An Analysis of the Themes of Refuting Rumors on China's Sina Weibo During the Period of Covid-19

Difan Guo
School of Journalism and Communication, Nanjing Normal University, Nanjing, China
guodifan@163.com
Diserahkan: 7 Juni 2021; Direvisi: 14 Juni 2021; Diterima: 15 Juni 2021

Abstract
From the end of 2019 to 2020, there were countless rumors on the Internet related to COVID-19 during the viral epidemic. This study analyzed how government Weibo, the official news release channel of government social media, refuted rumors on China's leading social media platform Sina Weibo during the COVID-19 pandemic outbreak in China. This study used the LDA topic model to model the Weibo text topic and obtain the topics of the rumors that the government Weibo defied. This study find that the five main topics of rumors presented in the anti-rumor Weibo are highly related to the operation of the social system, disease prevention and treatment, and social security.

Keywords: Anti-rumor, LDA, Covid-19, Subject Analysis.

Abstrak
Dari akhir 2019 hingga 2020, ada banyak rumor di Internet terkait COVID-19 selama epidemic virus. Studi ini menganalisis bagaimana Weibo pemerintah, saluran rilis berita resmi media sosial pemerintah, membantah rumor di platform media sosial terkemuka Tiongkok Sina Weibo selama wabah pandemi COVID-19 di Tiongkok. Penelitian ini menggunakan model topik LDA untuk memodelkan topik teks Weibo dan mendapatkan topik rumor yang ditentang pemerintah Weibo. Studi ini menemukan bahwa lima topik utama rumor yang disajikan di Weibo anti-rumor sangat terkait dengan pengoperasian sistem sosial, pencegahan dan pengobatan penyakit, dan jaminan sosial.

Kata Kunci: Anti-rumor, LDA, Covid-19, Analisis Subjek.
INTRODUCTION

In a public health emergency, media communication is as essential as laboratory analysis and epidemiology (WHO, 2020). Effective media communication plays a decisive role in providing emergency information, pacifying the public, and saving lives. In 2020, an estimated 3.6 billion people worldwide are using social media, and the trend is increasing (Clement, 2020). Social media’s role in daily health behavior education and emergency and recovery actions in public health emergencies has been proven (Thackeray, Neiger, & Smith, 2012). Medical institutions and public health management departments have also realized the importance of social media. In public health events, such as epidemic outbreaks, social media have been actively used to release information, monitor disease trends, and communicate with the public (Fung & Wong, 2013). During the COVID-19 pandemic, the World Health Organization, the US Center for Disease Control, and the Chinese Center for Disease Control and Prevention have released information promptly through official social platforms.

Social media platforms are critical communication tools for engaging the public in many health issues. The public can use social media to not only obtain information about various health issues from health organizations and news outlets but also to actively participate in web-based discussions with peers and influencers to exchange perspectives about these topics (Limaye, et al., 2020). On the other hand, the adverse problems of social media in public health incidents have also received attention. Due to the lack of social media fact-checking, social media is prone to become a channel for the dissemination of misinformation and rumors (Merchant, 2020; Merchant & Lurie, 2020). During the COVID-19 pandemic, a large amount of false information was spread on social media platforms and was recognized by the World Health Organization as an "infodemic" (WHO, 2020). In the few months since the first COVID-19 cases, a wide range of misinformation has spread across traditional media and social media, resulting in what WHO calls an infodemic (excessive amounts of misinformation, disinformation, and rumors that make it difficult to identify reliable sources of information). Misinformation can be lethal, as evidenced by the propagation of premature evidence suggesting that chloroquine is an effective treatment for COVID-19 (Limaye, et al., 2020).

In response to the false information and rumors, the World Health Organization's risk communication and social media teams closely followed the media, refuted the most common rumors that may harm public health based on evidence-based information, and provided public health information and advice through many social media channels, including Twitter and Facebook.

In China, after the outbreak of the COVID-19 pandemic, many rumors also appeared on commonly used social media platforms, such as WeChat and Weibo. Led by the China Center for Disease Control and Prevention, the disease control department and relevant government departments began to publish information to refute rumors on official websites and social media platforms. Among them, on Sina Weibo, which is one of China’s most influential social media platforms, local governments use government Weibo accounts that represent official news speaking channels to publish anti-rumor information as the official position.

