

Regional patterns and networked behavior: interdisciplinary methods for identifying *bots* in the 2020 elections in Brazil

João Guilherme Bastos dos Santos^{1,2}, Arthur Ituassu³, Sérgio Lifschitz³, Yago Cury³, Thayane Guimarães⁷, Diego Cerqueira⁶, Debora Albu⁸, Redson Fernando⁴, Julia Hellen Ferreira⁷, Maria Luiza Mondelli⁵

¹Instituto Nacional de Ciência e Tecnologia em Democracia Digital (INCT.DD)
Salvador, BA – Brasil

²Rooted in Trust (RiT)/Internews

³Departamento de Comunicação Social e Departamento de Informática
Pontifícia Universidade Católica do Rio de Janeiro (PUC-RJ)
Rio de Janeiro, RJ – Brasil

⁴Instituto de Tecnologia e Sociedade (ITS Rio)
Rio de Janeiro, RJ – Brasil

⁵Laboratório Nacional de Computação Científica (LNCC)
Petrópolis, RJ – Brasil

⁶Universidade Federal do Rio de Janeiro (PESC/COPPE - UFRJ)
Rio de Janeiro, RJ – Brasil

⁷Universidade Federal Fluminense (UFF)
Niterói, RJ – Brasil

⁸London School of Economics
Londres – Inglaterra

santos.jgb@gmail.com, ituassu@puc-rio.br, demotech@itsrio.org

Abstract. *To improve the detection of automated behavior on social networks in the Brazilian scenario, this work applies five methods of bot identification - analysis of (i) networks, (ii) sentiment, (iii) profile, (iv) temporal, and (v) crossing of these analyses - to find possible bots posting on the comments of tweets from candidates for the mayors' election in the capitals Rio de Janeiro, São Paulo, Recife, Porto Alegre and Fortaleza in 2020. After identification, we followed two strands of analysis (namely, network and lexical). We seek to better understand if there are different types of bots in each region, different network performances and vocabularies. Among the preliminary results obtained, we highlight the multiplicity of types of automation, understanding of regional variations and the role of bots in the Brazilian electoral scenario.*

Keywords. *Complex networks; Network science; Elections; Social bots.*

Resumo. *Para apoiar a detecção de comportamento automatizado em redes sociais no cenário brasileiro, o presente trabalho aplica cinco métodos de identificação de bots - análises de (i) redes, (ii) sentimento, (iii) perfil, (iv) temporal e (v) cruzamento dessas análises - a possíveis bots nos comentários de tweets dos candidatos à prefeitura das capitais Rio de Janeiro, São Paulo, Recife, Porto Alegre e Fortaleza em 2020. Após identificação, seguimos duas frentes de análise (a saber, de redes e lexical). Buscamos com isso entender melhor se há tipos de bots diferentes em cada região, diferentes atuações em rede e vocabulários. Dentre os resultados preliminares obtidos, destacamos a multiplicidade de tipos de automatização, compreensão das variações regionais e atuação de bots no cenário eleitoral brasileiro.*

Palavras-Chave. *Redes complexas; Ciência de redes; Eleições; bots sociais.*

1. Introdução

O papel das plataformas de redes sociais na dinâmica democrática é paradoxal: muitas vezes utilizadas como um canal que viabiliza, por exemplo, maior participação popular nos processos decisórios e no debate político, elas também têm sido utilizadas como forma de influenciar a opinião pública [Almeida et al. 2020] em direção a pautas pouco democráticas. O mesmo ocorre com a atuação de contas com comportamento automatizado, os *bots*. Eles são parte de estratégias que agravam esse tipo de influência indevida no processo democrático, mas igualmente úteis para reagir nesse novo cenário, auxiliando o monitoramento e resposta em tempo real. A discussão acerca desse tema ganhou mais centralidade a partir das eleições presidenciais nos Estados Unidos e o Brexit, ambos em 2016, bem como temores sobre operações de influência entre diferentes Estados, impulsionando o desenvolvimento de ferramentas e técnicas para identificar esse tipo de comportamento.

Há, no entanto, um problema muitas vezes ignorado nesse debate. A definição de *bots* enquanto conta com “comportamento atípico” esbarra no fato de que cada país possui um “comportamento típico” e cultura online peculiares, tornando a definição de quais comportamentos serão taxados como atípicos ou inautênticos em uma questão não apenas técnica, mas também política. Portanto, considerando o fato de diferentes países possuírem padrões de uso, contextos e gramáticas diversas, existe um desafio em avaliar se os métodos desenvolvidos no exterior são capazes de cobrir as especificidades do cenário brasileiro. A opacidade de parte dos identificadores de *bots* com relação aos métodos de detecção, a dificuldade em aplicar ao cenário nacional critérios feitos para outras realidades e a falta de clareza sobre que tipos de *bots* estão sendo discutidos em cada trabalho limitam as investigações sobre o tema.

