Understanding Political Polarization on Social Platforms by Jointly Modeling Users, Connections and Multi-modal Post Contents in Heterogeneous Graphs

Hanjia Lyu, Jiebo Luo
University of Rochester, Rochester, New York, United States
hlyu5@ur.rochester.edu, jluo@cs.rochester.edu

Abstract

In this study, we investigate the political polarization in Twitter discussions about inflation during the COVID-19 pandemic using a dataset composed of more than 20,000 vetted tweets posted by over 8,000 unique Twitter users from July to November 2021. Our analysis shows the timing of the volume changes in online discussions roughly corresponds to the dates when the U.S. Bureau of Labor Statistics released the Consumer Price Index (CPI). The usage of the hashtags varies across left- and right-leaning users. Left-leaning users tend to discuss more diverse subtopics while right-leaning users focus on blaming U.S. President Joe Biden for the inflation. Unlike previous studies, we characterize political polarization by jointly modeling user information, their connections, and the multi-modal post contents in a heterogeneous graph. By mapping the node embeddings into a two-dimensional space, we find there is clear segregation between left- and right-leaning users. Although using only one of the features or the concatenation does not improve the performance of modeling, we find notable differences in user descriptions, topics, images, and levels of retweet activities. Our framework shows wide applicability and can be further extended by integrating other modalities of data and more relations. The findings have implications for understanding online political polarization and designing mitigation policies for potentially negative outcome of extreme polarization.

Introduction

A 2016 Pew Research study (Gottfried and Shearer 2016) found that of all U.S. adults, 67% use social media platforms with 44% using the platforms to discover news. Social media is found to shape political discourse in the U.S. and the whole world (O’Connor et al. 2010; Tumasjan et al. 2010; Conover et al. 2011b; Weld, Glenski, and Althoff 2021). As discussed in (Conover et al. 2011b), the formation of such political polarization is not necessarily a serious problem, but the concern is that the opinions may become increasingly extreme when they are shared, viewed and discussed in a homogeneous community. For example, on March 16, 2020, the former U.S. President Donald Trump posted a tweet calling COVID-19 “Chinese Virus”. After that, this term was frequently used among the right-leaning Twitter users, and in most cases, they were carried with negative sentiment against China and Asian people (Lyu et al. 2020; Chen et al. 2021). More importantly, the harm caused by biased and false news has a substantial negative impact across the entire society (Weld, Glenski, and Althoff 2021). Fake or extremely biased news regarding COVID-19 vaccines is found to be negatively correlated with the state-level COVID-19 vaccine uptake (Lyu, Zheng, and Luo 2021). The growing political divide in the U.S. also extends to the #StopAsian-Hate movement with Trump followers being less supportive to the Asian-American victims of the racially motivated hate crimes (Lyu et al. 2021a).

In this study, we attempt to understand the political polarization on Twitter using the discussion about inflation during the COVID-19 pandemic as a case study. Since December 2019 when the first case of COVID-19 was reported in Wuhan, China (Carvalho, Krammer, and Iwasaki 2021), the COVID-19 pandemic has infected almost 250 million people and caused over five million deaths as of November 7, 2021. More importantly, the suppression policies (e.g., social distancing) that aim to prevent the spread of the virus have led to a decrease in economic activities and thus impacted the global economy (Ozili and Arun 2020). At the early stage of the pandemic, deflation was observed (Christensen, Gamble, and Zhu 2020). With more and more people being vaccinated, countries started to lift lockdown restrictions and rebuild the economy. Brunnermeier et al. (2020) predicted that the redistribution of the focus of the economy sectors can cause excessive inflation in the long run. Banerjee et al. (2020) found that the pandemic increased the inflation risks in both advanced economies and emerging market economies. Using the inflation beliefs reported in the New York Fed’s Survey of Consumer Expectations, Armantier et al. (2021) found that household inflation expectations responded slowly and mostly at the short-term horizon. The extreme opinions about inflation may result in real-world impacts such as increasing spending in the present and decreasing spending in the future (Yadav and Shankar 2014). With the interaction (Rosita 2020; Loxton et al. 2021) between inflation and panic buying which has been observed during the pandemic (Zhang, Lyu, and Luo 2020), the inflation can become even more severe. According to a Yahoo
News/YouGov poll[4] about 77 percent Americans indicate that inflation is affecting their lives. Political polarization on inflation is also observed. About 90 percent of Republicans assign President Biden at least some blame for inflation while only 28 percent of Democrats do. The debate around inflation is also prevalent on online social platforms. Large-scale social media data (e.g., publicly available user profile and behavioral data) and abundant multi-modal information provide us with the opportunity to better understand the political polarization on inflation and in general.

## Related Work

Most prior work on learning political polarization on social platforms focuses on user behavior and content [Hosseini-mardi et al. 2021][Borge-Holthoefer et al. 2015][Gruzd and Roy 2014][Guarino et al. 2020][Conover et al. 2011b][Waller and Anderson 2021][Urman 2020]. Borge-Holthoefer et al. (2015) analyzed the Egyptian political polarization on Twitter by a set of manually labeled hashtags and the retweet network constructed by hand-labeled seed users. Hosseini-mardi et al. (2021) identified several distinct communities of news consumers, including “far-right” and “anti-woke” by examining the consumption of radical content on YouTube. Conover et al. (2011b) investigated how social media shape the networked public sphere and facilitate communication between communities with different political orientations. In particular, they used the 2010 U.S. congressional midterm elections as a case study and focused on the retweet and mention networks. Waller and Anderson (2021) argued that quantifying the online communities in a purely behavioral fashion may suffer from the biases that result from using self-reported data, expert labels and survey-based methods. They chose to develop a neural-embedding methodology to quantify the positioning of online communities along social dimensions by leveraging large-scale patterns of aggregate behavior. However, the selection of social dimensions is still highly subjective to expert knowledge. In addition, these studies paid little attention to user characteristics and information of other modalities apart from the text. In contrast, our study intends to combine both the user information and multi-modal content information.

