VOL.40 Nº3 2023
https://doi.org/10.22491/rmct.v40i3.11622.pt

Detecção de Fake News nas Redes Sociais Virtuais: uma revisão dos métodos baseados em dados de propagação

Argus Antonio Barbosa Cavalcante[1], Paulo Márcio Souza Freire[1], Ronaldo Ribeiro Goldschmidt[1], Claudia Marcela Justel[1]
[1]**Instituto Militar de Engenharia**
Rio de Janeiro, RJ, Brasil
argus@ime.eb.br

RESUMO: A popularização das redes sociais virtuais aumentou significativamente a divulgação intencional de informações falsas, as chamadas Fake News. Esse tipo de notícia pode causar uma grande variedade de impactos negativos à sociedade, o que reforça a necessidade do desenvolvimento de métodos computacionais que possam detectar esse tipo nocivo de notícia. Entre os métodos computacionais existentes, destacam-se aqueles que utilizam os dados de propagação da notícia, tendo em vista sua relação com a natureza disseminava das informações existentes nas redes sociais. Assim, este trabalho apresenta uma revisão do estado da arte dos métodos de detecção de Fake News baseados em dados de propagação da notícia. Ao introduzir os conceitos básicos relacionados a esses métodos, além de apresentar e discutir suas principais características, espera-se fornecer subsídios para a criação de novos trabalhos na área.

PALAVRAS-CHAVE: Redes Sociais Virtuais. Informações falsas. Fake News. Dados de propagação. Métodos computacionais.

ABSTRACT: The popularization of virtual social media has significantly increased the intentional dissemination of false information, known as Fake News. This type of news can cause a wide range of negative impacts on society, reinforcing the need for developing computational methods that can detect this harmful type of news. Among the existing computational methods, those that use news propagation data have been highlighted given their relation with the disseminative nature of information on social media. Thus, this study describes a state-of-the-art review of Fake News detection methods based on news propagation data. By introducing the basic concepts related to these methods and describing and discussing their main characteristics, the authors intend to provide subsidies for the creation of new research in this area.

KEYWORDS: Virtual Social Networks. Disinformation. Fake News. Propagation data. Computational tools.

1. Introdução

Embora a divulgação de notícias esteja historicamente ligada aos meios tradicionais de comunicação, como jornais, revistas e televisão, recentemente as redes sociais virtuais (RSVs) se consolidaram como um importante meio de divulgação de notícias para os usuários [1]. Entre os motivos para essa mudança, pode-se destacar a relevância e a popularidade das RSVs na atualidade, bem como a existência de um processo simplificado de publicação de conteúdo nessas redes [2]. Esse processo, combinado a grande variedade de mecanismos de interação com o conteúdo publicado, acaba por impulsionar a audiência das publicações e consequentemente das próprias RSVs.

Nesse contexto, além das notícias reais, baseadas na apuração jornalística dos fatos, tipicamente feita pela imprensa, é possível também, encontrar diversos exemplos de informações falsas divulgadas de forma intencional, chamadas de *Fake News* [2]. Diversos episódios ocorridos nos últimos anos atraíram atenção para esse tipo de informação nas RSVs, entre os quais podemos destacar:

- Durante e logo após a ocorrência do furacão Katrina em 2005, diversas notícias falsas foram divulgadas com temas que foram desde a origem e explicação do fenômeno atmosférico, até quais eram as áreas afetadas e os impactos gerados na vida das pessoas[1];
- Por ocasião dos atentados terroristas à Maratona de Boston, ocorrida nos Estados Unidos em 2013,

1 Cf.: https://www.theguardian.com/us-news/2015/aug/16/hurricane-katrina-new-orleans-looting-violence-misleading-reports

informações falsas foram disseminadas, gerando medo e apreensão sobre o fato[2];

- Nas eleições presidenciais americanas de 2016, disputadas pelo candidato republicano Donald Trump e a candidata democrata Hillary Clinton, *Fake News* foram difundidas com o objetivo de distorcer fatos e opiniões relacionados a temas da campanha em favor de um ou outro candidato[3];- Mais recentemente, durante a pandemia do covid-19, houve divulgação de informações falsas com temas que abrangeram a origem da doença, o potencial de uso e aplicação de remédios, além da eficácia das vacinas[4].

Episódios como esses podem gerar diferentes impactos na sociedade, como, por exemplo, o descrédito da imagem pública de pessoas e instituições, problemas envolvendo a segurança pública e até mesmo um eventual sentimento coletivo de insegurança [3]. Sua ocorrência frequente atrai a atenção da sociedade, despertando, dessa forma, o interesse da academia e da indústria em entender e analisar a ocorrência desse fenômeno.