Sendo assim, este trabalho se propõe a analisar a existência de comportamento automatizado em perfis do Twitter que interagiram textualmente (i.e. comentários) com candidatos às Prefeituras das capitais em que houve segundo turno durante as eleições brasileiras de 2020. Essas capitais são Rio de Janeiro, São Paulo, Recife, Porto Alegre e Fortaleza. Para apoiar a análise dos perfis, utilizamos o Pegabot, uma ferramenta para detecção de comportamento automatizado que busca trazer mais transparência ao debate

sobre o tema e ao processo de identificação desses perfis. O Pegabot traz um conjunto de critérios que nos permitem explorar as diferentes funções do uso de automação, ampliando a discussão para além da separação binária de perfis entre *bot* ou não *bot*. Vale mencionar que a ferramenta foi criada considerando características e padrões de uso do cenário brasileiro e foge do escopo deste trabalho fazer uma análise comparativa entre ferramentas de detecção de *bots*. A partir dos resultados obtidos, realizamos dois outros tipos de análises: análise de redes e lexical. Procuramos entender, tanto em termos de conteúdo quanto em termos de ação coordenada em rede, as características do comportamento automatizado no Brasil. Além disso, buscamos avaliar se é possível identificar variações nas funções e tipos de usuários automatizados de acordo com cada capital. Neste artigo, estendemos as análises apresentadas em [Santos et al. 2021], incluindo resultados que exploram tanto a visão regional quanto a visão por candidato, além de trazer novos aspectos sobre o conteúdo dos tweets. Vale mencionar que este trabalho continua sendo parte de um esforço mais amplo e ainda em andamento, que envolve a inclusão de todas as capitais brasileiras e análise de *retweets*.

Além desta seção de introdução, o presente trabalho está organizado como segue: a Seção 2 descreve a metodologia utilizada, a Seção 3, os resultados obtidos, e a 4 apresenta trabalhos relacionados. Por fim, a Seção 5 conclui o trabalho com algumas considerações finais e possibilidades futuras de pesquisa.

2. Metodologia

Neste trabalho, buscamos identificar alguns aspectos da atuação de perfis automatizados em eleições no Brasil, que incluem, por exemplo: (i) as potenciais diferenças na interação de *bots* com os perfis dos diferentes candidatos analisados; (ii) a existência de variações nos vocabulários de acordo com o alinhamento político dos alvos ou região do país; (iii) a existência de tipos e redes de *bots* com comportamentos distintos; e (iv) se há indícios que indiquem ações coordenadas em rede mobilizadas por *bots* e os eventuais reflexos dessas ações nas interações da rede.

Para investigar esses aspectos, dividimos a metodologia do trabalho em três principais etapas: (i) coleta de dados, (ii) aplicação de algoritmos de identificação de comportamento automatizado utilizados pelo Pegabot e (iii) cruzamento, análise de rede e lexical. Cada uma das etapas é descrita com mais detalhes nas subseções a seguir.

2.1. Coleta de dados

Para a coleta dos dados acerca dos *tweets*, foi utilizada a ferramenta ePOCS Twitter Crawler (eTC)¹[Rodriguez et al. 2019], desenvolvida pelos laboratórios de P&D BioBD e COMP da PUC-Rio. A eTC permite a extração de dados e metadados históricos do Twitter, coletando *tweets* sobre determinados assuntos dentro de um período de tempo especificado. A eTC é composta por uma aplicação Web e uma aplicação *standalone*, responsável por fazer o *crawling* e extrair os dados. Por meio da interface Web, os usuários agendam a extração de informações sobre *tweets*, especificando para isso o termo da busca e as datas.

¹<https://etc.biobd.inf.puc-rio.br/>

Para o presente trabalho, foram realizadas duas buscas para cada candidato. O período delimitado das buscas variou entre 26 de setembro e 30 de novembro de 2020 e as coletas foram feitas em duas etapas, uma primeira fase em dezembro de 2020 e uma segunda fase em agosto e setembro de 2021. A busca foi realizada para os candidatos às Prefeituras (2º turno) das capitais Rio de Janeiro, São Paulo, Recife, Porto Alegre e Fortaleza em 2020. Utilizamos o comando “from:@username” para coletar os tweets com solicitações de “retweets” e “replies”. Não obtivemos resultados na coleta apenas nos casos de Tião Bocalom e Socorro Neri, candidatos em Rio Branco (AC).

Assim, para cada *tweet* destas buscas, a eTC implementa um algoritmo que seleciona uma amostra de até 100 usuários que comentaram postagens dos candidatos e, em seguida, agrupa todos estes usuários contando a frequência com a qual cada um apareceu. O limite de 100 usuários foi estabelecido para viabilizar a análise respeitando os limites da API do Twitter e do processamento de dados para identificação de *bots* então disponível. Vale mencionar que os dados necessários à análise, processados apenas para fins de pesquisa, foram agregados, ou seja, qualquer dado referente a usuários específicos foi anonimizado e descartado após uso, garantindo a proteção aos dados e adequação às exigências da Lei Geral de Proteção de Dados (LGPD). As planilhas com os tweets, obedecendo as restrições de divulgação do Twitter, estão publicamente acessíveis².

