Quantifying polarization within subreddit communities using distance-based measures

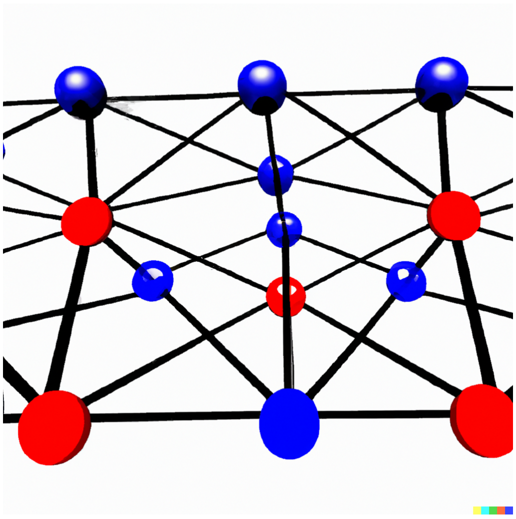
Christian Rasmussen, chur@itu.dk
Aidan Stocks, aist@itu.dk
João Vieira da Silva joap@itu.dk

Supervisors: Luca Rossi & Michele Coscia
Monday 15th May, 2023

IT-University of Copenhagen
Rued Langgaards Vej 7, 2300 København

Abstract

In recent years, the discourse circling around political polarization in the general public has been on the rise, with increased attention from the mainstream media on how the internet has affected the change. With this perceived increase in polarization, the importance of measuring and understanding this phenomenon has also increase. Within the study of polarization, an assortment of measures exist to attempt to quantify polarization. Though these measures have helped deepen our understanding of polarization, not all measures capture how direct communication between people affects our understanding of polarization. `Reddit.com` is a website which is often criticized for helping create echo chambers, by encouraging like-minded people to only interact with those harboring the same opinions as themselves which help greatly contribute to an increase in polarization. Here, we propose a method, which seeks to deepen our understanding of how direct communication affects polarization, which works within the technological framework presented by `Reddit.com`. This method can help support the existing polarization measures, while being aware that the output of this method is meant to be interpreted and used in combination with the existing methods within the research field.



Introduction

In recent years, the rise of social media as a prevalent platform for communication has prompted an increase in research on political polarization and its relationship to social media. The 2016 United States presidential election and the Covid-19 pandemic have further amplified public awareness of polarization in politics and society [FA08]. Online forums provide researchers with easily traceable sources of polarized views and beliefs. While there is existing literature concluding that political polarization is both increasing and decreasing, the issue is continuously discussed and researched when assessing online communication and its effect on society [Wal16] [Mic22].

Pre-existing literature has shown that communication on social media can limit exposure to counter-attitudinal news sources, resulting in increased polarization by creating echo-chambers where the same opinions and views circulate. This leads to individuals in these echo-chambers never being challenged or opposed to their views. It has also been shown that the emergence of more partisan media has helped increase polarization across society at large[Lev20] [Pri13].

In political science, the term ”polarization” can be subdivided into two sub-terms: ideological and affective polarization. Ideological polarization presumes that reduced dialogue between people with differing views leads to a divergence in ideological beliefs, while more dialogue results in more aligned ideological beliefs[Lel16]. On the other hand, affective polarization is more occupied with in-group favoritism and out-group hostility which leads to an increasing in structural polarity.[DL19]

The existing literature has not only shown that the Generalized Euclidean (GE) distance can be used to estimate distances between groups of nodes in a network [Cos20], but that it also works when applied to a social structure, with an attached ”opinion component” to apply this measure to an instance of political polarization [HDC23].

In this paper, we aim to build upon the existing framework for quantifying ideological polarization, by incorporating measures that seek to characterize aspects of a network which signify affective polarization. This is accomplished by examining the relationship between what online users post about, and how they respond to each other's posts. This approach helps quantify the relationship between ideological and affective polarization between two vectors in the same network, providing a better understanding of polarization in a given network.

With this study, the focus was specifically on social networks from `Reddit.com`, as the website's structure of numerous sub-boards (subreddits) each assigned to a different topic with its own or overlapping community of users, may implicate an inherent polarized structure. Furthermore, it may be inferable that some subreddits invite more polarized interactions than other subreddits. This is due to how some subreddits exist to post about a specific topic, whereas other subreddits focus more on having a discussion about a topic.

Method

Definition

When analyzing the structural polarity of a network, it is important to capture two specific components of polarization and their interplay [HDC23]. Those two being:

1. The opinion component, a numerical assessment of a person's political alignment, which in a balanced network with no polarization is centered and is heavily skewed towards opposite ends in a highly polarized network.

2. The Structural component, which characterize the social connections between people in communication, by conveying the flow of discussion and communication through edges in a network. If a network has no connections, sub-communities or cliques, then there is no opinion

homophily and no polarization. On the contrary, if a network has a lot of cliques and small communities, then the network is heavily polarized [BP21].

By adding weights to the edges connecting people in a network, we can construct a third component. We call this the *Relational Component Score*:

3. The *Relational Component Score*, used to analyze the correlation between the opinion-and structural components, to incorporate how communication between potentially disagreeing people affects polarization in a network.

Data Collection

The dataset utilized in this project consisted exclusively of select subreddits, most centered on various political topics, but also a couple of non-political subreddits for comparison. The choice of these subreddits was motivated by the inherent polarization of political opinions and discussions, which provides a suitable context for assessing polarization in social structures, and some alternative subreddits to make comparisons.

To collect the data, we employed the community developed PRAW (Python Reddit API Wrapper) tool, which facilitated the extraction of data directly from multiple subreddit pages.

