

Classification and influence of Twitter bots in 2019 Canadian federal elections at British Columbia.

Ana Cecilia Leon Morales

June 22, 2020

This work was developed at The University of British Columbia under the supervision of professor Ruben H. Zamar.

Summary

Social bots can affect online communication among humans. Previous research suggests that bots are present on political discussions and social movements [4] [6] [2]; a qualitative study of bots in Canadian politics found that these bots appeared to have a limited influence on the 2015 Canada federal election. We aim to analyze the presence of Twitter bots, their classification, and influence on the political discussions about the 2019 Canadian federal elections at British Columbia. To accomplish this goal, we collected data from three main topics which are: elections, political parties, and party leaders from June 21st to October 21st of 2019 by accessing the Standard Twitter API, and we performed two clustering analyses according to the characteristics of suspicious accounts and the content of their issued tweets.

1 Motivation

The popularity of Twitter, which positions it within the top 20 most popular social networks worldwide as of October 2019 with 330 million monthly active users [1], has attracted many automated programs which are known as “Bots”; these programs have appeared to be a double-edged sword for the social network.

Bots can autonomously carry out actions such as tweet or re-tweet, as well as follow, un-follow, or directly message other Twitter accounts. Beneficial uses of bots include the transmission of useful information such as global news, the automatic generation of interesting or creative content, and the automatic response to Twitter users through direct messages. Conversely, malign bots spread spam or malicious content and distort the shared information.

As Twitter became an important tool for protests, political conversations, and people mobilizations, the damage to shared information in this social media increased significantly. In September 2017, Twitter’s executives informed the Intelligence Committee of the United States that they have found evidence that signaled Russia was behind some fake and automatized accounts that were active during the 2016 presidential election. Studies [5] have found that about 20% of the activity on Twitter related to such elections came from bots suspicious accounts.

In contrast to the United States, it appears that the influence of Twitter bots on Canada is limited. A study from 2014 [5] found that just under half of Canadians had visited a government website. Even fewer had friended or followed a political actor on Facebook (6 percent) or Twitter (4 percent). In this study, it appears that not only do Canadians avoid politicians online, but they also avoid politics of all sorts since only 18 percent of Canadians had signed a petition, posted a political message on Facebook (14 percent) or had retweeted political content (3 percent).

This previous qualitative research [5] on the usage of political bots states that Twitter bots had a limited impact in the social media discourse during 2015 Canada federal election; it identifies four main types of bots in the Canadian political ecosystem which are presented in the following scheme.

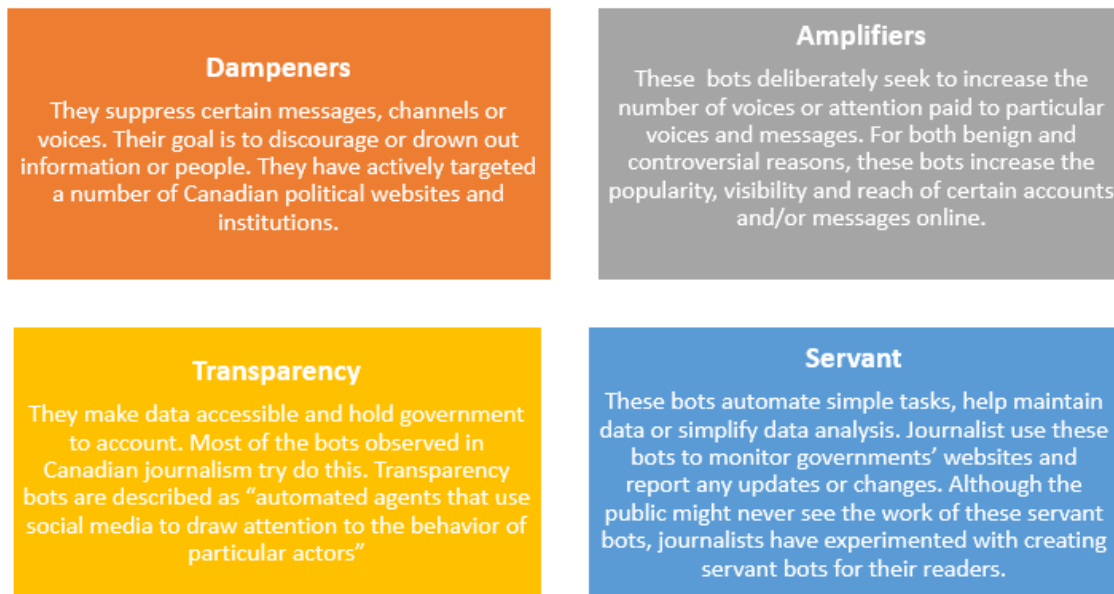


Figure 1: Identified types of bots that were present in political discussions about the 2015 Canada federal elections according to McKelvey and Dubois [5].

A more detailed description of each type of bot, as well as lists of suspicious accounts that were identified through this study are provided in [5]. In what follows we present some examples of these types of bots that were presented in [5].

- **Dampener.** An example of these bots behavior is told by a professor of law at The University of British Columbia. His testimonial at [5] tells that after issuing a tweet about the trending #GoodRiddanceHarper, which celebrated the resignation of ex-Prime Minister Stephen Harper, he received a negative reply that quickly received over 1,000 likes in few hours. He found out that bots amplified this negative reply to discourage him from tweeting.
- **Amplifier.** According to [5], the account *hashtag_cdnpoli* is an example of an amplifier bot. This account retweets Canadian politics news with the #cndpoli hashtag, and it issues tweets more than 10 times per day. Another example of these bots was observed during the 2017 provincial elections in British Columbia, when a suspicious amplifier account, *ReverendSM*,

issued content mainly about the incumbent Christy Clark of the Liberal Party with accusations of corruption.

- **Transparency.** The automated account *gccaedits* is a known example of this type of bot. This account issues a tweet every time an anonymous edit to Wikipedia is made from a Government of Canada IP address. This type of bots was scarcely found in [5].
- **Servant.** These bots are different from the others in the sense that they are referred to automated programs that perform tasks beyond the publishing of information through social media. For example, Kathleen Wynne, who was the Communications Office of Ontario Premier until 2018, managed her Facebook page so that it automatically removed posts that contained any word from a list of banned words. Another application of this kind of bots out of the political context in Canada is by providing assistance to manage the problem of online child exploitation by identifying suspicious images on public websites.

2 Objectives

We aim to analyze the presence of Twitter bots and the information they disseminated at British Columbia during the 43rd Canadian general election through the following goals:

- To describe the time evolution of tweets with political discourse in the months leading to the 2019 Canadian federal elections.
- To study the presence of bots, classify, and characterize them according to their account characteristics.
- To analyze the types of tweets that are issued by each type of bot according to their content, and investigate if bots can be classified under the classes: ‘Dampeners’, ‘Amplifiers’ and ‘Transparency’, which were proposed in the qualitative research of 2015 Canada federal elections [5].

