FakeSpread: Um framework para análise de propagação de fake news na Web

Anderson Cordeiro[1], Jonice Oliveira[1], Livia Ruback[2]

1 - Universidade Federal do Rio de Janeiro (UFRJ)
2 - Universidade Federal Rural do Rio de Janeiro (UFRRJ)

andersoncordeiro@iprj.uerj.br, jonice@dcc.ufrj.br, liviaruback@ufrrj.br

Abstract. *The advancement of social media as a communication tool drives the dissemination of news, making it faster, scalable and consequently making it difficult to assess the veracity of information, making fake news a true phenomenon with increasingly serious developments in the political, social and human context. Identifying the source of a fake news and understanding the paths that make it popular is an important strategy to combat the fake news industry. The objective of this work is to analyze the spread of fake news on the Web, using graph theory metrics, considering its network structure. We present as a case study an analysis of the fake news propagation by websites listed by the CPMI of fake news, proposed by the Brazilian Congress.*

Resumo. *O avanço das mídias sociais como ferramenta de comunicação impulsiona a disseminação de notícias, tornando-a mais rápida, escalável e consequentemente dificultando a avaliação quanto a veracidade de uma informação, fazendo com que as fake news sejam um fenômeno com desdobramentos cada vez mais sérios no contexto político, social e humano. Identificar a procedência de uma notícia falsa e compreender os caminhos que a tornam popular é uma importante estratégia de combate à indústria de fake news. O objetivo deste trabalho é analisar a propagação de fake news na Web, a partir de métricas da teoria dos grafos, considerando a sua estrutura de rede. Apresentamos como estudo de caso uma análise da propagação de fake news por sites listados pela CPMI das fake news, proposta pelo congresso brasileiro.*

1. Introdução

Com o avanço das mídias sociais, a produção e o envolvimento midiático atingiu uma escala sem precedentes. Por sua natureza distribuída e descentralizada, as mídias sociais possuem atributos extremamente favoráveis para a propagação de fake news. Como resultado, quando uma informação falsa se propaga por esses meios, pode ter impactos sociais, políticos e econômicos, muitas vezes irreversíveis. No Brasil, a propagação de *fake news* já foi responsável por tragédias como no caso ocorrido no município de Guarujá/SP, onde a dona de casa Fabiane Maria de Jesus, de 33 anos, foi linchada até a morte por moradores que a confundiram com o retrato falado de uma suposta sequestradora de crianças [Ribeiro 2014], e também esteve presente em momentos políticos importantes como a campanha para o primeiro turno das eleições de 2018 no Brasil, onde uma grande quantidade de notícias falsas

circulou por sites e redes sociais e em aplicativos de troca de mensagens [Tardaguila & Mares 2018] ou na votação da abertura do processo de Impeachment da então presidenta Dilma Rousseff, onde três das cinco notícias mais compartilhadas no Facebook eram falsas [Lavarda, Sanchotene & Silveira 2016].

Mesmo com a crescente atuação de agências de fact-checking e iniciativas de *letramento digital*, que visam oferecer informações educativas sobre a verificação de notícias, os esforços colaborativos de combate às fake news não conseguem impedir a proliferação de fontes de notícias falsas, devido a forma como os links são compartilhados nas redes, dificultando a identificação da natureza dos conteúdos em circulação, cenário ideal para a difusão das fake news. Segundo Chen, Conroy e Rubin (2015): "Em redes sociais como o Facebook, um artigo do The New York Times se apresenta da mesma forma que um artigo do The Onion 2, e qualquer um pode vir com o certificado de confiança do amigo que o compartilhou".

Algumas abordagens para combater fake news utilizam técnicas computacionais, como o sistema *The Truthy,* cujo propósito é rastrear e identificar notícias falsas no contexto político [Ratkiewicz et al 2011] e a plataforma *Hoaxy* [Shao et al. 2016], que coleta notícias de diversas mídias sociais através de crawlers e APIs, armazenando-as em bancos de dados para aplicação de algoritmos de classificação. São poucas as soluções computacionais, porém, que visam identificar a procedência de uma notícia, como os trabalhos de Duong et al (2017) ou Baeth e Aktas (2018), sendo essa uma das forças motrizes por trás do surgimento de grande parte das fake news. Conhecer toda a cadeia existencial de uma fake news desde quando ela foi criada até tornar-se popular, bem como entender seu caminho percorrido é uma importante estratégia de combate à indústria de fake news, que lucra e gera influência por meio da desinformação [Gdi 2019].