To measure the correlation between polarization and opinion, we required data that encompassed both the opinion and sentiment components of users in a network. With this objective in mind, user relationships were characterized based on their polarity or opinion, encoded through the titles of all their posts, and the sentiment component, represented by the sentiment of their comments on other users' posts.

For all subreddits, the same parameters for data collection were used:

1. Top 1000 posts of the month are collected using the Reddit built-in "Top Month" sorting method.
2. Top 200 highest scoring comments per post are collected. All scraped comments being top-level comments, meaning that they are comments made directly to the post and not replies to other comments.

NLP processing

The natural language data collected in this project are processed in a pipeline of text transformations to help make the data more easily interpretable and meaningful in our analysis:

1. Sentences are transformed to exclusively lowercase letters
2. Expand Contractions (can't → cannot)
3. Removal of certain abbreviations
4. Removal of nonessential text (URLs, special characters, punctuation, non-ascii characters)
5. Other minor transformations such as internet slang to common language (w/e → whatever)
6. Post titles for users with multiple posts are concatenated into one long text

When determining the opinion component of a user, a pretrained huggingface module sentence-transformer was utilized. The module is initialized using the *bert-base-nli-mean-tokens* model which maps a given sentence into a vector of 768 dimension. The model is fine-tuned to read sentences and semantically analyze the comparison between sentences. The output of the used model does provide a precise mapping of the semantic meaning of a sentence, but for this project, a 1-dimensional score was required to represent the opinion component. To obtain this score, the 768-dimensional vector was reduced to a single dimension using PCA (principal component analysis) while retaining as much information as possible [RG19].

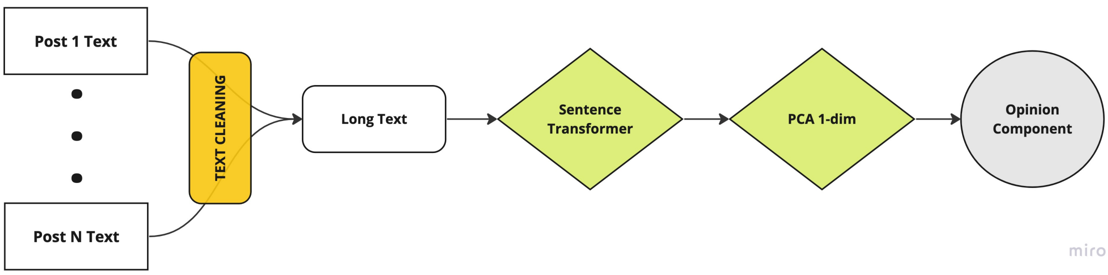


Figure 1: Illustration showing the full pipeline for how titles of the sampled posts are processed in order to create the opinion component.

Network Construction

For the construction of the networks, it was decided that each node would represent a unique user, and each edge an interaction where one user had commented on another user's post.

During creation, each node was assigned an attribute of the user's opinion component (the 1-dimensional embedding of their compiled post titles).

Edges between nodes in the network were constructed based on the direction of communication between users. Specifically, an edge was created from node A to node B if A had commented on a post made by user B.

For the *Relational Component Score* the edges needed weights, which were acquired by using a pretrained VADER [BKL09] sentiment classification model, which assigned a numerical sentiment score between $[-1, 1]$ to natural language sentences. This specific model was chosen after testing several different pretrained sentiment analysis models, after which VADER was chosen as it seemed to most accurately capture the sentiments of the textual data. The pipeline for how the edges are made is shown below in Figure 2.

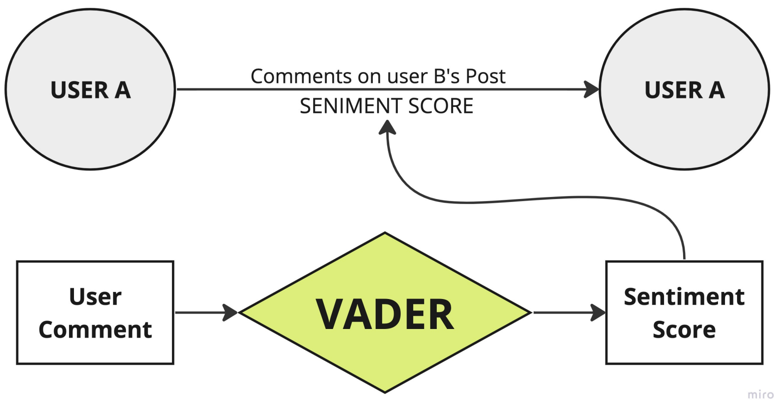


Figure 2: Illustration of edge weights are created and implemented in the network.

The choice to use pretrained models was made as creating and fine-tuning custom embedding and sentiment algorithms is outside the scope of this project and could elicit inferior results.

To summarize, the network details are as follows:

- Node - Represents a unique user that has made at least one post and left at least one comment on another user's post.
- Node Attribute - A users compiled post titles, transformed to word embeddings and reduced to a single dimension. (The opinion component).
- Edge (directed) - Represents that node A (user A) left a comment on node B's (user B's) post. (Edges are kept as directed, but for the purpose of using GE, they do not need to be directed)
- Edge Attribute - Sentiment of the comment text (ranging from $[-1, 1]$)

The existing framework adapted from [HDC23] works with the constraint that it must be fed a fully connected graph, which led to only the largest connected component within each network being used.

These networks contain the nodes and opinion components which use the GE-algorithm to calculate the ideological distance scores. To calculate the Relational Component scores these networks need to be modified, so to keep track of which networks are which we call these first networks *G1*.

Using Generalized Euclidean

The opinion component adapted from [HDC23] follows a set of network specific constraints and takes certain inputs. In this subsection we will outline the basics of how the distance measure is adapted in our project and how it works.

The biggest difference in the implementation of the measure is that we reject the imposed convention that opinions must be bounded between -1 and $+1$, and instead use the absolute score which we adapt directly from our NLP-pipeline.

