Partisan press on Twitter: Clustered semantics, attention, and diffusion dynamics

Keywords: News partisanship, Attention Dynamics, Time Series Clustering, Topic Modeling, Information Diffusion

Partisanship plays a vital role in the public sphere of online social media due to selective information exposure and dissemination (Shin & Thorson, 2017; Stroud, 2010). Previous studies have shown that partisan bias in media is ubiquitous and influential for online social network users (Bakshy et al., 2015), leading to polarized news consumption, affects, and perceptions (Guess et al., 2021; Törnberg, 2022). Media with distinct partisanship affiliations deliver distinct attitudes towards social events (e.g., presidential election), orient public attention and set agendas towards different directions (e.g., gun control vs. drug control), and inform or fact-check the general population with distinct focuses. Previous studies found that different news on Twitter has different diffusion dynamics that vary across cascade depth, size, breadth, and structural virality (Goel et al., 2016; Vosoughi et al., 2018). Therefore, it is important to understand the extent to which partisanship in news media impacts the broadcasting contents, and how these contents draw the public's attention across time and its corresponding information diffusion processes.

Hypothesis

In the current study, we hypothesize that:

(H1) Media on Twitter with distinct partisanship affiliations will have different topic dynamics across time.

(H2) Media on Twitter with distinct partisanship affiliations will have different public attention (i.e., impression) dynamics across time.

(H3) Media on Twitter with distinct partisanship affiliations will have different diffusion cascade structural dynamics with respect to (H3a) depth, (H3b) size, (H3c) breadth, and (H3d) structural virality across time.

Attention dynamics

To test H1 and H2, a time-varying news media dataset is collected. News media are labeled with their corresponding partisanship affiliation (left, left-center, center, right-center, right), which is retrieved from Allsides[1] (an online community that labels media with partisan bias). For each news media, we collected their historical tweets from Dec-01-2022 to Feb-22-2023 using Twitter's official API. Attention dynamics were constructed as time series of aggregated impression metrics over days (Figure A). A moving average smoothing technique was applied to filter out weekly fluctuations. Later, the time series of impressions were clustered using a K-shape clustering algorithm (Paparrizos & Gravano, 2015), which separates the time series into two clusters (Table 1). A Chi-Square test was applied to test H1. The result of the chi-square test indicates that the clusters can significantly separate the attention dynamics of media with distinct partisanship ($X^2(4, N = 65) = 9.69$, $p = 0.046$; Figure 3). In addition, we found that the clusters can best separate the left-center and right-center but not the right and left partisanship. Lastly, we found that different clusters of attention dynamics have different trends, and peak at different times. Specifically, the cluster of left media has an increasing trend while the right media have a decreasing trend. In addition, left media has two major peaks at time step 20 and step 45, which corresponds to Mccarthy Kevin's speaker votes and the MSU school shooting event, while the right side of media does not receive comparable attention during these days.

---

[1] https://www.allsides.com/media-bias/ratings

Semantic analysis

We analyzed news semantics of partisan media, which might lead to diverse coverage of social issues and agenda-setting. We applied an advanced topic clustering method (BERTopic; Grootendorst, 2022) to every posted tweet for each media. A total of 3,456 topics were assigned to 394,768 collected tweets. Then, the labeled tweets were aggregated for each media, thus media were represented by a bag-of-topics. A TF-IDF representation was constructed for each media, which encodes the frequency and importance of news topics covered by each news media. Consequently, we applied a K-centroid clustering algorithm (Park & Jun, 2009) were applied to cluster different media, and a Chi-Squared test was applied to test H2. The result of the chi-square test indicates that the clusters based on topic representation can significantly separate the media with distinct partisanship ($X^2(4, N = 65) = 17.893$, $p = 0.001$; Figure 3). The semantic topic dynamics were plotted in Figure 4.

Cascades analysis

The data collection is ongoing. We are currently collecting retweeting relationships among retweeters of the tweets posted by news media. Then the cascading tree will be constructed, and the structural metrics of the cascading tree will be measured and compared between media with different partisanship. We expect to complete data collection and analysis before the conference meeting.

