# Measuring Media Partisanship during Election: The Case of 2019 Indonesia Election

Maulana, Ardian and Situngkir, Hokky

Dept. of Computational Sociology, Bandung Fe Institute, Dept. of Computational Sociology, Bandung Fe Institute

1 May 2020

# Measuring Media Partisanship during Election: The Case of 2019 Indonesia Election

Ardian Maulana

[ai@compsoc.bandungfe.net]
Dept. Computational Sociology
Bandung Fe Institute

Hokky Situngkir

[hs@compsoc.bandungfe.net]
Dept. Computational Sociology
Bandung Fe Institute

**Abstract**

Analysis of media partisanship during election requires an objective measurement of political bias that frames the content of information conveyed to the audience. In this study we propose a method for political stance detection of online news outlets based on the behavior of their audience in social media. The method consists of 3 processing stages, namely hashtag-based user labeling, network-based user labeling and media classification. We applied this methodology to the tweet dataset related to the 2019 Indonesian general election, to observed media alignments during the election. Evaluation results show that the proposed method is very effective in detecting the political affiliation of twitter users as well as predicting the political stance of news media. Over all, the stance of media in the spectrum of political valence confirms the general allegations of media partisanship during 2019 Indonesian election. Further elaboration regarding news consumption behavior shows that low-credibility news outlets tend to have extreme political positions, while partisan readers tend not to question the credibility of the news sources they share.

**Keywords**: news media network, label propagation algorithm, twitter, election, media partisanship

## 1. Introduction

The rapid development of online media and social media in recent years has radically changed the way people consumes information. Survey shows that 63% of people read news digitally [1], while social networks, such as Twitter and Facebook, are places where people share and discuss the latest news. The combination of online news media and social media strengthens the role of news outlets as gatekeepers of information in relation to the formation of public opinion [2], [3].

Neutrality of news media is difficult to maintain at the time of the election. This has increasingly become a public concern given the ability of news media to influence individual choices, which can have an impact on the final outcome of an election. Scientific efforts to examine the partisans' behavior of news outlets during election are constrained by the lack of data about ideological stance of news media [4]–[7]. In fact, majority of news outlets do not openly express their political positions on various issues [5]. Generally, media alignment is reflected through content, terminology, and arguments used in framing reported issues. However, it is difficult to objectively measure the political biases contained in the media frame. The alternative approach is to infer the stance of media based on the political affiliation of their audiences. This approach is based on the assumption that people naturally tend to to interact with information adhering to their preferred narrative [6]–[8].

Social networks like Twitter are very rich in information related to user behavior, e.g. tweet content, who is followed, hashtags used. These information then can be used to identify users' political affiliations, as well as political leaning of news outlets they read. In this study we propose a method for political stance detection of online news outlets based on the behavior of their audience in social media. The method consists of 3 processing stages, as follow: (i) Hashtag-based user labeling: we use a number of political hashtags, i.e. hashtags that are strongly associated with certain political groups, to infer political affiliations of users of these hashtags; (ii) Network-based user labeling: we expand the number of tagged users using Label Propagation Algorithm; (iii) Media classification: at this stage we use polarity rule to identify the political stance of news outlets based on the political affiliation of their audiences.

We applied this methodology to the tweet dataset related to the 2019 Indonesian general election, to observed media alignments during the election. In doing so, we also report news consumption patterns on Twitter in relation to the credibility and partisan behavior of news sources. This paper is structured as follows: sections two and three discuss data and analysis methods, the results of the analysis will be shown in section 4, while the final section will discuss a number of conclusions and contributions of this study.

## 2. Data

We conducted the data collection process from 27 March to 19 May 2019, which covered the campaign period, general elections (April 17, 2019), vote recapitulation and the announcement period (May 21, 2019). Table 1 shows the descriptive statistics of the data used in this study. Tweet data was extracted from Twitter using the DMI-Tcat application [9] based on a number of keywords related to the candidates, namely: (i) Candidate I (Prabowo - Sandiaga Uno): prabowo,sandiaga uno ; (ii) Candidate II (Joko Widodo - KH. Maaruf Amin): joko widodo, jokowi, maaruf amin, kiyai maaruf.