The adaptation of Generalized Euclidean distance we use, calculates the Moore-Penrose pseudo-inverse of the opinion vector, this method of calculating distances between nodes has previously proven to work well within susceptible-infected-susceptible or susceptible-infected-recovered (SIR) models, to simulate infection spread.[HDC23] This method of calculating polarization works to effectively calculate how many "steps" nodes in a network are separated by. In our work, we directly adapt the coded functions on our own networks.

The next section will show how we adapt this code to create the Relational Component.

Network Modifications for Relational Component Score

When creating the Relational Component, the formulation for calculating the opinion component from [HDC23] was used in the following order:

Let G be an undirected networkx graph, where each node v in G has an opinion score $o(v)$ and each edge (u, v) has a sentiment score $s(u, v)$. Let $L(G)$ be the line graph of G, where each edge in G becomes a node in $L(G)$ and each node in G becomes an edge in $L(G)$.

Consider two dictionaries $D1(G)$ and $D2(G)$, let $D1$ contain a node for each edge in $L(G)$, let $D2(G)$ contain the absolute difference in opinion between every pair of nodes.

Calculate the effective resistance matrix of $L(G)$ denoted as $LQ(G)$.

Pass the two dictionaries $D1(G)$ and $D2(G)$ along with $L(G)$ and $LQ(D)$ into the supplied correlation function adapted from [HDC23].

The resulting score becomes the correlation between ideological distance and edge sentiment, which is the final score for each of the parsed networks. The score being a float between -1 and $+1$, with

-1 meaning that node pairs with high distance have negative edge sentiments, and $+1$ meaning that node pairs with high distance have positive edge sentiments.

Results

Subreddit	Scrape Date	#Nodes	#Edges	Average Shortest Path Length	Ideological Distance	Relational Component Score
Antiwork	April 17th	248	333	3.553742	95.982375	0.216976
Cooking	May 7th	317	499	3.793875	61.603669	-0.361591
Democrats	April 23rd	113	441	2.431416	38.203162	0.063780
UK Politics	April 23rd	303	1261	2.033047	39.539015	-0.008701
Political Discussion	April 23rd	55	95	2.744781	22.785946	-0.042481
Politics	March 23rd	326	874	3.287192	44.463131	-0.031926
Republican	April 23rd	52	116	2.643288	24.030078	0.035106
News	April 24th	264	463	3.657363	56.999132	-0.141636
World News	April 23rd	143	206	3.640599	48.764655	0.092681

Table 1: Statistics of the networks created from the given subreddits.

From Table 1 above it becomes apparent how much the sizes of the networks differ, and how some networks such as "UK Politics" have much more interaction (edges) than others.

The constructed networks for each subreddit end up varying in size due to the following factors:

- Users being included only if they have both made posts and comments. Some networks are more heavily affected by this than others.
- No graphs are naturally fully connected. Not all users interact with other users, which leads to them being removed from the final network. This varies randomly across subreddits.
- Subreddits varying in popularity. While all selected subreddits are relatively popular, there is still a varying amount of activity in each of them so there are not always 200 top-level comments to scrape on the less popular posts.
- Deleted comments. Some scraped comments contain only 'NaN' values if they were deleted. On Reddit, a comment can be deleted by the user themselves, a moderator of the subreddit, or a Reddit admin, but will remain in the comment section shown as "[deleted]". Due to how some subreddits are more strictly moderated than others, some subreddits will have more deleted comments than others.

These size differences could potentially lead to certain technical faults, like ideological distance or the *Relational Component Score* being directly affected by other variables, although we find that to not be the case due to the results of the project.

We find that within this sample of networks, there is a correlation between the ideological distance of a network and the *Relational Component Score*. As we note that *Antiwork* and *Cooking* have the two highest absolute values in both polarity measurements.