2.2. Identificação de comportamento automatizado

Para a detecção de comportamento automatizado, foi utilizada a nova versão do Pegabot³, projeto do Instituto de Tecnologia e Sociedade do Rio de Janeiro (ITS Rio) e do Instituto Equidade & Tecnologia desenvolvido em 2018. O algoritmo do Pegabot utiliza como fonte de dados as informações públicas dos perfis no Twitter e critérios específicos para reconhecer padrões comportamentais. O objetivo da análise é identificar características que ajudem a determinar se o perfil apresenta comportamentos similares aos de contas automatizadas. Abaixo estão descritos resumidamente os critérios adotados pela ferramenta.

Análise do perfil do usuário: entre as informações públicas dos perfis consideradas, destacamos: o nome do perfil do usuário e quantos caracteres ele possui, quantidade de perfis seguidos (*following*) e seguidores (*followers*), texto da descrição do perfil, número de postagens (*tweets*) e favoritos.

Análise de rede: é feita a coleta de uma amostra de até 200 *tweets* mais recentes publicados na linha do tempo do usuário, identificando *hashtags* e menções utilizadas. O índice de rede busca compreender se o usuário está, por exemplo, encaminhando mensagens de *spam* para uma *hashtag* específica. Perfis com grande quantidade de interações e comportamento monotemático quanto às *hashtags* utilizadas tendem a ter uma pontuação superior nesse tipo de análise.

Análise de sentimentos: a amostra de até 200 *tweets* mais recentes publicados pela conta é usada para identificar a neutralidade do perfil. Isso é feito a partir de uma pontuação de valência atribuída a cada uma das palavras dos *tweets* coletados. A

²<https://tinyurl.com/2xx9nd8j>

³<https://pegabot.com.br/>

classificação se baseia em um dicionário pré-estabelecido de palavras e pontuações, a partir do qual é calculada a pontuação média para a quantidade de palavras positivas, negativas e neutras utilizadas pelo usuário. Quanto mais neutro, menor a chance de ser considerado um *bot*.

Análise temporal: considera a frequência de postagem de *tweets* para a amostra coletada, a data de criação da conta e o total de *tweets* publicados pelo perfil a fim de verificar se o usuário tem uma alta atividade de publicações.

Após coletar as informações, o Pegabot processa e transforma os dados recebidos em variáveis que compõem o cálculo final de probabilidade, a Análise Total. Para cada um dos critérios são atribuídos pesos e cálculo do resultado final, que é feito de forma determinística, indica o percentual de probabilidade de um perfil do Twitter ser um *bot*. Quanto maior o percentual, maior a chance da conta ser automatizada.

2.3. Cruzamento e análise de rede e lexical

Os dados obtidos nas etapas anteriores alimentam três tipos de análise de comportamento coordenado: análise de redes (diferente do índice de redes da etapa anterior, essa análise se refere a algoritmos de análise de redes complexas) envolvendo (a) *clusters* de *hashtags*, (b) *clusters* de perfis com citações mútuas e (c) análise lexical. Vale destacar que registros de retweets, que trazem repetições idênticas de textos já presentes na amostra, não foram considerados para esse cruzamento.

Através de uma expressão regular para raspagem, registramos as *hashtags* e perfis citados em cada *tweet* coletado, identificando que outras postagens compartilham as mesmas características. A partir desses dados, construímos dois grafos de rede, um em que perfis que mencionam outros perfis estão conectados a esses perfis, e outro em que *tweets* contendo *hashtags* conectam seu autor às *hashtags* utilizadas (bem como a outros autores que as utilizam). Esse último grafo segue uma estrutura bipartite, onde os vértices representam *hashtags* e perfis. Cada menção a um perfil ou *hashtag* conta como uma única aresta entre os vértices envolvidos. A partir dos *clusters* detectados nesse grafo por meio do método de Louvain para detecção de comunidades, podemos definir: (i) como a rede de perfis pode nos ajudar a compreender a distribuição de *bots* nessa rede; e (ii) quais são as diferenças entre os grupos identificados.

No que diz respeito à análise lexical dos *tweets*, o conteúdo das mensagens passa por uma interpretação centrada na identificação de tópicos e vocabulários (clusterização hierárquica descendente e análise fatorial de correspondência). O objetivo dessa análise é identificar os vocabulários dos diferentes grupos de *bots* participando ativamente das publicações sobre as eleições de 2020 nas campanhas das prefeituras selecionadas. Como base para comparação entre léxicos, utilizamos a combinação entre os critérios analisados pelo Pegabot. Todas as combinações possíveis dos critérios foram consideradas, buscando identificar a pluralidade dos tipos de *bots*.

O resultado final desses três níveis de análise mutuamente independentes passa por uma etapa de validação e avaliação cruzada. Ou seja, os achados obtidos através de um nível (lexical, por exemplo) são validados por meio de resultados similares provenientes de outro (análise de redes). Acreditamos que os resultados obtidos segundo a

metodologia proposta podem aprimorar a detecção e análise de comportamento coordenado automatizado, além de diferenciar os tipos de atores envolvidos. Como ferramentas, utilizamos a linguagem R para consolidar os dados obtidos e para a construção da rede e parte da análise lexical, o Gephi (versão 0.9.1) para análise de redes e elaboração das imagens apresentadas na Seção 3, e a interface IRaMuTeQ para apoiar a geração de imagens na análise lexical.