Devido ao grande volume e velocidade de publicações nas RSVs, e ainda, diante da dificuldade de avaliar manualmente a veracidade das notícias, torna-se cada vez mais comum o uso de métodos computacionais para o combate a esse tipo de conteúdo [2]. Entre os principais métodos para a detecção de *Fake News*, aqueles relacionados ao uso dos dados de propagação da notícia, recentemente se destacam por sua relação intrínseca com a natureza disseminativa das informações existentes nas RSVs. Esses métodos utilizam os dados relacionados às interações dos usuários que ocorrem a partir da publicação da notícia, como no caso do compartilhamento de uma publicação, um comentário ou outros tipos de reação ao conteúdo publicado [1].

Apesar dos resultados promissores obtidos pelos métodos de detecção de *Fake News* que utilizam os dados de propagação da notícia, até onde foi possível observar, ainda são poucos os trabalhos de revisão do estado da arte que exploraram [4] suas principais características. Dessa forma, este trabalho de pesquisa pretende abordar alguns aspectos importantes não abordados por trabalhos anteriores.

Além da presente introdução, o artigo contém três seções adicionais. A *Seção 2* apresenta breve fundamentação teórica de assuntos inerentes ao tema da revisão. Na *Seção 3*, são apresentados os trabalhos relacionados e o modelo utilizado para a sua comparação. Finalmente, na *Seção 4* são apresentadas as conclusões dessa revisão, além de algumas considerações adicionais.

2. Conceitos Básicos

2.1 Redes Sociais Virtuais

Nos últimos anos, as RSVs se popularizaram de forma massiva, fazendo com que milhões de usuários as utilizem todos os dias [5]. As RSVs são compostas por pessoas, entidades, empresas e organizações que se conectam e interagem entre si, em um ambiente colaborativo que oferece diversas facilidades de comunicação para seus usuários. Dada a grande variedade de RSVs existentes, houve uma segmentação dessas redes em função do conteúdo publicado e do público que pretendem atingir. Por exemplo, enquanto o conteúdo do TikTok[5] consiste principalmente de vídeos e é voltado para o público jovem, o Linkedin[6] tem o foco no mundo corporativo, sendo amplamente utilizado por empresas e profissionais das mais variadas indústrias.

As relações nas RSVs se estabelecem com base em princípios de influência social, similares aqueles existentes nas redes sociais não virtuais [6]. Quando pessoas estão conectadas em uma RSV, torna-se possível influenciar o comportamento e as decisões de outras pessoas. Essa influência pode ser observada

2 Cf.: https://www.bbc.com/news/world-us-canada-39930236

3 Cf.: https://www.washingtonpost.com/news/the-fix/wp/2018/01/03/how-hillary-clinton-might-have-inspired--trumps-fake-news-attacks/

4 Cf.: https://edition.cnn.com/2021/11/09/media/kaiser--covid-misinformation/index.html

5 Cf.: https://www.tiktok.com/

6 Cf.: www.linkedin.com

em diversas situações, como, por exemplo, nas opiniões emitidas a respeito de produtos comprados, na demonstração de posições políticas, nas informações sobre atividades realizadas por um determinado usuário, ou mesmo em um lugar por ele visitado. No contexto das RSVs, as pessoas normalmente tomam decisões com base em suas inferências a partir do que outras pessoas fizeram. Esse comportamento é conhecido como *herding behavior* [6] e ocorre na forma de um processo sequencial de tomada de decisão dos usuários de uma rede. A popularidade das RSVs, somada à influência social entre seus usuários, reforçam a necessidade do estudo das *Fake News* nessas redes.

2.2 Informações Falsas

O estado da arte caracteriza os diferentes tipos de informações falsas como *Misinformation* e *Disinformation* [2]. No primeiro tipo, enquadram-se as informações que são acidentalmente falsas, em consequência de erros. Já no segundo tipo, estão as informações conhecidamente falsas, divulgadas de forma intencional. Ou seja, a intencionalidade se estabelece como critério fundamental utilizado para distinguir os dois tipos de informações falsas. Nesse contexto, para que uma notícia seja considerada uma *Fake News* sua divulgação de forma intencional deverá ser confirmada.

