



Master's thesis
Master's Programme in Data Science

Analysing Controversy on Twitter via Graph Embeddings

Andrei-Daniel Comănescu

September 7, 2020

Supervisor(s): Michael Mathioudakis

Examiner(s): Michael Mathioudakis
Kai Puolamäki

UNIVERSITY OF HELSINKI
FACULTY OF SCIENCE

P. O. Box 68 (Pietari Kalmin katu 5)
00014 University of Helsinki

Tiedekunta — Fakultet — Faculty		Koulutusohjelma — Utbildningsprogram — Degree programme	
Faculty of Science		Master's Programme in Data Science	
Tekijä — Författare — Author			
Andrei-Daniel Comănescu			
Työn nimi — Arbetets titel — Title			
Analysing Controversy on Twitter via Graph Embeddings			
Työn laji — Arbetets art — Level		Aika — Datum — Month and year	Sivumäärä — Sidantal — Number of pages
Master's thesis		September 7, 2020	51
Tiivistelmä — Referat — Abstract			
<p>Social networks represent a public forum of discussion for various topics, some of them controversial. Twitter is such a social network; it acts as a public space where discourse occurs. In recent years the role of social networks in information spreading has increased. As have the fears regarding the increasingly polarised discourse on social networks, caused by the tendency of users to avoid exposure to opposing opinions, while increasingly interacting with only like-minded individuals. This work looks at controversial topics on Twitter, over a long period of time, through the prism of political polarisation. We use the daily interactions, and the underlying structure of the whole conversation, to create daily graphs that are then used to obtain daily graph embeddings. We estimate the political ideologies of the users that are represented in the graph embeddings. By using the political ideologies of users and the daily graph embeddings, we offer a series of methods that allow us to detect and analyse changes in the political polarisation of the conversation. This enables us to conclude that, during our analysed time period, the overall polarisation levels for our examined controversial topics have stagnated. We also explore the effects of topic-related controversial events on the conversation, thus revealing their short-term effect on the conversation as a whole. Additionally, the linkage between increased interest in a topic and the increase of political polarisation is explored. Our findings reveal that as the interest in the controversial topic increases, so does the political polarisation.</p> <p>ACM Computing Classification System (CCS): Human-centered computing → Collaborative and social computing → Collaborative and social computing design and evaluation methods → Social network analysis Mathematics of computing → Discrete mathematics → Graph theory → Spectra of graphs Computational methodologies → Machine Learning</p>			
Avainsanat — Nyckelord — Keywords			
graph embedding, social networks, graphs, controversy analysis			
Säilytyspaikka — Förvaringsställe — Where deposited			
Muita tietoja — Övriga uppgifter — Additional information			

Contents

1	Introduction	1
2	Related Work	5
2.1	Echo Chambers	5
2.2	Analysis of Twitter Interactions	7
3	Background	11
3.1	Estimating Political Ideologies	11
3.1.1	Bayesian Point Estimation	12
3.2	Graph Embeddings	13
3.2.1	Laplacian Eigenmaps	13
3.2.2	GraphSAGE	14
4	Experiments	17
4.1	Data Description	18
4.1.1	Daily Graph Embeddings	19
4.1.2	Assigning Political Ideology	20
4.1.3	Data Overview	22
4.2	Trends and External Events Analysis	26
4.2.1	Synthetic Data	28
4.2.2	Inter-cluster Distance	31
4.2.3	Intra-cluster Distance	33
4.2.4	k%-Closest Neighbours Political Ideology Correlation	34
4.2.5	Closeness of Politically Aligned Users	37
4.3	Analysis of User Activity Levels	39
4.3.1	Inter-cluster Distance	41
4.3.2	Intra-cluster Distance	41
4.3.3	k%-Closest Neighbours Political Ideology Correlation	42
4.3.4	Closeness of Politically Aligned Users	43

5	Conclusions	45
5.1	Limitations and Future Work	45
	Bibliography	47

1. Introduction

Social media are platforms where a multitude of online discussions, about a variety of topics are debated. Multiple types of social media exist, with social networks being one of them. Their main distinctive characteristic from other types of social media is their emphasis on user' interactions with one another. The content that exists on these platforms is generated as a result of the interactions among users. Twitter is such a social network; it acts as a public space on which discourse occurs. In the recent years, social media, and social networks in particular have become more relevant in the spread and discussion of information.

Previous studies have shown that on social media, users have a strong preference to interact with like-minded individuals, thus leading them to be predominately exposed to views and opinions that are in accord with their own (see e.g. [7, 20, 21]). This exposure to only views that are in line with ones own is detrimental as a whole, as it enforces a belief system without any kind of counterbalance to check its overall validity. The tendency of users to become isolated from opposing opinions, while interacting with like-minded individuals, was studied for various forms of social media such as blogs (see e.g., [23, 29]), online newspapers (see e.g., [18, 26]) and social networks (see e.g. [2, 3, 7, 21]). Given the widespread prevalence of this phenomenon in social media, linked to its ever increasing popularity, can lead one to fear for an ever more increasingly divided society. Hence, the divide created by controversial topics should be analysed. This would allow both society and policy makers to take measures in the eventuality in which the divide between sides is ever more increasing.

In this work we analyse the evolution, over a number of years, of the discourse on Twitter in regards to controversial topics, from the prism of political polarisation. The analysed controversial topics are not political in nature in on themselves. By that we mean that the debate around these topics does not stem from political affiliation, but instead, it stems from the divergence in values and beliefs regarding the manner in which the opposing sides of the controversy think that their society should function.

For our analysis, for each controversial topic, we create daily graphs using the interactions and the underlying structure of the conversation as a whole. In the daily graphs, the users are depicted as nodes and the interactions as edges. These daily

graphs are then used to create daily graph embeddings. These embeddings highlight the divide across the pro and anti sides of the conversation. As our analysis is based on political ideology, and not on the pro and anti stances found in a controversial topic, we estimate the political affiliation of the users. This is done based on the overall structure of the conversation. Based on the estimated political affiliation, a user can either be left-leaning or right-leaning.

Using the graph embeddings, and the political ideology of the users, we are able to observe the long-term trends in the conversation's polarisation. We analyse the shifts in polarisation when the conversation is faced with external (i.e., they did not originate from the Twitter interactions) events relevant to the controversial topic. Also, we look at users tendencies when activity increases by observing the relationship between user activity levels and the overall polarisation of the conversation. We also look at the degree in which the left-leaning users and the right-leaning users match the pro and the anti side of a debate. In that regard we observe that each side of the political divide favours a distinct side of the controversial debate.

Previous works have been done to analyse controversy on Twitter (see e.g., [14, 20, 30]), with the closest related work being that of Garimella et al. [20]. In their paper, they analyse each topic from a pro or anti said topic perspective. E.g., if the controversial topic was *abortion*, then the two analysed sides were the pro-abortion side and the anti-abortion one. Their analysis was performed for multiple years of Twitter interactions for multiple controversial topics. The data used in this work was provided by them.

Even though our analysis stems from the same original data, we are analysing a different facet of these controversies by looking at them mainly from the prism of political ideology. Our analysis also differs from theirs in the manner in which the users' interactions are modeled. They rely on daily retweet graphs and daily reply graphs while we create daily graphs from all the interactions. The overall structure of our daily graphs also differs; in their daily graphs they include only the interactions from that day. On the other hand, our daily graphs consider all the interactions among users, for all the days of the conversation. In our daily graphs, each edge that models the interactions between two users is weighted in accordance to the number of interactions between said two users during that day. If no interactions have occurred, then a dummy weight is used instead. This allows us to analyse the daily debate structure from the perspective of the whole conversation. Their analysis is performed on graphs while ours relies on graph embeddings. To the best of our knowledge, we are not aware of any other work that relies on graph embeddings to analyse long-term controversial topics on social networks.

This work is structured as follows: in Chapter 2 we examine works related to the

analysis of controversy on social media, with a particular emphasis on social networks. Chapter 3 explores the main theoretical constructs that are used for the experimental part of this work. In Chapter 4 we present the preprocessing performed on the data and the analysis of Twitter interactions. Conclusions are drawn in Chapter 5.

2. Related Work

This chapter examines studies related to the analysis of controversy on social media, more specifically, the polarisation of social media when it comes to controversial topics. To the best of our knowledge, the works related to the aforementioned topic model the social media interactions with the help of graphs. Subsequently, the bulk of the analysis revolves around the structural properties of the resulting graphs.

2.1 Echo Chambers

Homophily represents the tendency of closely resembling individuals to associate with one another. In the case of networks, homophily manifests itself as the tendency of close nodes to also exhibit values close to one another [21]. When one refers to social media, echo chambers represent situations in which users end up being exposed to only content that is in accord with their views. One could regard echo chambers as a form of homophily. Two distinct components typically form an echo chamber, the view on a given subject and the medium in which it is shared [21]. In the case of social media, the view is represented by content, while the medium is the platform on which the sharing occurs. If the platform is a social network such as Twitter or Facebook then the content is represented by tweets and posts, while when the platform is a blog or a newspaper the articles act as content.

Multiple studies analyse the presence and the characteristics of echo chambers on various forms of social media. For example, the presence of echo chambers in blogs is studied in [23, 29], their presence in the context of online news is studied by [18, 26], while they are discussed in the context of social networks in [2, 3, 7, 21]. The aforementioned studies tend to discuss political echo chambers in a binary setting where the opposing views are depicted by either liberals and conservatives, or by *Democrats* and *Republicans*.

In [23], the paper determines the overall nature of blogs as echo chambers. This is done by analysing the comments corresponding to the world's most popular blogs. They notice that when comments take an ideological stance on a blog's article they are overwhelmingly in support of the author's view. When it comes to blog readers, the

majority of them consume blogs that cater to their political ideology [29]. As a whole, blog readers are more involved in politics than their non-blog reading counterparts and they show a higher degree of political polarisation. In general, political blogs are divided by cross-ideological lines [29].

The study of echo chambers when it comes to online newspapers revolves around the political ideology of the publication that created the news piece. In the case of [18], the paper relies on the web browsing activity of users to determine with which online newspapers they engage with frequently. For each user, the accessed political articles are separated into news and opinion pieces, the polarity of said articles is then determined based on the overall polarity of the publication that hosts them. By measuring the political divide between distinct users, they determine that articles accessed either via web-search or via social media are more divisive than those accessed by directly visiting the news website. When it comes to web browsing activity, most of the traffic to the news websites originated from direct access. Hence, the effect of echo chambers was somewhat mitigated. A controlled environment was used by [26] to determine that users preferred to read news coined to belong to news organizations of a political ideology similar to their own. This behaviour was found not only in the case of politically polarising topics but also in the case of non-controversial ones. It was also revealed that the users that were more active politically were more likely to prefer only partisan news outlets when compared to their less politically engaged counterparts.

More and more of our exposure to news is mediated through online social networks. At the same time, due to the advancements in technology, the risks have increased of them creating information bubbles that feed into a personalized narrative that foster further network engagement. Facebook and Twitter are two such social networks in which the existence and the particularities of echo chambers were explored. In the case of Facebook, the paper [2] explores echo chambers through the prism of information sharing. Their paper determines that Facebook users, even though they might be exposed to content belonging to both sides of the political aisle, tend to share only the pieces on news that correspond to their political ideology. They observe that this kind of behaviour occurs in the same fashion on Twitter as well. Also, by using the Twitter interactions, they determine that politically active users consider the news publications that are far from them, from a political ideological perspective, to be more biased than those closer ideologically to their views. Political homophily is studied by [3] by looking at what news are shared and consumed across various Facebook friends networks. This is accomplished by measuring the exposure and the engagement of users to content that pertains to a political ideology different from their own. Even though the engagement is significant, they observe that echo chambers are still present. The

presence of significant cross-ideological engagement can be explained by the structural form of the social networks. The Facebook social networks form different political homophily patterns than mediums such as Twitter or blogs [3]. In the case of Facebook, the connections depend on a multitude of offline social factors, wherein the case of Twitter and blogs, the users tend to mainly aggregate around their topics of interest.

In the case of Twitter, the presence of echo chambers, for both politically polarised and non-polarised topics, was explored by [7]. The paper discovers that the political conversations were mostly held among users belonging to the same political ideology while the non-political discussions saw user engagement from across the political divide. They noticed that the engagement of liberals in conversations across the political divide was significantly higher than that of their political counterparts - the conservatives. That happened both for political topics and non-political ones, though for the latter, the gap between conservative and liberal engagement was noticeably reduced. Even though the conversation across the political divide was higher for non-controversial topics, the rate was lower than it would have been had there been no political divide to begin with. The paper [7] suggests that even though social networks have a homophilic nature, the echo chambers do not end up ensuring that information does not permeate to the opposing view. The social networks are dynamic, leading to a widening of the political gap in the case of polarising topics while serving as a means of inter-ideological conversation for the non-controversial subjects. In the work of Garimella et al. [21], echo chambers are analysed both in terms of the information that their users create and their users receive. The paper's deals with both politically polarised and non-polarised networks based on Twitter retweets. Users belonging to an echo chamber are analysed both in term of the information that they create and the information that they receive. The analysis is performed on a large amount of data with the results indicating the prevalence of political echo chambers on Twitter. In the paper [21], the authors also study the relative positioning of users who consume and produce content from both echo chambers, thus theoretically closing the political gap, and the users who while consuming from both sides of the debate end up producing content for only one side. The former type of user is marginalized both in terms of content appreciation and network positioning while the latter has a more relevant position in the network when compared with the average user, both in terms of centrality and in terms of their content appreciation.

2.2 Analysis of Twitter Interactions

The previous subsection already mentions some forms of Twitter interaction analysis that were performed by [2, 7, 21]. In this subsection, we further explore the subject by

looking at some papers that focus on the interactions around controversial topics, be them political or not, for a given time period.

An analysis, through the prism of sentiment analysis, of the debate generated by controversial topics on Twitter, is presented by [36]. In the paper, the data analysed is from the months prior to the U.S. state of California ballot initiative of November 2012. The ballot was composed of 11 initiatives that dealt with various public issues. Throughout the paper, the users' behaviour is studied via their ideological position, estimated via sentiment analysis, taken by them in regards to controversial topics. The authors notice the preference of users to spread information with those with whom they agree. This is in direct contrast with the users' sparse debate with the opposing side and their tendency not to alter their opinion in the situations in which such a cross-opinion debate takes place. A significant time delay was also noticed between retweets and mentions; the time delay between when a user receives a post and when they retweet it is significantly smaller than the delay between when they are mentioned in a post and when they replay to said mention.

In the work of Conover et al. [14], Twitter interactions for the period right before the U.S congressional mid-term elections of 2010 are analysed. The used data spans some six weeks and is modelled as two interactions networks, one composed from the retweets and one from the mentions. It is shown by the paper that the retweet network exhibits a bi-cluster structure in which the left-leaning users are clearly divided from their right-leaning counterparts. The mention network presents no cluster structure thus showing no clear divide between ideologically opposing users. Most interactions in this network are across party lines. The authors [14] determine that this happens because interactions are provoked across political ideological lines by the insertion of opposing political views into the communication channels used by politically opposing users. One such way of provoking interactions is by using hashtags attributed to one political ideology in a message pertaining to their ideological counterparts. It is also noted that users that use hashtags that are considered to be politically neutral are more likely to engage in conversation across the political aisle.

The work of Morales et al. [30] both proposes a solution for the measurement of political polarisation and tests its validity by using Twitter interactions. The results of the analysis performed is then scrutinized using information external to Twitter. As a result of said scrutiny, the authors conclude the validity of their findings. Their proposed measurement of political polarisation is accomplished in two steps. First, the opinions of the analysed population are estimated, then their degree of polarisation is measured. Populations are deemed to be more polarised when they tend to aggregate in clearly opinion-divided groups of equal size. The opinions of the whole population are estimated by propagating the opinions of some relevant users throughout the network,

hence the opinion of the users depend on the opinion of their neighbours. Using the aforementioned polarisation model and Twitter data from around the time of death of Hugo Chávez, the former president of Venezuela, the paper [30] observes the social discourse before, during and after the announcement of his death. For each day of interactions, the retweets are used to form a weighted network. They notice that in the days prior to the announcement of the president's death the conversation was politically polarised while during the day of the announcement the political polarisation was not noticeable in the network's structure. The day after the announcement saw a return to a politically polarised network structure, this polarisation only increasing in the following days until it reached its peak a couple of days after the announcement; after that, the conversation remained polarised but the conversation shifted towards new topics, such as the interim new president.

