

A Polarization Approach for Understanding Online Conflicts in Times of Pandemic: A Brazilian Case Study

Carlos Kamienski
Federal University of ABC
cak@ufabc.edu.br

Lucas Mazim de Sousa
Federal University of ABC
lucas.mazim@ufabc.edu.br

Claudio Penteado
Federal University of ABC
claudio.penteado@ufabc.edu.br

Denise Hideko Goya
Federal University of ABC
denise.goya@ufabc.edu.br

Daniel di Genova
Federal University of ABC
daniel.genova@ufabc.edu.br

Fabricio Olivetti de Franca
Federal University of ABC
folivetti@ufabc.edu.br

Diogo Fornaziero Ramos
Federal University of ABC
diogo.ramos@ufabc.edu.br

Flavio Horita
Federal University of ABC
flavio.horita@ufabc.edu.br

Abstract

As society becomes digitalized, online social networks tend to be primary places for debate but can turn into a battlefield for imposing conflicting narratives. Automating the identification of online conflicts is a challenge due to difficulties in defining antagonist communities and controversial discussions. Here, we propose a polarization approach for understanding Twitter conflicts in Brazil during the COVID-19 pandemic, where a small group of polarizers influences a larger group of polarizees according to their ideological leaning. Polarizers are automatically identified by centrality metrics in following, retweet, and reply networks and manually labeled as leftists, rightists, or undefined. We collected and analyzed the polarization of 21 potentially conflicted political events in Brazil. Our results show that polarizers adequately represent the polarization of events, the traditional media is giving way to a new breed of tweeters, and retweet and reply play different roles within a conflict that reflects their polarization level.

1. Introduction

Society is steadily becoming digital, where computers, networks, algorithms, and data currently mediate most interactions among people [1]. During the 2020 Coronavirus pandemic, digital communication has significantly increased due to the need for social distancing. Similarly, conflicts have been migrating to the digital world, especially when physical interactions are impossible or inconvenient. The process of high polarization in social networks mediating the political debate leads to social fragmentation and difficulty in consensus building. The absence of well-established information filters (gatekeepers) in social media in a digital society plays a crucial role in magnifying

conflicts. In polarized conflicts, the debate moves from the field of political projects to a contentious moral dispute, focusing on right vs. wrong, true vs. false, and good vs. evil. Such minimum consensus is necessary for policies to combat enormous societal threats such as the COVID-19 pandemic.

Brazil is experiencing significant growth in the number of Internet and social media users in the last years, including more than 75% of the population in 2019 [2]. In such a society, the discussion moves naturally to the online world, giving voice to the otherwise unheard minorities. Social media has played a vital role in the Brazilian political processes during President Rousseff's impeachment in 2015 and 2016, and the election of president Bolsonaro in 2018 [3].

Polarization in the public political debate worsens when a conflicted view of the world leads individuals to align their beliefs in extreme positions [4], mostly characterized by strong opinions frequently expressed as uncivil language, scolding, exaggerations, denials of acknowledged facts, or lies [5, 6]. Although loud and visible, a relatively small group of activists frequently leads the polarized discourse and influences other people to follow them in increased conflicting situations in social networks. In this paper, these activists are called polarizers, also referred to as antagonists in the literature [5].

Humans browsing a social network can quickly identify opposite views, understand criticisms and tensions, and characterize this situation as a conflict. On the other hand, the automation of the understanding and characterization of conflicts in social networks is still a challenge due to inherent difficulties in defining conflicting communities, controversial discussion topics, and contrasting views, let alone detecting figures of speech, irony, and sarcasm.

This paper proposes a polarization approach for understanding online conflicts in social networks and applies it in a multi-event case study during the 2020

COVID-19 pandemic in Brazil. In this context, we propose the following research question: *RQ) How can an online conflict be understood by the polarization of users in online social networks?*. For answering this question, we developed a methodology where we: collect Twitter data for potentially conflicting events (Twitter stands out as an open and reliable platform [7]); identify and label a small number of highly influential users called polarizers that exert influence over a larger group of polarizees; compute following, retweet and reply polarization; automatically extend the polarizer group, and finally; repolarize users. The process of selecting polarizers is based on building retweet and reply networks for the datasets and computing centrality metrics. We applied this approach in 21 datasets collected from Twitter based on events with greater repercussions in Brazil during the 2020 pandemic, from which we selected 147 polarizers.

Our results show that polarization metrics can be successfully used to characterize online conflicts. The group of polarizers, even small and receiving no more than 30% of retweets and replies, adequately represents the polarization of the event datasets. We confirmed the polarizers' representativeness by automatically extending the polarizer group and repolarizing all users, where we verified insignificant changes in the polarization levels. The traditional media is losing its prime time in mediating social media debates, giving way to otherwise unknown tweeters and alternative media websites. Furthermore, retweets have been mostly used for reinforcing particular ideological views within the so-called echo chambers, whereas replies are the mechanisms used to exchange ideas among different groups. The key contributions of this paper are:

1. The development of a new methodology for understanding conflicts in online social networks, based on a polarization approach that divides users into two groups: polarizers and polarizees;
2. A validation of this methodology in the context of the 2020 Coronavirus pandemic crisis in Brazil.

The remainder of this paper is organized as follows. Section 2 introduces related work and our research approach is introduced in Section 3. Section 4 presents the results obtained by employing our approach, while Section 5 discusses our findings. Section 6 draws some conclusions and provide some future lines of works.

2. Related Work

Due to the digitization of societies, digital communication has been considered for several application domains. For instance, Chen & Du [8]

used data gathered from a stock forum in China, called Guba.com.cn, to predict the trading volume and stock prices, whereas Meske et al. [9] examined the usefulness of social network within organizational settings and how it improved the digital work in terms of knowledge sharing and engagement.