Vale ressaltar que as *Fake News* encontradas nas RSVs fazem parte de um sistema composto por diversos participantes e que extrapola essas redes. Entre os participantes mais relevantes estão a imprensa, as agências de checagem de fatos, as pessoas publicamente expostas, os poderes legislativo e judiciário, além dos próprios usuários das RSVs [2]. Embora nem todos os tipos de informações falsas sejam necessariamente notícias, conforme citado anteriormente, a imprensa possui um papel importante no sistema de divulgação de informações falsas. Em algumas situações, o processo tradicional do trabalho jornalístico é descaracterizado pelas informações falsas, com a substituição da apuração cuidadosa dos fatos e a impossibilidade das partes envolvidas expressarem suas opiniões (muitas vezes de natureza contraditória). Com o objetivo de complementar o trabalho da imprensa no contexto das *Fake News*, surgiram as agências de checagem de fatos. Essas agências também monitoram as informações divulgadas nos meios tradicionais de comunicação, mas se concentram naquelas que circulam nas RSVs [1]. As informações analisadas, tipicamente, encontram-se em formato de artigos, notícias ou mesmo discursos e afirmações oriundos, majoritariamente, de pessoas públicas, dando preferência a assuntos de interesse geral que tenham ganhado destaque recente. Como resultado do seu trabalho, é obtida uma avaliação da informação analisada, com base no processo de checagem dos fatos que é realizado por suas equipes.

As pessoas publicamente expostas, tais como celebridades, políticos, influenciadores sociais e outras personalidades de destaque são, comumente, tanto alvos quanto disseminadores de informações falsas, tendo em vista que as informações relacionadas a essas pessoas obtêm, normalmente, maior alcance, são mais compartilhadas e discutidas [7]. Os poderes legislativo e judiciário possuem papel central nesse sistema. Com o aumento da divulgação de informações falsas nas RSVs, surgiu a necessidade de estudar e propor leis que possam punir os responsáveis pela sua divulgação, especialmente quando causam algum tipo de dano às partes envolvidas. Tanto no poder judiciário brasileiro, quanto em outros países, existem atualmente discussões a esse respeito. Na falta de leis específicas, o poder judiciário hoje, utiliza de leis gerais, como, por exemplo, as leis de danos materiais e morais decorrentes da eventual agressão à honra das vítimas de informações falsas [8].

2.3 Combate Automático às Fake News

Além do tratamento manual que pode ser realizado por uma agência de checagem de fatos, uma outra possibilidade tem sido muito explorada recentemente,trata-se do uso de recursos computacionais com o objetivo de detectar automaticamente as *Fake News*, evitando, dessa forma, a necessidade de detecção manual ou complementando-a quando necessário [1]. Essa alternativa tem se mostrado amplamente viável, em especial por conta do avanço no uso de métodos computacio-

nais por meio da aplicação de técnicas de mineração de dados e aprendizado de máquina [9], além do aumento da capacidade de processamento dos computadores que suportam sua execução.

No contexto do combate automático às *Fake News*, duas principais funcionalidades se consolidam: a detecção e a intervenção das *Fake News* [4]. Na detecção temos, em resumo, um problema de classificação binária. Considerando a avaliação de uma dada notícia, o resultado será verdadeiro caso seja uma *Fake News*, ou falso caso contrário. Visando evitar os efeitos nocivos da divulgação das *Fake News*, temos a funcionalidade de intervenção, na qual torna-se possível interromper a propagação de uma notícia determinada.

3 Estado da Arte

3.1 Seleção dos trabalhos

Para analisar o estado da arte dos métodos para detecção de *Fake News* que utilizam os dados de propagação da notícia, foi definido um conjunto de critérios para seleção dos trabalhos relacionados. Inicialmente, foram definidas as bases de artigos a serem consultadas. Devido a sua popularidade em temas relacionados à computação, foram selecionadas: Instituto de Engenheiros Eletricistas e Eletrônicos (IEEE)[7], *Association for Computing Machinery* (ACM)[8] e Science Direct[9]. Definidas as bases de artigos, procedeu-se a identificação de uma questão inicial de busca composta pelas palavras-chave relacionadas ao tema de propagação de informações falsas divulgadas de forma intencional, resultando em *"Fake News"+ propagation*. Além da questão de busca, foi utilizado o critério de identificar o sistema de "qualis" classificação dos artigos, privilegiando, sempre que possível, aqueles no estrato superior. Foram considerados trabalhos escritos após 2015 devido a carência de trabalhos publicados antes dessa data de corte.

7 Cf.: https://ieeexplore.ieee.org/

8 Cf.: http://dl.acm.org

9 Cf.: https://www.sciencedirect.com

3.2 Trabalhos Relacionados

Nesta seção serão apresentados os trabalhos selecionados para revisão nos quais foram desenvolvidos métodos computacionais para a detecção de Fake News com base em dados de propagação da notícia. Cada trabalho é descrito de forma resumida, incluindo os dados utilizados, o tipo de aprendizado, *datasets* e melhores resultados obtidos nos experimentos realizados por esses trabalhos. Detalhes adicionais desses trabalhos podem ser consultados por meio das suas referências:

- *News credibility evaluation on microblog with a hierarchical propagation model* [10]: os autores se baseiam na tríade notícia, usuário e comunidade para detecção de *Fake News*, propondo, como parte de seu trabalho, a criação de um modelo hierárquico que identifica sub-eventos dentro das notícias, de maneira a descrever seus detalhes. O problema da detecção automática é tratado como um problema de otimização de grafo, no qual uma ótima solução é proposta. Foram utilizados os dados de conteúdo do texto, tópicos de *hashtags*, links, entre outros. Esse trabalho se baseia na aprendizagem supervisionada. Os *datasets* utilizados são SW-2013 e SW-MH370. O melhor resultado obtido foi 0.889 (Acurácia).
- *CSI: A hybrid deep model for Fake News detection* [11]: o processo de detecção se divide em três partes: *Capture, Score e Integrate* (CSI). O primeiro módulo é baseado na resposta e no texto, por meio de uma rede neural recorrente, *long short-term memory* (LSTM), para capturar um padrão temporal de atividades do usuário sobre o artigo e a representação *doc2vec* do texto gerado nessa atividade. O segundo, usa uma rede neural para aprender as características da fonte baseado no comportamento dos usuários de acordo com suas interações, gerando um score por meio de um grafo. Os dois módulos são integrados com o terceiro para caracterizar, ou não, o artigo como *Fake News*. Foram utilizados dados dos usuários, sem especificar exatamente quais. Esse trabalho se baseia na aprendizagem supervisionada. Os *datasets* utilizados são Twitter e Weibo. O melhor resultado obtido foi 0.892 (Acurácia).

- *Tracing fake-news footprints: Characterizing social media messages by how they propagate* [12]: esse método utiliza a proximidade dos nós e as dimensões sociais para inferir uma representação dos usuários da rede social para, posteriormente, representar e classificar os caminhos de propagação de mensagens propagadas pelos referidos usuários. Foram utilizados os dados da estrutura de propagação da notícia. Esse trabalho se baseia na aprendizagem supervisionada. O *dataset* utilizado é uma adaptação do Twitter. O melhor resultado obtido foi 0.9380 (F-Measure).
- *Fake news detection on social media using geometric deep learning* [13]: explora padrões de propagação das Fake News com o uso de *Geometric Deep Learning* (generalização das técnicas de *deep learning* para dados não euclidianos como grafos que utilizam redes neurais convolucionais), utilizando quatro tipos de informações: perfil do usuário, atividade do usuário, rede e conteúdo. Foram utilizados os dados de configurações de perfil, linguagem, descrição do perfil, geolocalização, usuário verificado, conexões sociais entre usuários, e outros. Esse trabalho se baseia na aprendizagem supervisionada. Os *datasets* utilizados são adaptações de Twitter15 e Twitter16. O melhor resultado obtido foi 92.7% (ROC AUC).
- *Hierarchical propagation networks for Fake News detection: Investigation and exploitation* [14]: desenvolveu o conceito de propagação hierárquica da rede para detecção de Fake News. Para construir essa estrutura hierárquica, se baseia em dois níveis: 1. Macro-level (análise topológica e temporal da rede) e 2. Micro-level (análise topológica, temporal e linguística da rede). Foram utilizados os dados da estrutura de propagação da notícia. Esse trabalho se baseia na aprendizagem supervisionada. Os *datasets* utilizados são Politifact e GossipCop. O melhor resultado obtido foi 0.863 (Acurácia).
- *Network-based Fake News detection: A pattern-driven approach* [15]: propõe-se a investigar os padrões de disseminação de Fake News. De maneira geral, esses padrões podem ser classificados em três grupos: 1. Padrões relacionados a notícia, 2. Padrões relacionados aos divulgadores das notícias e 3. Padrões de relacionamento entre os divulgadores da notícia. Foram utilizados os dados da estrutura de propagação da notícia. Esse trabalho se baseia na aprendizagem supervisionada. Os *datasets* utilizados são Politifact e BuzzFeed. O melhor resultado obtido foi 0.929 (Acurácia).
- *GCAN: Graph-aware co-attention networks for explainable Fake News detection on social media* [16]: utiliza o conteúdo do *tweet* original e a sequência de *re-tweets* correspondentes para detectar a notícia como "*fake*" ou "*not fake*". Em primeiro lugar, extrai informações dos usuários de seus perfis e interações, aprendendo o *embedding* do conteúdo original. Depois, usa redes neurais convolucionais (CNN) e recorrentes (RNN) para aprender a representação da propagação dos *re-tweets*, construindo um grafo que modela as possíveis interações entre os usuários. Finalmente, estabelece a correlação entre o tweet original e sua propagação, gerando uma predição binária baseada nessa correlação. Foram utilizados os dados de geolocalização, usuário verificado, entre outros. Esse trabalho se baseia na aprendizagem supervisionada. Os *datasets* utilizados são Twitter15 e Twitter 16. O melhor resultado obtido foi 0.9084 (Acurácia).
- *Graph neural networks with continual learning for Fake News detection from social media* [17]: os autores consideram que as informações falsas e as verdadeiras se propagam de forma diferente na internet. Além das informações do fluxo, utilizam oito atributos das publicações (usuário verificado do Twitter, *timestamp* da criação do usuário, número de seguidores, número de amigos, número de listas, número de favoritos, número de status e *timestamp* dos tweets). Faz uso de redes neurais em grafos (GNN) para diferenciar padrões de notícias *fake* e não *fake*. Foram utilizados os dados de usuário verificado, data de criação de perfil, entre outros. Esse trabalho se baseia na aprendizagem supervisionada. Os *datasets* utilizados são FakeNewsNet (Politifact e GossipCop). O melhor resultado obtido foi 0.853 (Acurácia).
- Unsupervised Fake News Detection: A Graph-based Approach [18]: Método composto por três etapas: 1. Mineração de bi-cliques, com base em notícias compartilhadas de forma síncrona geran-