An analysis on Twitter data that starts in late 2011 and spans circa five years, thus conferring the study the ability to explore the long term dynamics of controversial topics in the context of social networks is provided by [20]. The paper focuses on socially-relevant controversial topics in the U.S. while also taking into account some that are deemed non-controversial to be presented as a comparison. Four controversial topics are explored in the analysis, these being Obamacare, abortion, gun control and fracking. The interactions for these topics were collected in such fashion that they would confer a balanced view of both sides of the debate. For each of these topics, the data is aggregated on a daily basis in two kinds of graphs, one based on the retweets among users, thus signifying endorsements and one based on the replies signifying discussion. The former is meant to model the bi-cluster nature of the controversy while the latter explores the communication across opposing views. The daily retweets graphs are aggregated to allow the discovery of two clusters, one for each side of the controversial debate. To measure the controversy between clusters, the random walk controversy measure, proposed in [22], is used. This measure relies on the assumption that a graph is partitioned into two sides, each containing authoritative users. It measures the likelihood of a random user to be exposed to content generated by an authoritative user from the opposing side. The paper [20] notes that for each analysed topic most users are active only during a fraction of the days. There is though a subsection of users whose activity in the debate spans most of the analysed time period; therefore, these users are considered to form the core of the network, representing the backbone of the debate. In the case of the controversial topics, they note that there is a direct correlation between the levels of controversy and the overall interest in the topic. Each cluster in the retweet network also has the tendency to close-up by having most of its interaction inside their side of the debate. When analysing the lexicon used through the tweets they notice that as the number of active user increases, the lexicon between

the two sides has a tendency to converge thus implying that both sides end up focusing on the same fundamental issues. The paper notes that long term controversial topics fade and consequently reenter the mainstream discourse due to external events, Hence, they look for each controversial topic, at a series of related events that are also linked to a relative spike in user activity. When looking at the spikes, the authors [20] note an overall increase in polarisation. The retweet network maintains its structural properties through these spikes in popularity - that is, there are a couple of central nodes to which most peripheral ones tend to link, thus suggesting the tendency of occasional users to support the views of long term authoritative ones. This is in line with their general observations in regards to user activity and network polarisation. Overall, even though controversial topics create a temporal spike in polarisation between the two sides of the debate, in the long-term the authors do not find conclusive evidence to suggest either an ascending or descending trend in regards to the overall degree of debate polarisation. When it comes to the non-controversial topics used as reference, the authors are unable to detect relevant levels of controversy regardless of user activity levels.

3. Background

This chapter explores the relevant theoretical constructs that are directly used in the experimental part of this work. Section 3.1 briefly discusses general means of assigning a political ideology and it continues by introducing the method later on used for that endeavour, that is *Bayesian Point Estimation*. In Section 3.2 we introduce the two types of graph embeddings used, the former, *Laplacian Eigenmaps*, being used for embedding the experimental data into a facile to interpret vectorial space while the latter, *GraphSAGE*, serving in node classification.

3.1 Estimating Political Ideologies

In Section 2.2, we already noted the work of Morales et al. [30] that employs their own method for assigning political polarities to members of a Twitter network. This is done by spreading the opinion, in our case that being the political ideology, of a select few nodes of the network to the unassigned ones. One can note that the method used for the propagation of opinions can be easily change with a different diffusion model, such as [28, 35, 41].

Conover et al. [13] uses circa 1000 users on four distinct strategies to assign political ideologies to Twitter users in order to determine their relative quality. By using the content of users' tweets a ground-truth is established by manually assigning to each user a political ideology. The users are labeled as right-leaning, left-leaning or ambiguous when their political orientation was uncertain. Three distinct linear SVMs are trained to predict the political classes of the users. TF-IDFs vectors based on the users' tweets content are used as features for the first SVM - it's accuracy ends up being the overall worse. The second SVM is trained using a feature matrix that marks the frequency of relevant hashtags used throughout the tweets corpus. For the training of the third one a feature matrix based on a latent semantic analysis of the hashtag feature matrix - de facto this representing a PCA dimensionality reduction performed on the hashtag frequency matrix - is employed; its results were nearly identical with those obtained via the training of the second SVM. The final method was based on the network's structure and on information diffusion; in the retweet graph, the labels

of some nodes were attributed. Then, through an iterative process, the graph nodes were labelled with the most frequent label of their neighbours; this process continuing until equilibrium was reached. This method had overall the best accuracy rate.

3.1.1 Bayesian Point Estimation

A method for estimating, at scale, the political polarisation of users on Twitter by using a Bayesian model is proposed in [6]. On Twitter, a user can follow another user. This means, that when a followed user posts something, the content posted by said user will appear on the home screen of the user who is doing the following. The proposed model [6], infers the political ideology of a given user based on the political ideologies of the users that they follow. In broad strokes, the author's reasoning when considering a user's following preferences as a valid choice when it comes to estimating one's ideology can be summarised in two main points: (i) the presence of homophily in social networks indicate the closeness of users that are similar; and (ii) users also prefer to be exposed to opinions that are in line with theirs, thus by following users with whom they are in agreement with, the information that they receive reinforces their beliefs.

The proposed model considers that the probability that user $i \in \{1, \dots, m\}$ follows user $j \in \{1, \dots, n\}$ from the same network is given by

$$P(y_{ij} = 1 | \alpha_j, \beta_i, \gamma, \theta, \phi_j) = \text{logit}^{-1}(\alpha_j + \beta_i - \gamma ||\theta_i - \phi_j||^2) \quad (3.1)$$

where $y_{ij} = 1$ when i follows j and 0 otherwise, α_j is j 's popularity, β_i is i 's political interest, $\theta_i \in \mathbb{R}$ and $\phi_j \in \mathbb{R}$ are the point estimations of i and j respectively and γ is a constant. With the exception of y_{ij} , all the previously mentioned parameters must be inferred. When parameters are assumed to be independent, the model is maximized by the likelihood function given by equation 3.2.

$$p(\mathbf{y} | \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\phi}, \gamma) = \prod_{i=1}^n \prod_{j=1}^m l_{ij}^{y_{ij}} (1 - l_{ij})^{1-y_{ij}} \quad (3.2)$$

where $l_{ij} = \text{logit}^{-1}(\alpha_j + \beta_i - \gamma ||\theta_i - \phi_j||^2)$

The joint posterior distribution, where α_j , β_i , θ_i and ϕ_j are drawn from normal distributions, is expressed in equation 3.3.

$$p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\phi}, \gamma | \mathbf{y}) \propto \prod_{i=1}^n \prod_{j=1}^m l_{ij}^{y_{ij}} (1 - l_{ij})^{1-y_{ij}} \\ \prod_{j=1}^m N(\alpha_j | \mu_\alpha, \sigma_\alpha) \prod_{i=1}^n N(\beta_i | \mu_\beta, \sigma_\beta) \prod_{i=1}^n N(\theta_i | \mu_\theta, \sigma_\theta) \prod_{j=1}^m N(\phi_j | \mu_\phi, \sigma_\phi) \quad (3.3)$$

A publicly available implementation of the method is provided by the author*. The implementation relies on a predefined list of relevant Twitter users (identified via their Twitter user ID) whose political ideology is already known. The Markov Chain Monte Carlo implementation does not yield estimates for an user unless they follow at least three users from the predefined list. The estimated values are continuous. Users with positive values are considered to be *Republicans* (right-leaning) while those with negative values are considered to be *Democrats* (left-leaning).

3.2 Graph Embeddings

A graph embedding depicts a lower-dimensional representation of a graph. This new representation is meant to preserve the properties of the original graph. Graph embeddings facilitate the analysis of the underlying data by providing both an overall faster way to perform computations and a simpler way of extracting relevant information [12].

3.2.1 Laplacian Eigenmaps

Laplacian Eigenmaps represent a lower-dimensional representation of a graph into a linear space [10]. By projecting the graph into a lower dimension, the embedding ends up highlighting the graph's intrinsic cluster structures. The methods used to obtain the embedding are closely related to those used in the case of Spectral Clustering (for which there are multiple variations, some are presented in [17, 45]). They mainly differ from one another in the sense that Spectral Clustering has one more step at the end of its implementation when compared with *Laplacian Eigenmaps*; in that additional step the resulting embedding is clustered (see [10, 17]). In this work we do not rely on the cluster structure provided by the embedding; instead we use the embeddings, due to their lower dimension representation, in conjunction with the clusters corresponding to the users' political ideologies, hence our choice of referring to the method via its former name.

For a graph G with n nodes, the *Laplacian Eigenmaps* are determined by using the Laplacian matrix, defined in equation 3.4.

$$L = D^{-1/2}(D - A)D^{-1/2} \quad (3.4)$$

where $A \in \mathbb{R}^{n \times n}$ is the graph's adjacency matrix - the value of an element from row $i \in \{1, 2, \dots, n\}$ and column $j \in \{1, 2, \dots, n\}$ is given by the weight of the edge

*https://github.com/pablobarbera/twitter_ideology

between nodes i and j (with 0 if no edge exists); and $D \in \mathbb{R}^{n \times n}$ is a diagonal matrix that holds at row i and column i , for $i \in \{1, 2, \dots, n\}$, the node degree of node i .

Using the Laplacian matrix, the eigenvectors corresponding to the smallest k eigenvalues are computed. In this work, all the used graphs are connected. This leads the smallest eigenvalue to be 0, and its corresponding eigenvector to be the unit vector multiplied with a constant [45], hence it does not contain useful information. From this point forward, when we are referring to the smallest k eigenvector, we are referring to the eigenvector corresponding to the k^{th} smallest non-zero eigenvalue, that is, eigenvalue $k + 1$.

We use the smallest k eigenvectors to form a matrix $X \in \mathbb{R}^{n \times k}$, where the eigenvectors are the columns of the matrix. Ng. et al. [32] propose for matrix X to be normed row-wise using the second norm, $X_{ij}^{\text{normed}} = X_{ij} / (\sum_{j=1}^k X_{ij}^2)^{1/2}$. The resulting X_{ij}^{normed} is the *Laplacian Eigenmaps* for the original graph G .

3.2.2 GraphSAGE

Hamilton et al. [24] introduce *GraphSAGE*, an inductive learning algorithm that employs a neural network for the training of node embeddings. For training node embeddings, the algorithm uses the features of neighbouring nodes. More exactly, during each step, a constant number of neighbours are sampled. Using features that explain the local node structure, while positioning them in the graph, allows the algorithm to train node embeddings that are capable of generalising for nodes that were not part of the training process.

Algorithm 1: *GraphSAGE's* forward propagation (from [24])

Result: hidden states z_v

Input : graph $G(V, E)$ with nodes $\forall v \in V$; node input features f_v ; K ;

weight matrices $\mathbf{W}^k, \forall k \in \{1, \dots, K\}$; aggregation functions

$\alpha_k, \forall k \in \{1, \dots, K\}$; non-linear activation function σ ;

$h_v^0 \leftarrow f_v, \forall v \in V$;

for $k = 1 \dots K$ **do**

for $v \in V$ **do**

$h_{\text{Neighbourhood}(v)}^k \leftarrow \alpha_k(h_u^{k-1}, \forall u \in \text{Neighbourhood}(v));$

$h_v^k \leftarrow \sigma(\mathbf{W}^k \cdot \text{Concatenate}(h_v^{k-1}, h_{\text{Neighbourhood}(v)}^k));$

end

$h_v^k \leftarrow h_v^k / \|h_v^k\|_2, \forall v \in V$;

end

$z_v \leftarrow h_v^K, \forall v \in V$;

Node embeddings are obtained via *GraphSAGE* during forward propagation (see

algorithm 1). Central to the algorithm is the concept of node traversal depth, named K , that expresses the distance in hops from a given starting node. During the algorithm, each step represents an increase in the number of hops until K is reached. For each step, the node's embedding is obtained by passing through an activation function the concatenation formed by aggregating the node's neighbourhood with its previous step embedding. The node embeddings for all nodes are then normed.

The other aspects of neural network training are in line with classical Deep Learning methods. When testing the quality of the algorithm, the paper shows that aggregating via pooling leads to some of the best results.

4. Experiments

In our experiments, we analyse three data sets comprised of daily tweets, each pertaining to a distinct topic that is deemed to be controversial in the U.S. These topics are *abortion* - referring to the divide between pro-abortion and anti-abortion supporters, *gun control* - referring to the divide between those that want stricter gun control and those who do not, and *healthcare* - referring to the debate around the state of the U.S. healthcare system that was reformed by the introduction of the *Patient Protection and Affordable Care Act*, also informally known as *Obamacare*. The presence of controversy was already analysed in these data sets by Garimella et al.[20]. They were kind enough to provide us the data for our analysis. For each controversial topic, its corresponding data set contains only a portion of all the existing Twitter interactions.

Even though our analysis stems from the same original data, we are analysing a different facet of these controversial topics and we are employing different methods to do that. In their work, they analyse each topic from a pro or anti said topic perspective. We instead are performing our analysis from the prism of political ideology, thus trying to determine how each side of the political divide situates themselves regarding the controversies. Their analysis is performed on graphs while ours is performed on graph embeddings (which were obtained from graphs). In said graphs, the users are marked as nodes while the interactions among users as edges. The manner in which we construct the graphs differs from their own. While they build two graphs for each day of the controversy: one from the retweets among users and one from the replies; we instead build a single graph, per day, per controversial topic, that contains all the interactions among users. The overall structure of our daily graphs also differs, in their daily graphs they only include that day’s interactions. On the other hand, our daily graphs consider all the interactions among users, for all the days of the conversation. In our daily graphs, each edge that models the interactions between two users is weighted in accordance to the number of interactions between said two users during that day. If no interactions have occurred, then a dummy weight is used instead. This allows us to analyse the daily debate structure from the perspective of the whole conversation.

This chapter explores the analysis performed on the aforementioned controversial topics. Section 4.1 describes the preprocessing performed on the data, with Sec-

Table 4.1: Largest connect component (*aggregated graph*) compared with all the connected components combined; the total number of connected components is provided as the last column to emphasis the irrelevance of the smaller components.

Topic	Nodes Largest	Edges Largest	Nodes all	Edges all	No. components
<i>abortion</i>	316 237	1 040 760	507 867	1 196 433	123 096
<i>guncontrol</i>	236 579	686 732	335 721	768 242	69 649
<i>healthcare</i>	164 423	663 732	221 604	713 940	43 779

tion 4.1.3 presenting an overview of the postprocessed data. In Section 4.2 we look at the presence of trends in the data and we analyse the effects of external events on the conversation. Finally, in Section 4.3, we look at the effects or increased user activity on the overall political polarisation of the conversation.

4.1 Data Description

For each topic, the data is divided into files, each containing the Twitter interactions for a different day. In the case of each file, each row corresponds to a tweet that was either stored as a tab-separated entry or as a JSON. The amount of information stored from each tweet varies, but we were able to get from each tweet the screen name of the sender and the screen name of the receiver. Thus, for each topic, we create a graph by considering the users, determined via their screen names, to be the nodes, and the interactions between two users to represent an undirected edge between their corresponding nodes, with the overall number of interactions between said two users representing the edge’s weight. For each topic, this leads to a graph composed of multiple disjointed connected components. From these connected components we select the largest one, as in each case it clearly incorporates most of the conversation. Each of these largest connected components represent the *aggregated graph* for their corresponding topic.

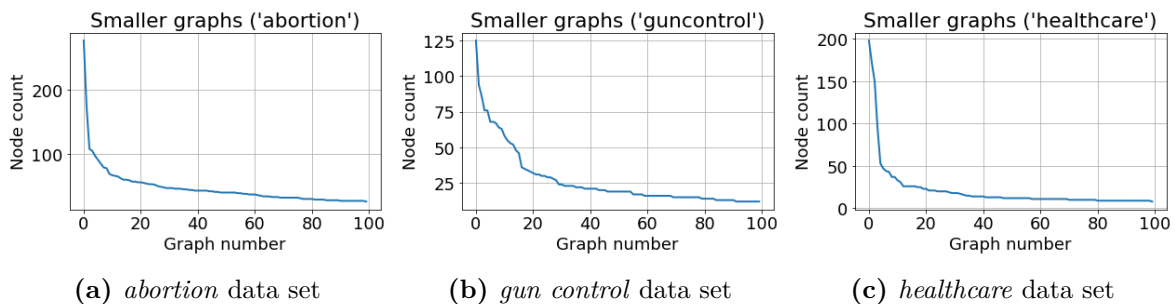


Figure 4.1: Number of nodes for the next largest 100 connected components

Table 4.1 and figure 4.1 confer an idea about the scale difference between the *aggregated graph* and all the other connected components. In the table the number of nodes and edges are provided, both for the *aggregated graph* and for all the connected components (including the largest one), thus highlighting its central role into the conversation. The number of nodes for the next largest 100 connected components is presented in figure 4.1. Table 4.1 provides a comparative view between the *aggregated graph* and all the components. The *aggregated graph* for the *abortion* controversial topic incorporates tweets from 2785 distinct days; the *gun control aggregated graph* incorporates tweets from 2411 distinct days, and the *aggregated graph* for the *healthcare* controversial topic aggregates tweets from 2579 distinct days.

