

Caracterização de *bots* no Twitter durante as Eleições Presidenciais no Brasil em 2018

Gabriel P. Nobre¹, Jussara M. Almeida¹, Carlos H. G. Ferreira^{1,2}

¹Universidade Federal de Minas Gerais

²Universidade Federal de Ouro Preto

{gnobre, jussara, chgferreira}@dcc.ufmg.br

Abstract. *In 2018 occurred the Brazil's presidential elections and it was widely spread on the media that bots were used in social media to share fake news and increase online support. That given, the goal on this paper is to characterize bot's behavior in Twitter during the electoral campaign. To do so, we identified 2,000 bot users, calculated behavior metrics and compared them with common users. After that, we were able to estimate their influence on the social network by quantifying influence metrics and the results suggested that they indeed affected the online discussions.*

Resumo. *Em 2018 ocorreram as Eleições Presidenciais no Brasil e foi frequentemente divulgado pela mídia que bots foram utilizados nas redes sociais para compartilhar fake news e aumentar o engajamento de candidatos. Com isso, o objetivo deste trabalho é caracterizar o comportamento desses usuários no Twitter durante as campanhas eleitorais. Para tanto, foram identificados mais de 2000 usuários bots, com métricas de atividade calculadas e comparadas com usuários comuns. Enfim, a influência de tais usuários foi quantificada e os resultados indicam que os bots afetaram as discussões online.*

1. Introdução

O sistema de interações humanas nas redes sociais tem frequentemente instigado estudos sobre como o compartilhamento de conteúdo *online* afeta e pode ser afetado por eventos no mundo real. Nesse contexto, torna-se necessário discutir a contribuição dessas mídias em acontecimentos sociopolíticos, como eleições, guerras civis, eventos esportivos e acidentes naturais.

Trabalhos recentes têm se dedicado a estudar a interação de contas automatizadas (usuários *bots*) em diferentes redes e contextos sociais. Eles estudam, por exemplo, detecção de *bots* [Abokhodair et al. 2015, Ferrara et al. 2014, Rizoïu et al. 2018, Dong and Liu 2018], como é o conteúdo compartilhado por *bots* [Caetano et al. 2017, Caetano et al. 2018], a interferência dos *bots* na dinâmica de compartilhamento de informações [Ratkiewicz et al. 2011, Rizoïu et al. 2018], bem como em eventos externos a Internet [Ferrara et al. 2014].

Dada a possibilidade de que contas automatizadas podem interferir em eventos sociais, surgiram trabalhos que estudam como esse fenômeno pode acontecer [Messias et al. 2013, Bence Kollanyi and Woolley 2016]. Por exemplo, a interferência de *bots* que pode ter acontecido durante as eleições presidenciais no Estados Unidos da

América em 2016 tem sido estudada recentemente [Bessi and Ferrara 2016] e um número de trabalhos concorda que *bots* afetam discussões *online* [Ferrara et al. 2014].

Com indícios de que atividades de usuários automatizados podem interferir em eventos sociais, o objetivo deste trabalho é estudar o comportamento de *bots* no Twitter durante as Eleições Presidenciais no Brasil em 2018. Esse evento político esteve em evidência na mídia devido à detecção de compartilhamento de *fake news* e da utilização de *bots* para propagá-las em plataformas como Twitter, WhatsApp e Facebook [Cristina Tardáguila and Ortellado 2018, Long 2018, Owen 2018]. Sendo assim, neste trabalho estudamos características de comportamento de *bots* e as comparamos com usuários comuns, além de calcular e analisar suas métricas de influência no Twitter.

Para guiar a realização do objetivo deste trabalho, foram definidas duas Questões de Pesquisa (QP):

- QP1: Usuários *bots* estavam presentes no Twitter em meio a discussões sobre as Eleições Presidenciais no Brasil em 2018? Se sim, como se compara a atividade de usuários *bots* com usuários comuns?
- QP2: Se sim, estes usuários *bots* influenciaram usuários comuns?

Para responder as QP's surgem três desafios. O primeiro desafio é coletar uma base de dados abrangente que possibilite as análises desejadas. Para tanto, coletamos cerca de 40 milhões de *tweets* que continham pelo menos o nome de um dos quatorze candidatos a presidência. O segundo é detectar *bots* na base de dados para comparar as métricas de atividade com usuários comuns. Com isso, classificamos 100 mil usuários da base coletada e os separamos em grupos de níveis de automação para, então, realizar as comparações de comportamento pela análise do resultado do cálculo de 18 métricas de atividade. O terceiro é relacionado ao cálculo de métricas de influência no Twitter nesse contexto. Calculamos duas diferentes métricas de influência inspiradas dos trabalhos relacionados e avaliamos se *bots* foram hábeis em influenciar usuários comuns. Para superar tais desafios, foi realizada uma pesquisa de trabalhos relacionados para identificar as técnicas a serem utilizadas.