An example of different node types and relation types on Twitter is illustrated in Figure 1. The network is composed of two types of nodes - users and tweets. The relation between a user and a tweet includes post, retweet, quote, mention, reply, and like. A user can follow another user and they can also follow each other. The tweet node has text information and sometimes it has other media information such as video, image, and link. Previous studies put an emphasis on the text. The user node has user characteristics such as number of followers, verified status, and profile description. The heterogeneity of information and complex relations among the nodes make it challenging to uncover insight. Recently, representing graph nodes in a low dimensional vector space based on factorization methods, random walks, and deep learning has been proposed to address this issue (Goyal and Ferrara 2018). Graph neural networks (GNN) have shown promising results in multiple applications including rumor detection (Huang et al. 2020), sentiment analysis (Huang and Carley 2019), recommender systems (Song et al. 2019), and ideology inference (Xiao et al. 2020). With respect to ideology inference, Xiao et al. (2020) emphasize on relations such as follow, retweet and mention. Our study differs from it since we focus on both context and relation. In other words, the major study target of [Xiao et al. 2020] is edge while ours is node and edge.

## Material and Method

### Dataset

This study leverages the publicly available tweets collected using the Twitter API[5]. To protect the privacy of the users in this study, we exclusively analyze the data and discuss the results in aggregate. Unlike Angelico et al. (2021) who used “inflation”, “inflationary”, “price”, “expensive”, “cheap”, and “expensive bill” as the search keywords, we only use one keyword - “inflation” and one hashtag - “#inflation” to collect the tweets that are about inflation because using their list of keywords may collect many false positive tweets. For example, “Pakistanis will surely pay the price of sending terrorists to Afghanistan.” would have been collected because of the word “price”. However, it is apparently not related to the inflation issue. Additionally, this study focuses on the political polarization in the United States. Therefore, we only include the Twitter users that are located in the United States by extracting the information from public user profiles. It is noteworthy that the tweets we collect may be posted by Twitter bots. To alleviate this issue, we only keep the Twitter users with profile images that have human faces, since lack of such imagery might be a sign of a Twitter bot account (Efthimion, Payne, and Proferes 2018). We use Face++[6] which analyzes a series of face attributes including age, gender, smile intensity. It has been applied in multiple social media studies [Lyu et al. 2021b].

[4] https://news.yahoo.com/poll-77-percent-of-americans-now-say-inflation-is-personally-affecting-them-and-57-percent-blame-biden-210739716.html [Accessed 2022-01-05]

[5] https://www.twepy.org/ [Accessed 2021-12-05]

[6] https://www.faceplusplus.com/ [Accessed 2021-12-05]
and can achieve a high accuracy in the age and gender inference of Twitter users [Jung et al. 2018]. To avoid ambiguous inference, we only infer the demographics of the users who have a profile image with only one human face. Although it is beyond the scope of this study to analyze the relationship between bots and online polarization, a more extensive experiment on detecting bots could have important implications for future work.

Political Affiliation

Golbeck and Hansen (2014) showed that Twitter users mostly follow accounts that share the same political ideology. Therefore, to infer the political affiliation of each Twitter user, we examine the number of Twitter accounts with known political leaning that this user follows. The accounts with known political leaning include media accounts, journalists and political figures. The political leanings of the media accounts and the journalists are judged and assigned by allsides.com[1] and politico.com[2] respectively. This set of Twitter accounts is roughly balanced with around 400 right-leaning and 400 left-leaning. Prominent politicians are excluded since Twitter users may just follow them because of their popularity and importance. Using the Twitter API, we collect the Twitter accounts that the Twitter users follow and count the number of right- and left-leaning accounts, denoted by $N_R$ and $N_L$, respectively. The political score $P_S$ is calculated as follows:

$$P_S = \frac{N_R - N_L}{N_R + N_L} \quad (1)$$

The political score has a range of $[-1,1]$ with 1 meaning the most extreme right and −1 meaning the most extreme left. To increase the robustness of inference, only the political scores of the Twitter users who follow at least five accounts from either side (i.e., right- and left-leaning) are calculated. In this way, we calculate the political scores of 8,823 Twitter users. Figure 2 shows the distribution of political scores. More Twitter users are left-leaning (as expected).

Graph Representation Learning

Graph definition. There are two types of nodes in our studies including users and tweets as shown in Figure 1. An edge between a user node and a tweet node is created if the user posts or retweets this tweet. To capture the user in-network interactions, we adopt the Heterogeneous Graph Neural Network (Het-GNN) (Zhang et al. 2019), which first uses random walks with restart to generate neighbors for nodes and capture the structural information, and then leverages bi-directional LSTM (Bi-LSTM) (Hochreiter and Schmidhuber 1997) and the attention mechanism (Veličković et al. 2018) to aggregate node features within each type and among types. Here we briefly describe the framework by Zhang et al. (2019).

We first fuse heterogeneous contents $C_v(v \in V)$ via a neural network $f_1$. In our study, the heterogeneous contents of the user node include the number of followers and profile description, etc. The heterogeneous contents of the tweet node include text and image. Let $x_i \in \mathbb{R}^{d_f \times 1}$ denote the $i$-th content in $C_v$ where $d_f$ is the content feature dimension. $x_i$ can be obtained using different feature embeddings methods as aforementioned. Therefore, the content embeddings $f_1(v) \in \mathbb{R}^{d \times 1}$ ($d$: content embeddings dimension).
The dimension of the final feature space is 128.