A propagação de fake news na Web por produtores de conteúdo se dá a partir de links entre os sites. Neste contexto, para analisar a propagação de fake news na Web não podemos desconsiderar que a Web é um grafo, em que os nós são as páginas web e as arestas direcionadas são os hiperlinks que levam de uma página à outra [Newman 2010; Kleinberg & Easley 2010]. Neste trabalho, propomos um framework para analisar a propagação de fake news na Web, considerando a sua estrutura de rede, e utilizando conceitos e métricas de teoria dos grafos. Nossa proposta tem o potencial de mensurar a popularidade e a influência de sites que disseminam fake news e detectar padrões em sua disseminação. Como estudo de caso, apresentamos uma análise da propagação de fake news por sites listados pela CPMI das fake news, instaurada pelo congresso brasileiro.

Na seção 2, apresentamos conceitos essenciais da teoria dos grafos e como eles se relacionam com a estrutura de rede da Web. Na seção 3, introduzimos o *FakeSpread*, um framework genérico para análise de propagação de fake news na Web. Na seção 4, apresentamos um estudo de caso que analisa as relações existentes entre os sites divulgados pela CPMI das fake news. A seção 5 apresenta conclusões e trabalhos futuros.

2. Fundamentação teórica

Nesta seção apresentamos conceitos da teoria dos grafos, ramo da matemática que estuda as relações entre os objetos de um determinado conjunto. Estas definições são essenciais para o entendimento da abordagem proposta pelo framework *FakeSpread*.

2.1 Grafos, nós e arestas

Um *grafo* é uma forma de especificar relacionamentos entre uma coleção de itens e é formado por *nós*, alguns deles conectados através de *arestas* (ligações entre os nós). Os grafos podem ser direcionados - nos quais as arestas têm uma direção - ou não direcionados - nos quais a direção não é importante. A distinção entre um grafo direcionado e um não direcionado é um aspecto importante no estudo da teoria dos grafos, que permite analisar a direção dos relacionamentos entre os nós. A Figura 1 mostra um exemplo de grafo direcionado, composto por 5 nós e 8 arestas. Os números indicados em cada uma das arestas indica seu *peso*, a quantidade de vezes em que a ligação ocorre.

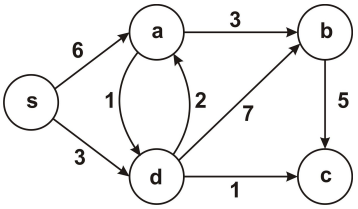

Figura 1. Exemplo de grafo com os nós e arestas com pesos

2.2 Links de entrada e saída, graus dos nós e pesos das arestas

Os *links de saída* de um nó são as arestas apontadas para um outro nó. Já os *links de entrada* para um nó são as arestas que apontam para ele. Na Figura 1, por exemplo, o nó *d* possui dois links de entrada (as arestas que partem dos nós *a* e *s*) e três links de saída (as arestas que chegam até os nós *a*, *b* e *c*). No caso da Web, os links de entrada de uma página são as páginas que apontam para ela e os links de saída são as páginas apontadas por ela.

O *grau* de um nó é uma métrica que representa o número de arestas conectadas a ele. Em uma rede social de amizade como o Facebook, por exemplo, o grau de um nó que representa um determinado perfil corresponde ao número de conexões associadas a este perfil, ou seja, o número de amigos conectados pela rede social. Geralmente, os nós com os maiores graus em uma rede - como o Facebook, a Web, entre outras - são aqueles que desempenham os papéis mais centrais e importantes na rede.

Em grafos não direcionados, o grau de um nó é apenas um número - que corresponde ao número de conexões com o nó. Porém, em grafos direcionados, deve-se computar os graus de entrada e graus de saída do nó - o número de arestas que apontam para ele e o número de arestas que ele aponta, respectivamente [Newman 2010]. No exemplo da Figura 1, o nó *d* tem grau de entrada 2 e grau de saída 3. No caso da Web, o grau de entrada de uma página é o número de páginas que apontam para ela e o grau de saída são as páginas para as quais ela aponta.