As seções seguintes deste trabalho estão organizadas da seguinte forma: A Seção 2 discute trabalhos relacionados. Na Seção 3, apresentamos a metodologia utilizada. A Seção 4 detalha os resultados experimentais obtidos. A Seção 5 discute os resultados obtidos relacionando-os com as QP's. E por fim, a Seção 6 conclui este estudo.

2. Trabalhos relacionados

Os trabalhos relacionados pesquisados podem se dividir em duas seções. A primeira se refere a trabalhos de identificação e caracterização de usuários *bots* no Twitter, bem como a influência desses usuários em eventos externos ao mundo *online*. A segunda seção reúne trabalhos que estudam métricas de influência de usuários no Twitter.

2.1. Identificação e caracterização de *bots*

Nesta seção de trabalhos relacionados relacionamos as pesquisas que estudaram mecanismos de detectar *bots* e apresentaram uma metodologia para caracterizar o comportamento de tais usuários.

A tarefa de identificar *bots* em uma rede social diversa como a de usuários do Twitter pode ser realizada por meio de diferentes abordagens. Com o objetivo de obter uma rápida detecção, [Howard and Kollanyi 2016] utilizou a métrica de número de *tweets* por dia de cada usuário e classificou como *bot* aqueles que passaram um limite mínimo. Em outro trabalho [Abokhodair et al. 2015], pesquisadores conseguiram detectar uma rede de *bots* manualmente. Já a pesquisa apresentada em [Dickerson et al. 2014] a detecção ocorreu com o desenvolvimento de uma ferramenta, denominada SentiBot, que se utiliza de métricas de *tweets* como sintaxe e semântica e, também, de característica do perfil do conta e imersão do usuário na rede do Twitter. Em outros trabalhos [Rizoiu et al. 2018, Caetano et al. 2018, Rizoiu et al. 2018], os pesquisadores utilizaram a Botometer API.

A Botometer API [Dong and Liu 2018] provê uma maneira de fácil utilização para classificar um usuário do Twitter como um usuário comum ou *bot*. Segundo os autores, são utilizadas mais de 1000 características das contas dos usuários em um modelo criado com o algoritmo Floresta Aleatória e retornando o CAP (do Inglês *Complete Automation Probability*, Probabilidade de Automação Completa) sendo necessário como entrada apenas o identificador de cada um. O CAP é um valor que varia de 0 a 1 e se refere a probabilidade de um usuário ser completamente automatizado, um *bot*. Os autores recomendam que usuários com CAP maior que 0.5 sejam classificados como *bots*.

O objetivo de muitos trabalhos é identificar *bots* e estudar como eles podem influenciar eventos externos à rede do Twitter. Tais trabalhos pesquisam esse assunto em diferentes contextos, como eventos sociais, políticos, eleições e guerras civis.

A influência de *bots* no Twitter é tópico de interesse em diferentes trabalhos [Ferrara et al. 2014, Howard and Kollanyi 2016, Ratkiewicz et al. 2011, Rizoiu et al. 2018]. Em [Rizoiu et al. 2018], os autores investigam como *bots* interferiram na propagação de *tweets* durante os debates nas Eleições Presidenciais nos Estados Unidos em 2016 e os caracterizaram de acordo com sua influência e polaridade política. O fenômeno de *astroturfing* foi estudado em [Ratkiewicz et al. 2011] com o objetivo de estudar como *bots* podem afetar a disseminação de informações sobre política nos Estados Unidos. A interação de *bots* com usuários comuns durante o referendo sobre a permanência do Reino Unido na União Europeia (*Brexit*) foi analisada em [Howard and Kollanyi 2016], e os autores concluíram que *bots* exerceram funções estratégicas e interferiram nas discussões *online*. No trabalho [Abokhodair et al. 2015] os autores revelaram uma rede de *bots* relacionada à Guerra Civil na Síria e calcularam a influência e a interação de *bots* com usuários comuns.

A Eleição Presidencial do Estados Unidos em 2016 foi um evento político de grande interesse acadêmico e tópico de diversas pesquisas. Em [Caetano et al. 2018], os autores desenvolveram um *framework* para identificar o comportamento de usuários durante a campanha eleitoral e os classificou em diferentes grupos em relação ao engajamento. Os pesquisadores em [Bence Kollanyi and Woolley 2016] identificaram *bots* e concluíram que tais usuários geraram um grande volume de dados. Já as conclusões obtidas em [Bessi and Ferrara 2016] expõe evidências de que *bots* podem ter influenciado o processo democrático nas eleições por potencialmente terem afetado a opinião pública por ocuparem posições centrais na rede de difusão de informações no Twitter. Nesse estudo, os autores estimaram que 15% dos usuários ativos eram *bots* e que eles foram autores por

19% do conteúdo gerado.