It is noteworthy that \( h \) denotes a feature transformer which can be identity or fully connected neural network, etc. Unlike Zhang et al. [2019] where they use linear transformation to map the input feature embeddings of different nodes which may vary in dimension into a shared feature space, we use the well-studied principal component analysis (Wold, Esbensen, and Geladi [1987]) for simplicity.

Next, we aggregate the neighbor information of node \( v \) via two steps: (1) using Bi-LSTM to aggregate embeddings of the neighbor nodes of the same type, and (2) aggregating the embeddings of the neighbor nodes of different types through the attention mechanism. By employing the random walk with restart strategy, we generate the t-type sampled neighbor set of \( v \in V \) as \( N_t(v) \). The aggregated t-type neighbor embeddings \( f_2(v) \in \mathbb{R}^{d \times 1} \) (\( d \): aggregated content embeddings dimension) are then calculated as follows:

\[
f_2(v) = \sum_{v' \in N_t(v)} \left[ \text{LSTM} \left\{ f_1(v') \right\} \oplus \text{LSTM} \left\{ f_1(v') \right\} \right]_{|N_t(v)|}
\]

Let \( P_V \) denote the set of node types in the graph. In our study, \( P_V = \{ V_{user}, V_{tweet} \} \). The output embeddings \( \mathcal{E}_v \in \mathbb{R}^{d \times 1} \) of node \( v \) are computed as follows:

\[
\mathcal{E}_v = \alpha_{v,v} f_1(v) + \sum_{t \in P_V} \alpha_{v,t} f_2(v) \tag{4}
\]

where \( \alpha_{v,v} \) represents the attention coefficients of (1) content embeddings of \( v \) or (2) the aggregated neighbor embeddings to node \( v \). These two sets of embeddings are grouped and denoted as \( F(v,i) \in \mathbb{R}^{d \times 1} \) which is defined as follows:

\[
F(v,i) = \begin{cases} f_1(v) & i = v \\ f_2(v) & i = t, \text{where } t \in P_V \end{cases}
\]

\[
F(v,i) \in P, \text{ where } P = \{ v \} \cup P_V
\]

Therefore, Eq. (4) is re-formulated as:

\[
\mathcal{E}_v = \sum_{i \in P} \alpha_{v,i} F(v,i) \tag{7}
\]

The attention coefficients \( \alpha_{v,i} \) are built by a single-layer feedforward neural network parametrized by a weight vector \( u \in \mathbb{R}^{2d \times 1} \), with the LeakyReLU nonlinearity. To make the coefficients comparable, we normalize them across all choices of \( i \) using the softmax function. Thus, the coefficients are expressed as:

\[
\alpha_{v,i} = \frac{\exp(\text{LeakyReLU}(u^T[F(v,v) \oplus F(v,i)]))}{\sum_{j \in P} \exp(\text{LeakyReLU}(u^T[F(v,v) \oplus F(v,j)]))} \tag{8}
\]

where \( ^T \) represents transposition and \( \oplus \) is the concatenation operation.

All the hyper-parameters are set as reported by Zhang et al. [2019]. The dimension of the final feature space is T28. For random walk, the walk length is 30 and window size is 5.

While there are other feature embeddings methods and heterogeneous graph representation learning methods that could potentially improve the performance of overall representation learning [Lv et al. 2021], our ultimate goal is not to perfectly characterize the user, behavior, tweet content and relation. Instead, we aim to improve our understanding of political polarization by capturing posting and retweeting behavior, and fusing user information and tweet content information, which we will show is a better framework than using only user information, only tweet content information, or simple concatenation (no relation). Additionally, if a tweet is associated with videos, they can be mapped into the feature space via different feature embeddings methods. The framework is the same. In our study, we only focus on text and image for simplicity.

**Results**

**Preliminary Analysis**

**General discussions.** Figure 5 presents the number of collected posts that are related to inflation. Every month, the U.S. Bureau of Labor Statistics releases the Consumer Price Index (CPI) of the previous month. We annotate the news release using vertical lines. A first observation is that the peaks of the number of posts roughly correspond to the dates of news release. Additionally, the scales of the peaks are rather related to the percent changes in CPI. On November 10, 2021, the U.S. Bureau of Labor Statistics reported that the CPI for all urban consumers increased 0.9 percent in October on a seasonally adjusted basis after rising 0.4 percent in September 8. On the same date, the number of posts related to inflation jumped to around 12,000.

**Food and energy categories.** Food and energy are two major expenditure categories of CPI. To obtain a better understanding of the items that people pay more attention to, we use a keyword search to identify tweets that mention either food or energy or both. The list of keywords is constructed based on the detailed expenditure categories of CPI provided by the U.S. Bureau of Labor Statistics [9]. Food includes food at home (e.g., cereals and meats) and food away from home (e.g., full service meals and snacks). The energy category is composed of energy commodities (e.g., fuel oil and propane) and energy services (e.g., utility). However, some of these categories are general. For example, people may not explicitly use “full service meals and snacks” to describe the food. To expand the keyword search list, we use word2vec [Mikolov et al. 2013] which maps words into a high-dimensional space where the vectors of the words are close to each other if the words are used in a similar way. For example, the cosine similarity between the tokens san_franisco and los_angeles is 0.67, but the cosine similarity between the tokens san_franisco and golden_gate is...
The word2vec model that was pre-trained on part of the Google News dataset (about 100 billion words) is employed to find the top 10 similar words of each word in the detailed expenditure categories of food and energy. The expanded keyword list is used to identify tweets that mention specific items. Figure 3 and Figure 4 show the number of posts that mention food and energy, respectively. On August 11, 2021, the U.S. Bureau of Labor Statistics reported that the CPI for all urban consumers increased 0.5 percent in July. There is a local peak in Figure 3, however, no peaks are observed in Figure 4. For the next three news releases, the local peaks for all tweets, food-related, and energy-related tweets roughly synchronize. Online discussion about inflation is more general at first, but it becomes more specific gradually, indicating a more specific understanding of inflation of the people. We also observe a spike in the number of food-related tweets on November 4, 2021, one week before the news release of CPI. On the same day, CNN discussed how badly inflation is hitting the middle class where CNN quoted a statement about the price increase in milk and this piece of news was retweeted multiple times.

