

Avaliando um Mecanismo de Consenso no Processo de Perícia de Desinformação através de Simulação

**Valdemar Vicente Graciano Neto, Luiza Martins de Freitas Cintra,
Pedro Henrique Campos Damacena, Acquila Santos Rocha, Vinícius Cunha
M Borges, Jacson Rodrigues Barbosa, Eliomar Araújo de Lima**

¹Instituto de Informática – Universidade Federal de Goiás (UFG)
Alameda Palmeiras, s/n - Chácaras Califórnia, Goiânia - GO - Brazil

{valdemarneto, jacson.rodrigues, eliomar.lima, vcmborges}@ufg.br

{cintraluiza, acquila, pedro.campos2}@discente.ufg.br

Abstract. *Misinformation is a latent contemporary concern, driven mainly by its dissemination on social media. Although artificial intelligence (AI) exhibits high precision and speed up the detection process, AI is still not fully reliable, requiring human participation in the verification process. Fact checkers are experts who play a crucial role, but the process can be slow and inefficient. In this sense, an R&D project is being developed in a partnership between a federal government agency and a university, creating a tool that speeds up the verification process, allowing multiple verifiers to work in parallel. This platform aims to reach a consensus on the veracity of posts, using a reliable consensus algorithm. The main contribution of this paper is to demonstrate, through simulations, the effectiveness of this algorithm, before deploying it. Preliminary results suggest that the consensus mechanism based on a weighted majority allows fact-checkers with more relevance to more significantly influence the evaluation result, although in some cases evaluators with low importance may override the importance of one or a few fact-checkers with more relevance.*

Resumo. *A desinformação é uma preocupação contemporânea latente, impulsionada principalmente pela sua disseminação nas redes sociais. Embora as técnicas de inteligência artificial (IA) acelerem o processo de detecção, tais técnicas ainda exigem a participação humana no processo de verificação. Checadores de fatos são peritos que desempenham um papel crucial, mas o processo pode ser lento e ineficiente. Neste sentido, um projeto de P&D está sendo desenvolvido numa parceria entre a ANATEL e a UFG, criando uma prova de conceito (PoC) que agilize o processo de verificação, permitindo que múltiplos checadores trabalhem em paralelo. A contribuição principal deste artigo é demonstrar, através de simulações, a eficácia do mecanismo de consenso que apoia o trabalho dos checadores antes de sua implantação. Resultados preliminares sugerem que o mecanismo de consenso baseado em maioria por ponderação possibilita que fact-checkers com mais relevância influenciem mais significativamente o resultado da avaliação, ainda que em alguns casos avaliadores com importância baixa possam anular a importância de um ou poucos fact-checkers com mais relevância.*

1. Introdução

A disseminação de conteúdo falso¹ é um fenômeno presente no uso das redes sociais. Segundo o Fórum Econômico Mundial, a desinformação está no topo dos riscos globais em 2024 [World Economic Forum 2024]. Embora os modelos de IA possam reconhecer informações potencialmente falsas, há uma proliferação de novas formas para criação e disseminação de conteúdo falso. Uma pesquisa do Gartner com mais de 200 consumidores entre julho e agosto de 2023² revelou que 53% dos consumidores acreditam que o estado atual das mídias sociais piorou. Mais ainda, atos delituosos gerados a partir do uso de técnicas de *DeepFake* [Westerlund 2019, Yu et al. 2021] podem elevar a disseminação de conteúdo falso a um patamar ainda mais preocupante.

No estado da prática, ainda que as técnicas de IA atinjam alta acurácia, os resultados não são 100% confiáveis. Afinal, uma única classificação equivocada por um algoritmo de IA sobre uma postagem importante poderia levar ao total descrédito, por parte da população, dos resultados entregues. Neste sentido, ainda é muito importante que haja o envolvimento do ser humano no processo de perícia de desinformação (*human-in-the-loop*). Nesse processo, os chamados ‘checadores de fatos’ (ou *fact-checkers*) são especialistas, geralmente associados a Agências de Checagens de Fatos, que realizam a verificação de postagens e entregam os resultados, divulgados por portais conhecidos da população, tais como o portal ‘Fato ou Fake’³. No entanto, esse processo de avaliação costuma ser feito por poucos peritos e pode ser moroso, visto que o levantamento de evidências que comprovem que a notícia é falsa pode levar até uma semana, um prazo inviável quando a disseminação de desinformação pode manipular a opinião em massa. Nesses casos, divulgar depois de uma semana que um conteúdo é falso pode ser tarde demais.