Em [Ferrara et al. 2014] os autores descrevem as possíveis consequências da existência de *bots* em redes sociais, como afetar o processo democrático alterando a percepção de usuários comuns. Os autores ainda identificaram que *bots* podem exercer influência significativa em usuários comuns e alertam para a dificuldade de estimar a população de *bots* na rede e, também, de identificar a interferência na propagação de conteúdo.

2.2. Métricas de influência

Nesta seção pesquisamos como medir métricas de influência no Twitter e identificamos que não existe um consenso, mas detectamos a tendência de relacionar influência com métricas de *retweet*.

O número de *retweets* como métrica de influência foi avaliado em [Anger and Kittl 2011, Ferrara et al. 2014, Rizoio et al. 2018], porém em [Anger and Kittl 2011] os autores desenvolveram uma métrica que considera o número de menções a outros usuários e em [Rizoio et al. 2018] foi calculada a probabilidade de um usuário ser *retweetado*. Em [Cha et al. 2010, Dubois and Gaffney 2014] os autores consideraram a centralidade dos usuários na rede de seguidores, mas o estudo em [Cha et al. 2010] concluiu que a popularidade de um usuário não está fortemente relacionada com suas métricas de propagação de conteúdo. Nos estudos [Stieglitz and Dang-Xuan 2012, Tinati et al. 2012], os autores geraram uma rede de *retweet* e estruturaram os perfis de comunicação de usuários por meio de suas métricas de *retweet*.

Os pesquisadores em [Abokhodair et al. 2015, Bakshy et al. 2011, Bessi and Ferrara 2016, Ratkiewicz et al. 2011] geraram grafos com conexões de *retweets* para estimar a influência de usuários. A interação entre humanos e *bots* foi descrita em [Bessi and Ferrara 2016], onde os autores concluíram que os *bots* têm sucesso em participar da rede de difusão de informações. Em [Abokhodair et al. 2015] os autores evidenciaram que uma rede de *bots* conseguiu acumular engajamento de usuários comuns e ressaltaram o crescimento dessa rede ao longo do tempo. O trabalho apresentado em [Bakshy et al. 2011] cita métricas de *retweet* e homofilia como consequência da influência de um usuário.

Os trabalhos [Anger and Kittl 2011, Messias et al. 2013] utilizam de APIs, como Klout e Twitalyzer, para calcular a influência de usuários. Essas APIs geram um medida de influência levando em consideração características do perfil dos usuários e, também, estimam o provável alcance do conteúdo gerado por uma conta.

É possível concluir pela análise dos trabalhos relacionados que *bots* podem afetar o processo democrático e, portanto, a detecção da presença de *bots* durante as Eleições Presidenciais no Brasil em 2018 [Cristina Tardáguila and Ortellado 2018, Long 2018, Owen 2018] eleva a necessidade de estudar a interação de humanos e *bots* nesse contexto.

O objetivo deste trabalho, portanto, é estudar o comportamento de *bots* a partir da detecção de contas automatizadas no Twitter e por meio do cálculo e comparação de métricas de comportamento. Além disso, avaliamos se tais usuários influenciaram usuários comuns no contexto sociopolítico, utilizando como estudo de caso um evento

apontado por grande parte da mídia brasileira e internacional, a eleição brasileira para presidente de 2018.

3. Metodologia

A metodologia utilizada neste trabalho é composta por duas etapas: A primeira se refere a coleta de dados. Já a segunda, refere-se ao cálculo de métricas dos usuários para detectar *bots* e para análise de influência exercida na rede. As próximas subseções detalham cada uma destas etapas.

3.1. Coleta de dados

A coleta de dados transcorreu-se de 11 de Setembro de 2018 a 5 de Novembro de 2018 com a utilização da Twitter Stream API [Twitter 2018]. Como palavras-chave para busca, utilizamos os nomes dos candidatos à Presidência da República conforme descrito na Tabela 1.

Tabela 1. Palavras-chave utilizadas para coleta de dados

Alvaro Dias
Cabo Daciolo, Daciolo
Ciro Gomes, Ciro
Eymael
Geraldo Alckmin, Alckmin
Guilherme Boulos, Boulos
Henrique Meirelles, Meirelles
Jair Bolsonaro, Bolsonaro
Joao Amoedo, Amoedo, Amoedo
Joao Goulart Filho, Joao Goulart
Luiz Inácio Lula Da Silva, Lula
Fernando Haddad, Haddad
Marina Silva, Marina
Vera Lucia

Como resultado da coleta, obtivemos mais de 300GB de dados contendo cerca de 40 milhões de *tweets*. Para examinar os dados, restringimos o período de análise do sétimo dia anterior ao quinto posterior ao primeiro e segundo turno das eleições sendo, portanto, 26 dias sob avaliação. Este conjunto de dados compreende-se em 15 milhões de *tweets* de mais de 2.3 milhões de usuários.