Reference

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Goel, S., Anderson, A., Hofman, J., & Watts, D. J. (2016). The Structural Virality of Online Diffusion. *Management Science*, *62*(1), 180–196. https://doi.org/10.1287/mnsc.2015.2158

Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure* (arXiv:2203.05794). arXiv. https://doi.org/10.48550/arXiv.2203.05794

Guess, A. M., Barberá, P., Munzert, S., & Yang, J. (2021). The consequences of online partisan media. *Proceedings of the National Academy of Sciences*, *118*(14), e2013464118. https://doi.org/10.1073/pnas.2013464118

Paparrizos, J., & Gravano, L. (2015). k-Shape: Efficient and Accurate Clustering of Time Series. *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, 1855–1870. https://doi.org/10.1145/2723372.2737793

Park, H.-S., & Jun, C.-H. (2009). A simple and fast algorithm for K-medoids clustering. *Expert Systems with Applications*, *36*(2, Part 2), 3336–3341. https://doi.org/10.1016/j.eswa.2008.01.039

Shin, J., & Thorson, K. (2017). Partisan Selective Sharing: The Biased Diffusion of Fact-Checking Messages on Social Media. *Journal of Communication*, *67*(2), 233–255. https://doi.org/10.1111/jcom.12284

Stroud, N. J. (2010). Polarization and Partisan Selective Exposure. *Journal of Communication*, *60*(3), 556–576. https://doi.org/10.1111/j.1460-2466.2010.01497.x

Törnberg, P. (2022). How digital media drive affective polarization through partisan sorting. *Proceedings of the National Academy of Sciences*, *119*(42), e2207159119. https://doi.org/10.1073/pnas.2207159119

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Figure A

Lineplot and the heatmap of time series of impressions across time for news media on Twitter