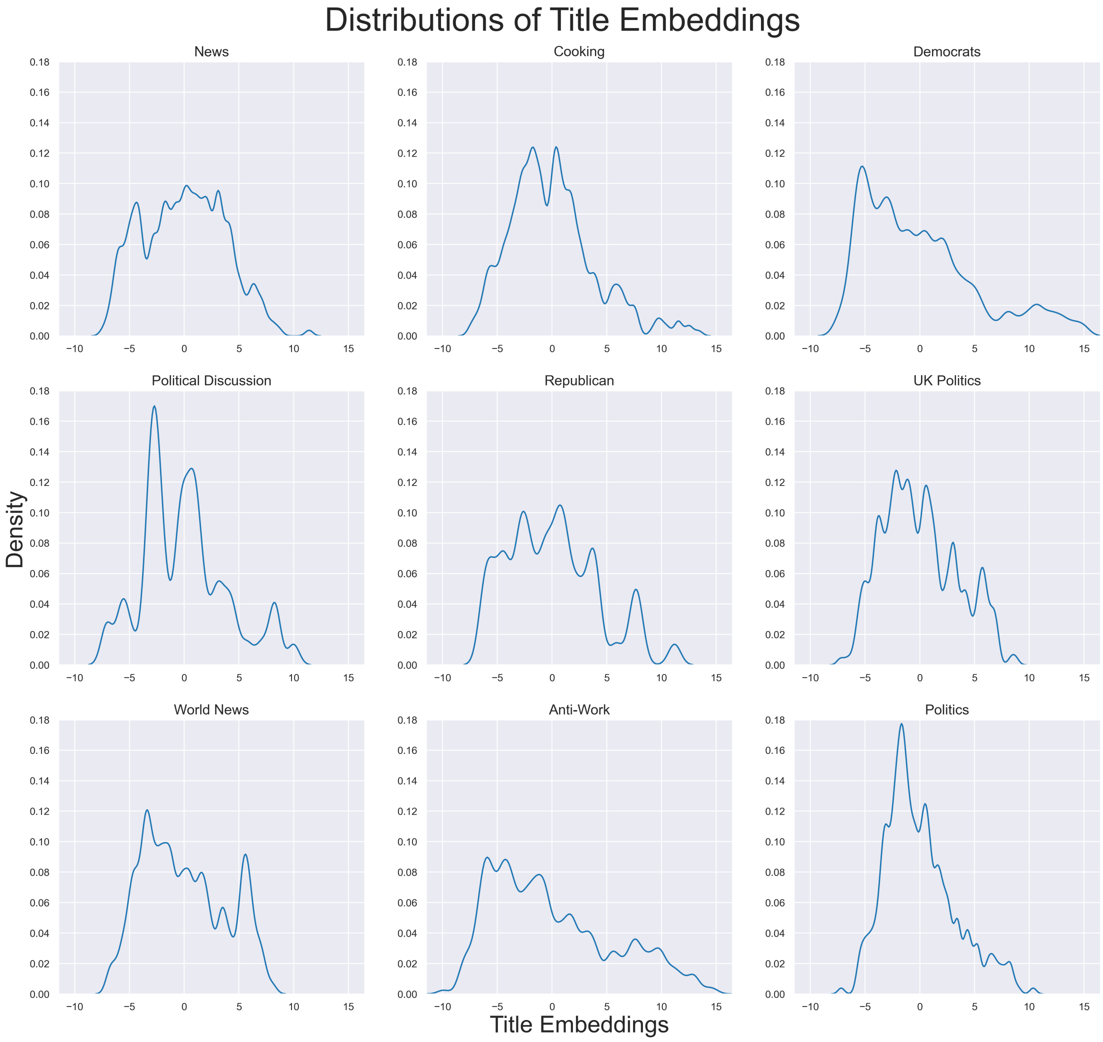


Figure 3: Illustration showing the distribution of the title embeddings for each subreddit network.

The title embedding distributions from figure 3 show how in some networks users tend to use more similar language than in other networks. This can be seen in figure 3 with how the *Politics* subreddit has a narrower and taller distribution compared to the *Antiwork* subreddit. In our case we can think of the embeddings as the interests of the poster, as it represents what it is they are posting about (since the embeddings characterize the language used in the post).

Also note, that some networks contain embedding score instances in the range of $[10, 15]$ where other networks seem to have an upper limit closer to 10. This tendency is not limited to networks within a certain theme, as it seems to be decided on a case-by-case basis. Since we reject the imposed convention of using a $[-1, +1]$ range for the opinion component, this is a result which we must live with as is, since we cannot reasonably tailor the ranges without having a thorough understanding of what each extreme of the scores represent.

Figure 4 shows the distribution of the comment sentiments scores from the Vader sentiment classifier. Generally, we observe even distributions, with some networks having flatter distributions (UK Politics) and some having high negative and positive (Political Discussion). Most networks contain a high number of neutral sentiments, most likely due to users responding to non-opinionated topics or simply engaging in factual discussions, where opinions don't always matter. Some commenters will also add to the discussion on a given post, by contributing with a non-opinionated statement, meant to add contextual information and explain a post, thereby interacting factually with a post rather than expressing a certain sentiment towards the post.

Some of this can also be explain through users interacting with a post without using emotive language, using wordings which downplay any strong sentiments. Online communication can reduce the level of elaboration in expression, which can affect how strong sentiment may be interpreted.

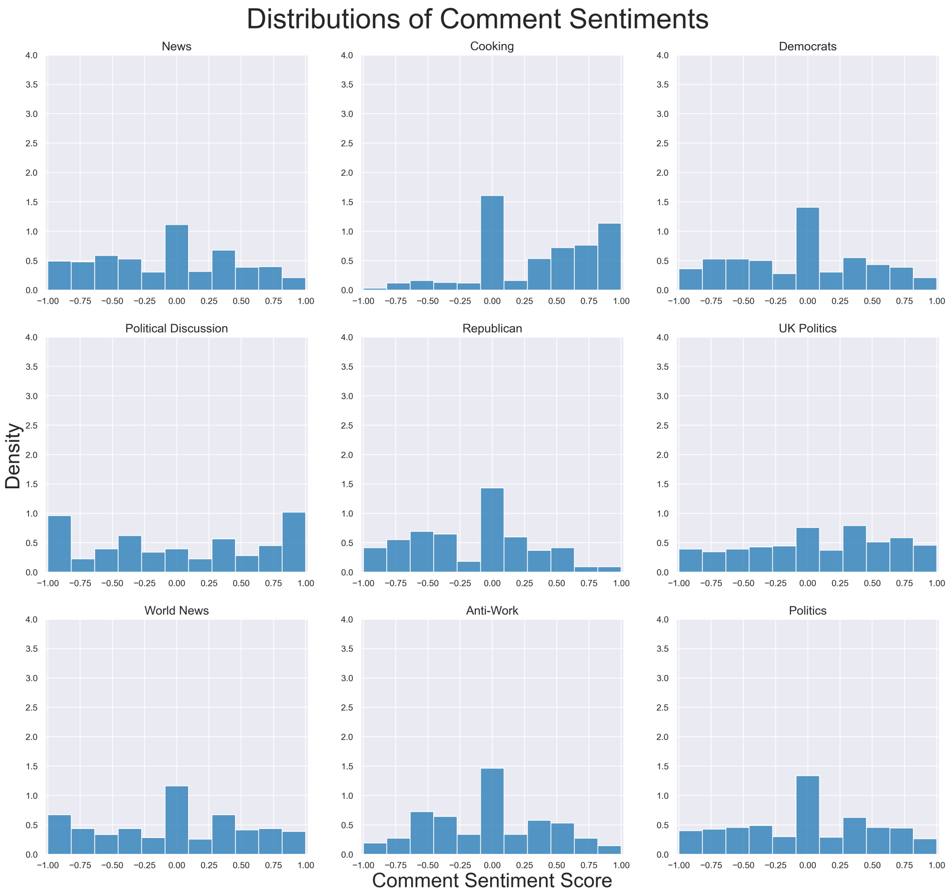
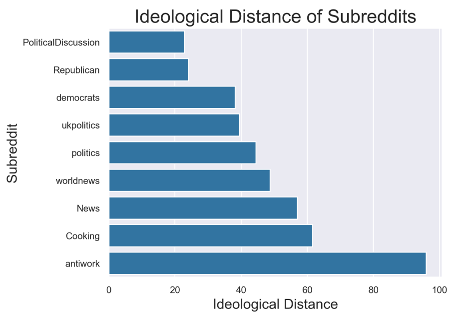
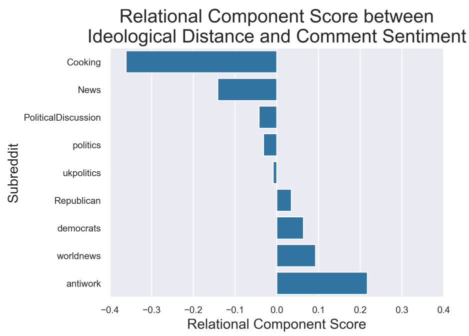