**Left and right participation.** If the political score is lower than or equal to 0, we label the user as left-leaning. If the political score is higher than 0, we label the user as right-leaning. Figure 3 and Figure 4 show the number of posts posted by left- and right-leaning users, respectively. The peaks from both timelines are similar, apart from November 4, 2021, when there is a peak of left-leaning users but no peak of right-leaning users. As aforementioned, on the same day, CNN reported a piece of news about inflation. According to allsides.com, the media bias of CNN is left. This is consistent with Conover et al. (2011) in that the network of political retweets exhibits a highly segregated partisan structure, with extremely limited connectivity between left- and right-leaning users.

**Hashtags.** We compare the most frequent patterns of hashtags used by left- and right-leaning users. Frequent pattern growth (FP-growth) algorithm (Han, Pei, and Yin 2000) is applied with a minimum support of 0.02. FP-growth is an efficient method for mining the complete set of frequent patterns by pattern fragment growth. Figure 5 presents the top 10 most frequent patterns of the usage of hashtags by left- and right-leaning users. The first patterns of both groups are “#inflation” which is within the expectation. An inspection of the tweets with “#bitcoin” suggests that some people think bitcoin can fix inflation. An example tweet is

> "#Inflation is a cancer that has been killing civilizations throughout history. #Bitcoin is the cure."

However, starting from the third most frequent pattern, there is clear segregation between the left- and right-leaning users. The left-leaning users show support to the Build Back Better Framework (#buildbackbetter) which is President Biden’s plan to rebuild the middle class. One example tweet is “Nobel-winning economists agreed in a recent letter: the #BuildBackBetter plan will reduce inflation by in-

---

1[https://code.google.com/archive/p/word2vec/](https://code.google.com/archive/p/word2vec/) [Accessed 2021-12-05]
highly related to the economy. Some users retweet the analyst reports of investment experts. For example, an analysis points out that today’s inflation is unusual and challenging for investors because the term inflation usually refers to a broad-based increase in prices across an economy that is attributed to the factors like monetary policy or an overheating growth rate instead of today’s situation where there are long delays for the materials and supplies that businesses and consumers need. In this sense, today’s inflation is more like “supply chain disruptions”.

Out of the top 10 most frequent patterns of the hashtags used by right-leaning users, six are related to President Joe Biden. A strong negative emotion is observed among the right-leaning users. They refer to inflation as “#bidenflation”:

“We Will Joey do ANYTHING today to ameliorate spiraling inflation? #Bidenflation”

They use #bidendelivers to blame President Joe Biden for causing the inflation:

“Here are all the things #BidenDelivers: High gas prices, Rapid inflation, 10+ million unfilled jobs, Chaos in the Middle East, Border crisis, Woke Military, Rising crime, Supply chain crisis, Critical Race Theory, Higher taxes, Refugee resettlement, IRS expansion.”

Offensive hashtags like #fjb and #letsgobrandon are also frequently used. #letsgobrandon is a political slogan that has been widely used as a minced oath for “Fuck Joe Biden” and was originated from a CNN interview with President Joe Biden when he was asked about inflation, which took about 20 seconds, during which Biden stood with his fists clenched and his elbows locked at a 90-degree angle as if piloting an imaginary jetpack. The right-leaning users think Biden tried to avoid answering questions about inflation.

Graph Representation and Political Polarization

In this section, we intend to analyze the user characteristics and tweet content posted by left- and right-leaning users. To better visualize the vectors of left- and right-leaning users, we use t-distributed stochastic neighbor embedding (t-SNE) which can visualize high-dimensional data by giving each data point a location in a two or three-dimensional map. t-SNE is capable of capturing the local as well as the global structures to reveal information of the presence of clusters. 

Six different representation strategies are applied.

- **User info**: The concatenation of the number of followers, friends, listed memberships, statuses, favorites, verification status and the sentence embeddings of the profile description is used.
- **Text**: Only the sentence embeddings of the tweet text content are used.
- **Image**: Only the image embeddings of the images that are associated with the tweets are used.
- **Text + Image**: The concatenation of text and image embeddings is used.
- **User info + Text + Image**: The concatenation of user information, text and image embeddings is used.
- **GNN-fused**: The graph representation learned by HetGNN is used.

Figure 5 shows the t-SNE visualization of the output of six different representation strategies. Figures 5a, 5e, and 5f show the user nodes, while Figures 5b, 5c, and 5d show the tweet node. The political score of the tweet is assigned by the average political scores of the users who post it. By comparing Figures 5a through 5f, we find that the representation strategies involving the text content perform better in modeling political polarization than the representation strategies using user information or image content only do. The small cluster in the right of Figure 5a stands for the users without descriptions. In Figures 5b, 5c, 5d, 5e, and 5f, the data points tend to be closer to the data points with the same colors, although the segregation levels vary. For instance, the relative distances between the red and blue clusters in Figures 5b and 5c are smaller than the distances in Figures 5e and 5f. Comparing Figures 5a and 5e, we find that adding the user information improves the performance of modeling political polarization. However, simple concatenation does not capture the relations among the users and tweets which can be reflected in Figures 5e and 5f where there is apparently greater segregation between the left- and right-leaning users using the GNN-fused representation strategy.