4.1.1 Daily Graph Embeddings

We will perform our analysis on a daily basis, thus we first need to build graphs that depict the daily interactions among Twitter users. In this section we describe the process for a single day, for a single controversial topic. The rest of the graphs, and subsequently, graph embeddings, are later on obtained in the same fashion.

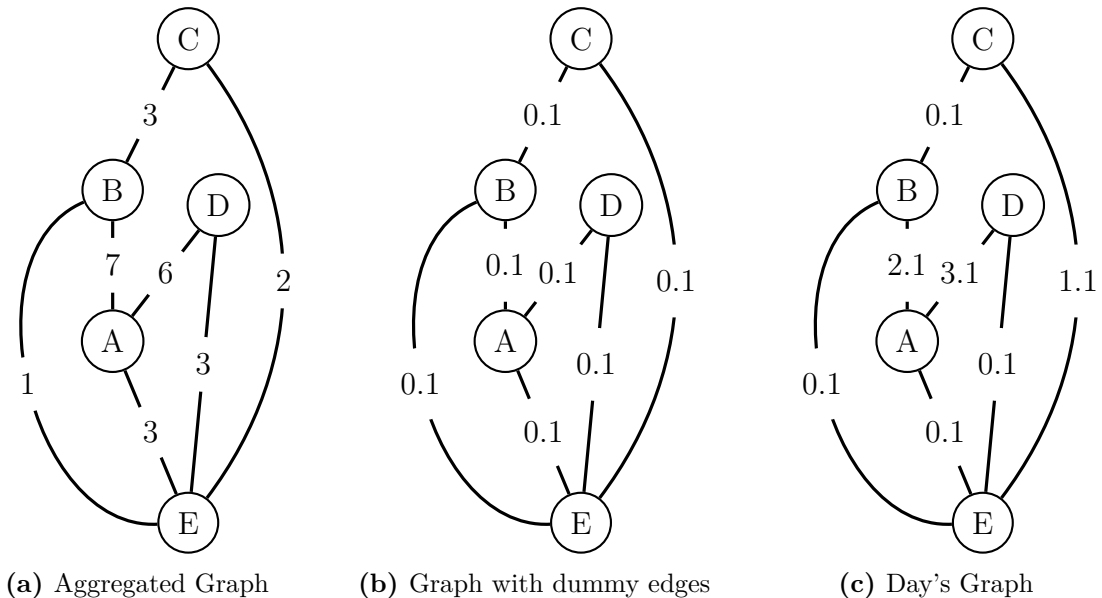


Figure 4.2: Toy example - building a daily graph: we start with the graph aggregated from multiple days depicted in (a), we then consider its edges as dummies (b); then using (b) we compute the daily graph. For a day with two interactions between *A* to *B*, one interaction between *C* and *E* and three interactions between *A* and *D* we have the resulting daily graph from (c).

To compute a daily graph for a controversial topic, we take its corresponding *aggregated graph* and we consider all its edges as dummy ones by setting their weights

to 0.1. We then look at all the interactions from the day in question and we add the number of interactions to the weight of the edges that connect the nodes where said interactions occurred. For each occurring interaction in a day we increment the weight of the edge that connects the two nodes between which the interaction occurred; we do this for all the interactions of the day. The resulting graph is the daily graph. We use the daily graph to compute its corresponding *Laplacian Eigenmaps*, composed from the smallest two eigenvectors, using the method described in Section 3.2.1. A toy example of how a daily graph is build starting from the aggregate one is provided in figure 4.2.

4.1.2 Assigning Political Ideology

We intend to estimate the political ideology of the users that form out *aggregated graphs* using the *Bayesian Point Estimation* depicted in Section 3.1.1. For the estimation of a user’s ideology, the method requires the Twitter user IDs of those on which the estimation is performed. We use the Twitter API*, that allows us to do 900 user-related requests per 15-minute batch, to match the user screen names based on which we have built the *aggregated graphs* to their corresponding user IDs. We then estimate the users political ideologies. That is done for each user by relying on the political ideologies of their neighbours in the graph. The use of neighbours’ polarities for ones political ideology estimation is a sound one: (i) homophily is present in social networks, thus the user and their neighbours tend to be similar; (ii) it was indicated (see [20]) that in the case of controversial topics there is a tendency of new users to link themselves to more authoritative users of the subject; (iii) ones social interactions represent a reflection of their own beliefs, for one would tend to interact with others that are exhibiting a similar line of thinking with their own - in the case of graphs that model social interactions this would be translated into the tendency of nodes to neighbour nodes with whom they tend to agree. Just a subsection of the users were neighbouring at least three relevant Twitter users, hence we were able to estimate the political ideology of just 78133 users from the *abortion aggregated graph*, 76983 users in the case of the *gun control* one and 70024 users from the *healthcare* one. The estimated ideologies were continuous values, with negative ones depicting *Democrats* and positive ones depicting *Republicans*; with higher magnitude-wise polarities corresponding to more polarised users - that is, users who were more partisan. In general terms, in the U.S., the *Republicans* can be viewed as a right-wing party, while the *Democrats* as a left-leaning one.

To determine the polarity ideology of the remaining users, we train a *GraphSAGE* model - using for it the implementation from the graph machine learning library StellarGraph [15]. We divide the already determined ideologies into classes that represent

*<https://developer.twitter.com/en>

different degrees of user political polarisation. For each political ideology the classes are divided such that the number of entries from one to another are of roughly equal size. Thus, we try four possible class divides: (i) four *Republican*-leaning classes and four *Democrat*-leaning ones, (ii) three and three, (iii) two and two, and (iv) a two class divide between *Republicans* and *Democrats*. The models were trained with the Adam optimizer and the aggregation function used by *GraphSAGE* was mean-pooling. When training binary classification models the loss function was binary cross-entropy and the activation function was the sigmoid. In the case of training multi-class classification models, the loss function was categorical cross-entropy and the activation function was softmax.

For each topic, the models are trained on their corresponding *aggregated graph*. We try three distinct ways of defining a node’s features. In the first method, a node’s features are given by concatenating its daily real node degree (the node degrees in which we ignore the influence produced by the dummy edge weights). Thus, a node’s feature vector in the case of the *abortion aggregated graph* has 2785 entries. For the *gun control aggregated graph*, a node’s feature vector had 2411 entries while for the *healthcare* one it has 2785 entries. Given the size of the *aggregated graphs*, the resulting features would significantly increase the computational costs of our models. Thus, we decide to perform a PCA dimensionality reduction on the features. For each topic we only keep the first 800 newly determined node degree based features. These explain some 96.58% of the original *abortion* node degree based features, 98.95% of the original *gun control* ones and 98.13% of the *healthcare* ones. From now on, we will refer to the PCA dimensionally reduced node degree features simply as node degree features. The second method defines the node features as the smallest 100 eigenvectors of their corresponding *aggregated graph*. The last method simply concatenates for each node the 800 node degrees with the smallest 100 eigenvectors.

The model is trained by keeping from the *aggregated graph* only the sub-graph containing the nodes labeled in accordance with their political ideology. The nodes are split into a training set (70%), a validation set (24%) and a testing set (6%), ensuring in each set the same percentage of each class as in the pre-split original set. The train, validation and testing of the model is performed on the sub-graph. After that, the trained model is used to predict the political ideology of the unlabeled nodes from the *aggregated graph*. Table 4.2 displays the various accuracy rates for the various models trained. Overall the best performing models are the binary classifiers that rely in a form or another on the node degrees. Given the additional overhead implied for the models that relies on the combined features, with no noteworthy accuracy benefit, we decide to use the binary prediction models that use as node features the node degrees in order to predict the political ideology of the unlabeled nodes. The training and

Table 4.2: Test accuracy rates for the trained models. The first four columns corresponds to models trained with the 800 node degrees as node features for various class sizes. The fifth uses the smallest 100 eigenvectors as node features and the last column presents the models in which the features were based on the concatenation of the 800 node degree with the smallest 100 eigenvectors.

Topic	Test accuracy rate [%]					
	N. degree	N. degree	N. degree	N. degree	Eigenvectors	Combined
	features	features	features	features	features	features
	8 classes	6 classes	4 classes	2 classes	2 classes	2 classes
<i>abortion</i>	65.28	72.17	80.70	93.44	66.24	93.62
<i>gun control</i>	69.17	72.63	83.98	95.53	88.18	95.60
<i>healthcare</i>	68.99	76.28	83.52	93.26	72.66	93.17

validation traces throughout their 20 epoch training is presented in figure 4.3.

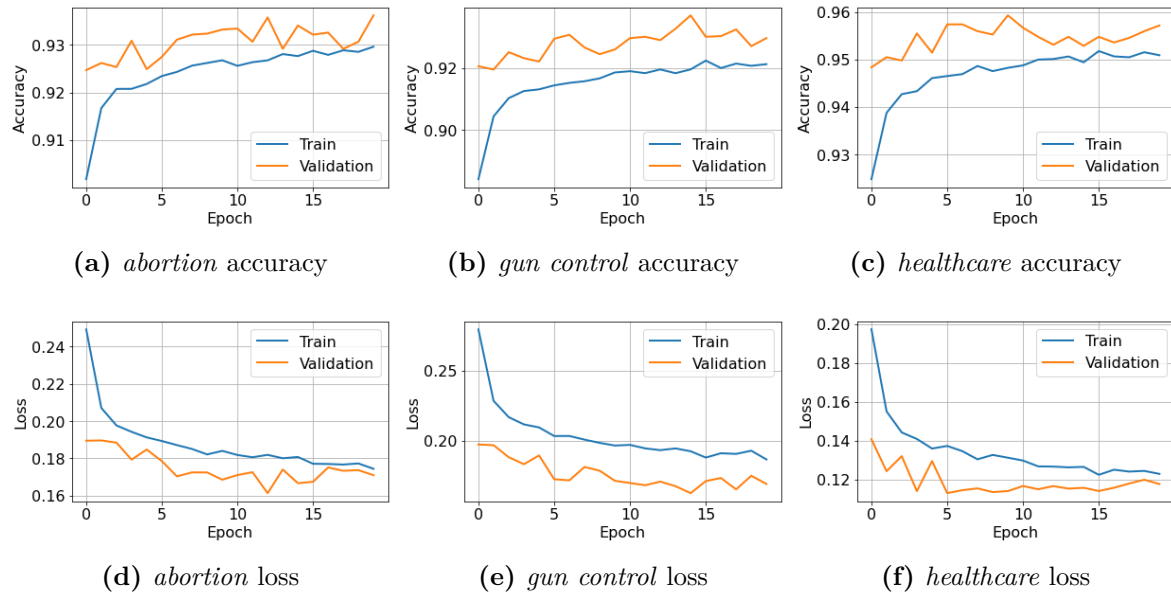


Figure 4.3: Train and validation accuracy rates (first row), and losses (second row) for the used binary political ideology classifiers

4.1.3 Data Overview

The *aggregated graph* for the *abortion* controversial topic incorporates tweets from 2785 distinct days, spanning from May 15, 2008 to October 9, 2016. For the *gun control* controversial topic the *aggregated graph* contains tweets for 2411 distinct days, from July 3, 2008 to October 9, 2016. The *aggregated graph* pertaining to the *healthcare* controversial topic has 2579 distinct days, from May 22, 2009 to October 9, 2016. For

Table 4.3: Percentage of user's from the *aggregated graph* that are either *Republican* or *Democrats* .

	<i>abortion</i>	<i>gun control</i>	<i>healthcare</i>
<i>Republicans [%]</i>	56.87	56.45	65.82
<i>Democrats [%]</i>	43.13	43.55	34.18

each topic, for the available daily tweets, the first years are quite sparse, both in terms of tweets per day and in terms of contiguous days in which the conversation occurs. Thus, the period on which we will focus our analysis spans from October 1, 2011 to October 9, 2016, for in this period not only the conversation occurs on a daily basis for each topic but also the overall user activity is much more higher when compared with the previous dates.

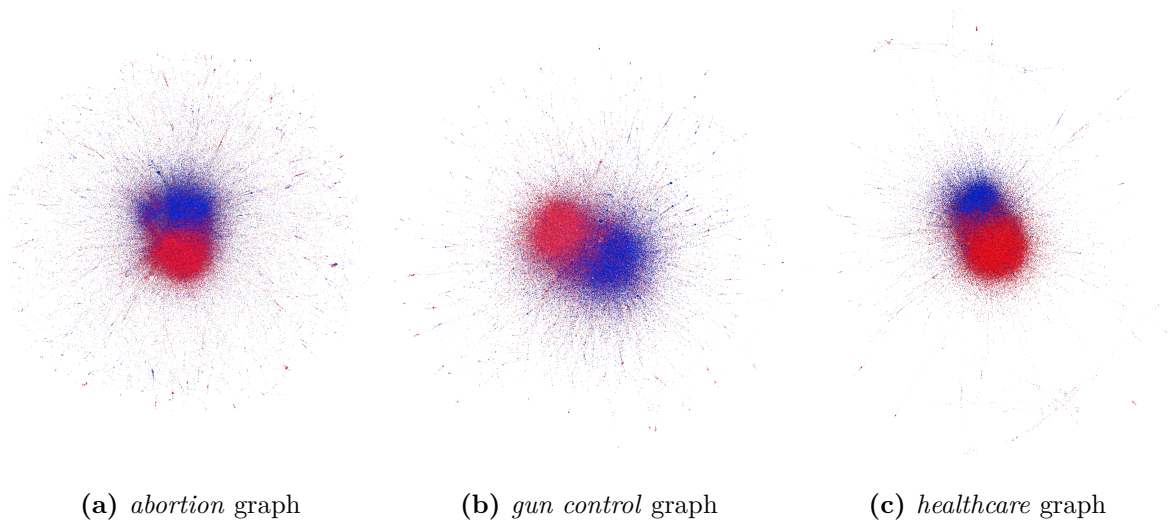


Figure 4.4: The *aggregated graphs* for the controversial topics. Gephi [8] was used for the visualisation, the used layout was ForceAtlas2 [27]; the structure revealed by the figures was obtained using only the graphs' properties, the political colouring was added afterwards. Some level of clustering along political ideology can be notice at a macro level by observing the colouring of the nodes (even though, inside each cluster there is a significant number of users that belong to the opposite political ideology).

Figure 4.4 depicts the structure of the *aggregated graphs*, the nodes colouring depicting the users political affiliation. The *Republicans* are coloured with red while the *Democrats* are coloured with blue. Table 4.3 presents the ratio's between the two political sides. We sort the days from our analysed period based on their number of daily users in ascending order. The sorted period is then divided into buckets, with each bucket being of roughly the same size, the numbering of the buckets corresponding to the level of user activity - from low to high. In figure 4.5 we present the percentage

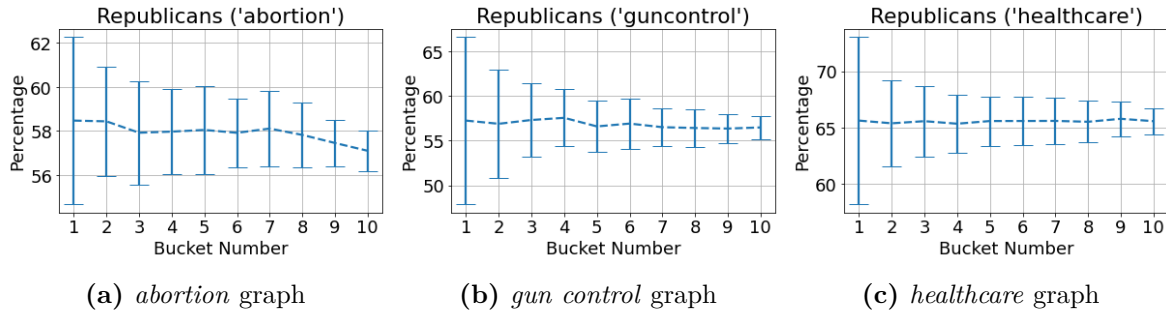


Figure 4.5: Percentage of the active user that are *Republicans*. The figure depicts the mean values from the 10 levels of user activity. The interval covered by the vertical line represents one standard deviation from the mean value.

of active users that are *Republicans* for the previously defined levels of activity. We notice that as the number of active user increases, the percentage of users of any political affiliation tends to get closer to the percentages portrayed by the *aggregated graph*.