do um conjunto *Seed* (grafo heterogêneo – vértices do tipo usuário e artigo); 2. Mineração de bi-cliques dos artigos restantes (sem a restrição de compartilhamento síncrono), com o uso de similaridade do texto da publicação e informações dos usuários; e 3. Mineração dos artigos restantes (fora dos bi-cliques), com o uso de similaridade do texto da publicação e informações dos usuários. Foram utilizados os dados da estrutura de propagação da notícia. Baseia-se na aprendizagem não supervisionada. Os *datasets* utilizados são FakeNewsNet (Politifact e GossipCop). O melhor resultado obtido foi 0.800 (Acurácia).

- *Fake news detection in social networks via crowd signals* [19]: Motivados por uma ferramenta na época, recém introduzida na rede social Facebook, nesse trabalho foram utilizados os chamados *crowd signals* para realizar a detecção das *Fake News*. Esses sinais são a opinião explícita dos usuários indicando que determinado conteúdo é uma *Fake News*. Seus autores desenvolveram um algoritmo que realiza a inferência Bayesiana capaz de aprender ao longo do tempo sobre a capacidade dos usuários em opinar sobre esse tipo de conteúdo. Foram utilizados os dados de opinião dos usuários. Baseia-se na aprendizagem semi-supervisionada. O *dataset* utilizado é Facebook. O melhor resultado obtido foi 0.9967 (Acurácia).
- *Fake news detection based on explicit and implicit signals of a hybrid crowd: An approach inspired in meta-learning* [20]: inspirado em [19], esse trabalho também utiliza os chamados *crowd signals* para realizar a detecção de *Fake News*. Porém, de forma oposta ao trabalho inicial, que utiliza a chamada opinião explícita dos usuários, emitida por meio de uma ferramenta voltada para esse fim –, seus autores exploraram o uso da opinião implícita dos usuários, obtida através de seu comportamento ao interagir com esse tipo de conteúdo. Esse trabalho combina, ainda, a opinião implícita dos usuários com a de máquinas (algoritmos), formando o chamado crowd híbrido. Foram utilizados os dados da opinião implícita dos usuários, além da quantidade de interações e o texto da notícia. Esse trabalho se baseia na aprendizagem supervisionada. Os *datasets* utilizados são Gossip, Politifact, Gossip2, FakeNewsSet e FakeBR. O melhor resultado obtido foi 0.9989 (Acurácia).
- *Early detection of Fake News on social media through propagation path classification with recurrent and convolutional networks* [21]: esse trabalho utiliza a classificação do fluxo de propagação criado pela notícia para realizar a detecção de forma antecipada. Essa classificação é realizada a partir da modelagem do fluxo de divulgação das notícias como uma série temporal multivariada, na qual cada dupla de valores é um vetor numérico representando as características do usuário que participou do processo de propagação. Foram utilizados dados do usuário como suas amizades e seguidores, além da idade e características de geolocalização desses usuários, entre outros. Esse trabalho se baseia na aprendizagem supervisionada. Os *datasets* utilizados são Twitter 15, Twitter16 e Weibo. O melhor resultado obtido foi 0.921 (Acurácia).
- *Propagation2vec: Embedding partial propagation networks for explainable Fake News early detection* [22]: os autores atribuem diferentes níveis de importância para os elementos de propagação das notícias, reconstruindo o conhecimento sobre o fluxo completo dessas notícias com base em fluxos parciais da propagação. O método utiliza mecanismos de atenção hierárquicos para enfatizar as informações mais importantes existentes nesses fluxos. Foram utilizados os dados do usuário como: se ele é verificado, seus seguidores, amizades, listas, favoritos, *tweets*, entre outros. Esse trabalho se baseia na aprendizagem supervisionada. Os *datasets* utilizados são Politifact, Gossipcop. O melhor resultado obtido foi 0.897 (Acurácia).
- *Detect rumors in microblog posts using propagation structure via kernel learning* [23]: cria inicialmente uma representação da propagação através de árvores para, na sequência, capturar padrões que diferenciam os tipos de rumores por meio da avaliação da similaridade de sua estrutura (kernel-based method). Embora descreva a detecção de rumores, a saída desse método tam-

