Measuring Alignment to Authoritarian State Media as Framing Bias

Timothy Niven
Doublethink Lab
Taipei, Taiwan
tim@doublethinklab.org

Hung-Yu Kao
Intelligent Knowledge Management Lab
National Cheng Kung University
Tainan, Taiwan
hykao@mail.ncku.edu.tw

Abstract

We introduce what is to the best of our knowledge a new task in natural language processing: measuring alignment to authoritarian state media. We operationalize alignment in terms of sociological definitions of media bias. We take as a case study the alignment of four Taiwanese media outlets to the Chinese Communist Party state media. We present the results of an initial investigation using the frequency of words in psychologically meaningful categories. Our findings suggest that the chosen word categories correlate with framing choices. We develop a calculation method that yields reasonable results for measuring alignment, agreeing well with the known labels. We confirm that our method does capture event selection bias, but whether it captures framing bias requires further investigation.

1 Introduction

Information warfare is of increasing concern in the current era, persecuted through various channels. Recently, social media has become a key focus, however traditional media continues to play an important role. As Cook (2020) notes, the economically powerful authoritarian Chinese Communist Party (CCP) is rapidly expanding its influence over media production and distribution channels globally, and is applying tactics developed for co-opting the Chinese diaspora in different countries to local mainstream media.

The scale of CCP investment may also signify the importance they attach to the role of traditional media (Cook, 2020). There have been some apparent successes. For example, a recent study significantly correlated access to CCP state media in Africa with favourable views of the CCP-state (Bailard, 2016). On the other hand, CCP media is not considered to be nearly as successful in the West. As an authoritarian party-state, CCP-branded media suffers from a large credibility deficit. As Varral (2020) notes, at China Global Television Network (CGTN) it is generally accepted internally that CGTN's influence is not what is could be given the amount of funding they receive. Nevertheless, given the scale of investment, it is possible that CCP state media are learning from their mistakes, refining their messaging, and iteratively improving their ability to influence foreign nations.

Of equal concern are efforts to overcome the key branding issue. As detailed by Brady (2017), the CCP's media strategy has long included the policy of "borrowing a boat to go on the ocean": strategic partnerships with foreign media organizations to spread their messaging. A recent example is advertorial agreements with Western media such as The Washington Post to run a China Daily supplement in their own newspapers (albeit with a small warning label). Such agreements allow the CCP to shake off their own brand and borrow from the credibility carefully developed by such newspapers. The "borrow a boat" policy has also been augmented with the "buy a boat" policy: strategic mergers and acquisitions of foreign media assets (Brady, 2017).

Proceedings of the 3rd NLP4IF Workshop on NLP for Internet Freedom, pages 11–21
Barcelona, Spain (Online), December 12, 2020

For example, Auckland's only Chinese-language 24-hour radio station, FM 90.6, was taken over by a subsidiary of China Radio International (CRI), and now sources all its news from CRI and its Australian subsidiary (Brady, 2017). With a bought and paid for boat, audiences do not even have the luxury of a small warning label on the content. Related to this last category are media outlets that are owned by agents loyal to the CCP, either ideologically or financially motivated. A prime example is Taiwan's *China Times*, a company of the Want-Want group who are proudly aligned to the CCP.[1]

Advertorial supplements from *China Daily*, subsidiaries of CCP state media, and media outlets run by loyal agents of the CCP are examples of *alignment* to authoritarian state media, the focus of our work. We define such alignment as effectively acting as an agent of that state: spreading their messaging, and working towards their goals. Alignment in this sense can therefore pertain to agents (media outlets or journalists), and also to text (news articles).

In our work with civil society organizations and journalists around the world, we have identified the usefulness of an NLP tool to detect and track media that are aligned to the CCP. The primacy of investigative journalism is not in question. In many cases the alignment of agents can be established by investigating the people and companies involved. However, such an NLP tool could be helpful in a number of ways. It could flag new and emerging media for investigation. It could detect the point in time when a media organization undergoes change in the bias of its content (which may be coincidental, for example, with when it was purchased by another company, flagging the company for investigation). It could be used to compare CCP influence across countries and time. It could hopefully also give insight into how the alignment is expressed - e.g., what issues and framings are important - which may help to understand the evolving nature of CCP information operations. Finally, it would ideally be able to provide hard statistical evidence to support claims of bias.