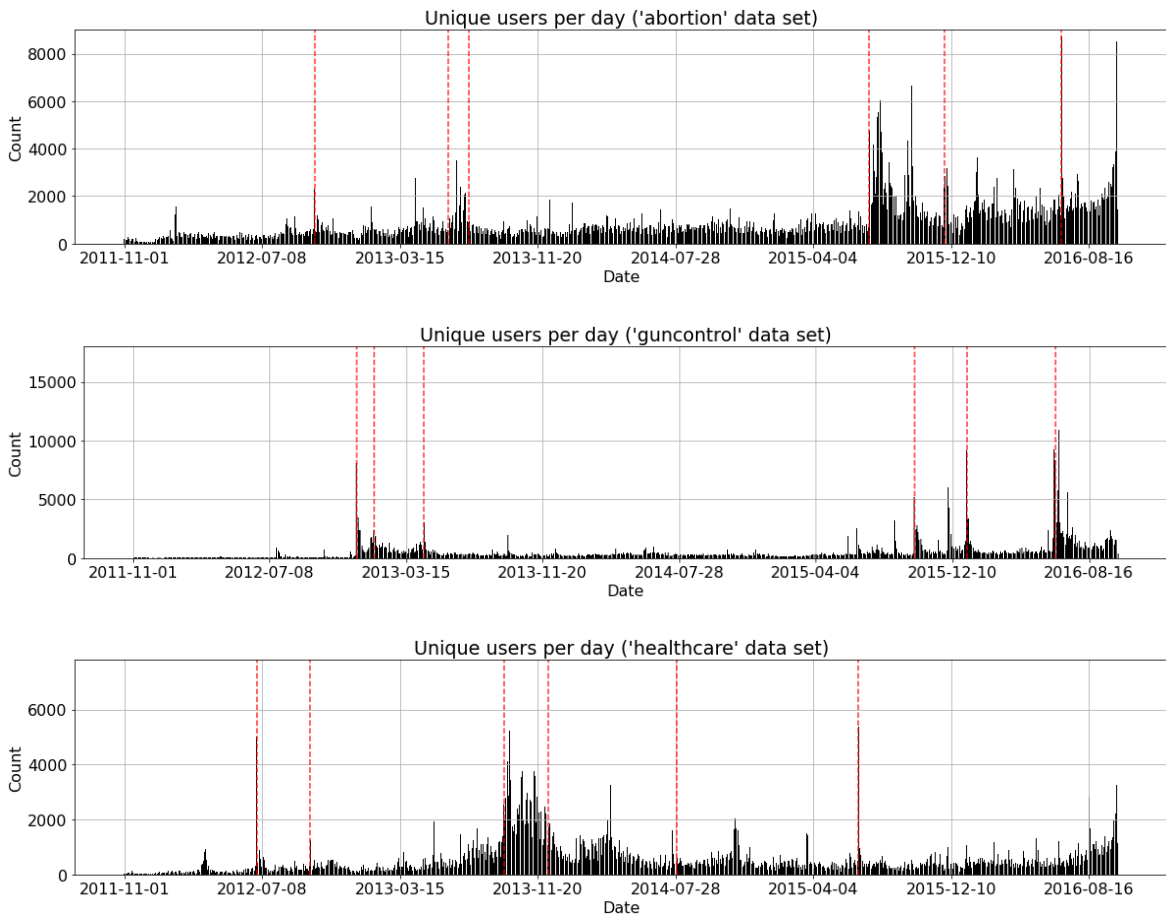


Figure 4.6: Active unique daily users; external events of interest are marked via dashed red lines

The active unique daily users during our period of interest are presented in figure 4.6. With horizontal dashed red lines we marked some relevant external events linked with the controversial topic (these events are an altered version of those highlighted in [20]). The events are part of the discourse regarding their corresponding controversial topic, but they are not something that stemmed from the Twitter interactions, hence why we named them external events. They will be part of the analysis presented in Section 4.2.

We are now going to give a summary of the highlighted external events. In the case of the *abortion* controversial topics they are:

- (i) October 11, 2012: During the 2012 U.S. presidential campaign, at the U.S. Vice Presidential Debate, the candidates for the position of Vice President mentioned their view abortion (see e.g. [33]).
- (ii) June 11, 2013: A new abortion bill is introduced in the Senate of the U.S. state of Texas (see e.g. [46]).
- (iii) July 18, 2013: the aforementioned bill is signed into law by then Texas Governor (see e.g. [9]).
- (iv) July 14, 2015: undercover videos about Planned Parenthood are released (see e.g. [37])
- (v) November 27, 2015: a mass shooting occurs at a Planned Parenthood clinic in Colorado Springs, U.S. (see e.g [31]).
- (vi) June 27, 2016: The U.S. Supreme Court strikes down the Texas abortion restrictions as a result of the ruling for *Whole Woman's Health v. Hellerstedt* (see [44]).

For the *gun control* controversial topic, the highlighted external events are:

- (i) December 14, 2012: Sandy Hook Elementary school mass shooting (see e.g. [25]).
- (ii) January 15, 2013: The U.S. State of New York signs new gun control legislation into law (see e.g. [34]). The legislation is in response to the previous mass shooting, mentioned at (i).
- (iii) April 17, 2013: The *Assault Weapons Ban of 2013* bill is defeated in the U.S. Senate (see e.g. [1]). The bill was a response to (i).
- (iv) October 1, 2015: Mass shooting in Oregon, U.S. (see e.g. [11]). In the follow up, then presidential candidate Hillary Clinton expresses her support for more gun regulations.

- (v) January 5, 2016: Then U.S. President Barack Obama presents executive actions on gun restrictions (see e.g. [40]).
- (vi) June 15, 2016: In the wake of the Orlando, Florida (U.S.) mass shooting (night between June 11 and June 12), a gun control filibuster is commenced by a Democratic Senator (see e.g. [19]). Tighter gun control regulations are proposed, these end up being shut down by the U.S. Senate on June 20 (see e.g. [16]). The spikes in activity for the aforementioned two related incidents are higher than for this event in itself.

The external events highlighted in the case of the *healthcare* controversial topic are:

- (i) June 28, 2012: U.S. Supreme Court upholds *Patient Protection and Affordable Care Act* by ruling in *National Federation of Independent Business v. Sebelius* (see [43])
- (ii) October 3, 2012: The *Patient Protection and Affordable Care Act* is discussed during the Presidential Debate for the U.S. 2012 Presidential elections (see e.g. [38]).
- (iii) September 20, 2013: U.S. House of Representatives Republicans pass bill that allows the government to function while defunding *Patient Protection and Affordable Care Act* (see e.g. [39]).
- (iv) December 10, 2013: Depicts one of three federal judges (the other two being in December 12) that were confirmed by then U.S. President Barack Obama in December 2013 (see e.g. [4]). This was done in the context in which the months surrounding the confirmations were marked by multiple disagreements between Republicans and Democrats around multiple topics, including *healthcare*.
- (v) July 30, 2014: The U.S. Republican dominated House of Representatives voted in favour to sue then President Barack Obama for delays in the *Patient Protection and Affordable Care Act* employer mandate (see e.g. [5])
- (vi) June 25, 2015: The ruling of the U.S. Supreme Court preserves subsidies (ruling for *King v. Burwell*, see [42]).

4.2 Trends and External Events Analysis

All the methods used in this section, and in Section 4.3, utilise for the analysis, the daily graph embeddings previously described. In this section, for each controversial topic, we try to determine if the political parties tend to gravitate toward the distinct sides of

the debate, that is if one political party supports one side on the controversy the other supports the opposite side. The *Laplacian Eigenmaps* have the tendency to highlight the naturally occurring cluster structures found in the *aggregated graphs*, and from the work of Garimella et al. [20] we know that these structures are indicative of the pro and anti stances on the controversial topics. In this work, instead of focusing on the cluster structures revealed by the *Laplacian Eigenmaps*, we instead focus on the structures depicted by the users' political ideologies. Thus, for a controversial topic, the presence of a clear divide between the two political ideologies would reveal that the distinct sides of the controversial topic are backed by users of distinct political ideologies. We analyse the presence of such divide, and its change over time, through the prism of the users' political ideology correlation with that of their $k\%$ -closest neighbours and through the closeness of politically aligned users.

At the same time, we look at the relative positioning of the two political sides during the highlighted external events. This is done by measuring the distance between average users of distinct political ideologies, and by measuring the mean distance among users of the same political view. From this point forward, we will refer to the distance between each side's average user as the inter-cluster distance. At the same time, the distance among users from the same side will be referred to as the intra-cluster distance. The inter-cluster distance and the intra-cluster distance are also used to detect the existence of trends in the overall level of conversation polarisation. Before we employ the inter and the intra-cluster distances on the daily graph embeddings, we test their efficiency, and observe their behaviour when the polarisation level changes, on synthetic data.

In broad strokes, this section observes no major long-term shift in polarisation for the analysed time period. There is albeit some small ascending trend in polarisation at the end of the analysed period for the *abortion* controversial topic. The findings are in line with those of Garimella et al. [20] that did not detect long-term significant changes in the overall level of controversy. It is revealed that the inter-cluster distance is able to detect the spikes in polarisation for the external events, as long as the events in question are reasonably politically divisive. Structure-wise, Sections 4.2.4 and 4.2.5 reveal that the political divide mainly follows along the controversial lines. This divide is manifested in the form of two main clusters, each one predominantly containing users of a single political ideology, situated near one another. Each of these clusters contain tight-knitted neighbourhoods of users belonging to the political ideology opposite to that of their own. These two main clusters tend to overlap the clusters indicated by the pro and anti divide. Throughout the whole analysed period, this structure did not suffer significant and long-term alterations, thus indicating the lack of an ascending or descending trend when it comes to the polarisation of the controversial topics.

4.2.1 Synthetic Data

We hypothesise that the inter-cluster distance between the two political sides increases as the political polarisation of the conversation increases. At the same time, we assume that the intra-cluster distance for each political affiliation decreases when the polarisation increases. These hypotheses are in line with previous findings that show that users have a tendency, when the controversy intensifies, to reduce their interactions with the opposite side, and to focus their interactions among other users who are in alignment with their own views (see [20]). We make these hypotheses by assuming that the users are usually involved in the conversation, and that the conversation in itself, is at least somewhat divided along political ideological lines. Occasional users are expected to be placed in the embedding at the periphery of their politically ideologically corresponding cluster. Due to that, we expect the inter-cluster distance to be less affected by the presence of occasional users than the intra-cluster distance. As the number of occasional users increases, we expect that the overall ability of the intra-cluster distance to measure polarisation to decrease.

Our experiments reveal that the inter-cluster distance increases as the interactions between the two sides decrease and that it increases as the in-cluster interactions increases. When it comes to the intra-cluster distance, we notice a distance decrease as the in-cluster interactions intensify and an increase when the interactions between the two cluster increase. Hence, our findings validate our hypotheses. Additionally, we notice that inter-cluster distance better highlights the difference between polarised and unpolarised situations than the intra-cluster distance.

Three distinct types of synthetic data sets were generated to test our hypotheses. The first two are based on weighted graphs while the last one is based on weighted ones.

Balanced Unweighted Clusters The first type of data set tests the hypotheses in an ideal setting. We form two clusters, each of 1000 nodes. For a number of n nodes in each cluster, chosen at random, we randomly chose for each of them n other nodes from the same cluster and we draw an edge between each of the resulting pairs. We will call n the intra-cluster degree. Then, we take at random a number of m nodes from one cluster and for each of them we will chose at random m nodes from the other cluster. We will connect these resulting pairs of nodes together. We will call m the inter-cluster degree.

We run experiments with different values assigned to n and m . For each n and m we run the experiment 10 times and we average over our findings. Also, for each experiment we get the inter-cluster and the intra-cluster distances using varying

numbers of smallest k eigenvectors in order to ensure that using only *two* is sufficient.

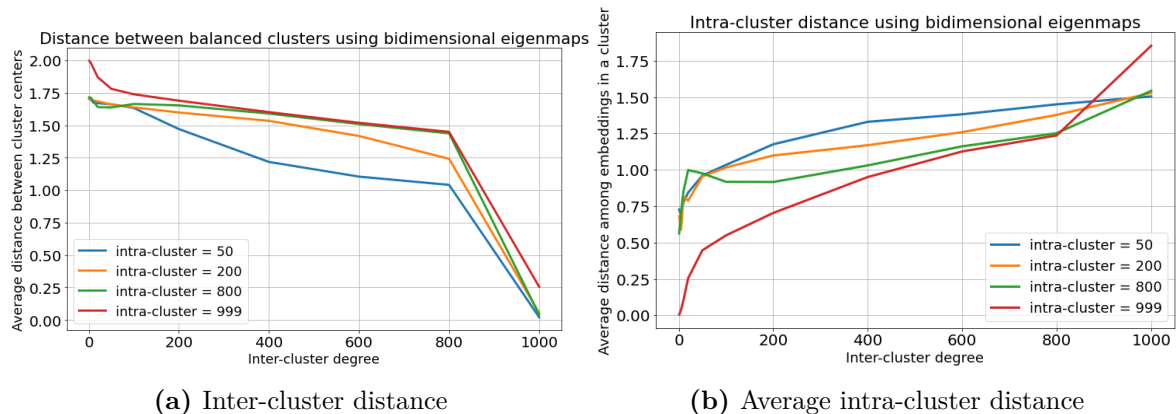


Figure 4.7: Inter and intra-cluster distances for the balanced unweighted clusters. For the intra-cluster distances the plot reveals the averaged values from the two clusters; we were able to do this because they were showing the same trends.

Figure 4.7 reveals that the inter-cluster distance decreases as the interactions between clusters increases. The inter-cluster distance also tends to increase when the number of in-cluster interactions increases. For the intra-cluster distance the trends indicate a rise in distance when the interactions between clusters increases and a fall in distance when the interactions inside the cluster increases. Figure 4.8 highlights the fact that there is no relevant trend change when the distance is computed using varying numbers of smallest k eigenvectors. Thus, there is no added advantage of using more than the smallest *two* eigenvectors.

Unbalanced Unweighted Clusters The second type of data set is identical to the first one with the caveat that we use 500 nodes per cluster and that n differs from one cluster to the other. We will display n as a pair of the form: (*n for the first cluster, n for the second cluster*). We perform these experiments to ensure that a cluster imbalance does not lead to disruptions in the previously observed trends.

We run experiments with different values assigned to n and m . For each n and m , we run the experiment 10 times and we average over our findings. Figure 4.9 reveals the same overall trends as previously observed.

Weighted Clusters We use this type of data set to simulate the presence of weights on edges, including dummy ones. In this setup each cluster is composed on 1000 nodes. Each node from each cluster is connected to some random nodes from their own cluster. The number of random nodes to which a node is connected is chosen at random, for each node, and it ranges from *zero* to 150. In the case of edges between clusters, the

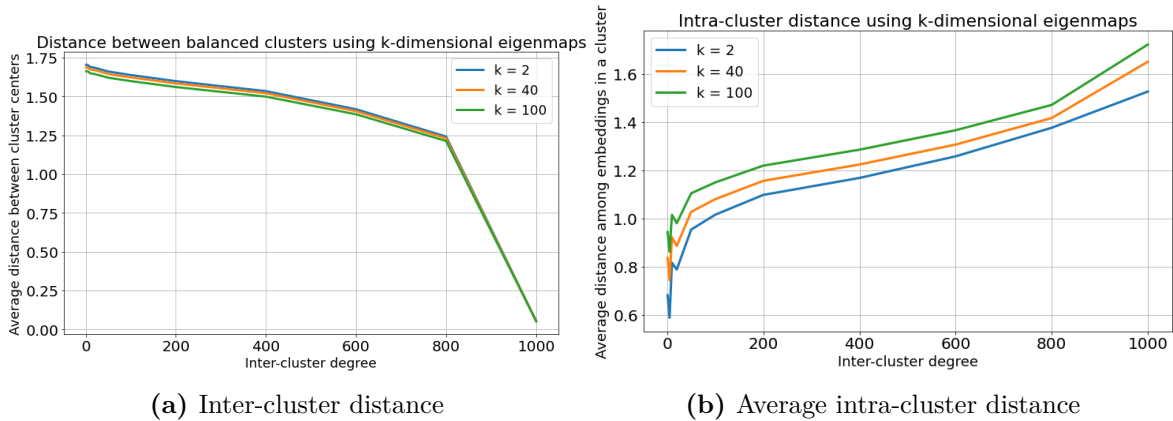


Figure 4.8: Inter and intra-cluster distances for the balanced unweighted clusters computed on embeddings composed from a different number k of eigenvectors. The intra-cluster degree for the graphs depicts in the plots is 200. For the intra-cluster distances the plot reveals the averaged values from the two clusters; we were able to do this because they were showing the same trends.

