

CHINESE MAINLAND NETIZENS' PERCEPTIONS OF THE HOKKAIDO EARTHQUAKE BASED ON CONTENT ANALYSIS OF SINA-WEIBO BLOGS

Graduate School of Engineering, Hokkaido University Student Member リアンヤオ シオン (Lianyao Xiong)
Faculty of Engineering, Hokkaido University Regular Member マイケル ヘンリー (Michael Henry)

1. Introduction

Lots of researches show the importance of social network system in facilitating to analyze people's responses towards major disasters¹⁾, and there are many disasters such as typhoons, floods and earthquakes in Japan recently causing huge negative impacts on the local tourism. Thus, it is very important to investigate the international perceptions towards Japan after disasters, especially for Chinese mainland visitors, as Chinese mainland tourists are the main part of the Japanese tourism market²⁾.

Data mining of public blogs is one method to understand the perceptions of Chinese mainland people. Sina-Weibo is known as Chinese Twitter with 203 million daily active clients, 94% of which are the mobile users according to the first quarter financial report from Sina-Weibo in 2019³⁾. Such a social network platform provides a powerful database for datamining. In this paper, a research about Chinese mainland netizens' perceptions towards the earthquake in Hokkaido on Sep 6th, 2018 was conducted based on Sina-Weibo. After retrieving relevant blogs about Hokkaido earthquake on the day of the disaster by web crawler, content analysis was performed to determine the main topics that Chinese mainland netizens were concerned about.

2. Data Collection and Processing

A web crawler based on python was used to mine the public blogs⁴⁾. Based on the URL of Sina-Weibo mobile end (<http://weibo.cn>), datamining relevant blogs to get the bloggers' ids, post contents, publication dates, comments and number of comments regarding the topic of the Hokkaido earthquake on 6th Sep, 2018 was conducted. In order to get as many representative blogs as possible, key words, which are top frequency words of the results from the first crawling based on key phrase 'Hokkaido earthquake', and common Weibo words were added to the searched key phrase. Then, comparison between the dataset acquired by crawling and data from the website was made to check the completeness of mining result, and the blogs matched well. Furthermore, in order to avoid the bias when applying content analysis regarding the only key phrase-based blogs, additional key phrase 'Japan earthquake' was added to cover the topic about Hokkaido earthquake on 6th Sep. Overall, after extracting some blogs which are not posted by Chinese mainland netizens, talk about some irrelevant things or only mention Hokkaido earthquake, a sample of 3675 blogs was acquired for analysis of the netizen's perceptions.

3. Analysis and Results

To figure out Chinese mainland netizens' perceptions about the earthquake in Hokkaido, content analysis about the fined 3675 microblogs was conducted based on KH Coder, a free software for text analysis⁵⁾. Before performing

analysis, some key words such as Hokkaido and earthquake are manually screened out since all the blogs are about the earthquake in Hokkaido. In this paper, attention is paid to the results of word frequency list and analysis of word co-occurrence. Some parts of the results are listed here to help to reach conclusions.

According to the list in Table 1, top frequency key words show the netizens' most concerns about the negative impacts brought about by this disaster such as casualties (injuries, death), building collapse, power outage, airport and landslide and some basic information of the earthquake such as the time, location and epicenter. Interestingly, the word typhoon also appeared many times because typhoon Jebi attacked Kansai area the day before the earthquake in Hokkaido.

Word co-occurrence analysis was also applied to these blogs. With the help of this operation, groups of words which have close association and similar appearance were extracted. Then, the overall images about netizens' focuses can be explored. As is shown in Figure 1, words in the circle of same color are described as one community with one solid line connected to each other revealing strongest co-occurrence. Two words linked with dashed lines are divided into different communities indicating the relationship less close than another word in the same community. And there are 6 communities showing the main perspectives of Chinese netizens.

Community 1(C1) is made up of some words referring to the description of epicenter. "Iburi", "North latitude", "East longitude", "middle east part" and "focal depth" are detailed information about the epicenter. "the Chinese Earthquake Network" is the source of these information.

Community 2 indicates the updating news provided by relevant department of Japan government such as they are trying to figure out the casualties. Additionally, attention was paid to whether the earthquake would cause tsunami.

Table 1 Top 15 terms and their frequency

Terms	Freq.	Terms	Freq.
人 (human)	2177	台风(typhoon)	691
时间(time)	1178	震源深度 (focal depth)	679
地区(area)	1169	房屋(houses)	626
停电 (power outage)	1068	死亡(death)	578
当地(local)	1028	山体滑坡 (landslide)	575
造成(cause)	972	关闭(close)	547
机场(airport)	934	受伤(injury)	535