A Agência Nacional de Telecomunicações (ANATEL, *sponsor* do projeto) e a Universidade Federal de Goiás (UFG) estão desenvolvendo um Projeto de Pesquisa e Desenvolvimento (P&D) para conceber uma prova de conceito (PoC) que permita agregar múltiplos checadores de fatos de modo que eles possam trabalhar de forma paralela em uma mesma notícia e chegar a um consenso sobre a falsidade ou veracidade de um conteúdo [Graciano-Neto et al. 2024a, Graciano-Neto et al. 2024b]. A PoC visa a acelerar o processo, entregando uma avaliação prévia com explicabilidade pelo mecanismo de IA, com evidências de falsidade ou veracidade, que serve de insumo para os checadores. Com isso, os checadores podem emitir suas opiniões com maior celeridade, alcançando consenso e entregando o resultado com menor tempo, reduzindo os danos causados pela disseminação de desinformação. O cerne deste processo é o mecanismo de consenso, isto é, um algoritmo que leva em consideração várias variáveis e, através de normalização, chega a um valor entre zero e um que representa quão falso ou verdadeiro é aquele conteúdo. O problema é que este algoritmo também precisa ser altamente confiável, visto que um cálculo equivocado poderia distorcer o resultado, afetando também a credibilidade

¹Os termos mais utilizados na literatura e no estado da prática por parte dos peritos de agências verificadoras de fatos são *desinformação*, *conteúdo falso* ou *postagens falsas*. *Notícias falsas* ou *fake news* mostram-se termos inadequados no vernáculo jornalístico, visto que denota-se por *notícia* algo produzido por entidade jornalística séria, não sendo passível de falsidade.

²<https://www.gartner.com/en/newsroom/press-releases/2023-12-14-gartner-predicts-fifty-percent-of-consumers-will-significantly-limit-their-interactions-with-social-media-by-2025>

³<https://g1.globo.com/fato-ou-fake/>

da ferramenta.

As principais contribuições deste trabalho incluem (i) um modelo de consenso adaptado de um existente na literatura (chamado *ProBlock* [Sengupta et al. 2021]) e (ii) o uso de simulação pode auxiliar na verificação de corretude de um mecanismo de consenso. Após verificado com simulação, o modelo será implantado na *blockchain* que apoia as atividades da PoC. Um modelo de simulação baseado em eventos discretos foi elaborado para testar cenários emblemáticos que podem gerar questionamentos por parte do *sponsor* e também dos usuários de uma ferramenta implementada sob a égide dessa PoC. Resultados preliminares mostraram que o mecanismo de consenso baseado em maioria por ponderação possibilita que *fact-checkers* com mais relevância (maior *expScore*) influenciem mais significativamente o resultado da avaliação. No entanto, dependendo do número de avaliadores com importância baixa e a força de votos deles, seria possível anular a importância de um ou poucos *fact-checkers* com maior relevância.

O restante deste artigo está organizado como segue. A Seção 2 apresenta os trabalhos relacionados. A Seção 3 apresenta o mecanismo de consenso; a Seção 4 apresenta a avaliação com Simulação, e a Seção 5 apresenta as considerações finais.

2. Fundamentos e Trabalhos Relacionados

Os termos *notícia falsa*, *desinformação* e *fake news* têm sido utilizados de forma intercambiável e ferramentas e abordagens têm sido criadas para lidar com este fenômeno [Morais et al. 2020, Nolasco and Oliveira 2021, Cavalcante et al. 2024]. O Projeto de Lei 2630-2020 (que possui a alcunha de *Lei das fake news*) define **desinformação** como “o conteúdo, em parte ou no todo, inequivocamente falso ou enganoso, passível de verificação, colocado fora de contexto, manipulado ou forjado, com potencial de causar danos individuais ou coletivos, ressalvado o (eventual) ânimo humorístico ou de paródia”.