Figure 4: Illustration showing the distribution of comment sentiments for each subreddit, where a sentiment of -1 is negative, +1 is positive, and 0 is neutral.

Most of the networks contain almost equal amounts of positive and negative-leaning comments, with the notable exception of Cooking, which almost exclusively contains neutral- or positive comments.



(a) Ideological distance of each subreddit.



(b) Relational Component Score of each subreddit.

Figure 5

Figure 5a shows the distance between the word embeddings of users that interact with each other. A higher ideological distance indicates that interacting users' posts are dissimilar, while a lower score indicates that interacting users' posts are similar. Interestingly some of the more similarly themed subreddits have similar ideological distances, like *worldnews* and "news", *ukpolitics* and *politics*. Antiwork also has a much higher ideological distance than other subreddits.

Figure 5b measures the correlation between the word embeddings of interacting users and the sentiment of the interacting comment. A positive score means that the greater the ideological distance, the more positive the comment sentiment, while a negative score means that the greater the ideological distance, the more negative the comment sentiment. Within our used framework a higher score means that users comment positively on reddit posts that are different to their own, and a lower score means that users comment negatively on posts different to their own. A score of zero means that how different a post is to a user's own posts has no effect on the sentiment of the comment they will leave.

From figure 5b it generally makes sense that the correlations are closer to zero, as we wouldn't expect how different a post is to a user's own post to have an extreme effect on the sentiment of the comment a user will leave. We can also see an interesting difference with the subreddits "News" and "World News", as they have similar ideological distances, but opposite Relational Component Scores.

Figure 6 shows that there is no real strong correlation between the two measures. When we created the measure, we expected to see some degree of correlation between a high embedding distance and negative comment sentiments and that closely connected networks would elicit positive sentiments.

By looking at specifically *antiwork* and *Cooking*, we notice an overall increase in ideological distance, results in a higher effect on the *Relational Component Score*, but mostly when the ideological distance is heavily increased is with these two networks. This also holds true for the two news-themed networks (worldnews News), but not to the same degree.

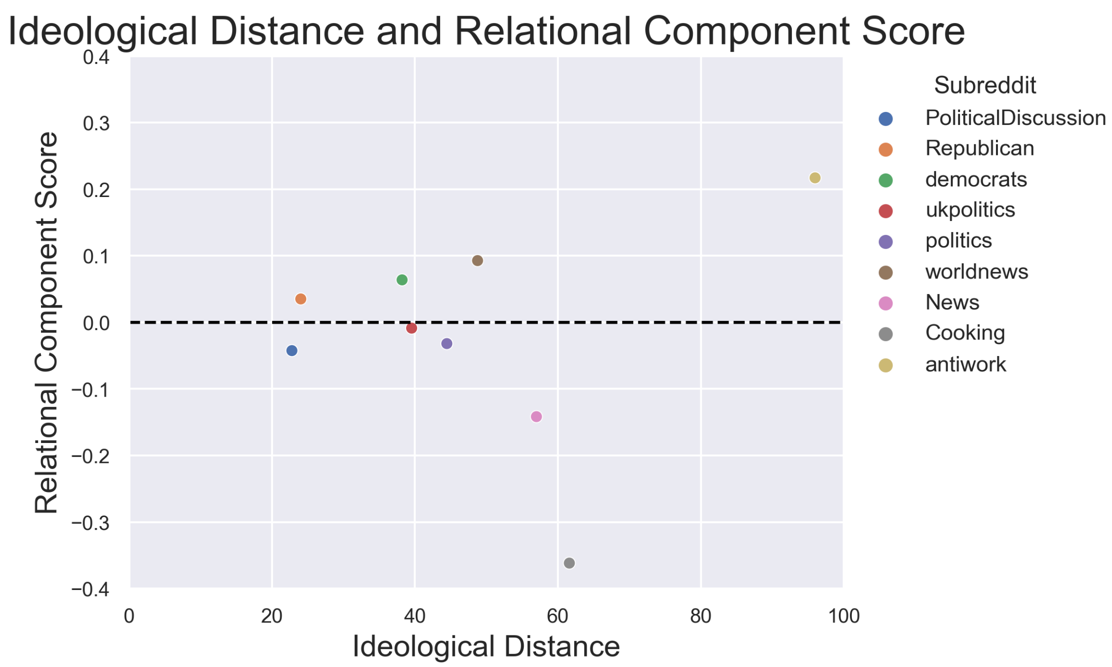


Figure 6: Illustration showing both the ideological distance and Relational Component Score of each subreddit.

