user experience and the detailed provisions of the platform censorship of social media companies A and B. Before elaborating on these limitations, the following is the basic introduction of these two companies. Company A is a social platform that revolves around Q&A and knowledge sharing. In the early stage, the platform's audience mainly aimed at people being higher educated, meanwhile their main product differentiation was rationality and objective discussion. After it had a certain influence, it absorbed more levels of users to expand its market share. Company B is a social media platform that focuses on video sharing of independent content creators. Initially, the audience mainly targets Japanese anime fans, and later, like Company A, it also chooses to expand its user base by absorbing more levels of users. Comparing these two, Company A has a more detailed community specification, which plays as the censorship standard, than Company B, because Company A used to treat rationality as their product differentiation. But on the extent of practical implementation, they show similar limitations and problems.

2.1 The Community Specification Contains Many Ambiguous Statements.

There are many ambiguous expressions in the community specification of both platforms, which contain elements requiring value judgment, a point that can make it very difficult to operate in practice. For example, the community specification clause 4.3 of Company A (The version of 2022.08.12), which titled is "*Content That May Raise Risks*", states that the words "*Deliberately cause confrontation, conflict, and intensification of conflicts between groups or individuals. Causing strife and confusion in the community through fabricated or mixed positions, stories, opinions, and gossip.*" will be deleted, because "*The above types of questions, answers, and articles can cause confrontation and tear the community apart, damaging the community atmosphere and trust.*" However, there is no easy way for censors (People who work in social media censorship) to determine whether or not conflicting views are causing confusion and strife in practice. Also, community atmosphere and trust are two criteria that are hard to quantify, since they exist within social media community relations, staff cannot easily determine the impact of individual statements on them. Similarly, in the community specification of Company B clause 6 (The version of 2022.08.13), titled "*Garbage Content*", states that "*Repetitive, purposeless or off-putting content*" would be treated as "*Garbage Content*" and deleted. But apparently, censors cannot easily determine whether a paragraph is a purposeless or off-putting content. It requires many subjective value judgments that are unable to convince users of the reason they were banned.

2.2 No Banning Details Given

When a particular statement is censored and banned on the Chinese social media platform, people seldom got detailed feedback or explanations. For example, when the content is banned in platform A, the only feedback received in most cases is "*Contains content that is prohibited by law or may cause controversy*". However, exactly which laws are violated and what may be in dispute will not be explained in detail. On the other hand, the efficiency of the platform's complaint channel is very slow, and even if you get customer service feedback, you can't know which specific content violates the rules and needs to be modified. This is even more serious in the B platform, which will use an automatic censorship system to manage the comment section of the video. The system will delete all comments with banned words without discrimination, regardless of the opinion expressed in the comment, and will not give the sender a message indicating that the comment has been deleted. Moreover, it has a credit system, when a user has been deleted too many times, the system will automatically hide his comments.

2.3 Differentiated Censorship Intensity

With the vague criteria mentioned in the first and second points, the strong and arbitrary review mechanism leads to the censor having the almost absolute power to interpret the vague space between these vague spaces. On the other hand, vague criteria also force the censor to use subjective judgment content. [4] This huge power and way of working cause the personal emotions and ideologies of the censor can largely affect the censorship results. For example, the community specification clause 3.2 of Company A and clause 5 of Company B clearly state that “*Content that incites ethnic hatred, ethnic discrimination, and undermines national unity is not allowed*”. But in reality, hate speech based on extreme Chinese nationalism and racism can openly appear in the comments. For example, insulting remarks about Japanese, Koreans, or African people are often made on the platform. In contrast, the community specification clause 2 of Company A and clause 1 of Company B emphasize that “*Pornographic and vulgar content*” is not allowed. The censorship intensity of this type of information is much stronger than the racism one. The automatic censor system can quickly find prohibited content containing keywords, so Chinese internet users usually use homonyms or unusual synonyms to refer to such information. [5] But for racism or extreme nationalism hate speeches, the response speed of the censorship system is much slower, even the censorship system can do so.

