

Article

Measuring Polarization in Online Debates

Teresa Alsinet , Josep Argelich , Ramón Béjar  and Santi Martínez 

INSPIRES Research Center, University of Lleida, Jaume II, 69, 25001 Lleida, Spain; josep.argelich@udl.cat (J.A.); ramon.bejar@udl.cat (R.B.); santi.martinez@udl.cat (S.M.)

* Correspondence: teresa.alsinet@udl.cat

Abstract: Social networks can be a very successful tool to engage users to discuss relevant topics for society. However, there are also some dangers that are associated with them, such as the emergence of polarization in online discussions. Recently, there has been a growing interest to try to understand this phenomenon, as some consider that this can be harmful concerning the building of a healthy society in which citizens get used to polite discussions and even listening to opinions that may be different from theirs. In this work, we face the problem of defining a precise measure that can quantify in a meaningful way the level of polarization present in an online discussion. We focus on the Reddit social network, given that its primary focus is to foster discussions, in contrast to other social networks that have some other uses. Our measure is based on two different characteristics of an online discussion: the existence of a balanced bipartition of the users of the discussion, where one partition contains mainly users in agreement (regarding the topic of the discussion) and the other users in disagreement, and the degree of negativity of the sentiment of the interactions between these two groups of users. We discuss how different characteristics of the discussions affect the value of our polarization measure, and we finally perform an empirical evaluation over different sets of Reddit discussions about diverse classes of topics. Our results seem to indicate that our measure can capture differences in the polarization level of different discussions, which can be further understood when analyzing the values of the different factors used to define the measure.

Keywords: Reddit; user-based model; polarization; local search optimization



Citation: Alsinet, T.; Argelich, J.; Béjar, R.; Martínez, S. Measuring Polarization in Online Debates. *Appl. Sci.* **2021**, *11*, 11879. <https://doi.org/10.3390/app112411879>

Academic Editors: Aida Valls and Karina Gibert

Received: 17 November 2021

Accepted: 12 December 2021

Published: 14 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Nowadays, there is growing controversy regarding the emergence of polarization in discussions on social networks, and the responsibility of companies in this problem. For example, Facebook researchers have studied the spread of divisive content on the platform [1,2]. One of their findings is that users with hyperactive engagement are far more partisan on average than normal users. Facebook has launched some initiatives to try to mitigate the factors that may be helping the spread of such content [3]. For example, they re-calibrated the prioritization of content in users' News Feeds to give more preference to posts of friends and family over news content, justified by the findings that suggest that people derive more "meaningful social interactions", or MSI, when they engage with people they know, although it is not clear what exactly comprises a meaningful social interaction for Facebook. Other actions include improving the detection of fake accounts and removing recommendations of pages that violate Facebook Community Standards or are rated false by their fact checkers.

However, the findings of Facebook suggest that completely fixing the polarization problem may be difficult. As they commented in an internal company presentation [4], Facebook algorithms exploit the human brain's attraction to divisiveness and tend to feed users with more and more divisive content, which seems to ensure that users will spend more time on the platform. Publishers and political parties reorient their posts toward outrage and sensationalism since this produces high levels of comments and reactions. Some political parties in Europe even told Facebook that the algorithm made

them shift their policy positions. However, because engagement is of course fundamental for the profit of the company, social network companies should be cautious to not react disproportionately when trying to mitigate polarization.

In spite of all these issues, it is worth noticing that it is not clear how much responsibility we can give to social networks on these problems. For example, there is some work that indicates that, at least for polarization in people's views of political parties, the increased use of social networks and the internet may not be necessarily increasing polarization [5]. Putting aside whether specific social network platforms are more or less responsible for polarization, it is clear that what companies, or policymakers, can do is to define more transparent ways to monitor such possible non-desirable behaviors so that we can decide to act only in situations where there is some objective measure that polarization is taking place and to a certain level of severity.

Previous work has studied the presence of polarization in different concrete examples, trying to analyze the relevant characteristics in these cases. For example, the works [6,7] studied the emergence of so-called echo chambers, where users are found to interact mainly only with the users that agree with them and to have very few interactions with other groups. However, online discussions in social networks can also show polarization where there are answers with negative sentiment between different groups, which can be considered the most worrying situation. For example, in [8] they studied hyperpartisanship and polarization in Twitter during the 2018 Brazilian presidential election. Their findings showed that there was more interaction within each group (pro/anti Bolsonaro) but there was also an interaction between both groups. Actually, there are also cases where the interaction between groups can be more relevant than those within groups, like in the case studied in [9] where the analysis of the 2016 U.S. elections on Reddit showed a significant number of interactions between pro-Trump and pro-Hillary supporters. It is worth noting that the existence of polarization in debates could be partially explained as a consequence of the spiral of silence theory [10]. In this theory, the public opinion about emotionally and morally laden issues is considered to be driven toward one opinion that receives the strongest support from people and mass media. In contrast, people with contrary opinions suffer isolation pressure from the supporters of the majority opinion, which makes them hide their opinions (or not defend them strongly). Thus, the existence of polarization around a certain topic can be explained as the result of a spiral of silence, where two strongly defended, but opposed, opinions tend to attract the attention of all the people and the isolation pressure forces them to join one of the sides. So, we may consider that, in fact, two opposing spirals emerge as a result. However, it is clear that around certain topics, in certain societies, there may exist a clear majority opinion that can make it difficult for the emergence of secondary publicly supported opinions. For example, in [11], it was studied the effect of governmental internet surveillance around the controversial topic of U.S. airstrikes against ISIS terrorists.

Our main focus in this work is to give a more clear and quantitative model for measuring polarization in an online debate such that this behavior can be monitored for generating a warning signal when necessary, that is, to detect communication patterns where users seem to interact positively only with a fixed group of users and negatively with the rest. Among the many online debating platforms that exist, in this work, we consider Reddit. Reddit (available at <http://www.reddit.com/>) is a social news aggregation, web content rating, and discussion website. Users submit content to the site, such as links, text posts, and images, which are then voted up or down by other members who, in turn, can comment on others' posts to continue the conversation. This online debating platform is widely used to create long and deep debates with comments and answers to comments, where, thanks to the almost unlimited text length of Reddit comments (40,000 characters), users can express their opinions more accurately, compared to other online debating platforms, such as Twitter which restricts the number of characters to 280.

As previous works seem to indicate that in a debate we can have interactions in two ways (within groups and between groups), an important aspect of our model is that it

allows us to quantify the relevance of the kinds of interactions present. Our model is inspired by the model used in [12] to identify supporting or opposing opinions in online debates, based on finding a maximum cut in a graph that models the interactions between users. In our case, as we are interested in quantifying polarization, we define a model that is based on a weighted graph and with labeled edges, where node weights represent the side of the user in the debate and edge labels represent the overall sentiment between two users. Then, given a bipartition of this graph, in our model, the polarization degree of the bipartition is based on how homogeneous each partition is and how negative the interactions are between both partitions. Finally, our measure of debate polarization is based on the maximum polarization we have in all the possible bipartitions of the graph.