Discussion

The scope of this project was to build upon the existing frameworks for how to measure polarity within a community complimentary to using the Generalized Euclidean measure. While the implementation of this method is made to assess any community structure with opinions and communication, the methods and results of this project, entail that the conclusions of the project are drawn within the specific boundaries of the project. This is in large part due to certain technical constraints presented by the data source, which means we must make certain amends when discussing the results of the project.

Main Technical Constraints

One of the biggest technical challenges of this project lie in utilizing the PRAW module to interface with the Reddit API. While this module offers accessible tools to facilitate data interaction and quantification, it relies on undisclosed mechanisms to accomplish these tasks. Consequently, setting a precise timeframe for data collection is partially unattainable, as the module only supports sorting mechanisms available on the Reddit website for requesting data within a timeframe.

Thus, the "Top (month)" sorting mechanism was employed uniformly across all subreddit pages, rendering it impossible to analyze posts from a particular subreddit during a designated time period, such as r/politics during the 2020 US presidential elections, when polarization would be anticipated to be discernibly high. This limitation significantly impacts the quality of the data utilized and complicates the characterization of specific aspects of the output data.

Challenges when trying to distil opinions

When capturing the Opinion component for this project, the use of Reddit-based data and the implementation of it does also carry some implications. The opinion component in its current form cannot directly capture the opinion of an individual, as it simply takes the combined titles of posts by a user and returns a numerical evaluation of the natural language. This approach does allow for cross-analysis of different topics/domains of subreddits, as it does not look for specific word-corpora related to politics but is able to read data from any subreddit based in any topic. Leaving the requirement of needing explicitly opinionated behind does allow for a broad implementation of this polarization method, but also means that it well never accurately capture the opinions of users.

This method of assigning the opinions to users also opens the doors for online "*trolls*" to affect the measure. If a bad actor is seeking to agitate users of a certain subreddit into responding negatively, they may post something which they expect to elicit strong reactions, even the poster may not hold the opinion detected through the post. The trouble deepens when considering the fact that the opinions of the commenters to a certain "*troll*" post may still manage to accurately capture the opinions of the commenters, yet it still leaves an interpretational uncertainty.

The opinion component to each of the edges in the network are also solely based on a pretrained sentiment analysis model, which has in no way been fine-tuned or fitted to the current "domain" of Reddit comments. This means that we rely heavily on how this model interprets the language fed, which we have no influence over and could have unforeseen consequences pertaining to how communication in online forums might differ to regular day-to-day language or written journalist language.

On Using PCA

When designing the study, we decided to implement a 1-dimensional PCA-reduced mapping of the computed sentence-embeddings, as the implemented GE-algorithm only takes 1-dimensional values. This meant that we by nature had to reduce a lot of the meaning derived from the sentences. When implementing PCA, the algorithm does retain as much variance within the data as possible, but when reducing a 768-dimensional object to a single dimension, we can expect to lose a lot of valuable information within the process. The sentence embeddings do work to characterize the most important aspects of its inputs, but in this case where we are not even 100 percent sure what aspect of the used language has the biggest effect on the embeddings, we do lose valuable information.

On Grouping Posts by User

We also need to address the fact, that the opinion component, assigned to each node within the networks, are grouped by each user, and contain one common embedding for every post by said user. This opens the doors for possible diminishing returns, as a user might balance themselves within the distribution of "opinions" by weighing heavily in both extremes. In this example we do not distinguish between users who have multiple posts or incorporate a balancing mechanism like a dampening score, since it is outside the scope of the project. Nonetheless, it still needs to be noted and taken into consideration when assessing the output of the project.

Challenges when assigning Opinions

It is also important to notice that a negative response to a post does not necessarily mean a disagreement between commenter and poster, as the poster themselves may also have a negative opinion about the topic of their own post. I.E., it is important to realise that the comment is a reaction to the post, and not the poster.

Post Title	Republicans are not pro life. And they need to be stopped immediately. They're anti humanity
Comment A	texas wants to be like australia but instead of animals trying to kill you it is republicans
Comment B	you guys do not do your research at all the bill does not ban water breaks

Table 2: Example of a post from Antiwork and two comments left on it.

Table 2 above shows an example where this distinction is important. The poster has a clear stance against republicans, but while comment A agrees with the poster, it has negative sentiment due to the negative language used. Furthermore, comment B disagrees with the poster and is classified as partially positive. This example shows how the comments lead to a visible disagreement on the post itself but does not necessarily imply that the two commenters have opposing ideologies.

Also note, that the data in this study is based on users who fit a very specific criteria within their community. It can be inferred that users who appear in these networks are all active on a fairly regular basis, as some subreddits have user requirements before posting and that for a user to appear in the network they need to both create a post and interact with another post. I.E., the opinions are sampled from only select active users, which means that the opinions expressed in the comments, may differ in different degrees to the overall sentiments from the subreddit as a whole.

The Relational Component Score

The main contribution of this project is what we define as *The Relational Component Score* which estimates how the direct interactions between nodes is correlated to the *Generalized Euclidean Distance* of the network. This component is implemented to capture a higher degree of affective polarization to work in tandem with *Generalized Euclidean.* From the results of testing the measure on the sampled networks, it becomes clear that the *Relational Component Score* is not directly affected by *GE.* As per 6 it is shown that there is no direct relation between the two. We can interpret a fair bit on the results though.