2.4 Unwritten Rules Besides Community Specification

In addition to the criteria already written on the platform’s community specification, there are unspoken rules that will not be indicated in the rules. For example, the censorship intensity under politics-related topics in Company A would be especially high, answers may be suddenly deleted or comment restricted, even if the content does not violate any written regulations. [4] In company B, the comment section of videos issued by official media is closed by default, only the so-called selected comments will be shown. Yet these censorship strategies are not written in the community specification, no other notification to users either. Thus, they are unspoken rules that exist in the platform.

3. The Factor of Limitations

Reviewing government regulatory records, it could show how the platform is influenced by certain regulations or events throughout the time relationship. After research, the author thinks that the reasons for these limitations can be understood in two ways, and it is the combined effect of both two factors that result in this situation. In addition to these two, there may be other factors that also have an impact on it, but in general, these two are the main factors.

3.1 Government Emphasis on the Control of Social Opinion

Compared to social media in other countries, China's management approach places more on top-down management measures, emphasizing that controversies, "misconceptions" and rumors on public platforms are harmful speech and must be eliminated [6], focusing on controlling the direction of social opinion [3]. And clearly state that platforms should be held accountable for the speech that appears in them. Typical examples are the banning of profanity or the promotion of pornography, gambling, and drug abuse. [4] Objectively speaking, such kind of censorship has some positive effects. However, under this "emphasis on using management to control speech to create a good Internet environment," many comments that criticize social phenomena and express negative emotions are gradually considered "harmful speech" that needs to be controlled. In order to comply with government policies and rationalize radical management measures, the platform has added many vague, subjective judgment-requiring censorship clauses to the community specification. But these extra terms of regulation are costly in terms of labor if they are carefully censored [7], thus profit-oriented companies will take different approaches to save costs. For example, in Company A, censors are required to process a large amount of text in a short time, generating a huge disparity between the amount of censorship ability and information. Eventually leads to a lack of banning details, inefficient

customer service feedback, and differentiated censorship. And in Company B, an automatic censorship system is used to quickly block content with keywords, to avoid complaints and negative feedback, users are not even alerted their messages are blocked. Under the dual pressure of policy and cost, platforms choose to respond with rough, imprecise censorship.

3.2 Campaign-style Enforcement Used in Social Media Regulation

Chinese government agencies adopt a "campaign-style enforcement" to regulate social media companies. Campaign regulation means that resources are suddenly pooled at specific times for large-scale, targeted enforcement as opposed to regular legal restrictions and inspections [8]. Although this approach has been very effective in dealing with environmental pollution [8], such swift justice could probably distort the criminal justice system by applying too much pressure [9]. This disadvantage would be vital in the area of speech censorship. China does not have a specific industry code and well-established departmental supervision in social media company management but uses an approach of suddenly "interviewing" some of them to make a cautionary example to the industry. For example, company A was abruptly "interviewed" by the Beijing Internet Information Office (Full name: Office of Network Security and Informatization Committee of the CPC Beijing Municipal Committee) in December 2021, the excuse is "repeatedly appearing content that is prohibited from publication or transmission by law or regulation". In the process of "interviewing", the entire social media platform is prohibited from commenting, thus losing a lot of new users. Moreover, its share price dropped by more than 8% and even more than 10% at one point. After the "interviewing", not only Company A has strengthened its censorship intensity of the whole platform and re-censored the content posted in the past, many other social media platforms have taken similar measures to protect themselves, including Company B. The automatic censorship system mentioned above was introduced at this time point. This campaign-style enforcement leaves platforms with a natural tendency to delete more, as tough censorship only affects their user experience, but being "interviewed" means dramatic swings in revenue and share price. In order to rationalize this harsh censorship mechanism, platforms have set up many vague community specification clauses. Those that cannot be elaborated in the community specification are generally evolved into unwritten rules.

