



UNIVERSIDADE FEDERAL DO AMAPÁ
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO
DEPARTAMENTO DE CIÊNCIAS EXATAS E TECNOLÓGICAS

ANÁLISE DO ALGORITMO FAKECHECK EM NOTÍCIAS REGIONAIS DO AMAPÁ

RODRIGO SANTOS BALIEIRO

Orientador: Jose Walter Cárdenas Sotil

MACAPÁ
OUTUBRO DE 2019

RODRIGO SANTOS BALIEIRO

**ANÁLISE DO ALGORITMO FAKECHECK EM NOTÍCIAS
REGIONAIS DO AMAPÁ**

Anteprojeto de TCC apresentado à Universidade Federal do Amapá como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Jose Walter Cárdenas Sotil

MACAPÁ
OUTUBRO DE 2019

RODRIGO SANTOS BALIEIRO

**ANÁLISE DO ALGORITMO FAKECHECK EM NOTÍCIAS
REGIONAIS DO AMAPÁ**

Anteprojeto de TCC apresentado à Universidade Federal do
Amapá como requisito parcial para obtenção do título de
Bacharel em Ciência da Computação.

Trabalho aprovado. Macapá, de de 2019.

Jose Walter Cárdenas Sotil
Orientador

Msc. Marco Antônio Leal da Silva
Universidade Federal do Amapá

Esp. Adeildo Telles da Silva
Universidade Federal do Amapá

MACAPÁ
OUTUBRO DE 2019

Resumo

Com a popularização do acesso às redes sociais e o consumo de notícias em tempo real, as *Fake News* surgem como um grande fator que influencia cada vez mais a opinião pública e as decisões corporativas. Frente à grande disseminação em massa de conteúdo falso, se tem a tentativa de detectar a veracidade dessas informações com ferramentas virtuais como o *FakeCheck* – da USP. Assim sendo, questiona-se a efetividade de ferramentas como a *FakeCheck*, analisando-a e submetendo-a a testes específicos.

Palavras-chave: Fake news, notícias falsas, complexidade de algoritmos.

Abstract

Keywords: Fake News. Algorithm Complexity. Virtual Bots.

Sumário

1 – Introdução	1
2 – Contexto	2
3 – Problemática	4
4 – Objetivos	5
4.1 Objetivo Geral	5
4.2 Objetivos Específicos	5
5 – Hipótese	6
6 – Justificativa	7
7 – Metodologia	8
8 – Cronograma	9
Referências	10

1 Introdução

Nos últimos anos do séc. XXI, com o advento do uso expressivo de redes sociais[1], tanto por empresas, órgãos públicos ou mesmo por populares, as notícias e informativos estão presentes continuamente em diversos aparelhos, tornando a comunicação intensiva e dinâmica.

Entretanto, em meio a inúmeros textos, imagens e vídeos que "bombardeiam" o cotidiano das pessoas existem interesses comerciais, ideológicos, religiosos, dentre outros[2] que, passam conteúdos manipulados ou falsos para que sejam consolidados, em que a verdadeira intenção seja o compartilhamento dessas informações para que haja uma grande repercussão.

Uma expressão em inglês que tornou-se bastante debatida nos últimos anos é a "Fake News", que em uma tradução livre, quer dizer Notícias Falsas. Ela ganhou notoriedade após as eleições presidenciais de 2016, nos Estados Unidos da América[3], em que o candidato Donald Trump foi eleito em uma campanha tomada por acusações e ofensas entre os presidenciáveis, mas não pessoalmente e sim através das redes sociais. Ambos evidenciaram o fato de que "bombardeamento" das *Fake News* teriam sido feitas por robôs digitais, os chamados *Bots*. No Brasil, na última eleição presidencial, em 2018, teve-se uma experiência semelhante.

Sobre os robôs digitais, que são desenvolvidos com inteligência artificial e automação, pode-se questionar a complexidade de desenvolvimento de robôs digitais, bem como seus efeitos, sendo que o desafio estaria em identificá-los para conter a propagação de notícias falsas e auxiliar na tomada de decisões e opinião das pessoas que têm acesso às informações digitais.

Em meio a ferramentas criadas para detecção de notícias e conteúdos falsos, o pesquisador Rafael Augusto Monteiro, da Universidade de São Paulo, desenvolveu um protótipo capaz de tomar como entrada um texto e devolver uma resposta se o mesmo é verdadeiro ou falso. Diante disso, o presente trabalho busca analisar a efetividade da pesquisa de Monteiro em um contexto específico de notícias regionais, analisando seu algoritmo.