The structure of the rest of the paper is as follows. In Section 2, we present our user-based model to represent Reddit debates, the User Debate Graph, where a node represents a user's opinion as to its whole set of comments in the debate. In Section 3, we define a measure to quantify the polarization in a debate that is based on a value defined over bipartitions of the user debate graph. In Section 4, we introduce a greedy local search optimization algorithm for finding the polarization of a debate, based on searching a bipartition with the highest possible polarization value. Finally, in Section 5, we perform an empirical evaluation of the polarization value obtained with our algorithm with different Reddit debates.

2. User-Based Model for Reddit

In this section, we present a computational model to represent a Reddit debate that allows us to study the polarization of the debates based on the interactions between the participants. In Reddit, debates are generated from a main (root) comment that contains a link to some news and a set of comments, where each comment, except for the root comment, answers exactly one previous comment, usually by another user or author. These answers between comments lead to positive, negative or neutral interactions, and their analysis will allow us to classify users into three groups: users that are in agreement with the root comment of the debate, users that are in disagreement with it, and undecided users.

In order to analyze the agreement and disagreement between user comments in Reddit debates, recently in [13], we defined an analysis system that represents a Reddit debate as a two-sided debate graph. In this graph, comments are divided into two groups, the ones that agree with the root comment of the debate, and the ones that disagree with it. The edges of the graph represent disagreement between the comments of the two groups. So, in this work, we use the two-sided debate model as a starting point to build the user-based model to study the polarization of the debates based on the interactions between the participants.

Following [13], we first formalize the notions of a comment and Reddit debate for a root comment.

Definition 1 (Reddit comment and debate). *A Reddit comment is a pair $c = (m, u)$, where m is the text of the comment and u is the user's identifier of the comment.*

Let $c_1 = (m_1, u_1)$ and $c_2 = (m_2, u_2)$ be two comments. We say that c_1 answers c_2 if c_1 is a reply to the comment c_2 .

Let $r = (m_r, u_r)$ be a comment such that m_r contains a link to some news. A Reddit debate on r is a non-empty set Γ of Reddit comments such that $r \in \Gamma$ and every comment $c \in \Gamma$, $c \neq r$, c answers a previous comment in Γ . We refer to r as the root comment of Γ .

Given the structure of a Reddit debate Γ on a root comment r , the next step is to extract the relationships between the comments in Γ . We represent Γ as a labeled tree, where each comment gives rise to a node, and edges denote answers between comments and are labeled with a value in the real interval $[-2, 2]$. The label for an edge (c_1, c_2) denotes the *sentiment* expressed in the text of the comment c_1 in response to the text of the comment c_2 so that the value -2 denotes a total disagreement and the value 2 a total agreement. We use the sentiment value 0 to denote both answers expressing a neutral position with respect

to the opinion expressed in c_2 , and answers expressing, at the same time, agreement with part of the opinion expressed in c_2 , and disagreement with another part of c_2 .

Definition 2 (Debate Tree). Let Γ be a Reddit debate on a root comment r . A Debate Tree (DebT) for Γ is a tuple $\mathcal{T} = \langle C, r, E, W \rangle$ such that the following is true:

- For every comment in Γ , there is a node in C ;
- Node $r \in C$ is the root node of \mathcal{T} ;
- If c_1 answers c_2 , then there is a directed edge (c_1, c_2) in E , and
- W is a labeling function of edges $W : E \rightarrow [-2, 2]$, where the value assigned to an edge denotes the sentiment of the answer, from highly negative (-2) to highly positive (2).

Only the nodes and edges obtained by applying this process belong to C and E , respectively.

Given a Reddit debate, we make its corresponding DebT using the Python Reddit API Wrapper (PRAW, available at <https://github.com/praw-dev/praw>) to download its set of comments, and then we evaluate the sentiment for an edge (c_1, c_2) in the DebT using the sentiment analysis software of [14] using the text of the comment c_1 , where the value assigned denotes the sentiment of the answer, from highly negative (-2) to highly positive (2).

In Figure 1, we show the DebT structure for a Reddit debate of the subreddit World News (r/worldnews) (<https://www.reddit.com/r/worldnews/comments/f62i35> accessed on 4 October 2020). Each comment is represented as a node, and each answer between comments is represented as an edge. The graph has 42 nodes and 41 edges since each comment responds exactly to one comment, except for the root comment. These 41 answers between comments are classified as agreement edges (painted in green color) and disagreement edges (painted in red color), and there is no neutral answer. The darkness of the color is directly proportional to the sentiment of the answer concerning the maximum value. The root comment of the discussion is labeled with 0, and the other comments are labeled with consecutive identifiers according to their generation order. Notice that the graph is acyclic since a comment only answers a previous comment in the discussion. Additionally, most of the answers are disagreement answers, and they seem to generate the longest discussion paths in the DebT. However, this model does not allow to infer directly neither the set of the most dominant users' opinions between the different users in the debate, nor the position of every user concerning the root comment since the information of each user is represented by multiple nodes and edges in the DebT.

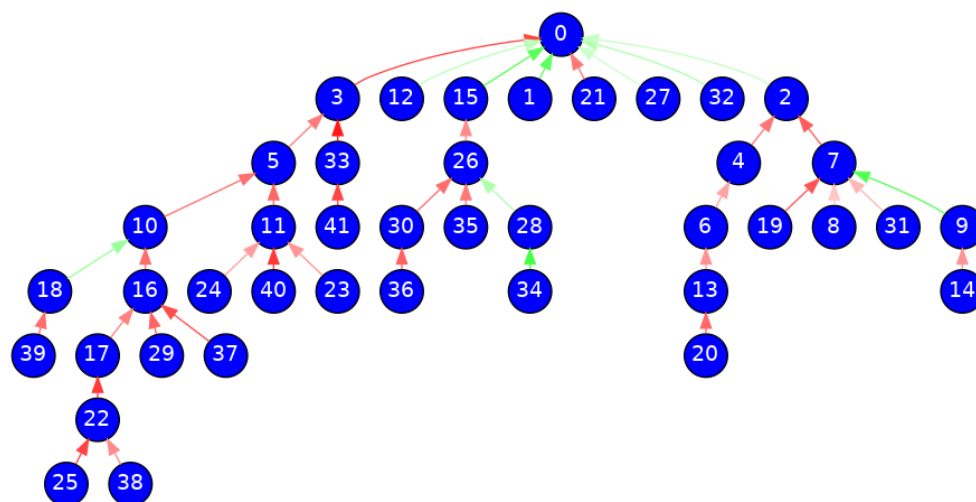


Figure 1. Debate tree for a Reddit debate of the subreddit World News.

Analyzing the relationships between the DebT nodes, we classify the comments into two groups: comments that support the root comment and comments that disagree with

it. With this objective, we extend the DebT structure with a labeling function of nodes denoting the side of each comment in the debate. We label the comments that support the root comment with 1 and the rest of the comments with -1 . We refer to this tree extension as a *two-sided debate tree*.