4. Possibility of Problem-solving

These limitations mentioned above have a huge impact on the discursive environment of Chinese society, the most important of which is that it greatly undermines the regulatory role of social opinion. When "sensitive" events occur, social media platforms do not dare to let go of the discussion, instead, adopt a blockade response program, prohibiting people from communicating and discussing with each other. It will be liberalized only after the official response; thus, the public loses their right to monitor the process. For example, the Xuzhou chained woman incident [10] and the Henan bank protest event [11]. And more seriously, some events have a low discussion and do not get official attention. This blockade will make it gradually lose exposure, and finally, the problem gets no attention and solution. For example, the Wuhan home loan problem [12]. On the other hand, it has shaped the way people communicate, as the bottom line is not clear, self-censorship becomes a crucial problem. As the trend of "not talking about political affairs" and "not talking about sensitive topics" spreads more and more, individuals will eventually lose their sense of social responsibility, and the concept of democracy in a country will be severely damaged. [13] Therefore, in the long run, reform is crucial, and the following are two possible solutions focusing on the phenomenon of social media and the factors of limitations respectively.

4.1 Side on Changing Phenomenon, Intervening Censorship Quality, and Accuracy

Starting from the phenomenon of limitations, the government can introduce relevant industry regulations to force companies to improve the quality of censorship, thus can optimize the discussion environment. In the logic above, it can be found that companies wanting to reduce the cost of

censorship is one of the key factors of many problems. [7] In the trade-off between the government's policy requirement for "*control of negative speech*" and the gradually rising cost of censorship and maintenance, companies choose to lower the quality of censorship and sacrifice their user experience, leading to the gradual deterioration of the discussion environment. This dilemma could be improved if the government issued regulations to force platforms to raise their review costs. But the problem with this approach is that it only optimizes the quality of censorship, but the nature of the problems of vague regulations, differential enforcement, and unwritten rules, those related to radical social opinion control, will not be changed. And while the benefits of optimizing the social media discussion environment are latent and very long-term, the economic impact of intervening in the market and raising business costs will immediately be reflected in the economic data. The market does not welcome such changes, and the government has no strong desire to reform because not many officials think the present censorship is too harsh, so it will be difficult to promote the implementation of an intervention.

4.2 Side on Changing Factors, Reforming Management Approach.

Focusing on the causes of the limitations, the government needs to change its management direction or management approach to some extent. First, in terms of management direction, it should be clear that in the current situation of international ideological warfare, the Chinese government will not easily change the general direction of speech censorship, at least in the short term, that is, the view that "top-down regulation on negative speeches is important" and the intention to protect ideological security [14]. It is worth recognizing that this does work in controlling the spread of information about crimes such as drugs, but it can easily expand into some normal discussions and critical remarks. Without changing the general direction, the government needs to clarify the boundaries involved to prevent criticizing and monitoring speeches from being misplaced as well. So, reforming the management approach is an available choice. By changing the campaign-style enforcement, and introducing regular government sectoral regulation with specific industry guidelines, thus companies can understand the complete specification content instead of judging the requirements and policies by guessing. If this happens then, platforms and individuals will not fall into a crisis of self-censorship and the problem of unwritten rules of review can be well resolved. However, the drawback with this approach is that the problem of the high cost of censorship will not change, the companies will still reduce the quality of censorship to save costs, and the issue of low feedback efficiency and differential enforcement will continue to disturb. Moreover, the lack of motivation to reform also exists similarly in this approach.

A combination of the two approaches might yield better results, but it would also mean higher government spending and social costs. Reviewing the limitations and factors, it all comes down to the idea that "top-down regulation of negative speech is important." It costs a lot to keep the public opinion under control and would take more when trying not to make misplaces. In this context, since the protection of the speech environment takes a back seat to the many social problems in China that require spending, the problems with China's social media censorship will continue for a long time.