From an NLP perspective, this research problem is novel (as far as we are aware). Yet by operationalizing alignment in terms of media bias we are able to draw upon a wealth of prior research. Our conceptualization of media bias follows the sociological model reported in Hamborg (2018). The types of bias we intend to address concern what is reported, and how it is reported. What is (and is not) reported is covered by *event selection bias*. How it is reported is covered by *word choice*, *labeling*, and *framing bias*. In this work we aim to measure framing bias, and find that our methodology also captures event selection bias. We ideally want an unsupervised solution that is applicable to as many languages (and therefore national contexts) as possible.

In the abstract, our methodology is to treat the CCP state media data as a "propaganda distribution." As news outlets in other countries will follow their own news cycle and political context, we need a way to factor out this part of their observed distribution. We then check the remainder for alignment in the direction of the CCP propaganda distribution.

In the present work, we concretely implement this methodology by using the frequencies of words in psychologically meaningful lexical categories drawn from the Linguistic Inquiry and Word Count (LIWC) dictionary (Pennebaker et al., 2015). We identify 15 categories that appear to relate to framing. For each chosen category, we calculate an indicator variable representing whether a media outlet differs from the Taiwanese mean in the direction of the CCP, and take the average of these indicators as a score. Our results are a strong match for the expected labels, and also capture the ideological leaning of the four Taiwanese media outlets studied. We therefore have reasonable confidence that our methodology captures the known bias in the data. However, an initial qualitative study is inconclusive regarding whether we have captured framing bias. We do, however, confirm that our methodology unexpectedly captures event selection bias. Therefore, a careful analysis of framing in our dataset should be conducted to confirm the nature of the bias our method is capturing.

[1]For example, consider this editorial in support of the annexation of Taiwan by the CCP-state: https://www.chinatimes.com/newspapers/20200524000371-260108?chdtv.

2 Related Work

Gentzkow and Shapiro (2006) offer one of the earliest computational methods for measuring framing bias in the U.S. national media context. They use the χ^2 statistic to compare the frequency of phrases used by Republicans and Democrats, selecting the top-k phrases for each as marking the bias distributions. We implemented this method on our data, but found it captured more differences in national contexts than ideological differences (e.g. "report from Beijing"). We therefore hypothesize that this methodology is not suitable for international bias measurements unless augmented by some other technique for factoring out national differences.

Field (2018) develop a multilingual method for detecting frames based on the coding in the Media Frames Corpus (Card et al., 2016). Their methodology could ostensibly be used to study international framing bias. They collected frame-specific lexicons in English using the PMI of words with the annotated frames, then translated the lexicons into the target language, using word embeddings to refine the lexicon through query expansion so as to more accurately capture the local linguistic and political context. Our experiments with this methodology did not prove as successful as our reported methodology, yielding more variable and noisy results. It is possible that the lexicon construction process is too noisy for our purpose. We also suspect that using a pre-defined set of high-level frames is not sensitive enough. It is possible that our method of using word categories that relate to framing as opposed to the high-level frames themselves is more flexible, allowing unannotated framings to be expressed in the relative frequencies of lower-level word categories.

Shahid et al. (2020) follow an approach most similar to ours, using the Moral Foundations dictionary (Graham et al., 2013) to analyze framing bias in an unsupervised manner. Their results appear reasonable, and are supported by other studies that suggest the usefulness of moral foundations in understanding framing bias (Fulgoni et al., 2016). One advantage of LIWC is that, to the best of our knowledge, it has been translated into more languages than the Moral Foundations dictionary. Our experience is that lexicon quality is important. Nevertheless, comparing our method to a Chinese version of the Moral Foundations dictionary is necessary future work.

More recent work has considered multi-label and multi-lingual frame detection in a few-shot setting with minimal labeled data (Akyürek et al., 2020). However, their data comes solely from articles about gun violence in the U.S. and also uses a pre-defined set of framings.

A fairly comprehensive plan for research into automated framing bias detection is offered by Hamborg (2020). However, this methodology still involves upstream inaccuracies and components of supervised learning - e.g., aligning entity mentions through coreference resolution. Although we support this research direction as well principled and promising, our initial investigation aims to be fully unsupervised for two reasons. The first is to avoid propagation of upstream errors as much as possible.[2] Second, due to our desire to work with as many languages (and therefore national contexts) as possible, we want to reduce our reliance on language-specific resources such as supervised coreference resolution classifiers.

Cremisini et al. (2019) released a dataset for analyzing bias in coverage of the Ukranian crisis including a range of media outlets from 43 countries. Including Russian state media also makes this work closer to our task of measuring alignment to authoritarian state media, although it only deals with a single issue. Application of our methodology to this dataset remains as future work.