bém aponta uma notícia como "*fake*" ou "*not fake*" (*non-rumor, false rumor, true rumor or unverified rumor*). Foram utilizados os dados do usuário como seus seguidores e amizades, status de verificação da conta, histórico de posts, entre outros. Esse trabalho se baseia na aprendizagem supervisionada. Os *datasets* utilizados são adaptações de Twitter15 e Twitter16. O melhor resultado obtido foi 0.750 (Acurácia).

3.2 Modelo comparativo

De forma a proceder com a análise do estado da arte, foram definidos três critérios de comparação entre os trabalhos selecionados. São eles: os dados de propagação utilizados, a indicação sobre o uso de dados de acesso restrito e a realização, ou não, da detecção de *Fake News* de forma antecipada. A escolha desses critérios visa aprofundar o entendimento sobre as principais características e limitações existentes nos trabalhos selecionados.

Existe grande variedade de dados de propagação utilizados pelo estado da arte. No trabalho de [4], foi estabelecida uma classificação desses dados considerando as seguintes categorias: contribuição, usuário, assunto, temporalidade e rede. A contribuição de um usuário diz respeito aos diferentes tipos de mídias (por exemplo, texto, áudio ou imagens) utilizados durante a propagação. Em relação ao usuário propagador da notícia, são obtidos os dados de seu perfil e sua reputação, que pode estar vinculada à sua capacidade de identificar, ou publicar *Fake News*. Outro aspecto relevante é o assunto da notícia, que visa explorar a relação da notícia em questão com assuntos populares e controversos. Em relação ao critério de temporalidade, utilizam-se dados sobre o tempo de cada uma das propagações da notícia, critério importante do ponto de vista da detecção automática de *Fake News*. Por último, temos também o critério da rede, no qual pretende-se identificar se foram utilizados dados dos relacionamentos de cada usuário que interagiu com uma dada notícia.

Com base na classificação exposta, a Tabela 1 sumariza a relação dos trabalhos selecionados com os tipos de dados de propagação utilizados. É possível observar que os tipos mais frequentemente utilizados são dados do usuário e sua rede de relacionamentos, além da temporalidade relacionada às interações que ocorreram durante a propagação da notícia. Os dados da contribuição também foram utilizados, ainda que em menor número.

Tabela 1 - Utilização de dados de propagação: C (Contribuição), U (Usuário), A (Rede), T (Temporalidade) e R (Rede).

Trabalho	C	U	A	T	R
[10]		X	X		
[11]		X		X	X
[12]		X		X	X
[13]		X		X	X
[14]	X	X		X	X
[15]	X	X			X
[16]	X	X		X	X
[17]	X	X		X	X
[18]	X	X		X	X
[19]		X			
[20]	X	X		X	X
[21]	X	X		X	X
[22]		X		X	X
[23]		X		X	X

Embora diversos dados de propagação sejam utilizados pelo estado da arte, sua aplicação prática pode ser inviabilizada diante de restrições de acesso a esses dados. Essas restrições se devem principalmente à consolidação do arcabouço jurídico da privacidade de dados [24], especialmente devido às leis como a *General Data Protection Regulation* (GDPR) e a Lei Geral de Proteção de Dados Pessoais (LGPD). Considerando o grande número de incidentes de segurança e a popularização do uso da internet em todo o mundo, houve grande debate sobre a necessidade de leis para auxiliar o combate aos crimes virtuais, que culminou na criação das referidas leis. Em resposta a esse movi-

mento, as RSVs criaram suas políticas de privacidade de dados, garantindo o tratamento adequado e restrição no acesso aos dados. Essas políticas se concentram principalmente em dados pessoais (por exemplo, nome, apelido, fotografia e localização geográfica) e dados sensíveis (religião, origem racial, entre outros exemplos). Os dados pessoais podem ser utilizados para identificar um usuário, já os dados sensíveis têm o potencial de causar discriminação as pessoas. Assim, trabalhos que dependem de dados de acesso restrito para realizar a detecção de *Fake News* possuem desvantagem significativa em relação a trabalhos que não dependam desses dados.

A Tabela 2 apresenta um resumo dos trabalhos relacionados em função da utilização dos dados de acesso restrito, sendo possível observar que há uma distribuição equivalente entre os trabalhos que dependem e aqueles que não dependem desses dados.

Tabela 2 - Utilização de dados de acesso restrito.