No estado da prática, Agências Verificadoras de Fatos (AVF) são aquelas que reúnem evidências para classificar a veracidade de uma postagem. Neste sentido, é importante diferenciar os termos *postagem* e *notícia*. O primeiro termo diz respeito a qualquer conteúdo postado em uma rede social ou outro veículo de comunicação cuja veracidade não se sabe e demanda averiguação. Uma notícia, por sua vez, é um conteúdo gerado por jornalistas, e entende-se que há primazia pela veracidade; portanto, prefere-se utilizar os termos *desinformação* ou *postagem falsa*. Uma postagem cuja veracidade deve ser determinada é analisada nas AVF por *fact-checkers* (checadores de fatos independentes). Eles atribuem pontuações com base no seu grau de confiança e autenticidade da postagem. Em seguida, os votos são ponderados com base nas credenciais/dados do *fact-checker*, como experiência, designação, afiliações, frequência e precisão.

Para levantar o estado da arte sobre mecanismos de consenso, foi realizada uma revisão de literatura *ad hoc* e exploratória em bases científicas através de mecanismos de buscas, como o *Google Scholar*. A busca teve o intuito de externalizar o estado da arte sobre a utilização de mecanismos de consenso para *fact-checkers* através da tecnologia *blockchain* na detecção e no combate às *fake news*. Dois conjuntos de palavras-chave foram definidos e empregados nas pesquisas: 1) *weighted majority voting blockchain fake news* e 2) *weighted majority voting blockchain fake news*.

Os resultados mostram que existem poucas abordagens que utilizam a técnica de

checagem de fatos [Torky et al. 2019, Jaroucheh et al. 2020, Sengupta et al. 2021, Rani et al. 2022, Boovitha et al. 2023, Choi and Kim 2023, Faridi et al. 2023, Tajrian et al. 2023, Bodaghi et al. 2023]. Dentre elas, duas técnicas se destacaram para o escopo do mecanismo de votação dos *Fact-Checkers*: **avaliação da credibilidade da fonte** e **avaliação da confiança no *fact-checker*** (avaliador).

Dentre as abordagens que focam em avaliar a credibilidade da fonte, Torky et al. (2019) [Torky et al. 2019], por exemplo, baseiam-se na avaliação de credibilidade da fonte da notícia com base na popularidade da fonte e conforme a categoria que lhe é inerente (e.g. revista online, *blogs*, etc.). Entretanto, nesse tipo de abordagem, uma pessoa influente que atue como fonte de informação (por exemplo influenciador digital, cantores, atletas, e outros) pode ter mais sucesso para difundir conteúdo falso. Dentre as abordagens de checagem manual por *fact-checkers* que focam na confiança do avaliador, o *ProBlock* propõe o algoritmo de consenso chamado Proof-of-Trust (PoT) [Sengupta et al. 2021], em que os *fact-checkers* são especialistas do domínio que avaliam a postagem, tais como jornalistas e/ou instituições com notório saber (*expert-based*). No que tange ao *ProBlock*, os votos dos *fact-checkers* são ponderados conforme critérios que qualificam os votos realizados, como por exemplo sua experiência, designação, organização a que pertence, frequência de avaliação e histórico de acurácia. Com base nesses critérios, o modelo matemático probabilístico (*ProBit*) é empregado para dar um veredito sobre a postagem (falsa, genuína ou inconclusivo) na abordagem *ProBlock* [Sengupta et al. 2021].

Considerando a robustez, precisão e confiabilidade entre as três abordagens avaliadas, a técnica embutida no modelo de consenso do *ProBlock* é a que se sobressaiu. Contudo, um desafio em aberto da proposta diz respeito a validação das informações sobre os critérios dos *fact-checkers* e a definição de um número mínimo/razoável de *fact-checkers* para validar uma notícia. Além disso, os especialistas geralmente comparam as afirmações com provas estabelecidas em validações anteriormente expressas ou relatadas. Portanto, a técnica empregada no *ProBlock* pode ter um tempo de latência considerável, atrasando a entrega do resultado sobre a notícia [Tajrian et al. 2023].