2 Contexto

O filósofo Platão[4] falava em suas obras que a dúvida era essencial para chegarmos até a verdade. Seguindo o raciocínio, em meio a inúmeras fontes de informação presentes no cotidiano das pessoas é importante reconhecer o que de fato é real ou apenas especulação ou, até mesmo, a mentira, o que pode gerar a desinformação e influenciar na tomada de decisões.

As redes sociais estão cada vez mais presentes entre as fontes de conteúdo provenientes da internet[5]. De acordo a pesquisa científica do Instituto Brasileiro de Geografia e Estatística (IBGE)[1], entre os anos de 2016 e 2017 o percentual de utilização da internet subiu de 69,3% para 74,9%, significando 3 em cada 4 domicílios brasileiros que possuem aparelhos conectados à internet, como tablets, celulares, smart TVs. Os usuários desses aparelhos consultam as redes sociais todos os dias para consumir informações de entretenimento e editoriais jornalísticos[5].

Fake News é um termo usado para definir notícias de conteúdo falso, enganoso, fabricado ou manipulado, cujas fontes ou a declaração citada do conteúdo da notícia não pode ser constatado como verdadeiro[6].

Uma vez conectadas à internet, as pessoas são expostas à uma série de fontes de informações que podem conter diversos assuntos generalizados ou especializados com um determinado fim, em que existe o agente produtor (origem), os usuários que compartilham (proliferação) e a maneira como são publicadas (tom)[7]. Geralmente as corporações governamentais, jornalísticas, publicitárias e comerciais se utilizam das redes sociais para apresentar seus conteúdos, entretanto, existe uma parte dessas fontes de informação produzida por pessoas mal intencionadas que podem praticar crimes virtuais, propagar ideologias, distorcer informações e confundir a realidade para quem recebe a informação[8].

Os pesquisadores Nir Kshetri e Jeffrey Voas[8] afirmam que as notícias falsas produzidas e mantidas de forma profissional têm consequências na economia, política e na sociedade, pois o consumo delas pode gerar um círculo vicioso em que o excesso da mentira torna-se verdade.

No ano de 2019 um ex-candidato à presidência da república no Brasil foi condenado por ter, supostamente, pago à empresas de mídia especializada, para impulsionarem notícias falsas que pudessem prejudicar a campanha de seu adversário, o que pode ser analisado como o poder público tem dado importância nos efeitos que o conteúdo digital pode causar quando manipulado profissionalmente[9, 10, 11].

A detecção de *fake news* passa por técnicas computacionais que compreendem

redes neurais, classificação de linguagem natural, engenharia reversa, autômatos, IA e diferentes tipos de busca. Os *Bots* da internet são programas que executam tarefas de forma automática. Inicialmente foram criados para facilitar respostas comuns a muitos questionamentos que as empresas recebiam, mas passaram a ser utilizados para criar perfis falsos e disseminar a desinformação nas redes sociais[12].

Com várias fontes de informação propagadas na internet existem cientistas e acadêmicos que desenvolvem protótipos de programas para a detecção de notícias falsas, utilizando diversas metodologias para obter um melhor resultado. Questiona-se a eficácia e a confiabilidade que os programas de detecção de "Fake News" conseguem.

3 Problemática

As corporações donas das redes sociais têm desenvolvido ferramentas e estratégias para detecções de notícias falsas, principalmente com colaboração da comunidade participante da própria rede social. Entretanto, conforme o teor e a forma em que tais notícias são repassadas, questiona-se a efetividade da detecção e neutralização das "Fake News" através dos algoritmos desenvolvidos, como é o caso da ferramenta desenvolvida pela USP, chamada de "FakeCheck".

4 Objetivos

4.1 Objetivo Geral

Verificar a eficácia do algoritmo de detecção de notícias falsas *Fake Check*, da USP.

4.2 Objetivos Específicos

O presente trabalho de pesquisa tem os seguintes objetivos específicos:

- Analisar a complexidade do algoritmo usada na ferramenta de detecção de Fake News "FakeCheck" nas modalidades propostas pelos desenvolvedores.
- Testar o algoritmo do software "FakeCheck" na detecção de notícias falsas com um banco de dados proposto.
- Comparar resultados de busca à "fake news" pelo programa "FakeCheck" e criar resultados estatísticos.
- Identificar os métodos de detecção de notícias falsas e compará-los pelos usados no programa "FakeCheck".