Table 1 Descriptive statistics of the dataset

| Statistics | Count |
|---|---|
| # of tweets | 13990975 |
| # of tweets with a Url | 667821 |
| # of hashtags | 74515 |
| # of unique users | 3958817 |

## 3. Method for political stance detection of online news outlets

The process of political stance detection consists of 3 stages: (i) hashtag-based user labeling; (ii) network-based user labeling); and (iii) media classification.

### 3.1 *Hashtag-based user labeling*

In order to analyze the political stance of news outlets we first find the stances of Twitter users. Twitter users usually use political hashtags in their tweets to express their support for the political message contained in the hashtag. Nowdays, political hashtags are a strategy commonly used to mobilize opinions, popularize candidates, or attack opponents. In this study we use this simple fact to identify the political affiliations of Twitter users. Figure 1 shows a histogram of the 10 most widely used political hashtags in the 2019 Indonesian elections.

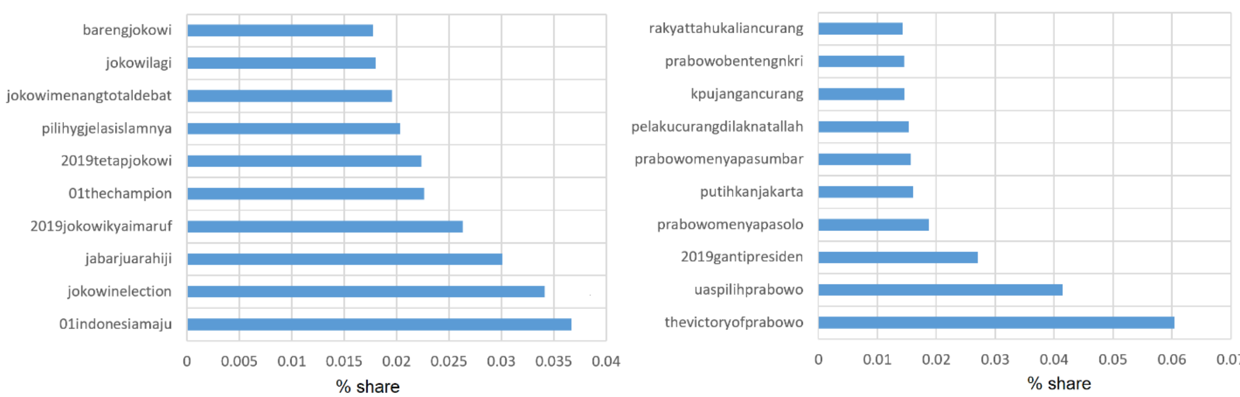


Figure 1: 10 most used political hashtags used by candidate : (a) Joko Widodo; (b) Prabowo

We label manually 1400 hashtag, which is 700 for each candidate, of the total 74515 hashtags recorded in the dataset. These hashtag has been used by at least 10 different twitter users. We apply polarity rule to infer the political affiliation of the users, as follow [10]:

$$V(u) = 2\frac{\frac{tf(u,C_O)}{total(C_O)}}{\frac{tf(u,C_O)}{total(C_O)}+\frac{tf(u,C_1)}{total(C_1)}} - 1 \qquad (1)$$

where $tf(u, C_o)$ is the number of times (user frequency) user $u$ use group of hashtag $C_o$ of candidate $i$, $total(C_o)$ is sum of frequency of hashtag usage by all users. Hashtag of other candidate are defined in similar fashion. The political valence value $V(u)$is in the range $-1 \le V(u) \le 1$. To ensure the user's political affiliation, we use a threshold value of ± 0.2, where users are lean to Prabowo if they have a valence score < 0.2, while lean to Joko Widodo if the valence score > 0.2. Table 2 shows that at this stage we are able to identify the political affiliation of 181,145 Twitter users, of which 89,000 are Jokowi's supporters and 92,145 are Prabowo's supporters.