The present work complements our existing work on CCP information operations (Tseng and Shen, 2020). The development of our tool could assist this research by identifying local agents (e.g. journalists, influencers) working on behalf of the CCP, which have been found to be an important component of CCP information warfare directed against Taiwan.

[2] As Chinese text does not separate words with a space, the first step in any NLP pipeline is deciding where to draw word boundaries. There is no perfect solution available, so a degree of upstream noise is inevitable, but does not appear crucial for our methodology.

Media Outlet	Label	Num. Documents	Num. Tokens
China Internet Information Center	CCP state Media	65,804	45,203,219
China Daily	CCP state Media	6,190	4,877,458
China Development Gateway	CCP state Media	544	487,252
China Times (CCP)	CCP state Media	543	632,425
Global Times	CCP state Media	24,585	15,263,884
National Party Media Information Public Platform	CCP state Media	39,556	23,431,048
Xinhua	CCP state Media	806	751,059
People's Daily	CCP state Media	102,158	64,067,371
Qiushi	CCP state Media	3,210	3,063,152
China Times (Taiwan)	Aligned to CCP	17,143	5,790,180
United Daily	Bias Towards CCP	17,275	6,137,210
Liberty Times	Bias Against CCP	6,704	2,015,226
Apple Daily	Oppose CCP	6,076	1,864,850

Table 1: The composition of our dataset. The label for Taiwan's *China Times* is well established, making it a good test case for our method. *United Daily* can be roughly considered ideologically closer to the CCP, but independent. *Liberty Times* can be considered ideologically opposed to the CCP. *Apple Daily* is a staunch opponent of the CCP. Note that the data we collected is filtered by a large list of keywords relating to politics.

3 Data

We scraped news articles from the media outlets listed in Table 1 from January 1st, 2020, to May 1st, 2020. The scraping process filtered for content related to politics using a large manually curated list of keywords. The outlets labeled CCP state media provide the "propaganda distribution" for bias calculations. We consider the articles from all CCP state media domains as a single distribution.[3] The Taiwanese media fall into three categories. First, we have Taiwan's *China Times* which is known to be working on behalf of the CCP. The three other domains provide coverage of the rest of the ideological spectrum: *United Daily* is ideologically closer to the CCP yet independent; *Liberty Times* is ideologically opposed to the CCP; *Apple Daily* is a staunch opponent.

For all media domains, we manually clean the scraped text to remove common template tokens, such as journalist credits and advertising. Word segmentation is then performed with jieba[4] augmented with the custom dictionary provided by Media Cloud (Chuang et al., 2014).[5]

4 Methodology

In this section we first motivate our use of LIWC categories, and our method of selecting which categories reliably relate to framing. We then outline our calculation method and procedure for statistical inference.

4.1 LIWC Categories and Framing

The psychologically meaningful word categories we use come from LIWC dictionary (Pennebaker et al., 2015). The LIWC dictionary has been carefully developed over a decade and a half and has been used in many psycholinguistic studies. We use the 2015 versions of the simplified and traditional Chinese LIWC dictionaries (Huang et al., 2012). Even though China and Taiwan share a common language, there are many important linguistic differences. Fortunately, these dictionaries were developed appropriately for their respective linguistic contexts, and represent more than a naive mapping between traditional and simplified characters.[6] Altogether there are 79 word categories from which we select 15 that appear to relate to framing.

[3]Future work is to investigate each outlet separately and pick apart what parts of the propaganda distribution they are responsible for.

[4]https://github.com/fxsjy/jieba

[5]https://github.com/berkmancenter/mediacloud

[6]The mean Jaccard similarity of the 79 word category sets across the two dictionaries is 0.88, with 0.15 standard deviation.

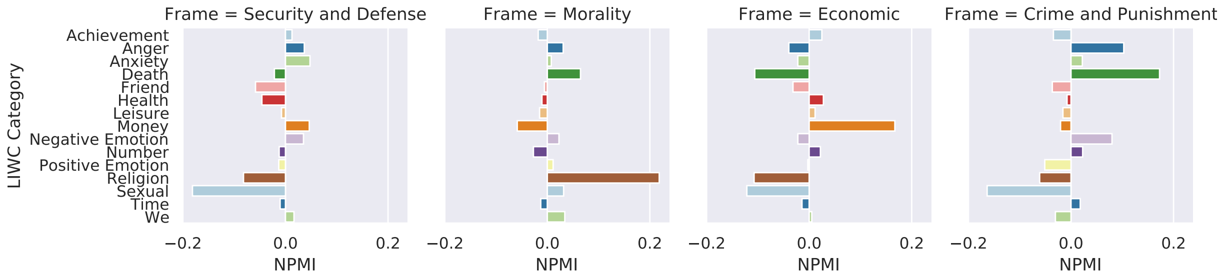

Figure 1: The NPMI of the 15 LIWC categories we selected with the annotated framings in the Media Frames corpus. These relationships appear reasonable and underwrite our assumption that the frequencies of using these categories in news articles express framing choices.