3.2. Cálculo de métricas de usuários

Após a coleta de dados, o passo seguinte foi a detecção de *bots* na base de dados para posterior caracterização e comparação com usuários comuns. Para tanto, decidimos utilizar Botometer API [Dong and Liu 2018], ferramenta comumente utilizada na tarefa de identificação de *bots* no Twitter.

A Botometer API recebe como entrada o identificador de um usuário e retorna o valor do CAP (do Inglês *Complete Automation Probability*, Probabilidade de Automação Completa). O CAP é um valor variável de 0 a 1 que estima a probabilidade de uma conta

de usuário do Twitter ser completamente automatizada, um *bot*. O valor de referência para considerar um usuário como *bot* é de 0.5. Sendo assim, consideramos usuários com CAP menor que 0.5 como usuários comuns e com valor maior que 0.5 como usuários *bot*. Em alguns casos, não foi possível identificar o CAP de usuários e, após uma verificação manual, reunimos esses usuários em um grupo com contas suspensas.

O período de análise de 26 dias agrupou *tweets* de 2.3 milhões de usuários distintos e, dada a relativa dificuldade em classificar todos eles, decidimos por selecionar os 20 mil usuários mais ativos (em número de *tweets* publicados) por dia. Com isso, foi possível calcular o CAP de 107,877 usuários. O número de usuários classificados como comuns, *bots* ou suspensos encontra-se na Tabela 2.

Tabela 2. Classificação de usuários por grupo

	Comum	Bot	Suspense	Total
Quantidade	101,254	2,031	4,602	107,887
Percentual	93.85%	1.88%	4.27%	100.00%

O trabalho que apresenta a Botometer API [Dong and Liu 2018] descreve as métricas mais estatisticamente importantes para o cálculo de automação de usuários do Twitter. Com isso, nos inspiramos nessas métricas para selecionar e calcular outras 18 métricas para cada um dos usuários classificados a fim de comparar o comportamento de *bots* com usuários comuns. Em seguida, calculamos o valor médio de cada métrica com um nível de confiança de 99%. Os resultados estão apresentados na Tabela 3.

Tabela 3. Valores médios de métricas de atividade por grupo de usuários

Métrica	Comum	Bot	Suspense
Amigos	1,148 ±34	604 ±98	784 ±65
CAP	0.029 ±0.001	0.662 ±0.006	
Contagem de <i>retweets</i> recebidos	104.79 ±18.98	10.18 ±7.49	39.32 ±25.43
<i>Favourites</i> †	14,458 ±224	1,029 ±375	12,858 ±832
<i>Hashtags</i> distintas publicadas	9.91 ±0.16	6.48 ±0.89	7.39 ±0.53
Idade da conta	1,727 ±10	1,352 ±70	1,536 ±48
Menções	20.34 ±0.43	7.79 ±1.76	21.04 ±2.79
<i>Retweets</i>	135 ±2	64 ±12	92 ±9
<i>Retweets</i> de <i>bots</i>	0.96 ±0.17	0.65 ±0.85	0.35 ±0.19
<i>Retweets</i> de comuns	100 ±18	9 ±6	36 ±23
<i>Retweets</i> de suspensos	3.18 ±0.51	0.32 ±0.18	2.90 ±1.82
<i>Statuses (tweets)</i> †	26,000 ±425	5,951 ±1,277	26,555 ±1,891
<i>Statuses (tweets)</i> por dia †	17.47 ±0.23	6.19 ±1.20	18.83 ±1.12
Seguidores	3,182 ±724	447 ±145	1,216 ±514
Taxa de <i>retweet</i>	0.75 ±0.002	0.71 ±0.024	0.69 ±0.012
<i>Tweets</i>	156 ±3	79 ±12	118 ±10
<i>Tweets</i> com <i>link</i>	135 ±2	70 ±11	99 ±9
Usuários distintos de <i>retweet</i>	65.74 ±0.97	38.96 ±5.57	52.77 ±3.75

† Valor da métrica se refere à atividade externa ao período de coleta

A literatura atual descreve múltiplas maneiras de se calcular a influência de

usuários no Twitter: de métricas simples, como contagem de *favourites*, a mais complexas, como avaliar a centralidade na rede, ou mesmo ferramentas dedicadas. Neste estudo utilizamos duas métricas para avaliar a influência de usuários: número de *retweets* e centralidade de usuários na rede de *retweets*.