3 Data description

From June 21st to October 21st of 2019, we collected a sample of tweets related to the Canadian federal election which were issued by Twitter accounts registered in British Columbia and surrounding areas according to the information provided by users. Data were accessed by consulting the Standard Twitter API through three queries that contain terms related to three main topics: the Canadian federal elections, political parties in the House of Commons of Canada, and leaders of these parties.

In addition to the information related to each tweet, the probability of Twitter accounts being a bot was estimated by Botometer, which is the most popular random forest classifier of Twitter accounts to detect Bots based on hundreds of features. These features are categorized into six main classes which are related to networks of the account, user, friends, temporal patterns in issued tweets, their content and sentiment.

4 Exploratory data analysis

We present an exploratory analysis of the previously described data.

The number of tweets that were collected from 113,960 different Twitter accounts through our three queries is as follows:

Query	Numer of tweets
Elections	247, 116
Parties	196,774
Leaders	663,610

Table 1: Number of total collected tweets by query.

4.1 Historic trends

4.1.1 Elections

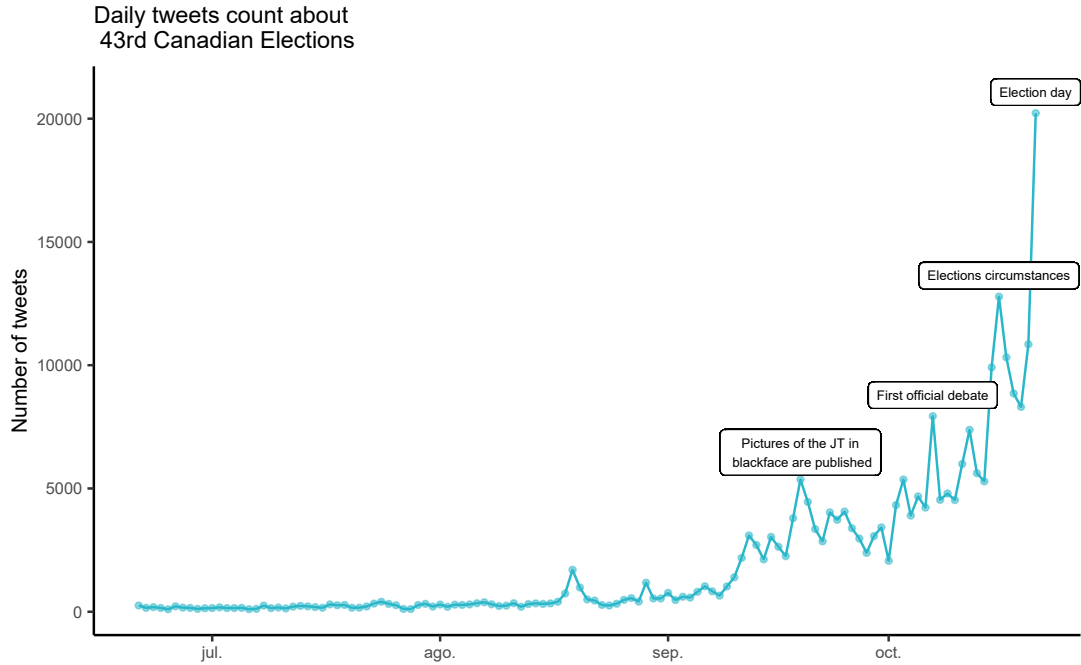


Figure 2: Trend of the number of tweets in British Columbia about the 2019 Canadian federal elections that were sampled.

A total of 247,116 tweets with content related to the 2019 Canadian federal elections in British Columbia (BC) was sampled, this quantity represents 22.31 % of the total number of collected tweets. As is noted in Figure 2, the number of tweets has an increasing trend. The number of tweets

starts increasing at the beginning of September and reaches its maximum of 20,225 on the election day; during this period, peaks match remarkable situations: on September 19th, controversial pictures of Prime Minister Justin Trudeau in blackface were published, while on October 7th the first official debate was held.

4.1.2 Parties

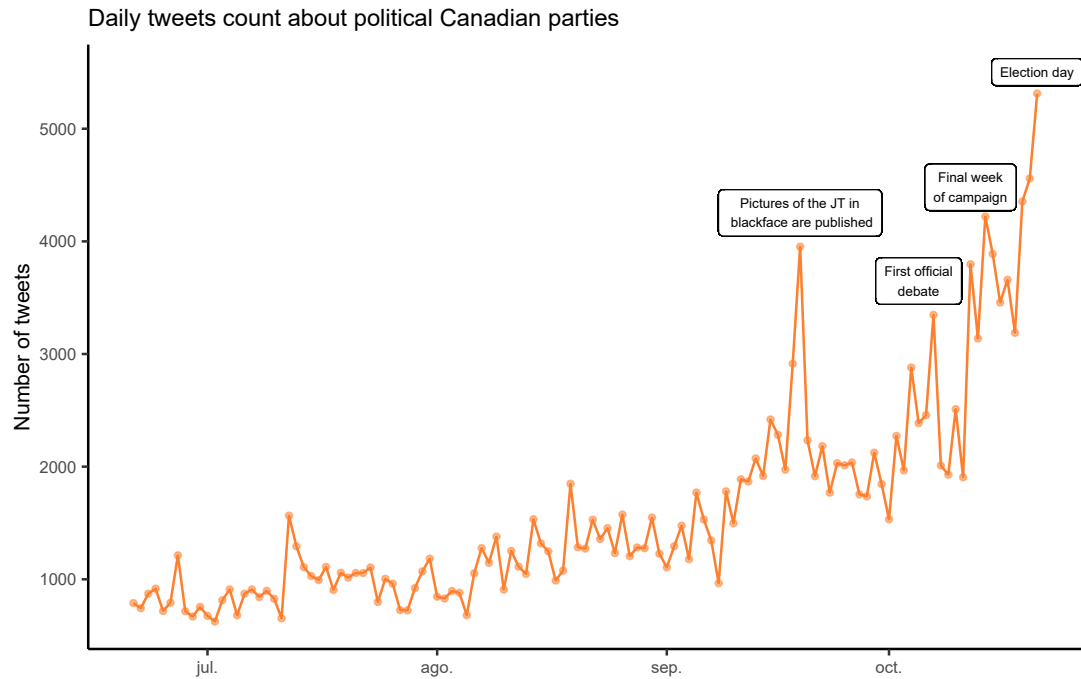


Figure 3: Trend of the number of tweets in British Columbia about the political Canadian parties that were sampled.