Table 2 Identification of users' political affiliations using hashtags and network-based labeling

| Label | *Hashtag* | *Network* |
|---|---|---|
| Label = 0 (pro-jokowi) | 89000 | 176109 |
| Label = 1 (pro-prabowo) | 92145 | 366134 |
| Total | 181145 | 542243 |

### *3.2 Network based user labelling*

A central assumption in this stage is that if a user retweets a tweet, they most likely agree or erdosed message contained in that tweet. Some empirical studies [7], [11], [12] showed that content consumption in social media is dominated by selective exposure (i.e., the tendency of users to ignore dissenting information and to interact with information adhering to their preferred narrative). This is means that individuals tend to selectively interact, that is, only with other individuals who share their political understanding. In this stage, we first construct undirected weighted retweet graph to represent interaction between Twitter users, where vertices represent users and directed relationships between vertices are formed if one user retweetes another user's posts. Table 3 shows descriptive statistics of this network, where the density value indicates that this network is a sparse network, where the largest component consisting of 542,243 nodes.

Table 3 Descriptive statistics of retweet network

| Statistics | Retweet network | *giant component* |
|---|---|---|
| # of nodes | 558801 | 542243 |
| # of edges | 4372893 | 4372706 |
| Density | | |
| Average degree | 15.651 | 16.1282 |
| # of component | 16397 | 1 |

Then, we apply the label propagation algorithm to identify the labels of each node in the network. This algorithm works iteratively to renew the label of each node based on the majority label of its neighbor node. This process is carried out until the labels of the majority of nodes no longer change [13], [14]. At this stage we use 153,990 labeled node identified in previous stage as seeds (nodes with fixed labels). The label propagation algorithm successfully identified the political afiliation of 388,253 users in the retweet network.

Table 4 Precision & recall score of label propagation algorithm

| Precision | Recall | Accuracy |
|---|---|---|
| 0.984853 | 0.985031 | 0.984955 |

The k-stratified cross (5-fold) validation model is implemented to validate the classification results [14]. We use 4/5 of the total node seeds as training data, while the remaining node is used to evaluate the algorithm performance. The evaluation results in table 4 show a prediction accuracy of 98%. This strengthens confidence in the performance of the classification algorithm that we use.

### *3.3 Media classification*

The political afiliation of twitter users obtained in the previous stage is then used to predict the political stance of news media during election. We determine the stance of a media based on the average political affiliation of Twitter users who are that media audiences (see eq. 1) [10]. In other words, the political alignments of a news outlet are measured by the extent to which these outlet become sources of information for one political group. The greater the audience share of an outlets is come from a particular political group, the stronger the association between the two. As such, the score of media alignments capture differences in the type of content, which covers topics and news frames, shared by partisan users.

We split the alignment scores into 5 equal size bands, as follows [8]:

$$S(v) = \begin{cases} -2 & if\ v \in [-1, -0.6] \\ -1 & if\ v \in [-0.6, -0.2] \\ 0 & if\ v \in [-0.2, 0.2] \\ 1 & if\ v \in [0.2, 0.6] \\ 2 & if\ v \in [0.6, 1] \end{cases} \qquad (2)$$

where $S(v) < 0$ means that the media tends to lean to Prabowo, $S(v) > 0$ means the media tends to lean Joko Widodo, and S (v) = 0 means that the media tends to be neutral or moderate news media.

## 4. Analysis

Figure 1a shows the daily number of articles and unique articles shared by users during the data collection period. The statistics of unique article become a proxy for the volume of articles published by the media outlets. In general, the two indicators do not show different dynamics. This indicates the influence of the media on the intensity of news consumption on social media. Although the daily volume has fluctuated, the dynamics clearly show an upward trend ahead of the election. This indicates election-related news as well as the reader's attention is increasing toward the election, which reaches its peak on election day, then shows a downward trend afterwards.

