

Analysis on the ABC NEWS Report on Huawei Based on Data Analysis

Zhecheng Xiao

MUCCI, Monash University, Melbourne, 3161, Australia

Keywords: Huawei, News Media, Data Visualization, Cluster, International Relations, Data Analysis.

Abstract: With global politics development, the confrontation between China and the United States will become the norm. As the most famous Chinese communication company, Huawei often becomes a victim of ideology. Huawei has been deeply involved in communications construction in Australia in the past decade. Therefore, as a traditional ally of the USA and UK, the Australian government's attitude towards Huawei is also an important aspect that can be studied. The different tendencies of media reports can reflect the international situation to a certain extent. Simultaneously, media act as the government's mouthpiece. Exploring the emotional changes in media can reflect the government's attitude to a particular time. This paper mainly studies how Australian media reports news related to Huawei. ABC News, one of Australia's largest media, was chosen as objectives, as the research object because its coverage is more comprehensive than other Australia media. Based on data analysis, the emotions contained by these news were examined expressed in these reports and produced word clouds. This paper selected 943 news about Huawei from ABC NEWS as its research objects to achieve the above goal. Python language was used to crawl. Through analyzing ABC's news data about their titles and contents, it is found that there is a positive correlation between the number of reports and events. At the same time, most of the reports on Huawei in 2019 and 2020 have a negative tendency.

1 INTRODUCTION

The U.S. government began sanctioning Huawei companies for the first time in January 2018. The U.S. government was firmly opposed to Huawei signing cooperation with U.S. telecom operator-AT&T and banning Huawei mobile phones from entering the U.S. market. Over the next few months, U.S. lawmakers continued to claim that Huawei phones had "back doors" and stolen user data, calling on allies to take the same measures to limit Huawei phones. As a result, in July of the same year, Australia took the lead in banning Huawei from participating in the 5G construction on national security grounds. Although Australia always supports the U.S., the Australian government's attitude towards Huawei has been wavering.

Huawei established its first local board of directors in Australia as early as 2011 as a pilot of its globalization strategy (Cai, 2015). Australia's council of directors' most important work is lobbying the government and building good relations with the local media community (Kania, 2018). Huawei even bought an Australian football team to increase its

influence in the Australian market. Huawei wanted to step into Australia's national broadband program in 2012, although the Australian government finally rejected it on national security grounds; Huawei also participated in constructing and 5G network construction. Huawei's success in the mobile phone market has triggered a strong rebound in the U.S. government and raised Australia's concerns. The natural differences in the political positions of Huawei companies and the Chinese government with Europe and the United States lead to the Australian government distrusting. The Australian government believes that the Huawei network lacks transparency and may be used by the Chinese government (Jennings, 2018).

Under these social conditions, a study of Huawei reporting changes in Australian media is of great value. The author takes ABC news as the research object, hoping to explore its reaction to the rejection of Huawei by the Australian government. In this article, the author will use news data to analyze how the Australian local media reported significant changes in multinational companies. For example, how the attitude of the same news media changes in the course of events, how the importance of

information dissemination has changed, how the content direction of information dissemination changes. Finding experiences and lessons from this news communication's characteristics and improving future cross-cultural business activities is also one of the project's problems.

2 FARMING THEORY

The frame theory of news is different from the traditional news value and goalkeeper theory. This theory can reflect the value judgment of reporters when choosing news and events (Brüggemann, 2014). The frame structure constitutes the fundamental thinking and presentation of NEWS. Therefore, by analyzing news reports' framework, the author can understand the connotation of reports more intuitively. The news reports on the website are mainly composed of a title, author, release time, news description, news text, keywords and pictures. The unified structure standard of news website constitutes the basic news framework and is also the basis of this analysis of news data.

3 RESEARCH METHOD

3.1 Data Capture and Cleaning

The author chose to use python to grab news because it is simpler and more operable than C++. For

example, data statistics and a certain degree of visualization can be implemented. However, in the actual operation process, the author found that only python cannot directly grab ABC news sites. Therefore, the author used the scape framework to help crack the site's anti-grabbing protection. General processes are shown in figure 1 (Fan, 2018).

The author used Request Object to get all the information of the whole page, including text, pictures, Cookie, client certificates, query strings, etc. Therefore, the web pages' layout information obtained from news pages contains a large amount of data that people do not need to use when analyzing problems, such as the code in the header file and the code related to the web pages' layout. Using Beautiful soap can help extract the data we want from the HTML files of a web page, such as URL, news headlines, news organizations.

The author got a table about Huawei report in ABC NEWS through the above data capture and cleaning method. At first, the author wanted to choose 2012 Huawei entering the Australian market as the starting point of the data. However, the author found that the sample size was too small to be representative. Finally, the author chose 2018, Huawei was banned from Australia's national broadband construction as the starting year and captured the data was until 2021. The lines of the table include the news title, release time, primary content. Finally, the total amount of data the author got is 1062; through manual screening, the final valid data' number is 943.