Furthermore, a growing line of studies analyze debates on online social networks as they would have a massive impact on public opinions. Through the analysis of Twitter data, [10] indicated that members of higher social hierarchies play a central role in the public debate as they receive most of the cross-ideological mentions; the general public was mostly ignored. Apart from that, 40% of cross-ideological tweets were disagreement, which shows that Twitter might be contributing to increasing the contradictory debate within societies. While, Primario et al. [11] examined the polarization on the 2016 US Presidential Debates in Twitter and argued on a reasonable similarity of virtual and face-to-face debates as in both cases follow the so-called U-Shape pattern, i.e., contradictory discussions start in high-levels of debate, decrease over time, and rise when reaching the end of argumentation. They also showed that the polarization also increases from a debate to another throughout the Presidential campaign, which reinforces the similarity behavior on virtual and face-to-face debates. On the other hand, Stewart et al. [12] suggested close attention to trolls, i.e., automated Twitter accounts with substantial influence on increasing the polarization of online conflicts and thus accentuating disagreements, even though evidence of benign bots has also been found [13].

Different means of social interaction within Twitter have been used to analyze and characterize the online social conflicts. In particular, retweeting has been concisely adopted as a proxy for influence and thus can be used to indicate polarization. For instance, Liu et al. [14] demonstrated that the communication patterns (e.g., retweet network) have a significant potential to identify the users' ideological leaning within a conflict. This fact was further explored by Morales et al. [4] that measured the political polarization by propagating the opinion of elite users (i.e., users with high-levels of popularity) to listeners (i.e., influenced users). Hong et al. [15] showed that politicians with extreme political ideologies have more Twitter followers than their moderate peers. In contrast, Gruzd et al. [16] suggested that even though Twitter increases the shared views, some cross-party debate among supporters of similar leaning parties were noticed.

Content analysis on reply discourse is another approach for understanding and measuring the virtual conflicts; therefore, the communication language

analysis might be suitable to indicate a particular level of polarization. Gruzd et al. [16] found evidence that interactions of distant ideological supporters - or antagonists [5] - are negative and hostage in nature, which ultimately would lead to a more polarized discussion. Hong [15] further recognized that online messages with radical views are more likely to be noticed and disseminated. All these hostile interactions are increasing due to the so-called echo chambers, i.e., users are exposed only to communities and users that support their views [16, 17, 18]. In contrast, Garimella et al. [19] measured the conflict as a controversy indicator on Twitter by building a conversation graph from the Tweets hashtags and showed that retweet and follow graphs, as well as a random walk, edge betweenness, and sentiment analysis measures were more effective when measuring controversies.

Even though previous work has addressed relevant and vital issues, this area still has many open research questions, such as using metrics from social media analytics that may improve the understanding of online conflicts. Metrics are essential to determine how much polarization there is in the ongoing discussions and perceive the phenomenon over a specified period.

3. Research Approach

This section first introduces our research approach, including the basic concepts and later presents the steps to carry out the experimental design, starting from data collection up to the analysis of results.

3.1. Basic concepts

We adopt existing concepts used in related work - seed users and polarization - as well as consolidate new ones, such as polarizer, polarizee, and repolarization. **Seed users** belong to the original user set to measure polarization, selected by building retweet and reply networks and computing centrality metrics. **Polarizer** specialize and extend seed users, influencing other users in the choice of retweeting and replying. Throughout this paper, the term polarizer is used to indicate what otherwise may be known as seed user. **Polarizee** is a common user influenced by polarizers in their retweet and reply behavior. **Polarization** indicates the level of how polarized a conflicted event is in terms of frequency of retweets and replies to polarizers whose political leaning is known (left/wright/undefined). Also, it means the level of how polarized are individual polarizees. **Repolarization** is a process of extending the initial polarizers by automatically labeling user leaning based on their polarization level.

Figure 1 depicts the relationship among seed users,

polarizers, and polarizees, where seed users are initially selected as candidates for polarizers. A process of repolarization extends the initial group of polarizers into a larger one. Polarizers influence other users, i.e. the polarizee, when they post particularly polarized tweets. In response to polarized tweets, polarizees follow, retweet, and reply the polarizes, which indicates their level of polarization.

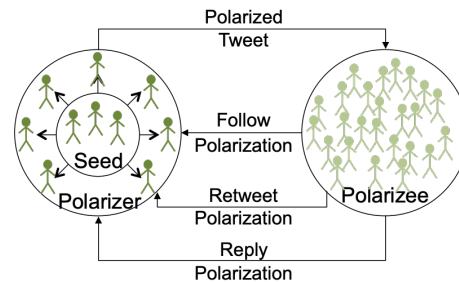


Figure 1: Basic concepts of Online Conflicts

3.2. Research Methodology

We developed a new methodology to characterize online conflicts from the user polarization, consisting of six major steps, as shown graphically in Figure 2. While we adopted existing concepts from the literature, we developed the whole sequence of steps, and also, the repolarization approach from step 6 is our original contribution. Here, we describe our general methodology and its usage for analyzing typical conflicts in Brazilian politics.

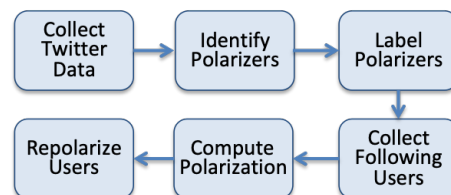


Figure 2: Research Methodology

Step 1: Collect Twitter data of potential conflicted events into different datasets. The Twitter Streaming API is the default data collection approach that returns meaningful datasets [7]. We developed a crawler that collects data matching specific hashtags chosen by 21 events with a higher impact on the news during the first semester of 2020 on the ongoing context of the COVID-19 pandemic.