Based on the spread of rumors on Chinese social media platforms and the refutation of rumors on government Weibo accounts during the COVID-19 pandemic, this study proposes the following main research questions: (1) During the outbreak of the COVID-19 pandemic, what is the overall situation of the rumor refutation of the government Weibo accounts? (2) Judging from the anti-rumor information published on the government Weibo accounts, what are the main topics of widespread rumors during the COVID-19 pandemic?

LITERATURE REVIEW

Several studies on Covid-19 news and disinformation have been conducted, particularly on the dynamics of information on social media platforms. Misinformation can have fatal consequences, as shown by the spread of premature evidence suggesting chloroquine is an effective treatment for
COVID-19. Given the COVID-19 pandemic's exponential expansion, the unchecked and quick spread of misinformation, driven mostly by social media, poses a serious public health problem for COVID-19 control and mitigation measures, as misinformation undermines public confidence, consensus, and subsequent action. (Limaye, et al., 2020) on their article explained that Unlike the litany of debunked misinformation perpetuated by anti-vaccine activists, our understanding of COVID-19 is evolving continuously. As such, what qualifies as misinformation will be sensitive to new scientific discoveries and insights, making it in turn, harder to eliminate misinformation. That is, distinguishing misinformation from legitimate information is, at present, a moving target. The stakes are exceptionally high. With the scale and reach of the pandemic increasing exponentially, there is an urgent need to establish practices to effectively disseminate current, accurate information and quickly identify and root out outdated guidance or misinformation. Government agencies' efforts to collaborate with social media behemoths Facebook, Twitter, and Instagram demonstrate how public–private collaboration to flag, fact-check, and even remove false or outdated information can be an effective way to prevent these networks from impeding and even worsening public health efforts. Eliminating disinformation can assist social media users in gathering and disseminating accurate information, allowing them to remain secure while also reducing risk to others. New tools, such as crowd intelligence-based disinformation detection, can help social media sites root out misinformation. 

A case study in Italy regarding misleading information about COVID-19 was conducted by Lovari (2020) through a research titled ‘Spreading Dis(Trust): Covid-19 Misinformation and Government Intervention in Italy’. The analysis focuses on the dissemination of COVID-19 disinformation in Italy, emphasizing the variables that have influenced pandemic communication. Italy has recently experienced a progressive erosion of trust in public institutions, as well as a general condition of information crisis in health and science. In this context, the politicization of health issues, as well as the growing use of social media to combat the Coronavirus "infodemic," has led the Italian Ministry of Health to take a strategic role in using its official Facebook page to mitigate misinformation and provide updates to online publics. Despite this prompt intervention, which increased the visibility and dependability of public health communication, it appears that coordinated efforts involving various institutions, media, and digital platform companies are still required to reduce the impact of misinformation, as using a multichannel strategy helps avoid increasing social and technological disparities during a crisis (Lovari, 2020).

On top of that, China has also had several studies related to the dynamic of COVID-19 information on many digital platforms. Using earlier research on how online stories spread, the study fills a vacuum by looking at the retransmission of COVID-19 rumors and correction messages on Sina Weibo, China's largest and most popular microblogging site. The study investigates the impact of rumor kinds, content aspects (including frames, emotion, and logic), and source characteristics (including follower size and source identification) on the chance of a COVID-19 rumor and its correction being shared. The study is not only increase scholarly understanding but also discloses how corrective messages might be produced to disprove cyber-rumors in specific cultural contexts by investigating the retransmission of rumors and their corrections on Chinese social media. The effects of source type were noteworthy.