Trabalho	Utiliza Dados de Acesso Restrito ?
[10]	Sim
[11]	Não foi possível determinar.
[12]	Não
[13]	Sim
[14]	Não
[15]	Não
[16]	Sim
[17]	Sim
[18]	Não
[19]	Não
[20]	Não
[21]	Sim
[22]	Sim
[23]	Sim

Um desafio imposto aos métodos de detecção automática de *Fake News*, tanto para aqueles que utilizam os dados de propagação ou, eventualmente, outros tipos de dados, é o fato de que a propagação da notícia pode demorar horas, dias, semanas e até mesmo anos para ser concluída. A partir do momento em que a notícia é publicada, ela pode ficar disponível de forma permanente. Ou seja, usuários que porventura tiverem acesso a esse conteúdo, continuam a poder interagir com ele e assim, podem criar novos ciclos de propagação da notícia. Considerando que a espera de um tempo maior dificulta a implementação de ações de intervenção que objetivam evitar os impactos das *Fake News*, alguns métodos de detecção baseados em dados de propagação buscam realizar a detecção de forma antecipada [1]. Nesses trabalhos, a detecção deve ser realizada nos estágios iniciais da propagação, isto é, antes que essa notícia tenha sido amplamente divulgada. Esse critério também foi considerado como parte da análise do estado da arte.

Observando a Tabela 3, torna-se possível verificar que o uso da detecção antecipada por parte dos métodos de detecção de *Fake News* ainda é incomum.

Tabela 3 - Realização da Detecção Antecipada.

Trabalho	Realiza a Detecção Antecipada ?
[10]	Não
[11]	Não
[12]	Não
[13]	Não
[14]	Não
[15]	Não
[16]	Não
[17]	Não
[18]	Não
[19]	Não
[20]	Não
[21]	Sim
[22]	Sim
[23]	Sim

4 Conclusão

Nos últimos anos, o uso das redes sociais virtuais para divulgação de notícias se torna cada vez mais frequente. No entanto, essas redes são também utilizadas para a divulgação das Fake News, que são notícias intencionalmente falsas. Neste artigo foi feita a revisão do estado da arte com foco nos trabalhos que utilizam um tipo particular de dados relacionados com a propagação das notícias. Embora populares, poucas revisões [4] foram feitas visando avaliar especificamente esse tipo particular de dados. Nesse contexto, este artigo fez uma atualização em relação ao trabalho de revisão anterior [4], destacando os trabalhos de detecção de Fake News que utilizam dados de propagação. O resultado desta revisão aponta algumas observações relevantes que merecem ser destacadas. Em primeiro lugar, é amplo o espectro de tipos de dados de propagação utilizados pelo estado da arte. Já em relação ao uso de dados de acesso restrito, os trabalhos se dividem de forma equilibrada entre aqueles que dependem ou não dessa restrição. Finalmente, a detecção antecipada de Fake News, que busca habilitar ações de intervenção após a detecção da notícia, ainda é pouco explorada.

Referências

[1] ZHOU, X.; ZAFARANI, R. A survey of Fake News: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys*, New York, v. 53, n. 5, p. 1-40, 2020.

[2] SHARMA, K.; QIAN, F.; JIANG, H.; RUCHANSKY, N.; ZHANG, M.; LIU, Y. Combating Fake News: A survey on identification and mitigation techniques. *ACM Transactions on Intelligent Systems and Technology*, New York, v. 10(3), n. 21, p. 1-42, Apr. 2019.

[3] WATTS, D. J.; ROTHSCHILD, D. M.; MOBIUS, M. Measuring the news and its impact on democracy. *Proceedings of the National Academy of Sciences*, Madison, v. 118, n. 5, e1912443118, 2021. DOI 10.1073/pnas.1912443118

[4] FREIRE, P. M.; GOLDSCHMIDT, R. R. Combate automático às Fake News nas mídias sociais virtuais. *Revista Militar de Ciência e Tecnologia*, Rio de Janeiro, v. 38 n. 2, p. 3-12, 2021.

[5] EASLEY, D.; KLEINBERG, J. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge: Cambridge University Press, 2010.

[6] EASLEY, D; KLEINBERG, J. *Networks, Crowds, and Markets*: Reasoning about a Highly Connected World. Cambridge: Cambridge University Press, 2010.

[7] KUMAR, S.; CHENG, J.; LESKOVEC, J. Antisocial behavior on the web: Characterization and detection. *In*: INTERNATIONAL CONFERENCE ON WORLD WIDE WEB COMPANION, 26., 2017, Perth. *Anais* […]. Republic and Canton of Geneva: International World Wide Web Conference Committee, 2017. p. 947–950.