3. Resultados

Neste trabalho, focamos em avaliar os resultados da perspectiva da análise de redes. Seguindo a metodologia apresentada na seção anterior, coletamos uma amostra de 293.008 tweets que mencionaram candidatos às Prefeituras das capitais Rio de Janeiro, São Paulo, Recife, Porto Alegre e Fortaleza em 2020. Esse conjunto de tweets foi publicado por 79.653 usuários distintos e, após analisar os usuários usando o Pegabot, identificamos que 32,4% deles pontuaram mais de 70% em pelo menos um dos critérios utilizados pela ferramenta para detecção de comportamento automatizado. A avaliação do algoritmo e comparação com outras ferramentas foge do escopo deste trabalho.

A partir desses dados, seguimos com as análises propostas na terceira etapa. As redes a seguir têm como base as relações entre usuários (conexão por menção) e a rede de usuários e *hashtags* (conexão entre usuários e *hashtags*). A primeira possui 32.987 vértices, conectados por 39.778 arestas de menções. A segunda possui 21.300 vértices, conectados por 26.419 arestas de utilização de *hashtags*. Aplicando o algoritmo de detecção de comunidades segundo o método de Louvain na rede de menções, identificamos 3.601 comunidades. A distribuição dos perfis (sem considerar se são *bots*) é altamente concentrada em 5 comunidades, conforme mostramos na Figura 1: em rosa, verde, azul, marrom e laranja (ids Com0, Com1257, Com3177, Com1 e Com1040, respectivamente). Juntas, essas comunidades reúnem 54,4% dos usuários. Buscando identificar a atuação dos usuários automatizados nessa rede, filtramos os usuários que tiveram uma pontuação de pelo menos 70% na Análise Total, valor que consolida os critérios analisados pelo Pegabot. O resultado, destacando *bots* em roxo na Figura 3 mostra onde estão localizados os *bots* e apontam para uma concentração em algumas comunidades específicas, em especial a comunidade que aparece em laranja no primeiro grafo (Com3177).

Com esse primeiro resultado, buscamos identificar os diferentes padrões de comportamento dos usuários identificados como *bots*. Seleccionamos as 12 comunidades em que *bots* estiveram mais presentes (ids Com3177, Com857, Com0, Com1039, Com1040, Com1276, Com251, Com1422, Com918, Com3299, Com721, Com1029) e analisamos cada um dos critérios do Pegabot de forma separada. Verificamos, como mostra o gráfico de dispersão na Figura 3, que algumas comunidades trazem apenas um tipo de *bots* (1422 e 1276, por exemplo) enquanto outras combinam diferentes tipos de *bots* (1040 e 1039). Isso indica que os *bots* não estão distribuídos de modo homogêneo entre os *clusters* identificados, mas se concentram em partes específicas da rede, com menções a atores em comum.

Ao analisar os padrões lexicais utilizando a decomposição de valores singulares (SVD) e clusterização hierárquica descendente, associado à comparação entre os comentários dos perfis, e transformação dessas relações em um dendrograma (método Rei-

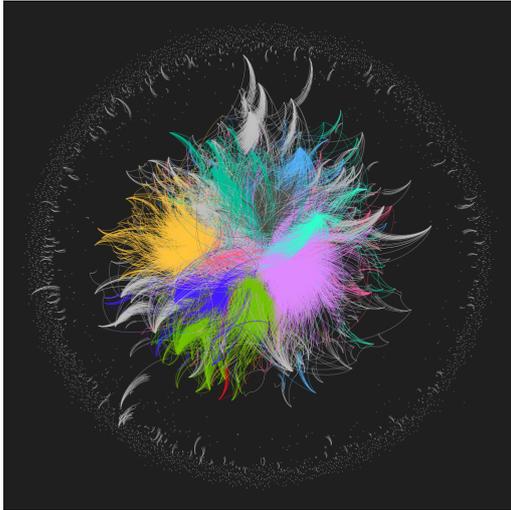


Figura 1. Comunidades da rede de menções

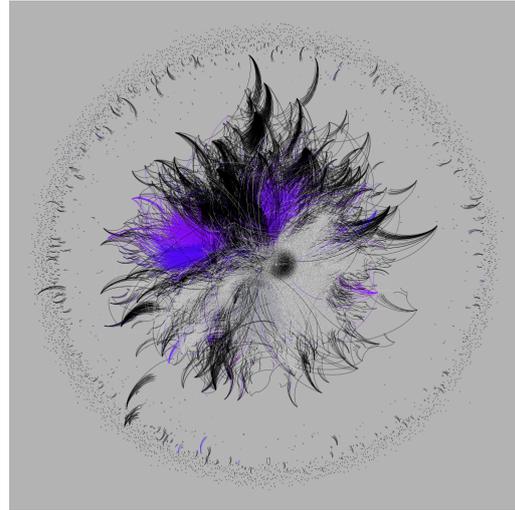


Figura 2. Interações mobilizadas por bots.

nert, aplicado através da interface para linguagem R IRaMuTeQ) percebemos os diferentes tópicos discutidos por usuários identificados como *bots* na rede de menções. Podemos perceber a pluralidade de léxicos nos *clusters* identificados, com oito vocabulários predominantes, como mostra a Figura 4: dois com vocabulário genérico (classes 2 e 7), um com elogios à gestão de candidatos do Rio do Janeiro (classe 8), três relacionados a São Paulo (classes 1, 3 e 4), um relacionado a Recife (classe 6) e um mais associado a Porto Alegre (classe 5).