In this study we only investigated 560 news outlets out of 700 outlets found in the dataset. Overall we only focus on domestic news media, which has been shared by 10 different Twitter users. Figure 1b shows that the statistics of news shares is dominated by mainstream media.

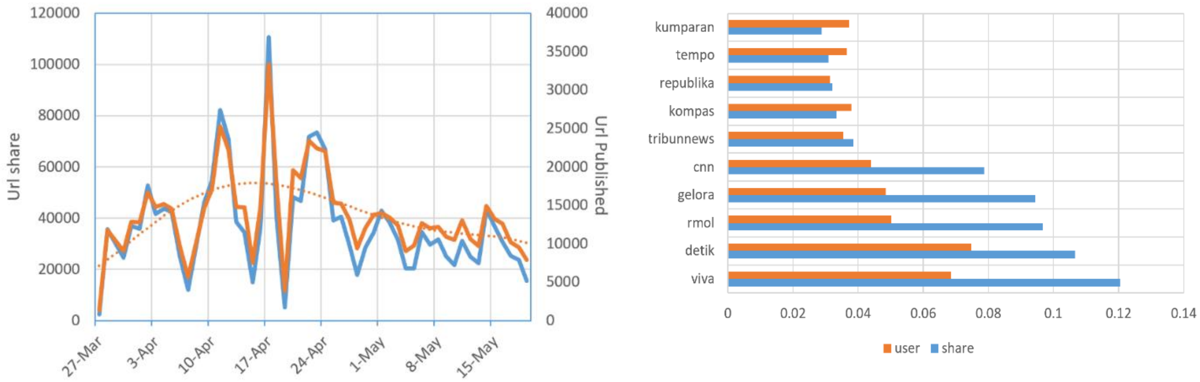


*Figure 2* Daily number of articles and unique articles shared by twitter users. Dotted lines are trend lines; (b) news outlets by number of *shared*

## 4.1 Political stance of news media outlets

Table 5 shows the number of news outlets that showed political alignments toward candidate. From this table, it is known that the proportion of Joko Widodo-leaning media is far greater than the number of news media in favor of Prabowo. However, Prabowo-leaning media is superior in terms of frequency of share and total users.

Table 5 Descriptive statistics of media stance

| | Total | # of Share | # of User |
|---|---|---|---|
| Media | 560 | 777990 | 101772 |
| pro Joko Widodo | 348 | 373932 | 39806 |
| pro Prabowo | 212 | 404058 | 61966 |

Figure 3a shows the distribution of media stance in the 2019 Indonesian elections. The bimodal pattern indicates the presence of media polarization where the majority of news outlets reside on the extreme side of the political spectrum. Meanwhile, Figure 3b shows the position of a number of mainstream media in the spectrum of political alignments. From this figure, it is known that political valence scores are able to capture the main differences between news outlets on each side of the spectrum, as follow:

- Majority of Islamic news media, such as **Republika**, **Portal-islam**, **Eramuslim**, **Kontenislam**, tend to favour Prabowo. This is not surprising because religious issues are very dominant in the 2019 Indonesian elections, and Prabowo was imaged as a representative of an Islamic group;
- The opposition news outlets which has criticized the Joko Widodo government for the past 5 years, such as **Viva**, **Rmol**, **Gelora**, has a valensi score on Prabowo;
- Most of mainstream news media, such as **Kompas**, **Detik** and **Tempo**, tend to support Joko Widodo. While some others such as **CNN**, **Merdeka**, tend to be politically neutral.