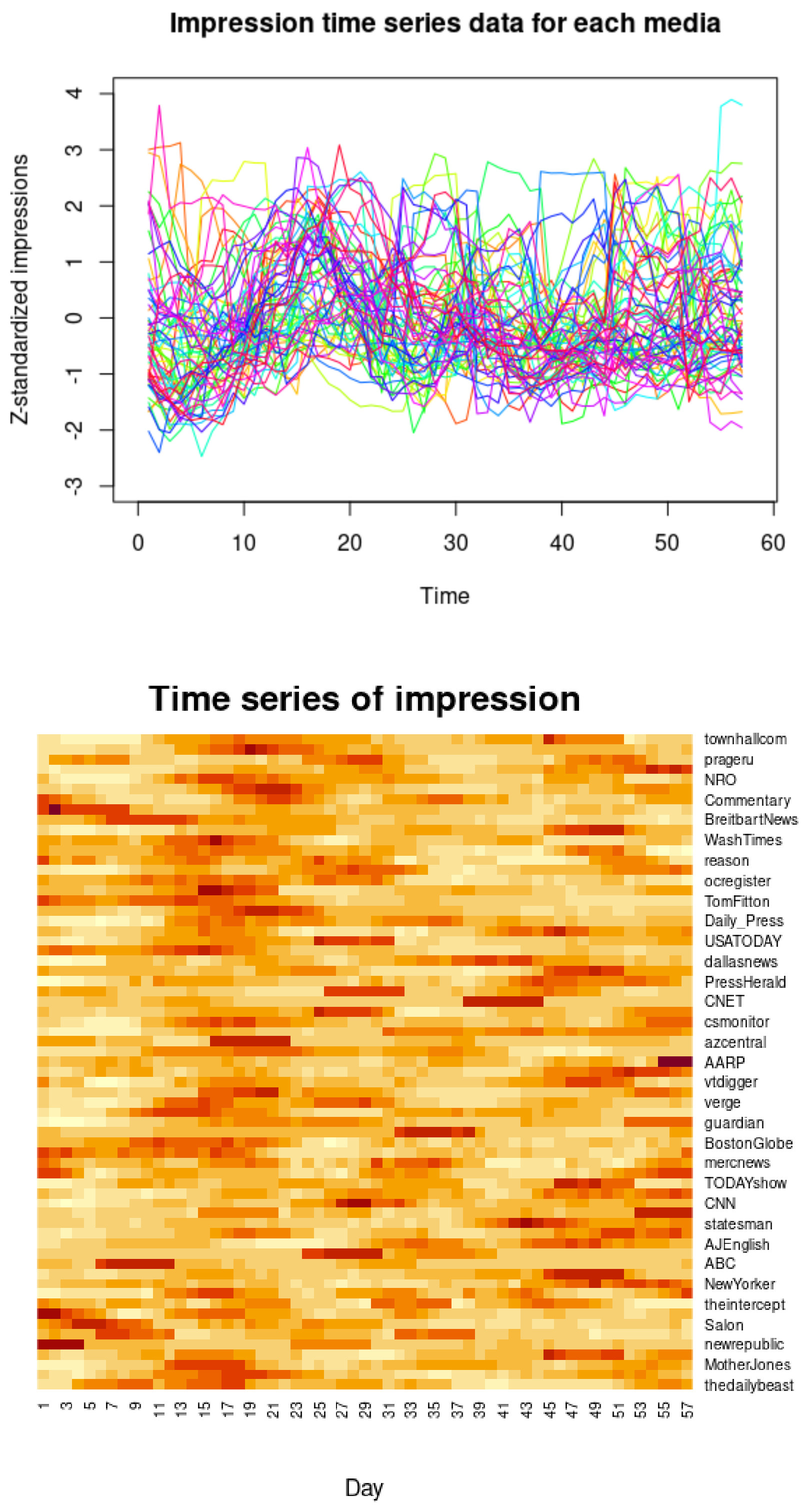


Figure 2

Attention Time series clusters with K-shape clustering method (Upper panel) and corresponding cluster centroid (Lower panel)

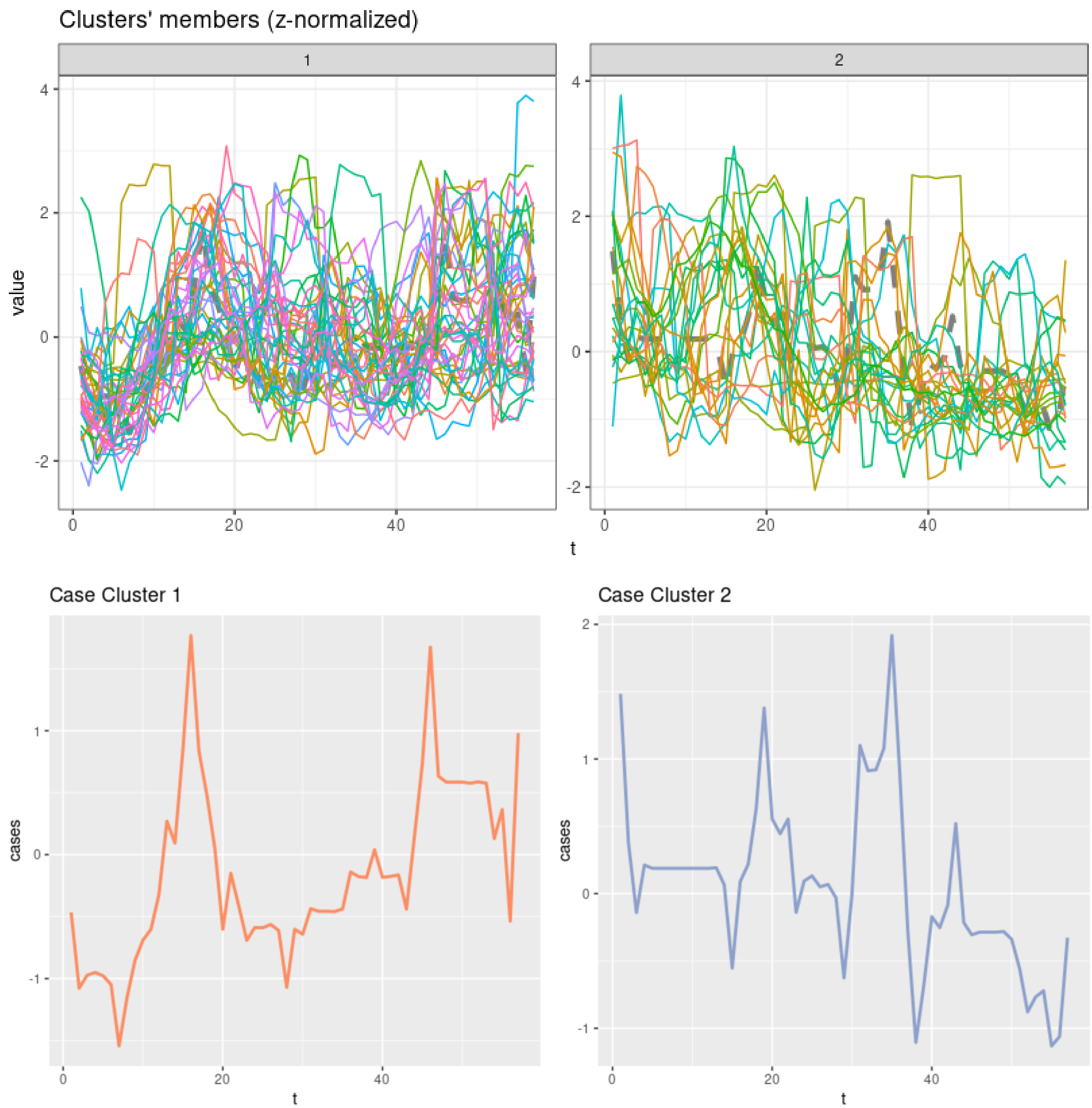
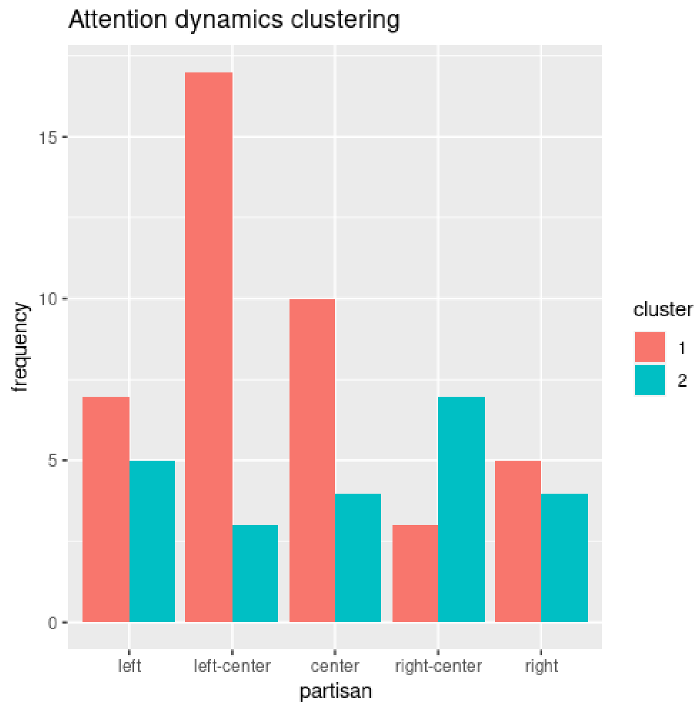