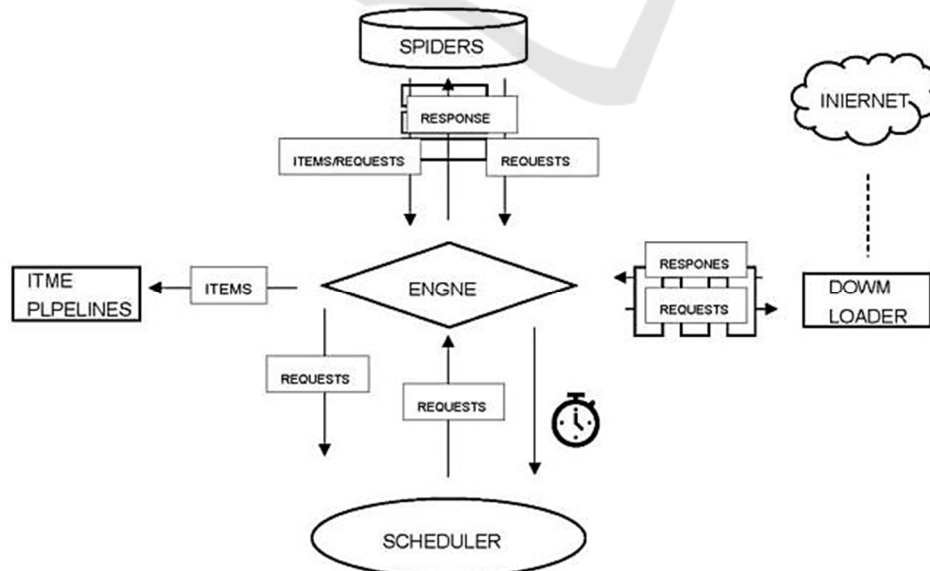


Figure 1: Data capture process.

3.2 Keyword Trend

The consistent structure of a web page makes crawling data possible. After grasping the keywords, the author can start analyzing the keywords. First, we can use the Term Frequency -Inverse Document Frequency (TF-IDF) method for keyword analysis. Some words are classified as keywords because they appear more frequently (Roelleke, 2013). Quantitative data is the purpose of TF-IDF calculation. After quantifying the data, the content of the data can be classified by keywords. News keywords can introduce emotional classification and category tagging. The author can then find whether the emotion of news is positive, neutral or negative by keyword clustering (Topkev, 2016). To judge the tendency of the report. In addition to using data to express news analysis results of news analysis, researcher can also use pictures, tables and colors to express.

On the choice of tools for data analysis and visualization, the author chose Python and Dychart. Dychart is a software that can easily visualize data, and the built-in preset template can greatly reduce the time of making visual charts. Although the author has captured the relevant headlines of the news in the data collection phase, and direct analysis of the title is the fastest and easiest way, the analysis of news headlines may not fully reflect the changes in Huawei by news reports. Hence, in the end, the author chose news content as the source of keyword capture.

After selecting the text as the source of the keyword, the author need to face the most intuitive problem: there are many recurring words. There is not much practical significance and does not play a decisive role in our research, such as punctuation and prepositions or conjunctions. Hence, the author use the python data cleaning function (data clean) to filter punctuation. At the same time, English words contain many changes in tense and plural. At this point, we need to use morphological repair. Lexicalization can transform words into general forms and express complete semantics. For the first time, this will clean up the keywords and then manually screen out unnecessary conjunctions such as "and", "or", "then". Add these words to the stop list and filter the keywords again to select the first 100 words with the highest frequency. Subsequent analysis will also be based on chosen keywords.

3.3 Word Cloud Map

Word cloud analysis is also the primary method of visualizing news data. Cloud images can convey information about the importance of words through

the size of terms. The word cloud image can also be distinguished by adding different colors to different words. Simultaneously, the word cloud map can intuitively show the important difference of keyword and display the key information of news text (Ponnambalam, 2019). By forming a word cloud, the author can filter out most meaningless textual information, and words with specific meanings can be retained. However, prepositions or meaningless function words still need to be cleared in advance.

The author chose to make word clouds to visualize keyword changes in 4 years. The word cloud can visually display keywords in news reports by the size and color of the text. Import the obtained keyword and word frequency data into Dychart. Using the word cloud template, simply adjust the text size and color to generate the corresponding word cloud picture. The word clouds of 2018-2021 to express ABC News reports' keywords can be got.

3.4 Text Clustering Analysis

Based on keywords and TF-IDF analysis, the news can be clustered to discuss the report's emotion. Hence, KMeans was used to classify the text. The author imported the KMeans Library in Python. In the final output, yellow represents that the article has positive emotions, and black represents that the article has negative feelings.

4 RESULTS

The researcher finds that as shown in figure 2. In July 2018, the Australian government took the lead in banning Huawei from participating in Australia's 5G network on national security grounds, and the number of news reports showed a rapid growth trend. In mid-2019, with the release of the Trump administration's Defense Authorization Act, the number of reports peaked at 96 and 94 in a month, which has been called a high concern for single media. In 2020, the Australian subsidiary of Huawei announced the termination of its ten-year sponsorship relationship with the Australian rugby team Canberra Raiders, which was also continued reported on ABC news for a while (Overton, 2020). Thus, ABC News reports on Huawei are positively related to the Australian and American governments' continued attention.