Outro importante atributo a ser observado são os pesos das arestas, representados por números na Figura 1 e que correspondem a quantidade de vezes em que um mesmo relacionamento entre o nó de origem e o nó de destino ocorreu. A aresta que liga o nó *d* ao nó *b*, por exemplo, tem peso 7, o que indica que a frequência que o nó *d* se conecta ao *b* é de 7.

3. O framework FakeSpread

Nesta seção apresentamos o framework *FakeSpread*, uma abordagem genérica para analisar a propagação de fake news na Web, considerando a sua estrutura de rede e utilizando conceitos e métricas de teoria dos grafos. Considerando a Web como um grafo, os nós são as páginas web, e as arestas direcionadas são os hiperlinks que levam de uma página à outra.

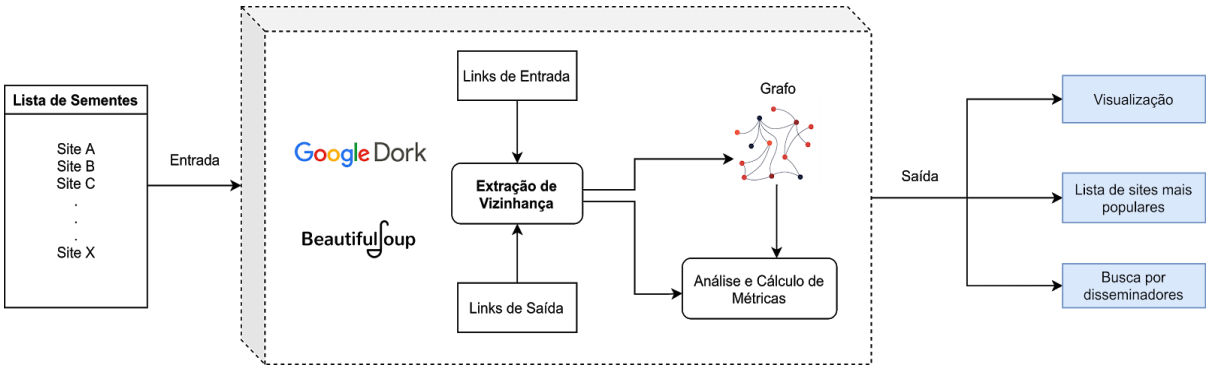

Figura 2. Arquitetura do framework *FakeSpread*

Como entrada, o framework recebe uma lista de *sementes*, isto é, sites que já propagaram ou são suspeitos de propagar fake news. A partir desta lista, o módulo de *Extração de Vizinhança* extrai os links de entrada e saída de cada site descobrindo quais são os nós mais próximos ao nó (site) observado. O reconhecimento da vizinhança de sites suspeitos ou confirmados de publicarem notícias falsas é fundamental para identificar a existência de uma rede onde sites fazem referência uns aos outros como forma de aumentar sua popularidade.

Os *links de saída* de um site são os sites que são apontados por ele e representam quais sites possuem alguma forma de apoio do site analisado ou quais referências o site analisado está utilizando para produzir ou replicar um conteúdo. A extração dos links de saída de um site pelo *FakeSpread* se dá a partir do acesso ao endereço principal, e da navegação entre as páginas internas, utilizando a biblioteca Beautiful Soup em Python. Para esta tarefa, o módulo de extração de vizinhança considera somente links apontando para outros sites de notícias e ignora links para outros recursos de mídia e publicidade. A identificação dos links de saída mostra somente quais sites são citados pelos sites *sementes.*

Já os *links de entrada* de um site são os sites que apontam para ele e tem um papel fundamental no rastreamento dos potenciais propagadores de fake news, pois mostram quais sites fazem referência aos sites *sementes*. A sua extração, porém, não é direta como a extração dos links de entrada. Para esta tarefa, o módulo de extração de vizinhança utiliza a biblioteca em Python GoogleWebSearch e técnicas de Google Hacking (ou Google Dorks), utilizadas para encontrar vulnerabilidades em sites a partir de buscas avançadas no Google. O módulo de extração de vizinhança pode escolher os operadores de busca avançados que forem mais convenientes, dentre a lista de operadores disponíveis pela plataforma[1].