The study finds the sources that the public deemed more legitimate than ordinary users were more effective in enhancing retransmission of both rumors and rumor-corrections. The role of news aggregators was prominent in the retransmission of rumor-correction messages: The likelihood of a news aggregator's post being shared was 53 times higher than that of ordinary users, while other legitimate sources such as government authority and traditional media were only about 15 times more or so. News aggregators also played a role, along with public figures. (mostly celebrities) in retransmissions of rumor posts. News aggregators on Weibo are distinguished from “bots” in Twitter in terms of legitimacy. While bots in Twitter could range from official news bots to spam
bots and fake news bots, the news aggregators have verified Weibo accounts that specialize in releasing a variety of news. According to Weibo policy, news aggregators have to maintain more than 1,000 followers and more than 50,000 views in 30 days to attain their verification status. Government and professional news media do not have to follow such requirements. As such, news aggregators are influential in the first place, and the findings of this study show that news aggregators can contribute to the retransmission of both false rumors and correction messages (Yunya, Hazel, & Yin, 2020).

In addition, based on a comparative study from her journal article entitled BLABLA, it was found that in the predominantly US social networks, represented by Twitter, participants were more concerned about the impact of COVID-19 on the economy and science and technology issues. Meanwhile, in the other social media sphere, represented by Weibo in China, politics and public health development were always the center of the discussion. A possible explanation for this is that the two governments applied different measures to control the spread of the virus during the early stages of the crisis. The study also found that combining the results of LDA and temporal analysis, such as for the issues of Domestic Situation and International Situation, which appear alternately on Twitter and Weibo, to some extent reveals that online discourse reflects the crisis lifecycle according to the different stage of COVID-19 occurring in China and the US. However, the differences in the level of discussion of positive issues such as Response and Measure and negative issues such as Worried Emotion in different countries reflect their differences in society, value, cultural background, health communication, and administration systems in the real world beyond the social media ‘sphere’ (Wen & Yi, 2021).

METHODS

We collected microblog data from the official Weibo accounts of Chinese governments at all levels with a strong influence on the Sina Weibo platform during the critical period after the outbreak of the COVID-19 pandemic. We collected microblog text, comments, release time, number of reposts, number of comments, and other data related to anti-rumor topics through data mining. We cleaned the collected data, carried out the topic modeling and topic analysis of related texts through the LDA topic model.

Sina Weibo evaluates the top 100 most influential government Weibo accounts every month based on four dimensions: communication power, service power, interaction power, and recognition (Government Weibo List Rules, 2020). In mid-January 2020, the COVID-19 pandemic broke out on a large scale in China, and the Chinese media began to pay attention to it. In response to the spread of many rumors regarding the outbreak, the China CDC conducted the first official rumor dismissal on January 18 (China Center for Disease Control and Prevention, 2020). Considering the above factors and according to the rankings of government Weibo accounts in January and February 2020, a total of 80 official Weibo accounts launched by governments at the municipal level and above that were listed during the two months were selected. Besides, the period for collecting data is set from 0:00 on January 18, 2020, to 23:59:59 on February 29, 2020, Beijing time.

We collected a total of 1,699 rumor-defying microblog texts and other data from the 80 selected accounts through the web crawler tool. After initial cleaning of the data by deleting microblogs containing inaccurate or incomplete content, 1242 valid microblog texts and other data were finally obtained. According to statistics, the number of comments to the 1,242 microblog texts was 52,878, and the number of reposts was 31,689 after cleaning.

Wuhan was the first outbreak area of the COVID-19 pandemic in China (Huang, Wang, & Li, 2020). After the outbreak, rumors related to Wuhan spread wantonly. As a platform for the information release of the Wuhan Municipal Government, its official Weibo account @武发布 has become an essential channel for dispelling rumors. We conducted statistics on the cleaned data and found that among the 80 government Weibo accounts, the number of rumor-defying microblog
posts published by @武汉发布 accounted for 8.45% of the total sample, amounting to 105. The number of comments on the 105 posts was 10,910, and the total number of reposts was 3966. Given the particularity of Wuhan's response to the COVID-19 pandemic and the significance of the number of rumors dispelling microblog posts and the comments published by @武汉发布, we paid special attention to it.

The topic model extracts appropriate text content topics and appropriate vocabulary to describe and model the text content, which is currently a typical natural language processing method (Hu & Chen, 2014). By establishing a topic model, the hidden structures and latent semantics can be discovered in a large text corpus, to realize the topic analysis of the text (Nicola, Giuseppe, & Ettore, 2013). We used the LDA topic model for topic modeling, and finally got the text content topic and essential vocabulary in this study. LDA is based on the bag of words model, which considers that documents and words are interchangeable, ignoring the order of words in the document and the order of the document in the corpus, thereby transforming text information into digital information that is easy to model (Blei, Ng, & Jordan, 2003). The characteristics of LDA give it excellent short text processing capabilities, and it is widely used in short text processing such as Twitter and Weibo (Griffiths & Steyvers, 2004).