Definition 3 (Two-sided debate tree). Let $\mathcal{T} = \langle C, r, E, W \rangle$ be a DebT for a Reddit debate Γ on a root comment r . A *two-sided debate tree* (SDebT) for \mathcal{T} is a tuple $\mathcal{T}_S = \langle C, r, E, W, S \rangle$ such that S is a labeling function of nodes $S : C \rightarrow \{-1, 1\}$ defined as follows:

- $S(r) = 1$;
- For all node $c_1 \neq r$ in C , $S(c_1) = 1$ if for some node $c_2 \in C$, $(c_1, c_2) \in E$ and either $S(c_2) = 1$ and $W(c_1, c_2) > 0$, or $S(c_2) = -1$ and $W(c_1, c_2) \leq 0$; otherwise, $S(c_1) = -1$.

To evaluate the side of a comment c_1 expressing a neutral position with respect the opinion expressed in c_2 , i.e., with $W(c_1, c_2) = 0$, we consider a similar approach to the proposal of Murakami and Raymond [12]. In their proposal, they consider neutral opinions to be (slightly) negative. The motivation for this is the realization of Agrawal et al. [15] that, in newsgroups, people respond more frequently when they disagree. In Reddit, for example, when users agree with a comment, they can simply upvote without leaving any response. When users disagree, on the other hand, they can downvote, but it is more likely that they expose their reasons. Thus, from this perspective, a neutral response would be expressing more disagreement than agreement.

Figure 2 shows the SDebT structure we obtain for the Reddit debate of Figure 1. The nodes (comments) that support the root comment are colored in cyan, while the rest are colored in navy blue.

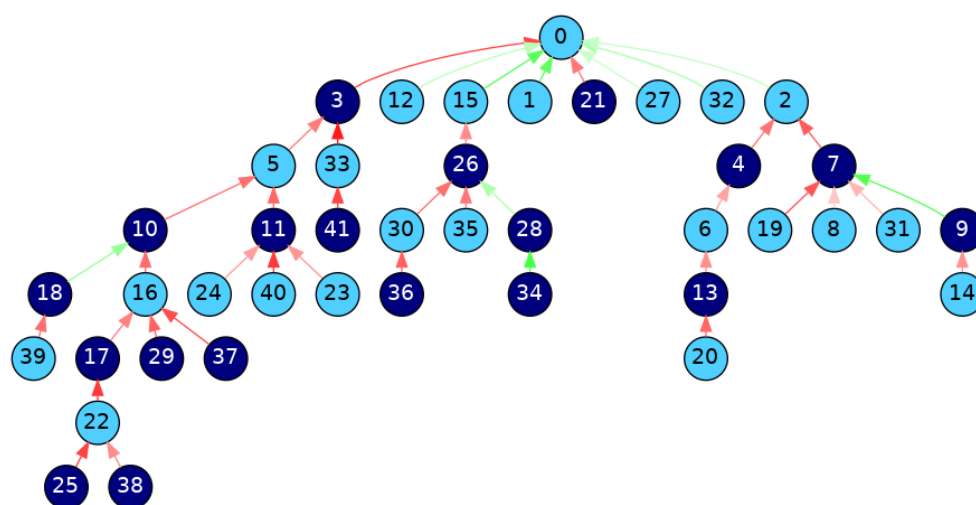


Figure 2. Two-sided debate tree for the Reddit debate of Figure 1.

At this point, we have classified the comments of a Reddit debate according to whether or not they offer support to the root comment and we have labeled the interactions between them according to the sentiment they express. So, to develop a quantitative model for measuring polarization in a Reddit debate, the next step is to introduce and investigate a suitable user-based model that allows us to represent the different interactions between the users to discover groups that either support or reject the root comment of the debate and how far these groups can be considered to be concerning each other.

To this end, once the comments are classified into two sides, in favor or against the root comment, we group comments by the user and we consider that the relationship between users' opinions of two users is defined from the agreement and disagreement relationships between the individual comments of these two users. When we represent a debate grouping comments by users, interactions between different users can give rise to

circular agreement and disagreement relationships, and the agreement or disagreement of a user concerning the opinion of another user in the debate should be defined by aggregating the set of single interactions that have occurred between them during the debate.

Given a debate Γ on a root comment r with users' identifiers $U = \{u_1, \dots, u_m\}$, we define the *opinion* of the user $u_i \in U$, denoted C_i , as the set of comments of u_i in Γ , except for the root comment r . We consider debates in which users are not self-referenced. That is, for all user $u \in U$ and each pair of comments $c_1 = (m_1, u, sc_1)$ and $c_2 = (m_2, u, sc_2)$, we assume c_1 does not respond to c_2 , nor c_2 to c_1 . Considering the particular case of root users (the users who post the root comment), we notice that when they do not participate in the debate, their opinion is empty, denoting that they have only posted the (root) news while staying passive throughout the debate. This is intentional since the root comment plays a special role in the debate, setting its topic. Thus, to be considered a "true participant" in the debate, the root user should contribute during the discussion. Notice that the Reddit platform itself distinguishes between root and non-root comments, as it provides two different global user metrics, Post Karma and Comment Karma, where the first one corresponds to the points achieved by posting interesting news (root comments) and the second one corresponds to the points achieved from non-root comments (debate generated on some root comment).

Next, we formalize the graph that we propose to represent user-based debates, called *User Debate Graph*, where the nodes are the users of the debate denoting their opinion concerning the root comment and the edges denote interactions between users mined from the prevailing sentiment among the aggregated comments of nodes. In addition, we define two weighting schemes: an *opinion weighting scheme* for nodes that associates every node of the graph with a side value representing the side of the user in the debate and an *interaction weighting scheme* for edges that associates every edge of the graph with a pair of values representing the overall sentiment of agreement or disagreement between users. For both schemes, we propose a *skeptical* approach based on stating that a user completely agrees or disagrees with a root comment or with another user if and only if one can be completely sure of it.

In a debate, a user can answer comments of different users, and thus, can agree or disagree with several users. This fact is represented in the User Debate Graph with a different edge for each user. However, if a user $u_i \in U$ answers several comments of a same user $u_j \in U$, the interaction between them is represented with a single edge in the User Debate Graph, and with a single sentiment value which is meant to be defined from the set of the sentiment of the answers of the user u_i to the user u_j , i.e., from the set of weights $\{W(c_1, c_2) \mid (c_1, c_2) \in E \text{ and } c_1 \in C_i \text{ and } c_2 \in C_j\}$. Moreover, a user can agree with part of a user's opinion and disagree with the rest. To reflect the possible ambivalence between the users' opinions, we also attach the edges of the User Debate Graph with a value that represents the ratio of interactions that agree with the users' opinions.

Retaining this information is crucial for the analysis of the debate since, depending on how we aggregate the sentiments, a single highly negative comment could outweigh several moderately positive ones (or vice versa). By storing both values (aggregated sentiment and ratio of positive answers), we can differentiate between consistent interactions with moderate opinions and interactions whose aggregated sentiment *seems* moderate but is a combination of inconsistent (positive and negative) responses.