Similar to the methodology of Field et al. (2018), we consider the normalized Pointwise Mutual Information (NPMI) of LIWC categories with the annotated frames of the Media Frames corpus (Card et al., 2016). In order to account for document length, we consider the event space in terms of the selection of individual words. For each document, we consider all its words to be associated with all frames ascribed by the annotators. Let x_f be the event that a word appears with a frame, f. Let y_c be the event that the selected word comes from an LIWC category, c. The NPMI of c with f is then calculated in the standard way as

$$\mathrm{NPMI}(c, f) = \frac{\log(p(y_c|x_f)/p(x_f))}{-\log_2(p(x_f, y_c))}, \tag{1}$$

where the denominator is self-information.

The LIWC dictionary involves some 79 word categories. Experiments with every word category did not yield meaningful results. It was therefore reasonable to hypothesize that only a subset of the LIWC categories are reliably associated with framing. Observing the NPMI values, we decided to choose a 0.01 absolute NPMI cutoff value with any framing, resulting in 15 categories. Figure 1 shows the NPMI of the chosen categories with a subset of frames. These results are inherently reasonable, which motivates their use in framing analysis. Some findings of note: the “Religion” category almost singularly defines the “Morality” frame in the same way that “Money” defines the “Economic” frame. It is expected that “Anxiety” and “Negative Emotion” are well correlated with “Security and Defense,” and that “Death” and “Anger” are most associated with “Crime and Punishment.” Given these results, we view this subset of LIWC categories as correlates of framing. We hypothesize that the stable and meaningful results we achieve, as outlined below, in part come from the flexibility of modeling framing with respect to these more general word categories. We now turn to describe our calculation method for framing bias.

4.2 Calculation Method

For a specific country, our methodology scores how often a media outlet deviates from the national mean of that country in the direction of the CCP. Formally, the CCP distribution is defined by chosen media outlets, $m^{(i)}_{\text{ccp}}$, in the set, $\mathcal{M}_{\text{ccp}}$, where $i = 1, 2, \ldots, |\mathcal{M}_{\text{ccp}}|$. These are to be compared to the national media outlets, $m^{(j)}_{\text{national}}$, in the set $\mathcal{M}_{\text{national}}$, where $j = 1, 2, \ldots, |\mathcal{M}_{\text{national}}|$. The media outlets under consideration are listed in Table 1.

We first calculate the mean frequencies of each LIWC category $c \in \mathcal{C}$ per document in the corpus collected for the CCP media outlets. Let $\mathcal{D}_{\text{ccp}}$ be the collection documents from all media outlets in $\mathcal{M}_{\text{ccp}}$. Let $F^{(c)}(d)$ be a function that counts the proportion of words from category c in document d. Then the CCP's mean is:

$$\mu^{(c)}_{\text{ccp}} = \frac{1}{|\mathcal{D}_{\text{ccp}}|} \sum_{d \in \mathcal{D}_{\text{ccp}}} F^{(c)}(d). \tag{2}$$

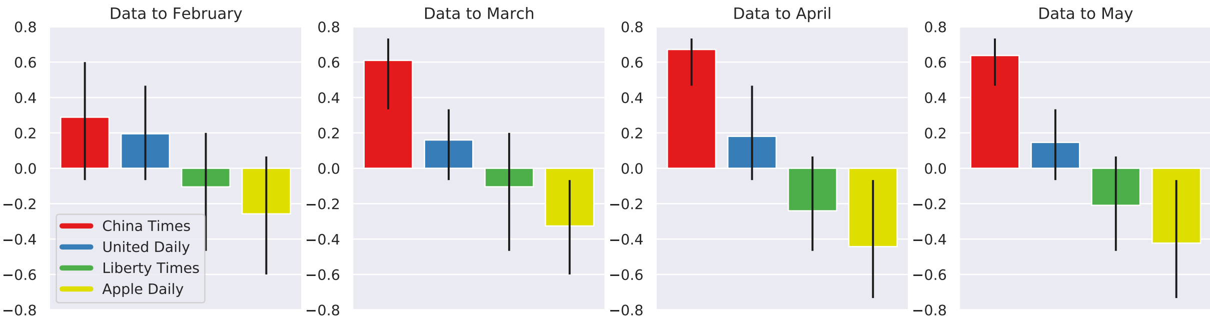

Figure 2: Results of our scoring calculation (Equation 5). Error bars represent 95% confidence intervals for the score given the bootstrap distributions. The data starts from January 1st, 2020, and continues to the first of the indicated month. After one month the results are uncertain, with all error bars crossing zero. By month two the results have already settled down into the qualitative pattern we expect, with *China Times* singled out as aligned to the CCP state media. The results for the other three media outlets also qualitatively align with our expectations given the known ideologies of each.