Step 2: Identify first polarizers based on centrality metrics in interaction networks. First, for each dataset, build two directed interaction graphs (networks), where users are the vertices (nodes), and retweets and replies are the edges (links). Specifically, if user A retweets or replies to user B, a directed edge is created from user A

to user B. Second, compute degree, indegree, outdegree, betweenness, and PageRank centrality metrics to identify the most influential users. Third, generate lists with the top 100 users for each dataset, interaction (retweet/reply), and centrality metric. Fourth, count the number of times each user appeared in the top 100 lists for retweet and reply networks separately and rank the 100 top users for the retweet and reply interactions. Finally, merge the retweet and reply top 100 lists to form a unique list of polarizers. We applied this step on 21 datasets, resulting in 42 networks (21 retweet and 21 reply ones) and 210 centrality lists. We tried different alternatives for computing the top polarizers, but this one showed itself to be more robust in terms of the group of polarizers it selects. Also, by using different centrality metrics, we captured users who play different roles in the network. For example, degree, indegree, outdegree, and PageRank centrality metrics capture the user's importance regarding the number of retweets and replies he/she sent and received. On the other hand, betweenness centrality capture users who intermediate the interaction among other users. Electing influential users to define the polarization is a common practice in this area, where they receive different names, such as antagonists [5], elite users [4], or seed users [20].

Step 3: Manually label polarizers into left, right, and undefined political leaning. Each unique polarizer selected in Step 2 must be individually labeled according to their activity on Twitter. For example, we labeled politicians according to their parties and personal ideology. Also, traditional media and traditional journalists were classified as undefined since they usually report a fact with less ideology interference. On the other hand, some otherwise relatively unknown users outside Twitter required an in-depth analysis of their original tweets, retweets, and replies according to their political leaning. These users are also classified into nine categories, which did not directly interfere in the polarization analysis but helped us understand our polarizers in terms of their ideology and activity.

Step 4: Collect following relationship: polarizee \rightarrow polarizer. After selecting and labeling the polarizers, use the Twitter Streaming API to collect data and establish the following \rightarrow follower network. Rather than collecting all polarizers' followers, we collected their following users who retweeted or replied ten times or more. The sum of polarizers' followers is almost 120 million users, whereas we collected around 40 million users in around 200 hours, respecting the API limits.

Step 5: Compute following, retweet, and reply polarization, using the Garimella and Weber [20] approach based on the Bayesian probability, which provides a way to update our beliefs in the evidence

of new data. The decision of following, retweeting, and replying may be considered a Bernoulli trial whose repetition leads to the Binomial distribution. The conjugate prior to the Binomial is the Beta distribution with two parameters α and β with the mean by $\alpha/(\alpha+\beta)$. We used the Beta distribution with initial uniform prior, i.e., $\alpha = \beta = 1$, where α represents left-leaning and β represents right-leaning. Therefore, for computing the three polarization metrics (following, retweet, and reply), increase the count of α or β every time a polarizee followed, retweeted, and replied a polarizer for each dataset. This process renders different means for different datasets since different sets of polarizers and polarizees participate in the events. Polarizers labeled as undefined do not change the polarization metric. The l metric reflects the political "leaning." When l is 0.5, polarizees have a null polarization, as they have the same number of counts to α and β . When l tends to 1, the user is highly polarized to the left and highly polarized to the right when l tends to zero. To unify left and right polarization under a single metric, compute polarization $p = 2 \cdot |0.5 - l|$, where $p = 0$ means $l = 0.5$ that means no polarization whatsoever. On the other hand, as p tends to 1 (it will never be 1 actually), it means that the polarizee is more polarized either to the left or right. We then compute the average of all polarizees p values for each dataset, which we call retweet, reply, and following polarization.

Step 6: Repolarize users. This step extends the original set of polarizers by a repolarization approach that automatically labels polarizees with $p \geq \delta$ according to their political leaning. After the first automated repolarization process, the polarization is recomputed for adding more retweets and replies. This process (automated 1) can be extended to n iterations (automated n), where the repolarization only stops when there is no increment in the number of polarizers between current and previous iterations. We tested different repolarization thresholds, and $\delta = 0.7$ showed to be adequate to our needs. Thus, polarizees with $l \leq 0.15$ are labeled right, and $p \geq 0.85$ are labeled left.

4. Results

4.1. Datasets

As explained in Step 1 of our methodology, we collected data related to potential conflicting events in the Brazilian political scenario during the 2020 pandemic, from which 21 datasets were considered adequate in terms of polarization to be individually

analyzed. Table 1 presents the selected datasets¹ along with some statistics of the number of users participating in the debate, number of retweets, and number of replies.

Table 1: Twitter Datasets

Dataset	Period	Conflict	Users	Retweets	Replies
Coronavirus	15/Mar/20	Pandemic	143,080	166,322	5,378
Speech 1	24/Mar/20	Pandemic	195,494	288,546	16,376
Speech 2	01/Apr/20	Pandemic	66,142	143,279	14,063
Speech 3	08/Apr/20	Pandemic	87,456	201,131	18,858
Fasting	03/Apr/20	Pandemic	124,197	271,001	45,257
Chloroquine	08/Apr/20	Pandemic	111,165	372,890	30,385
Supreme	15/Apr/20	Pandemic	40,291	83,124	12,058
Ministry	20/May/20	Pandemic	121,801	221,166	24,173
Mandetta	16/Apr/20	Political Controversy	143,067	222,769	21,416
Moro	24/Apr/20	Political Controversy	546,185	1,487,111	232,954
Duarte	07/May/20	Political Controversy	98,605	176,113	9,174
Video	12/May/20	Political Controversy	56,904	140,374	21,334
Heleno	22/May/20	Political Controversy	25,256	33,713	2,836
Universities	10/Jun/20	Political Controversy	31,553	50,640	2,361
Weintraub	18/Jun/20	Political Controversy	158,380	324,542	17,597
Impeachment	20/Apr/20	Bolsonarism	73,942	298,547	77,466
Trench	04/May/20	Bolsonarism	40,275	132,349	36,049
Police	27/May/20	Bolsonarism	173,522	744,378	55,765
Fascist	03/Jun/20	Bolsonarism	56,673	63,428	10,951
Business	07/May/20	Other	12,789	14,089	807
Queiroz	18/Jun/20	Other	232,042	496,340	32,714