Chamam atenção os elogios concentrados na classe 8 (que aparecem fortemente associados ao *cluster* 3177, que concentra a maior parte dos *bots*) ajudando a entender a relevância do Rio de Janeiro na comparação entre regiões. Ao mesmo tempo, há um conjunto de palavras de ataques, relacionadas aos ataques sofridos por Guilherme Boulos (invasor, vagabundo, bandido, laranjal e etc.). Ainda na rede de menções, há diferenças claras entre os tipos e frequência de *bots* identificados de acordo com o candidato em questão, como indicado na Figura 5.

Novamente, há diferenças claras na distribuição dos *bots* de acordo com a capital, considerando tanto o volume quanto o tipo de *bot*. A relevância do *cluster* voltado a elogios e o fato desses elogios estarem associados a textos direcionados a candidatos do Rio de Janeiro ajuda a entender a relevância do pleito na utilização de *bots* formando redes de menções na campanha de 2020. Esse cenário, no entanto, se restringe à rede de menções, não sendo uma característica geral da utilização de *bots* no Twitter quando consideramos *hashtags*, por exemplo. Na rede de *hashtags* os *bots* não se concentram em um único *cluster*, mas se espalham por diversas *hashtags*, muitas vezes sem outros conteúdos textuais. A seguir, duas versões do grafo composto por *hashtags* (em que postagens com as mesmas *hashtags* se conectam), uma com *clusters* considerando todos os interagentes e outra destacando em amarelo apenas os *bots*.

Há uma preponderância de São Paulo quando consideramos as *hashtags* (diferen-

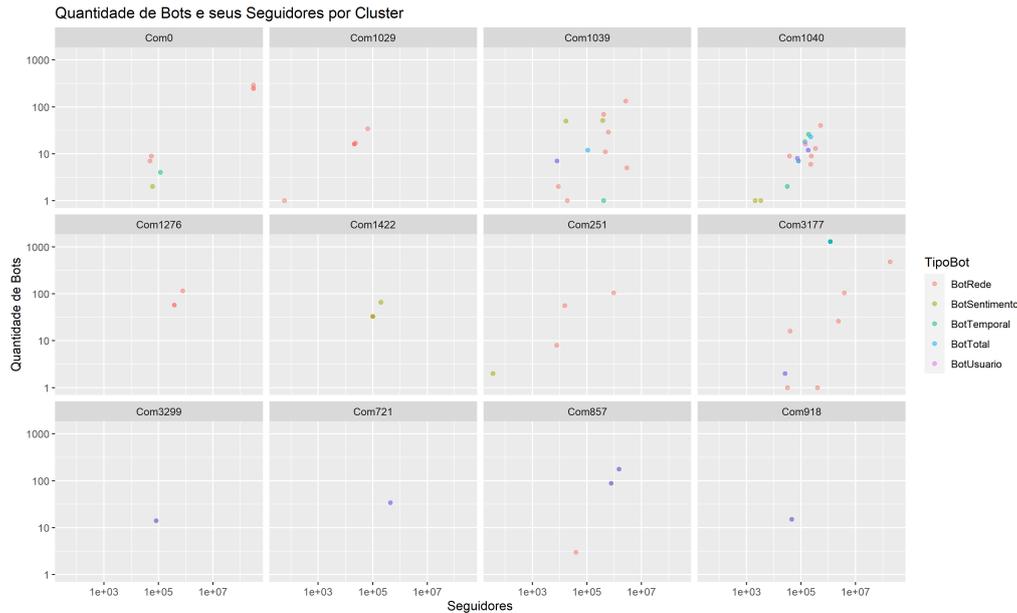


Figura 3. Quantidade de bots e volume de seguidores por comunidade.

temente da rede completa feita por menções entre usuários). Isso é um indicador de que existem diferenças na atuação e estratégias utilizadas por atores por trás de perfis automatizados em cada região e candidato. A identificação das *hashtags* se dá pela mesma base textual da identificação de menções. Como exposto anteriormente, há um vocabulário de ataque, particularmente contra o candidato Guilherme Boulos, relacionado aos termos “invasor”, “bandido” e “laranjal”. Isso é coerente com *hashtags* como ‘Laranjal-DoBoulos’ na rede de *hashtags*. No entanto, diferentemente da consideração das menções como critério de conexão, os *clusters* encontrados a partir de *hashtags* nos auxiliam em distinções temáticas, agrupando *hashtags* que versam sobre o mesmo tema, mas não concentram os possíveis bots em locais específicos da rede. Seja no tipo de bot identificado (que varia de acordo com o *cluster*), nas estratégias textuais (elogios no Rio de Janeiro e ataques em São Paulo), ou ainda nas ferramentas utilizadas (menções ou *hashtags*), há diferenças consideráveis na utilização de bot nos cenários encontrados. Mais do que isso, um identificador que foque em apenas um tipo de bots poderia distorcer completamente nossa leitura sobre esse cenário.