The GNN-fused representation strategy roughly captures three major groups of the left- and right-leaning users: the red clusters in the left, the mixture of blue and red data points in the middle, and the blue clusters in the right. For better presentation, we highlight these three groups in Figure 5e. It is also noteworthy that multiple local sub-clusters are formed within the left-leaning users. This is similar to the usage of hashtags where left-leaning users talk about multiple topics (e.g., China, stocks, the Build Back Better Framework), but a large body of hashtags of right-leaning users is only about the negative opinions against Biden. We also highlight two sub-clusters in Figure 5a for case study. For simplicity, we refer to the users...
Figure 5: t-SNE visualization of the output of six different representation strategies. The color indicates the political scores. The bluer the color is, the lower the political score is (more left-leaning). The redder the color is, the higher the political score is (more right-leaning). Each data point represents one user node. The political score of the tweet is assigned by the average political scores of the users who post it. Panel (a) shows the t-SNE output of the embeddings of only user information. Panel (b) shows the t-SNE output of the sentence embeddings of the tweet text content. Panel (c) shows the t-SNE output of the image embeddings of the images that are associated with the tweets. Panel (d) shows the t-SNE output of the concatenation of text and image embeddings. Panel (e) shows the t-SNE output of the concatenation of user information, text and image embeddings. Panel (f) shows the t-SNE output of node embeddings learned by the GNN-fused method. Panels (g) and (h) show the identical t-SNE output in Panel (f). They are plotted for a clearer presentation. Area 1 in Panel (g) indicates the clusters of right-leaning users. Area 2 in Panel (g) indicates the clusters of both left- and right-leaning users. Area 3 in Panel (g) indicates the clusters of left-leaning users. Areas 4 and 5 in Panel (h) indicate separate sub-clusters of left-leaning users.

of Areas 1, 2, 3 in Figure 5g as Right, Middle, and Left, respectively. We refer to the users of Areas 4 and 5 in Figure 5h as Left (sub-cluster 1) and Left (sub-cluster 2), respectively.

The political polarization is less severe in the Middle cluster. The average political scores of the users of Right, Middle, and Left are 0.66 (SD=0.45), -0.30 (SD=0.67), and -0.79 (SD=0.28), respectively. Left-leaning users (N=2,271) are two times more than right-leaning users in the Middle cluster. Additionally, the political polarization in the Middle cluster is less severe with the average political scores of the left-leaning users in Middle being -0.71 (SD=0.26) and right-leaning users in Middle being 0.62 (SD=0.30).

User description is a good indicator of political affiliations. To understand the differences among the users of Right, Middle, Left, Left (sub-cluster 1) and Left (sub-cluster 2), we visualize the user information and the multi-modal post content. Figure 6a shows the word cloud of the user descriptions of the users in each group. Interestingly, “love” is observed in the descriptions of all three groups. The words of the Left and Right clusters are associated with the political affiliation. For instance, there are “blm”, “resist”, and “democrat” in the Left cluster, while “maga”, “conservative”, “trump”, and “christian” in the Right cluster. Left-leaning users are found to be more engaged in online civil right movements such as “#BlackLivesMatter” (Panda, Siddarth, and Pal 2020) and “#StopAsianHate” (Lyu et al. 2021a). “resist” represents the American liberal political movement that protested the presidency of Donald Trump. “Make America Great Again” or “MAGA” is a campaign slogan popularized by Donald Trump. With respect to “christian”, Pew Research Center (2015) found there are proportionally more Christians among conservatives than among democrats and Democratic leaners. Moreover, there are fewer political keywords in the descriptions of Middle making it harder to classify left- and right-leaning users using only user information which is consistent with our findings in Figure 5. By comparing the word clouds of the user descriptions of Left (sub-cluster 1) and Left (sub-cluster 2), we find they have different focuses with “blm” in Left (sub-cluster 1) and “resist” in Left (sub-cluster 2).

Topics vary across different political affiliation. We apply a Latent Dirichlet Allocation (LDA) model [Blei, Ng, 2003] on the user descriptions. The LDA model is trained on the user descriptions of each cluster. The topics with the highest probability in each cluster are shown in Figure 6b. The topics vary across different political affiliation. For instance, the topics with high probability in the Left cluster are related to civil rights movements such as “#BlackLivesMatter” and “#StopAsianHate”. The topics with high probability in the Right cluster are related to political movements such as “MAGA” and “conservative”.

[1] https://en.wikipedia.org/wiki/The_Resistance_(American_political_movement) [Accessed 2022-01-09]
[2] https://en.wikipedia.org/wiki/Make_America_Great_Again [Accessed 2022-01-09]
and Jordan 2003) to the tweets of each group to extract the topics. The hyper-parameters are selected based on a grid search and the coherence score. By assigning the dominant topic label to each tweet, we obtain the topic distributions. We refer to the topic with the highest proportion as the most popular topic among the tweets of each group. Figure 6b shows the word clouds of the most popular topics. The size of the word is proportional to the weight assigned by each LDA model. Combining the findings regarding the usage of hashtags, we find that “border” and “crisis” in the Right cluster are related to the tweets that express negative opinions against President Joe Biden. “build_back”, as one of the keywords in the most popular topic in the Left cluster refers to the Build Back Better Framework. Users in Middle describe and discuss inflation in a more general way. “pandemic”, “supply”, “labor” and “shortage” are related to the factors that are considered to have effects on inflation. The keywords in Left (sub-cluster 1) and Left (sub-cluster 2) suggest a different emphasis on the topics. The left-leaning users in Left (sub-cluster 1) pay more attention to food, gas and tax while the left-leaning users in Left (sub-cluster 2) discuss more about the policy.

