How Presidential Candidates and Parties Tweeted During the 2018 Brazilian Election

Gabriel Lima

School of Computing, KAIST, Daejeon, South Korea

Abstract. The 2018 Brazilian Presidential Election was marked by an increase in the importance of online political campaigns, allowing Jair Bolsonaro, which focused his campaign on social media, to win the election without almost any time on television. This short paper analyzes how tweets from the leading presidential candidates and their parties were distributed during the time when political television propaganda was live and how similar their tweets are to each other's, finding personal tweets as a possible reason for the success of Bolsonaro. It also investigates how candidates were cited by their opponents and what they talked the most about during the campaign, showing Bolsonaro and PT (Partido dos Trabalhadores / Workers' Party) as the main characters of the 2018 Brazilian Presidential Election.

Keywords: Politics · Brazil · Presidential · Election · Twitter · 2018

Introduction

In the year of 2018, Brazil had its ninth direct election since the end of its military dictatorship in 1985. From 2002 to 2016, Brazil was governed by a left wing party, PT (Partido dos Trabalhadores / Workers' Party), and was able to consistently grow economically and socially during Luis Inácio Lula da Silva (Lula)'s terms (2002-2010). His successor, Dilma Rousseff, was not able to continue his success, troubled by unemployment and inflation, and suffered a political impeachment in 2016. Even though PT was able to advance Brazil in many areas, its government was stained by corruption scandals. The 2018 election was expected to be a turning point in Brazilian politics, shifting the power from the left to the right wing, as a remedy to the country's problems. For a more complete analysis of the political situation in Brazil, Chapters 1 and 2 of the first author's short paper "The Empowerment of Liberalism in Brazil"[1] explain the political context more in depth.

Especially after Dilma's impeachment, Jair Bolsonaro, a right wing candidate, was able to rise as the main opposition to the government. He led the polls leading the election alongside Lula, who PT tried to set as its candidate. However, halfway through the campaign, Lula was prohibited from running for president, since he is in jail under corruption charges. PT, then, chose Fernando Haddad as its candidate, using Lula's image to boost Haddad's voting intentions. This election marked the end of consecutive left wing presidencies in Brazil after Bolsonaro defeated Haddad, with over 55% of the

---

[1] https://thegcamilo.github.io/assets/liberalismbr.pdf

votes, in the second round of the 2018 Brazilian Election and took office in the start of 2019.

Television propaganda has always played a big role in politics. Television is often the easiest way for a politician to show his proposals and ideas to the general public. In Brazil, during presidential campaigns, candidates have time on TV for propaganda proportional to the number of congressman in his/her party coalition. In the 2018 election, Jair Bolsonaro changed how politics is done in Brazil by focusing his campaign on social media, due to the fact that he only had 8 seconds on TV out of the 12 minutes dedicated to political campaigns. Geraldo Alckmin, an opposing candidate, had almost 6 minutes of propaganda and was only able to achieve fourth place.

This short paper analyzes a Twitter dataset composed by tweets of the leading candidates and parties in the 2018 presidential election. We first examine the distribution of these tweets over time, findings peaks during televised debates. We also inspect the similarity between the candidates and parties' tweets, analyzing how Bolsonaro's social media usage was different from other candidates. In order to understand how much a candidate was targeted by his opponents during the election, we examined how much each candidate and party was talked about by his/its opponents. Finally, we inspect on what each candidate focused his social media campaign by analyzing the most frequent words used on their tweets.

Related Work

Twitter is a social media where users can share their thoughts through tweets, messages with a maximum of 280 characters. Users can interact with each other by retweeting (sharing), replying and liking other users' tweets. Twitter has roughly 330 million active monthly users and is often used as a platform where researchers can analyze user behaviors and interactions. Research about Twitter include predicting stock market [1], measuring user influence [2] and understanding gender and race inequalities [5].