O uso de simulação no domínio de *fake news* não é uma novidade [Jiang and Porter 2022]. Há inclusive trabalhos que fazem simulação de mecanismos de consenso [Oyinloye et al. 2023]. Entretanto, não foram encontrados outros trabalhos que tenham avaliado o mecanismo de consenso adaptado do *ProBit* e utilizando DEVS. Dentre as abordagens analisadas, a abordagem *ProBlock* mostrou-se mais apropriada e viável para o escopo do projeto, uma vez que fornece uma detecção mais precisa e confiável da notícia. O algoritmo de consenso elaborado foi inspirado nos conceitos propostos pelo *Problock* com as devidas adaptações, como mostrado na seção a seguir.

3. Modelo de Consenso

O mecanismo de consenso foi inspirado na abordagem *ProBlock* [Sengupta et al. 2021], que propõe uma solução de consenso chamada *Proof of Trust (PoT)*. Nela, os *fact-checkers* são especialistas do domínio (jornalistas e/ou instituições com notório saber) que avaliam a postagem. Uma postagem cuja veracidade deve ser determinada é votada por *fact-checkers* autorizados que atribuem pontuações com base no seu grau de autenticidade e confiança sobre ela. Em seguida, os votos são ponderados com base nas credenciais/dados do *fact-checker*, como experiência, designação, afiliações, frequência e

precisão. Portanto, nós empregaremos um modelo de votação por maioria ponderada.

O *ProBit* [Shao et al. 2020] é o modelo matemático probabilístico clássico que analisa os votos ponderados para prever sua precisão, que leva em consideração credenciais dos *fact-checkers* e os seus votos para gerar uma pontuação final. O modelo de votação por maioria ponderada é implementado calculando a pontuação do especialista/*fact-checker* (*expScore*) baseada na interpretação dos *fact-checkers* sobre a falsidade do conteúdo (voto), nas estatísticas estáticas e dinâmicas de sua carreira e uma pontuação baseada na confiança do verificador sobre seu voto na notícia. O *expScore* é atualizado para cada ciclo de revisão, que é uma pontuação inteira que consiste em uma base estática e dinâmica para cada *fact-checker*.

A componente estática do *expScore* consiste numa pontuação baseada na análise da experiência (*yrsScore*), organização (*orgScore*) e designação (*desScore*) dos *fact-checkers* referida na Tabela 1. A pontuação relativa é dada a cada critério e ajuda a criar o perfil do especialista, o que ajuda a atribuir maior responsabilidade e confiança com o voto. A Equação 1 que descreve a componente estática é:

$$expScore = yrsScore + orgScore + desScore \quad (1)$$

Tabela 1. Métricas da Componente Estática do *ExpScore*.

Experiência	Organização	Designação
0 a 5 anos - 1	Freelancer - 1	Jornalista Jr. - 1
5 a 10 anos - 2	Regional - 2	Jornalista - 2
mais de 10 anos - 3	Local - 3	Sr. Jornalista - 3
	Nacional - 4	Executivo - 4
	Internacional - 5	Sr. Executivo - 5
<i>yrsScore</i>	<i>orgScore</i>	<i>desScore</i>

A variável *confidenceScore* é fornecida pelo próprio *fact-checker* e indica seu nível de confiança em seu voto. Portanto, ela tem um caráter auto-avaliativo. A *confidenceScore* segue a seguinte escala de valores: 1: Não confiante, 2: Confiante, 3: Altamente Confiante.

A variável *newsVote* é de fato o voto do *fact-checker* e seus valores podem variar conforme a escala de -2 até +2, indicando a análise do *fact-checker* sobre se o conteúdo é genuíno ou falsa, em que -2 é Completamente Falso, -1 é Parcialmente Falso, 0 é Neutro, +1 é Parcialmente Verdadeira, e +2 é Completamente Verdadeira.