Attempting to find some logic in the results of the *Relational Component Score*, we can assess how similar the scores for each of the politically themed networks end up. We find that most of these networks (democrats, Republican, politics, ukpolitics, PoliticalDiscussion) fall within the range of

$[-0.042481, 0.063780]$, which we can attribute to the similar topics from these networks. On the Contrary, the two major News-themed subreddits (News, worldnews) have a higher difference in the *Relational Component Score* $[-0.141636, 0.092681]$ than any of the previously mentioned.

We speculate that this difference in outcome is in some degree affected by the posts from the dataset invites inherently positive or negative opinions. While the topics presented and discussed in politically themed networks may invite differing opinions, that might be different for news themed networks.

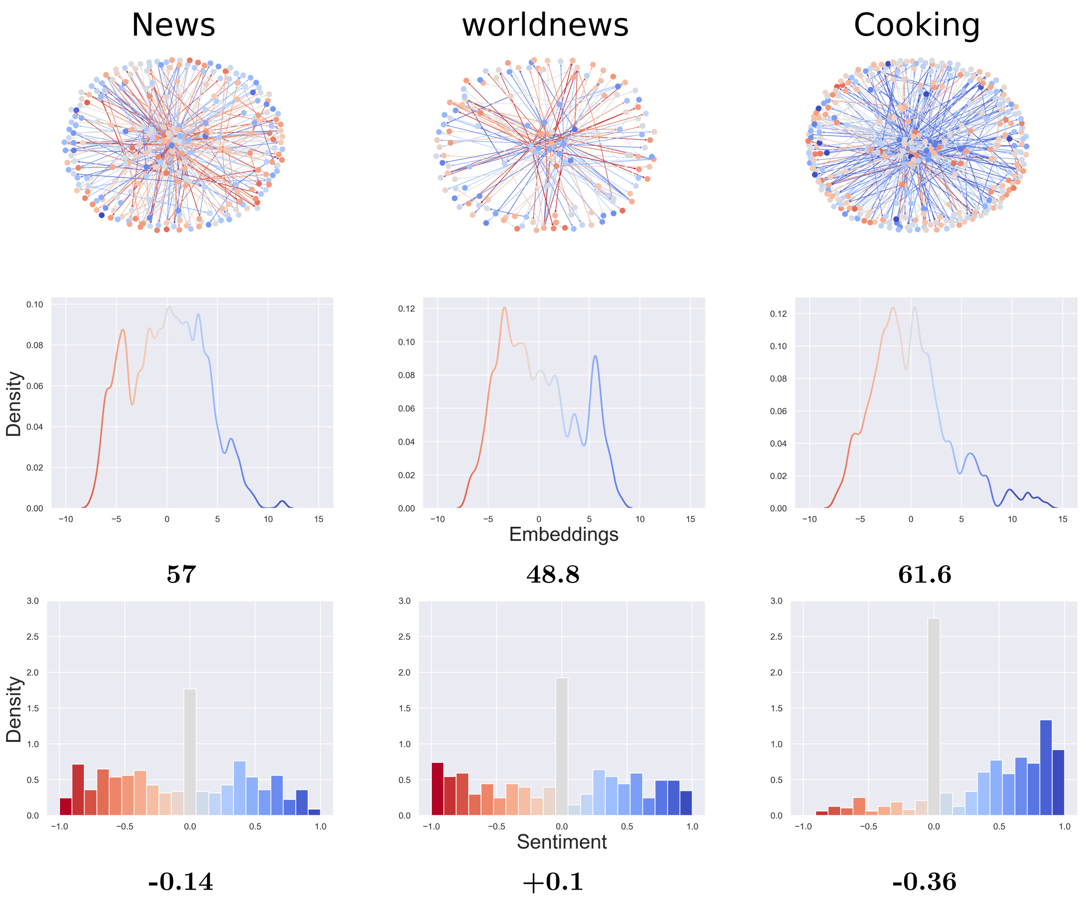


Figure 7: General stats and visualizations for 3 different networks. **Row 1**: Graph visualizations of each network, colors for nodes and vectors are color-coded to match the plots in row 2 and 3. **Row 2**: Density-plot showing the distribution of the NLP-based embeddings, below each density-plot is the calculated General Euclidean score for each network. **Row 3**: Histograms showing the distributions of comment sentiments for each of the networks, below each histogram is the calculated *Relational Component Score* for each network.

Figure 7 compares the distributions of sentiments, embeddings and the structure of three select networks, to show how the effect these have on GE and RCS are difficult to readily interpret and characterize. None of the networks seem to have any real cliques which makes visualizing how polarized the different networks are difficult.

The most significant interpretation we can make from Figure 7, is by assessing Cooking and concluding that it gains a fairly low RCS from the assumption that negative sentiment edges occur more frequently when sentiment difference is higher than in worldnews or News. We can make certain assumptions about how this could be, like saying that users who are interested in different food cultures or cooking styles, may have different opinions towards each other's interests. Yet this example also shows the interpretational barrier presented by using the embedding scores in their current iteration, as we can't fully grasp what aspect of the language used they capture most. This is even harder to interpret when the topics discussed don't invite a strictly two-sided argument, like in cooking. The motivation for a post in `r/cooking` may be broader than in for example `r/Politics`.

Going back to figure 3, we can see how all of the subreddits embedding distributions differ, such as *Politics* and *Antiwork*. In this example, the difference here could potentially be due to an important political event occurring in the US. If an important political event occurred, it could potentially lead to many users on *Politics* posting about it, which would lead to the embeddings being more similar. Without going into the subreddit and manually reading the posts it is uncertain whether the embeddings have been affected by such an event, or if the users on the subreddit naturally talk about more similar topics compared to *Antiwork*.