Politics, especially during elections, is a frequent topic on Twitter, since candidates utilize their accounts to interact with their electors and to also promote their campaigns and proposals. The electorate also has the opportunity to voice their opinion about all aspects of the elections. Previous works have analyzed how twitter messages correlate with important events during political campaigns [3], how tweets' sentiment analysis can predict a political party popularity [6] and candidates' tweeting strategies during elections [4]. Other authors studied how Twitter can be used to predict election outcomes [7] and how political discourse on Twitter is correlated to the real world public opinion [8]. This short paper tries to expand the work on political tweets by analyzing the similarity between opposing political candidates' tweets and these candidates' campaign strategies in the context of the 2018 Brazilian Presidential Election.

Data

The Twitter dataset used in this analysis is composed by 9332 tweets acquired through the Twitter REST API (using the Tweepy library) and GET requests. The tweets collected range from August 31st 2018, the day when political propaganda was set live

on TV, until October 28th 2018, the final day of the election. Out of 13 candidates that ran for president in the election, we chose to collect tweets from the top 4 candidates (Jair Bolsonaro, Fernando Haddad, Geraldo Alckmin and Ciro Gomes) and Lula, the Brazilian ex-president that is now in jail. Lula was set to run for president throughout almost half of the campaign; however, he was prohibited from doing so due to his incarceration, setting Fernando Haddad as his successor.

We also collected the tweets from the candidates' parties: PSL (Partido Social Liberal / Social Liberal Party), PT (Partido dos Trabalhadores / Workers' Party), PSDB (Partido da Social Democracia Brasileira / Brazilian Social Democracy Party) and PDT (Partido Democrático Trabalhista / Democratic Workers' Party). Due to API restrictions, the full text of tweets were also acquired through web scraping using the tweet id obtained through the Twitter API. The tweet distribution is shown in Table 1.

Candidate	N	%	Party	N	%
Jair Bolsonaro	698	7.5	PSL	634	6.8
Fernando Haddad	1152	12.3	PT	2153	23.1
Geraldo Alckmin	606	6.5	PSDB	966	10.3
Ciro Gomes	960	10.3	PDT	301	3.2
Lula	1862	20	-	-	-
					Total: 9332

Table 1: Number of tweets by candidate and his respective party in the dataset. The candidate and his respective party are aligned. Lula is also from PT.

The text preprocessing was done as follows: all tweets were 1) transformed to lower case tweets, 2) all links, mentions, hashtags and emojis were removed and 3) every symbol other than letters and numbers were removed. Finally, after the 4) tokenization of the tweet, 5) stopwords and the RT (retweet) symbol were also removed. It is important to note that for the analysis of citations between candidates and parties, this preprocessing was not performed since the search was done over the whole tweet.

Time Distribution over Time

Each tweet obtained through the Twitter API has its $created\_at$ field, which indicates when the tweet was posted. Figure 1 shows the distribution of tweets over time, starting from August 31st 2018 until October 28th 2018.

Since the candidates' accounts are not mainly controlled by the candidates themselves but by their communication team, most of the tweeting peaks are either during or right after debates on TV. These tweets intend to show what the candidates talked about and how they did during the debate and what their proposals are. A tweeting peak was also expected after Bolsonaro's stabbing since it marked an important breaking point during the campaign; however, this peak did not happen, as shown in Figure 1.

After the 1st round took place, to the right of the second dotted line, Ciro and Alckmin, third and fourth place respectively, stopped tweeting as frequently as during
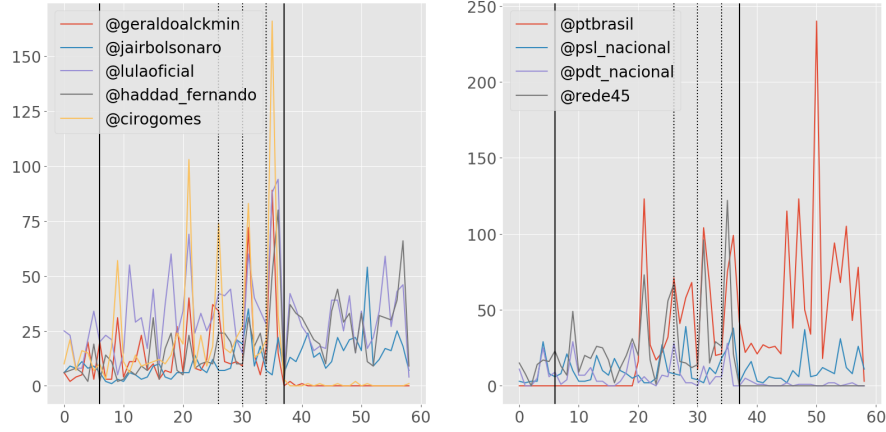
Fig. 1: Number of tweets over time by candidate (left) and party (right). The dotted lines represent the three major television debates and the first and second straight lines indicate when Bolsonaro was stabbed, which was a marking point during the campaign, and when the first round took place, respectively.