We calculate the means for the national context similarly, but remove the media outlet under consideration from the set in order to stop outliers from having an oversized impact on the mean (as it is outliers we want to detect, this increases the sensitivity of our method). Let $\mathcal{D}_{national/m}$ be the set of documents for all $m \in \mathcal{M}_{\text{national}/m}$. Then the national mean is:

$$\mu^{(c)}_{\text{national}/m} = \frac{1}{|\mathcal{D}_{\text{national}/m}|} \sum_{d \in \mathcal{D}_{\text{national}/m}} F^{(c)}(d). \tag{3}$$

The final ingredient in our score is the category means for the media outlet under investigation, m. Let $\mathcal{D}_m$ be the set of documents for m. Then the mean is:

$$\mu^{(c)}_{m} = \frac{1}{|\mathcal{D}_{m}|} \sum_{d \in \mathcal{D}_{m}} F^{(c)}(d). \tag{4}$$

Our final bias score for m, s_m, is then the average number of categories for which the difference from the national mean is in the direction of the CCP:

$$s_m = \frac{1}{|\mathcal{C}|} \sum_{c \in \mathcal{C}} \mathbb{1}\left[\text{sign}(\mu^{(c)}_{\text{national}/m} - \mu^{(c)}_{m}) = \text{sign}(\mu^{(c)}_{\text{national}/m} - \mu^{(c)}_{\text{ccp}})\right] \tag{5}$$

4.3 Statistical Inference

We favour non-parametric statistical inference through bootstrapping (Efron and Tibshirani, 1986). Specifically, as our data involves a time series, we use the tapered block bootstrap (Kunsch, 1989) in order to preserve the dynamics of the news cycle through time. We use the python implementation provided by the recombinator package.[7]

5 Results

For each domain to be scored, we perform 1,000 bootstrap samplings of the whole dataset, and report the distribution of results. The 95% confidence interval is calculated in terms of percentiles of the bootstrap distribution. We are interested to know how stable the results are with varying amounts of data. We therefore repeat the bootstrap scoring procedure with varying time periods. All scores start at the January 1st, 2020, and extend an integer number