The LDA model can be represented by a directed probability graph (Figure 1). Its model structure has three levels. In this model, $\alpha$ and $\beta$ are parameters of the corpus layer. $\alpha$ reflects the relative strength of the implicit topics in the document collection, and $\beta$ reflects the probability distribution of all implicit topics themselves. $\theta_d$ and $\phi_k$ are variables at the document level, $\theta_d$ representing the topic probability distribution of the document with serial number d and $\phi_k$ representing the probability distribution of the feature words in a particular topic. M represents the number of texts in the document set, and K represents the number of topics in the document set. $z_{dn}$ and $w_{dn}$ are variables at the feature word level, $z_{dn}$ representing the nth word in the document with serial number d, which belongs to observable variables; $w_{dn}$ represents the topic of $z_{dn}$, which belongs to hidden variables. N represents the number of feature words contained in a document.

![Figure 1. Schematic diagram of LDA](image)

In practice, we cleaned the data of effective Weibo texts, and used LDA model for analysis. We used 90% of the data as the training set and 10% of the data as the test set. Since the data were not of many orders of magnitude, it is more accurate to use approximately 10% of the data to evaluate the effect of the model. By setting, the 10 most frequent words were displayed under each topic, and the number of topics was used as a variable to change the output result by changing the number of topics. After the parameter adjustment test, we found that when the number of topics was set to 5, the clustering effect between topics was the best.

RESULT AND DISCUSSION

We used the LDA topic model to model the topic of the post-cleaning Weibo text and output the top ten keywords with the highest probability. Moreover, we summarized each topic's title based on the output keywords and found that topic clustering was the best when outputting five topics, which better displayed the hidden topic information in Weibo. To make the topic model results easy to
After the outbreak of the COVID-19 pandemic, misinformation and rumors about the epidemic were spread on major social media. It is possible to know the status of the government's management of rumors and the contents of the rumors that spread during this period. From the perspective of social psychology, rumors are a kind of collective behavior influenced by information dissemination. When the social system cannot satisfy the normal life expected by members of society, it is easy for pressure to build and collective behavior to be produced. By extracting the topics of rumors during the COVID-19 pandemic, it is possible to know the focal points that can influence society members to maintain regular social order during this period.

From the results regarding the subjects extracted by the LDA model, the most prominent aspect of the rumors during this period is the content of the "social system operation." There is a contradiction between people's uncertainty about the development of the epidemic and the actual demand for social participation. Coupling with the lack of channels and capabilities to obtain
authoritative information, collective pressure can easily break out, and rumors about the social system's operation can quickly arise. The rumors about "disease prevention" and "disease treatment" are also prominent, reflecting people's uncertainty about the new coronavirus, new coronavirus pneumonia (Han, Wang, Zhang, & Wang, 2020), and the medical system related to treatment. In the face of the sudden emergence of infectious disease, the existing medical methods and systems' coping ability is also a difficult problem for the national medical system. For members of society, the unknown in terms of "disease prevention" and "disease treatment" can easily lead to panic during the epidemic. Therefore, rumors in this area can easily break out (Yin, Lv, Zhang, Xia, & Wu, 2020).

The rumors under the theme of "emergency notice on epidemic" mainly spread misinformation on emergency measures taken by various departments of the social system to face the epidemic. Judging from the consequences of the spread of such rumors, they can easily disrupt the order of the social system (Barry, McGinty, Vernick, & Webster, 2013). These rumors also reflect people's expectations for the emergency response of various departments of the social system and indicate the lack of emergency measures the corresponding departments implement to face the epidemic.

CONCLUSION

The content of the rumors under the theme of "social governance" focuses on the control of illegal and criminal activities by public security and other law enforcement agencies during the epidemic. From the content point of view, some of these types of rumors are concentrated on the ineffective enforcement of the movement of people in the epidemic area, and some are concentrated on the out-of-control criminal behavior during the epidemic. On the one hand, this reflects the possibility of the lack of authoritative information on social governance. On the other hand, it indicates the panic of members of society about whether personal safety and social security can be guaranteed during the epidemic.