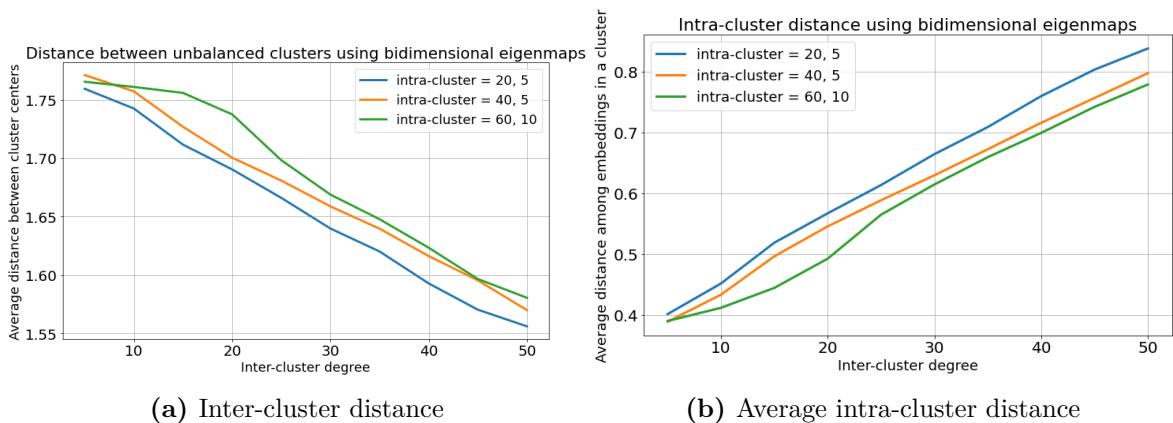


Figure 4.9: Inter and intra-cluster distances for the unbalanced unweighted clusters. For the intra-cluster distances the plot reveals the averaged values from the two clusters; we were able to do this because they were showing the same trends.

nodes from one cluster are connected to m nodes chosen at random from the other cluster. m is a number chosen at random for each node, it can range from *zero* to 30. The weight of each edge is chosen at random from a number in between *one* to 20.

We simulate the presence of dummy edges in configurations that either imply high polarisation (most edges in between the two clusters are dummies) or low polarisation (most edges in each cluster are dummies). We perform 10 runs and then we average the results, each run is performed for multiple configurations. A run is defined as: (i) we form at random the weighted graph as described above and we compute its embedding - then we compute the inter and the intra-cluster distances; (ii) for each of the config-

urations, for edges chosen at random, in a percentage specified by the configuration, we replace their weights with dummy ones (0.1), then we compute their corresponding embedding that is then used to obtain the inter and the intra-cluster distances. In figure 4.10 we display the relative percentage difference (given by equation 4.1) between the weighted graphs, with no dummy edges - which we consider to be mildly polarised and thus acting as a baseline, and the polarised graphs (the first three bars), and between the weighted graphs and the unpolarised ones (the last three bars). The results are consistent, in terms of polarisation changes and in terms of corresponding inter and average intra-cluster distance changes (again, both clusters behaved the same thus we were able to average their values). Percentage-wise, the changes between the baseline and the unpolarised graphs are more relevant in the case of inter-cluster distance while the changes between the baseline and the polarised graphs are more noticeable for the intra-cluster distance. Given that the baselines are already mildly polarised, we consider the inter-cluster distance to be better at highlighting the difference between polarised and unpolarised situations.

$$Difference [\%] = \left(\frac{\text{compared distance}}{\text{baseline distance}} - 1 \right) \cdot 100\% \quad (4.1)$$

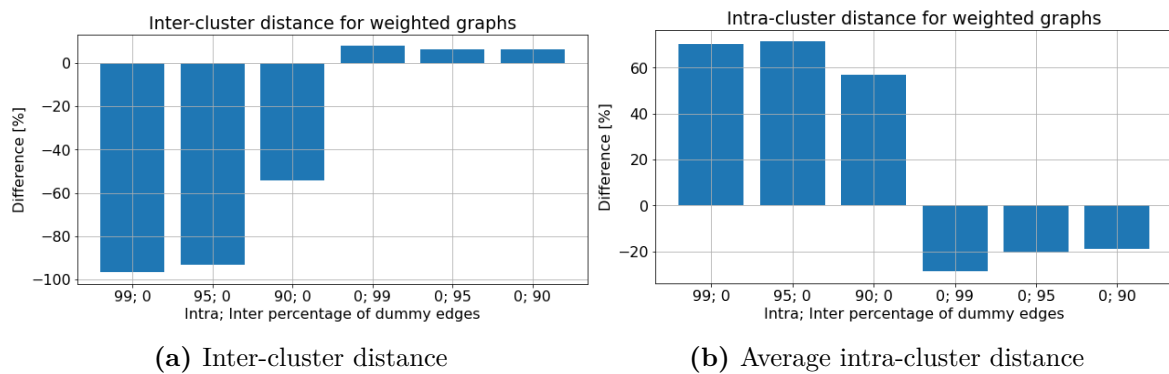


Figure 4.10: Relative percentage difference between the distances obtained via the original graphs and the distances obtained by converting various edges into dummy ones. The numbers on the x -axis are the percentages of intra and respectively inter-cluster edges that are converted into dummy ones. For the inter-cluster distance, negative values mark a decrease in polarisation while positive ones an increase. In the case of the intra-cluster distance, the opposite is true.

4.2.2 Inter-cluster Distance

The inter-cluster distance for the analysed time period is presented in figure 4.11. We notice that for each of the controversial topics, the conversation suffers from occasional

short lived spikes of polarisation. After the spikes, the polarisation level swiftly returns to a value close to its pre-spike polarisation levels. For most of the timeline, the distance reveals that the overall long-term level of political polarisation stays the same; in the case of the *abortion* debate we notice a slight increase in the polarisation levels in the last few month of our analysed period.

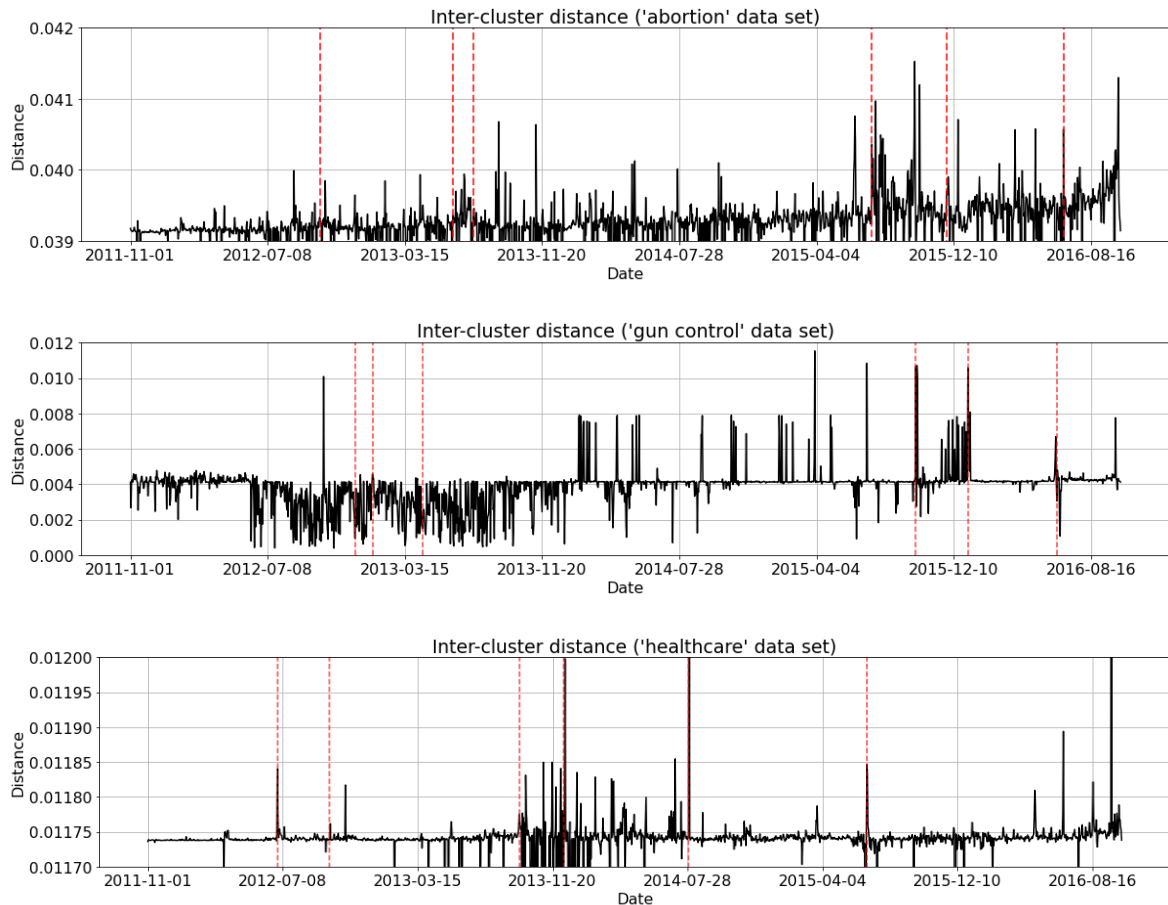


Figure 4.11: Inter-cluster distance, we mark with horizontal red dashed lines the selected external events

When looking at the external events we notice that in the case of the *abortion* controversial topic, the only one who is not marked by a significant spike in inter-cluster distance, thus an increase in polarisation, is (i) - the Vice Presidential Debate in which *abortion* was mentioned. The inter-cluster distance is somehow higher in that day than in the surrounding ones but the values are not significant. This indicates, that the debate gained some traction, but that it did not produce a relevant level of political divide along the pro and anti abortion supports. It is worth mentioning that (ii) and (iii) represent a period in time surrounding the same event - the U.S. state of Texas new *abortion* legislation. Thus, it was to be expected for the peaks in polarisation to occur in between the two dashed lines.

For the *gun control* controversy, we notice significant levels of polarisation for the last three events. In the case of the first three, it was to be expected for the polarisation levels to be low: (i) depicts a school mass shooting, (ii) marks the signing into law of new *gun control* regulation in the U.S. in the state where the mass shooting occurred, (iii) represents a bipartisan rejection in the U.S. Senate of tighter *gun control* legislation. It is noteworthy to mention that for (vi), the spike in polarisation is attributed to the mass shooting at a night club, and not to the filibuster held in the wake of the shooting.

In the case of the *healthcare* controversy, all the marked events (with (iii) and (iv) conveying more of a tentative time interval marked by multiple disagreements in the U.S. Congress around multiple issues, along with *healthcare* was one of them) show a significant increase in polarisation. (ii) shows the least amount of polarisation - it corresponds to the U.S. Presidential Debate where *healthcare* was discussed. Given that the method works as expected we are tempted to say that Presidential Debates do not spark an increase in polarisation when measuring controversial topics, whose controversy does not stem from political ideology. We can speculate that this occurs, because during a Presidential Debate, a multitude of topics are discussed in a short span of time. Thus, we assume that different sections of each political side will tend to focus on different aspect of the debate.

4.2.3 Intra-cluster Distance

The average intra-cluster distance for the analysed time period is presented in figure 4.12. There were no noteworthy differences between the intra-cluster distances of the two sides, so for brevity's sake, we averaged the values of the two intra-cluster distances and depicted them as a single plot per controversial topic. Trend-wise, we do not see an overall long-term increase or decrease in political polarisation.

By looking at the external events, we notice that our experimental observation, derived from the synthetic data experiments, does not hold. Thus, we are unable to link external events with increases in polarisation - that would be marked by drops in intra-cluster distance. We suspect that the presence of occasional users, that leads to an increase in the average distance among users, affects the intra-cluster distance much more than the closing up of the usually active users, that leads to a decrease in the average distance among users. At this point, we deem the results to be inconclusive. The appropriateness of the method, in its current form, to measure the changes in polarisation will be discussed in Section 4.3.2.

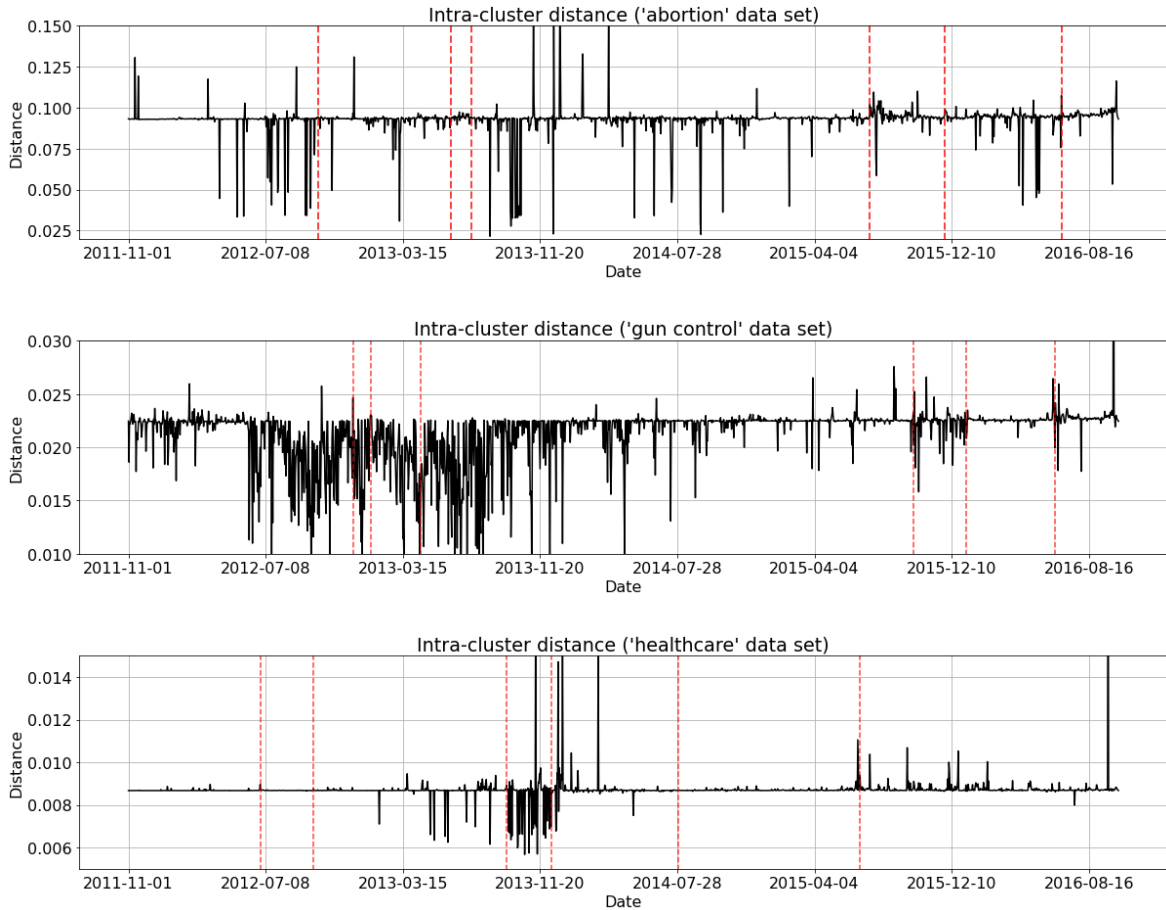


Figure 4.12: Average intra-cluster distance, we mark with horizontal red dashed lines the selected external events

4.2.4 $k\%$ -Closest Neighbours Political Ideology Correlation

In this subsection, and in Section 4.2.5, we look, from a political ideology perspective, at the long-term trends in the internal structure of the daily graph embeddings for each controversy. To accomplish this, for each day analysed, we select from the daily embedding only the part that corresponds to the active users. In this method, for each user, we take its political ideology and the mode of the political ideologies of its surround $k\%$ -closest neighbours. As one can recall, for all the *aggregated graphs*, a subsection of the users' political ideologies were determined via *Bayesian Point Estimation*, by relying on the political ideologies of their neighbours in the graphs. A user's political ideology was determined only if they were neighbouring at least *three* influential users whose political ideology was already predefined. Thus, this allowed us to determine the political ideology of only a subsection of users, hence obliging us to employ *GraphSAGE* to estimate the rest of the political ideologies. This could lead one to believe that the manner in which the political ideologies were estimated would

significantly affect the results observed in this subsection, and in Section 4.2.5. Such a belief would not be in line with the homophilic tendencies of social networks. For example, it is highly unlikely for a user to be a *Democrat* if they are mostly directly neighbouring (thus interacting) with influential *Republicans*. Thus, the observations to follow reflect valid findings when it comes to the political polarisation of our analysed controversial topics.