5 Hipótese

Os algoritmos vêm apresentando uma alta complexidade tanto na criação de robôs virtuais automáticos para propagar notícias falsas quanto na busca e detecção dos mesmos, tornando um processo de conflito pela identificação e neutralização de ações mal-intencionadas.

6 Justificativa

A busca por atividades virtuais maliciosas é uma prática antiga, desde a existência dos ataques virtuais, seja por meio de vírus, *worms*, *trojans*, *adwares*, dentre outras ferramentas criadas para prejudicar os usuários de aparelhos eletrônicos. Os antivírus foram criados para tal tarefa.

Com o passar dos anos e aumento significativo no acesso à internet, os golpes e os softwares maliciosos foram aperfeiçoados com novos recursos apresentados pela área de Tecnologia da Informação, passando a conter automação, inteligência artificial e computação em nuvem a fim de atingirem seus objetivos.

Com tantos recursos disponíveis para prejudicar o usuários das tecnologias da informação, tem-se um desafio da comunidade científica em identificar, prevenir e banir todo conteúdo malicioso com ferramentas de contra-ataque e estratégias de divulgação preventivas. Todavia necessita-se estudar e analisar essas ferramentas para avaliar sua eficácia e um melhor acesso a todo conteúdo consumido na internet, sobretudo nas redes sociais.

Se os resultados de uma pesquisa acerca da eficácia na busca por notícias falsas apresentarem resultados tanto positivos ou negativos, esta pesquisa poderá apresentar uma contribuição para as pesquisas tecnológicas.

7 Metodologia

Como o que se pretende fazer seja estudar tanto os softwares que impulsionam as notícias falsas quanto os que combatem-nas, uma metodologia será entrevistar alguns responsáveis pelo projeto que desenvolve sistemas web e de aplicativos.

O tipo de entrevista que se pretende é do tipo subjetiva, em que os questionamentos serão acerca do desenvolvimento dos softwares, bem como sua eficiência.

Uma comparação pode ser feita entre as ferramentas de detecção de *fake news* para analisar os resultados.

8 Cronograma

O presente trabalho é pautado na realização das atividades listadas, como segue:

1. Revisão bibliográfica;
2. Investigação dos objetivos;
3. Identificação de recursos técnicos;
4. Entrega do anteprojeto;
5. Análise do algoritmo.
6. Teste do algoritmo com banco de dados próprio.
7. Apresentação final do trabalho de pesquisa para a conclusão do curso.

As atividades acima são efetuadas conforme o cronograma descrito na Tabela 1.

Tabela 1 – Cronograma

Etapa	nov.	dez.	jan	fev.	mar
1	x	x	x	x	x
2	x	x			
3	x	x	x	x	
4				x	x
5					x
6			x	x	x
7					

Referências

- [1] I. B. de Geografia e Estatística, “Pnad contínua tic 2017: Internet chega a três em cada quatro domicílios do país.”
- [2] R. Christofletti, “Padrões de manipulação no jornalismo brasileiro: Fake news e a crítica de perseu abramo 30 anos depois,” vol. 12, Rumores, 2018.
- [3] T. Lee, “The global rise of “fake news” and the threat to democratic elections in the usa,” vol. 22, pp. 15–24, Emerald Publishing Limited, 2019.
- [4] M. Zingano, *Platão & Aristóteles*, vol. 2. Odysseus, 2002.
- [5] K. Shu, H. Liu, J. Han, L. Getoor, W. Wang, J. Gehrke, and R. Grossman, *Detecting Fake News on Social Media*. Morgan & Claypool, 2019.
- [6] S. A. Khan, M. H. Alkawaz, and H. M. Zangana, “The use and abuse of social media for spreading fake news,” in *2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*, pp. 145–148, June 2019.
- [7] S. B. Parikh, V. Patil, and P. K. Atrey, “On the origin, proliferation and tone of fake news,” in *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, pp. 135–140, March 2019.
- [8] N. Kshetri and J. Voas, “The economics of “fake news”,” *IT Professional*, vol. 19, pp. 8–12, November 2017.
- [9] R. B. Delfim, “Tse multa campanha de haddad por notícias contra bolsonaro.”
- [10] M. Oliveira, “Tse multa campanha de haddad por impulsionar notícias contra bolsonaro na internet.”
- [11] L. Lellis, “Tse multa coligação de haddad por impulsionar conteúdo contra bolsonaro.”
- [12] E. Koike and S. Nishizaki, “Software analysis of internet bots using a model checker,” in *2013 International Conference on Information Science and Cloud Computing Companion*, pp. 242–245, Dec 2013.