Measuring the Influence of Twitter Bots during the 2018 US Midterm Election

Abstract

This study uses 6.48M tweets and 1.6M profile data collected from Twitter using 109 keywords during the 2018 US Midterm Elections to determine the influence bots had. Data was collected between October 31, 2018 (1 AM UTC) and November 7, 2018 (12 AM UTC). Machine Learning techniques were used on this data to separate bots and humans. Afterward, the analysis found that 1% of the top bots influenced the creation of 11% of the tweets. Additionally, for URLs which occurred more than ten times in our dataset, the total influence percentage of bots was double their population percentage. Here, influence is a number given to each tweet which measures "the expected number of time the tweet is retweeted – direct retweets or descendants in a diffusion scenario – over all possible diffusion scenarios associated." (Rizoiu et al. 2018) By this definition, a tweet T_2 is said to be influenced by a tweet T_1 if T_2 is a retweet of T_1. A tweet can be influenced by multiple tweets. In the previous example, if T_3 retweets the retweet made by T_2, it is influenced by both T_1 and T_2. Influence percentage is calculated from the percentage of influence a user had calculated on aggregate from the complete dataset.

1 Introduction

Twitter is a social media where short messages get exchanged. Messages shared by a user are shown to their followers and to followers of those who share the message. These messages are called Tweets. Sharing can be done by retweeting (sharing) or replying to the tweet. When Alice replies/retweets, a tweet made by Bob, Alice and Bob's followers will see the tweet. In this way, messages spread. Having massive followers, getting many retweets or getting retweeted by people with massive followers can make a tweet viral.

People use social media, at least in part, to form an opinion about lifestyle, health, politics, and purchase (Varol et al. 2017). Cost of advertising on Twitter can reflect the value of a Tweet. In 2018 Q4 the median Cost Per Mile (CPM) - the cost for reaching 1000 people was - \$5.93 while the median Cost Per Click (CPC) was \$0.4. The median Click Through Rate (CTR) was 1.55% (Adstage 2019). Due to the associated cost and possible underlying problems, advertising is not always feasible. Various other techniques have been devised and are also combined with advertisements to reach a big audience. One such technique is the use of bots. Social bots are account controlled wholly or partially by a computer algorithm. These bots can generate content and interact with human users often imitating humans (Ferrara et al. 2016). A US special counsel investigation found that Internet Research Agency (IRA), a St Petersberg based company, created Twitter accounts of US personas and operated a botnet (a network of bots) to amplify content. During the US Presidential Election, they explicitly supported or opposed political candidate (Luceri et al. 2019; Mueller 2019). These propaganda bots were found to be high volume, multichannel, rapid, continuous, repetitive, inaccurate, and inconsistent (Paul and Matthews 2016).

Studies have explored political polarization (Munger 2017; Rizoiu et al. 2018; Gruzd and Roy 2014; Bovet and Makse 2019) , origin of bots (Zannettou et al. 2019b; 2019a) and the spread of fake news (Vosoughi, Roy, and Aral 2018; Shao et al. 2018; Lazer et al. 2018; Bovet and Makse 2019; Grinberg et al. 2019) on twitter. Most studies focus on 2016 Election (Bovet and Makse 2019; Rizoiu et al. 2018; Bessi and Ferrara 2016; Howard, Woolley, and Calo 2018; Howard, Kollanyi, and Woolley 2016). One studied the 2018 Midterm (Deb et al. 2019). Few studies perform a large scale quantitative study of the role played by bots during a significant event. Using ML techniques to detect bots and measure their success/failure during a significant event, this study attempts to add to the small body of literature.

1.1 Summary and Findings

- 10.5% of the users in the datasets were bots. They made 12.6% of the tweets and created 22% of the total influence.
- On Average, bots were twice as influential as humans
- Top 1% of the bots influenced the creation of 11% of retweets in the dataset
- In average a single URL was shared 27.84 times in 1.97 cascades
- The influence percentage of bots in URL shared was similar to their population. However, for URLs which were shared more than ten times, their influence percentage was twice their population.

- Bots dominate humans in the use of conservative words. Humans dominate bots in the use of slangs like yall, bc, and bunch.
- Tweets made by bots was retweeted by bots 2.87% or 120k times. Tweets made by bots were retweeted by humans 6.34% or 265k times. Tweets made by humans were retweeted by bots 9.14% or 382k times. Finally, most of the interaction was between humans and humans i.e., 81.64% or 3.4M times.
- Most of the codes and notebooks are shared publicly in https://github.com/warproxxx/2018Midterm to help future researchers and to ensures transparency.

*

2 Related Work

2.1 Social Media and Bots

People are spending more and more time on social media. Globally, people spend 2 hours and 23 minutes on average in Social Media. 40% people use social media to stay up to date with news and current events (GlobalWebIndex 2019). 68% of American adults get their news from social media while 42% find news on social media to be mostly accurate (Center 2018).

Twitter is not the most popular social media. 12% of US adults get news from Twitter compared to Facebook's 43% (Center 2018). However, Twitter is the number one platform for government leaders. 97% of all UN members have an official presence in Twitter (Twiplomacy 2018). Furthermore, Twitter API provides easy access to data. Consequently, Twitter has been widely studied (Rizoiu et al. 2018; Grinberg et al. 2019; Bovet and Makse 2019; Morstatter et al. 2018; Munger 2017; Gruzd and Roy 2014; Zannettou et al. 2019a; Howard, Kollanyi, and Woolley 2016).

Social Media have been used for positives like democratizing online discussion, organizing civil movements (González-Bailón, Borge-Holthoefer, and Moreno 2013), augumenting public health (Dredze 2012), forecasting (Asur and Huberman 2010; Nguyen, Shirai, and Velcin 2015; Liu et al. 2016), forming social connections (Ellison, Steinfield, and Lampe 2007), and for other greater goods (Moorhead et al. 2013; Househ, Borycki, and Kushniruk 2014). However, recently, more focus is being put on the negatives with high focus on disinformation and bots (Forelle et al. 2015; Bradshaw and Howard 2017; Marwick and Lewis 2017).