A vizinhança extraída a partir dos links de saída e de entrada compõem o grafo com as conexões entre os sites sementes e as suas vizinhanças. O grafo é armazenado pelo *FakeSpread* como uma lista de conexões do tipo (*siteA*, *siteB*, *peso*). Tal tripla indica que o *siteA* aponta para o *siteB* e a conexão se repete *peso* vezes (ver seção 2.2). Dessa forma, se o *siteA* aponta para o *siteB* 3 vezes, teremos a tripla (*siteA*, *siteB*, *3*).

A partir do grafo de conexões gerado, o módulo de *Análise e cálculo de métricas* computa as métricas descritas na seção 2.2. Os graus dos nós, por exemplo, podem servir de indicativo de relevância de um site que propaga fake news. O grau de saída de um site representa o número de sites que ele aponta e o grau de entrada corresponde ao número de sites que apontam para ele. Quanto maior é o grau de entrada, maior é a chance da fake news

[1] http://www.googleguide.com/advanced_operators_reference.html

ser amplamente disseminada. Já os pesos das arestas indicam quantas vezes um *siteA* aponta para o *siteB* o que pode indicar uma relação de confiança ou parceria entre sites que propagam fake news.

4. Estudo de caso: A CPMI das Fake News

Nesta seção, instanciamos o framework *FakeSpread* em um estudo de caso que analisa a propagação de fake news por sites listados na CPMI das Fake News, instaurada pelo congresso brasileiro em setembro de 2019 para "investigar os ataques cibernéticos que atentam contra a democracia e o debate público; a utilização de perfis falsos para influenciar os resultados das eleições 2018" [BRASIL 2019].

Utilizamos como referência o relatório que identificou 47 sites de notícias falsas que receberam verba de propaganda pela Secretaria de Comunicação Social da Presidência da República (Secom) no período de junho e julho de 2020. Tal relatório, identificou canais considerados inadequados para veiculação de campanha patrocinada do governo, sendo o objetivo principal veicular informações sobre a Reforma da Previdência, que foi distribuída por meio da plataforma Adsense, que é um serviço de publicidade do Google, e que gera lucro por clique ou visualização. A Tabela 1 mostra os 22 sites apontados pelo relatório que ainda estão disponíveis para acesso.

Tabela 1. Sites propagadores de notícias falsas, segundo relatório da CPMI das Fake News

factschasers.com	imprensaviva.com	publicabrasil.com	notibras.com
gazetadopovo.com.br	revistaforum.com.br	politicanarede.com	conservadorismodobrasil.com.br
clickpolitica.com.br	opiniaocritica.com.br	gospelprime.com.br	direitapolitica.com
estadodedireita.com.br	antropofagista.com.br	diariodobrasil.org	tercalivre.com.br
jornaldacidadeonline.com.br	estudosnacionais.com	opiniaocritica.com.br	
vejaisso.com	diariodocentrodomundo.com.br	conexaopolitica.com.br	

A lista de sites apresentada na Tabela 1 representa a entrada do framework *FakeSpread*, ou seja, a lista de *sementes*, a partir das quais o framework extrai a vizinhança e calcula as métricas (ver seção 3).

Extração da vizinhança

Para este estudo de caso, o módulo de extração de vizinhança utilizou as duas abordagens: a extração de links de saída e de links de entrada. Os links de saída foram extraídos a partir do acesso ao endereço principal de cada um dos sites semente e da navegação entre as páginas internas, utilizando a biblioteca Beautiful Soup. Nesta etapa, foram descobertos 542 novos sites, que não estavam na lista de sementes e que podem ser relevantes no entendimento de como as fake news se propagaram a partir delas, pois são apontados por elas. Os links de saída são armazenados pelo módulo em triplas da forma (*semente*, *novosite, peso*).

Já os links de entrada foram extraídos a partir de uma busca avançada pelo GoogleWebSearch utilizando o operador *allintext*, que busca por páginas web que contenham a expressão procurada no corpo do texto. A string utilizada neste estudo de caso foi [*allintext:"fonte: endereco_semente" -fake -farsas -youtube -facebook -twitter -medium -falsa -desmente -falso -mentira -boato*], que indica que estamos buscando sites que contenham a

string “fonte: *endereco_semente*” e que não contenham as strings listadas a seguir, para desconsiderar as páginas publicadas por agências de fact-checking que desmentem as fake news. Escolhemos este template para a busca avançada pois identificamos que muitos sites reproduzem o conteúdo completo de uma fake news publicado em um outro site e ao final da página acrescentam a string “fonte: endereco_semente”. Nesta etapa, consideramos os 50 primeiros resultados da busca por relevância e foram identificados mais 968 novos sites. Estes sites representam potenciais propagadores de fake news, pois fazem referência aos sites sementes. Os links de entrada são armazenados pelo módulo em triplas da forma (*novosite, semente, peso*) e compõem, junto com os links de saída, o grafo de conexões.