Definition 4 (User Debate Graph). Let Γ be a Reddit debate on a root comment r with users' identifiers $U = \{u_1, \dots, u_m\}$ and let $\mathcal{T}_S = \langle C, r, E, W, S \rangle$ be a SDebT for Γ . A User Debate Graph (UDebG) for \mathcal{T}_S is a tuple $\mathcal{G} = \langle \mathcal{C}, \mathcal{E}, S, W \rangle$, where:

- \mathcal{C} is the set of nodes of \mathcal{G} defined as the set of users' opinions $\{C_1, \dots, C_m\}$; i.e., $\mathcal{C} = \{C_1, \dots, C_m\}$ with $C_i = \{(m, u_i, sc) \in \Gamma \mid (m, u_i, sc) \neq r\}$, for all users $u_i \in U$.
- $\mathcal{E} \subseteq \mathcal{C} \times \mathcal{C}$ is the set of edges of \mathcal{G} defined as the set of interactions between different users in the debate, i.e., there is an edge $(C_i, C_j) \in \mathcal{E}$, with $C_i, C_j \in \mathcal{C}$ and $i \neq j$, if and only if for some $(c_1, c_2) \in E$ we have that $c_1 \in C_i$ and $c_2 \in C_j$.

- S is an opinion weighting scheme for \mathcal{C} that expresses the side of users in the debate based on the side of their comments. We define S as the mapping $S : \mathcal{C} \rightarrow [-1, 1]$ that assigns to every node $C_i \in \mathcal{C}$ the value

$$S(C_i) = \frac{\sum_{c \in C_i} S(c_i)}{|C_i|}$$

in the real interval $[-1, 1]$ that expresses the side of the user u_i with respect to the root comment, from strictly disagreement (-1) to strictly agreement (1), going through undecided opinions (0).

- \mathcal{W} is an interaction weighting scheme for \mathcal{E} that expresses both the ratio of positive interactions between the users' opinions and the overall sentiment between users by combining the individual sentiment values assigned to the responses between their comments.

Let \oplus be an aggregation operator of dimension n on the real interval $[-2, 2]$, i.e., a bounded, monotonic, symmetric, and idempotent mapping $\oplus : [-2, 2]_n \rightarrow [-2, 2]$. We define \mathcal{W} as the mapping $\mathcal{W} : \mathcal{E} \rightarrow ([0, 1] \times [-2, 2])$ that assigns to every edge $(C_i, C_j) \in \mathcal{E}$ the pair of values $(p, w) \in ([0, 1] \times [-2, 2])$ defined as follows:

$$p = \frac{|(c_1, c_2) \in E \cap (C_i \times C_j) \text{ with } W(c_1, c_2) > 0|}{|(c_1, c_2) \in E \cap (C_i \times C_j)|} \text{ and}$$

$$w = \oplus_{\{(c_1, c_2) \in E \cap (C_i \times C_j)\}} W(c_1, c_2),$$

where p expresses the ratio of positive answers from the user u_i to the user u_j in the debate, and w expresses the overall sentiment of the user u_i regarding the comments of the user u_j , from highly negative (-2) to highly positive (2).

Only the nodes and edges obtained by applying this process belong to \mathcal{C} and \mathcal{E} , respectively.

Notice that if a user only responds to the news of the debate (the root comment r), the user is mapped in the UDebG to a node in \mathcal{C} with zero output degree denoting that the user starts no discussion with other users. In addition, users whose comments do not generate answers from other users are represented with nodes whose input degree is zero. Therefore, isolated nodes may appear in the UDebG that correspond to users who have neither generated nor participated in the debate, that is, users that have only answered to the news and whose opinions can be considered to be accepted by all users since they have not been discussed yet.

In the UDebG, each node denotes a user's opinion, and relationships between nodes are mined from the prevailing sentiment among the aggregated comments of those nodes. For instance, if we instantiate the aggregation operator \oplus to the minimum function, we obtain a pessimistic interpretation of the degree of agreement and disagreement among users since it represents the fact that in a debate with multiple negative and positive responses from one user to another, a pessimistic analysis focuses on the most negative response and in the case of multiple positive responses, on the softest positive response. Similarly, the maximum operator corresponds to an optimistic interpretation and the mean operator to a more realistic interpretation based on an intermediate weight. Furthermore, the ratio of positive answers is either 1 or 0, only for strictly agreement or disagreement relationships, respectively, between users' opinions; thus, the ratio of positive answers is a value in $(0, 1)$ only for ambivalent relationships.

Finally, since the edges between nodes in the UDebG can reflect ambivalent relationships between users' opinions, in our current user-based model, we do not consider indirect (transitive) relationships between users, and thus, the UDebG shows interactions between users only if there is enough evidence of such relationships through direct answers between authors' comments. Moreover, the UDebG for a given SDebT may contain cycles, and these cycles provide fundamental information about the different relationships that are established from the interactions between different users.

Figure 3 shows the UDebG structure we obtain for the two-sided debate tree of Figure 2. Each user's opinion is represented as a node, and each relationship between them is represented as an edge. The graph has 27 nodes and 31 edges; the overall sentiment for edges is evaluated using the mean operator, which leads to three agreement relationships, with the rest being disagreement. The edges are colored in green and red to denote agreement and disagreement relationships between users' opinions, respectively, and the darkness of the color is directly proportional to the ratio of positive and negative answers among analyzed users. The nodes are colored with different colors to denote the side of users' opinions in the debate based on the weighting scheme \mathcal{S} . The nodes colored in cyan correspond with users whose opinion supports the root comment, i.e., nodes $C_i \in \mathcal{C}$ with $\mathcal{S}(C_i) > 0$, the nodes colored in navy blue denote disagreement opinions, i.e., $\mathcal{S}(C_i) < 0$, and white nodes denote undecided opinions, i.e., $\mathcal{S}(C_i) = 0$. In this case, the darkness of the color reflects the proportion of user comments for and against the root comment.

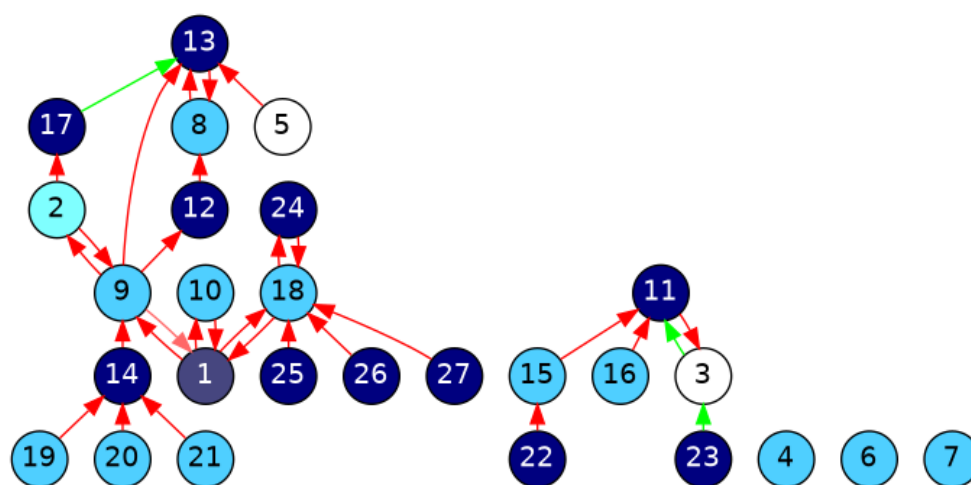


Figure 3. User Debate Graph for the two-sided debate tree of Figure 2.