Different types of bots exist in social media. A simple bot will post predetermined messages at predetermined intervals (Haustein et al. 2016). Bots are also used to increase followers and make an account appear popular (Cresci et al. 2015). Botnets are a network of bots working together. The word Botnet is also used to refers to a network of compromised computer. But, in this report, the word botnet refers to a network of social media bots working together to make an influence and should not be confused with the other definition. Many botnets use a hybrid human/automation approach (Grimme, Assenmacher, and Adam 2018). Botnets have been used to promote spam (Ferrara 2018) (which seem to shifted to social media due to the effectiveness of spam filter (Gao et al. 2010; Chu, Widjaja, and Wang 2012; Ferrara 2018)), manipulate stocks (Ferrara 2015), manipulate elections (Morstatter et al. 2018), and numerous other purposes (Abokhodair, Yoo, and McDonald 2015). Botnets were detected as early as 2010 US Midterm Election (Mitter, Wagner, and Strohmaier 2014). Over time, their use and study have increased. Studies have detected and analyzed bots in the 2017 German Federal Election (Morstatter et al. 2018) ,2017 French Presidential Election (Ferrara 2017) and during various US Elections (Mitter, Wagner, and Strohmaier 2014; Bovet and Makse 2019; Rizoiu et al. 2018; Bessi and Ferrara 2016; Howard, Woolley, and Calo 2018; Howard, Kollanyi, and Woolley 2016; Deb et al. 2019). In Elections, they have been used to support candidates (Luceri et al. 2019), attack people (Mueller 2019) and spread fake news (Vosoughi, Roy, and Aral 2018; Grinberg et al. 2019). Political bots have been used beyond elections. (Stewart, Arif, and Starbird 2018) found that Russian bots infiltrated both right and left-leaning communities and spread different narratives (Mueller 2019).

In Twitter, bots share content in multiple channel (Paul and Matthews 2016). This activity is in line with literature which shows that information from multiple sources appears more trustworthy (Harkins and Petty 1981). Bots attack people. Attacking trustworthiness has been shown to diminish the credibility of the Original Poster (Pornpitakpan 2004). Bots are continuous, repetitive, but most of the times, they are also far from reality (Paul and Matthews 2016). The high volume seems to ensure early exposure and (Petty et al. 1994) shows that early first impression is more likely to be accepted by the brain. People have confirmation bias reinforcing their views. However, low credibility sources can have persuasive power even to unbiased or neutral people. When remembering, information gets disassociated from the source in a phenomenon known as sleeper effect (Underwood and Pezdek 1998; Paul and Matthews 2016). As a boost to sleeper effect, people are 31% more likely to remember what they see on twitter, compared to the normal web (Park 2018).

2.2 Bot Detection

Botometer, the only openly accessible bot detection system used in academia (Davis et al. 2016; Yang et al. 2019), has been widely used (Rizoiu et al. 2018; Yang et al. 2019; Shao et al. 2018). Botometer's Machine Learning algorithm provides an account-level bot classification. Botometer classifies account partially and completely automatized as bots. It has some false positives , and most notably, its detection of organizational accounts as bots has been highly criticized (Varol et al. 2017) (Botometer 2019). However, even with its false positive, Botometer has no transparent competitor and remains the most accurate bot detection system in academia.

Botometer uses more than a thousand features created from temporal activity, network structure, content analysis, sentiment analysis, and user profile data to determine a bot score (Davis et al. 2016; Yang et al. 2019). (Bessi and Ferrara 2016) illustrated that profile customization, geographical metadata, and activity statistics provided the strongest signals for bot detection. (Ferrara 2017) and (Kudugunta and Ferrara 2018) used various machine learning techniques on these features. (Ferrara 2017) obtained 93% accuracy with an AUC-ROC score of 92% in their best model using Random Forest classifier. (Kudugunta and Ferrara 2018) used 3,000 labeled examples to train a system with an AUC greater than 99% using Adaboost Classifier with Over Sampling and Undersampling of data using the SMOTENN algorithm. (Kudugunta and Ferrara 2018) also presented a tweet level classification using contextual LSTM classifier with tweet metadata and deep learning.

Most studies show that 10-20% in social media are bots. In earlier studies, the number was on the higher side. 18% of users in (Bovet and Makse 2019)'s study were bots. A more recent study by (Deb et al. 2019) states that their number may have dropped.

2.3 Influence Detection

(Zannettou et al. 2019b) uses Hawkes Process to determine the influence Iranian and Russian bots had on pushing URLs in 4 social media platforms. They found that Russian trolls were extraordinarily influential and efficient in spreading URLs. (Shao et al. 2018) found that bots amplify URLs in early moments before an article goes viral, while (Ferrara 2018) found the same for spam. Bots target users with many followers through replies and manipulation. This method was very efficient.

(Shao et al. 2018) also found that articles spread mostly through tweets and retweets and much lesser from replies. Their study showed that people do not discriminate between resources shared by humans and bots. (Shao et al. 2018) and (Grinberg et al. 2019) found super spreaders. Moreover, (Grinberg et al. 2019) found that 1% of individuals accounted for 80% of fake news exposure. (Varol et al. 2017) also found that 2% user accounts were responsible for 60% of the conversation.

(Rizoiu et al. 2018) found that social bots were 2.5 times more influential than humans. They introduce a scalable algorithm for estimating user influence in retweet cascades. For each tweet in an information cascade, it uses time of post and the number of followers to determine the probability of whether a tweet is a retweet of another. They then find influence which can be used to find out how many users the tweet possibly influenced to retweet. It was tested successfully in an artificial social media. Other studies have attempted to measure influence before this. (Weng et al. 2010) used eigenvector centrality of the connection to measure influence, but it is not scalable into big cascades. There are other methods like (Rodriguez, Balduzzi, and Schölkopf 2011; Cho et al. 2013; Linderman and Adams 2014) but they either have scaling issue or require full diffusion graph, which Twitter does not provide.