The “Canadian parties” topic corresponds to the less popular with a total of 196,774 tweets of related content, which represents 17.78 % of the total number of sampled tweets in BC. Similarly to the tweets about elections, according to Figure 3, the number of tweets about political parties has an increasing trend that reaches its maximum on the election day with 5,313 tweets, peaks of this trend match the same remarkable situations of the tweets about elections such as the controversial pictures of Justin Trudeau and the first official debate, as well as the final week of the campaign.

4.1.3 Leaders

The largest quantity of collected tweets corresponds to content related to the six leaders of each party at the House of Commons on 2019, which are Justin Trudeau from liberals, Andrew Scheer from conservatives, Jagmeet Singh from New Democratic Party, Yves-François Blanchet from Bloc Québécois, Elizabeth May from Green Party, and Maxime Bernier from People’s party.

As it is noted in Figure 4, the number of collected tweets has an increasing trend, where Justin Trudeau is the most popular leader followed by Andrew Scheer, with an exception on October 3rd and 4th when the number of tweets about Andrew Scheer is bigger than Justin Trudeau; these dates match with the first French-language debate in which Scheer was questioned about his personal views on abortion rights and his party's environmental policies. The third place in popularity is interchanged among Elizabeth May, Maxime Bernier, and Jagmeet Singh, while the leader of Bloc Québécois was the less mentioned in tweet contents issued by users registered at British Columbia.

Similarly to tweets about elections and parties, there are some peaks that match dates corresponding to the first official debate and the day when pictures of Justin Trudeau were published, which is the day when this leader presents its maximum number of tweets with a total of 15,892.

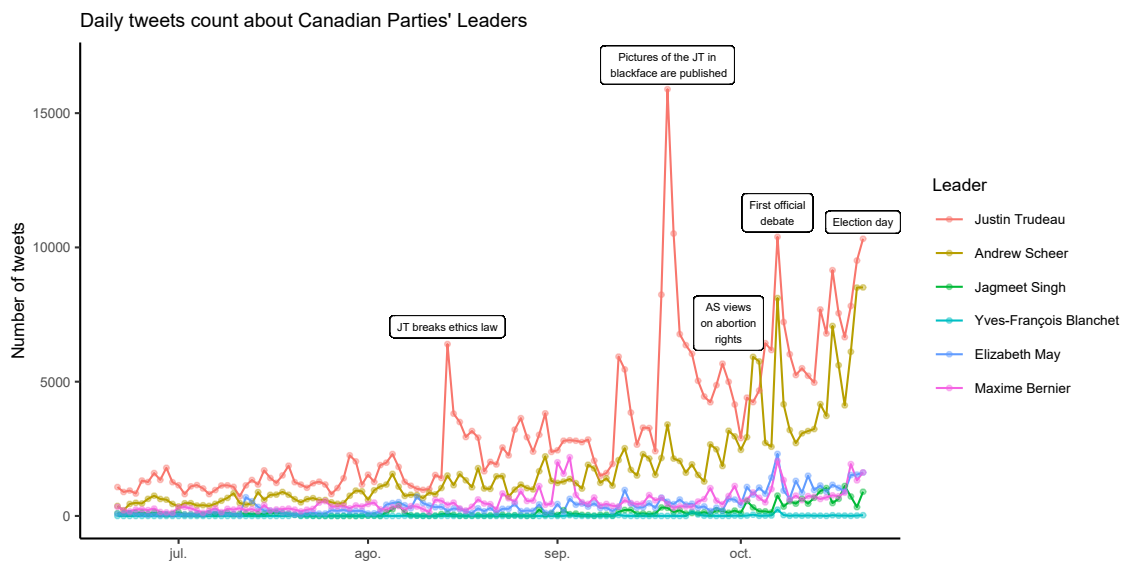


Figure 4: Trend of the number of sampled tweets in British Columbia with content about the leaders of political Canadian parties.

4.1.4 Twitter accounts

The number of different Twitter accounts, which issued content about one or more of the considered topics, has an increasing trend as it is shown in Figure 5. Data about Twitter accounts were collected on a weekly basis according to the norm ISO 8601 where each week starts on Mondays and finishes on Sundays. It is worth to remark that the week 2019-43 only contains information about the 21st October, which is the day of the election and that has a total of 31,251 accounts, while the number of accounts of previous weeks corresponds to the cumulative counts of the whole week.

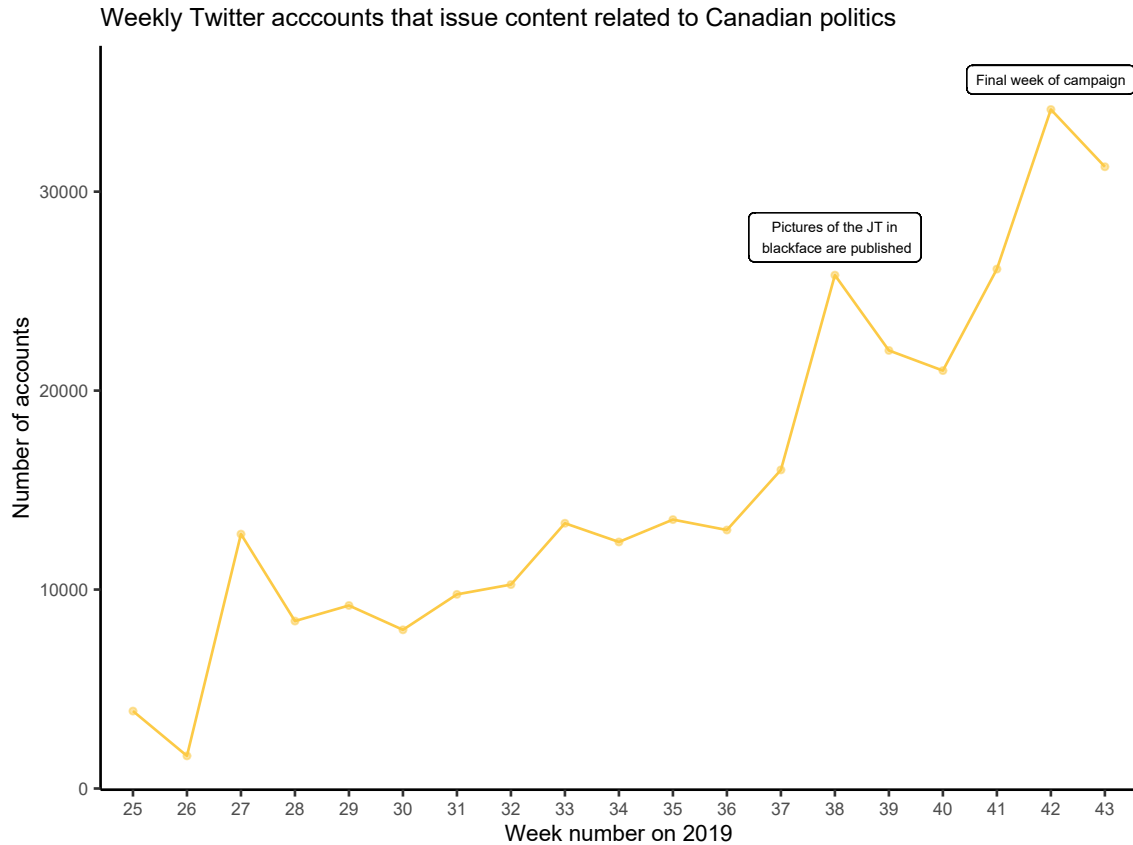


Figure 5: Trend of the number of users in British Columbia that issued content about 2019 Canadian elections, parties and leaders.