Source and content of images imply political affiliation. The images that are related to these five groups are in general different. We manually read the unique images of these five groups (84 for Right, 359 for Middle, 85 for Left, 7 for Left (sub-cluster 1) and 7 for Left (sub-cluster 2)), select the representative ones and plot them in Figure 4. Almost all groups post images of political figures (e.g., Joe Biden, Donald Trump) which is within our expectation based on the previous discussions on hashtags and topics. In Left (lower left) and Middle (lower right), we observe more portraits while in Right (upper left) there are more sarcasm images of Joe Biden. Right-leaning users also share images with text (lower left in Right) to blame President Joe Biden for his policies. Interestingly, users of Right share more eye-catching images. After manual inspection, we find users in Left share screenshots of long paragraphs that explain the reason for inflation (upper left and right in Left). Users in Middle use images of items that can reflect inflation (lower left in Middle) and geographic maps (upper right in Middle) more often. However, users in Right use images with text describing the price increase (upper right in Right) more frequently. These images intend to depict inflation more vividly. Furthermore, users in Middle share images of charts as users in Left do, however, the sources of this type of image in Left are mainly Federal Reserve Economic Data and the U.S. Bureau of Labor Statistics, while the sources of this type of image in Middle are mainly media and financial companies. Moreover, the screenshots of media interviews or reports are observed in Left (lower right) and Right (lower left). The images of Left (sub-cluster 1) and Left (sub-cluster 2) share similar properties of the images of Left, however, no notable differences are observed due to the limited sample sizes.

Adding relation information improves the understanding of political polarization. In previous discussions, clearer segregation in Figure 5 than in Figure 5 shows that capturing relations (retweet relation in our study) improves the modeling of political polarization. To better understand this mechanism, we calculate the number of users per unique tweet in each group. This value indicates the number of retweet activities. The higher this value is, the more users retweet the same tweet, suggesting more retweet activities. There are more retweet activities in the Left cluster (4.69 users per unique tweet) and Right cluster (3.04 users per unique tweet) than in the Middle cluster (1.10 users per unique tweet). Fewer retweet activities indicate a sparser network on which GNN performs poorly due to the limited knowledge gain from less representative neighbors (Jia et al. 2020).

Discussion and Conclusion

In this study, we first show that the discussions about inflation on Twitter are most heated when the U.S. Bureau of Labor Statistics reported the CPI of the previous month (Fig. 3a). Additionally, some local peaks of the number of posts roughly correspond to media coverage. The number of posts of left- and right-leaning users (Fig. 3a and Fig. 3b) share a similar pattern with the general number of posts (Fig. 3c), with an outlier when many left-leaning users retweeted a report of a left-biased media while right-leaning users did not.

By diving deeper into the content of the tweets of left- and right-leaning users, we find there is partisan segregation. Using the most frequent pattern of hashtags as an indicator, we observe that the left-leaning users discuss inflation in a more diverse way where they may explain the causes for today’s inflation, support Biden’s plan that is believed to address inflation, and point out that inflation is not an issue only in the U.S. Compared to the left-leaning users, the right-leaning users’ topics are more condensed to six most frequently used hashtags targeting President Joe Biden. All of these six hashtags show negative sentiment toward Biden. Align with previous studies (Lyu et al. 2020; Chen et al. 2021), the left-leaning users show a higher degree of hierarchical analytical thinking when they talk about political issues than the right-leaning users do. We have also observed that the left-leaning users discuss the inflation in China. This form of action was detected in multiple previous research (Lewandowsky, Jetter, and Ecker 2020, Roberts and Roberts 2018, Timmer et al. 2018) that people try to divert attention from controversial issues. Therefore, this action of referring to China by the left-leaning users could be a response to the blame on Biden by the right-leaning users. We also conduct a temporal analysis on hashtags, but the findings are similar. Thus we only report our findings in an aggregate manner.

By jointly learning user information, their connections and the multi-modal post contents in the heterogeneous Twitter network, we achieve a better performance of modeling political polarization (Figure 5). First, we find that the left- and right-leaning users describe themselves differently. They use politics-related keywords such as “blm”, “resist”,...
Figure 6: User information and the multi-modal post content that are related to the users of Right, Middle, Left, Left (sub-cluster 1) and Left (sub-cluster 2). Panel (a) shows the word cloud of the user descriptions of the users in each group. The size of the word is proportional to the frequency of the appearance. Panel (b) shows the word cloud of the most popular topic of each group. The size of the word is proportional to the weight assigned by the LDA model. Panel (c) lists the representative images that are shared by the users of each group.

and “maga”. However, consistent with the findings of Pen-nacchiotti and Popescu (2011), profile descriptions do not contain enough good-quality information to be directly used for user classification purposes (Figure 5). Second, by comparing the most popular topics among the left- and right-leaning users, we find the major semantic themes are different where the left-leaning users pay more attention to the Build Back Better Framework proposed by Biden while the right-leaning users focus on blaming Biden for inflation and a series of other societal issues, suggesting that the content of political discourse on Twitter remains partisan (Conover et al. 2011b). Conover et al. (2011a) found that applying topic modeling to the content of users’ tweets can help identify hidden structures in the data strongly associated with political affiliation, but did not find that topic detection improves prediction performance which is in line with our findings in Figures 5 and 5. Third, we have also found discrepancies in the images that the left- and right-leaning users use to describe inflation where the left-leaning users tend to share screenshots of paragraphs that explain the causes of inflation while the right-leaning users convey the idea of inflation using more eye-catching images. Fourth, the partisan clusters are more easy to detect if there are many retweet activities.

There are a few of limitations in our study. First, although the current keyword list we use to collect the relevant tweets produce a high precision rate, future work can expand the keyword list and build a classifier to filter out the irrelevant tweets in an attempt to achieve a good balance between precision and recall. Second, using the follower relationship on Twitter to infer the political scores is easy to implement, however, to obtain better inference, traditional survey-based methods can be leveraged to further reduce the noise. Third, since the quality of graph representation learning will have an enormous impact on the utility of political polarization modeling, we intend to refine our methods including more extensive experiments for selecting the feature embeddings, designing and implementing a better heterogeneous graph neural network framework, and fine-tuning the hyper-parameters. Fourth, we only capture retweet relations, while an extension to following and mentioning relations could be an important direction of future work. Finally, the current study focuses on political polarization on Twitter, however, modeling political polarization across different social platforms such as Facebook, Reddit, and YouTube by using the discussions around other economic and societal issues would further demonstrate the applicability and generalizability of our proposed heterogeneous graph-based framework.