3 Methods

3.1 Data Collection

Twenty-nine manually identified keywords, and names of all 53 House candidates, and 27 Senate candidates were used to collect live data from Twitter. The manually defined keywords were:

• 2018Midterms	• resist
• 2018MidtermElections	• VOTE
• Election2018	• GAEarlyVoting
• ElectionDay	• EarlyVoting
• MAGA2018	• plus1
• MAGA	• IVoted
• Trump2020	• WinBlue
• AmericaFirst	• WinRed
• TheResistance	• BlueWave
• WalkAwayFromRepublicans2018	• RedWave
• VoteThemOut	• republican2018
• 2018Senate	• democrat2018
• VoteRed	• Republican
• Voteblue	• Democrat
• WalkAwayFromDemocrats2018	

Table 1: Manual Keywords

Twitter provides Streaming API, Search API, and a premium Firehose API to provide access to tweets. Streaming API provides limited access to live data (termed 1% API in the early days) while Firehose API provides complete data. However, due to the associated costs, Firehose API was not an option in this study. (Morstatter et al. 2013) compares the Streaming API with Firehose API. They found that Streaming API was nearly as good as a random sample of Firehose API when the dataset was large enough. Although the 1% API was not as good as a 1% random sample from the Firehose API in all of their tests, it estimated the top hashtags correctly when the data was large enough. It recommended the creation of specific parameters and using a large sample. Due to the possible issues in Streaming API, (Bessi and Ferrara 2016) recommends using Twitter Search API.

This study uses the Streaming API with many keywords aiming to collect a large dataset. While collecting data, information about the tweeter and the tweet was collected. Collected information about a tweet is:

1. **Timestamp:** UTC Timestamp in which the post was made
2. **ID:** Post ID provided by twitter using their snowflake algorithm
3. **Text:** Text in the tweet and the parent tweet if the tweet is a reply
4. **User:** Username of the tweeter/retweeter/replier
5. **Replies:** The number of replies the tweet has received. During live collection, this value is zero.
6. **Retweets:** The number of retweets the tweet has received. During live collection, this value is zero.
7. **Likes:** The number of likes the tweet has received. During live collection, this value is zero.
8. **Reply To ID:** The ID of the parents tweet if this is retweet or a reply.
9. **Response Type:** Either Tweet or Retweet or Reply

Profile data were collected along with the tweets

1 **Username:** Username of the user who posted the tweet
2 **Location:** Binary if geolocation is enabled
3 **Is Verified:** A binary if the profile has been verified
4 **Total Tweets:** Total number of tweets created by the user
5 **Total Following:** Total accounts following
6 **Total Followers:** Total Followers
7 **Total Listed:** Number of times a user has been listed
8 **Total Status:** Total Status
9 **Total Likes:** Total likes the user has received
10 **Has Background:** Binary if an account has a background
11 **Is Protected:** Binary if an account is protected
12 **Profile Modified:** Binary if a profile has been modified

These parameters are same as the ones used by (Kudugunta and Ferrara 2018) and (Ferrara 2017). Same data was collected because the Machine Learning architecture used for bot detection is same.

These parameters were used to collect Twitter data from October 31, 2018 (1 AM UTC) to November 7, 2018 (12 AM UTC). Twitter API was used to re-scrape the details and update the number of likes, replies, and retweets of top 100 thousand tweets in the dataset on July 18, 2019.

3.2 Data Processing

First, irrelevant tweets were manually removed. VOTE was one of the keywords, and other people, most notable of which were the followers of music bands, had used it too. Keywords were created to find and remove these irrelevant data.

SentiStrength and VADER Sentiment analysis was used. SentiStrength (Thelwall et al. 2010) is used to annotate short, informal tweet like texts (Bessi and Ferrara 2016). It can capture positive and negative emotions at an accuracy of 60.6% and 72.8%. SentiStrength gives a positive and negative score between 0 and 4. The positive sentiment is subtracted from the negative like in (Bessi and Ferrara 2016) to get a whole number. VADER is a rule-based sentiment classification designed for use in social media data. VADER was found to outperform individual human raters with an F1 accuracy of 0.96 compared to 0.84 for humans (Hutto and Gilbert 2014).

3.3 Bot Detection

Detecting bots on the wild is more complicated than against a validation set (Bovet and Makse 2019; Varol et al. 2017; Ferrara et al. 2016). The collected dataset had 1.7M users. Although Botometer is publicly available, getting information for 1.7M data would be expensive. So the following steps were used to train a bot detection system:

- (Cresci et al. 2017) provided a labeled account level bot or not data with relevant features that included the ones we collected. These features were used to train a Random Forest Classifier and detect bots. (Cresci et al. 2017)'s data is old, and bots evolve daily. So further processing was required.
- The trained model was used in our dataset, which contains the same features, to select 40k highly probable bots.
- 40k highly probable bots and 32k random accounts were selected, and bot detection was done through Botometer. Botometer returns CAP score and bot score. CAP score denotes the probability of being a bot. Studies have used bot score of 0.5 to determine bots. CAP score of 0.3 is equivalent to the score 0.5 used by researchers before. (Deb et al. 2019) used a threshold of 0.3 after the introduction of new changes as mentioned in (Yang et al. 2019).
- Grid search was performed with thresholds of 0.1, 0.2, and 0.3 for humans, 0.3, 0.5, 0.7, and 0.8 for bots and Random Forest, Neural Network and SMOTENN as possible ML algorithms.
- It was determined that 0.5 was the optimal bot threshold, 0.3 the optimal human threshold and SMOTENN the optimal ML algorithm. Using this threshold, we selected 513 bots whose CAP score was higher than 0.5 , +27k clean accounts whose CAP score was smaller than 0.3 for training and testing our data. The remaining 44k values that lied between these were used as a broader test set.

We obtained an accuracy of 85% with AUC of 0.85 in the training set, 87% accuracy with an AUC of 0.8 in the smaller test set and accuracy of 85% with an AUC of 0.8 in the broader test set when compared with Botometers prediction. This model detected 10.5% of the accounts as bots. The accuracy was lesser than in (Kudugunta and Ferrara 2018) likely due to the evolution of bots or because we compared with Botometer instead of manually labeling like in the study.