Each dataset, represented by a nickname and the data collection date, was classified into four categories of conflicts. *Pandemic* refers to datasets with direct involvement in the spread of the coronavirus, speeches of the Brazilian president Jair Bolsonaro, repercussions about the use of chloroquine for treating COVID-19, as well as struggles involving the Brazilian Ministry of Health. *Political controversy* may also relate to the pandemic, but include datasets where the debate was permeated by an intense political background involving the Brazilian government. *Bolsonarism* includes datasets where the discussion was more related to the nature of the Bolsonaro government. *Other* includes two datasets related to a demonstration of businesspeople demanding the opening of the country and a crisis involving the Bolsonaro family.

Our methodology rendered 156 unique users after merging the retweet and reply top 100 lists. Among the first polarizers, nine accounts were not found during the manual labeling process in Step 3 because they were either banned or removed so that, in practice, we

¹A more detailed description of each dataset is provided in the supplementary materials

worked with 147 users. Since 21 polarizers were labeled as undefined (i.e., neutral leaning), the polarization approach considered in practice 126 polarizers, 99 rightists, and 27 leftists, i.e., a predominance of the right-wing in the current political debate in Twitter.

Table 2 depicts some numbers resulting from the labeling and classification processes. Some categories are self-explanatory, such as journalists, politicians, traditional media, and artist/athlete. Institution includes the Brazilian Supreme Court and Senate official Twitter accounts. Alternative media include highly biased websites that support one side of the political spectrum. Traditional media accounts are classified as undefined since they do not explicitly support either side.

Table 2: Polarizer distribution: categories and leaning

Category	Left	Right	Undefined	Total
Journalist	5	10	6	21
Politician	10	24	0	34
Traditional Media	0	1	12	13
Alternative Media	2	4	0	6
Artist/Athlete	1	2	1	4
Tweeter	6	49	0	55
Political Group	1	5	0	6
Institution	0	0	2	2
Other	2	4	0	6
Total	27	99	21	147

Tweeter is the most common category and includes users with no prominence outside Twitter, but that has a high impact on this social network. One noticeable characteristic is that most top-ranked tweeters belong to the right. The political group category includes users representing a group of people with great ideological content but has no formal relation to any political party. In contrast, Garimella and Weber [20] selected a shorter list of 39 seed users among politicians and traditional media from both left and right wings manually. In the USA, traditional media is classified into left and right, unlike our approach during the 2020 COVID-19 pandemic in Brazil.

Table 3 shows a sample with the top 10 retweet polarizers², which is dominated by rightist tweeters. It is fundamental to mention that the 5th user of RT Rank is not present as this account was not available anymore. Most users also appear in the reply rank.

On the other hand, the top 10 reply polarizer list (not shown here) has more politicians and traditional media. The right-wing is still the majority, but there are two undefined and one left-wing. We found that reply is a mechanism where discussions among different user profiles happen, lead by politicians and traditional media. Most top 10 users in the reply rank also appear in the retweet rank.

²The complete list of 147 polarizers can be found in the supplementary material.

Table 3: Sample of top 10 retweet polarizers

Polarizer	Leaning	Profile	Following	Followers	RT Rnk	RP Rnk
taoquei1	Right	Tweeter	1,168	282,595	1	73
BrazilFight	Right	Political Group	15,806	154,417	2	5
Jouberth19	Right	Tweeter	847	107,607	3	13
leandroruschel	Right	Journalist	409	410,347	4	64
LeiaRachor	Right	Tweeter	29,247	49,199	6	
Dom_Andolini	Right	Tweeter	296	1,957	7	82
carlosczarzjr	Right	Tweeter	30,642	105,448	8	
allantercalivre	Right	Journalist	687	360,210	9	87
FabioTalhari	Right	Tweeter	39,542	43,909	10	46
BlogdoNoblat	Undefined	Journalist	338	1,168,985	11	11

4.2. Retweeting and Replying

For the 21 selected datasets, we compute the retweet and reply behavior of the polarizees, i.e., when polarizees retweet or reply polarizers. To understand the representativeness of the polarizer group to measure polarization, we counted the percentage of retweets and replies targeted the polarizers in each dataset, depicted by Figure 3. On average, polarizees retweeted and replied to polarizers around 30% of the time. On average, 70% of the retweets and replies targeted non-polarizers, and this discovery leads us to expand the polarizer group by automatically labeling polarizees as left and right into the polarizer set. In some datasets, polarizers received more than 40% of the retweets (i.e., Chloroquine, Supreme, Heleno, and Police) and more than 40% of the replies (i.e., Speech 2, Speech 3, Video, Impeachment, and Trench). We found out that the higher the proportion of retweets and replies, the higher the polarization.