A *confidenceScore* é matematicamente combinada com a variável *newsVote* para calcular o *cummVote*, ou seja, o voto cumulativo. A Equação 2 descreve a operação matemática para cálculo do *cummVote*.

$$cummVote = confidenceScore \cdot newsVote \quad (2)$$

Como mostrado na equação 3, a variável *cummVote* é então ponderada com o *expScore*, resultando na variável *weightedVote* e portanto, o modelo de votação por maioria ponderada é implementado. O *expScore* é normalizado para um valor decimal entre 0

e 1 para efeito de ponderação. *expScore* é considerado como o peso e o *cummVote* é considerado como o voto no sistema de votação ponderada.

$$weightedVote = cummVote \cdot expScore \quad (3)$$

A votação ponderada (*finalVote*) é um conjunto de votos ponderados *weightedVote* de todos *fact-checkers* que avaliaram a notícia, conforme mostrado na equação.

$$finalVote = [weightedVote_1, weightedVote_2 + weightedVote_3 + \dots + weightedVote_x], \quad (4)$$

O modelo recebe como entrada o conjunto de votos ponderados e processa a média, desvio padrão, o parâmetro de máxima verossimilhança e função de distribuição cumulativa (*Cummulation Distribution Function*, CDF) da distribuição normal. Com base nesses valores, a análise probabilística dos votos é feita para determinar a probabilidade de a postagem levada em consideração ser genuína ou falsa. A probabilidade é dada na escala de valor entre 0 e 1. Se valor for maior que 0,56, a notícia é avaliada como verdadeira, se for menor ou igual a 0,45, o conteúdo é avaliado como falso, se o valor ficar entre 0,46 e 0,55, a avaliação do conteúdo é considerado inconclusivo.

4. Avaliação com Simulação

O **objetivo** deste estudo foi formalizado com base na técnica GQM (*Goal-Question-Metric*) [Basili 1993]: *Analisar o mecanismo de consenso para fins de testá-lo no que diz respeito a diferentes configurações de votos do ponto de vista dos fact-checkers no contexto de fake news.* Seguiu-se também preceitos de protocolos inspirados no trabalho de [de França and Travassos 2016]. Foi realizada uma família de experimentos com simulação, **testando diversos cenários de votação**: com *3 fact-checkers*, *4 fact-checkers* e *5 fact-checkers*. O modelo de simulação foi implementado em DEVS⁴ e implantado na ferramenta MS4Me⁵. Fact-checkers foram modelados como modelos atômicos que enviam seus votos a um contabilizador (também modelo atômico), que juntos formam o modelo acoplado. O componente estático do algoritmo de consenso de votos possui três variáveis principais: *ExpScore*, que corresponde à reputação do *fact-checker* (variando de 1 a 5, onde 5 é a maior reputação), *confiança do verificador no voto dado* (variando de 1 a 3, onde 3 é a maior confiança) e o *voto dado* (que varia de 1 a 5, onde 1 é totalmente falso e 5 é totalmente verdadeiro).

Dadas as variáveis e os valores que podem assumir, é possível obter 75 combinações (5x3x5) diferentes de reputação x confiança x voto para cada *fact-checker*. Assim, ao pensar em avaliar todas as combinações possíveis entre as variáveis para cada *fact-checker*, temos uma explosão combinatorial que torna inviável a verificação de todas as possibilidades. Percebe-se que o número de combinações a serem testadas seria sempre 75^N , onde N é o número de *fact-checkers* envolvidos na perícia de uma única postagem. Para um número modesto (3 *fact-checkers*), isso demandaria verificar

⁴Devido a restrições de espaço, toda a documentação da implementação da simulação e experimentação estão disponíveis em link externo <https://ww2.inf.ufg.br/~valdemarneto/projects/consensoFakeNews.html>.

⁵<https://rtsync.com/ms4me>

421.875 combinações diferentes. Sendo assim, devido à inviabilidade de testar todas as combinações, para garantir uma cobertura razoável, adotou-se uma abordagem da prática de teste de software: a priorização de casos de teste [Naik and Tripathy 2011]. Este tipo de teste, inspirado na técnica de teste de software caixa-preta Análise do valor limite (ou *Boundary-Value Analysis*) [Iavernaro et al. 1998], corresponde a testar casos limítrofes (chamados neste trabalho de '*casos de fronteira*') ao invés de testar absolutamente todos os casos.

A estratégia adotada nesta avaliação é fazer primeiro o teste do componente estático para, posteriormente (estudo futuro), realizar o teste dinâmico. A ideia é implementar um número pequeno de *fact-checkers* e selecionar um subconjunto de casos de teste para priorizar. Foi definida uma escala de valores definidos para importância do revisor e a força do seu voto. A escala dos valores referentes a importância do verificador conforme o valor calculado da *expScore* correspondente ao componente estática:

- 1 a 3: Baixa.
- 4 a 7: Média.
- 8 a 10: Alta.
- 11 a 13: Máxima.