Broader Impact and Ethical Consideration

Our study provides a better understanding of the political polarization in Twitter discussions about an economic and societal issue, which can help improve the strategy for policymakers in both the government and social platforms to mitigate or even prevent the potential negative outcomes of extreme political polarization. All data collected for this study are publicly available through the official Twitter API. Ag-
aggregated data is available upon request in an attempt to promote future work and at the same time minimize possibility of unethical use and privacy risks.

References

Angelico, C.; Marcucci, J.; Miccoli, M.; and Quarta, F. 2021. Can we measure inflation expectations using Twitter? Bank of Italy Temi di Discussione (Working Paper) No 1318.

Armanitier, O.; Koşar, G.; Pomerantz, R.; Skandalis, D.; Smith, K.; Topa, G.; and Van der Klauw, W. 2021. How economic crises affect inflation beliefs: Evidence from the Covid-19 pandemic. Journal of Economic Behavior & Organization 189: 443–469.

Banerjee, R. N.; Mehrotra, A.; Zampolli, F.; et al. 2020. Inflation at risk from Covid-19. Technical report, Bank for International Settlements.

Blei, D. M.; Ng, A. Y.; and Jordan, M. I. 2003. Latent dirichlet allocation. the Journal of machine learning research 3: 993–1022.

Borge-Holthoefer, J.; Magdy, W.; Darwish, K.; and Weber, I. 2015. Content and network dynamics behind Egyptian political polarization on Twitter. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, 700–711.

Brunnermeier, M.; Merkel, S.; Payne, J.; and Sannikov, Y. 2020. COVID-19: Inflation and Deflation Pressures. CESifo Area Conferences, mimeo.

Carvalho, T.; Krammer, F.; and Iwasaki, A. 2021. The first 12 months of COVID-19: a timeline of immunological insights. Nature Reviews Immunology 21(4): 245–256.

Chen, L.; Lyu, H.; Yang, T.; Wang, Y.; and Luo, J. 2021. Fine-grained analysis of the use of neutral and controversial terms for COVID-19 on social media. In International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation, 57–67. Springer.

Christensen, J. H.; Gamble, J. M.; and Zhu, S. 2020. Coronavirus and the Risk of Deflation. FRBSF Economic Letter 2020(11): 1–5.

Conover, M. D.; Gonçalves, B.; Ratkiewicz, J.; Flammini, A.; and Menczer, F. 2011a. Predicting the political alignment of twitter users. In 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing, 192–199. IEEE.

Conover, M. D.; Ratkiewicz, J.; Francisco, M.; Gonçalves, B.; Menczer, F.; and Flammini, A. 2011b. Political polarization on twitter. In Fifth international AAAI conference on weblogs and social media.

Efthimion, P. G.; Payne, S.; and Proferes, N. 2018. Supervised machine learning bot detection techniques to identify social twitter bots. SMU Data Science Review 1(2): 5.

Golbeck, J.; and Hansen, D. 2014. A method for computing political preference among Twitter followers. Social Networks 36: 177–184.

Gottfried, J.; and Shearer, E. 2016. News use across social media platforms 2016.

Goyal, P.; and Ferrara, E. 2018. Graph embedding techniques, applications, and performance: A survey. Knowledge-Based Systems 151: 78–94.

Gruzd, A.; and Roy, J. 2014. Investigating political polarization on Twitter: A Canadian perspective. Policy & internet 6(1): 28–45.

Guarino, S.; Trino, N.; Celestini, A.; Chesa, A.; and Riotta, G. 2020. Characterizing networks of propaganda on twitter: a case study. Applied Network Science 5(1): 1–22.

Han, J.; Pei, J.; and Yin, Y. 2000. Mining frequent patterns without candidate generation. ACM sigmod record 29(2): 1–12.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 770–778.

Hochreiter, S.; and Schmidhuber, J. 1997. Long short-term memory. Neural computation 9(8): 1735–1780.

HosseiniMardi, H.; Ghasemian, A.; Clauset, A.; Mobius, M.; Rothschild, D. M.; and Watts, D. J. 2021. Examining the consumption of radical content on YouTube. Proceedings of the National Academy of Sciences 118(32).

Huang, B.; and Carley, K. M. 2019. Syntax-aware aspect level sentiment classification with graph attention networks. arXiv preprint arXiv:1909.02606.

Huang, Q.; Yu, J.; Wu, J.; and Wang, B. 2020. Heterogeneous graph attention networks for early detection of rumors on twitter. In 2020 International Joint Conference on Neural Networks (IJCNN), 1–8. IEEE.

Jia, X.; Zhao, H.; Lin, Z.; Kale, A.; and Kumar, V. 2020. Personalized image retrieval with sparse graph representation learning. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2735–2743.

Jung, S.-G.; An, J.; Kwak, H.; Salminen, J.; and Jansen, B. J. 2018. Assessing the accuracy of four popular face recognition tools for inferring gender, age, and race. In Twelfth international AAAI conference on web and social media.

Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems 25: 1097–1105.

Lewandowsky, S.; Jetter, M.; and Ecker, U. K. 2020. Using the president’s tweets to understand political division in the age of social media. Nature communications 11(1): 1–12.

Lin, H.; Ma, J.; Cheng, M.; Yang, Z.; Chen, L.; and Chen, G. 2021. Rumor Detection on Twitter with Claim-Guided Hierarchical Graph Attention Networks. arXiv preprint arXiv:2110.04522.