[7] https://github.com/InvestmentSystems/recombinator

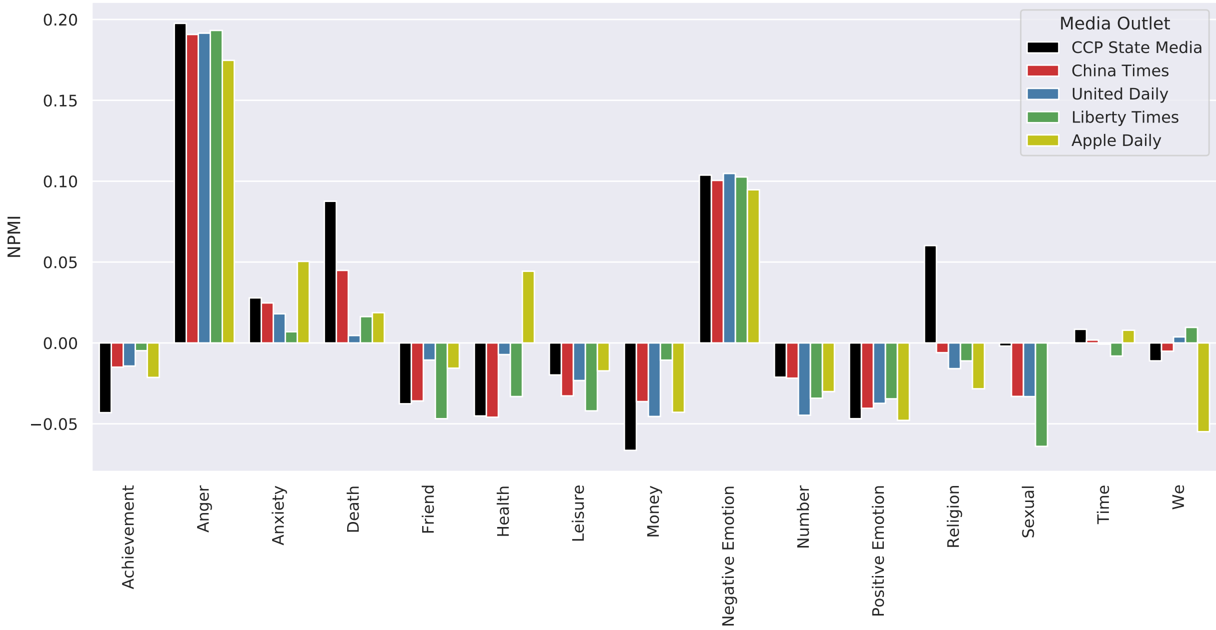

Figure 3: NPMI of LIWC categories with sentences including the word "protest." It stands out that the CCP prefers the *Death* and *Religion* categories. Randomly sampling 20 sentences with "protest" and these categories reveals that when the CCP report on protest in our dataset, they discuss murder, crime, and damage to property. However, they also talk about protests in the US, suggesting that this statistical signal is also picking up event selection bias.

of months, up to May 1st, 2020. Figure 2 shows the results. With two months of data the expected qualitative pattern is observed: *China Times* is strongly aligned to the CCP, the lower bound of the confidence interval well above zero; *United Daily* is independent but ideologically closer; *Liberty Times* is ideologically divergent; *Apple Daily* is strongly divergent from the CCP. The pattern appears to be stabilizing over time.

Limitations: Although the results appear to be stabilizing, we believe this length of time is still short enough to leave uncertainty about the reliability of our methodology. The results should also be reproduced in different national and linguistic contexts for more thorough verification.

6 Analysis

Our results suggest we have measured bias in our dataset in accordance with the expected labels. However, our case that we have captured framing bias rests on the findings of Section 4.1: that the LIWC categories we chose correlate well with the annotations in the Media Frames corpus. This case needs to be strengthened by a careful look at frames in our data. We conducted a cursory analysis that is suggestive yet inconclusive about framing bias. However, it does confirm that our methodology captures event selection bias.

We identified sentences that include at least one instance of the word "protest." We then calculated the NPMI of LIWC categories with these sentences following equation (1), however at the sentence not document level.[8] The distribution over LIWC categories is given in Figure 3. We observe that the CCP prefers the categories "Death" and "Religion" compared to Taiwanese media (with *China Times* biasing in their direction). Given how these categories correlate with frames as observed in Figure 1, these would appear to indicate the CCP frames protest in terms of "Crime and Punishment" and "Morality" in our dataset. To check this hypothesis, we

[8]We found the document level did not yield reasonable results, which is to be expected, as a document can discuss a number of topics and concepts. Furthermore, some CCP state media have the habit of reporting a number of different news items in a single document.

Media Outlet	p(America)	p(Protest)	p(America|Protest)
CCP state Media	0.094	0.007	0.410
China Times (Taiwan)	0.091	0.018	0.140
United Daily	0.106	0.025	0.048
Liberty Times	0.069	0.019	0.082
Apple Daily	0.098	0.043	0.036

Table 2: Statistics of coverage of "America" and "Protest" in our dataset. The CCP is far less likely to discuss protest, but when they do, they talk about protests in the U.S. almost half the time. The *China Times* strongly biases in this direction relative to the Taiwanese mean. If a media outlet follows CCP frames, we would expect our methodology to also indirectly capture event selection bias.

randomly sampled 20 sentences containing "protest" for inspection. We observed that most of the sentences are indeed about murder, death, and criminal behavior, and some highlight the negative impacts of protest on society. However, to fully address this question a more principled analysis of framing conducted by political scientists is required.

We also observed that the majority of these sentences were about protests in the United States after the murder of George Floyd. We therefore hypothesized that we are also measuring event selection bias. Table 2 shows the statistics of covering "America" and "protest" in our dataset, calculated as any document with at least one sentence mentioning either, or both. We observe that, while the CCP is far less likely to discuss protest than the Taiwanese media, when they do, it is nearly half the time about protest in the United States. *China Times* strongly biases in this direction, relative to the Taiwanese distribution. It is reasonable to suggest that, if a media outlet follows CCP framing on a specific issue, detection of similar framing over a period of time may also indirectly capture event selection bias. We leave a more thorough investigation of this question to future work.