A força do voto é representada pela variável *cummVote*, seguindo a escala de valores abaixo:

- 0: Nulo.
- 1 a 3: Fraco (Positivo) ou -1 a -3: Fraco (Negativo).
- 4 a 6: Forte (Positivo) ou -4 a -6: Forte (Negativo).

A implementação atual do mecanismo de consenso incluiu a componente estática de ponderação dos votos dos *fact-checkers*. Para os dados experimentais, alguns exemplos de casos levando em consideração a importância do *fact-checker* e da força do voto do *fact-checker* foram gerados de forma aleatória. Todos os resultados aqui apresentados consideram 5 *fact-checkers*. O caso Extremo determina que um dos *fact-checkers* tem importância e/ou força do voto com valores altos e os demais *fact-checkers* baixos. Os Casos médios apresentam um maior equilíbrio entre importância e força do voto dos *fact-checkers*. Esses dois tipos de casos ajudam a avaliar como o mecanismo de consenso inspirado no ProBlock impacta no resultado da avaliação conforme a relevância do *fact-checker* e o conteúdo avaliado.

A Figura 1 ilustra o primeiro exemplo que mostra três combinações de votos de três revisores sobre três notícias com um dos revisores significativamente mais relevante. No conjunto de barras à esquerda do gráfico, temos um *fact-checker* com a menor importância (relevância) em azul (menor valor de voto ponderado, isto é, importância baixa e voto fraco) manifestando um voto negativo, portanto influenciando pouco a CDF que resulta no valor aproximadamente de 0.78; portanto, os votos dos dois *fact-checkers* com mais importância em laranja e verde define a avaliação da notícia como verdadeira. No conjunto de barras do meio do gráfico, há dois *fact-checkers* com a menor importância em azul e laranja (menor valor de voto ponderado, ou seja, menor credibilidade) manifestando votos negativos, também influenciando pouco a CDF que resulta no valor aproximadamente de 0.68. Portanto, o *fact-checker* com mais importância influencia de forma muito significativa o resultado da avaliação devido ao seu valor de voto ponderado possuir

muito mais peso. Isso fica ainda mais evidente no conjunto de barras da direita do gráfico, em que os dois *fact-checkers* com a menor importância em laranja e verde (menor valor de voto ponderado) manifestaram votos positivos, mas com o *fact-checker* mais relevante manifestando voto negativo, o que resultou no valor aproximadamente de 0.39 (postagem avaliada como falsa).

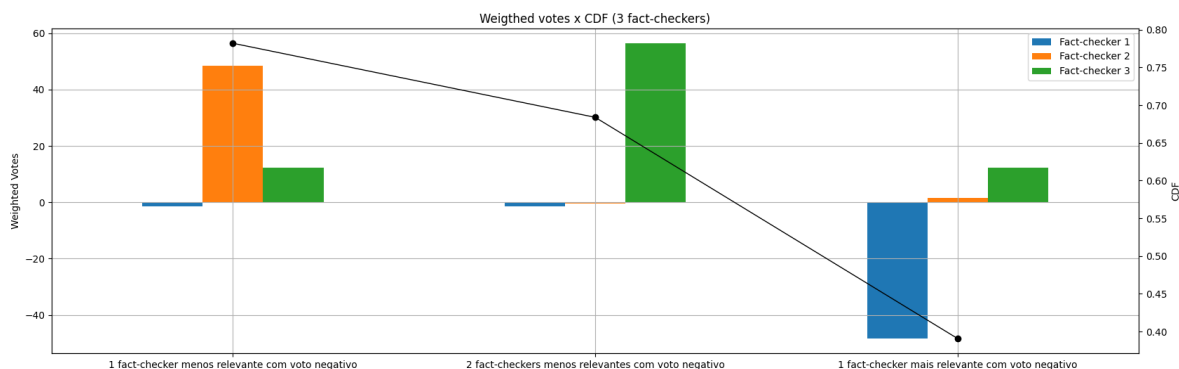
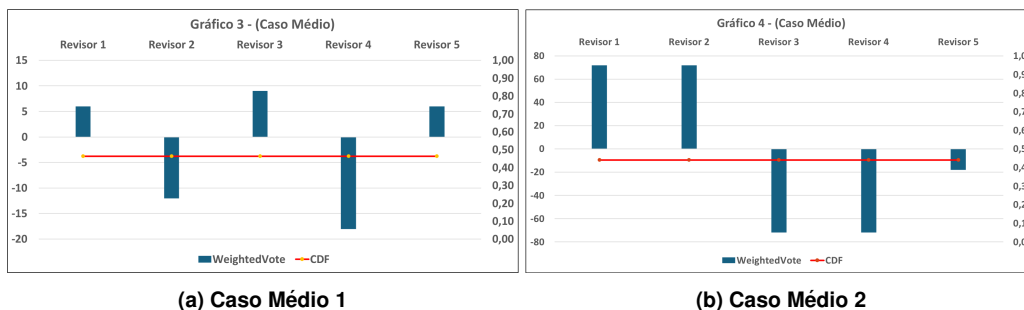