3.4 Cascading and Influence Detection

A cascade starts when a Tweet is made. Any retweet or response which involve that Tweet belongs to that cascade. If Bob retweets Alice's tweet, and Eve retweets Bob's retweet, the tweets made by Alice, Bob, and Eve will belong in the same cascade. The Twitter API will show Eve and Bob as a direct descendant of Alice without any mention that Eve retweeted Bob's retweet.

(Du et al. 2013) defines influence as the average number of users who get in contact with the content created by a user *u*. However, the number of people reached, and the diffusion graph is not provided by Twitter. (Rizoiu et al. 2018) define the influence of a user over a retweet cascade as "the expected number of time the tweet is retweeted – direct retweets or descendants in a diffusion scenario – over all possible diffusion scenarios associated." Influence of a user is defined as the sum of the influence of tweets authored by a user. As their algorithm is used, this definition is used. (Rizoiu et al. 2018) uses time and number of followers to estimate a user influence. Using this mechanism, a high influence score can be provided to highly connected users who never start diffusions and to active retweeters with little followership. (Rizoiu et al. 2018) tested their algorithm on an artificial social network with 1000 users and found that the influence calculated had a Spearman correlation coefficient of 0.88 with the actual influence.

In (Rizoiu et al. 2018)'s algorithm, for each tweet in a cascade, the probability of it being a descendant of each previous tweet is calculated using a softmax function. Mathematically, probability that a j_{th} tweet is a retweet of i_{th} tweet is measured by :

$$p_{ij} = \frac{m_i e^{-r(t_j - t_i)}}{\sum_{k=1}^{j-1} m_k e^{-r(t_j - t_k)}}$$

where,

t_j-t_i is used as exponential decay between the timing of original tweet and that of the retweet,

r is a hyperparmeter which they found to be 6.8×10^{-4},

m is the number of followers

Then for every tweet in a cascade, pairwise influence is calculated as:

$$m_{ij} = \begin{cases} \sum_{k=i}^{j-1} m_{ik} p_{kj}^2 & ,\mathrm{i} < \mathrm{j} \\ 1 & ,\mathrm{i} = \mathrm{j} \\ 0 & ,\mathrm{i} > \mathrm{j} \end{cases} \quad (1)$$

Then the total influence of a node is the sum of the pairwise influence score m_{ij} over all subsequent nodes. For a derivation of this, the original study (Rizoiu et al. 2018) and its citations should be referred.

(Rizoiu et al. 2018) tested the validation of their algorithm when they had access to full data. We did not. So random cascades were synthesized using the following algorithm to test the validity:

- An aggregate database D was created. In the database D, values of D_u was set to all possible usernames and $D_u{}^a$ and $D_u{}^s$ set to 0.
- For 500000 cascades, following steps was repeated:
  - Length of cascade, n was randomly selected from the all possible size of cascades we captured. Then n random values were selected from our cascade that included username, time and followers count.
  - Influence was calculated for the selected n values. 1% of values were randomly selected from the n, and influence was calculated for it too.
  - The total sample influence, $D_u{}^s$ and actual influence $D_u{}^a$ in the dataset D was updated for a user u, by adding the calculated value with the current values of $D_u{}^s$ and $D_u{}^a$.

The influence was then converted to a user wise percentage, and correlation was calculation between the actual and predicted value. A high correlation was observed. This is a duct tape algorithm and should be treated as such. It would have been better to repeat the original process as done by (Rizoiu et al. 2018) through the use of 100k synthesized cascades, but it required the creation of huge synthesized Adjacency Matrix. Due to limitations in computing, this was not possible.

4 Data Analysis

4.1 Data Captured

After cascading and re-scraping, the actual size of the top cascades on Twitter was compared with the size we had. This comparison has issues. There is a selection bias in this data as only the top tweets were re-scraped. Additionally, we re-scraped metrics from twitter in July 2019, nine months after the collection. Tweets may have been removed, and new retweets may have been made. However, this analysis is still included for future comparisons.

In our 329606 cascades, we had re-scraped the details of 154603 parent tweets. We compared the total retweet and replies we had (when re-scraped) with the total retweets and replies on Twitter. After summing the total tweet and retweet amount in Twitter and comparing it with the sum we had managed to captured, it was found that we had captured 9.64% of data in each cascade.

The biases explained above, and the use of many keywords may be possible reasons for this high capture.

4.2 Exploratory Data Analysis

After the removal of irrelevant data, 6.5M Tweets by 1.7M users remained. Bots made 12.6% of the Tweets. We removed stopwords to compare the words used by bots and humans. We selected words which appear more than 0.03% times in the dataset (which was 45759 times for humans and 6538 for bots). Wordcloud was created to visualize the differences.

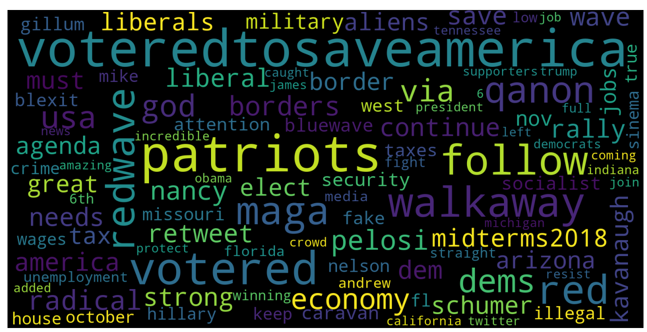


Figure 1: Word used more more often by bots

From Table 2 and Figure 1, we can see that bots dominate humans in the use of conservative words.

Table 3 and Figure 2 shows that humans dominate bots in the use of slang.

In Figure 3, a histogram was created to visualize the difference in Sentiment difference between Human and Bots. The figures are very similar and do not show any significant differences.

The 329606 cascades in the dataset was analyzed to calculate human-bot interaction. The following results was observed:

- Tweets made by bots was retweeted by bots 2.87% or 120k times.
- Tweets made by bots were retweeted by humans 6.34% or 265k times.