Using No Null Model

In this project we focus specifically on implementing the *Relational Component Score* in social networks solely based on communication on Reddit. While the results of this project do seem to capture this aspect of polarization, the results must be analyzed in isolation, as it is outside the scope of this project to create any null models and test how this component reacts to synthetic data. This project is meant to specifically analyze and validate this component on Reddit-based networks.

Concluding remarks

Through the development and implementation of the *Relational Component Score* with this specific Reddit-based database, we conclude that though the developed component does capture some aspect of polarization, which is not captured in previously developed frameworks, there is a thick interpretational barrier to the results of this project. The output of the *Relational Component Score* must be taken in pair with a qualitative assessment of the data which the component is used to analyze. We display this in our own data collection, as the output of the *Relational Component Score* adapts itself differently to the domains it is implemented in. The *Cooking* network displays this, as the output of the *Relational Component Score* seems to suggest that the network is more polarized than the politically themed networks, which would seem counter intuitive.

Overall, as this project shows, polarization as a field of study, encompasses many different aspects and the prospect of quantifying polarization can present many challenges. This project seeks to support the existing understanding of polarization and explore another method of quantifying polarization through this specific framework, yet we ultimately conclude that the methods developed in this project need much interpretation to help strengthen our understanding of polarization.

From the results of this project, we conclude that applying this method of characterizing polarization in social networks should be explored. This project has some major constraints due to the framework in which the experiments presented are conducted. If the model was tested on a network with a proven degree of structural polarization, where opposing views are guaranteed, assessing to which degree the *Relational Component Score* helps deepen our understanding of polarization may have been clearer. From the results of this project, we must speculate and make our best guesses.

Acknowledgements

Author Contributions: L.R. and M.C designed the study, A.S., C.R. and J.S. performed the analysis, A.S., C.R. and J.S. prepared the figure. A.S., C.R. and J.S. wrote the manuscript.

Data and materials availability: All code and data are available in the supplementary code repository. [Aid23]

References

[Aid23] Joao Viera da Silva Aidan Stocks Christian Hugo Rasmussen. *Bachelor Project*. `https://github.com/hugzito/bachelor_project`. 2023.

[BKL09] Steven Bird, Ewan Klein, and Edward Loper. *Natural language processing with Python: analyzing text with the natural language toolkit.* " O'Reilly Media, Inc.", 2009.

[BP21] Delia Baldassarri and Scott E. Page. "The emergence and perils of polarization". In: *Proceedings of the National Academy of Sciences* 118.50 (2021), e2116863118. DOI: `10.1073/pnas.2116863118`. eprint: `https://www.pnas.org/doi/pdf/10.1073/pnas.2116863118`. URL: `https://www.pnas.org/doi/abs/10.1073/pnas.2116863118`.

[Cos20] Michele Coscia. "Generalized Euclidean Measure to Estimate Network Distances". In: *Proceedings of the International AAAI Conference on Web and Social Media* 14.1 (May 2020), pp. 119–129. DOI: `10.1609/icwsm.v14i1.7284`. URL: `https://ojs.aaai.org/index.php/ICWSM/article/view/7284`.

[DL19] James N Druckman and Matthew S Levendusky. "What Do We Measure When We Measure Affective Polarization?" In: *Public Opinion Quarterly* 83.1 (May 2019), pp. 114–122. ISSN: 0033-362X. DOI: `10.1093/poq/nfz003`. eprint: `https://academic.oup.com/poq/article-pdf/83/1/114/28683759/nfz003\_suppl\_online\_appendix.pdf`. URL: `https://doi.org/10.1093/poq/nfz003`.

[FA08] Morris P. Fiorina and Samuel J. Abrams. "Political Polarization in the American Public". In: *Annual Review of Political Science* 11.1 (2008), pp. 563–588. DOI: `10.1146/annurev.polisci.11.053106.153836`. eprint: `https://doi.org/10.1146/annurev.polisci.11.053106.153836`. URL: `https://doi.org/10.1146/annurev.polisci.11.053106.153836`.

[HDC23] Marilena Hohmann, Karel Devriendt, and Michele Coscia. "Quantifying ideological polarization on a network using generalized Euclidean distance". In: *Science Advances* 9.9 (2023), eabq2044. DOI: `10.1126/sciadv.abq2044`. eprint: `https://www.science.org/doi/pdf/10.1126/sciadv.abq2044`. URL: `https://www.science.org/doi/abs/10.1126/sciadv.abq2044`.

[Lel16] Yphtach Lelkes. "Mass Polarization: Manifestations and Measurements". In: *Public Opinion Quarterly* 80.S1 (Mar. 2016), pp. 392–410. ISSN: 0033-362X. DOI: `10.1093/poq/nfw005`. eprint: `https://academic.oup.com/poq/article-pdf/80/S1/392/6864582/nfw005.pdf`. URL: `https://doi.org/10.1093/poq/nfw005`.

[Lev20] Ro'ee Levy. "Social Media, News Consumption, and Polarization: Evidence from a Field Experiment". In: *Mass Communication eJournal* (2020).

[Mic22] Luca Rossi Michele Coscia. "How minimizing conflicts could lead to polarization on social media: An agent-based model investigation". In: (2022). URL: `https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0263184`.

[Pri13] Markus Prior. "Media and Political Polarization". In: *Annual Review of Political Science* 16.1 (2013), pp. 101–127. DOI: `10.1146/annurev-polisci-100711-135242`. eprint: `https://doi.org/10.1146/annurev-polisci-100711-135242`. URL: `https://doi.org/10.1146/annurev-polisci-100711-135242`.

[RG19] Nils Reimers and Iryna Gurevych. "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, Nov. 2019. URL: `http://arxiv.org/abs/1908.10084`.

[Wal16] Cass R. Sunstein Walter Quattrociocchi Antonio Scala. In: (2016). URL: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2795110`.