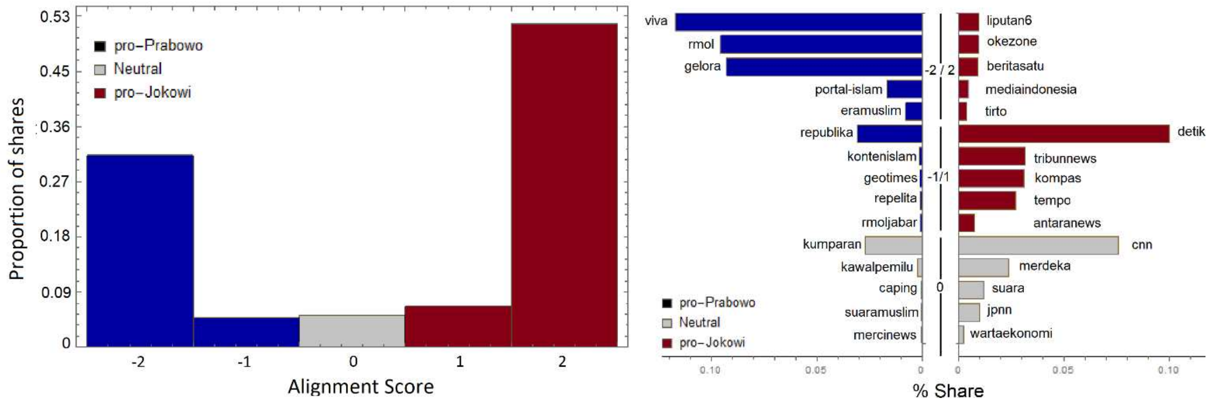


Figure 3 Media stance in the 2019 Indonesian elections (a) distribution of political valency scores; (b) Position of mainstream media.

The heat map shown in Figure 4 illustrates the relationship between the political alignments of news media and their credibility. In this study we use Alexa Rank [15] as a proxy to assess the credibility of a media. As shown in Figure 4a, the most of mainstream media with high credibility ratings have a neutral valence score or tend to favor Joko Widodo, while Prabowo-leaning media generally have moderate or low credibility. In addition, most of low-credibility media tend to have extreme political valence scores. In other words, low-credibility media tend to be more partisan than media with high credibility. From Figure 4b we know that, for all political valence scores, the intensity of information dissemination originating from low-credibility media is relatively not much different compared to high-credibility media. This means that partisan readers tend not to question the credibility of the news sources they share. We highlight this empirical fact in relation to the rise of false news during the election and the potential of low-credibility media as sources of misinformation.

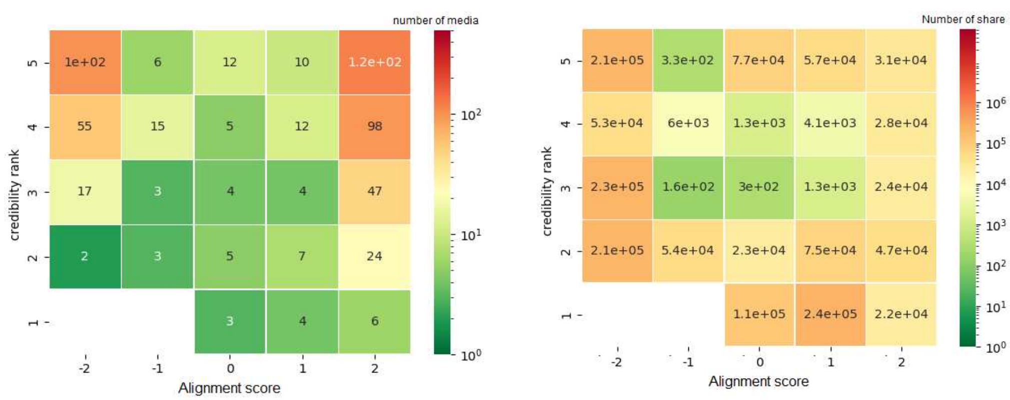


Figure 4. Heat map between political valence score vs credibility ranking by: (a) number of media; (b) number of shared article.