Análise e Cálculo das métricas

O módulo de *Análise e Cálculo das métricas*, a partir do grafo de conexões, pode gerar uma visualização do grafo de conexões, gerar uma lista dos sites mais populares do grafo (a partir dos graus de cada um dos nós), fazer uma busca pelos potenciais disseminadores (através dos links de entrada das sementes), entre outros.

Para a visualização do grafo de conexões, o módulo faz uma integração com o software Gephi. A Figura 3 mostra um subconjunto do grafo deste estudo de caso. Para a representação do layout do grafo, foi utilizado o algoritmo de Yifan Hu multinível [Hu 2005] que combina um modelo de força direcionada com uma técnica de grafos que reduz a complexidade dos mesmos. O grafo completo consiste em 1510 sites (nós) e 1884 links (arestas) e pode ser acessado em http://fakepedia.org/mapa.

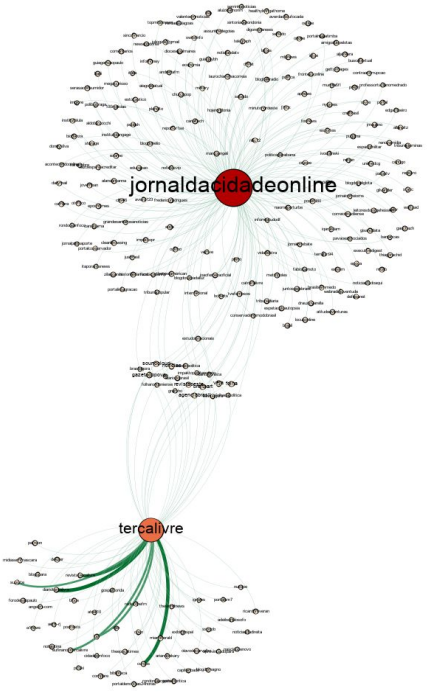

Figura 3. Clusters relacionados à dois dos sites mais populares

Os nós em destaque na Figura 3 (jornaldacidadeonline e tercalivre) representam dois dos sites mais referenciados no grafo, recebendo o maior número de links de entrada. Neste grafo não estão representados os links de saída de cada site. Os nós estão agrupados formando *clusters*, grupos de nós com alguma similaridade ou conexão. As cores e o tamanho de cada nó são proporcionais ao seu grau de entrada, ou seja, quanto mais links o nó recebe, maior e com uma cor mais forte ele é representado no grafo. Da mesma forma, o peso das

arestas é representado pela espessura das linhas e é proporcional à quantidade de conexões entre os nós.

Apontado na CPMI das Fake News como um dos sites que mais recebeu impressões de campanhas publicitárias do governo, o Jornal da Cidade Online também recebeu verba publicitária do Banco do Brasil segundo levantamento apurado pelo movimento Sleeping Giants Brasil [Siqueira 2020], que busca alertar grandes marcas sobre a divulgação de campanhas publicitárias em sites de conteúdo impróprio, com ênfase no combate a desinformação.

Já o Terça Livre, cujo proprietário foi alvo de uma operação da Polícia Federal, que investiga o financiamento de grupos que promovem atos contra a democracia [Shalders 2020], possui características diferentes se comparado ao Jornal da Cidade Online. Apesar de possuir um grau de entrada menor do que o Jornal da Cidade Online, o site Terça Livre possui arestas mais espessas com outros nós de seu cluster, como mostra a Figura 3. Essas arestas indicam o quão recorrente é a referência de um site a outro. Os nós cujas arestas possuem maior peso com o site Terça Livre (possuem mais links de referência) são: *eufinanciotercalivre.com.br*, *diariotercalivre.com.br* e *revistatercalivre.com.br*, todos com o domínio sob a mesma propriedade do site *tercalivre.com.br*.