the 1st round. As expected, Bolsonaro and Haddad, the candidates that were in the 2nd round, kept a high number of daily tweets, alongside Lula, which was supporting Haddad.

Tweet Similarity

In order to understand how each candidate's tweets are similar to others', we calculated the cosine similarity between vector representations of a tweet. For the similarity between two candidates, we averaged the similarity of both candidates' tweet combinations.

For the tweet vector representation, a 50-dimensional Word2Vec pre-trained model was used[2]. A tweet containing N words $w_i, i \in [1, n]$ has a vector representation following Equation 1, where $embed$ is the word embedding function.

$$Tweet = \frac{1}{N} \sum_{i=1}^{N} embed(w_i) \tag{1}$$

The similarity results are shown in Table 2. For each candidate, we calculated his tweet self similarity and his tweet similarity in relation to another candidate and we applied a Tukey HSD significance test using a p-value of 0.05. All values are significant at such level of confidence.

In Table 2, Bolsonaro has the lowest self similarity, caused by the fact that his Twitter account has both more personal tweets, thanking supporters, and official tweets,

---

[2] http://nilc.icmc.usp.br/embeddings

Candidate/Candidate	Alckmin	Bolsonaro	Lula	Haddad	Ciro
Alckmin	0.465	0.404	0.431	0.431	0.419
Bolsonaro	0.404	0.392	0.397	0.395	0.383
Lula	0.431	0.397	0.437	0.431	0.403
Haddad	0.431	0.395	0.431	0.429	0.405
Ciro	0.419	0.383	0.403	0.405	0.404

Table 2: Average similarity between each candidates' tweets. All values are statistically significant at p = 0.05.

promoting his campaign. An evidence of this fact is "obrigado", Portuguese word for "thank you", being a highly frequent word in Bolsonaro's tweets. Bolsonaro also has the lowest average similarity across all opposing candidates, due to his personal tweets and his complete different political stance (all other candidates are from parties in the left to central political spectrum, while Bolsonaro is a right wing candidate) .

Lula, Haddad and Alckmin have high and equal similarity between each other. In case of Haddad and Lula, this is expected since they are both from the same party (PT) and focused on promoting the same candidate, Haddad. However, PSDB, Alckmin's party, is known to be the biggest PT opposition (even though PSDB is not a right wing party), running against them for presidency since 1994. This high similarity is caused by the fact that PSDB and PT are two of the biggest and long standing parties in Brazil and have similar online campaign strategies. A further analysis on how these candidates, that at first glance are political opponents, tweet similarly is possibly an important future work.

We also calculated the similarity between each party. The results are presented in Table 3. All values are significant at p = 0.05 using the Tukey HSD significance test.

Party/Party	PT	PSL	PDT	PSDB
PT	0.407	0.419	0.421	0.426
PSL	0.419	0.473	0.435	0.453
PDT	0.421	0.435	0.454	0.447
PSDB	0.426	0.453	0.447	0.474

Table 3: Average similarity between each parties' tweets. All values are statistically significant at p = 0.05.

As shown in Table 3, PSL, Bolsonaro's party, has one of the highest self similarity. Its lowest similarity is with PT, its main opponent. However, these similarity results do not accurately reflect all intricacies in Brazilian politics, mainly because these parties' accounts have the same objective of promoting their candidate and thus, show similar tweet content.

Lastly, we calculated the cosine similarity between each candidate and party combination. Table 4 presents the results.