Word	Percentage Difference
voteredtosaveamerica	169.715447
patriots	161.121157
votered	130.514988
walkaway	107.832423
follow	106.300268
red	99.139168
maga	95.622477
redwave	95.294118
qanon	86.301370
dems	65.538736
usa	60.909091
via	57.861635
god	56.432247
economy	51.132075

Table 2: Words used more often by bots

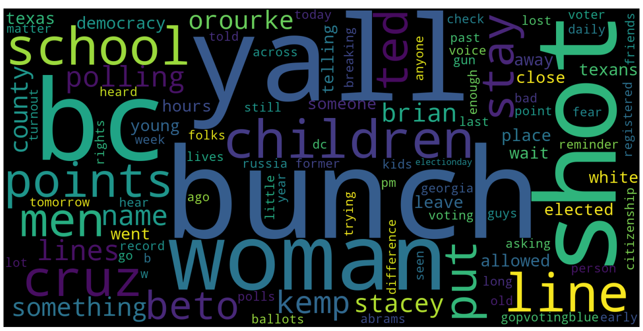


Figure 2: Word used more more often by humans

Word	Percentage Difference
bunch	58.440047
yall	52.095130
bc	45.844156
shot	45.454545
woman	43.218954
children	41.602787
line	41.316348
school	41.313559
cruz	39.701074
points	39.238095
men	39.154161
put	35.545906
stay	35.414725
ted	35.111111

Table 3: Words used more often by humans

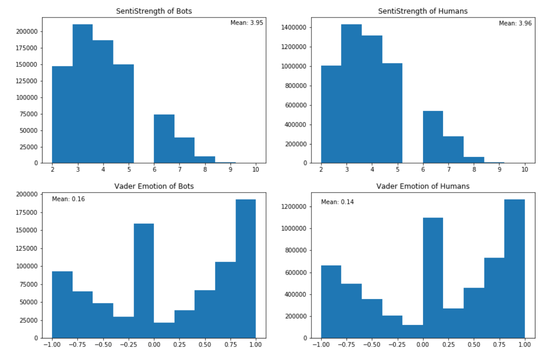


Figure 3: Histogram of Tweet Sentiment

- Tweets made by humans were retweeted by bots 9.14% or 382k times.
- Tweets made by humans were retweeted by humans 81.64% or 3.4M times.

Next, cascades started by bots, and humans were compared with each other. The cascades were grouped by their length to properly visualize the differences. In Figure 4a we can see that percentage amount of bots was high for smaller bot started cascades. It becomes smaller as the size increases. Opposite was observed for humans in Figure 4b. It might possibly mean that humans and mostly followed by humans and bots by bots. Both join once a tweet becomes famous.

In Figure 6 we look at the successful tweets for the percentage intervals. From the figure, we can see that the number of bots is high in first one percentage for successful tweets in case of bots and is high for humans in a decreasing amount. This analysis provides further support to the assertion made in previous paragraph.

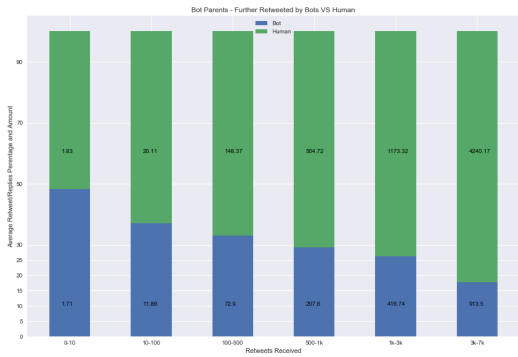
4.3 Retweet Influence Analysis

The average influence of a human was 4.06, half that of a bot, which was 8.12. This finding is similar to (Rizoiu et al. 2018)'s finding who found that bots were 2.5 times more influential than humans. Then influence for each user was converted into a percentage to determine how influential a user was.

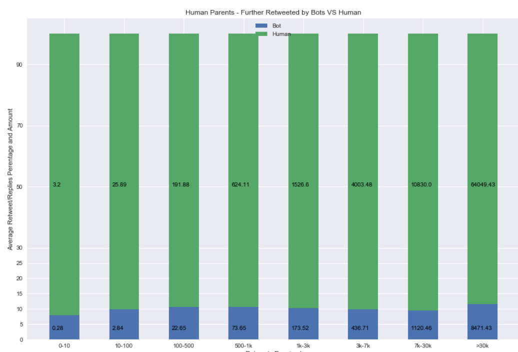
Figure 6 unsurprisingly shows that the most influential user, who was possibly responsible for the creation of as much as 0.93% of the tweets was the POTUS, @realDonaldTrump.

Then, we find the difference in influence made by the top 1% of all users, top 1% of the bots, and top 1% of the Humans.

Figure 7, shows that the top 1% of bots were responsible for 75% of influence made by bots. Top 1% of humans were responsible for 65% of the influence. The top 20% of bots are responsible for nearly 98% of the influence. The top 20% of humans are responsible for 64% of total influence.



(a) Bots Parents



(b) Humans Parents

Figure 4: Further Retweet

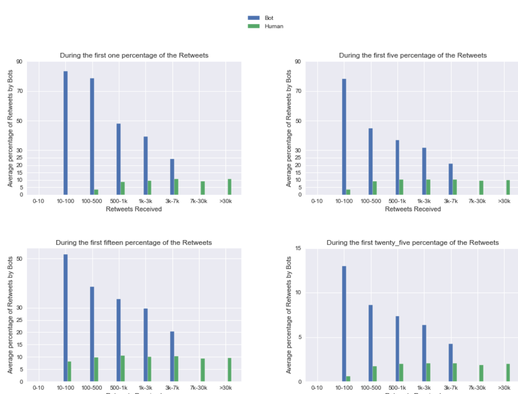


Figure 5: At Different Percentage Intervals

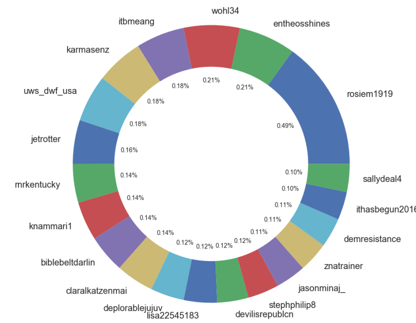
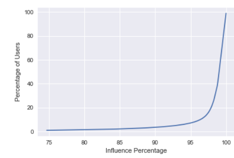
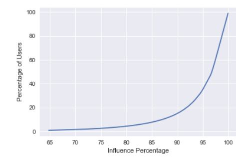