O número de *retweets* é uma métrica que quantifica o número de vezes que um usuário foi *retweetado* por outro no período em análise. O trabalho [Cha et al. 2010] demonstrou que tal métrica tem relação direta com o engajamento e o alcance do conteúdo gerado por um usuário e, portanto, indica influência. Os valores na Tabela 3 para o número de *retweets* revelam que o usuário comum médio é o tipo mais influente na rede, seguido pelo usuário suspenso e pelos *bots*. Essa métrica, porém, não atesta se *bots* foram eficientes em influenciar usuários comuns.

Para tanto, foi necessário determinar se *bots* foram *retweetados* por usuários comuns. A métrica denominada *retweets* de comuns na Tabela 3 na coluna *Bot* indica o valor médio de *retweets* de usuários comuns recebidos por *bots*. O valor obtido (9 ± 6) indica que um *bot* médio foi capaz de influenciar usuários comuns.

A segunda métrica de influência é a centralidade de *bots* no grafo gerado pela conexão de *retweets*, em que cada usuário é representado por um nó e as arestas são não direcionais. Para calculá-la, utilizamos a metodologia descrita em [Bessi and Ferrara 2016] que aplica o algoritmo *k-core*. Tal algoritmo funciona com a obtenção de subgrafos de outro grafos com a seleção de nós com um grau (*k*) mínimo. Intuitivamente, é possível considerar que usuários mais influentes ocupam posições mais centrais na rede. Sendo assim, obtivemos os subgrafos para valores crescentes de *k* e calculamos o percentual de usuários comuns, *bots* e suspensos em cada um deles. Com esses resultados, incluímos também na análise os usuários com classificação desconhecida.

4. Resultados experimentais

Esta seção de resultados experimentais se divide em duas subseções. Primeiro, se refere a comparação de comportamento de *bots* com usuários comuns e, segundo, analisamos a influência exercida por *bots* na rede.

4.1. Comparação de comportamento de *bots* e usuários comuns

Os resultados apresentados na Tabela 2 indicam a presença de cerca de 2 mil usuários automatizados no Twitter em meio as discussões sobre os candidatos à Presidência da República. Esse resultado ressalta a necessidade de investigar o comportamento de tais usuários e sua interação com usuários comuns durante as Eleições Presidenciais no Brasil em 2018.

Os resultados na Tabela 3 indicam que, quando comparado com um usuário comum médio, um usuário *bot* médio:

- Possuem contas mais novas
- Interagem com menos *hashtags*
- *Retweetam* menos usuários distintos
- *Favouritam* menos *tweets*
- Possuem menos seguidores e menos amigos
- Mencionam menos outros usuários

- *Retweetam* menos, mas em proporção parecida
- São menos *retweetados*
- Publicam menos *tweets*
- Publicam menos *tweets* com *link*

A análise das métricas de atividade dos usuários suspensos permite observar resultados ambíguos. Os valores das métricas de número de *favourites*, menções, *statuses* e *statuses* por dia são mais próximas aos resultados de usuários comuns do que *bots*. Porém, os valores de *hashtags* distintas publicadas, número de amigos e número de *retweets* se aproximam mais aos de usuários *bots*.

A existência dessa ambiguidade sugere a presença de *bots* classificados como usuários suspensos. Na tentativa de identificá-los, treinamos um modelo de classificação utilizando o algoritmo Floresta Aleatória. Para treinar o modelo, consideramos 17 métricas (desconsideramos o CAP) listadas na Tabela 3. O modelo gerado apresentou uma acurácia de 87% e F1 Score de 0.78, resultando na classificação de 507 *bots* dentre os 4,602 usuários com contas suspensas. Esse resultado fortalece a suspeita da presença de *bots* dentre os usuários classificados como suspensos.

Os valores médios para as métricas de atividade são úteis para comparação mas não representam a distribuição dos valores para cada grupo de usuários. Para tanto, geramos os gráficos de CDF (do Inglês *Cumulative Distribution Function*, Função Distribuição Acumulada) para quatro das métricas apresentadas na Tabela 3: idade da conta, CAP, número de *favourites* e taxa de *retweet*. Os gráficos gerados estão apresentados nas Figuras 1, 2, 3 e 4 a seguir:

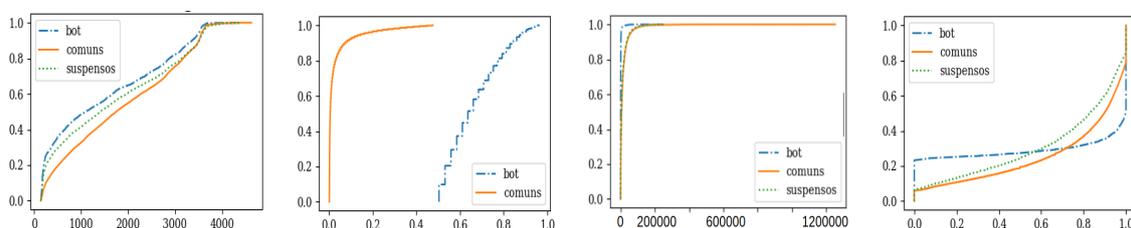


Figura 1. Idade

Figura 2. CAP

Figura 3. *Favourites*

Figura 4. Taxa de *retweet*

Na Figura 1, observa-se que os *bots* são contas de usuários mais recentes do que usuários comuns. A CDF para o CAP na Figura 2 ilustra que grande parte dos usuários comuns apresentam valores próximos a 0 enquanto que para *bots* os valores estão melhor distribuídos na faixa entre 0.5 e 1. A Figura 3 indica que *bots* possuem menos interações por *favourites* do que usuários comuns. Já a Figura 4 mostra que a taxa de *retweet* (contagem de *retweets* dividido pelo número de *tweets*) obtida para os usuários *bots* concentra-se perto dos valores de 0 e de 1, indicando uma distribuição bimodal.

4.2. Influência exercida por *bots*

Os resultados obtidos para as métricas de influência de *bots* estão apresentados nas Figuras 5, 6 e 7 a seguir.

A Figura 5 apresenta o resultado do percentual de cada classificação de usuários em relação ao subgrafos gerados do grafo de *retweets*. Os resultados mostram que com

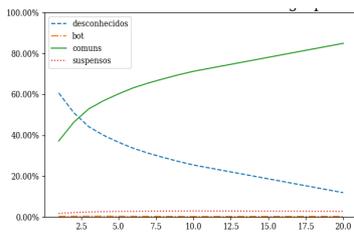


Figura 5. Distribuição de usuários por k -core

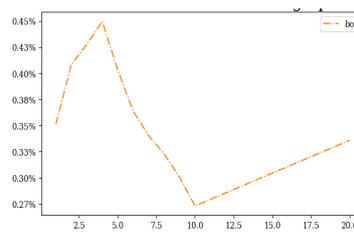


Figura 6. Distribuição de bots por k -core

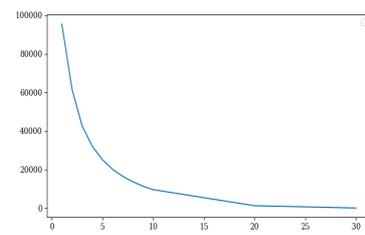


Figura 7. Número de usuários por k -core

o crescimento de k o número de usuários com classificação desconhecida diminui e o número de usuários com classificação conhecida aumenta e indicam, portanto, que os usuários mais influentes foram classificados.

A Figura 6 detalha o percentual de *bots* nos subgrafos com o crescimento de k . Os resultados indicam que os *bots* se mantiveram presentes nos subgrafos gerados, ou seja, alguns deles possuíam grande centralidade na rede. Sendo assim, é possível considerar que esses resultados fortalecem os indícios de que usuários automatizados influenciaram usuários comuns, dado que *bots* ocuparam posições de influência no grafo gerado pela rede de *retweets*.

A Figura 7 exibe a quantidade de usuários por subgrafos gerados. Esse resultado revela que o valor diminui exponencialmente com o aumento de k , ocorrendo grande restrição de usuários quando subgrafos são selecionados sequencialmente.

5. Discussão dos resultados

Os resultados experimentais foram obtidos a fim de responder as Questões de Pesquisa definidas. Para tanto, foram coletados cerca de 40 milhões de *tweets* com referência aos candidatos à Presidência da República nas Eleições no Brasil de 2018, identificamos usuários *bots* e comparamos sua atividade com usuários comuns. Além disso, calculamos duas métricas para avaliar a influência de *bots* sobre usuários comuns.

A QP1 se refere a detecção de *bots* nas discussões no Twitter sobre as Eleições Presidenciais. Conforme descrito na seção de Resultados Experimentais, utilizamos a Botometer API [Dong and Liu 2018] para classificar mais de 100 mil contas do Twitter e detectamos a presença de cerca de 2 mil *bots*. Tal resultado indica, portanto, que os *bots* estavam presentes nas discussões *online*. Além disso detectamos cerca de 4,5 mil usuários com contas suspensas e, ao gerar um modelo de classificação com o algoritmo de Floresta Aleatória, identificamos outros 507 usuários suspeitos como *bots*. Esse resultado sugere que o Twitter suspendeu alguns dos usuários *bots*.

A CDF para o valor do CAP, disponível na Figura 2, mostra que cerca de 80% dos usuários comuns possuem valores pequenos, ou seja, a maior parte deles demonstra um comportamento humano. O 20% restante obteve CAP crescente, sugerindo que podem ser usuários com baixo nível de automação. Essa mesma CDF revela valores diversos de CAP para os *bots*, sinalizando que tais usuários possuem comportamento diverso na rede.