Candidate/Party	PT	PSL	PDT	PSDB
Alckmin	0.417	0.454	0.436	0.466
Bolsonaro	0.386	0.419	0.397	0.406
Lula	0.418	0.433	0.434	0.439
Haddad	0.414	0.436	0.431	0.437
Ciro	0.393	0.426	0.417	0.421

Table 4: Average similarity between candidate's and party's tweets.

An interesting result in Table 4 is how Lula's and Haddad's tweets are different from PT's, their party, tweets. The similarity values between Lula/Haddad and PT are the lowest across all other parties. This is mainly caused by the intention of the candidates to separate themselves from the negative image that PT currently has in Brazil. PT has been involved in many corruption scandals in the last years and its identity has been stained negatively.

Bolsonaro again has the lowest average similarity, because of the different and diverse content that he tweets. His similarity with PT, his main opposition, is the lowest overall. It is also importance to note the high similarity between Alckmin and his party's, PSDB, tweets.

Finally, Figure 2 shows some graphical representations of the similarity values between candidates and parties.
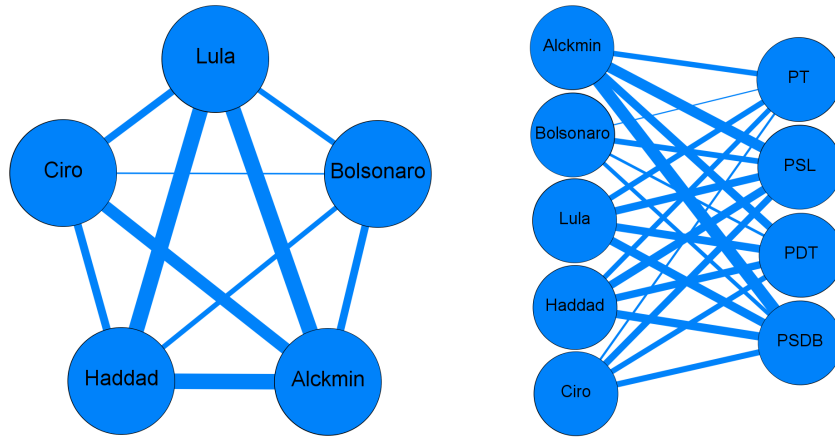

Fig. 2: Similarity graph between candidates and parties. The thickness of each edge is proportional to the similarity between the nodes (candidates/parties). In the left image, note the high similarity in the Lula/Haddad/Alckmin triangle and the low similarity between Bolsonaro and other candidates. In the right image, note the low similarity between PT and all candidates.

Citations between Candidates and Parties

In order to analyze how much a candidate/party was talked about by a certain candidate, we obtained the number of tweets in which a candidate/party was cited, either through his Twitter id or commonly used name, by each candidate. Tables 5 and 7 show the absolute number and the percentage of times a candidate and party was talked about by each candidate. Figure 3 shows the results presented in Tables 5 and 7 and explained below.
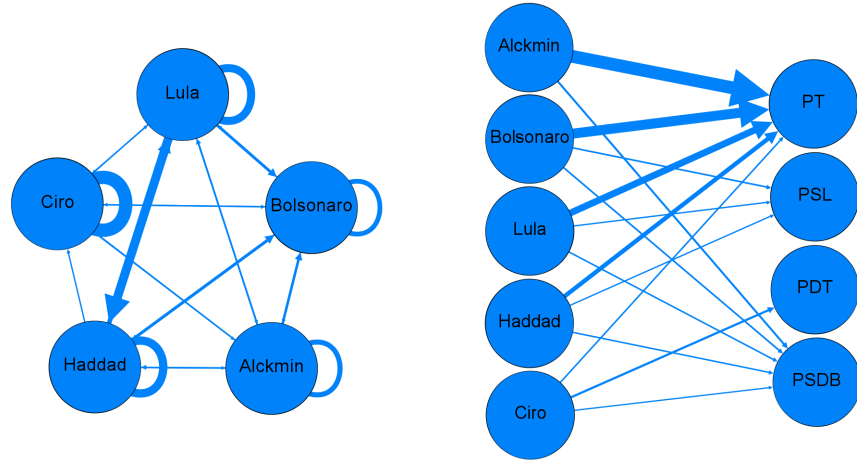

Fig. 3: Conversation graph between candidates and candidates/party. The thickness of the edges are proportional to the percentage of times the candidate cited the user. The direction of the edges is from the writer to the cited user. Note the thick edges between Lula and Haddad and the edges toward PT.