4.2 Presence of bots by query

The probabilities of Twitter accounts to be a bot were estimated at the end of each week through the public API of Botometer (<https://botometer.iuni.iu.edu/#!/api>), which is the most well known publicly available classifier for bot detection, using the packages `httr`, `xml2`, `RJSONIO`, `mongolite`, `jsonlite`, and taking the function `botcheck()` provided by Joey Marshal (<https://github.com/marsha5813/Botcheck>) as an initial reference to create our own code.

Botometer [3] is a machine learning classifier that is the result of a joint project of the Network Science Institute and the Center for Complex Networks and Systems Research at Indiana University. This classifier was trained on more than 5.6 million tweets to classify Twitter accounts according to their estimated probability of being a bot based on approximately 1,200 features that feed seven different classifiers. These classifiers correspond to the overall classifier and six classifiers for each subclass of features, which are categorized as follows:

- **Network.** It is comprised of features that capture information diffusion patterns. It consists of statistical features like degree distribution and centrality measures of various networks based on retweets, mentions, hashtags, etc.
- **User.** This set of features consists of data related to the Twitter account including language, geographic locations, creation time, etc.
- **Friends** It is comprised of features that are built from descriptive statistics relative to social contacts of the account, such as the median, moments, and entropy of the distributions of their number of followers, friends, and so on.
- **Temporal** This set of features captures timing patterns of content generation and consumption, such as the inter-tweet time distribution.
- **Content** It is comprised of features that are based on linguistic cues computed through natural language processing, especially part-of-speech tagging.
- **Sentiment** This set of features is built using general purpose and Twitter specific sentiment analysis algorithms including algorithms that consider happiness scores, emoticon scores, the three dimensions that describe emotional states according to a psychological model, which are called arousal (how pleasant emotion is), dominance (what is the dominant nature of the emotion) and valence (the intensity of emotion), etc.

The reason to estimate the probabilities at the end of each week is because of the classifier extracts features from both the analyzed Twitter account and from up to the last 200 issued tweets, so probabilities change along time. As Figure 6 illustrates, accounts with lower probabilities of being a bot keep consistent along the period of study while accounts with higher estimated probabilities tend to be less stable, although there are accounts with consistent high probability of being a bot such as the account *MichelleCasti*.

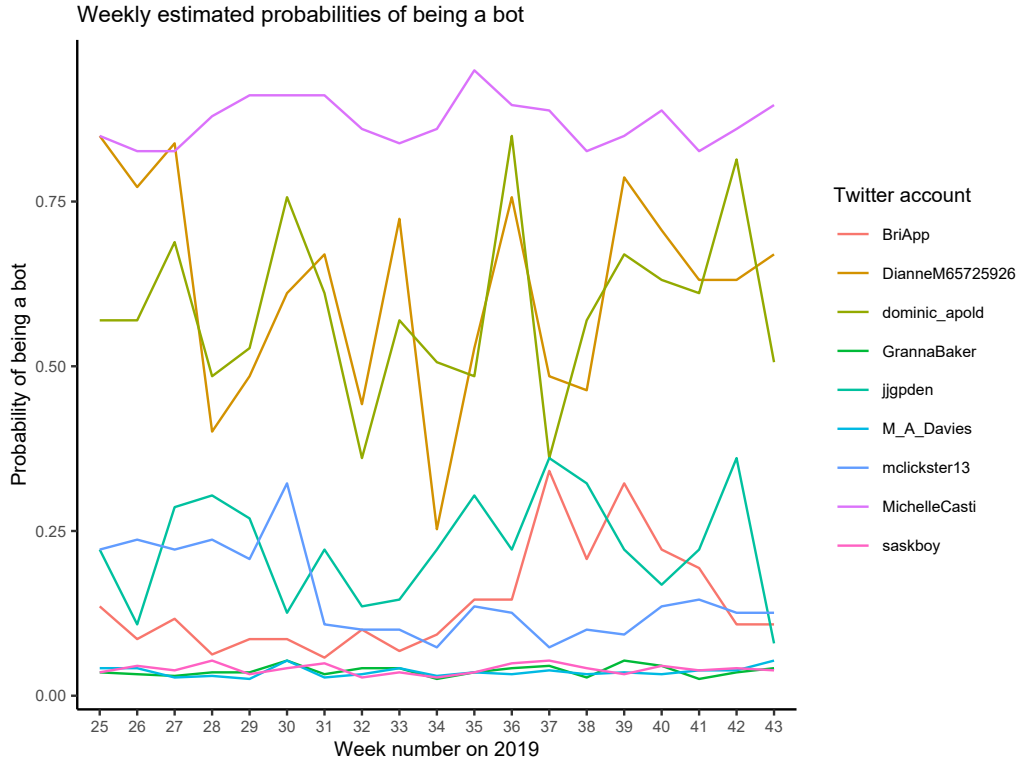


Figure 6: Trend of the estimated probabilities of being a bot of some sampled Twitter accounts that issued content related to elections, parties, or leaders every week during the period of study.

We classify Twitter accounts into five categories according to their estimated probability of being a Bot, where the class “missed” represents accounts that could not be analyzed because its information was restricted or because the account had been eliminated before the end of the corresponding week.

4.2.1 Elections

During the period from July 22nd to October 21st, as Figure 7 shows, the most of tweets about elections are issued by accounts with a probability of being a bot lower than 25%. Around 5% of the tweets come from accounts with a probability of being a bot greater than 50%, while tweets issued by the category of accounts with an estimated probability greater than 75% represent the smaller percentage of the total collected tweets. This distribution is consistent along this period of study both in a monthly basis as table 2 shows, and in a daily basis as Figure 21 (see Appendix) illustrates.