3. Debate Polarization

Given a User Debate Graph $\mathcal{G} = \langle \mathcal{C}, \mathcal{E}, \mathcal{S}, \mathcal{W} \rangle$, we propose a model to measure the level of polarization in the debate between its users. We identify two characteristics that a polarization measure should capture. First, a polarized debate should contain a bipartition of \mathcal{C} into two sets (L, R) such that the set L contains mainly users in disagreement, the set R contains mainly users in agreement, and both sets should be similar in size. The second ingredient is the sentiment between users of L and R . A polarized discussion should contain most of the negative interactions between users of L and users of R , whereas the positive interactions, if any, should be mainly within the users of L and within the users of R .

We believe that both factors are relevant, as the existence of only the first characteristic can indicate simply a set of two *echo chambers*, where within each group, everybody agrees with the rest of the group but nobody pays attention to the other group. So, negative interactions between users of the two sides are the second ingredient for the right measurement of the polarization, such that the more negative the interactions between users of L and R , the higher the polarization we interpret between both sides. In our measure, the echo chambers case is considered as a case with middle polarization.

To capture these two characteristics with a single value, we define two different measures and their combination in a final one referred to as the *bipartite polarization level*. First, given a bipartition (L, R) of \mathcal{C} , we define its level of consistency and how balanced the sizes of L and R are as follows:

$$SC(L, R, \mathcal{G}) = LC(L, \mathcal{G}) \cdot RC(R, \mathcal{G}),$$

where $LC(L, \mathcal{G})$ evaluates the disagreement strength of the comments in L (regarding the root comment) and $RC(R, \mathcal{G})$ the agreement strength of the comments in R . They are defined as follows:

$$LC(L, \mathcal{G}) = \frac{\sum_{C_i \in L, \mathcal{S}(C_i) \leq 0} -\mathcal{S}(C_i)}{|\mathcal{C}|} \text{ and}$$

$$RC(R, \mathcal{G}) = \frac{\sum_{C_i \in R, \mathcal{S}(C_i) > 0} \mathcal{S}(C_i)}{|\mathcal{C}|}.$$

The value of $SC(L, R, \mathcal{G})$ lies in the interval $[0, 0.25]$. The minimum value (0) is achieved when either there are no users in L with $\mathcal{S}(C_i) < 0$ or no users in R with $\mathcal{S}(C_i) > 0$. The maximum value (0.25) is achieved when half of the users are users in disagreement with $\mathcal{S}(C_i) = -1$ and are found in L , and the other half are users in agreement with $\mathcal{S}(C_i) = 1$ and are found in R .

Secondly, the sentiment of the interactions between users of different sides is defined as follows:

$$SWeight(L, R, \mathcal{G}) = \frac{\sum_{e_i \in \mathcal{E} \cap ((L \times R) \cup (R \times L))} -c(p(e_i)) \cdot w(e_i)}{|\mathcal{E}|} + 2,$$

with $c(p(e_i)) = 2(p(e_i) - 0.5)^2 + 1/2$, and where $p(e_i)$ and $w(e_i)$ denote the values of p and w , respectively, in $\mathcal{W}(e_i) = (p, w)$. It is worth noticing that the value of $SWeight(L, R, \mathcal{G})$ lies in the interval $[0, 4]$:

- The minimum value (0) is achieved when all the interactions $e_i \in \mathcal{E}$ have $\mathcal{W}(e_i) = (1, 2)$ (pure positive interactions with maximum sentiment value), and all of them are found between L and R .
- The middle value (2) is achieved when the sum of negative values in the summatory of the numerator equals the sum of the positive ones. So, this represents a situation where positive and negative interactions between both partitions compensate for each other. As a special case, observe that if there are no interactions of any type between them (an echo chamber situation), then we also achieve this middle value.
- The maximum value (4) is achieved when all the interactions $e_i \in \mathcal{E}$ have $\mathcal{W}(e_i) = (0, -2)$ (pure negative interactions with minimum sentiment value), and all of them are found between L and R .

Finally, we combine the $SC(L, R, \mathcal{G})$ and the $SWeight(L, R, \mathcal{G})$ measures to define the *bipartite polarization level* of a given bipartition (L, R) as follows:

$$BipPol(L, R, \mathcal{G}) = SC(L, R, \mathcal{G}) \cdot SWeight(L, R, \mathcal{G}).$$

So, given the range of values for the previous measures, the value of $BipPol(L, R, \mathcal{G})$ lies in the interval $[0, 1]$:

- The minimum value (0) is achieved when either $SC(L, R, \mathcal{G}) = 0$ or $SWeight(L, R, \mathcal{G}) = 0$. If this happens with $SC(L, R, \mathcal{G}) = 0$ and $SWeight(L, R, \mathcal{G}) > 0$ then in L , we have only neutral and positive users, at least one user in R , and some interactions between L and R that are not extremely positive (such that $SWeight(L, R, \mathcal{G})$ can be positive). So, we consider that not having users in disagreement in L with $\mathcal{S}(C_i) < 0$ gives the lowest polarization level in the debate. We can also have $SC(L, R, \mathcal{G}) > 0$ and $SWeight(L, R, \mathcal{G}) = 0$ if both L and R contain nodes of both types (with $\mathcal{S}(C_i) < 0$ and with $\mathcal{S}(C_i) > 0$), and all the edges are found between L and R and have $\mathcal{W}(e_i) = (1, 2)$. This implies that the more edges we have in \mathcal{E} , the more different pairs of vertices (C_i, C_j) we will have with the same sign in $\mathcal{S}(C_i)$ and $\mathcal{S}(C_j)$, but in different sides of the partition. So, in this second case, a big quantity of edges implies a small value for $SC(L, R, \mathcal{G})$, when $SWeight(L, R, \mathcal{G}) = 0$.

- The middle value ($1/2$) can of course be achieved for many combinations of these two factors, but there are two canonical cases that reflect well the intended meaning of this middle case. They are the cases where one of the two factors has its maximum value and the other has its middle value. This can happen in two cases. First, this happens when $SC(L, R, \mathcal{G}) = 0.25$ and $SWeight(L, R, \mathcal{G}) = 2$. That is, when users in disagreement and users in agreement are perfectly balanced between L and R , as we discussed before when talking about the behavior of $SC(L, R, \mathcal{G})$, and either there are no interactions between L and R (echo chambers situation), or positive interactions between them are canceled by the negative ones (so we can say that, overall, the interactions are neutral between them). Second, this also happens when $SC(L, R, \mathcal{G}) = 0.25 \cdot 0.25$ and $SWeight(L, R, \mathcal{G}) = 4$. These values can be obtained when in L , the sum of negative $S(C_i)$ values represents 0.25 of the total number of comments, and analogously for R with the positive values, (so the partitions are maximally heterogeneous), and all the interactions $e_i \in \mathcal{E}$ have $\mathcal{W}(e_i) = (0, -2)$ and are found only between L and R . That is, both partitions have a balanced representation of users in agreement and disagreement, and we have pure negative interactions between users in agreement of one partition and the users in disagreement of the other one.
- The maximum value (1) is achieved when both factors have its maximum value. That is, perfect balance between users in disagreement in L and users in agreement in R ($SC(L, R, \mathcal{G}) = 0.25$) and all the interactions have $\mathcal{W}(e_i) = (0, -2)$ and are found only between L and R . So, this extreme case represents a situation where there are two perfect homogeneous sides in the debate (agreement and disagreement sides) with sentiments that are maximally opposed ($+1$ and -1) and the only talking between them is to criticize the other side with maximum strength, so there are no positive answers between the two sides. This can be considered a very extreme situation because even in the most tense debates in a social network, one usually can find some positive answers, although these positive answers may be concentrated within each side, or not all the negative answers will have the extreme value (-2).