Loxton, M.; Truskett, R.; Scarf, B.; Sindone, L.; Baldry, G.; and Zhao, Y. 2020. Consumer behaviour during crises: preliminary research on how coronavirus has manifested consumer panic buying, herd mentality, changing discretionary
spending and the role of the media in influencing behaviour. *Journal of risk and financial management* 13(8): 166.

Lv, Q.; Ding, M.; Liu, Q.; Chen, Y.; Feng, W.; He, S.; Zhou, C.; Jiang, J.; Dong, Y.; and Tang, J. 2021. *Are We Really Making Much Progress? Revisiting, Benchmarking and Refining Heterogeneous Graph Neural Networks*, 1150–1160. New York, NY, USA: Association for Computing Machinery. ISBN 9781450383325.

Lyu, H.; Chen, L.; Wang, Y.; and Luo, J. 2020. Sense and sensibility: Characterizing social media users regarding the use of controversial terms for covid-19. *IEEE Transactions on Big Data*.

Lyu, H.; Fan, Y.; Xiong, Z.; Komisarchik, M.; and Luo, J. 2021a. Understanding Public Opinion Toward the #StopAsianHate Movement and the Relation With Racially Motivated Hate Crimes in the US. *IEEE Transactions on Computational Social Systems*.

Lyu, H.; Wang, J.; Wu, W.; Duong, V.; Zhang, X.; Dye, T. D.; and Luo, J. 2021b. Social media study of public opinions on potential COVID-19 vaccines: informing dissent, disparities, and dissemination. *Intelligent medicine*.

Lyu, H.; Zheng, Z.; and Luo, J. 2021. Both Rates of Fake News and Fact-based News on Twitter Negatively Correlate with the State-level COVID-19 Vaccine Uptake. *arXiv preprint arXiv:2106.07435*.

Mikolov, T.; Chen, K.; Corrado, G.; and Dean, J. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

O’Connor, B.; Balasubramanyan, R.; Routledge, B. R.; and Smith, N. A. 2010. From tweets to polls: Linking text sentiment to public opinion time series. In *Fourth international AAAI conference on weblogs and social media*.

Ozili, P. K.; and Arun, T. 2020. Spillover of COVID-19: impact on the Global Economy. *Available at SSRN 3562570*.

Panda, A.; Siddarth, D.; and Pal, J. 2020. COVID, BLM, and the polarization of US politicians on Twitter. *arXiv preprint arXiv:2008.03263*.

Pennacchiotti, M.; and Popescu, A.-M. 2011. A machine learning approach to twitter user classification. In *Fifth international AAAI conference on weblogs and social media*. Pew Research Center. 2015. 2014 U.S. Religious Landscape Study.

Reimers, N.; and Gurevych, I. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Roberts, M.; and Roberts, M. E. 2018. *Censored*. Princeton University Press.

Rosita, R. 2020. Panic buying in the Covid–19 pandemic era in Indonesia. *INTERNATIONAL JOURNAL OF MULTI SCIENCE* 1(07): 60–70.

Song, W.; Xiao, Z.; Wang, Y.; Charlin, L.; Zhang, M.; and Tang, J. 2019. Session-based social recommendation via dynamic graph attention networks. In *Proceedings of the Twelfth ACM international conference on web search and data mining*, 555–563.

Timmer, A.; Sery, J.; Connable, S. T.; and Billinson, J. 2018. A tale of two paranoias: a critical analysis of the use of the paranoid style and public secrecy by Donald Trump and Viktor Orbán. *Secrecy and Society* 1(2): 3.

Tumasjan, A.; Sprenger, T.; Sandner, P.; and Welpe, I. 2010. Predicting elections with twitter: What 140 characters reveal about political sentiment. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 4.

Urman, A. 2020. Context matters: political polarization on Twitter from a comparative perspective. *Media, culture & society* 42(6): 857–879.

Van der Maaten, L.; and Hinton, G. 2008. Visualizing data using t-SNE. *Journal of machine learning research* 9(11).

Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; and Bengio, Y. 2018. Graph Attention Networks. In *International Conference on Learning Representations*.

Waller, I.; and Anderson, A. 2021. Quantifying social organization and political polarization in online platforms. *Nature* 1–5.

Wang, K.; Shen, W.; Yang, Y.; Quan, X.; and Wang, R. 2020. Relational graph attention network for aspect-based sentiment analysis. *arXiv preprint arXiv:2004.12362*.

Weld, G.; Glenski, M.; and Althoff, T. 2021. Political Bias and Factualness in News Sharing across more than 100,000 Online Communities. *arXiv preprint arXiv:2102.08537*.

Wold, S.; Esbensen, K.; and Geladi, P. 1987. Principal component analysis. *Chemometrics and intelligent laboratory systems* 2(1-3): 37–52.

Xiao, Z.; Song, W.; Xu, H.; Ren, Z.; and Sun, Y. 2020. Timme: Twitter ideology-detection via multi-task multi-relational embedding. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2258–2268.

Yadav, S.; and Shankar, R. 2014. Inflation Expectation and Consumer Spending in India: Evidence from Consumer Confidence Survey. *Reserve Bank of India—Publications*.

Zhang, C.; Song, D.; Huang, C.; Swami, A.; and Chawla, N. V. 2019. Heterogeneous graph neural network. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 793–803.

Zhang, X.; Lyu, H.; and Luo, J. 2020. Understanding the hoarding behaviors during the COVID-19 pandemic using large scale social media data. *arXiv preprint arXiv:2010.07845*.

Zhang, Z.; Chen, T.; Zhou, Z.; Li, J.; and Luo, J. 2018. How to become Instagram famous: Post popularity prediction with dual-attention. In *2018 IEEE international conference on big data (big data)*, 2383–2392. IEEE.