4. Trabalhos relacionados

Contas seguindo comandos automatizadas que compartilham conteúdo nas redes sociais são também conhecidas como bots sociais. Esses bots são controlados por aplicações (*software*), responsáveis por gerar conteúdo de forma artificial, imitando comportamento humano e estabelecendo interações com usuários autênticos [Ruediger et al. 2017]. Discussões sobre a existência e comportamento de bots em redes sociais se popularizaram nos últimos anos com o surgimento de estudos apontando a interferência dessas contas na criação de movimentos não espontâneos (operações de influência) visando influenciar o debate político. Parte relevante do esforço de compreensão desse fenômeno envolve a

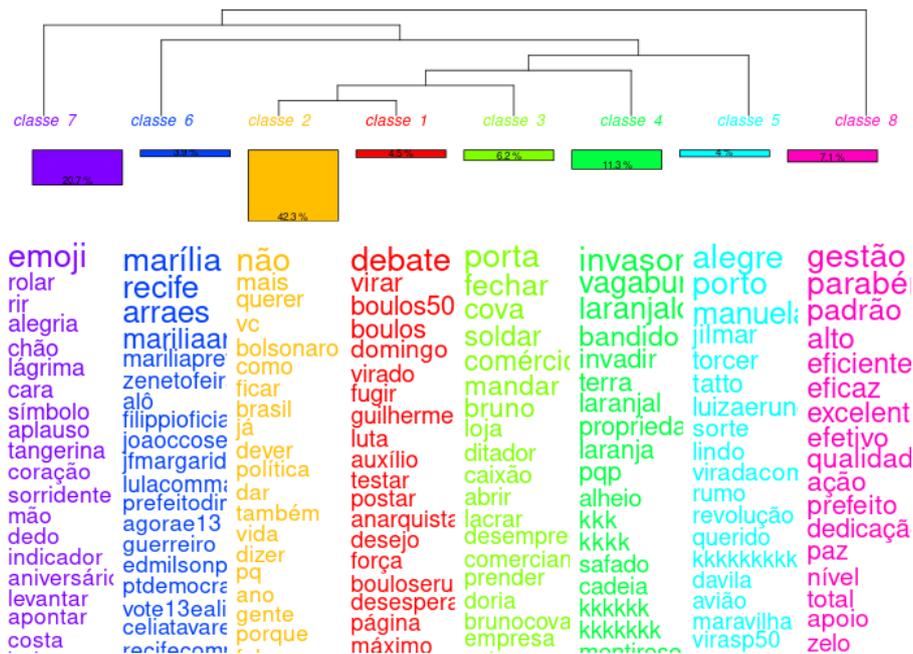


Figura 4. Vocabulários predominates em cada *cluster*.

utilização de detectores capazes de apontar quais perfis têm maior probabilidade de serem *bots*.

Sobre as eleições nos Estados Unidos de 2016, por exemplo, Heredia et al. [2018] fazem uma análise sobre a influência que *bots* exerceram na opinião pública, considerando mais de 700 mil usuários no Twitter. Os autores utilizam a ferramenta Botometer para identificação de *bots* associando-a à análise de sentimento através de redes neurais convolucionais. Dentre os resultados, os autores mostram o papel dos *bots* influenciando positiva ou negativamente as candidaturas analisados. Também dentro do tema eleições, Morstatter et al. [2018] analisam o papel exercido por *bots* nas eleições de 2017 na Alemanha. Usando dados do Twitter, eles identificaram que pelo menos 11% dos usuários analisados eram *bots*. Os autores também fazem uma análise de redes para identificação de comunidades, a fim de entender como ocorre o fluxo de informação entre elas. Rofrío et al. [Rofrío et al. 2019] apresentam um estudo sobre a presença de *bots* nas eleições de 2017 no Equador. Analisando 30 mil *bots* eles evidenciam como essas contas foram utilizadas para favorecer determinados candidatos.

Focado no contexto brasileiro, Ruediger et al. [2017] analisam a presença de *bots* no Twitter a partir de um método que considera o intervalo de publicação de *tweets* e a utilização de plataformas de automatização para publicação de conteúdo. No relatório, os autores apresentam estudos de caso que incluem temas relacionados, por exemplo, às eleições de 2014 e debates políticos correntes. Através de mapas de interações entre usuários, o trabalho pôde identificar como ocorre o comportamento de contas que publicam *tweets* de forma automatizada.

No que diz respeito aos mecanismos existentes para detecção de *bots*, a ferramenta Botometer [Varol et al. 2017] é uma das pioneiras e tem como base a utilização de