Na QP1 questiona-se a comparação do comportamento de *bots* e usuários comuns. Para tanto, calculamos o intervalo de confiança de 18 métricas de atividade para cada uma das classificações de usuários e os apresentamos na Tabela 3. Os resultados indicam

diferenças nas métricas atividade e, portanto, do comportamento entre *bots* e usuários comuns. Por exemplo, é possível observar que os *bots* entraram mais recentemente na rede que usuários comuns (também observável na Figura 1), são menos ativos (métricas de *tweets* e *retweets*) e tem menos engajamento de outros usuários (métricas de amigos, seguidores e *retweets* recebidos). Além disso, a Figura 4 indica um comportamento bimodal na distribuição da métrica de taxa de *retweets*, não observável em usuário comuns. Esse resultado sugere a existência de três grupos de *bots*: os que geram conteúdo, os que promovem conteúdo por meio de *retweets* e, em menor número, os que exercem as duas atividades.

É observável também que, em comparação ao usuário comum médio, um *bot* médio publica menos *hashtags* distintas, *retweeta* menos usuários e tem menos *favorites*. Tais resultados são indícios de que os *bots* tem menos engajamento na rede do que usuários comuns e, além disso, tem dificuldade em seguir conteúdos e usuários diversificados.

Para avaliar se *bots* e usuários comuns se interessaram pelos mesmos assuntos analisamos as *hashtags* compartilhadas. Para tanto, criamos dois *rankings* de *hashtags* compartilhadas, um para *bots* e um para usuários comuns, e calculamos a relação entre eles utilizando o coeficiente de correlação de *Spearman* e obtivemos ρ igual a 0.20. Ainda detectamos que, dentre as 5 mil *hashtags* mais publicadas pelos *bots*, 654 não estavam presentes nas 5 mil mais publicadas por usuários comuns. Esses resultados indicam uma fraca correlação entre os dois *rankings* mas indica que os dois grupos de usuários estavam interessados nos mesmos tópicos.

Na QP2 questiona-se o objetivo é determinar se *bots* influenciaram usuários comuns. Para tanto, calculamos métricas de influência e detectamos que usuários comuns *retweetaram* *bots* e que usuários automatizados ocuparam posições centrais na rede (como visto na Figura 6). Esses resultados indicam que *bots* influenciaram usuários comuns. De forma geral, os resultados indicam que os *bots* nesse período mantiveram baixa atividade de publicação e engajamento (se comparados com usuários comuns) mas conseguiram se manter ativos na rede social e ainda influenciar outros usuários.

Os resultados obtidos são indícios da influência contas automatizadas nas discussões *online* no Twitter durante as Eleições Presidenciais no Brasil em 2018. De fato, tal constatação tem forte relação com outros trabalhos recentes que evidenciam a utilização de *bots* para influenciar discussões *online* por motivos políticos [Abokhodair et al. 2015, Bessi and Ferrara 2016]. Com isso, os resultados neste trabalho podem ser utilizados como evidência e, também, incitar outros estudos sobre a detecção de *bots* e as possíveis consequências na sociedade. Nesse sentido, estudos [Abokhodair et al. 2015, Bessi and Ferrara 2016, Ferrara et al. 2014, Rizoju et al. 2018] que descrevem consequências negativas para a sociedade e para o processo democrático. Sendo assim, o conhecimento desse contexto deve ser amplamente estudado e divulgado, a fim de evitar tais efeitos danosos à sociedade.

6. Conclusão

Os objetivos deste trabalho são detectar a presença de *bots* no Twitter durante as Eleições Presidenciais no Brasil em 2018, comparar o comportamento deles com usuários comuns e estimar se eles influenciaram outros. Concluimos que cerca de 2% dos usuários mais

ativos no período de análise indicavam alto nível de automação e que eles foram eficientes em influenciar usuários comuns. Sendo assim, os resultados obtidos realçam a necessidade de investimento na detecção de *bots* em redes sociais e, também, no estudos dos impactos causado por *bots* no mundo real.

Em trabalhos futuros pretendemos usar informações temporais baseadas no comportamento dos *bots*, contribuindo para a detecção desses perfis automatizados. Além disso, esperamos estudar a participação de usuários automatizados no compartilhamento de conteúdo no Twitter.