4. Finding the Most Polarized Partition

In this section, we present a greedy local search optimization algorithm for finding a bipartition (L, R) of $\mathcal{G} = \langle \mathcal{C}, \mathcal{E}, \mathcal{S}, \mathcal{W} \rangle$ with a high value for $BipPol(L, R, \mathcal{G})$ that follows the same basic idea of common greedy algorithms for the maxcut problem, with the goal to find a partition with a polarization value that is close to the highest one. It seems very unlikely that we can have exact polynomial-time algorithms for our problem, as it can be shown that the maxcut problem can be reduced to a slight generalization of our optimization problem.

The algorithm follows these steps:

1. Initialize vertices in either L or R . The most obvious idea is to order vertices by polarity, and then assign the negative ones to L and the rest to R . Although this clearly gives the best possible value for $SC(L, R, \mathcal{G})$, other assignments could make improvements to the value of $SWeight(L, R, \mathcal{G})$. That is, it could be that the best partition (with respect to its value of $BipPol(L, R, \mathcal{G})$) is not a partition where all users' opinions of the same polarity are found on the same side of the partition because the value of $SWeight(L, R, \mathcal{G})$ can be significantly higher when some users' opinions with different polarity are placed on the same side. This class of users would be users that are, to some extent, inconsistent, in the sense that they seem to support one side of the debate, but at the same time, they seem to attack a significant number of users' opinions of the same side.

To consider these kinds of partitions as possible solutions for our search algorithm, we propose two possible ways to initialize the partition (L, R) :

- g0: Place each user's opinion uniformly at random in either L or R . This way, we do not initially force a bias toward consistent partitions.

- g1: Place each user's opinion C_i randomly in L with probability $P_L = \frac{1-S(C_i)}{2}$ and in R with probability $1 - P_L$. In this second way, only users with a strict disagreement ($S(C_i) = -1$) or a strict agreement ($S(C_i) = 1$) will have a unique possible side on the initial partition.
2. Perform local improvement (regarding the value of *BipPol*) steps: at each iteration, find a good candidate from L (to be moved to R) and a good candidate from R (to be moved to L).

The pseudocode is shown in Algorithm 1.

Algorithm 1: Finding a local optimal solution for a bipartition of $\mathcal{G} = \langle \mathcal{C}, \mathcal{E}, \mathcal{S}, \mathcal{W} \rangle$ with high bipartite polarization.

Input: $\mathcal{G} = \langle \mathcal{C}, \mathcal{E}, \mathcal{S}, \mathcal{W} \rangle$
Output: a bipartition (L, R) of \mathcal{G} with high bipartite polarization
 $(L, R) := \text{getInitBPart}(\mathcal{G})$ // Get Initial bipartition
 $\text{improving} := \text{true}$
 $\text{steps} := 0$
while ($\text{improving} \ \&\& \ \text{steps} \leq \text{maxsteps}$) **do**
 $\text{improving} := \text{false}$
 if $\exists v \in L, \text{BipPol}(L \setminus v, R \cup v, \mathcal{G}) > \text{BipPol}(L, R, \mathcal{G})$ **then**
 $(L, R) := (L \setminus v, R \cup v)$ // improved bipartition
 $\text{improving} := \text{true}$
 if $\exists v' \in R, \text{BipPol}(L \cup v', R \setminus v', \mathcal{G}) > \text{BipPol}(L, R, \mathcal{G})$ **then**
 $(L, R) := (L \cup v', R \setminus v')$ // improved bipartition
 $\text{improving} := \text{true}$
 $\text{steps} := \text{steps} + 1$
return (L, R)

Figure 4 shows the result of applying Algorithm 1 using the UDebG of Figure 3 as input. Nodes in red color are the nodes in the L set, and nodes in blue color are the nodes in the R set. The algorithm used to compute the initial partition is g1, and after this initialization, the algorithm performs a few steps depending on the randomness of g1. These steps are to reallocate nodes 1, 2, 3 and 5 to the sets that increase more the $\text{BipPol}(L, R, \mathcal{G})$ value, in case they are not already there. In this example, the reallocation of other nodes does not increase this value. We can observe that the only nodes within the same set that disagree (red arrows) between them are nodes 2 and 9 in the R set (blue nodes), and 3 and 11 in the L set (red nodes). All the other disagreement relations are between nodes of different sets, and all the agreement relations (green arrows) are between nodes of the same set. The final value of $\text{BipPol}(L, R, \mathcal{G})$ is 0.5161, being that the other values $LC = 0.4296$, $RC = 0.4419$, and $S_{weight} = 2.718$, as we will see in the next section, are common values for debates in the subreddit World News.

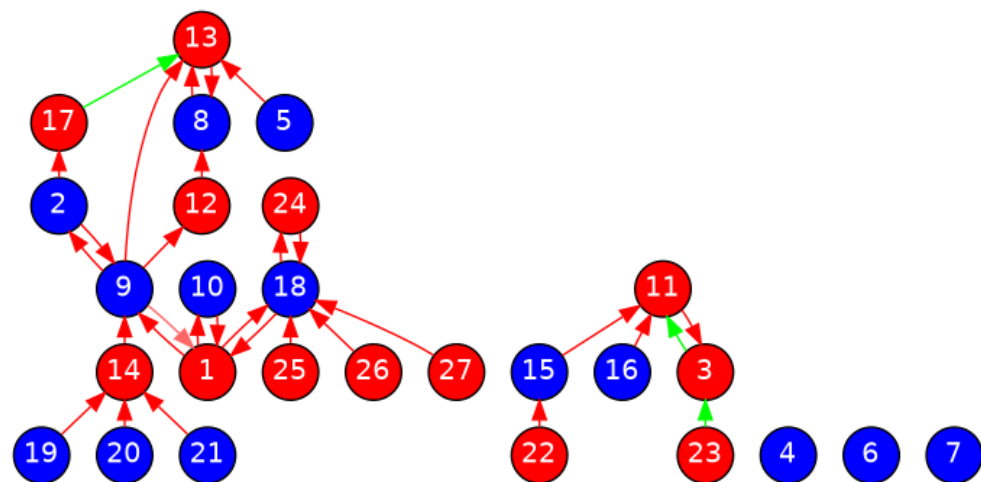


Figure 4. Bipartite graph showing sets L (red nodes) and R (blue nodes) after running Algorithm 1 using the UDebG of Figure 3 as input.