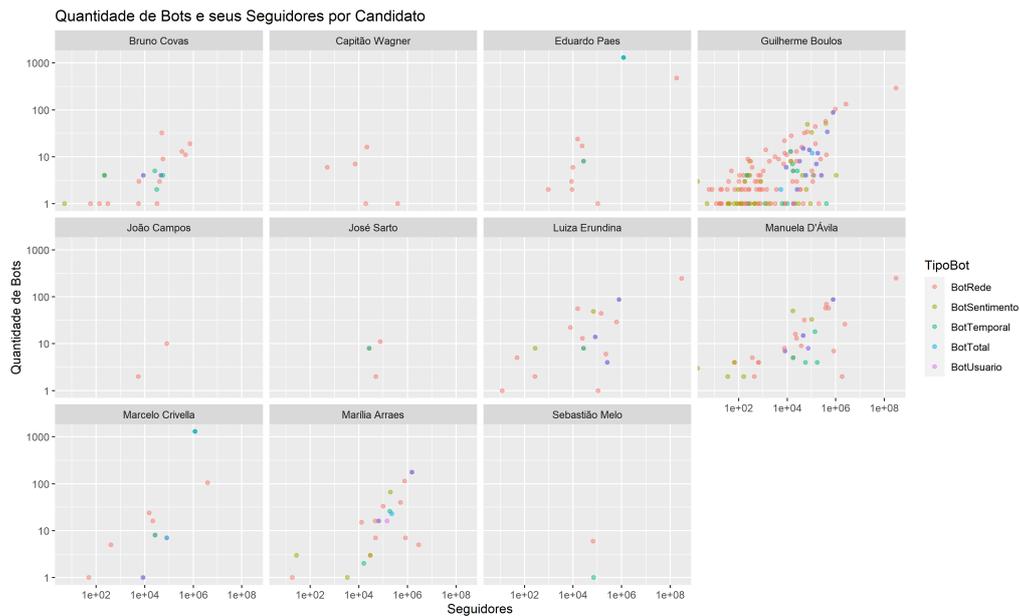


Figura 5. Quantidade de *bots* e volume de seguidores por candidato.

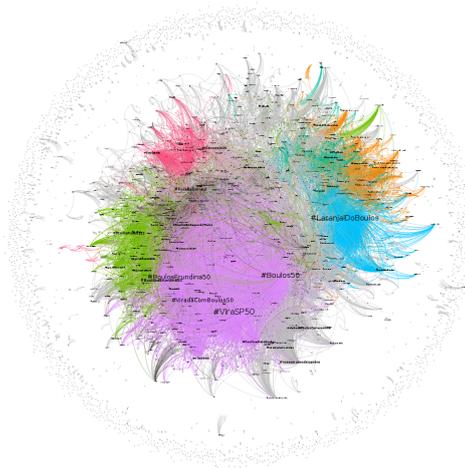


Figura 6. Menções à *hashtags*.

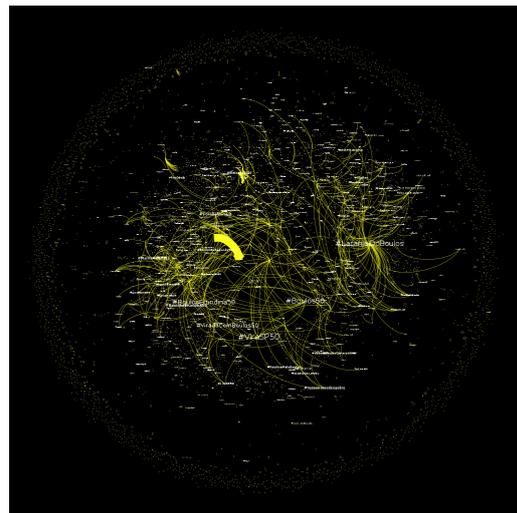


Figura 7. Menções à *hashtags* mobilizadas por *bots*.

algoritmos de aprendizagem de máquina para identificação de contas suspeitas no Twitter. Seu modelo para detecção de *bots* foi construído a partir de mais de mil atributos e metadados das contas, considerando como base de avaliação uma amostra classificada de usuários ativos e falantes de inglês. A ferramenta BotSentinel [Bouzy 2018] é um outro exemplo que também utiliza modelos de aprendizagem de máquina para detecção de *bots*, porém focado na identificação de atividades que violam as políticas de uso estabelecidas pelo Twitter. Em [Leite et al. 2020], os autores implementam um modelo baseado em regras para a identificação de *bots*. Os resultados obtidos apontam para distinções no comportamento desses usuários, mas não consideram campos textuais, que podem ser re-

levantantes para compreender sobre as funções dos *bots*. Vale mencionar que não é parte do escopo deste trabalho avaliar e comparar os resultados obtidos por diferentes ferramentas de detecção de *bots*, mas explorar sobre a importância de considerar um grande número de critérios na análise mais ampla sobre a atuação desses perfis na rede.

Em comparação às demais abordagens mencionadas, a metodologia proposta neste trabalho busca incluir múltiplos critérios para a identificação de *bots* no contexto brasileiro, de forma mais transparente. Utilizamos o Pegabot por ser uma ferramenta pensada no contexto brasileiro e que fornece resultados para diferentes critérios: usuário, redes, sentimento e temporal. Buscamos entender as especificidades de cada um desses critérios de identificação de *bots*, obtidos a partir de dados extraídos das contas analisadas do Twitter, bem como as características dos comportamentos coordenados. Propomos a realização tanto de uma análise estrutural de rede, quanto da análise lexical que considera o conteúdo que é compartilhado por contas com comportamento total ou parcialmente automatizado. Considerando também o cenário das eleições em diferentes capitais, entendemos que esses dois tipos de análise nos permitem, por exemplo, explorar as diferenças na interação de *bots* com os perfis analisados e os reflexos do comportamento coordenado nos possíveis *clusters* de *bots* e nos padrões de vocabulário em cada pleito.