Referências

- 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, CSCW 2015, Vancouver, BC, Canada, March 14 - 18, 2015*, pages 839–851.
- Anger, I. and Kittl, C. (2011). Measuring influence on twitter. In *I-KNOW 2011, 11th International Conference on Knowledge Management and Knowledge Technologies, Graz, Austria, September 7-9, 2011*, page 31.
- Bakshy, E., Hofman, J. M., Mason, W. A., and Watts, D. J. (2011). Everyone’s an influencer: quantifying influence on twitter. In *Proceedings of the Forth International Conference on Web Search and Web Data Mining, WSDM 2011, Hong Kong, China, February 9-12, 2011*, pages 65–74.
- Bence Kollanyi, P. N. H. and Woolley, S. C. (2016). Bots and automation over twitter during the u.s. election.
- Bessi, A. and Ferrara, E. (2016). Social bots distort the 2016 U.S. presidential election online discussion. *First Monday*, 21(11).
- Caetano, J. A., Almeida, J. M., and Marques-Neto, H. T. (2018). Characterizing politically engaged users’ behavior during the 2016 US presidential campaign. In Brandes, U., Reddy, C., and Tagarelli, A., editors, *IEEE/ACM 2018 International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2018, Barcelona, Spain, August 28-31, 2018*, pages 523–530. IEEE Computer Society.
- Caetano, J. A., Lima, H. S., dos Santos, M. F., and Marques-Neto, H. T. (2017). Utilizando análise de sentimentos para definição da homofilia política dos usuários do twitter durante a eleição presidencial americana de 2016.
- Cha, M., Haddadi, H., Benevenuto, F., and Gummadi, P. K. (2010). Measuring user influence in twitter: The million follower fallacy. In *Proceedings of the Fourth International Conference on Weblogs and Social Media, ICWSM 2010, Washington, DC, USA, May 23-26, 2010*.
- Cristina Tardáguila, F. B. and Ortellado, P. (2018). Fake news is poisoning brazilian politics. whatsapp can stop it. <https://www.nytimes.com/2018/10/17/opinion/brazil-election-fake-news-whatsapp.html>. Accessed: 2018-12-02.
- Dickerson, J. P., Kagan, V., and Subrahmanian, V. S. (2014). Using sentiment to detect bots on twitter: Are humans more opinionated than bots? In *2014 IEEE/ACM Inter-*

- national Conference on Advances in Social Networks Analysis and Mining, ASONAM 2014, Beijing, China, August 17-20, 2014*, pages 620–627.
- Dong, G. and Liu, H. (2018). Feature engineering for machine learning and data analytics.
- Dubois, E. and Gaffney, D. (2014). The multiple facets of influence: Identifying political influentials and opinion leaders on twitter. *American Behavioral Scientist*, 58(10) 1260.
- Ferrara, E., Varol, O., Davis, C. A., Menczer, F., and Flammini, A. (2014). The rise of social bots. *CoRR*, abs/1407.5225.
- Howard, P. N. and Kollanyi, B. (2016). Bots, #strongerin, and #brexit: Computational propaganda during the UK-EU referendum. *CoRR*, abs/1606.06356.
- Long, C. (2018). How false news haunted the brazilian elections. <https://slate.com/technology/2018/10/brazil-election-fake-news-whatsapp-facebook.html>. Accessed: 2018-12-02.
- Messias, J., Schmidt, L., Oliveira, R., and Benevenuto, F. (2013). You followed my bot! transforming robots into influential users in twitter. *First Monday*, 18(7).
- Owen, L. H. (2018). What to know about whatsapp in brazil ahead of sunday's election. <http://www.niemanlab.org/2018/10/what-to-know-about-whatsapp-in-brazil-ahead-of-sundays-election/>. Accessed: 2018-12-02.
- Ratkiewicz, J., Conover, M. D., Meiss, M. R., Goncalves, B., Flammini, A., and Menczer, F. (2011). Detecting and tracking political abuse in social media. In Adamic, L. A., Baeza-Yates, R. A., and Counts, S., editors, *Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, July 17-21, 2011*. The AAAI Press.
- Rizoiu, M., 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 U.S. presidential debate. In *Proceedings of the Twelfth International Conference on Web and Social Media, ICWSM 2018, Stanford, California, USA, June 25-28, 2018.*, pages 300–309.
- Stieglitz, S. and Dang-Xuan, L. (2012). Political communication and influence through microblogging-an empirical analysis of sentiment in twitter messages and retweet behavior. In *45th Hawaii International International Conference on Systems Science (HICSS-45 2012), Proceedings, 4-7 January 2012, Grand Wailea, Maui, HI, USA*, pages 3500–3509. IEEE Computer Society.
- Tinati, R., Carr, L., Hall, W., and Bentwood, J. (2012). Identifying communicator roles in twitter. In Mille, A., Gandon, F. L., Misselis, J., Rabinovich, M., and Staab, S., editors, *Proceedings of the 21st World Wide Web Conference, WWW 2012, Lyon, France, April 16-20, 2012 (Companion Volume)*, pages 1161–1168. ACM.
- Twitter (2018). Twitter developer. <https://developer.twitter.com/en.html>. Accessed: 2018-12-02.