Red Bots Do It Better: Comparative Analysis of Social Bot Partisan Behavior

Luca Luceri*  
University of Applied Sciences and Arts of Southern Switzerland, and University of Bern  
Manno, Switzerland  
lucaluceri@supsi.ch

Ashok Deb  
USC Information Sciences Institute  
Marina del Rey, CA  
ashok@isi.edu

Adam Badawy  
USC Information Sciences Institute  
Marina del Rey, CA  
badawy@isi.edu

Emilio Ferrara  
USC Information Sciences Institute  
Marina del Rey, CA  
emiliofe@usc.edu

ABSTRACT
Recent research brought awareness of the issue of bots on social media and the significant risks of mass manipulation of public opinion in the context of political discussion. In this work, we leverage Twitter to study the discourse during the 2018 US midterm elections and analyze social bot activity and interactions with humans. We collected 2.6 million tweets for 42 days around the election day from nearly 1 million users. We use the collected tweets to answer three research questions: (i) Do social bots lean and behave according to a political ideology? (ii) Can we observe different strategies among liberal and conservative bots? (iii) How effective are bot strategies?

We show that social bots can be accurately classified according to their political leaning and behave accordingly. Conservative bots share most of the topics of discussion with their human counterparts, while liberal bots show less overlap and a more inflammatory attitude. We studied bot interactions with humans and observed different strategies. Finally, we measured bots embeddedness in the social network and the effectiveness of their activities. Results show that conservative bots are more deeply embedded in the social network and more effective than liberal bots at exerting influence on humans.

KEYWORDS
social media, political elections, social bots, political manipulation

INTRODUCTION
During the last decade, social media have become the conventional communication channel to socialize, share opinions, and access the news. Accuracy, truthfulness, and authenticity of the shared content are necessary ingredients to maintain a healthy online discussion. However, in recent times, social media have been dealing with a considerable growth of false content and fake accounts. The resulting wave of misinformation (and disinformation) highlights the pitfalls of social media networks and their potential harms to several constituents of our society, ranging from politics to public health.

In fact, social media networks have been used for malicious purposes to a great extent [11]. Various studies raised awareness about the risk of mass manipulation of public opinion, especially in the context of political discussion. Disinformation campaigns [2, 5, 12, 14–16, 21, 23, 25, 29] and social bots [3, 4, 20, 22, 24, 28, 30, 31] have been indicated as factors contributing to social media manipulation.

The 2016 US Presidential election represents a prime example of the significant perils of mass manipulation of political discourse. Badawy et al. [1] studied the Russian interference in the election and the activity of Russian trolls on Twitter. Im et al. [17] suggested that troll accounts are still active to these days. The presence of social bots does not show any sign of decline [10, 31] despite the attempts from social network providers to suspend suspected, malicious accounts. Various research efforts have been focusing on the analysis, detection, and countermeasures development against social bots. Ferrara et al. [13] highlighted the consequences associated with bot activity in social media. The online conversation related to the 2016 US presidential election was further examined [3] to quantify the extent of social bots activity. More recently, Stella et al. [26] discussed bots’ strategy of targeting influential humans to manipulate online conversation during the Catalan referendum for independence, whereas Shao et al. [24] analyzed the role of social bots in spreading articles from low credibility sources. Deb et al. [10] focused on the 2018 US Midterms elections with the objective to find instances of voter suppression.

In this work, we investigate social bots behavior by analyzing their activity, strategy, and interactions with humans. We aim to answer the following research questions (RQs) regarding social bots behavior during the 2018 US Midterms election.

*Also with USC Information Sciences Institute.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

The Web Conference ’19, May 14, 2019, San Francisco, CA  
© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-1-4503-9999-9/18/06... $15.00  
https://doi.org/10.1145/1122445.1122456
RQ1: Do social bots lean and behave according to a political ideology? We investigate whether social bots can be classified based on their political inclination into liberal or conservative leaning. Further, we explore to what extent they act similarly to the corresponding human counterparts.

RQ2: Can we observe different strategies among liberal and conservative bots? We examine the differences between social bot strategies to mimic humans and infiltrate political discussion. For this purpose, we measure bot activity in terms of volume and frequency of posts, interactions with humans, and embeddedness in the social network.

RQ3: Are bot strategies effective? We introduce four metrics to estimate the effectiveness of bot strategies and to evaluate the degree of human interplay with social bots.