To describe the content of the tweets by month, we split tweets by words and give the *Term frequency-inverse document frequency* score to each of their words. Tf-idf is a commonly used score for extracting features in a bag-of-words model, and it represents the importance of a word for a document over a corpus of documents, which in this case corresponds to the set of tweets about elections. We provide a more detailed description of this score calculation in Section 5.1.2

Distribution of tweets by the probability of being a bot

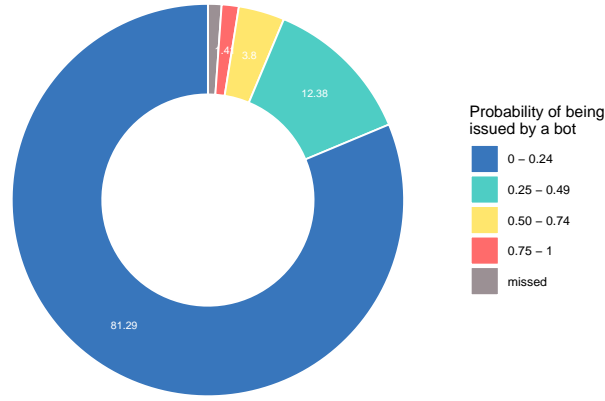


Figure 7: Proportion of tweets about elections that are issued by accounts classified according to their probability of being a bot.

Category of issuing account	July - August	August - September	September - October
Probability < 0.25	82.23	81.58	81.15
Probability $\in [0.25, 0.50)$	11.78	12.40	12.42
Probability $\in [0.50, 0.75)$	3.74	4.08	3.73
Probability $\in [0.75, 1)$	1.50	1.31	1.43
Missed	0.75	0.63	1.27

Table 2: Proportion of tweets about elections that are issued by accounts classified according to their probability of being a bot by month.

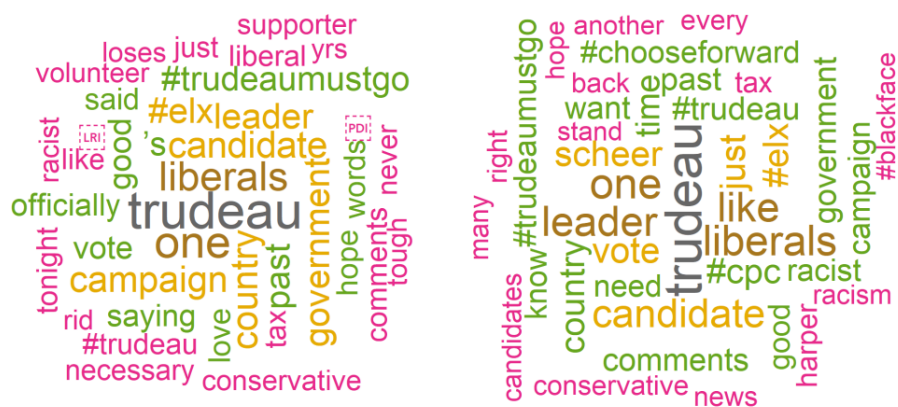
Figure 8 shows the highest scored words of tweets with content about elections according to two categories which are “Likely bot” and “Likely human”; these classes correspond to the probability of being issued by a bot greater or equal to 75% and less than 75%, respectively. During the period from July 22nd to August 21st, most of these words are similar between categories, however, there are some terms, like “bernier”, that appear only in the category “Likely bot”. In the following periods from the end of August to the day of elections, “truedau” is consistently one of the most important terms in both classes of accounts.

In general, similar terms are present into the content of tweets generated by both types of account; this fact raises the question if the accounts are well classified, or if there is a spreading information pattern from bots content to the rest of accounts, or if bot accounts emulate the content that is issued by “likely human” accounts.

Period: July - August



Period: August - September



Period: September - October

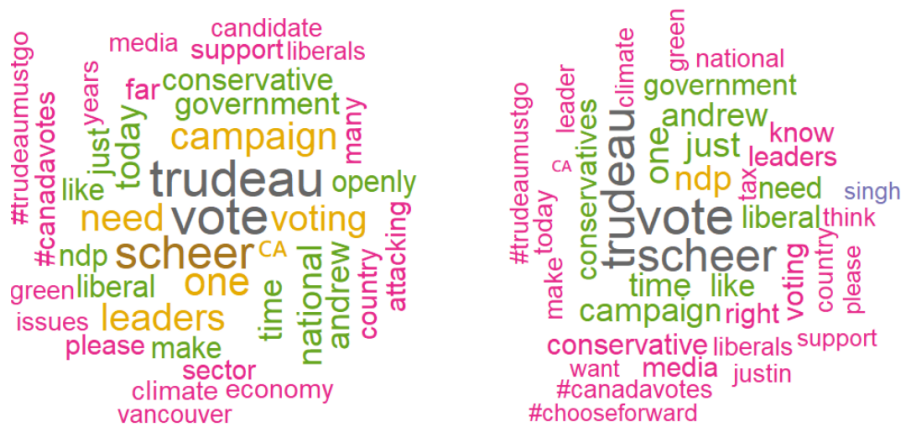


Figure 8: Cloud of highest scored words into tweets about elections from July 22nd to October 21st.
 Left-hand side clouds correspond to tweets issued by accounts belonging to category “Likely bot”;
 Right-hand side clouds correspond to accounts belonging to the “Likely human” class.

4.2.2 Parties

During the period from July 22nd to October 21st, most of the tweets about Canadian parties are issued by accounts with a probability of being a bot lower than 25% as Figure 9 exhibits. Around 5% of tweets come from accounts with a probability of being a bot greater than 50%, while tweets issued by accounts with an estimated probability greater than 75% represent the smaller percentage of the total collected tweets. This distribution is consistent along this period of study on a monthly basis as table 3 shows; the percentages of tweets by account category are similar to the corresponding percentages of tweets about elections. Furthermore, as Figure 22 illustrates (see Appendix) the percentages of tweets issued by each type of account keep consistent on a daily basis.

Distribution of tweets by the probability of being a bot

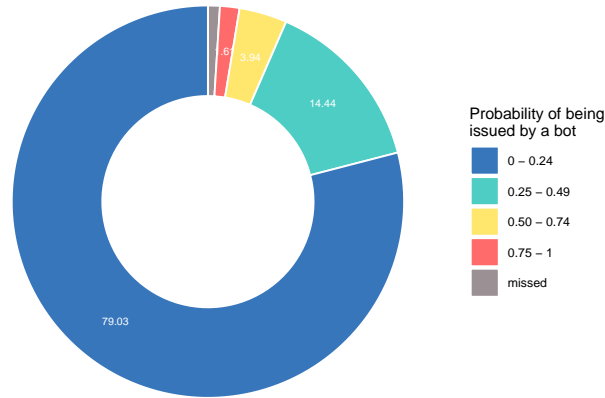


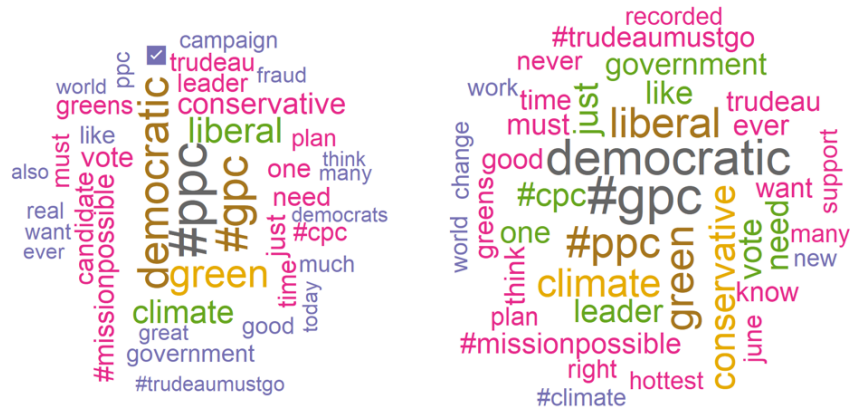
Figure 9: Proportion of tweets about Canadian parties that are issued by accounts, classified according to the probability of being a bot.