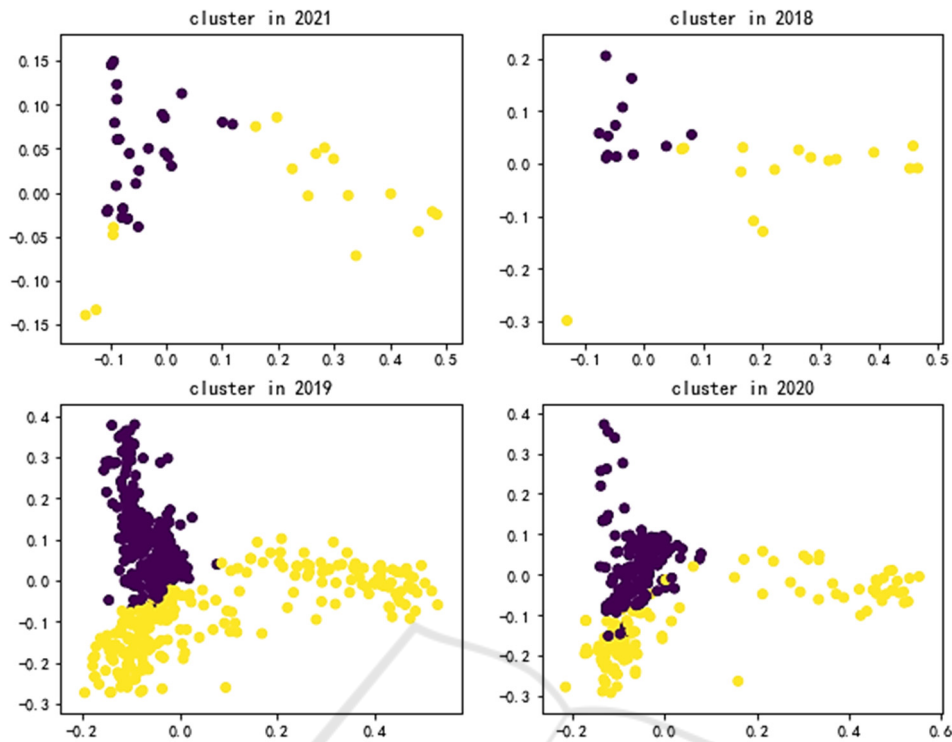


Figure 7: Cluster analysis.

of the results are shown in the figure above. Researcher find that ABC News reports on Huawei are positive most of the time. However, there will still be a lot of adverse sentiment reports in 2019 and 2020. The different changes may be caused by the Sino US trade war and the Australian government’s dialogue measures.

5 CONCLUSION

This research finds that news reports’ number and tendency are related to the international situation. There are related essential events at the peak of each report.

In the process of data analysis, there are still many areas that need to be improved. For example, although the author imported a stop word, the exported keyword still contained some meaningless words. This is because the stop words were manually filtered, and there were some omissions in the selection. In addition, the code when writing programs can be optimized. If future research can calculate the similarity between different documents, the study will become more complete.

REFERENCE

- A. Topkev. Discourse analysis and media attitudes: the representation of Islam in the British press [Review of Discourse analysis and media attitudes: the representation of Islam in the British press]. *Journal of Multilingual and Multicultural Development*, 2016, 37(2), 216–218. Routledge. <https://doi.org/10.1080/01434632.2014.973654>
- E. Kania, & Australian Strategic Policy Institute. *Huawei and Australia’s 5G Network* (pp. 04-05, Rep.). Australian Strategic Policy Institute. 2018. doi:10.2307/resrep23079.
- K. Ponnambalam, & LinkedIn.com. *Text Analytics and Predictions with Python Essential Training*. 2019.
- M. Brüggemann. Between Frame Setting and Frame Sending: How Journalists Contribute to News Frames. *Communication Theory*, 2014, 24(1), 61–82. <https://doi.org/10.1111/comt.12027>
- P. Jennings, & Australian Strategic Policy Institute. *Huawei and Australia’s 5G Network* (pp. 25-26, Rep.). Australian Strategic Policy Institute. 2018. doi: 10.2307/resrep23079.15
- P. Overton. Huawei quits Raiders: Telco giant Huawei has blamed the trade dispute between Australia and China for its decision to cancel its major sponsorship deal with the Canberra Raiders. In *Nine News*. 2020.

- T. Roelleke, & Morgan & Claypool Publishers. Information retrieval models foundations and relationships. Morgan & Claypool. 2013.
- Y. Fan. Design and Implementation of Distributed Crawler System Based on Scrapy. IOP Conference Series. Earth and Environmental Science, 2018, 108(4), 42086. <https://doi.org/10.1088/1755-1315/108/4/042086>
- Y. Cai. How does Huawei "go out". 21st Century Business Review (1), 2015, 30-31.