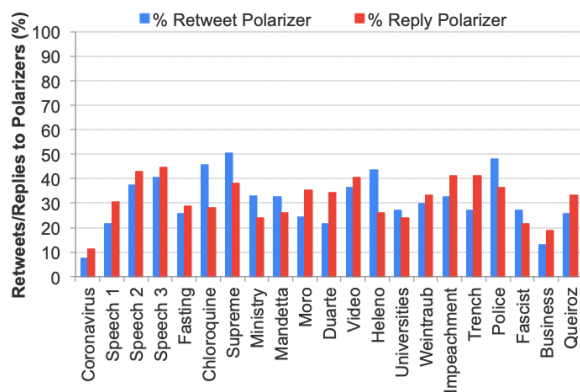


Figure 3: Proportion of retweets and replies to polarizers

Figure 4 shows the proportion of retweets and replies from the polarizees to the polarizers according to their political leaning: left, right, and undefined. The first finding is that right-wing polarizers capture most retweet and reply engagement. Rightists received almost 100% in some datasets, namely Impeachment and Trench, as demonstrated in Figure 4. Four datasets

presented a majority of retweets to left polarizers, i.e., Speech 1, Heleno, Business, and Queiroz, which leftists mostly criticized the Brazilian government. Undefined polarizers did not generate significant engagement for polarizees, except for the Coronavirus dataset with 35%, at the beginning of the pandemic.

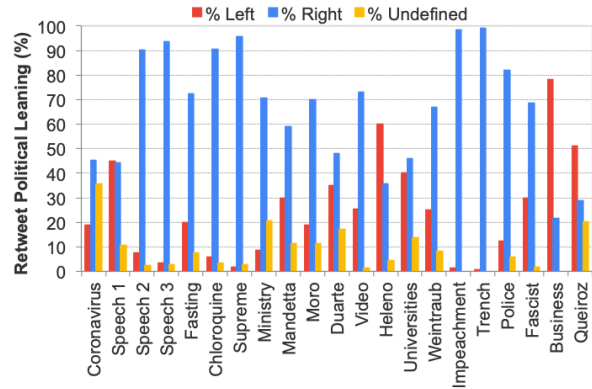


Figure 4: Proportion of retweets to polarizers with political leanings: left, right, and undefined

Right-wing polarizers engaged in a higher proportion of polarizees also for replies, shown in Figure 5, but with a less strong predominance for most datasets. Moreover, there was an inversion of proportionality between undefined and left in the debated mediated by replies, where undefined polarizers received the most replies for the Duarte and Business datasets. The Duarte event derives from the involvement of the ex-Secretary of Culture Regina Duarte in an adverse situation during an interview with CNN Brazil. In contrast, the business event attracted a lower presence of polarizers and a higher presence of journalists and traditional media. The higher proportion of undefined polarizers in replies is a direct consequence of the fact that the traditional media is more present in discussions of ideas.

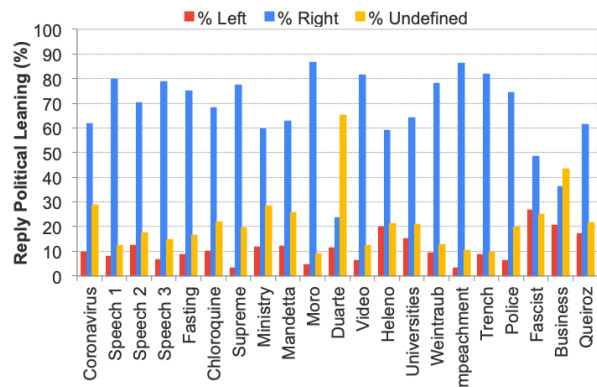


Figure 5: Proportion of replies to polarizers with political leanings: left, right, and undefined

Figure 6 further highlights the predominance of right polarizers in the debate, showing the Empirical Cumulative Density Function (ECDF) for retweets in the Police Dataset. According to the definition of leaning in Step 5 of section 3, the ECDF shows that around 45% of the polarizees of this dataset have leaning 0.5, which means that they did not retweet anyone, retweeted only undefined polarizers, or retweeted both sides the same number of times. Also, 40% of the polarizees have leaning below 0.5, indicating a polarization to the right. Only 15% have polarization to the left with leaning above 0.5.

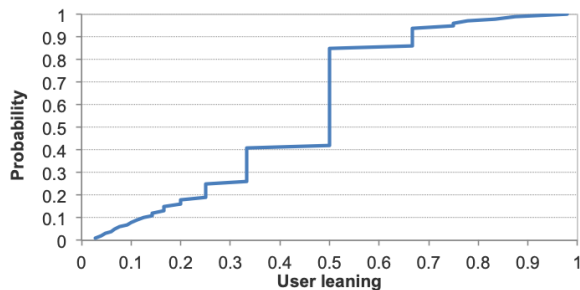


Figure 6: ECDF of user retweet leaning - Police Dataset

4.3. Polarization Results

The average retweet polarization of all datasets is 0.62, considering polarizees who retweeted ten or more times (shown in Figure 7), and 0.43 including all polarizees (not shown). A first finding is the relationship of retweeting behavior to political leanings (Figure 4) and the retweeting polarization (Figure 7). The error bars show the standard deviation for the three polarization charts.

A second finding is that the higher one of the sides prevails among the others, the higher the polarization. There is no direct relationship between the proportion of retweets/replies to political leaning and the polarization since the polarization of a dataset is computed as the average polarization of its polarizees. The behavior of polarizees may change considerably among datasets. In other words, the retweeting polarization depends on the mix of highly polarized and less polarized users. For example, the highest retweeting polarization (0.78) was achieved by the Police dataset, wherein around 80% of the retweets targeted rightist polarizers. On the other hand, the Trench dataset has 99% of retweets targeting rightist polarizers but with a retweet polarization of 0.64. Undefined polarizers also influence the level of polarization, even though they do not directly count for the polarization. They consume retweets that could go to either side otherwise, thus smoothing the polarization.