Category of issuing account	July - August	August - September	September - October
Probability < 0.25	79.00	80.10	78.40
Probability $\in [0.25, 0.50)$	14.60	13.90	14.70
Probability $\in [0.50, 0.75)$	4.00	3.80	4.00
Probability $\in [0.75, 1)$	1.60	1.50	1.70
Missed	0.70	0.70	0.65

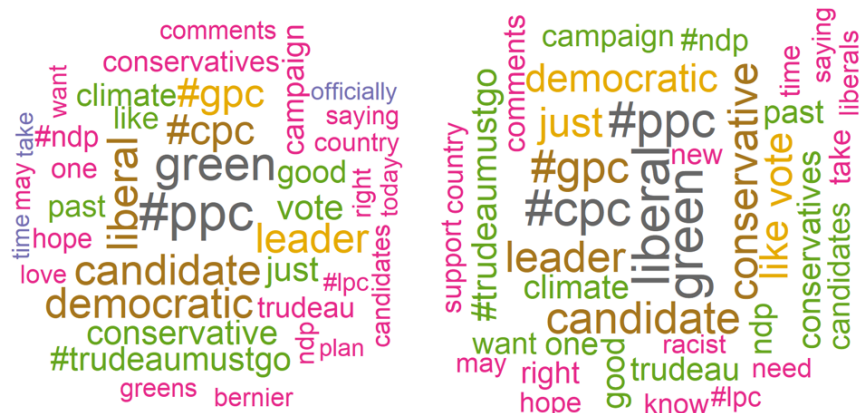
Table 3: Proportion of tweets about Canadian parties that are issued by accounts, classified according to the probability of being a bot by month.

Figure 10 shows the highest scored words of tweets with content about Canadian parties according to categories “Likely bot” and “Likely human”. In general, most of these words are similar among categories, however, there are some terms that are most frequent into the class of “likely bot” such as the official hashtag of Peoples’ party “#ppc” during the three months of study, and such as the terms “conservative” and “democratic”, which are present during the last period of study.

Period: July - August



Period: August - September



Period: September - October

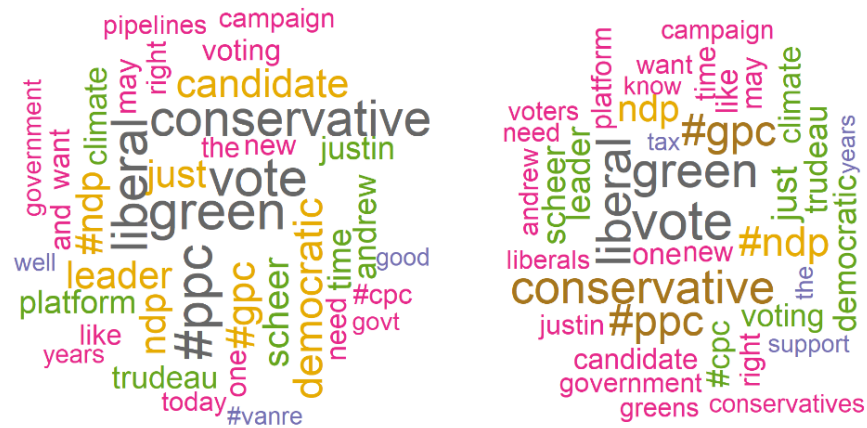


Figure 10: Cloud of highest scored words into tweets about Canadian parties from July 22nd to October 21st. Left-hand side clouds correspond to tweets issued by accounts belonging to category “Likely bot”; Right-hand side clouds correspond to accounts belonging to the “Likely human” class.

4.2.3 Leaders

Figure 11 shows the percentages of tweets with content about Canadian parties leaders during the period from July 22nd to October 21st according to the type of issuing account. It is noted that the most of tweets about the six leaders are issued by accounts with a probability of being a bot lower than 25%; around 5% of tweets come from accounts with a probability of being a bot greater than 50%, while tweets issued by accounts with an estimated probability greater than 75% represent the smallest percentage of collected tweets. This distribution is similar to the corresponding percentages of tweets about elections and Canadian parties, and it is consistent in a daily basis from three months before election day as Figure 23 (see Appendix) shows.

Distribution of tweets by the probability of being a bot

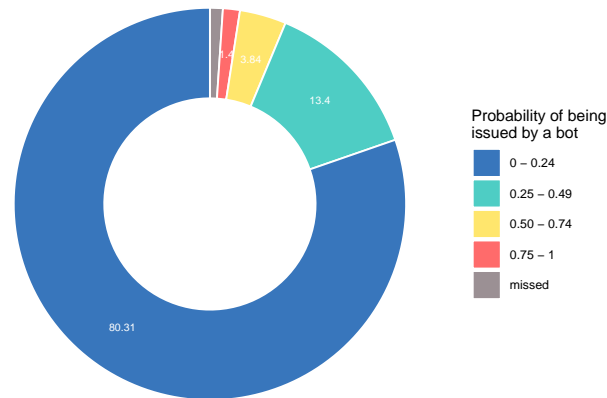


Figure 11: Proportion of tweets about leaders of Canadian parties that are issued by accounts, classified according to the probability of being a bot.

We calculate the percentage of tweets according to four categories of issuing accounts during the last month of election campaigns including the election day, and we display these percentages according to the mentioned leaders into tweets as Figure 12 shows. It is noted that tweets with the higher probabilities of being issued by a bot have content about the leader of the Peoples' Party, Maxime Bernier, while the leader of Conservative Party, Andrew Scheer, is mentioned by accounts that mostly have a small probability of being a bot: 82.3% of his tweets belong to accounts with the lowest probability of being a bot.