We leverage Twitter to capture the political discourse during the 2018 US midterm elections. We collected 2.6 million tweets for 42 days around election day from nearly 1 million users. We then explore collected data and attain the following findings:

- We show that social bots are embedded in each political side and behave accordingly. Conservative bots abide by the topic discussed by the human counterpart more than liberal bots, which in turn exhibit a more provocative attitude.
- We examined bots’ interactions with humans and observed different strategies. Conservative bots stand in a more central social network position, and divide their interactions between humans and other conservative bots, whereas liberal bots focused mainly on the interplay with the human counterparts.
- We measured the effectiveness of these strategies and recognized the strategy of conservative bots as the most effective in terms of influence exerted on human users.

DATA

In this study, we use Twitter to investigate the partisan behavior of malicious accounts during the 2018 US midterm elections. For this purpose, we carried out a data collection from the month prior (October 6, 2018) to two weeks after (November 19, 2018) the day of the election. We kept the collection running after the election day as several races remained unresolved. We employed the Python module Twyton to collect tweets through the Twitter Streaming API using the following keywords as a filter: 2018midtermelections, 2018midterms, elections, midterm, and midtermelections. As a result, we gathered 2.7 million tweets, whose IDs are publicly available for download. From this set, we first removed any duplicate tweet, which may have been captured by accidental redundant queries to the Twitter API. Then, we excluded all the tweets not written in English language. Despite the majority of the tweets were in English, and to a lesser degree in Spanish (3,177 tweets), we identified 59 languages in the collected data. Thus, we inspected tweets from other countries and removed them as they were out of the context of this study. In particular, we filtered out tweets related to the Cameron election, the Democratic Republic of the Congo election, the Biafra call for Independence, democracy in Kenya (#democracyKE), to the two major political parties in India (BJP and UPA), and college midterm exams. Overall, we retain nearly 2.6 millions tweets, whose aggregate statistics are reported in Table 1.

| Statistic | Count |
|-----------|-------|
| # of Tweets | 452,288 |
| # of Retweets | 1,869,313 |
| # of Replies | 267,973 |
| # of Users | 997,406 |

METHODOLOGY

Bot Detection

Nowadays, bot detection is a fundamental asset for understanding social media manipulation and, more specifically, to reveal malicious accounts. In the last few years, the problem of detecting automated accounts gathered both attention and concern [13], also bringing a wide variety of approaches to the table [7, 8, 19, 27]. While increasingly sophisticated techniques keep emerging [19], in this study, we employ the widely used Botometer.2

Botometer is a machine learning-based tool developed by Indiana University [9, 28] to detect social bots in Twitter. It is based on an ensemble classifier [6] that aims to provide an indicator, namely bot score, used to classify an account either as a bot or as a human. To feed the classifier, the Botometer API extracts about 1,200 features related to the Twitter account under analysis. These features fall in six broad categories and characterize the account’s profile, friends, social network, temporal activity patterns, language, and sentiment. Botometer outputs a bot score: the lower the score, the higher the probability that the user is human. In this study we use version v3 of Botometer, which brings some innovations, as detailed in [31]. Most importantly, the bot scores are now rescaled (and not centered around 0.5 anymore) through a non-linear re-calibration of the model.

In Figure 1, we depict the bot score distribution of the 997,406 distinct users in our datasets. The distribution exhibits a right skew: most of the probability mass is in the range [0, 0.2] and some peaks can be noticed around 0.3. Prior studies used the 0.5 threshold to separate humans from bots. However, according to the re-calibration introduced in Botometer v3 [31], along with the emergence of increasingly more sophisticated bots, we here lower the bot score threshold to 0.3 (i.e., a user is labeled as a bot if the score is above 0.3). This threshold corresponds to the same level of sensitivity setting of 0.5 in prior versions of Botometer (cf. Fig 5 from [31]).

According to this choice, we classified 21.1% of the accounts as bots, which in turn generated 30.6% of the tweets in our data set. Overall, Botometer did not return a score for 35,029 users that corresponds to 3.5% of the accounts. We used the Twitter API to further inspect them. Interestingly, 99.4% of these accounts were suspended by Twitter, whereas the remaining percentage of users protected their tweets turning on the privacy settings of their accounts.