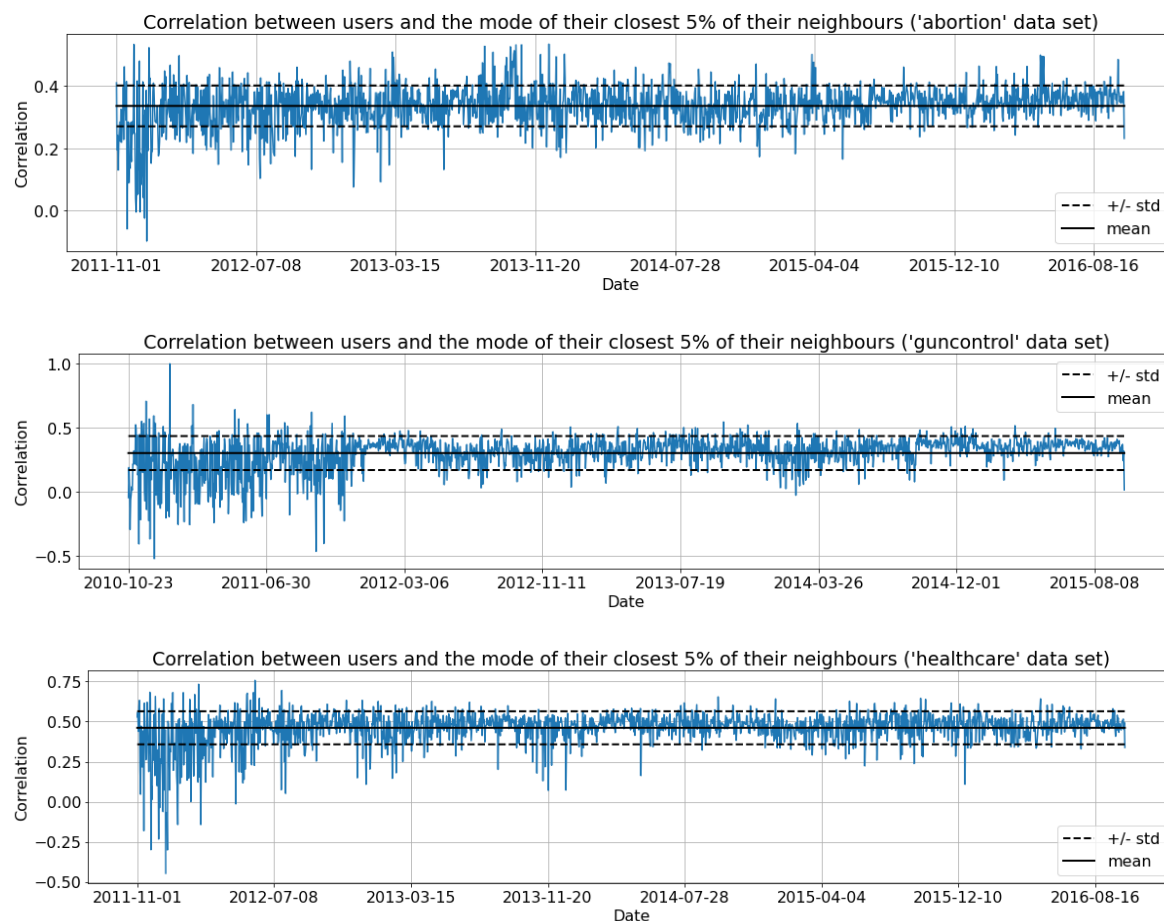


Figure 4.13: The correlation between the users political ideology and that of their 5%-closest neighbours

We take pairs of political ideologies, formed from a user's political ideology and the mode of the political ideologies of its surrounding $k\%$ -closest neighbours, for all the daily active users, and we compute the correlation between them. This represents the correlation for a given day. We compute the correlations for all the days in our analysed time period. The same thing is done for the correlations between the users' political ideologies and the mode of the political ideologies of the users' percentage-based rings of closest neighbours. For a given user, we consider a rings of closest neighbours to be composed of all the users whose distance is in the interval given by the limits of the

ring. Given that we consider each ring to start where a neighbourhood ends, and to end where the next neighbourhood begins (with the exception of the ring containing the highest interval), we pair in our analysis each neighbourhood with the ring that starts where it ends. From this point forward, we will simply refer to the aforementioned structures as correlation neighbourhoods and correlation rings. We pair, for each user, the correlation neighbourhoods with their corresponding correlation rings, because we want to see how the correlation is affected by the increase in distance from each user, thus revealing to us the shifts in relative position, from the perspective of political polarisation, among the active users.

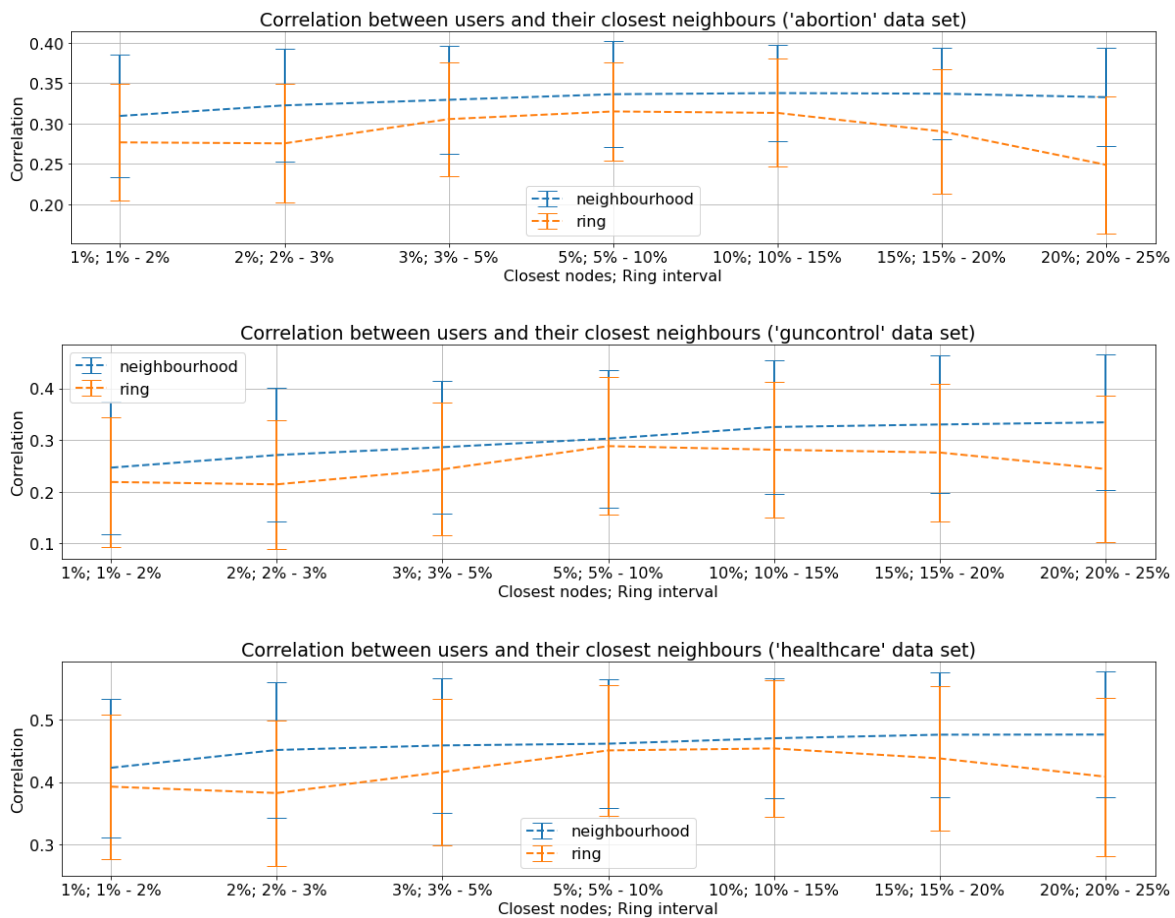


Figure 4.14: The correlation between the users political ideology and the mode political ideology of their closest neighbours (and respectively, their rings of neighbours). Each vertical line gives the mean value and a standard deviation interval around the mean, for the whole analysed time period, for the specified percentage of neighbours and the specified percentage-based interval that forms its corresponding ring. On the x -axis both aforementioned percentages are given as: *percentage of closest neighbouring points; percentage-based interval of neighbouring points*.

The correlation can be interpreted as follows: (i) if the two political sides would have been clearly segregated one from another, both the correlation neighbourhoods

and the correlation rings would have had a tendency to be closer to one; (ii) if the two political sides would have intermingled at random with one another the correlation would have been near zero; (iii) the closer we are to a correlation of 0.5, the closer we are to a balanced scenario compared to the previously mentioned two, that is, that we would have an intermingling of mini-clusters, with each mini-cluster being of the same political ideology.

Figure 4.13 depicts the correlation neighbourhoods based on the closest 5% of their neighbours. It indicates to us, that for all controversial topics, the internal structure pretty much stays the same through the whole analysed period of time. There are some occasional spikes, but in general, the structure being that of mini-clusters based on political ideology intermingling one another. For various other neighbourhood sizes, and for various ring intervals, the findings are similar and are summarised in figure 4.14. In that figure, each vertical line gives the mean value and the interval given by a standard deviation around the mean. Each mean and each standard deviation are for the whole analysed period. As a general rule, we see that the correlation rings return somewhat smaller values than the correlation neighbourhoods. This only further enforces our view, that is, that the structure of the embeddings resembles an intermingling of mini-clusters based on political ideology.

4.2.5 Closeness of Politically Aligned Users

In this subsection, we continue to look, from a political ideology perspective, at the long-term trends in the internal structure of each controversial topic. To this end, the employed measures are similar, the difference stemming from the fact that now we are looking at the percentages of active users of each political ideology that are predominantly surrounded by like-minded individuals instead of looking at the overall correlation for all the active users. For simplicity's sake, we refer to the similarly defined structures as those from Section 4.2.4, as percentage neighbourhoods and percentage rings.

Figure 4.15 depicts the percentage neighbourhood, for both *Republicans* and *Democrats*, based on the closest 5% of their neighbours. We see no clear trend in our analysed period, thus indicating that throughout the whole period the internal structure does not suffer major alterations. For the various other neighbourhood sizes, and for the various ring intervals, the findings are similar and are summarised in figure 4.16 in the case of *Republicans*, and in figure 4.17 in the case of *Democrats*. In these figures, each vertical line represents the interval given by a standard deviation around the mean value - which is located where the horizontal dashed line stems from. We notice that when the values increase for one side, they decrease for the other. The

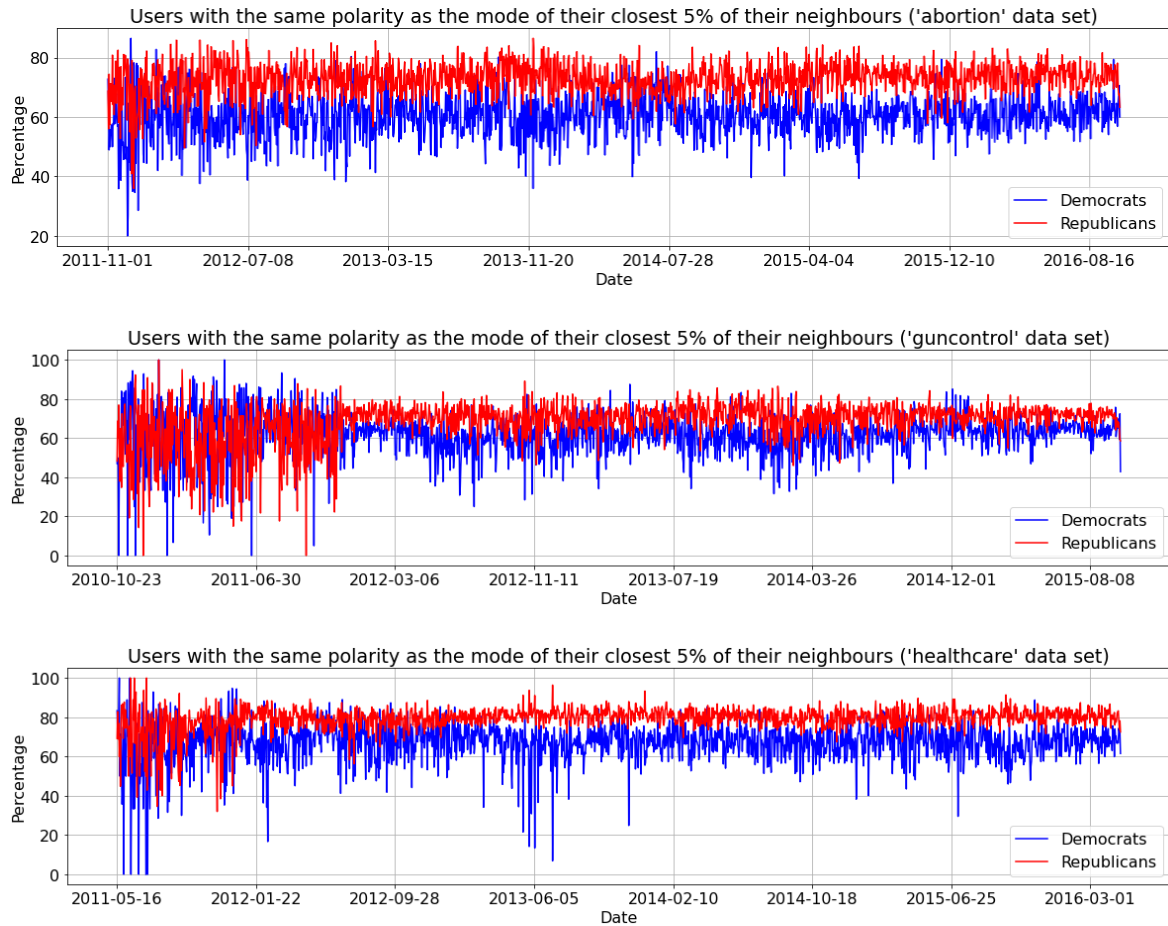


Figure 4.15: The percentage of users of the same political ideology as most of their 5%-closest neighbours

overall high percentage values combined with the findings from the previous subsection (Section 4.2.4), indicate the presence of two main agglomerations of mini-clusters of similar political ideologies. Thus, for each political ideology, we have a main cluster, composed on the mini-clusters specified in Section 4.2.4. Each of these main clusters contain a significant amount of mini-clusters of opposing political ideology. The main clusters are near one another. This indicates, that the bird's-eye view of the *aggregated graphs* that was displayed in figure 4.4, is also present in the embedding, and that this structure does not significantly change throughout the whole analysed period. It is noteworthy to mention that the two tentative clusters noticed in each of the *aggregated graphs*, depicted in figure 4.4, were highlighting the divide between the pro and anti sides found in each controversial topic. The presence of the two main clusters based on political ideology, each containing a significant amount of mini-clusters of opposing political ideology, indicate that for each controversial topic, the political divide mainly overlaps the pro and anti divide.