By/Who	Alckmin	Bolsonaro	Lula	Haddad	Ciro
Alckmin	159 (26.2%)	53 (8.7%)	20 (3.3%)	16 (2.6%)	7 (1.2%)
Bolsonaro	3 (0.4%)	169 (24.2%)	6 (0.9%)	18 (2.6%)	1 (0.1%)
Lula	2 (0.1%)	202 (10.8%)	870 (46.7%)	987 (53.0%)	3 (0.2%)
Haddad	1 (0.1%)	136 (11.8%)	265 (23.0%)	565 (49.0%)	5 (0.4%)
Ciro	2 (0.2%)	11 (1.1%)	4 (0.4%)	0 (0.0%)	806 (84.0%)

Table 5: Absolute number and percentage of times a candidate was cited by each candidate.

Even though at first glance, Lula and Haddad seem to be the most noted candidates in Table 5, they were mostly cited by themselves. Table 6 shows the absolute number of citations of each candidate excluding when they were cited by someone from their own party. Bolsonaro was the most cited candidate, because all opponents considered him as

the candidate to defeat, the one with the highest chances of winning. He was the central candidate of this election.

Candidate	Number of Citations
Alckmin	8
Bolsonaro	402
Lula	30
Haddad	34
Ciro	16

Table 6: Absolute number of times a candidate was cited by each candidate that was not in his own party.

By/Who	PT	PSL	PDT	PSDB
Alckmin	106 (17.5%)	0 (0.0%)	0 (0.0%)	7(1.2%)
Bolsonaro	92 (13.2%)	4 (0.6%)	0 (0.0%)	5 (0.7%)
Lula	155 (8.3%)	3 (0.2%)	0 (0.0%)	9 (0.5%)
Haddad	54 (4.7%)	2 (0.2%)	0 (0.0%)	5 (0.4%)
Ciro	5 (0.5%)	0 (0.0%)	14 (1.5%)	1 (0.1%)

Table 7: Number and percentage of times a party was cited by each candidate.

Analyzing Table 7, it is possible to understand how Bolsonaro and Alckmin used PT's current negative image in their favor. They often talked about PT, connecting the party to words such as "unemployed" and "corrupt", in order to tackle the unemployment and corruption problems during the last years of PT' presidency.

Most Frequent and Relevant Words

We calculated the frequency of each word resultant of the tweet preprocessing explained previously for each candidate and party. After analyzing the 15 most frequent words for each user present in this dataset, we decided to only present the frequent words that are related to the political discussion on Twitter, excluding verbs (e.g. vamos), numbers (e.g. milhões) and other parts of speech that are not relevant (e.g. forte). Table 8 shows the resultant words for each user translated to English.

Every candidate, but Ciro and his respective party, cited Bolsonaro frequently. This is an evidence that Bolsonaro was the central character in the 2018 election.

Alckmin's main campaign strategies, according to Table 8, are attacking PT and Bolsonaro, and remembering his followers about his work as the governor of the state of São Paulo, the biggest in the country. PSDB follows similar patterns to his candidate's tweets.

Bolsonaro often thanks for and sends hugs (usual greeting in Brazil) to his supporters' support while attacking PT and showing his position as against the previous government.

User	Words
Alckmin	brazil, pt, country, bolsonaro, geraldo, alckmin, paulo, state, round
Bolsonaro	brazil, bolsonaro, thank you, pt, hug, jair, support, against
Lula	lula, haddad, brazil, people, fernando, country, president, bolsonaro, democracy
Haddad	haddad, brazil, fernando, lula, people, bolsonaro, country, democracy, president
Ciro	ciro, brazil, country, thank you, change, support, people, debate, proposals
PT	haddad, brazil, bolsonaro, people, lula, education, country, president, democracy
PSL	bolsonaro, brazil, jair, captain, country, thank you, pt
PDT	ciro, brazil, gomes, question, debate, proposals, country, pdt, president, block
PSDB	brasil, pt, country, candidate, congressman, federal, paulo, state, health, bolsonaro

Table 8: Relevant frequent words by candidate and party translated to English. Note that "obrigado" was translated to "thank you", two words.