The reply polarization, shown by Figure 8, is less

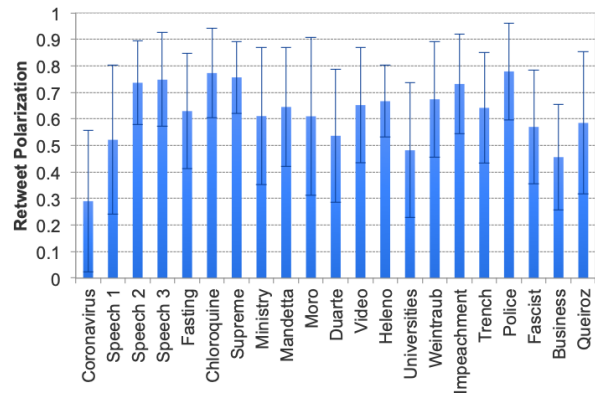


Figure 7: Retweet Polarization

intense than the retweet polarization. The average for all datasets is 0.49 for polarizees who replied ten or more times and 0.36 for all polarizees. It also shares a relationship with the proportion of replies going to left, right, and undefined, shown by Figure 5. Again, the reply polarization is lower due to the higher level of discussion involving different sides, which is less common in retweets. However, the standard deviation of the reply polarization is higher compared to retweet.

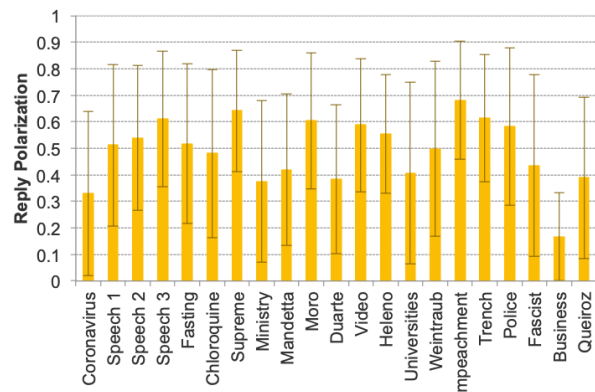


Figure 8: Reply Polarization

Figure 9 shows the following polarization with an average of 0.7 for all datasets, higher than for retweet and reply. The standard deviation again reached a higher level. Users are generally more polarized regarding following other users, somewhat less polarized for retweeting, and even less polarized for replying.

When analyzing a dataset in further detail, Figure 10 shows the number of retweets received by all users (polarizers and polarizees) filtering by their polarization for the Police dataset, which duly represents the other 20 datasets. The chart demonstrates that users within the polarization range between 0.5 and 0.8 received around 50% of all retweets and that higher polarized users are rarely retweeted. In contrast, results also indicate a different user retweeting behavior, where most retweets,

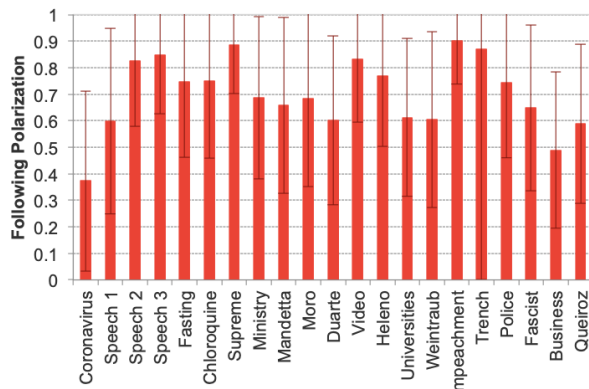


Figure 9: Following Polarization

about 80%, come from users with polarization between 0.6 and 1, whereas almost 40% of all retweets come from users with polarization 0.9 and higher.

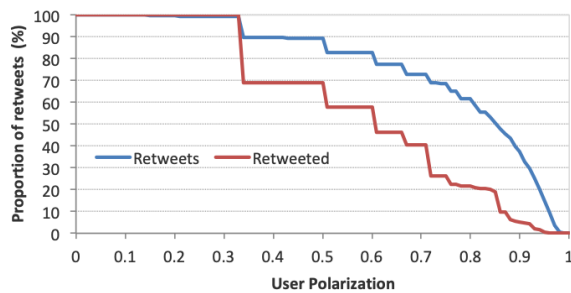


Figure 10: Proportion of retweets received by users according to the polarization metric - Police Dataset

Figure 11 depicts the effect of filtering a minimum number of retweets and replies to compute the polarization, from 0 to 50 for the Police dataset. We filtered the dataset in four different ways, i.e., the number of retweets sent (Retweets), number of retweets received (retweeted), number of replies sent (Replies), and number of replies received (replied). The figure shows that the higher the number of times users retweet and reply, the higher their polarization (with retweet higher than reply according to previous results). The average polarization of the dataset increases quickly from 0.52 retweet and 0.39 reply to 0.80 and 0.69, respectively, from zero to ten, and increases more smoothly afterward. On the other hand, the higher the number of times a user is retweeted and replied, the lower their polarization. After filtering users who retweeted and replied zero and one times, where the polarization increases, it slowly decreases. The explanation for this behavior is that polarizees very prolific in retweeting/replying are highly polarized. Contrarily, users who are frequently retweeted and replied do not often retweet and reply to other users. Consequently, as the polarization metric is computed

over the number of retweets and replies made by a user, the users who cause a higher impact on others (who are most frequently the polarizers) have a lower polarization. This behavior might be counter-intuitive, but it comes from the very nature of the metric.