ACKNOWLEDGEMENT

My heartfelt thanks to the organizers of the 3nd Jogjakarta Communication Conference and Jurnal Audiens.

REFERENCES

Barry, C. L., McGinty, E. E., Vernick, J. S., & Webster, D. W. (2013). After Newtown—public opinion on gun policy and mental illness. New England journal of medicine, 368(12), 1077-1081.

Blei, D., Ng, A., Jordan, M. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3 993-1022.

China Center for Disease Control and Prevention. (2020, Januari 18). Don't believe these 5 rumors about Wuhan viral pneumonia. Retrived from: http://www.chinacdc.cn/yyrdgz/202001/t20200118_211265.html.

Clement, J. (2020, July 19). Number of global social media users 2010-2021. Statista website. Accessed July 19, 2020. https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users.

Fung, I., Wong, K., (2013). Efficient use of social media during the avian influenza A(H7N9) emergency response. Western Pacific Surveillance and Response Journal, 4(4) 1–3.

Government Weibo List Rules. (2020, March 2). Weibo Data Center. Retrived from: https://gov.weibo.com/rank/rule/index?type=1.

Griffiths, T., Steyvers, M. (2004). Finding scientific topics. Proceedings of the National Academy of Sciences of the United States of America, 101(1):5228-5233.

Han, X., Wang, J., Zhang, M., & Wang, X. (2020). Using social media to mine and analyze public opinion related to COVID-19 in China. International Journal of Environmental Research and Public Health, 17(8), 2788.
Hu, J., Chen, G. (2014). Mining and Evolution of Content Topics Based on Dynamic LDA. Library and Information Service, 58(02):138-142.

Huang, C., Wang, Y., Li, X. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet. doi:10.1016/s0140-6736(20)30183-5.

Limaye, R. J., Sauer, M., Ali, J., Bernstein, J., Wahl, B., Barnhill, A., et al. (2020). Building trust while influencing online COVID-19 content in the social media world. Lancet Digit Health, e277-e278.

Lovari, A. (2020). Spreading (Dis)Trust: Covid-19 Misinformation and Government. Media and Communication, 458–461

Merchant, R. (2020). Evaluating the Potential Role of Social Media in Preventive Health Care. JAMA, 323(5) 411-412.

Merchant, R., Lurie, N. (2020) Social Media and Emergency Preparedness in Response to Novel Coronavirus. JAMA, 323(20):2011–2012.

Nicola, B., Giuseppe, M., Ettore, R. (2013). Probabilistic topic models for sequence data. Machine Learning, 93 (1):5-29.

Thackeray, R., Neiger, B., Smith, A. (2012). Adoption and use of social media among public health departments. BMC Public Health.

WHO. (2020, April 20). WHO ad-hoc online consultation on managing the COVID-19 infodemic. Retrieved from: https://www.who.int/teams/risk-communication/infodemic-management/who-ad-hoc-online-consultation-on-managing-the-Covid-19-infodemic.

WHO. (2020, July 19). Effective Media Communication during Public Health Emergencies: A WHO Handbook. Retrieved from: http://apps.who.int/iris/bitstream/10665/48511/1/WHO_CDS_2005.31_eng.pdf?ua=1.

Yin, F., Lv, J., Zhang, X., Xia, X., & Wu, J. (2020). COVID-19 information propagation dynamics in the Chinese Sina-microblog. Mathematical Biosciences and Engineering, 17(3), 2676-2692

Wen, D., & Yi, Y. (2021). Cross-Platform Comparative Study of Public Concern on Social Media during the COVID-19 Pandemic: An Empirical Study Based on Twitter and Weibo. Int. J. Environ. Res. Public, 1-20.

Yunya, S., Hazel, K. K., & Yin, L. (2020). The “Parallel Pandemic” in the Context of China: The Spread of Rumors and Rumor-Corrections During COVID-19 in Chinese Social Media. American Behavioral Scientist, 1-23.