## 5. Conclusion

In this study we use partisan behavior of media audiences on Twitter to identify political alignments of news media during 2019 Indonesian elections. The identification method we proposed is carried out in 3 stages, as follow: (i) Identification of the political affiliations of twitter users based on the political hashtag they used in their tweet; (ii) Identification of the political affiliations of twitter users based on their interaction networks using the label propagation algorithm; (iii) Identification of the political alignments of news media based on the political affiliation of its audience. Evaluation results show that the proposed method is very effective in detecting the political affiliation of twitter users as well as predicting the political stance of news media. The position of media in the spectrum of political valence confirms the general allegations of media partisanship during the election. Further elaboration regarding news consumption behavior shows that low-credibility news outlets tend to have extreme political positions, while partisan readers tend not to question the credibility of the news sources they share.

**Acknowledgments**

The author thanks researchers at the Bandung Fe Institute for motivation and discussion. Thanks are extended to Dadan Suhandana BFItechnology for assistance in data processing.

**Reference**

[1] N. Nic, D. A. L. Levy, and R. K. Nielsen, "Reuters Institute Digital News Report 2018," *Reuters Inst. Digit. News*, 2018, doi: 10.1017/CBO9781107415324.004.

[2] P. J. Shoemaker, T. P. Vos, and S. D. Reese, "Journalists as gatekeepers," in *The Handbook of Journalism Studies*, 2009.

[3] H. Allcott and M. Gentzkow, "Sosial media and fake news in the 2016 election," *J. Econ. Perspect.*, vol. 31, no. 2, pp. 211–236, 2017.

[4] T. Groeling, "Media Bias by the Numbers: Challenges and Opportunities in the Empirical Study of Partisan News," *Annu. Rev. Polit. Sci.*, 2013, doi: 10.1146/annurev-polisci-040811-115123.

[5] P. Barberá and G. Sood, "Follow Your Ideology: A Measure of Ideological Location of Media Sources," *Eff. Br. mindfulness Interv. acute pain Exp. An Exam. Individ. Differ.*, 2015, doi: 10.1017/CBO9781107415324.004.

[6] C. Becatti, G. Caldarelli, R. Lambiotte, and F. Saracco, "Extracting significant signal of news consumption from social networks: the case of Twitter in Italian political elections," *Palgrave Commun.*, 2019, doi: 10.1057/s41599-019-0300-3.

[7] E. Bakshy, S. Messing, and L. A. Adamic, "Exposure to ideologically diverse news and opinion on Facebook," *Science (80-. ).*, 2015, doi: 10.1126/science.aaa1160.

[8] P. Stefanov, K. Darwish, A. Atanasov, and P. Nakov, "Predicting the Topical Stance of Media and Popular Twitter Users," *arXiv*, 2019.

[9] E. Borra and B. Rieder, "Programmed method: Developing a toolset for capturing and analyzing tweets," *Aslib J. Inf. Manag.*, 2014, doi: 10.1108/AJIM-09-2013-0094.

[10] M. Conover, J. Ratkiewicz, and M. Francisco, "Political polarization on twitter.," *Icwsm*, 2011, doi: 10.1021/ja202932e.

[11] M. Del Vicario *et al.*, "The spreading of misinformation online," *Proc. Natl. Acad. Sci. U. S. A.*,

2016, doi: 10.1073/pnas.1517441113.

[12] F. Zollo *et al.*, “Debunking in a world of tribes,” *PLoS One*, 2017, doi: 10.1371/journal.pone.0181821.

[13] U. N. Raghavan, R. Albert, and S. Kumara, “Near linear time algorithm to detect community structures in large-scale networks,” *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, 2007, doi: 10.1103/PhysRevE.76.036106.

[14] A. Badawy, E. Ferrara, and K. Lerman, “Analyzing the digital traces of political manipulation: The 2016 Russian interference Twitter campaign,” 2018, doi: 10.1109/ASONAM.2018.8508646.

[15] A. Internet, “Keyword Research, Competitor Analysis RESEARCH & Website Ranking,” 2019. .