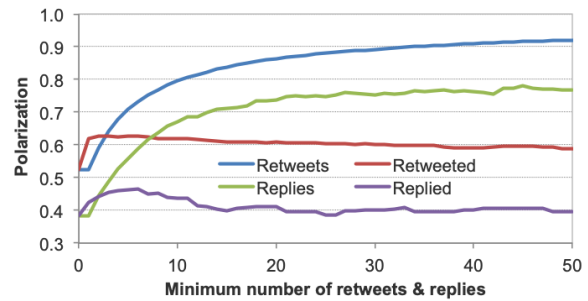


Figure 11: Effect of a minimum number of retweets and replies to compute polarization - Police Dataset

Finally, we tried to understand our polarizers' representativeness to compute the polarization metric, even though they received, on average, less than 50% of the retweets and replies. By computing the repolarization and including many polarizers, we found out that polarizers correctly represent the datasets. Table 4 demonstrates the effect of repolarization by two automated approaches, automated 1 (i.e., one repolarization iteration) and automated n (i.e., n repolarization iterations). From the first 126 polarizers with a definite leaning, most datasets considerably increased this metric, e.g., Moro and Police. However, this higher number of polarizers did not change the polarization metric significantly. Most datasets had a small increment of less than 1%, with the peak for the Impeachment dataset with a 3.09% increase from the manual (original) labeling to the automated n approach. This result indicates that our approach for selecting polarizers based on centrality metrics is robust and representative of the datasets since adding new polarizers did not change significantly.

5. Discussion

The Brazilian political scenario has been increasing in polarized opinions in the last few years but showed a considerable rise during the 2020 pandemic. Our research methodology and results revealed insights from which significant lessons could be learned. Our approach to analyzing conflicts makes it possible to follow and understand Brazil and other countries' political debate. Based on the literature and qualitative analysis of our polarizers, we have the underlying assumption that the higher the polarization of an event, the higher the probability of this event representing a conflict between sides competing to impose narratives.

Table 4: Retweet repolarization Effect

Dataset	Manual Auto-1	Manual Auto-n	Auto-1 Auto-n	Polarizer Auto 1	Polarizer Auto n
Coronavirus	0.01%	0.01%	0.00%	159	159
Speech 1	0.31%	0.40%	0.08%	1,435	1,542
Speech 2	0.66%	0.74%	0.07%	3,250	3,468
Speech 3	0.74%	0.83%	0.09%	4,975	5,350
Fasting	1.86%	2.60%	0.73%	3,424	4,323
Chloroquine	0.65%	0.76%	0.11%	11,074	11,642
Supreme	0.93%	0.97%	0.05%	2,513	2,739
Ministry	0.54%	0.56%	0.02%	2,952	3,124
Mandetta	0.36%	0.37%	0.01%	2,288	2,393
Moro	2.03%	2.47%	0.43%	17,390	20,596
Duarte	0.27%	0.28%	0.01%	1,153	1,206
Video	1.18%	1.30%	0.11%	2,513	2,895
Heleno	0.00%	0.00%	0.00%	262	263
Universities	0.21%	1.64%	1.43%	240	408
Weintraub	0.42%	0.47%	0.05%	4,739	4,964
Impeachment	2.36%	3.09%	0.72%	6,704	8,221
Trench	1.86%	2.23%	0.37%	1,935	2,345
Police	1.81%	1.89%	0.08%	21,948	24,426
Fascist	0.00%	0.00%	0.00%	282	283
Business	0.00%	0.00%	0.00%	128	128
Queiroz	0.29%	0.46%	0.16%	4,670	5,011

Polarizers: The set of polarizers revealed the polarization of our datasets. It generated stable polarization levels even though polarizers received, on average, 30% of retweets and replies. Automated repolarization using the polarization level to label new polarizers did not present significant differences from the original polarizer set. A low number of highly influential users, the polarizers, can establish the dynamics and ideological leaning of debates [10, 4].

Polarizer selection and labeling: Our methodology requires a customized selection of polarizers, which must be recalculated and tuned for different situations. For example, as new datasets are added, new users may achieve central positions, and old ones may be less active, or even accounts may be banned or deleted. Also, different subjects, not only politics, may involve different sets of polarizers. The selection of polarizers based on betweenness centrality is a CPU-intensive process for large datasets. The good news is that the attraction of retweets and replies mostly characterizes polarizers, so that degree is an adequate predictor of other centrality metrics. On the other hand, the polarizer labeling is currently a manual process that must be automated in the future, using different methods such as sentiment analysis and machine learning.

Retweet, reply, and following polarization: The number of times polarizees retweet and reply polarizers

revealed to be a correct predictor of their polarization, rather than the frequency an original tweet is retweeted and replied. This phenomenon happens due to the nature of the polarization metric and the behavior of polarizers and polarizees. Firstly, polarization is based on the number of retweets and replies, which must be repeated dozen times for the metric to yield high value. On the other hand, polarizers are retweeted and replied to many times, but they are not particularly prolific in retweeting and replying. Users are generally more polarized regarding following other users, less polarized for retweeting, and even less for replying. In general, the interaction patterns among users (follow, retweet, reply) have a significant potential to identify their ideological leaning within a conflict. Retweets are mostly used for reinforcing a particular ideological view within the so-called echo chambers. Replies are used to exchange ideas within the same homogeneous group, but also including users of opposing groups [10, 4, 15].

Role of users in online conflicts: In this new digital society, the traditional media loses influence, giving way to new actors, such as alternative media websites or otherwise anonymous tweeters. In most event datasets, most retweets and replies targeted rightist polarizers' tweets. Only four out of 21 datasets presented a prevalence of retweets to leftist polarizers and none to undefined ones. As for replies, two datasets presented the highest number to the undefined polarizers (mostly traditional media) and none to the left. Traditional media has a higher attraction on replies, used to debate ideas among a more heterogeneous user base. In the current highly polarized Brazilian political scenario, most polarizers belong to the right, and most retweets and replies go to rightist polarizers. The traditional media was mostly labeled as undefined because the current Brazilian political situation is so polarized that small biases may be classified as neutral. The traditional media is frequently blamed as the source of all evil and accused of supporting the other side by both sides.

Action of bots: We found out that almost 40% of all retweets come from users with polarization 0.9 and higher who were not retweeted. This behavior may indicate bots' action as it is in line with previous works on the topic [12, 13]; however, such analysis can be recognized beyond the scope of this work and should be considered as future work.