Figura 1. CDF e Voto Ponderado (WeightedVote) de 3 revisores variando aleatoriamente o voto

A Figura 2(a) mostra cinco *fact-checkers* e seus votos ponderados em azul. A linha vermelha mostra o cálculo do CDF para os votos. É possível verificar que os revisores possuem os votos mais equilibrados. Os revisores 1, 3 e 5 são revisores de classificação de importância baixa e voto fraco, que estão manifestando que a notícia é verdadeira. Já os revisores 2 e 4, que estão em menor número e que possuem uma classificação de importância igualmente baixa mas com voto forte, fazem com que o cálculo do CDF chegue a 0,46, ou seja, a notícia se torna inconclusiva pela força do voto dos revisores 2 e 4, mesmo que em menor número. A Figura 2(b), por sua vez, mostra votos mais equilibrados onde os revisores 1 e 2 são de importância máxima e voto forte, que classificam a notícia como verdadeira; enquanto os revisores 3 e 4 são de igual importância e voto, porém classificam a notícia como falsa e finalmente o revisor 5 possui importância baixa e sua força de voto é classificada como forte. Com isso, o CDF é de 0,44, sendo a postagem falsa e o voto do revisor 5 determinante para o desempate na simulação.



(a) Caso Médio 1

(b) Caso Médio 2

Figura 2. Casos Médios

Ameaças à Validade e Limitações. [de França and Travassos 2016] mencionam que, quando se utiliza simulação, a validade experimental está fortemente associada à validade do modelo de simulação em si. Os autores lembram que a simulação melhora a

validade de **construto** e **interna**. Do ponto de vista estrutural, o modelo apresentado representa precisamente os componentes típicos de um processo de votação sobre a veracidade de postagens; do ponto de vista de parâmetros, os valores foram extraídos da literatura especializada; e do ponto de vista empírico, novamente, tem-se a contrapartida com os parâmetros do mundo real, o que mitiga as ameaças listadas de construto e interna. As validades **externa** e de **conclusão** devem ser alcançadas com a aplicação de testes estatísticos adequados sobre as saídas do modelo. Para isso, o estudo utiliza vários tamanhos de amostras de entrada para exercitar o modelo. Ressalta-se, ainda assim, que testes estatísticos podem não ser efetivos, visto que o número de repetições das simulações podem alterar consideravelmente a significância estatística que se possa alcançar.

5. Considerações Finais

Este artigo apresentou resultados de uma avaliação de um mecanismo de consenso com base em simulação. O mecanismo será implantado em uma PoC para checagem de postagens em um projeto de P&D entre a UFG e a ANATEL. A simulação contribuiu por permitir exercitar diferentes casos que combinam força de voto e importância de *fact-checkers* (revisores). O mecanismo de consenso mostra que, na maioria das vezes, *fact-checkers* com mais relevância (maior *expScore*) influenciam mais significativamente o resultado da avaliação. Cenários mais raros também foram testados, mostrando que, dependendo dos valores de importância e força de voto dos demais avaliadores, o resultado pode ser diferente do que foi aferido pelo(s) *fact-checker(s)* com maior importância. Trabalhos futuros incluem (i) aumentar o número de *fact-checkers* no estudo; (ii) investigar meios de modular os votos e priorizar *fact-checkers* de maior relevância, (iii) criar mecanismos de incentivo para atrair *fact-checkers* com maior importância e (iv) realizar validações com dados reais de redes sociais.