[8] FLUMIGNAN, W. G. G. As Fake News à luz da legislação brasileira. *Revista Científica Disruptiva*, Recife, v. 2 n. 2, p. 145-161, 2020.

[9] GOLDSCHMIDT, R.; PASSOS, E.; BEZERRA, E. *Data Mining*. 2. ed. Rio de Janeiro: Elsevier Brasil, 2015.

[10] JIN, Z.; CAO, J.; JIANG, Y. G.; ZHANG, Y. News credibility evaluation on microblog with a hierarchical propagation model. *In*: IEEE INTERNATIONAL CONFERENCE ON DATA MINING, 14., 2014, Shenzhen. *Anais* […]. Piscataway: Institute of Electrical and Electronics Engineer, 2024. p. 230–239.

[11] RUCHANSKY, N.; SEO, S.; LIU, Y. CSI: A hybrid deep model for Fake News detection. *In*: INTERNATIONAL CONFERENCE ON INFORMATION AND KNOWLEDGE MANAGEMENT, 26., 2017, Singapore. *Anais* […]. New York: Association for Computing Machinery, 2017. p. 797–806.

[12] WU, L.; LIU, H. Tracing fake-news footprints: Characterizing social media messages by how they propagate. *In*: INTERNATIONAL CONFERENCE ON WEB SEARCH AND DATA MINING, 11., 2018, Marina Del Rey. *Anais* […]. New York: Association for Computing Machinery, 2018. p. 637-645.

[13] MONTI, F.; FRASCA, F.; EYNARD, D.; MANNION, D.; BRONSTEIN, M. M. Fake News detection on social media using geometric deep learning. Appleton: *ICLR*, 2019.

[14] SHU, K.; MAHUDESWARAN, D.; WANG, S.; LIU, H. Hierarchical propagation networks for Fake News detection: Investigation and exploitation. *Proceedings of the International AAAI Conference on Web and Social Media*, [*s. l.*], v. 14, n. 1, p. 626-637, 2019.

[15] ZHOU, X.; ZAFARANI, R. Network-based Fake News detection: A pattern-driven approach. *SIGKDD Explorations*, [*s. l.*], v. 21, n. 2, p. 48-60, 2019.

[16] LU, Y.-J; LI, C.-T. GCAN: Graph-aware co-attention networks for explainable Fake News detection on social media. *In*: ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 58., 2020 , Online. *Anais* […]. Kerrville: Association for Computational Linguistics, 2020. p. 505–514.

[17] HAN, Y.; KARUNASEKERA, S.; LECKIE, C. *Graph neural networks with continual learning for Fake News detection from social media*. Nova York: Cornell GANGIREDDY, S. C. R. P. D.; LONG, C.; CHAKRABORTY, T. Unsupervised Fake News detection: A graph-based approach. *In*: HYPERTEXT AND SOCIAL MEDIA, 31., 2020, Online. *Anais* […]. New York: Association for Computing Machinery, 2020. p. 75–83.

[18] TSCHIATSCHEK, S.; SINGLA, A.; GOMEZ RODRIGUEZ, M.; MERCHANT, A.; KRAUSE, A. Fake news detection in social networks via crowd signals. In: COMPANION PROCEEDINGS OF THE THE WEB CONFERENCE 2018, *Anais* […]. Republic and Canton of Geneva: International World Wide Web Conferences Steering Committee, 2018. p. 571-524.

[19] SOUZA FREIRE, P. M.; MATIAS DA SILVA, F. R.; GOLDSCHMIDT, R. R. Fake news detection based on explicit and implicit signals of a hybrid crowd: An approach inspired in meta-learning. *Expert Systems with Applications*, [*s. l.*], v. 183, 115414, 2021.

[20] LIU, Y.; WU, Y.-F. Early detection of Fake News on social media through propagation path classification with recurrent and convolutional networks. Thirty-Second AAAI Conference on Artificial Intelligence, 32., 2018, New Orleans. *Anais* […]. Palo Alto: AAAI Press, 2018. p. 354-361.

[21] SILVA, A.; HAN, Y.; LUO, L.; KARUNASEKERA, S.; LECKIE, C. Propagation2vec: Embedding partial propagation networks for explainable Fake News early detection. *Information Processing Management*, [*s. l.*], v. 58, n. 5 102618, 2021.

[22] MA, J.; GAO, W.; WONG, K. Detect rumors in microblog posts using propagation structure via kernel learning. In THE 55TH ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 55., 2017, Vancouverv. 1. *Anais […]*. Kerrville: Association for Computational Linguistics, 2017. p. 708–717.

[23] FINKELSTEIN, M.E.; FINKELSTEIN, C. Privacidade e lei geral de proteção de dados pessoais. *Revista de Direito Brasileira*, São Paulo, v. 23, n. 9, p. 284-301, 2019.