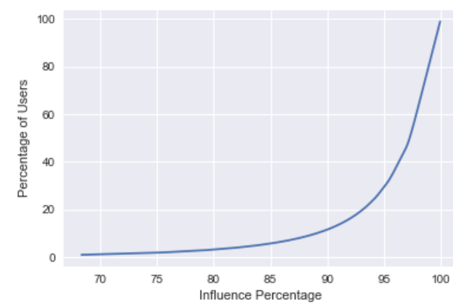
Figure 6: Most influential users



(a) Bots



(b) Humans



(c) All

Figure 7: Percentage of top users and their influence percentage

1% of the bots are responsible for 11% of the total influence. Out of 4.8M data in our cascade, this can mean that bots were responsible for the creation of 500k human tweets. However, this evaluation cannot be accurately made without further research. When calculating influence, we could calculate the influence of a bot or a human in the diffusion graph multiple times.

In Figure 8 the influence score of node 1 will be 4, 2 will be 2, 3 will be 1, 4 and 5 will be 0. If 1 and 2 are bots, the total bot influence score will be 6. But 1 created 2. Although this might cancel out over a large data when taken in percentage, we still refrain from making that conclusion without further tests.

4.4 URL Influence Analysis

Before performing URL influence calculation, short URLs were resolved, and twitter.com links were removed. Some URL could not be resolved (possibly due to removal) and they were kept as they are, mostly in t.co domain. 652,054 unique URLs belonging to 157,192 unique domains were

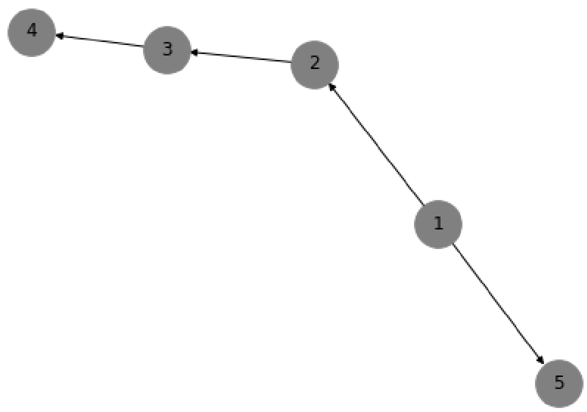


Figure 8: Diffusion Graph

Domain Name	Frequency
t.co	127717
vote.gop	96127
thehill.com	57476
cnn.com	54129
youtube.com	51347
wtxl.com	38124
nytimes.com	35863
foxnews.com	35277
instagram.com	31634
breitbart.com	30290
pscp.tv	27702

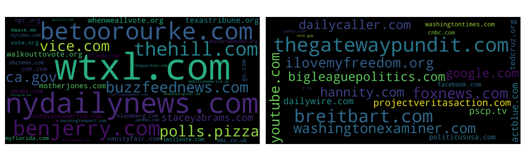
Table 4: Most Common URLs

shared in our dataset by 652,054 unique users 1.9 Million times. Ten most popular domains in our dataset with their frequency are in Table 4

The top 10 domains for humans and bots, apart from t.co were:

Wordcloud was then created to find human dominant and bot dominant domains. URL which appeared at least 0.03% of times which is 493 times for humans and 77 times for bots were used as before.

Various cascades could share the same URL. On average, a URL was shared in 1.97 different cascades. A single URL was shared 27.84 times in 1.97 cascades on average. Next, we compare the average influence bots had on the URL cascades with the number of bots. On first look, the influence looked similar to the percentage. Their correlation was 0.93



(a) Human Dominant (b) Bots Dominant

Figure 9: Dominant Domains

Human	Bots
vote.gop	vote.gop
thehill.com	youtube.com
cnn.com	foxnews.com
youtube.com	breitbart.com
wtxl.com	thegatewaypundit.com
nytimes.com	pscp.tv
foxnews.com	cnn.com
instagram.com	instagram.com
breitbart.com	thehill.com
pscp.tv	facebook.com

Table 5: Most Common Human and Bots URL

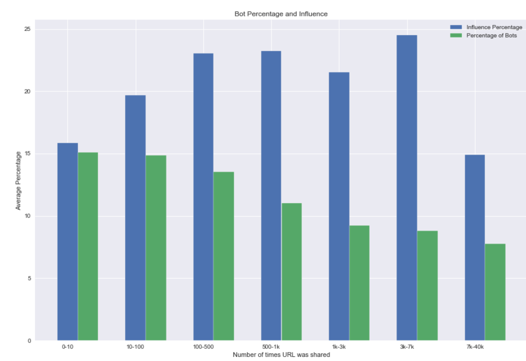


Figure 10: Percentage of Bots and Percentage of their influence

(pval=0) and the total percentage of bot influence was 1.09 times the total bots percentage. However, on further evaluation average was found to deceptive.

The average appears similar because the influence percentage is higher for unsuccessful URLs (i.e., when they are shared less than ten times). But we can see in Figure 10 that influence bots had was much more than their amount in bigger cascades.

This is better visualized in Table 6. The ratio is increasing as the size increases except for the last value. The difference in the last interval might be due to small sample size.

Interval	Size	Average Influence	Average Size	Ratio
0-10	98709	15.83	15.1	1.05
10-100	15309	19.67	14.84	1.33
100-500	1879	23.02	13.53	1.7
500-1k	209	23.24	11.02	2.11
1k-3k	122	21.53	9.23	2.33
3k-7k	25	24.49	8.82	2.78
7k-40k	5	14.92	7.77	1.92

Table 6: Percentage of Bots and their Influence

5 Limitations

- The analysis was performed on a subset of data
- Validation of influence detection algorithm in a subset was not done in a proper academic way
- The accuracy of the bot detection algorithm is not manually tested.