Political Ideology Inference

In parallel to the bot detection analysis, we examine the political leaning of both bots and humans in our dataset. To classify users based on their political ideology, we rely on the political leaning of the media outlets they share. We make use of a list of partisan media outlets released by third-party organizations, such as AllSides3

1https://github.com/A-Deb/midterms
2https://botometer.iuni.iu.edu/
3https://www.allsides.com/media-bias/media-bias-ratings
we compute precision and recall by reiterating the validation of the model. We apply a stratified cross-validation (5-fold) to a set composed of 38,920 seed accounts. We train the algorithm using 80% of the seeds and we evaluate the performance on the remaining 20%. Finally, we use label propagation to classify the remaining accounts, which exhibit an unbalance in favor of liberal bots. Further, we examine social bot partisanship and, accordingly, we analyze the combination of the outcome from the bot detection algorithm with the content generated by social bots. Thereby, we propose the following metrics:

- **Retweet Pervasiveness** $(RTP)$ measures the intrusiveness of bot-generated content in human-generated retweets:
  \[
  RTP = \frac{\text{no. of human retweets from bot tweets}}{\text{no. of human retweets}}
  \]
- **Reply Rate** $(RR)$ measures the percentage of replies given by humans to social bots:
  \[
  RR = \frac{\text{no. of human replies to bot tweets}}{\text{no. of human replies}}
  \]
- **Human to Bot Rate** $(H2BR)$ quantifies human interaction with bots over all the human activities in the social network:
  \[
  H2BR = \frac{\text{no. of humans interaction with bots}}{\text{no. of humans activity}}
  \]
  where the numerator counts for human replies/retweets to/of bots generated content, while the denominator is the sum of the number of human tweets, retweets, and replies.
- **Tweet Success Rate** $(TSR)$ is the percentage of tweets generated by bots that obtained at least one retweet by a human:
  \[
  TSR = \frac{\text{no. of tweet retweeted at least once by a human}}{\text{no. of bots tweets}}
  \]

### RESULTS

Next, we address the research questions discussed in the Introduction. We examine social bot partisanship and, accordingly, we analyze bots’ strategies and measure the effectiveness of their actions.

#### RQ1: Bot Political Leaning

The combination of the outcome from the bot detection algorithm and the political ideology inference allowed us to identify four groups of users, namely Liberal Humans, Conservative Humans, Liberal Bots, and Conservative Bots. In Table 2a, we show the percentage of users per group. Note that percentages do not sum up to 100 as each political ideology inference was not able to classify every user, or Botometer did not return a score, as we previously mentioned. In particular, we were able to assign a political leaning to 63% of bots and 67% of humans. We find that the liberal user population is almost three times larger than the conservative counterpart. This discrepancy is also present, but less evident, for the bot accounts, which exhibit an unbalance in favor of liberal bots. Further,
we investigate the suspended accounts to inspect the consistency of this result. The inference algorithm attributed a political ideology to 63% of these accounts, which in turn show once again the liberal advantage over the conservative faction (45% vs. 18%).

Figure 2 shows two \( k \)-core decomposition graphs of the retweet network. In a \( k \)-core, each node is connected with at least \( k \) other nodes. Figures 2a and 2b capture the 10-core and 25-core decomposition, respectively. Here, nodes represent Twitter users and link represent retweets among them. We indicate as source the user that retweeted the tweet of a target user. Colors represent the political ideology, with darker colors (red and blue) being bots and lighter colors (cyan and pink) being human users; size represents the in-degree. The graph is visualized using a force-directed layout [18], where nodes repulse each other, while edges attract their nodes. In our setting, this means that users are spatially distributed according to the amount of retweets between each other. The result is a network naturally split into two communities, where each side is almost entirely populated by users with the same political ideology. This polarization is also reflected by bots, which are embedded, with humans, in each political side. Two facts are worth noting: (i) as \( k \) increases, the left \( k \)-core appears to disrupt, while the right \( k \)-core remains well connected; and, (ii) as \( k \) increases, bots appear to outnumber humans, suggesting that bots may populate areas of the retweet network that are more central and better connected.

Next, we examine the topics discussed by social bots and compare them with the human counterparts. Table 3 shows the top 20 hashtags utilized by liberal and conservative bots. We highlight (in bold) the hashtags that are not present in the top 50 hashtags used by the corresponding human group to point out the similarities and differences among the groups. In this table, we do not take into account general hashtags (such as #elections, #midterms, #democrats, #liberals, #VoteRed(or Blue)ToSaveAmerica, and #Trump) as (i) the overlap between bot and human hashtags is noticeable when these terms are considered, and (ii) we aim to narrow the analysis to specific topics and inflammatory content, inspired by [26]. Moreover, we used an enlarged subset of hashtags for the human groups to further strengthen the differences and, at the same time,
RQ2: Bot Activity and Strategies

In this Section, we investigate social bot activity based on their political leaning. We explore their strategies in interacting with humans and the degree of embeddedness in the social network.