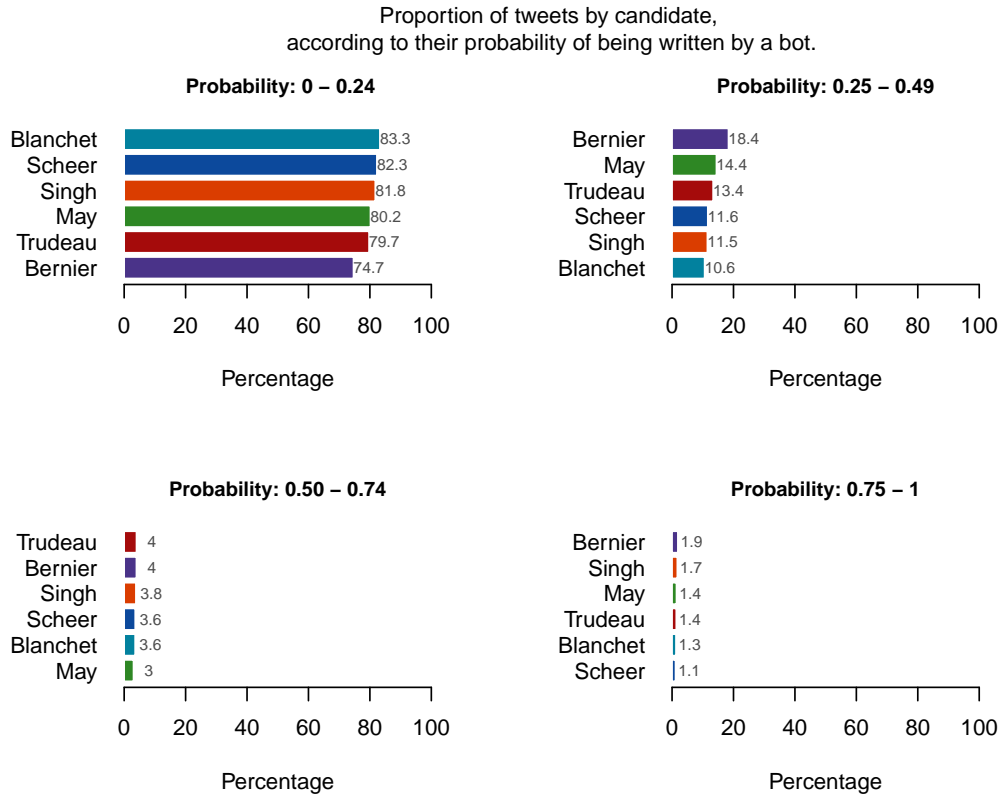


Figure 12: Percentages of tweets issued by account type and mentioned leader from September 22nd to October 21st.

5 Clustering

To accomplish our goal of investigating if accounts that are denominated by us as “likely bots”, and also as ‘bots’ or ‘suspicious accounts’ in what follows, can be classified into the categories that were proposed by McKelvey and Dubois in 2017 (omitting the servant type), we perform a clustering analysis to the set of tweets about elections that were issued from September 22nd to October 21st and their corresponding accounts.

We apply the standard k-means to our first set of features, which consists of a subset of 18 variables that describe account characteristics, these characteristics are provided by the standard Twitter API (see the Appendix to know the set of all characteristics). Then, we study the content of tweets that are issued by each type of account. We do this through the application of the sparse hierarchical clustering proposed in [8] to our second set of features, which captures the content of tweets. In this way, we characterize the kind of content that is issued by each type of account and compare our findings to the results of McKelvey and Dubois.

According to Table 4, this group consists of 802 accounts. This group of accounts issued 1765 tweets; we use the content of these tweets to build a set of features based on a bag of words model, which has a total of 152 included words that reached a tf-idf score of at least 25¹. We pick the tuning parameter for regularization $s = 2.77$, this generates the highest Gap statistic among different values of s and a set of 15 words with non-zero weights as Figure 16 shows.

Table 13 (see Appendix) presents words that are used as features for clustering the tweets belonging to this type of account. As we illustrate in Figure 16, there are 15 words with non-zero weights that separate tweets according to their content, it is noted that some of the words with non-zero weights are related to economy, health, and energy.

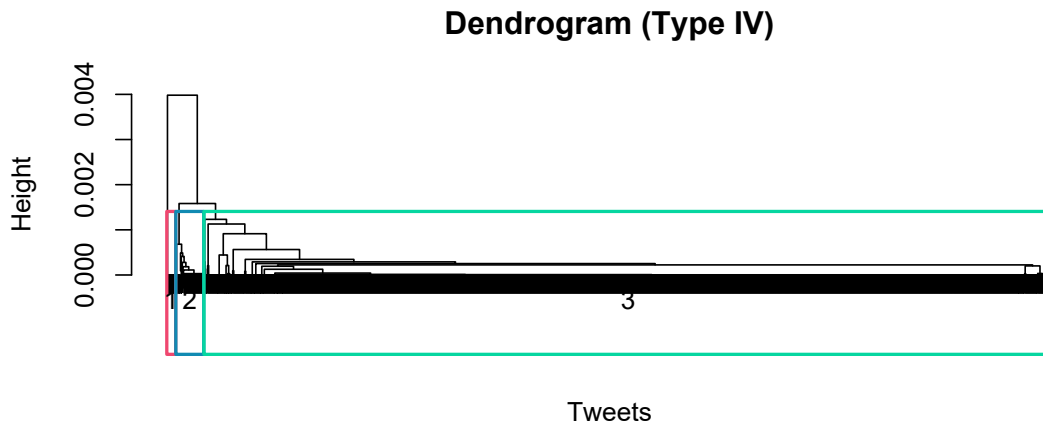


Figure 20: Dendrogram generated by the sparse hierarchical clustering of tweets that were issued by accounts of Type IV (“Not amplified retweeters”).

We separate tweets into three clusters of sizes 18, 56, and 1691, respectively, according to the dendrogram that is shown in Figure 20. In what follows we show a sample of tweets belonging to each group. It is noted that Cluster 2 consists mainly of tweets about the debate that was held on October 7th. Although we do not detect a clear pattern in the contents of Cluster 3, we note the presence of terms related to the green party, the new democratic party, and the people’s party. Contrary to the conservative leader, who appears in the most relevant words of all types of account, and to Justin Trudeau, who is mentioned in tweets of all types of account with mostly negative messages, these not amplified terms just appear into the most important words of this group and “Repliers”.

¹This threshold corresponds to the 99.95% percentile of scores, and it was chosen based on the RAM memory of the computer that was available at the moment of the analysis. This threshold can be diminished when a computer with a bigger RAM memory capacity is employed for doing the analysis.