6 Further Work

Various directions are possible to go in future:

- Analysis was done by considering all bots as a single entity. Although easy, this distinction has many flaws. In futures, attempts should be made to divide different types of bots and study their agenda
- As Twitter attracts some demographics more than others, using random Twitter data is not representative of the nation or the voting population. Careful clustering can be done to create a representative data to get representative influence in the future.
- The validation of the influence detection algorithm can be done in a better way. Moreover, Machine Learning techniques can be used to make a better influence detection algorithm to predict the diffusion edges. Honeypot accounts can be used to aid in this purpose.
- Twitter Firehose API or Search API can be used in the future to have complete access to the data. Apart from Firehose and Search API, a mechanism for finding Twitter posts using Twitter's snowflake algorithm looks promising (Baumgartner 2019)
- Reddit can be analyzed along with Twitter to create a better spread graph. A bot detection system can be made for Reddit too. Combining multiple sources, a complicated long term study can be conducted to measure the success/failure of bots in changing people's opinion.

References

Abokhodair, N.; Yoo, D.; and McDonald, D. W. 2015. Dissecting a social botnet: Growth, content and influence in twitter. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 839–851. ACM.

Adstage. 2019. Q4 2018 paid media benchmark report.

Asur, S., and Huberman, B. A. 2010. Predicting the future with social media. In *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology-Volume 01*, 492–499. IEEE Computer Society.

Baumgartner, J. 2019. Reconstructing twitter's firehose.

Bessi, A., and Ferrara, E. 2016. Social bots distort the 2016 us presidential election online discussion. *First Monday* 21(11-7).

Botometer. 2019.

Bovet, A., and Makse, H. A. 2019. Influence of fake news in twitter during the 2016 us presidential election. *Nature communications* 10(1):7.

Bradshaw, S., and Howard, P. 2017. Troops, trolls and troublemakers: A global inventory of organized social media manipulation.

Center, P. R. 2018. News use across social media platforms 2018.

Cho, Y.-S.; Galstyan, A.; Brantingham, P. J.; and Tita, G. 2013. Latent self-exciting point process model for spatial-temporal networks. *arXiv preprint arXiv:1302.2671*.

Chu, Z.; Widjaja, I.; and Wang, H. 2012. Detecting social spam campaigns on twitter. In *International Conference on Applied Cryptography and Network Security*, 455–472. Springer.

Cresci, S.; Di Pietro, R.; Petrocchi, M.; Spognardi, A.; and Tesconi, M. 2015. Fame for sale: Efficient detection of fake twitter followers. *Decision Support Systems* 80:56–71.

Cresci, S.; Di Pietro, R.; Petrocchi, M.; Spognardi, A.; and Tesconi, M. 2017. The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race. In *Proceedings of the 26th international conference on World Wide Web companion*, 963–972. International World Wide Web Conferences Steering Committee.

Davis, C. A.; Varol, O.; Ferrara, E.; Flammini, A.; and Menczer, F. 2016. Botornot: A system to evaluate social bots. In *Proceedings of the 25th International Conference Companion on World Wide Web*, 273–274. International World Wide Web Conferences Steering Committee.

Deb, A.; Luceri, L.; Badaway, A.; and Ferrara, E. 2019. Perils and challenges of social media and election manipulation analysis: The 2018 us midterms. In *Companion Proceedings of The 2019 World Wide Web Conference*, 237–247. ACM.

Dredze, M. 2012. How social media will change public health. *IEEE Intelligent Systems* 27(4):81–84.

Du, N.; Song, L.; Rodriguez, M. G.; and Zha, H. 2013. Scalable influence estimation in continuous-time diffusion networks. In *Advances in neural information processing systems*, 3147–3155.

Ellison, N. B.; Steinfield, C.; and Lampe, C. 2007. The benefits of facebook "friends:" social capital and college students' use of online social network sites. *Journal of computer-mediated communication* 12(4):1143–1168.

Ferrara, E.; Varol, O.; Davis, C.; Menczer, F.; and Flammini, A. 2016. The rise of social bots. *Communications of the ACM* 59(7):96–104.

Ferrara, E. 2015. Manipulation and abuse on social media by emilio ferrara with ching-man au yeung as coordinator. *ACM SIGWEB Newsletter* (Spring):4.

Ferrara, E. 2017. Disinformation and social bot operations in the run up to the 2017 french presidential election. *First Monday* 22(8).

Ferrara, E. 2018. Measuring social spam and the effect of bots on information diffusion in social media. In *Complex Spreading Phenomena in Social Systems*. Springer. 229–255.

Forelle, M.; Howard, P.; Monroy-Hernández, A.; and Savage, S. 2015. Political bots and the manipulation of public opinion in venezuela. *arXiv preprint arXiv:1507.07109*.

Gao, H.; Hu, J.; Wilson, C.; Li, Z.; Chen, Y.; and Zhao, B. Y. 2010. Detecting and characterizing social spam campaigns. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, 35–47. ACM.

GlobalWebIndex. 2019. 2019 q1 social flagship report.

González-Bailón, S.; Borge-Holthoefer, J.; and Moreno, Y. 2013. Broadcasters and hidden influentials in online protest diffusion. *American Behavioral Scientist* 57(7):943–965.

Grimme, C.; Assenmacher, D.; and Adam, L. 2018. Changing perspectives: Is it sufficient to detect social bots? In *International Conference on Social Computing and Social Media*, 445–461. Springer.

Grinberg, N.; Joseph, K.; Friedland, L.; Swire-Thompson, B.; and Lazer, D. 2019. Fake news on twitter during the 2016 us presidential election. *Science* 363(6425):374–378.

Gruzd, A., and Roy, J. 2014. Investigating political polarization on twitter: A canadian perspective. *Policy & Internet* 6(1):28–45.

Harkins, S. G., and Petty, R. E. 1981. The multiple source effect in persuasion: The effects of distraction. *Personality and Social Psychology Bulletin* 7(4):627–635.

Haustein, S.; Bowman, T. D.; Holmberg, K.; Tsou, A.; Sugimoto, C. R.; and Larivière, V. 2016. Tweets as impact indicators: Examining the implications of automated "bot" accounts on t witter. *Journal of the Association for Information Science and Technology* 67(1):232–238.