Table 2b depicts the number (and percentage) of tweets generated by each group. Despite the group composed of conservative bots is the smallest in terms of number of accounts, it produced more tweets than liberal bots and closely approaches the number of tweets generated by the human counterpart. The resulting tweet per user ratio shows that conservative bots produce 7.4 tweets per account, which is more than twice the ratio related to the liberal bots (3.5), almost the double of the human counterpart (3.9), and nearly three times the ratio of liberal humans (2.5).

To investigate the interplay between bots and humans, we consider the previously described retweet network. Figure 3 shows the interaction among the four groups. We maintain the same color mapping described before, with darker color (on the bottom) representing bots and lighter color (on top) indicating humans. Node size is proportional to the percentage of accounts in each group, while edge size is proportional to the percentage of interactions between each group. In Figure 3a, this percentage is computed considering all the interactions in the retweet network, while in Figure 3b we consider each group separately, therefore, the edge size gives a measure of the group propensity to interact with the other groups. Consistently with Figure 2, we observe that there is a limited amount of interaction between the two political sides. The majority of interactions are either intra-group or between groups of the same political leaning. From Figure 3b, we can observe that the two bot factions adopted different strategies. Conservative bots balanced their interactions by retweeting group members 43% of the time, and the human counterpart 52% of the time. On the other hand, liberal bots mainly retweeted liberal humans (71% of the time) and limited the intra-group interactions to the 22% of their retweet activity. Interestingly, conservative humans interacted with the conservative bots (28% of the time) much more than the liberal counterpart (16%) with the liberal bots. To better understand these results and to measure the effectiveness of both the strategies, in the next Section we evaluate the four metrics introduced earlier in this paper.

Finally, we examine the degree of embeddedness of both humans and bots within the retweet network. For this purpose, we first compute different network centrality measures, and then we adopt the k-core decomposition technique to identify the most central nodes in the graph. In Table 4, we show the average out- and in-degree centrality for each group of users. Out-degree centrality measures the quantity of outgoing links, while in-degree centrality considers the number of incoming links. Both of these measures are normalized by the maximum possible degree of the graph. Overall, conservative groups have higher centrality measures than the liberal ones. We can notice that conservative bots achieve the highest values both for the out- and in-degree centrality. To further investigate bots embeddedness in the social network, we use the k-core decomposition. The objective of this technique is to determine the set of nodes deeply embedded in a graph. The k-core is a subgraph of the original graph in which every node has a degree equal to or greater than a given value k. We extracted the k-cores from the retweet network by varying k in the range between 0 and 30. Figure 4 depicts the percentage of liberal and conservative users as a function of k. We can notice that, as k grows, the fraction of conservative bots increases, while the percentage of liberal bots remains almost stationary. On the human side, the liberal fraction drops with k, whereas the conservative percentage remains approximately steady. Overall, conservative bots sit in a more central position in the social network and are more deeply connected if compared to the liberal counterpart.
RQ3: Bot Effectiveness

In this Section, we aim to estimate the effectiveness of bot strategies and measure to what extent humans rely upon, and interact with the content generated by social bots. We examine the effect of bot activities by means of the four metrics described in Section Bot Activity Effectiveness. We evaluate each political side separately, thus, we compare the interaction between bots and humans with the same leaning. In Table 5, we depict the results for each group of bots. Diverse aspects are worthy of consideration. We can observe that conservative bots are significantly more effective than the liberal counterpart. Although the TSRs of the red and blue bots are comparable, the gap between the two groups, with respect to the other metrics, is significant. To carefully interpret this result, it should also be noticed that (i) the TSR is inversely proportional to the number of tweets generated by bots, and (ii) conservative bots tweeted more than the liberal counterpart, as depicted in Table 2b. Overall, conservative bots received a larger degree of interaction with (and likely trust from) human users. In fact, conservative humans interacted with the bot counterpart almost twice with retweets (RTP), and more than three times with replies (RR) if compared to the liberal group. Finally, the H2BR highlights a remarkable amount of human activities that involve social bots: almost one in four actions performed by conservative humans goes towards red bots.

CONCLUSIONS & FUTURE WORK

In this work, we conducted an investigation to analyze social bots activity during the 2018 US Midterm election. We showed that social bots are embedded in each political wing and behave accordingly. We observed different strategies between conservative and liberal bots. Specifically, conservative bots stand in a more central position in the social network and abide by the topic discussed by the human counterpart more than the liberal bots, which in turn exhibit an inflammatory attitude. Further, conservative bots balanced their interaction with humans and bots of the red wing, whereas liberal bots focused mainly on the interplay with the human counterpart.