PSL, Bolsonaro's party, also thanks Bolsonaro's supporters, connects the candidate's image to the army (by using Bolsonaro's army rank) and attacks PT's image.

On Lula's and Haddad's side, they often cite themselves and talk about Bolsonaro. They portray themselves as the people's candidate, an image that PT has carried throughout many years. Lastly, a campaign strategy used by PT and its candidates was to show Bolsonaro as a threat to democracy. Brazil suffered from a military dictatorship from 1964 to 1985, and Bolsonaro's opponents often correlated his position as an army's captain to the dictatorship and the lack of democracy during that time. PT's strategy on Twitter was similar to Lula's and Haddad's, with the addition of education to the frequent words list: access to university level education was one of the areas where Lula invested the most during his terms.

Ciro and his party, PDT, had its tweeting peaks during TV debates and often talked about them and Ciro's proposals. Ciro also has some personal tweets thanking his supporters.

Conclusion

The 2018 Brazil Presidential Election was marked by the rise of social media in the political campaign and the shift from the left to a right wing government. Bolsonaro was able to win the election without almost any televised propaganda time, focusing his campaign on social media. He was the central character of this election, being cited by most candidates. Bolsonaro also applied a more personal approach to his tweets, talking to his supporters on Twitter.

My analysis also showed similarities between Alckmin, Haddad and Lula's, candidates from two of the biggest parties in Brazil, tweets. PT was the central party during this election, since most opposing candidates tried to use the current negative image associated with the party, while the party's candidates tried to move away from this image. PT also used Lula's positive image with a big part of the Brazilian population to transfer votes to Haddad. Lula was leading the polls with almost double the number

of votes (39%) of Bolsonaro (19%) before his candidacy got blocked by the Supreme Court. Haddad, when he was included in the polls, had only 4% of the votes. By the end of the first round, Haddad received over 29% of the votes. This strategy was somewhat successful, allowing Haddad to go to the second round of the election and to compete head to head against Bolsonaro, losing in the second round with 44.87% of the votes.

References

1. Bollen, J., Mao, H., Zeng, X.: Twitter mood predicts the stock market. Journal of computational science **2**(1), 1–8 (2011)
2. Cha, M., Haddadi, H., Benevenuto, F., Gummadi, P.K., et al.: Measuring user influence in twitter: The million follower fallacy. Icwsm **10**(10-17), 30 (2010)
3. Jungherr, A., Schoen, H., Jürgens, P.: The mediation of politics through twitter: an analysis of messages posted during the campaign for the german federal election 2013. Journal of Computer-Mediated Communication **21**(1), 50–68 (2016). https://doi.org/10.1111/jcc4.12143, bluehttp://dx.doi.org/10.1111/jcc4.12143
4. Kušen, E., Strembeck, M.: Politics, sentiments, and misinformation: An analysis of the twitter discussion on the 2016 austrian presidential elections. Online Social Networks and Media **5**, 37–50 (2018)
5. Messias, J., Vikatos, P., Benevenuto, F.: White, man, and highly followed: Gender and race inequalities in twitter. CoRR **abs/1706.08619** (2017), bluehttp://arxiv.org/abs/1706.08619
6. Sharma, S., Shetty, N.P.: Determining the popularity of political parties using twitter sentiment analysis. In: Satapathy, S.C., Tavares, J.M.R., Bhateja, V., Mohanty, J.R. (eds.) Information and Decision Sciences. pp. 21–29. Springer Singapore, Singapore (2018)
7. Tumasjan, A., Sprenger, T.O., Sandner, P.G., Welpe, I.M.: Predicting elections with twitter: What 140 characters reveal about political sentiment. Icwsm **10**(1), 178–185 (2010)
8. Yaqub, U., Chun, S.A., Atluri, V., Vaidya, J.: Analysis of political discourse on twitter in the context of the 2016 us presidential elections. Government Information Quarterly **34**(4), 613–626 (2017)