Househ, M.; Borycki, E.; and Kushniruk, A. 2014. Empowering patients through social media: the benefits and challenges. *Health informatics journal* 20(1):50–58.

Howard, P. N.; Kollanyi, B.; and Woolley, S. 2016. Bots and automation over twitter during the us election. *Computational Propaganda Project: Working Paper Series*.

Howard, P. N.; Woolley, S.; and Calo, R. 2018. Algorithms, bots, and political communication in the us 2016 election: The challenge of automated political communication for election law and administration. *Journal of information technology & politics* 15(2):81–93.

Hutto, C. J., and Gilbert, E. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth international AAAI conference on weblogs and social media*.

Kudugunta, S., and Ferrara, E. 2018. Deep neural networks for bot detection. *Information Sciences* 467:312–322.

Lazer, D. M.; Baum, M. A.; Benkler, Y.; Berinsky, A. J.; Greenhill, K. M.; Menczer, F.; Metzger, M. J.; Nyhan, B.; Pennycook, G.; Rothschild, D.; et al. 2018. The science of fake news. *Science* 359(6380):1094–1096.

Linderman, S., and Adams, R. 2014. Discovering latent network structure in point process data. In *International Conference on Machine Learning*, 1413–1421.

Liu, T.; Ding, X.; Chen, Y.; Chen, H.; and Guo, M. 2016. Predicting movie box-office revenues by exploiting large-scale social media content. *Multimedia Tools and Applications* 75(3):1509–1528.

Luceri, L.; Deb, A.; Badawy, A.; and Ferrara, E. 2019. Red bots do it better: Comparative analysis of social bot partisan behavior. In *Companion Proceedings of The 2019 World Wide Web Conference*, 1007–1012. ACM.

Marwick, A., and Lewis, R. 2017. Media manipulation and disinformation online. *New York: Data & Society Research Institute*.

Mitter, S.; Wagner, C.; and Strohmaier, M. 2014. A categorization scheme for socialbot attacks in online social networks. *arXiv preprint arXiv:1402.6288*.

Moorhead, S. A.; Hazlett, D. E.; Harrison, L.; Carroll, J. K.; Irwin, A.; and Hoving, C. 2013. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of medical Internet research* 15(4):e85.

Morstatter, F.; Pfeffer, J.; Liu, H.; and Carley, K. M. 2013. Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. In *Seventh international AAAI conference on weblogs and social media*.

Morstatter, F.; Shao, Y.; Galstyan, A.; and Karunasekera, S. 2018. From alt-right to alt-rechts: Twitter analysis of the 2017 german federal election. In *Companion Proceedings of the The Web Conference 2018*, 621–628. International World Wide Web Conferences Steering Committee.

Mueller, R. S. 2019. Report on the investigation into russian interference in the 2016 presidential election.

Munger, K. 2017. Don't@ me: Experimentally reducing partisan incivility on twitter. *Unpublished manuscript*.

Nguyen, T. H.; Shirai, K.; and Velcin, J. 2015. Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications* 42(24):9603–9611.

Park, M. 2018. Expanding in-stream video ads to more advertisers globally.

Paul, C., and Matthews, M. 2016. The russian "firehose of falsehood" propaganda model. *Rand Corporation* 2–7.

Petty, R.; Cacioppo, J.; Strathman, A.; and Priester, J. 1994. To think or not to think? exploring two routes to persuasion in: Shavitt s, brock tc, editors. persuasion: Psychological insights and perspectives.

Pornpitakpan, C. 2004. The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of applied social psychology* 34(2):243–281.

Rizoiu, M.-A.; Graham, T.; Zhang, R.; Zhang, Y.; Ackland, R.; and Xie, L. 2018. # debatenight: The role and influence of socialbots on twitter during the 1st 2016 us presidential debate. In *Twelfth International AAAI Conference on Web and Social Media*.

Rodriguez, M. G.; Balduzzi, D.; and Schölkopf, B. 2011. Uncovering the temporal dynamics of diffusion networks. *arXiv preprint arXiv:1105.0697*.

Shao, C.; Ciampaglia, G. L.; Varol, O.; Yang, K.-C.; Flammini, A.; and Menczer, F. 2018. The spread of low-credibility content by social bots. *Nature communications* 9(1):4787.

Stewart, L. G.; Arif, A.; and Starbird, K. 2018. Examining trolls and polarization with a retweet network. In *Proc. ACM WSDM, Workshop on Misinformation and Misbehavior Mining on the Web*.

Thelwall, M.; Buckley, K.; Paltoglou, G.; Cai, D.; and Kappas, A. 2010. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology* 61(12):2544–2558.

Twiplomacy. 2018. Twiplomacy study 2018.

Underwood, J., and Pezdek, K. 1998. Memory suggestibility as an example of the sleeper effect. *Psychonomic Bulletin & Review* 5(3):449–453.

Varol, O.; Ferrara, E.; Menczer, F.; and Flammini, A. 2017. Early detection of promoted campaigns on social media. *EPJ Data Science* 6(1):13.

Vosoughi, S.; Roy, D.; and Aral, S. 2018. The spread of true and false news online. *Science* 359(6380):1146–1151.

Weng, J.; Lim, E.-P.; Jiang, J.; and He, Q. 2010. Twitterrank: finding topic-sensitive influential twitterers. In *Proceedings of the third ACM international conference on Web search and data mining*, 261–270. ACM.

Yang, K.-C.; Varol, O.; Davis, C. A.; Ferrara, E.; Flammini, A.; and Menczer, F. 2019. Arming the public with artificial intelligence to counter social bots. *Human Behavior and Emerging Technologies* 1(1):48–61.

Zannettou, S.; Bradlyn, B.; De Cristofaro, E.; Stringhini, G.; and Blackburn, J. 2019a. Characterizing the use of images by state-sponsored troll accounts on twitter. *arXiv preprint arXiv:1901.05997*.

Zannettou, S.; Caulfield, T.; Setzer, W.; Sirivianos, M.; Stringhini, G.; and Blackburn, J. 2019b. Who let the trolls out?: Towards understanding state-sponsored trolls. In *Proceedings of the 10th ACM Conference on Web Science*, 353–362. ACM.