5. Empirical Evaluation

In this section, we present an empirical evaluation to detect if our optimization algorithm can find distinctive polarization in debates obtained from six representative subreddits with different topics: Halloween, travel, movies, Bitcoin, politics, and worldnews. Intuitively, we expect to have debates with low polarization in the Halloween and travel topics, those with higher polarization in the movies and Bitcoin topics, and those with the highest polarization in the politics and worldnews topics. From each subreddit, we downloaded debates from the last month, excluding the most recent debates that probably are still active, with comments per debate ranging from 50 to 500, and with a limit of 100 debates per subreddit.

Table 1 shows the average results for the polarization value (*BipPol*) obtained with our algorithm for each of the six sets of debates using the algorithm with the two initialization methods, so we have two rows per subreddit.

Table 1. Results with different initialization algorithms with debates from several subreddits. The numbers shown are average values over the debates of each subreddit.

Subreddit	#Debates	$ C $	$ E $	Alg.	<i>BipPol</i>	<i>LC</i>	<i>RC</i>	<i>SWeight</i>
Halloween	74	78.16	30.09	g0	0.4053	0.3681	0.5603	2.1594
				g1	0.4087	0.3685	0.5654	2.1643
travel	49	75.10	60.57	g0	0.4246	0.4835	0.4005	2.3808
				g1	0.4245	0.4836	0.4004	2.3808
movies	100	146.51	123.23	g0	0.4785	0.4826	0.4189	2.3930
				g1	0.4783	0.4825	0.4189	2.3926
Bitcoin	100	135.01	95.82	g0	0.4879	0.5332	0.3862	2.4360
				g1	0.4883	0.5333	0.3863	2.4367
politics	100	205.93	160.72	g0	0.5104	0.5713	0.3511	2.5742
				g1	0.5105	0.5712	0.3511	2.5752
worldnews	100	148.40	163.36	g0	0.5141	0.4950	0.4076	2.5732
				g1	0.5142	0.4949	0.4076	2.5732

In any case, the maximum number of iterations is fixed to the number of users in the debate graph. To analyze the reasons behind the *BipPol* values obtained, we also show the corresponding average values for the factors *LC*, *RC* (that make the *SC* value), and *SWeight*. The table also shows the total number of debates in each set, and the average values for the number of users and number of edges in their corresponding user debate graphs. The average results for the different subreddits show some differences in the *BipPol* values

such that we can divide the six subreddits into three groups: subreddits with the lowest *BipPol*, the middle one, and the highest one.

In the first group (Halloween and travel), we observe average *BipPol* values around 0.4. However, looking into the value of the different factors, we can explain why there is an slightly higher *BipPol* value for the travel subreddit. For the Halloween subreddit, we observe that the agreement users group is more relevant (*RC* value around 0.56 in contrast to a *LC* value around 0.36) and a *SWeight* value around 2.15 (close to the neutral value for the global sentiment between users in agreement and disagreement). By contrast, for the travel subreddit, the *LC* value is higher than the *RC* value, but they are more balanced (0.48 versus 0.40), and the *SWeight* value is around 2.38. So now, negative interactions between both groups are more significant. These differences make sense if one looks into the nature of the discussions in both these subreddits. In the Halloween case, discussions are mainly friendly, discussing ideas about costumes and Halloween decorations. However, in the travel case, we find some controversies in many discussions about certain touristic destinations. Regarding the differences between using either the *g0* or *g1* initialization method in the algorithm, we observe no significant differences in the final value obtained; the only difference we observe is that when using the uniform random method (*g0*), the number of iterations needed by the algorithm can be about ten times higher, but in all the debates we have solved, the algorithm has never reached its maximum number of allowed iterations. So, even if the method (*g0*) makes the algorithm make more iterations, it seems that the quality of the solution obtained is similar to the one with *g1*.

In the second group (movies and Bitcoin), we observe average *BipPol* values around 0.47. The increase with respect to the first group can be explained due to slightly more balanced values for *LC* and *RC* (close to the 0.5 value), although in the more hot topic of Bitcoin, the disagreement group value (*LC*) is more significant, and also due to increased values for *SWeight*, but again, this is more significant in the Bitcoin subreddit. It is worth noticing that a higher value for *SWeight* seems to be correlated with a higher size of the debate (in the values of $|C|$ and $|\mathcal{E}|$). This is consistent with what we commented earlier about the tendency to respond more frequently when people disagree.

In the last group (politics and worldnews), we observe the highest *BipPol* average values, around 0.51. Here the values of *LC* and *RC* show a balance similar to the one of the second group (politics is similar to Bitcoin, and worldnews is similar to movies). However, the difference lies in the increase in the *SWeight* values, that again seems to be correlated with an increase in the size of the debates. The existence of the highest values for the *SWeight* in this last group of subreddits is somehow natural, as many of the topics discussed are highly controversial.

To further investigate the characteristics of the debates of these different subreddits, we also looked into the results obtained for the debates of each group with the minimum, median, and maximum *BipPol* values within the group. Table 2 shows these results, with three rows per subreddit, showing the results for the minimum, median, and maximum cases.

Looking at the results for the minimum, we observe that clearly the lowest *BipPol* values are obtained in the Halloween and travel subreddits (0.069 and 0.101, respectively), in contrast to the minimum *BipPol* value for the other subreddits, which is at least 0.29. The median values show differences similar to the ones we observed for the average value in Table 1, from the median value of 0.41 for Halloween to the median value of 0.51 for worldnews, so the differences in the minimum value are more significant. The maximum values are at least 0.56 (the maximum for Halloween), but the other maximum values are similar, around 0.6. Curiously, the highest maximum value is observed in a debate from the travel subreddit (0.68). So, these results indicate that one can find polarized debates on many topics, but when looking at the average values of the previous table, we can conclude that the average polarization tends to be higher for certain topics.

Table 2. Results for the min/median/max values of *BipPol* of the same debates of Table 1.

Subreddit	DebateID	$ \mathcal{C} $	$ \mathcal{E} $	<i>BipPol</i>	<i>LC</i>	<i>RC</i>	<i>SWeight</i>
Halloween	qdg24n	27	27	0.0692	0.0370	0.9246	2.0219
	qg6q98	84	39	0.4156	0.5765	0.3294	2.1890
	q71wxm	41	4	0.5626	0.3780	0.5935	2.5074
travel	n1uxcn	16	25	0.1019	0.6771	0.0625	2.4072
	muovcd	34	15	0.4379	0.4706	0.3725	2.4979
	mu25ze	27	16	0.6857	0.5556	0.4296	2.8730
movies	q7g2fh	135	156	0.3400	0.4547	0.3728	2.0054
	q1vw6l	125	111	0.4809	0.4680	0.4320	2.3786
	q8975v	87	86	0.5939	0.5594	0.4103	2.5874
Bitcoin	mmx2og	32	23	0.2953	0.7585	0.1875	2.0761
	mwglkl	110	62	0.4912	0.4818	0.4498	2.2668
	mtux4x	43	12	0.6107	0.5349	0.4651	2.4546
politics	qhm2di	139	153	0.4318	0.5454	0.3120	2.5374
	qcrq1v	216	201	0.5082	0.4961	0.3997	2.5630
	qerwa6	217	141	0.5927	0.5668	0.3944	2.6513
worldnews	morqje	78	91	0.3754	0.4983	0.2949	2.5553
	mq580o	195	257	0.5175	0.4521	0.4434	2.5816
	morpw5	59	68	0.6393	0.5036	0.4669	2.7186

6. Conclusions and Future Work

In this work, we introduce a quantitative model for measuring polarization in an online debate such that this behavior can be monitored for generating a warning signal when the debate polarization reaches some threshold value.