Finally, we inspected the effectiveness of these strategies and recognized the strategy of the conservative bots as the most effective. However, these results open the door to further interpretation and discussion. Are conservative bots more effective because of their strategy or because of the human ineptitude to distinguish their nature? This, and related analysis, will be expanded in future work.

REFERENCES

[1] Adam Badawy, Emilio Ferrara, and Kristina Lerman. 2018. Analyzing the Digital Traces of Political Manipulation: The 2016 Russian Interference Twitter Campaign. In Int. Conference on Advances in Social Networks Analysis and Mining. 258–265.

[2] Adam Badawy, Kristina Lerman, and Emilio Ferrara. 2018. Who Falls for Online Political Manipulation? arXiv preprint arXiv:1808.03284 (2018).

[3] Alessandro Bessi and Emilio Ferrara. 2016. Social bots distort the 2016 US Presidential election online discussion. First Monday 21, 11 (2016).

[4] Olga Boitkach, Sam Jackson, Jeff Hemsley, and Sikana Tanupabrungsun. 2018. Automated Diffusion? Bots and Their Influence During the 2016 US Presidential Election. In International Conference on Information. Springer, 17–26.

[5] Alexandre Bovet and Hernán A Makse. 2019. Influence of fake news in Twitter during the 2016 US presidential election. Nature communications 10, 1 (2019), 7.

[6] Leo Brennan. 2001. Random forests. Machine learning 45, 1 (2001), 5–32.

[7] Nikan Chavoshi, Hossein Hamoobi, and Abdullah Mueen. 2016. DeBot: Twitter Bot Detection via Warped Correlation. In ICDM. 817–822.

[8] Zhouban Chen and Devika Subramanian. 2018. An Unsupervised Approach to Detect Spam Campaigns that Use Botnets on Twitter. arXiv preprint arXiv:1804.05232 (2018).

[9] Clayton Allen Davis, Onur Varol, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. 2016. Botomret: A system to evaluate social bots. In Proceedings of the 25th International Conference Companion on World Wide Web.

[10] Nathaniel Persily. 2017. The 2016 US Election: Can democracy survive the internet? Journal of democracy 28, 2 (2017), 63–76.

[11] Emilio Ferrara. 2015. Manipulation and abuse on social media. ACM SIGWEB Newsletter Spring (2015), 41–44.

[12] Mathieu Jacomy, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian. 2014. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. PloS one 9, 6 (2014), e98679.

[13] Engels Guo, Tonjon Njamen, and Joshua Tucker. 2019. Less than you think: Prevalence and predictors of fake news dissemination on Facebook. Science Advances 5, 1 (2019), eaau4586.

[14] Philip N Howard, Gillian Boshover, Bence Kollanyi, Sami Hassinen, Lisa-Maria Neudert, and Sune Lehmann. 2017. Evidence of Complex Contagion of Information in Social Media: An Experiment Using Twitter Bots. PloS one 12, 9 (2017), e0184148.

[15] Dietram A Scheufele and Nicole M Krause. 2019. Science audiences, misinformation, and fake news. PNAS 116, 1 (2019), 300–301.

[16] Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Filippo Menczer, and Alessandro Flammini. 2017. Online human-bot interactions: Detection, estimation, and characterization. In Int. AAAI Conference on Web and Social Media. 280–289.

[17] Samuel C Woolley and Douglas R Guilbeault. 2017. Computational propaganda in the United States of America: Manufacturing consensus online. Computational Propaganda Research Project (2017), 22–36.

[18] Kai Shi, Ami Sliva, Suileang Wang, Jialiang Tang, and Huai Lu. 2017. Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter 19, 1 (2017), 22–36.

[19] Maximo Stollen, Emilio Ferrara, and Manlio De Domenico. 2018. Bots increase exposure to negative and inflammatory content in online social systems. Proceedings of the National Academy of Sciences 115, 49 (2018), 12435–12440.

[20] Yasameh Shamsi, Afshin Vakilzadeh, Tapan K Grover, Sumeet S Thakur, and Ravi Rastogi. 2016. Bot detection using machine learning. In SDM. 143–154.

[21] Hadi Zeeb, Ayman Elzataty, Shvigelys Kuznetsova, and Emilio Ferrara. 2017. Social bots detection. In Interactive Learning Environments. 375–392.