As conexões e novos sites descobertos neste estudo de caso podem servir como base para uma investigação mais detalhada e que considera a rede de conexões entre os sites investigados pela CPMI. Além disso, o grafo gerado pode ser combinado com outros grafos gerados em outros cenários (por exemplo, outras listas de sites que propagam fake news divulgadas pela CPMI), de forma a se criar uma rede mais ampla de grafos. Tais grafos, combinados, podem explicitar conexões entre sites que divulgam fake news, como os vizinhos mais próximos aos nós semente, que também são populares.

5. Conclusão

Este trabalho teve como objetivo apresentar o framework *FakeSpread*, uma abordagem genérica para analisar a propagação de fake news na Web, que extrai os links de entrada e saída de um site para identificar outros sites com os quais ele se relaciona, considerando a sua estrutura de rede e utilizando conceitos e métricas de teoria dos grafos. Como estudo de caso, foi utilizado o relatório divulgado pela CPMI das Fake News que divulgou uma lista com 47 sites de notícias falsas que receberam verbas originadas de campanhas publicitárias do governo. A visualização do grafo gerado pelo framework demonstrou a presença de clusters de diferentes tamanhos na rede, destacando os sites mais populares e revelando novos sites, não indicados pela CPMI.

Como trabalhos futuros, pretende-se analisar os caminhos realizados por notícias que veiculam pela rede identificada, incluir novas métricas de análise de redes, utilizar o framework nos novos sites descobertos e integrar a solução com a Fakepedia, uma plataforma de fact checking que utiliza crowdsourcing e media literacy, possibilitando exibir dados de proveniência de fake news, disponibilizar a visualização dos grafos gerados e permitir a interação com as redes por meio de filtros e consultas.

Referências

Baeth, M. J., & Aktas, M. S. (2018). Detecting misinformation in social networks using provenance data.

BRASIL. Senado Federal. Comissão Parlamentar Mista de Inquérito - Fake News. 2019. Disponível em: https://legis.senado.leg.br/comissoes/comissao?0&codcol=2292. Acesso em: 06 ago. 2020.

Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. Proceedings of the Association for Information Science and Technology, 52, 1-4.

Duong, C. T., Nguyen, Q. V. H., Wang, S., & Stantic, B. (2017). Provenance-Based Rumor Detection. Databases Theory and Applications, 125–137.

Gdi, Global Disinformation Index (2019). The Quarter Billion Dollar Question: How is Disinformation Gaming Ad Tech?

Hu, Y. (2005). Efficient, high-quality force-directed graph drawing. Mathematica Journal, 10(1):37–71.

Kleinberg, J., Easley, D. Networks, Crowds, and Markets: Reasoning about a Highly Cambridge University Press, 2010.

Newman, M. J. Networks: An Introduction. [S.l.]: Oxford University Press, 2010.

Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Patil, S., Flammini, A., & Menczer, F. (2011). Truthy: Mapping the spread of astroturf in microblog streams. Proceedings of the 20th International Conference Companion on World Wide Web (WWW '11), 249–252.

Ribeiro, A. G. (2014). Mulher morta após boato em rede social é enterrada em Guarujá, SP. Disponível em <https://glo.bo/2DZXDuO> Acesso: 10.ago.2020.

Sanchotene, C., Machado da Silveira, A. C., & de Lima Lavarda, S. (2017). Quando as notícias mais compartilhadas são falsas: a circulação de boatos durante a semana do impeachment no Facebook. Comunicação & Informação, 20(3), 99-112.

Shalders, A. (2020). O que se sabe sobre o inquérito que levou às buscas e quebras de sigilo contra apoiadores de Bolsonaro. Disponível em <https://bbc.in/30WZYzC>. Acesso: 10.ago.2020.

Shao, C., Ciampaglia, G. L., Flammini, A., & Menczer, F.(2016). Hoaxy: A Platform for Tracking Online Misinformation.

Shu, K., Bernard, H. R., & Liu, H. (2018). Studying Fake News via Network Analysis: Detection and Mitigation. Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining, 43–65.

Siqueira, A. (2020). Sleeping Giants Brasil: "Governo dissemina ódio com dinheiro público" Disponível em <https://bit.ly/341AiDW>. Acesso: 10.ago.2020.

Tardaguila, C., Mares, C. (2018). Dez notícias falsas com 865 mil compartilhamentos: o lixo digital do 1º turno. Disponível em <https://bit.ly/3iIaSzc>. Acesso: 10.ago.2020.