Applications: While the primary purpose of this paper is to generate knowledge by shedding light on polarized behaviors on Twitter, this understanding may be used for practical applications as well. The two most clear applications are in enhancing public policy and preventing uncivil behavior in social media. Authorities may use this methodology to understand polarized

user behavior and improve the democratic processes that shape our society. Social network providers may also apply our methodology to improve surveillance mechanisms on abusive users and build new online services where respect and healthy discussions prevail.

Acknowledgment

This research is funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), grant number 2018/23022-3.

6. Conclusion

We presented a polarization approach built on descriptive measures of online conflicts on social networks, i.e., following, retweet, and reply rates. A multi-context case study during the 2020 Coronavirus pandemic in Brazil showed that polarization levels are higher in the following frequencies than retweet and reply frequencies. Also, retweets support ideological view within echo chambers, whereas replies are used to exchange ideas. We demonstrated that our methodology based on network centrality is feasible for selecting polarizers, which are then utilized to compute the polarization. From a practical perspective, we revealed high polarization in Twitter-based political conflicts in Brazil, which may have contributed to the ineffective policies for combating the Coronavirus pandemic [21].

Future work will automatically identify bots and remove them from the polarization metrics, as they can increase the level of polarization [12], and influence the disagreements on online conflicts [13]. Also, the algorithmic labeling of polarizers is a crucial topic for automating the identification of conflicts. Finally, a discourse and network analysis should be carried out to investigate if a high polarization reflects a high probability of a conflict event between at least two sides competing to impose their narratives.

References

- [1] T. Dufva and M. Dufva, "Grasping the future of the digital society," *Futures*, vol. 107, pp. 17–28, 2019.
- [2] CETIC.BR, "ICT households in Brazil," *Regional Center for Studies on the Development of the Information Society (Cetic.BR)*, 2019.
- [3] F. O. de França, D. H. Goya, and C. L. de Camargo Pentead, "User profiling of the twitter social network during the impeachment of brazilian president," *Social Network Analysis and Mining*, vol. 8, no. 1, p. 5, 2018.
- [4] 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*, vol. 25, no. 3, p. 033114, 2015.
- [5] S. Post, "Polarizing communication as media effects on antagonists. understanding communication in conflicts in digital media societies," *Communication Theory*, vol. 29, no. 2, pp. 213–235, 2019.
- [6] S. Sobieraj and J. M. Berry, "From incivility to outrage: Political discourse in blogs, talk radio, and cable news," *Political Communication*, vol. 28, no. 1, pp. 19–41, 2011.
- [7] V. Ayora, F. Horita, and C. Kamienski, "Social networks as real-time data distribution platforms for smart cities," in *10th Latin America Networking Conference*, 2018.
- [8] Z. Chen and X. Du, "Study of stock prediction based on social network," in *2013 International Conference on Social Computing*, pp. 913–916, 2013.
- [9] C. Meske, K. Wilms, and S. Stieglitz, "Enterprise social networks as digital infrastructures - understanding the utilitarian value of social media at the workplace," *Information Systems Management*, vol. 36, no. 4, pp. 350–367, 2019.
- [10] Z. Liu and I. Weber, *Is Twitter a Public Sphere for Online Conflicts? A Cross-Ideological and Cross-Hierarchical Look*, pp. 336–347. 2014.
- [11] S. Primario, D. Borrelli, L. Iandoli, G. Zollo, and C. Lipizzi, "Measuring polarization in twitter enabled in online political conversation: The case of 2016 us presidential election," in *2017 IEEE International Conference on Information Reuse and Integration (IRI)*, pp. 607–613, 2017.
- [12] L. G. Stewart, A. Arif, and K. Starbird, "Examining trolls and polarization with a retweet network," in *WSDM workshop on Misinformation and Misbehavior Mining on the Web (MIS2)*, pp. 1–6, 2018.
- [13] F. Brachten, M. Mirbabaie, S. Stieglitz, O. Berger, S. Bludau, and K. Schrickel, "Threat or opportunity? - examining social bots in social media crisis communication," in *Australasian Conference on Information Systems (ACIS)*, pp. 1–10, 2018.
- [14] Z. Liu and I. Weber, *Is Twitter a Public Sphere for Online Conflicts? A Cross-Ideological and Cross-Hierarchical Look*, pp. 336–347. 2014.
- [15] S. Hong and S. H. Kim, "Political polarization on twitter: Implications for the use of social media in digital governments," *Government Information Quarterly*, vol. 33, no. 4, pp. 777 – 782, 2016.
- [16] A. Gruzd and J. Roy, "Investigating political polarization on twitter: A canadian perspective," *Policy & Internet*, vol. 6, no. 1, pp. 28–45, 2014.
- [17] H. Takikawa and K. Nagayoshi, "Political polarization in social media: Analysis of the "twitter political field" in japan," in *2017 IEEE International Conference on Big Data (Big Data)*, pp. 3143–3150, 2017.
- [18] N. Gillani, A. Yuan, M. Saveski, S. Vosoughi, and D. Roy, "Me, my echo chamber, and i: Introspection on social media polarization," in *2018 World Wide Web Conference, WWW '18*, p. 823–831, 2018.
- [19] K. Garimella, G. D. F. Morales, A. Gionis, and M. Mathioudakis, "Quantifying controversy on social media," *Trans. Soc. Comput.*, vol. 1, no. 1, 2018.
- [20] V. R. K. Garimella and I. Weber, "A long-term analysis of polarization on twitter," in *Eleventh International AAAI Conference on Web and Social Media*, 2017.
- [21] C. R. Sunstein, *# Republic: Divided democracy in the age of social media*. Princeton University Press, 2018.