7 Conclusion

We are surprised by how well our calculation method matches the known bias labels. Lexical category approaches have some well known difficulties, particularly that they do not consider syntax - e.g., a negative emotion word following a negation may not express negative emotion. Furthermore, it may be that following CCP framing does not necessarily represent the CCP's interests in different national contexts. We observe an example of this: the *China Times* uses "we" in sentences containing "China" far more than any other Taiwanese newspaper - but the CCP state media does not. Despite these complications, with enough data our methodology appears to work qualitatively well and yield stable results. Drilling down to consider framing around specific issues also appears to yield qualitatively reasonable insights consistent with our expectations. Our hypothesis is therefore that the aforementioned complications are effectively smoothed over by enough data.

Our assumption that the LIWC categories capture framing, and therefore that our methodology captures framing bias, is based on the observation of meaningful correlations of these categories with the annotations in the Media Frames corpus. However, we also saw in our qualitative analysis that we are picking up event selection bias. A closer examination will be necessary to establish the extent to which our calculation method is picking up framing and event selection bias. It will be necessary to observe data over a longer period of time in order to validate the observed stability of our conclusions. We must also be apply it to different linguistic and national contexts, over many more media organizations. We should also compare our methodology to another apparent correlate of framing: the Moral Foundations Dictionary. The LIWC resource is also a constraint on the number of languages we can handle. Currently there are 15 languages covered by LIWC. Therefore, future work is to find a way to cover more languages, ideally in an unsupervised manner that does not rely on such resources.

Our calculation method is very simple, and can almost certainly be improved. For example, we have found that using vectors of our chosen LIWC category frequencies and taking the first principal component also yields a surprisingly coherent ideological ordering of media, including the different CCP state media outlets. Vector-based calculation methods are likely to be more sensitive to the statistics than our binary indicator variables, and may also facilitate more complicated calculations, such as comparing to multiple, different propaganda distributions. It also makes more sense to consider frames as occupying regions in the 15-dimensional space of frequencies induced by our LIWC category selections (see Figure 1), as opposed to considering the categories separately.

How might our methodology be applied to other authoritarian states? Our results suggest that collecting two months of data from the authoritarian state media, and from media outlets of interest with unknown labels, should be enough to apply our scoring calculation. It will be interesting to explore other potential applications of our methodology, such as detection of content farms (Tseng and Shen, 2020), or networks of Twitter accounts operated by the CCP. We make our code publicly available to facilitate subsequent research.[9]

In the introduction we outlined a number of goals for the NLP tool we have set out to build. How does our methodology help to those ends? With two months of data we could tag an emerging or unknown media outlet as potentially part of CCP information operations. We have not modeled the time dimension so far, so we could not detect when a media outlet undergoes a change in alignment. We could compare the media outlets within and across countries, but not across time. Drilling down around specific concepts, events, or entities does facilitate analysis of how alignment is expressed, as our analysis of protest suggests. Our calculation method can be considered a form of statistical evidence, however should be further validated as outlined above.

We note an important weakness in our methodology. Taking the difference from the average national media will fail where the national media is almost entirely aligned to the CCP. For example, it is estimated that in Australia 95% of the Chinese language media is under CCP influence (Munro and Wen, 2016). A similar situation exists in New Zealand (Brady, 2017). In such contexts our methodology would require modification. It would be most desirable to seek a more absolute measure of alignment, but it may be difficult to develop one that is automatic and works in every national political context.

For future work, we intend to explore the use of word embeddings with techniques developed in the gender bias literature to try and identify directions in the CCP state media embedding space that capture the propaganda distribution. Initial experiments are promising. For example, the distance between "Taiwan" and "country" is much larger in the CCP embedding space then in the Taiwan (minus *China Times*) embedding space, with *China Times* falling quite far outside the Taiwanese distribution in the direction of the CCP.[10] It may be possible to build a list from official CCP framings of particular issues such as Taiwan (Brady, 2015). However, such a static approach would be less ideal than one that could be automatically learned and updated daily. We leave these challenges for future work.

Acknowledgements

We would like to thank the reviewers for their highly useful feedback.

References

Afra Feyza Akyürek, Lei Guo, Randa Elanwar, Prakash Ishwar, Margrit Betke, and Derry Tanti Wijaya. 2020. Multi-label and multilingual news framing analysis. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8614–8624, Online, July. Association for Computational Linguistics.

[9]https://github.com/doublethinklab/nlp4if2020p

[10]Comparing distances in different embedding spaces is know to be problematic (Gonen et al., 2020), however we offer this result as suggestive.

Catie Snow Bailard. 2016. China in africa: An analysis of the effect of chinese media expansion on african public opinion. *The International Journal of Press/Politics*, 21(4):446–471.