Tweets from Cluster 1
1. Imagine federal leaders openly attacking Quebec's aviation industry, or Ontario's auto sector. It wouldn't happen. And nor should it. So why are three federal leaders openly attacking Alberta's energy sector? They are undermining national unity. (18 times)
Sampled tweets from Cluster 2
1. I'm so excited for the Leaders' Debate tonight - come join me, my team, and LPC President @suzcowan this afternoon for a debate watching party! RSVP and info:
2. Canada's federal leaders have spent the most time visiting Ontario, Quebec and B.C. during the first half of the 2019 election campaign
3. The two-hour election debate saw federal party leaders clash over ethics, climate change and the economy. But no one immediately emerged as the clear winner.
4. Crosshairs on people's heads. How has this become part of our political discourse? This election has seen images of nooses, swastikas on signs, racist rhetoric from leaders, slurs thrown at candidates - Canada, we are not as safe from all this as we keep pretending we are.
5. Today's Conservative leaders say we can't transition our economy to meet the greatest existential threat of our time. Seth Klein asks: Where is the courage and imagination of their predecessors? #cdnpoli #elxn43
6. The exact right answer by Scheer. I'm not interested in what ex foreign leaders are saying. I'm focused on what Canadians are saying and focused on finishing strong. #elxn43
7. How one can have a debate without Fife or Hébert or Akin or Kapelos interrogating the Leaders is beyond me #cdnpoli #elxn43 #cdnmedia
8. Folks on Twitter are ranking their positive/negative feelings towards the leaders during the #CanadaDebates2019 #leaders-debate2019. For anyone wondering who's winning ; impressing Canadians - wow - @theJagmeetSingh is crushing it. #elxn43
9. Ahead of #elxn43, I invited federal leaders to discuss the key issues of housing, transit, ; opioids with me. All agreed except for Andrew Scheer. When I read his platform I knew why. It's clear to me that Andrew Scheer would be worse than Stephen Harper. #cdnpoli #vanpoli
10. The one good thing about the cancellation of the federal leaders' Munk Debate on foreign policy is that Canadians were spared the embarrassment of watching their federal party leaders make excuses for themselves... #elxn43
Sampled tweets from Cluster 3
1. My choice was reaffirmed during the campaign, as I have watched the NDP spread misinformation about Green candidates, and even ask us to step aside for them. #cdnpoli #elxn43 #Burnaby #NorthVan
2. A good summary of Green policies! #TeamGreen #GreenWave #elxn43
3. Scheer continues to support controversial candidates including Port Moody-Coquitlam Conservative candidate Nelly Shin (also mentioned in the article). #BCpoli #Cdnpoli #Elxn43
4. Mr Trudeau, In the interests of fairness; informing (Canadian flag emoji) about one of the most important issues of our day, would you please RT this tweet with its links to the our short film, @OveraBarrelDoc
5. Trudeau and Morneau use hand gestures to illustrate how much influence women actually have in the Trudeau government. #elxn43 #canadavotes2019
6. TRUDEAU flew in 60,000 Syrian Islamists - housed them and gave them cheques TRUDEAU opened our borders with one tweet - free housing/cash/medical This poor First Nations Canadian has to go to the UN to plead for drinking water (confused face emoji) #TrudeauMustGo #elxn43
7. Huge overflow crowd in Victoria #yyj for Jagmeet Singh and South Island #NDP candidates! #elxn43 #UpRiSingh
8. Tomorrow we have our election in Canada. Andrew Scheer and his Conservative party aren't much different from the GOP. Its tax breaks for the rich, limiting immigration, questioning #ClimateChange, opposing abortion and same sex marriage. I'll be voting for Justin Trudeau.(Grinning Face with Smiling Eyes Emoji)(Thumbs up emoji)
9. Quebec's mandatory private drug insurance system is the most expensive way to provide drug coverage to all citizens. It would cost Canadians \$5 billion more per year than we are paying today. #pharmacare #elxn43 #cdnpoli #vote4care
10. The support doesn't count until it's in the ballot box. Keep working hard folks. I NDP is on the rise and could win much of British Columbia, Atlantic Canada, Ontario #Ndp #uprising #cdnpoli #elxn43

Table 8: Sample of tweets that belong to each cluster according to the dendrogram of tweets that were issued by Not amplified retweeters.

6 Conclusion

To analyze the presence of bots into the politic discourse at BC during the period of study that starts three months before the day of Canadian Federal elections and which ends on October 21st of 2019, we have collected a sample of tweets that were issued during this period and whose content is related to three main topics: the Federal Canadian elections, the political parties, and their leaders. We have applied an implemented random forest classifier called Botometer, which is the most known machine learning algorithm to detect bots in Twitter. Using this tool we have estimated that the percentage of accounts with a probability of being a bot greater or equal to 50% that issued content about these topics is around 5%, while the estimated percentage of accounts with a probability of being a bot greater than 74% corresponds to less than 2% of the number of sampled issuing accounts. In addition, by weekly estimating these probabilities, we have found that the probabilities estimated by Botometer that are smaller than 25% are less variable than higher probabilities.

We have characterized suspicious accounts, which are denominated by us as “likely bots”, into four groups that we have named as “Repliers”, “Popular amplifiers”, “Retweeter amplifiers” and “Not amplified retweeters” according to their characteristics through the analysis of 18 variables that describe characteristics of Twitter accounts. We have summarized the type of content that each group of accounts tends to issue through the application of the sparse hierarchical cluster proposed in [8], and we have found that three of these groups, which are the “Repliers”, the “Popular amplifiers” and the “Retweeter amplifiers”, exhibit the same behavior as the “Amplifiers” of McKelvey and Dubois. However, we have not detected groups of accounts that follow the same behavior as “Dampeners” or “Transparency bots”.

Twitter accounts classified as “Repliers” resemble some characteristics of “dampeners” like the fact that the tweet they issue are replies to other accounts, and that their content is the most liked among the four classes of accounts; however, after analyzing the content of tweets that were issued by “Repliers” we have not found negative or discouraging discourse, we have actually found that their content encourages Twitter users to vote during the election day. Regarding the content issued by “Popular amplifiers”, we have found that it mainly consists of negative discourse against Justin Trudeau, and against Andrew Scheer to a lesser extent. We have detected that the main topics in the discourse of “Retweeter amplifiers” are related to the debate. For this category, we also have found an amplified tweet that was originally issued by Jason Kenney. Moreover, we have found that most of the accounts belong to the category of “Not amplified retweeters”, and although we have not detected a general pattern to describe the content of these tweets, we have found that the presence of the Green, People’s and New Democratic parties in this set of tweets is more notable.

Through this analysis, we have found that bots exhibit different complexity levels in their operating mechanism. For example, accounts belonging to “Repliers” are mostly verified and issue their own content, which might correspond to a more complex mechanism than the used by accounts of type “Retweeter amplifiers”, which do not emit their own created content. Additionally, “Retweeter amplifiers” appear to have a more complex mechanism than the “Not amplified retweeters” because the former tends to retweet content that has been commonly retweeted, while latter does not. This suggests that the “Not amplified retweeters” are inefficient bots regarding the task of amplifying their content.