5. Conclusões e trabalhos futuros

O presente trabalho avança ao analisar e viabilizar comparações sobre o uso de *bots* e propaganda computacional nas eleições em diferentes capitais brasileiras, auxiliando a superação de vieses regionais das pesquisas sobre *bots* no país. Segundo Woolley e Howard [Woolley and Howard 2018], “propaganda computacional” se refere ao uso de algoritmos, automação e curadoria humana para deliberadamente distribuir informação enganosa nas mídias sociais.

Identificamos não apenas a participação de *bots* em algumas capitais pouco exploradas (notadamente Porto Alegre e Recife), mas também como a utilização de diferentes critérios para identificação de *bots* poderia descrever esse cenário de modos consideravelmente diferentes (a seleção do tipo de *bots* a ser considerado pode dar preponderância a algumas capitais em detrimento de outras). Isso confirma a necessidade de uma perspectiva que una essas diferentes análises e possibilite definições mais precisas sobre tipos de *bots* em questão, e não apenas identificações genéricas.

A maior clareza quanto às diferenças entre os resultados de cada uma dessas análises de comportamento automatizado e as relações dos *bots* com a discussão eleitoral brasileira no Twitter são ganhos fundamentais para proteção de nossos processos eleitorais. Embora diferentes identificadores de *bots* apresentem leituras diversas sobre a probabilidade, quantidade e atuação indevida desses perfis automatizados no cenário político, investigadores ainda não têm acesso a dados que indiquem o motivo dessas divergências entre identificadores e, portanto, a confiabilidade de suas próprias análises baseadas nessas identificações. Dessa forma, nossa agenda de pesquisa buscou jogar luz sobre a relevância dessa questão, trazer mais segurança e confiabilidade às conclusões dos atores envolvidos na discussão sobre o tema.

Vale mencionar que este é um trabalho em andamento e que, além de nos aprofun-

darmos nas análises aqui apresentadas, vislumbramos a ampliação do estudo para todas as capitais do país, bem como a possibilidade de recomendações para o combate ao uso indevido de propaganda computacional na eleição presidencial de 2022 e de um mecanismo que apoie o monitoramento desse tipo de propaganda computacional durante o pleito. Além disso, um esforço para construção de uma base de dados de perfis classificados, considerando o cenário brasileiro, tem sido feito no contexto do projeto Pegabot, assim como o aprimoramento da ferramenta, a fim de possibilitar uma avaliação mais detalhada e aprofundada da acurácia e corretude do algoritmo de classificação.

Agradecimentos

Agradecemos a União Europeia pelo financiamento do projeto Pegabot. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES).

Referências

- Almeida, Y. L., Rubin, F. S., de Faria Alvim, A. C., Dias, V. M. F., and dos Santos, R. P. (2020). O uso das redes sociais para interferir nas democracias: Um mapeamento sistemático da literatura. In *Anais do IX Brazilian Workshop on Social Network Analysis and Mining*, pages 178–183. SBC.
- Bouzy, C. (2018). Botsentinel: <https://botsentinel.com/>.
- Leite, M., Guelpele, M., and Santos, C. (2020). Um modelo baseado em regras para a detecção de bots no twitter. In *Anais do IX Brazilian Workshop on Social Network Analysis and Mining*, pages 37–48. SBC.
- Rodriguez, A., Sava, P. S., Ituassu, A., and Lifschitz, S. (2019). Sistema web crawler para coleta automática de tweets, persistência em bancos de dados e análises estatísticas. In *Proceedings Companion 34o Simpósio Brasileiro de Bancos de Dados (SBBDD), Workshop de Trabalhos de Alunos de Graduação (WTAG)*, pages 325–332. SBC.
- Rofrío, D., Ruiz, A., Sosebee, E., Raza, Q., Bashir, A., Crandall, J., and Sandoval, R. (2019). Presidential elections in ecuador: Bot presence in twitter. In *2019 Sixth International Conference on eDemocracy eGovernment (ICEDEG)*, pages 218–223.
- Ruediger, M. A., Grassi, A., Freitas, A., Contarato, A. d. S., Taboada, C., Carvalho, D., Ferreira, H., Silva, L. R. d., Lenhard, P., Bastos, R., et al. (2017). Robôs, redes sociais e política no brasil: estudo sobre interferências ilegítimas no debate público na web, riscos à democracia e processo eleitoral de 2018. FGV DAPP.
- Santos, J. G., Ituassu, A., Lifschitz, S., Guimarães, T., Cerqueira, D., Albu, D., Fernando, R., Ferreira, J., and Mondelli, M. L. (2021). Das milícias digitais ao comportamento coordenado: métodos interdisciplinares de análise e identificação de bots nas eleições brasileiras. In *Anais do X Brazilian Workshop on Social Network Analysis and Mining*, pages 187–192. SBC.
- Varol, O., Ferrara, E., Davis, C., Menczer, F., and Flammini, A. (2017). Online human-bot interactions: Detection, estimation, and characterization. In *Proceedings of the international AAAI conference on web and social media*, volume 11, pages 280–289.

Woolley, S. C. and Howard, P. N. (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press.