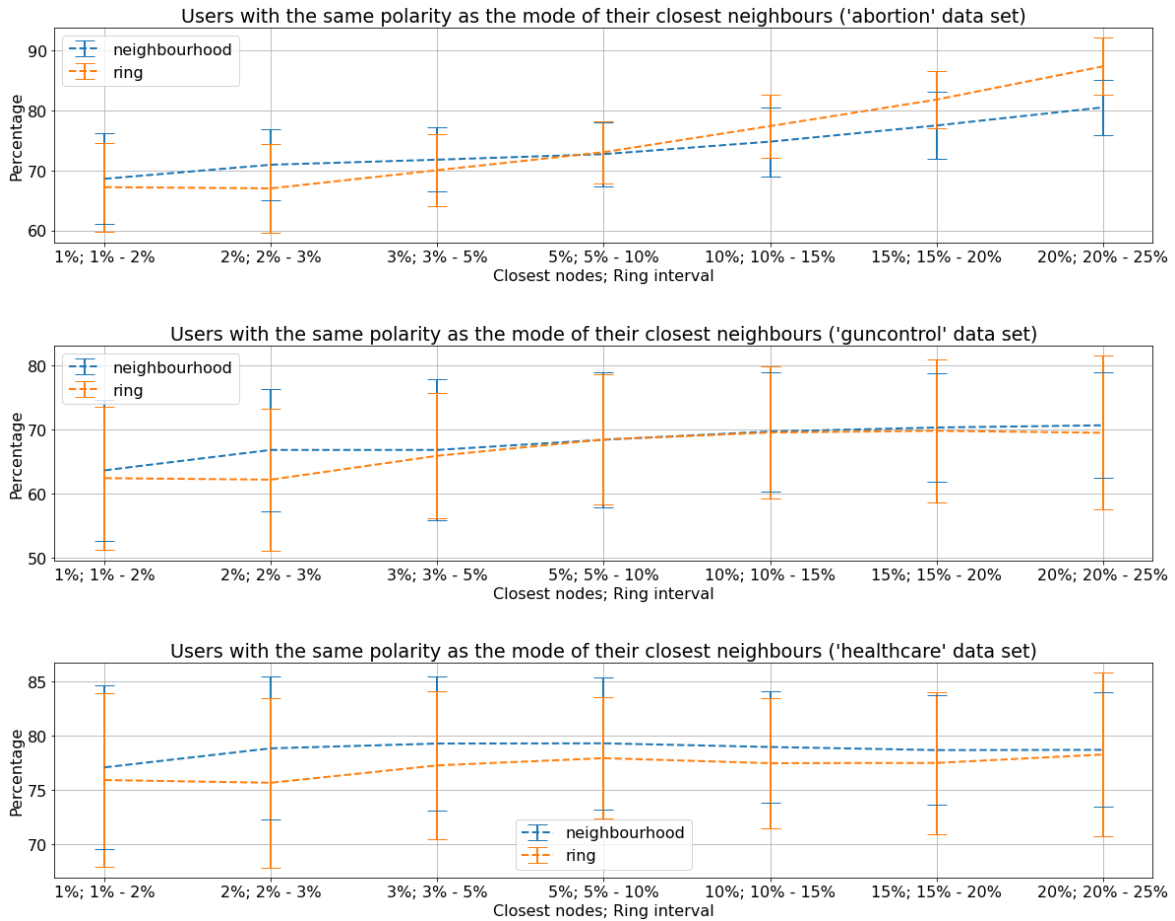


Figure 4.16: The percentage of *Republicans* that are closest to mostly (the mode of the closest neighbours - and respectively, their rings of neighbours) users of the same political ideology. Each vertical line gives the mean value and a standard deviation interval around the mean, for the whole analysed time period, for the specified percentage of neighbours and the specified percentage-based interval that forms its corresponding ring. On the x -axis both aforementioned percentages are given as: *percentage of closest neighbouring points; percentage-based interval of neighbouring points*.

4.3 Analysis of User Activity Levels

In this section, we try to determine if an increase in user activity leads to an increase in political polarisation. To accomplish this, we take the long-term measurements determined in the previous section, we sort them in ascending order based on the number of active users for their corresponding day and we divide them into 10 buckets of roughly equal size. In the case of the correlations (Section 4.2.4) and in the case of the politically aligned percentage of users (Section 4.2.5), we aggregate the neighbourhoods and their corresponding rings into two sets of buckets per controversial topic: the former for their means and the latter for their standard deviations. A bucket, from the set of buckets containing means, will contain the means for all the neighbourhoods, and

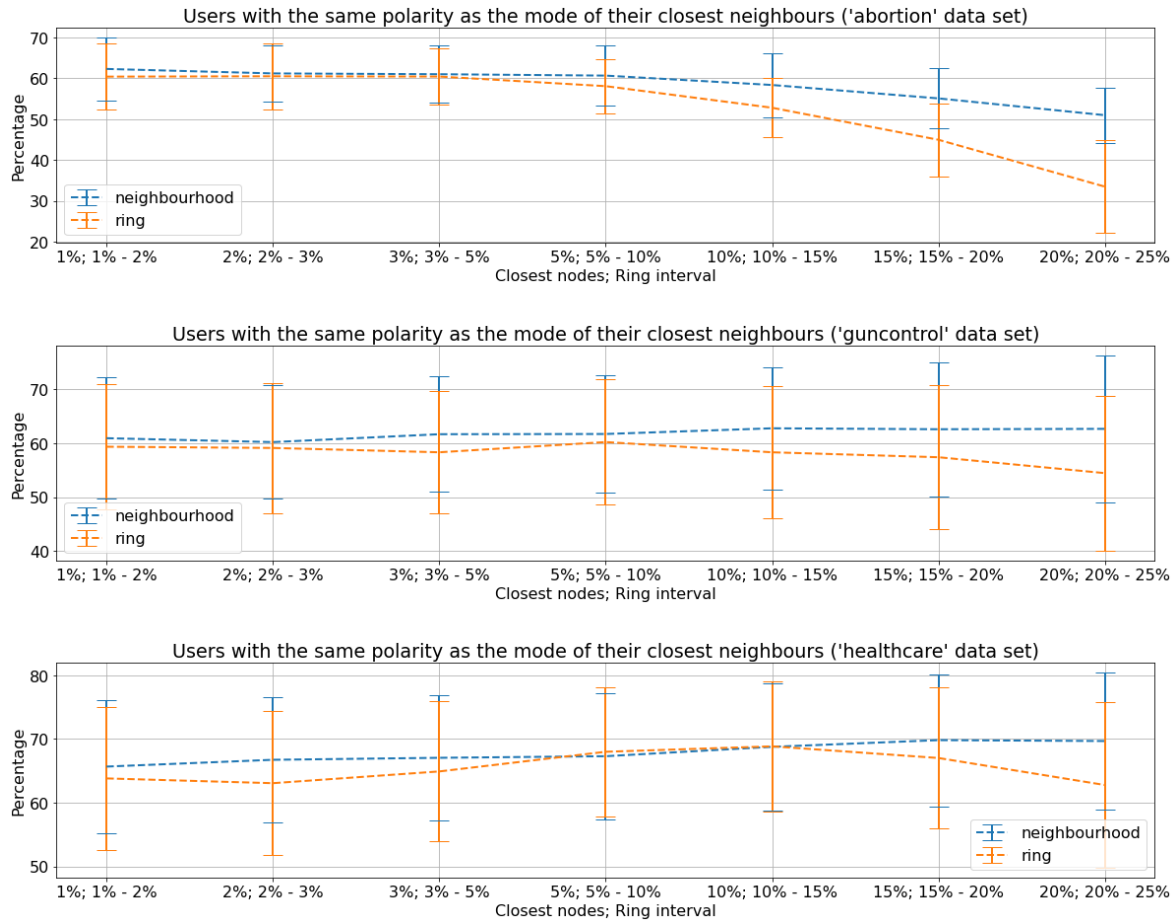


Figure 4.17: The percentage of *Democratic* that are closest to mostly (the mode of the closest neighbours - and respectively, their ring of neighbours) users of the same political ideology. Each vertical line gives the mean value and a standard deviation interval around the mean, for the whole analysed time period, for the specified percentage of neighbours and the specified percentage-based interval that forms its corresponding ring. On the x -axis both aforementioned percentages are given as: *percentage of closest neighbouring points; percentage-based interval of neighbouring points*.

for their corresponding rings, at the bucket's user activity level. The buckets for the standard deviation will be the same with the only caveat that they will contain standard deviations and not means. The plots will depict the mean values, the maximum ones and the minimum ones from each bucket.

In broad strokes, this section reveals that an increase in user activity is linked with an increase in political polarisation (this is in line with the findings of Garimella et al. [20]). The inter-cluster distance reveals some slight indication in this regard in the case of *abortion* and *gun control*, while the intra-cluster distance confers additional reasons to deem it unfit to detect political polarisation given the nature of the analysed data. The clear indications are provided by the $k\%$ -closest neighbours political ideology correlations and by the closeness of politically aligned users. These

two measures indicate the general tendency of users of a similar political ideology to shift towards one another when more users engage in the conversation. Thus, the users from each main cluster predominantly of the same political ideology shift towards one another, while the communities of users of opposing political ideology, that are located inside these clusters, tend to do the same, but around their politically defined neighbourhoods.

4.3.1 Inter-cluster Distance

Figure 4.18 displays the inter-cluster distance for multiple levels of user activity. The vertical lines depict one standard deviation distance from the mean value. In the case of *abortion* and *gun control* there is a slight increase in the inter-cluster distance when the number of unique active users increases, thus indicating a slight increase in political polarisation. For *healthcare*, the inter-cluster distance does not reveal any clear trend in political polarisation. The slight increases in inter-cluster distances are not enough on their own, for us do conclude, that an increase in the users' activity levels leads to an increase in political polarisation, but they are a sign in that direction.

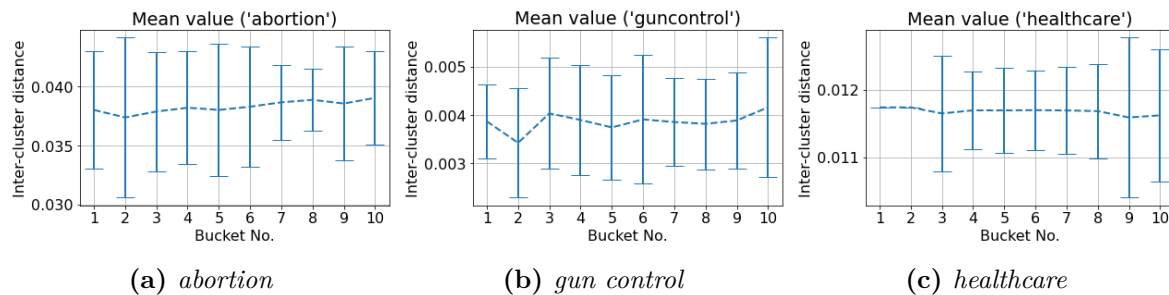


Figure 4.18: Inter-cluster distance for distinct numbers of unique active users. Higher bucket numbers correspond to higher levels of user activity.

4.3.2 Intra-cluster Distance

Figure 4.19 displays the intra-cluster distance for multiple levels of user activity. The vertical lines depict one standard deviation distance from the mean value. In this case, with the exception of intra-cluster distance for the *Democrats* when it came to the *healthcare* controversial topic, there is a slight increase in the inter-cluster distance when the number of unique active users increases. In Sections 4.2.5 and 4.3, it was revealed that, each controversial topic contains two main clusters based on political ideological affiliation. In these clusters, there are communities whose political polarisation is in opposition to the one predominantly found in their side of the debate (in terms of

pro and anti a controversial topic). Also, occasional users, of whom we assume are a noticeable amount when the conversation is polarised, have the tendency to connect themselves to authoritative users (see [20]). Thus, they end up being embedded at the periphery of the clusters determined by the interactions among usually active ones. The findings in this subsection, the structure of the daily graph embedding over time, the tendency of occasional users and the inconclusive results from Section 4.2.3, lead us to consider that, in its current form, this method is unable to determine the occurring changes in political polarisation.

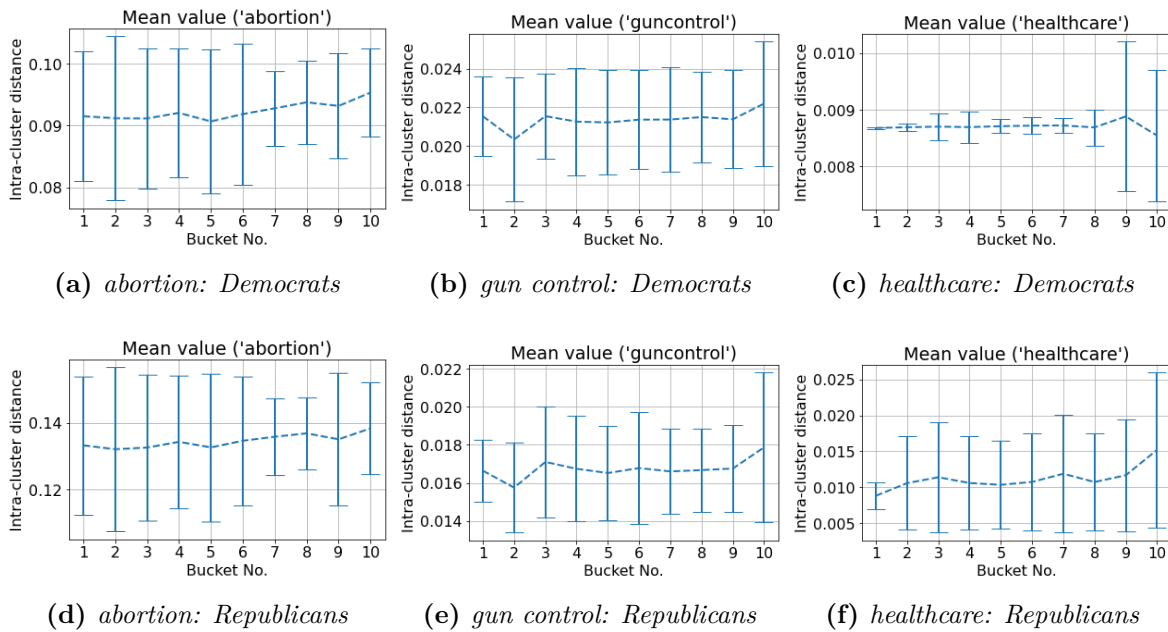


Figure 4.19: Intra-cluster distance for distinct numbers of unique active users. Higher bucket numbers correspond to higher levels of user activity.

4.3.3 $k\%$ -Closest Neighbours Political Ideology Correlation

Figure 4.20 depicts the correlation neighbourhoods and the correlation rings for multiple levels of user activity. For a vertical line, the lower cap indicates the minimum value taken in that bucket while the upper cap depicts the maximum value obtained. We notice that as the level of activity increases so does the correlation in the correlation neighbourhoods. At the same time the standard deviation decreases significantly. The correlation between the users and their correlation rings also increases, but it consistently stays, mean value wise, under the values provided by their corresponding correlation neighbourhoods. These findings indicate a user tendency to shift towards more tightly-knitted communities of like-minded individuals. Hence, these measurements indicate that an increase in user activity leads to an increase in political polarisation.

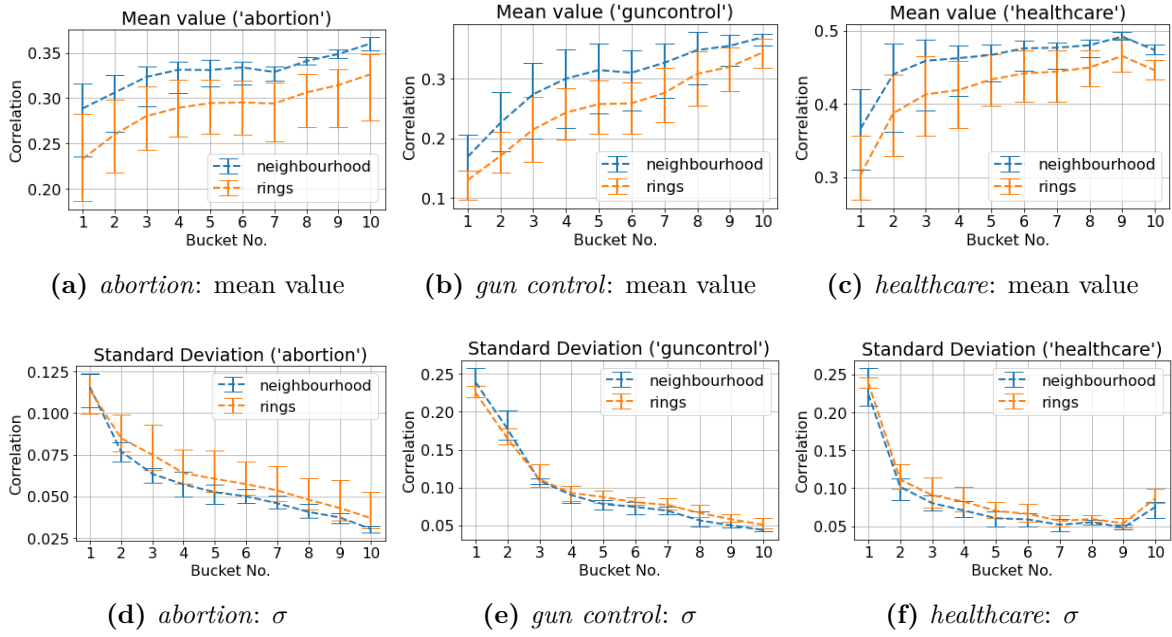


Figure 4.20: the correlation between the users political ideology and the most frequently occurring political ideology for each user’s closest $k\%$ -neighbours, and respectively the percentage-based rings of neighbours. Higher bucket numbers depict higher levels of user activity.