Figure 3

Frequency of partisan media for different clusters based on impression time series (Upper) and based on semantic topic representations (Bottom).



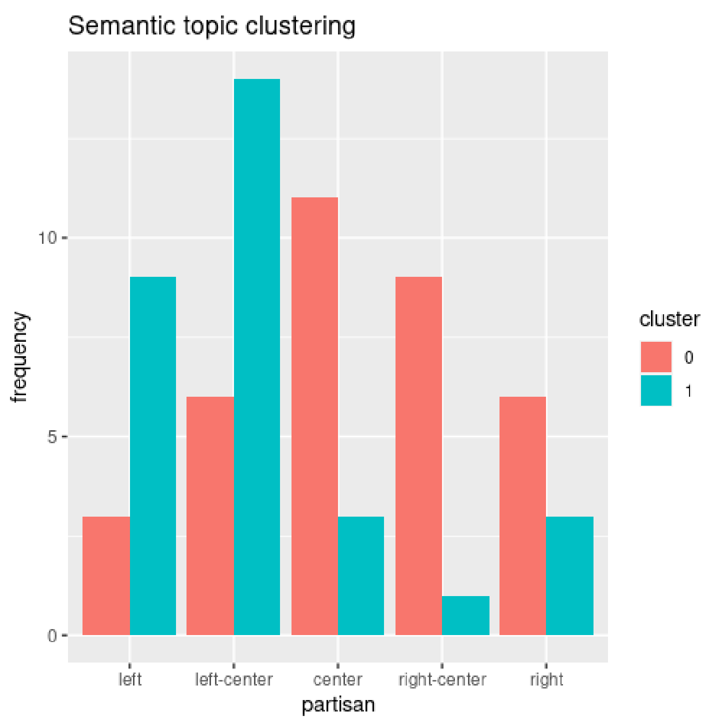


Figure 4
Topic frequency dynamics of media on Twitter.