Among the many online debating platforms that exist, in this work, we consider Reddit; we model a Reddit debate as a weighted graph and with labeled edges, where the weights of nodes represent the side of the users' opinions in the debate and the labels of edges represent the overall sentiments between users' opinions. Given a bipartition of this graph, we quantify the degree of polarization of the bipartition by measuring how homogeneous each partition is and how negative the interactions are between both partitions. Then, our quantitative model is based on the maximum polarization we have in all the possible bipartitions of the graph, and it is computed with a greedy local search optimization algorithm.

Finally, we develop an empirical evaluation of the debate polarization value obtained with our algorithm with different Reddit debates. The results indicate that our quantitative model captures differences in the polarization level of different discussions, which can be further understood when analyzing the values of the different factors used to define the measure.

As future work, we plan to study the suitability of other schemes for the aggregation of individual interactions between two users, when computing the sentiment weighting scheme, as well as different features to capture the degree of polarization of users' opinions in online debates. For example, to quantify the weight that corresponds to each node of the graph, it can be interesting to consider parameters available on the Reddit platform, such as the user's Karma and the comment score, since active users that post interesting comments can be considered more relevant than users who are not very active or whose comments are valued or rated with negative scores. Moreover, we would also like to consider the identification of other communication patterns that may not be considered as polite, such as, for example, when users are *ratioed* on an online discussion. That is, users whose comments are ignored, but their replies receive much more attention. For a discussion about getting ratioed on Twitter see <https://www.popbuzz.com/internet/social-media/ratioed-meaning-twitter/>.

Author Contributions: Conceptualization, T.A., J.A., R.B. and S.M.; methodology, T.A., J.A., R.B. and S.M.; software, T.A., J.A., R.B. and S.M.; validation, T.A., J.A., R.B. and S.M.; formal analysis, T.A., J.A., R.B. and S.M.; investigation, T.A., J.A., R.B. and S.M.; data curation, T.A., J.A., R.B. and S.M.; writing—original draft preparation, T.A., J.A., R.B. and S.M.; writing—review and editing, T.A., J.A., R.B. and S.M.; visualization, T.A., J.A., R.B. and S.M.; supervision, T.A., J.A., R.B. and S.M.; project administration, T.A., J.A., R.B. and S.M.; funding acquisition, T.A., J.A., R.B. and S.M. All authors have contributed equally to the work. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Spanish Project PID2019-111544GB-C22 (MINECO/FEDER), by the European Union’s Horizon 2020 Research and Innovation Program under Grant Agreements 723596, 768824, 764025 and 814945, and by 2017 SGR 1537.

Institutional Review Board Statement: We choose to exclude this statement since the study do not involve humans and animals.

Informed Consent Statement: We choose to exclude this statement since the study do not involve humans.

Data Availability Statement: We choose to exclude this statement since the study did not report any data.

Acknowledgments: The authors would like to thank the anonymous reviewers for providing helpful comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Newton, C. How to think about polarization on Facebook. *The Verge*, 28 May 2020.
2. Horwitz, J.; Seetharaman, D. Facebook Executives Shut Down Efforts to Make the Site Less Divisive. *The Wall Street Journal*, 26 May 2020.
3. Rosen, G. Facebook: Investments to Fight Polarization. 2020. Available online: <https://about.fb.com/news/2020/05/investments-to-fight-polarization/> (accessed on 27 May 2020).
4. Hagey, K.; Horwitz, J. Facebook Tried to Make Its Platform a Healthier Place. It Got Angrier Instead. *The Wall Street Journal*, 16 September 2021.
5. Boxell, L.; Gentzkow, M.; Shapiro, J.M. *Cross-Country Trends in Affective Polarization*; National Bureau of Economic Research: Cambridge, MA, USA, 2020.
6. Bessi, A.; Zollo, F.; Vicario, M.D.; Puliga, M.; Scala, A.; Caldarelli, G.; Uzzi, B.; Quattrocioni, W. Users Polarization on Facebook and Youtube. *PLoS ONE* **2016**, *11*, e0159641. [CrossRef] [PubMed]
7. Vicario, M.D.; Vivaldo, G.; Bessi, A.; Zollo, F.; Scala, A.; Caldarelli, G.; Quattrocioni, W. Echo chambers: Emotional contagion and group polarization on Facebook. *Sci. Rep.* **2016**, *6*, 37825. [CrossRef] [PubMed]
8. Recuero, R.; Soares, F.B.; Gruz, A.A. Hyperpartisanship, Disinformation and Political Conversations on Twitter: The Brazilian Presidential Election of 2018. In Proceedings of the Fourteenth International AAAI Conference on Web and Social Media, ICWSM 2020, Atlanta, GA, USA, 8–11 June 2020; Choudhury, M.D., Chunara, R., Culotta, A., Welles, B.F., Eds.; AAAI Press: Menlo Park, CA, USA, 2020; pp. 569–578.
9. Morales, G.D.F.; Monti, C.; Starnini, M. No echo in the chambers of political interactions on Reddit. *Sci. Rep.* **2021**, *11*, 2818. [CrossRef] [PubMed]
10. Noelle-Neumann, E. The Spiral of Silence a Theory of Public Opinion. *J. Commun.* **2006**, *24*, 43–51. [CrossRef]
11. Stoycheff, E. Under Surveillance: Examining Facebook’s Spiral of Silence Effects in the Wake of NSA Internet Monitoring. *J. Mass Commun. Q.* **2016**, *93*, 296–311. [CrossRef]
12. Murakami, A.; Raymond, R. Support or Oppose? Classifying Positions in Online Debates from Reply Activities and Opinion Expressions. In Proceedings of the COLING 2010, 23rd International Conference on Computational Linguistics, Beijing, China, 23–27 August 2010; Huang, C., Jurafsky, D., Eds.; Chinese Information Processing Society of China: Beijing, China, 2010; pp. 869–875.
13. Alsine, T.; Argelich, J.; Béjar, R.; Martínez, S. Measuring user relevance in online debates through an argumentative model. *Pattern Recognit. Lett.* **2020**, *133*, 41–47. [CrossRef]
14. Manning, C.D.; Surdeanu, M.; Bauer, J.; Finkel, J.; Bethard, S.J.; McClosky, D. The Stanford CoreNLP Natural Language Processing Toolkit; In Proceedings of the Association for Computational Linguistics (ACL) System Demonstrations, Baltimore, MD, USA, 22–27 June 2014; pp. 55–60.
15. Agrawal, R.; Rajagopalan, S.; Srikant, R.; Xu, Y. Mining newsgroups using networks arising from social behavior. In Proceedings of the 12th International Conference on World Wide Web, Budapest, Hungary, 20–24 May 2003; pp. 529–535.