Anne-Marie Brady. 2015. Unifying the ancestral land: The ccp's "taiwan" frames. *The China Quarterly*, 223:787–806.

Anne-Marie Brady. 2017. Magic weapons: China's political influence activities under xi jinping.

Dallas Card, Justin Gross, Amber Boydstun, and Noah A. Smith. 2016. Analyzing framing through the casts of characters in the news. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1410–1420, Austin, Texas, November. Association for Computational Linguistics.

Jason Chuang, Sands Fish, David Larochelle, William P. Li, and Rebecca Weiss. 2014. Large-scale topical analysis of multiple online news sources with media cloud.

Sarah Cook. 2020. Beijing's global megaphone: The expansion of chinese communist party media influence since 2017. Retrieved August 26 2020 from https://freedomhouse.org/report/special-report/2020/beijings-global-megaphone.

Andres Cremisini, Daniela Aguilar, and Mark A. Finlayson. 2019. A challenging dataset for bias detection: The case of the crisis in the ukraine. In Robert Thomson, Halil Bisgin, Christopher Dancy, and Ayaz Hyder, editors, *Social, Cultural, and Behavioral Modeling*, pages 173–183, Cham. Springer International Publishing.

Bradley Efron and Robert Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statist. Sci.*, 1(1):54–75, 02.

Anjalie Field, Doron Kliger, Shuly Wintner, Jennifer Pan, Dan Jurafsky, and Yulia Tsvetkov. 2018. Framing and agenda-setting in russian news: a computational analysis of intricate political strategies. *CoRR*, abs/1808.09386.

Dean Fulgoni, Jordan Carpenter, Lyle Ungar, and Daniel Preoţiuc-Pietro. 2016. An empirical exploration of moral foundations theory in partisan news sources. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 3730–3736, Portorož, Slovenia, May. European Language Resources Association (ELRA).

Matthew Gentzkow and Jesse M Shapiro. 2006. What drives media slant? evidence from u.s. daily newspapers. Working Paper 12707, National Bureau of Economic Research, November.

Hila Gonen, Ganesh Jawahar, Djamé Seddah, and Yoav Goldberg. 2020. Simple, interpretable and stable method for detecting words with usage change across corpora. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 538–555, Online, July. Association for Computational Linguistics.

Jesse Graham, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P. Wojcik, and Peter H. Ditto. 2013. Moral foundations theory. In *Advances in Experimental Social Psychology*, pages 55–130. Elsevier.

Felix Hamborg, Karsten Donnay, and Bela Gipp. 2018. Automated identification of media bias in news articles: an interdisciplinary literature review. *International Journal on Digital Libraries*, pages 1–25.

Felix Hamborg. 2020. Media bias, the social sciences, and NLP: Automating frame analyses to identify bias by word choice and labeling. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*, pages 79–87, Online, July. Association for Computational Linguistics.

Chin-Lan Huang, Cindy K Chung, N. Hui, Yi-Cheng Lin, Yi-Tai Seih, Wei-Chuan Chen, Ben C. P. Lam, Michael H. Bond, and J. W. Pennebaker. 2012. The development of the chinese linguistic inquiry and word count dictionary. *Chinese Journal of Psychology*, 54:185–201.

Hans R. Kunsch. 1989. The jackknife and the bootstrap for general stationary observations. *Ann. Statist.*, 17(3):1217–1241, 09.

Kelsey Munro and Phillip Wen. 2016. Chinese language newspapers in australia: Beijing controls messaging, propaganda in press. Retrieved August 31 2020 from https://www.smh.com.au/national/chinese-language-newspapers-in-australia-beijing-controlsmessaging-propaganda-in-press-20160610-gpg0s3.html.

James Pennebaker, Ryan Boyd, Kayla Jordan, and Kate Blackburn. 2015. The development and psychometric properties of liwc2015. 09.

Usman Shahid, Barbara Di Eugenio, Andrew Rojecki, and Elena Zheleva. 2020. Detecting and understanding moral biases in news. In *Proceedings of the First Joint Workshop on Narrative Understanding, Storylines, and Events*, pages 120–125, Online, July. Association for Computational Linguistics.

Poyu Tseng and Puma Shen. 2020. The chinese infodemic in taiwan. Retrieved October 11 2020 from https://medium.com/doublethinklab/the-chinese-infodemic-in-taiwan-25e9ac3d941e.

Merriden Varrall. 2020. Behind the news: Inside china global television network. Retrieved August 26 2020 from https://www.lowyinstitute.org/publications/behind-news-inside-china-global-television-network.