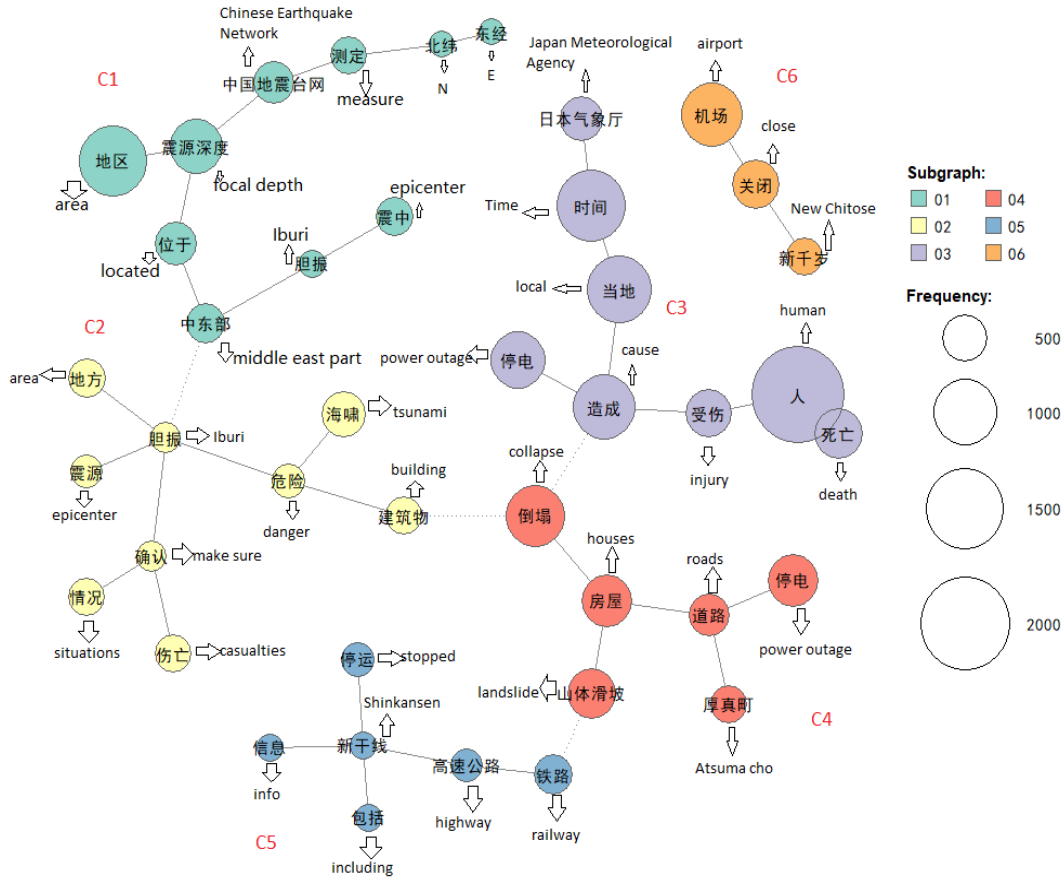


Figure 1 Result of major words' co-occurrence network

Community 3 and 4 are combinations of terms related to the impact of earthquake on the local people such as the casualties, the power outage, the collapse of houses, the damage to roads and the landslide providing a clear and detailed relationship of feature words from the high frequency list. Besides, there are also some information released by the Japan Meteorological Agency such as the local time of the disaster.

Community 5 is about the state of land transportation including trait words Shinkansen, highway and railway which are less frequent than the description of the airport.

Community 6 is consisted of key words airport, close and New Chitose indicating that the netizens were concerned about the situation of the main international airport, the New Chitose airport, which would directly affect some tourists schedule resulting in the strand of tourists due to the close of the airport.

4. Conclusion

To conclude, the topic regarding the earthquake in Hokkaido became very hot on the day when disaster happened gaining great concerns from Chinese mainland netizens. According to the results of word frequency list and analysis of word co-occurrence, their main focuses are summarized into following aspects: 1) the casualties with regard to community 3; 2) the secondary disasters, such as damage to lifeline engineering and landslide fully explained

by community 3,4,5,6; and 3) basic information about this earthquake illustrated by community 1. These images about the disaster in Hokkaido would negatively influence the perception of choosing it as tourist destination for Chinese visitors. On the contrary, this could be eased by active disaster relief measures since the reported operations taken by local government is also a focus point as for the result from community 2.

References

- 1) Yan Qu, P. F. Wu and Xiaoqing Wang, "Online Community Response to Major Disaster: A Study of Tianya Forum in the 2008 Sichuan Earthquake," 2009 42nd Hawaii International Conference on System Sciences, Big Island, HI, 2009, pp. 1-11.
- 2) Ctrip: Report on Chinese Tourists traveling to Japan in 2018, www.199it.com/archives/795489.html
- 3) First Quarter Financial Report of Sina-Weibo in 2019, finance.sina.com.cn/roll/2019-05-23/doc-ihvviews4086253.shtml
- 4) Lin Chen and Fang Ren, "Sina-Weibo Data Crawler Program Design Based on Python" 2016 9th ed Information System Engineering.
- 5) K. Higuchi, A two-step approach to quantitative content analysis: KH coder tutorial using Anne of Green Gables (part I), *Ritsumeikan Social Science Review* 52 (3) (2016) p77-91