4.3.4 Closeness of Politically Aligned Users

Figure 4.21 presents the percentage neighbourhoods in the case of *Republicans*; in figure 4.22 the percentage neighbourhoods for the *Democrats* are revealed. For a vertical line, the lower cap indicates the minimum value taken in that bucket while the upper cap depicts the maximum value obtained. As the user activity levels increases, the standard deviation decreases, while the mean value increases; this happens for all the controversial topics. In the case of *Democrats* the mean value increases are more modest than those experienced by the *Republicans*. This was to be expected given that, in our *aggregated graphs*, more users are identifies as *Republicans* than they are as *Democrats*. The findings indicate a tendency of users to shift towards users of a similar political ideology. Hence, the users from each main cluster predominantly of the same political ideology shift towards one another, while the communities of users of opposing political ideology, that are located inside these clusters, tend to do the same, but around their politically defined mini-clusters. Therefore, the increase in the level of user activity marks an increase in political polarisation.

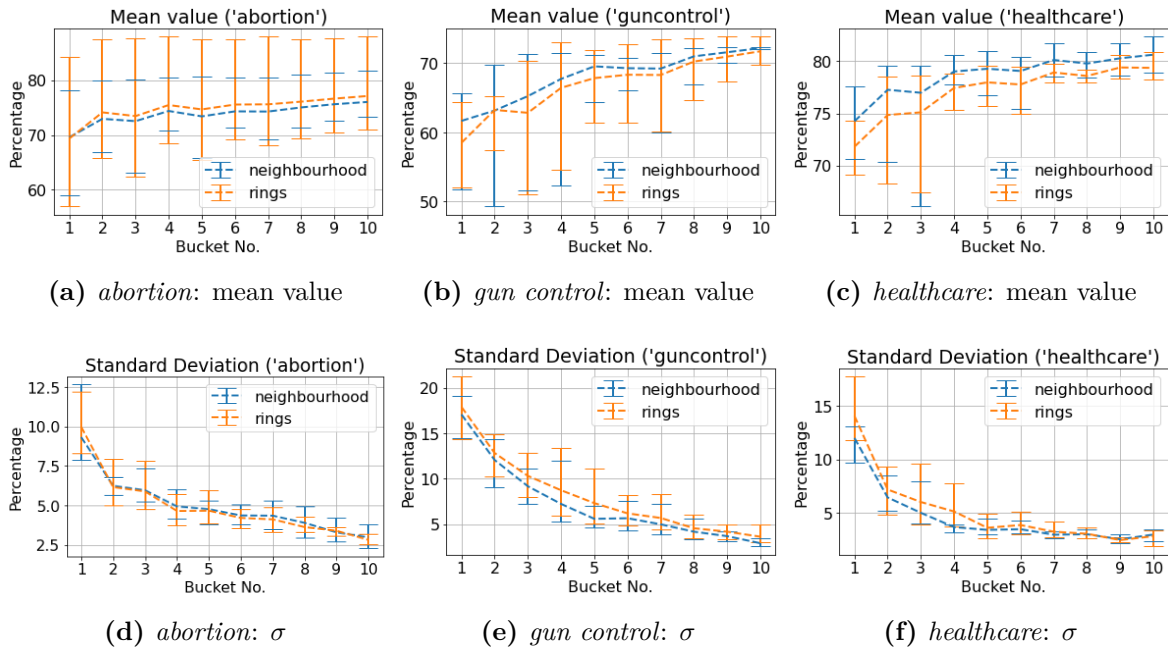


Figure 4.21: The percentage of *Republicans* that are closest to mostly (most frequent closest k%-neighbours) users of the same political ideology. Higher bucket numbers depict higher levels of user activity.

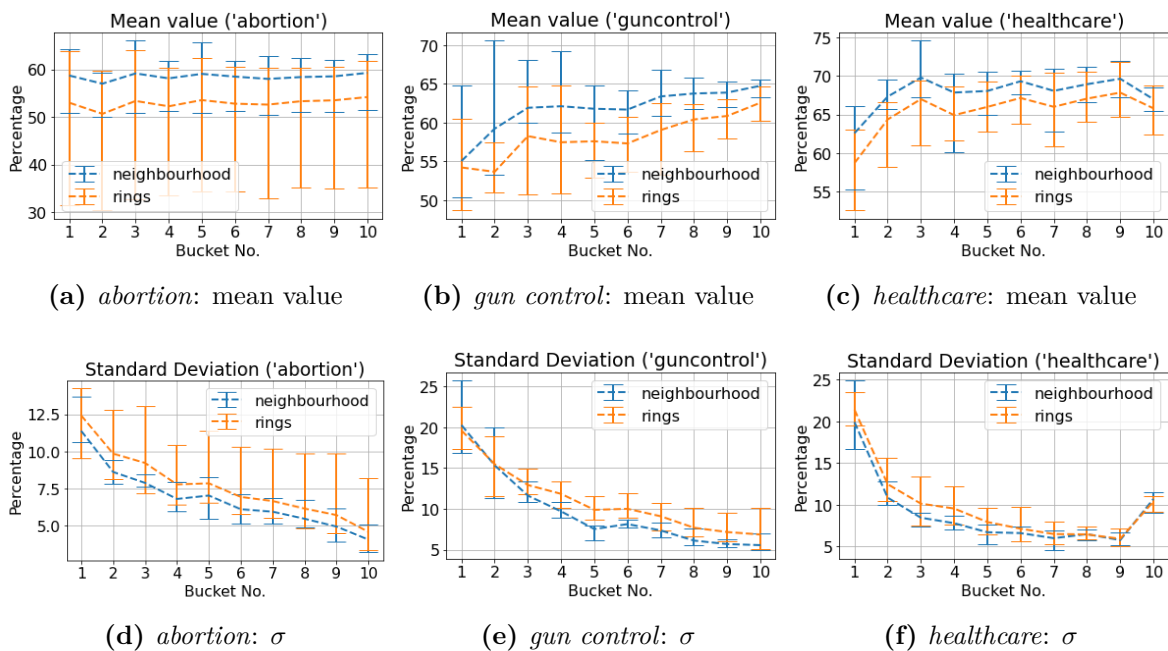


Figure 4.22: The percentage of *Democrats* that are closest to mostly (most frequent closest k%-neighbours) users of the same political ideology. Higher bucket numbers depict higher levels of user activity.

5. Conclusions

In this work, we introduced methods, based on graph embeddings, for analysing the political polarisation of Twitter interactions that pertained to controversial topics. This analysis revealed that external events, linked to the controversial topics, and that are reasonably politically polarising, cause spikes in the conversation’s overall level of political polarisation. The effect of these spikes is short lived, and after them, the conversation returns to pre-spike political polarisation levels. The trends revealed, that for the analysed period, the conversations political polarisation tended to stagnate. Structure-wise, we noticed that the pro and anti sides of a controversy have a tendency to match with the political divide between the two distinct political ideologies. Thus, each political ideology dominates one side on each controversy while forming tight-knitted groups in the other side of the controversy. Hence, even though there is a matching between political lines and controversy lines, for all the analysed topics, there are still relevant segments from each political ideology that are in disaccord with the predominant stance of their political partisanship. By analysing the methods in the context of user activity levels, we were able to determine that the controversial topics become more politically polarised as the activity increases.

5.1 Limitations and Future Work

All the methods, with the exception of intra-cluster distance, revealed relevant information about the polarisation of the controversial topics. In the case of the intra-cluster distance, the method worked on synthetic data while failing to do so on the real one. The method assumed a much clearer divide between the two political sides than the one found in the real data. A future work could see the retrofitting of this method so that it would take into account the revealed structure. This could be done by computing the intra-cluster distance, on each cluster defined by the controversial divide, only on the users whose political ideology is in line with the most prevalent political ideology found in the cluster in question. For the users who, based on their political ideology, leaves them in minority in the cluster, the intra-cluster distance could be measured for each neighbourhood. One would expect that the neighbourhood-based intra-cluster

distances to be comparable to one another, thus one could average over them, hence revealing the average intra-cluster ideological polarisation of the users who are not in line with the mainstream view of their political ideology.

Overall, the analysed methods required the use of all the data from the analysed interval. Thus, as the size of the analysed interval increases, so does the overall computational costs and the required computational times. The inter-cluster distance measures polarisation in a relative fashion, thus one could not take two time intervals, compute for each of them their inter-cluster distances separately and then compare the inter-cluster distances between the two interval. Therefore, a future work could give the definition of an absolute interpretation for the inter-cluster distances so that, two distinct time intervals, computed separately, could be compared one to another. This could be done by pivoting the inter-cluster distances from a time interval to the political ideological stances of influential individuals.

Bibliography

- [1] ABC News. How 15 Democrats helped tank the 2013 assault weapons ban. <https://abcnews.go.com/Politics/15-democrats-helped-tank-2013-assault-weapons-ban/story?id=50275295>, last accessed on 2020-08-29.
- [2] J. An, D. Quercia, and J. Crowcroft. Partisan sharing: Facebook evidence and societal consequences. In *Proceedings of the second ACM conference on Online social networks*, pages 13–24, 2014.
- [3] E. Bakshy, S. Messing, and L. A. Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015.
- [4] Ballotpedia. Federal judges nominated by Barack Obama. https://ballotpedia.org/Federal_judges_nominated_by_Barack_Obama, last accessed on 2020-08-29.
- [5] Ballotpedia. Obamacare lawsuits. https://ballotpedia.org/Obamacare_lawsuits, last accessed on 2020-08-29.
- [6] P. Barberá. Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data. *Political analysis*, 23(1):76–91, 2015.
- [7] P. Barberá, J. T. Jost, J. Nagler, J. A. Tucker, and R. Bonneau. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological science*, 26(10):1531–1542, 2015.
- [8] M. Bastian, S. Heymann, M. Jacomy, et al. Gephi: an open source software for exploring and manipulating networks. *Icwsm*, 8(2009):361–362, 2009.
- [9] BBC. Texas abortion bill is signed into law. <https://www.bbc.com/news/world-us-canada-23368980>, last accessed on 2020-08-29.
- [10] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.

-
- [11] Business Insider. Hillary Clinton: Oregon mass shooting is 'just beyond my comprehension'. <https://www.businessinsider.com/hillary-clinton-oregon-shooting-2015-10?r=USIR=T>, last accessed on 2020-08-29.
- [12] H. Cai, V. W. Zheng, and K. C.-C. Chang. A comprehensive survey of graph embedding: Problems, techniques, and applications. *IEEE Transactions on Knowledge and Data Engineering*, 30(9):1616–1637, 2018.
- [13] M. D. Conover, B. Gonçalves, J. Ratkiewicz, A. Flammini, and F. Menczer. 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*, pages 192–199. IEEE, 2011.
- [14] M. D. Conover, J. Ratkiewicz, M. R. Francisco, B. Gonçalves, F. Menczer, and A. Flammini. Political polarization on twitter. *Icwsn*, 133(2011):89–96, 2011.
- [15] C. Data61. Stellargraph machine learning library, 2018.
- [16] Deutsche Welle. US Senate rejects tighter gun control measures after Orlando. <https://www.dw.com/en/us-senate-rejects-tighter-gun-control-measures-after-orlando/a-19344358>, last accessed on 2020-08-29.
- [17] Á. Fernández Pascual et al. Advanced methods for dimensionality reduction and clustering: Laplacian eigenmaps and spectral clustering. Master's thesis, 2010.
- [18] S. Flaxman, S. Goel, and J. M. Rao. Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, 80(S1):298–320, 2016.
- [19] France24. Democratic senator stages nearly 15-hour filibuster to force vote on gun control. <https://www.france24.com/en/20160616-democratic-senator-wages-anti-gun-filibuster-orlando-shooting>, last accessed on 2020-08-29.
- [20] K. Garimella, G. De Francisci Morales, A. Gionis, and M. Mathioudakis. The effect of collective attention on controversial debates on social media. In *Proceedings of the 2017 ACM on Web Science Conference*, pages 43–52, 2017.
- [21] K. Garimella, G. De Francisci Morales, A. Gionis, and M. Mathioudakis. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the 2018 World Wide Web Conference*, pages 913–922, 2018.
- [22] K. Garimella, G. D. F. Morales, A. Gionis, and M. Mathioudakis. Quantifying controversy on social media. *ACM Transactions on Social Computing*, 1(1):1–27, 2018.

- [23] E. Gilbert, T. Bergstrom, and K. Karahalios. Blogs are echo chambers: Blogs are echo chambers. In *2009 42nd Hawaii International Conference on System Sciences*, pages 1–10. IEEE, 2009.
- [24] W. Hamilton, Z. Ying, and J. Leskovec. Inductive representation learning on large graphs. In *Advances in neural information processing systems*, pages 1024–1034, 2017.
- [25] History Channel. Sandy Hook school shooting. <https://www.history.com/this-day-in-history/gunman-kills-students-and-adults-at-newtown-connecticut-elementary-school>, last accessed on 2020-08-29.
- [26] S. Iyengar and K. S. Hahn. Red media, blue media: Evidence of ideological selectivity in media use. *Journal of communication*, 59(1):19–39, 2009.
- [27] M. Jacomy, T. Venturini, S. Heymann, and M. Bastian. Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. *PloS one*, 9(6):e98679, 2014.
- [28] D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146, 2003.
- [29] E. Lawrence, J. Sides, and H. Farrell. Self-segregation or deliberation? blog readership, participation, and polarization in american politics. *Perspectives on Politics*, 8(1):141–157, 2010.
- [30] A. J. Morales, J. Borondo, J. C. Losada, and R. M. Benito. Measuring political polarization: Twitter shows the two sides of venezuela. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(3):033114, 2015.
- [31] New York Times. Colorado Springs Planned Parenthood mass shooting. <https://www.nytimes.com/2015/11/28/us/colorado-planned-parenthood-shooting.html?smid=pl-share>, last accessed on 2020-08-29.
- [32] A. Y. Ng, M. I. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in neural information processing systems*, pages 849–856, 2002.
- [33] NPR. Biden Ryan vice presidential debate during the 2012 u.s. elections. <https://www.npr.org/2012/10/11/162754053/transcript-biden-ryan-vice-presidential-debate>, last accessed on 2020-08-29.

-
- [34] Reuters. New York enacts gun-control law, first since Newtown attack. <https://www.reuters.com/article/us-usa-guns-newyork/new-york-enacts-gun-control-law-first-since-newtown-attack-idUSBRE90E06L20130115>, last accessed on 2020-08-29.
- [35] Z. Ruan, G. Iniguez, M. Karsai, and J. Kertész. Kinetics of social contagion. *Physical review letters*, 115(21):218702, 2015.
- [36] L. M. Smith, L. Zhu, K. Lerman, and Z. Kozareva. The role of social media in the discussion of controversial topics. In *2013 International Conference on Social Computing*, pages 236–243. IEEE, 2013.
- [37] The Atlantic. Planned Parenthood undercover video release. <https://www.theatlantic.com/politics/archive/2015/07/planned-parenthood-abortion/398558/>, last accessed on 2020-08-29.
- [38] The Commission on Presidential Debates. October 3, 2012 Debate Transcript. <https://www.debates.org/voter-education/debate-transcripts/october-3-2012-debate-transcript/>, last accessed on 2020-08-29.
- [39] The Mercury News. House passes bill to keep government running but defund Obamacare. <https://www.mercurynews.com/2013/09/20/house-passes-bill-to-keep-government-running-but-defund-obamacare/>, last accessed on 2020-08-29.
- [40] The White House. FACT SHEET: New Executive Actions to Reduce Gun Violence and Make Our Communities Safer. <https://obamawhitehouse.archives.gov/the-press-office/2016/01/04/fact-sheet-new-executive-actions-reduce-gun-violence-and-make-our>, last accessed on 2020-08-29.
- [41] J. Török and J. Kertész. Cascading collapse of online social networks. *Scientific reports*, 7(1):1–8, 2017.
- [42] US Supreme Court. *King v. Burwell*, volume 576, docket no. 14-114.
- [43] US Supreme Court. *National Federation of Independent Business v. Sebelius*, volume 567, citation 567 U.S. 519.
- [44] US Supreme Court. *Ruling for Whole Woman’s Health v. Hellerstedt*, volume 579, docket no. 15-274.
- [45] U. Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416, 2007.

-
- [46] VoteScan. timeline for texas senate bill 5 of the 83rd texas legislature. <https://justfacts.votesmart.org/bill/16928/44981/138338/increases-abortion-facility-requirements-and-prohibits-abortions-after-20-weeks>, last accessed on 2020-08-29.