5. Conclusion

In summary, this paper examines the limitations of China's public social media censorship and its factors and solutions, especially the relationship between limitations and government policies. For now, the problems in China's public social media censorship will persist, mainly because the government lacks the incentive to solve them and their impact receives less attention compared to other social issues. In the future, these problems may even be exacerbated as policies are tightened.

It is worth mentioning that this paper has many limitations in the research methodology, because of the lack of information sources. For example, when investigating social media censorship limitations, it would be helpful to have data on the number of deletions, or a thesaurus of banned words, but the platform is unlikely to publish relevant information, so most data used is from

estimated data given by third parties. Even more third parties now very difficult to collect relevant data, because most Chinese social media accounts require real-name authentication, there are not enough accounts to conduct large-scale experiments anymore. Another example is the automatic censorship system of Company B. The platform itself does not directly announce its existence, it is all gradually discovered by many users in the process of using it. So, the elaboration of relevant content is possible to make mistakes.

References

- [1] King, G., Pan, J., Roberts, M. E. How Censorship in China Allows Government Criticism but Silences Collective Expression. *American Political Science Review*, 2013, 107(2): 326–343.
- [2] Roberts, M. E. *Censored: Distraction and diversion inside China's Great Firewall*. Princeton University Press, 2018.
- [3] Xu, B., Albert, E. *Media censorship in China*. Council on Foreign Relations. February 17, 2017. Retrieved August 24, 2022, from <https://www.cfr.org/background/under/media-censorship-china>
- [4] Jason Q. Ng. *Tracking censorship on WeChat's public accounts platform*. The Citizen Lab. August 26, 2020. Retrieved August 26, 2022, from <https://citizenlab.ca/2015/07/tracking-censorship-on-wechat-public-accounts-platform/>
- [5] King, G., Pan, J., Roberts, M. E. Reverse-engineering censorship in China: Randomized experimentation and participant observation. *Science*, 2014, 345(6199).
- [6] Garrick, J., Bennett, Y. C. *China's socialist rule of law reforms under Xi Jinping*. Routledge. 2016. ISBN 9781317354161
- [7] Fedasiuk, R. *Buying silence: The price of internet censorship in China*. Jamestown. January 12, 2021. Retrieved August 20, 2022, from <https://jamestown.org/program/buying-silence-the-price-of-internet-censorship-in-china/>
- [8] Jia, K., Chen, S. Could campaign-style enforcement improve environmental performance? evidence from China's Central Environmental Protection Inspection. *Journal of Environmental Management*, 2019, 245, 282–290.
- [9] Wang, P. Politics of crime control: How campaign-style law enforcement sustains authoritarian rule in China. *The British Journal of Criminology*, 2020, 60(2), 422–443. <https://doi.org/10.1093/bjc/azz065>
- [10] Kuo, L., Li, L.. *Saga of the chained mother of eight continues to roil china, inspiring rare social activism*. The Washington Post. February 25, 2022, Retrieved August 20, 2022, from <https://www.washingtonpost.com/world/2022/02/25/xuzhou-chained-woman-china/>
- [11] BBC. *Henan: China to repay customers after Mass Bank protests*. BBC News. July 12, 2022. Retrieved August 20, 2022, from <https://www.bbc.com/news/business-62130744>
- [12] Danxu, Y. *Chinese home buyers are refusing to pay their mortgages: Is there a way out?* ThinkChina. July 20, 2022. Retrieved August 20, 2022, from <https://www.thinkchina.sg/chinese-home-buyers-are-refusing-pay-their-mortgages-there-way-out>
- [13] Luqiu, L. R. The cost of humour: Political satire on social media and censorship in China. *Global Media and Communication*, 2017, 13(2), 123–138.
- [14] Zhouxian, Z., Zhidong, X. Information Technology Development Trend and Ideological Security. *Political Work Journal*, 2015, 2015(3):2