Agradecimentos. Os autores agradecem à ANATEL e à Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG) por financiarem esta pesquisa.

References

- Basili, V. R. (1993). Applying the goal/question/metric paradigm in the experience factory. *Software quality assurance and measurement: A worldwide perspective*, 7(4):21–44.
- Bodaghi, A., Schmitt, K. A., Watine, P., and Fung, B. C. (2023). A literature review on detecting, verifying, and mitigating online misinformation. *IEEE Transactions on Computational Social Systems*.
- Boovitha, D., Abirami, M., Gunavathi, S., Revathi, N., and Rubavarshini, S. (2023). Fake media detection based on natural language processing and blockchain approaches. *South Asian Journal of Engineering and Technology*, 13(1):69–82.
- Cavalcante, A. A. B., Freire, P. M. S., Goldschmidt, R. R., and Justel, C. M. (2024). Improving implicit crowd signals based fake news detection on social media: A time-aware method for early detection. In *20th SBSI 2024*, pages 7:1–7:9, Juiz de Fora, Brazil. ACM.
- Choi, N. and Kim, H. (2023). Dds: Deepfake detection system through collective intelligence and deep-learning model in blockchain environment. *Applied Sciences*, 13(4):2122.
- de França, B. B. N. and Travassos, G. H. (2016). Experimentation with dynamic simulation models in software engineering: planning and reporting guidelines. *Empirical Software Engineering*, 21(3):1302–1345.

- Faridi, A. R., Singh, R., Masood, F., and Salmony, M. Y. (2023). Machine learning based novel framework for fake news detection and prevention using blockchain. In *10th INDIACom*, pages 751–755. IEEE.
- Graciano-Neto, V. V., Barbosa, J., Lima, E., Carvalho, S., and Venzi, S. (2024a). A Blockchain-based and AI-Endorsed Mechanism to Support Social Networks on Fake News Containment. In *XIII BraSNAM*, pages 207–213, Brasília/DF. SBC.
- Graciano-Neto, V. V., Barbosa, J. R., de Lima, E. A., de Freitas Cintra, L. M., Venzi, S., and Kassab, M. (2024b). Establishing a blockchain-based architecture for fake news detection. In *SBCARS*, pages 1–10, Curitiba/PR. SBC.
- Iavernaro, F., La Scala, M., and Mazzia, F. (1998). Boundary values methods for time-domain simulation of power system dynamic behavior. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 45(1):50–63.
- Jaroucheh, Z., Alissa, M., Buchanan, W. J., and Liu, X. (2020). Trustd: Combat fake content using blockchain and collective signature technologies. In *44th COMPSAC*, pages 1235–1240. IEEE.
- Jiang, Y. and Porter, M. D. (2022). Simulating fake news dissemination on twitter with multivariate hawkes processes. In *IEEE BigData*, pages 3597–3606.
- Morais, J. I. d., Abonizio, H. Q., Tavares, G. M., da Fonseca, A. A., and Jr, S. B. (2020). A multi-label classification system to distinguish among fake, satirical, objective and legitimate news in brazilian portuguese. *iSys - Brazilian Journal of Information Systems*, 13(4):126–149.
- Naik, K. and Tripathy, P. (2011). *Software testing and quality assurance: theory and practice*. John Wiley & Sons.
- Nolasco, D. and Oliveira, J. (2021). Topical rumor detection based on social network topic models relationship. *iSys - Brazilian Journal of Information Systems*, 14(2):05–27.
- Oyinloye, D. P., Teh, J. S., Jamil, N., and Teh, J. (2023). Sim-p—a simplified consensus protocol simulator: Applications to proof of reputation-x and proof of contribution. *IEEE Internet of Things Journal*, 10(6):5083–5094.
- Rani, P., Jain, V., Shokeen, J., and Balyan, A. (2022). Blockchain-based rumor detection approach for covid-19. *Journal of Ambient Intelligence and Humanized Computing*, pages 1–15.
- Sengupta, E., Nagpal, R., Mehrotra, D., and Srivastava, G. (2021). Problock: a novel approach for fake news detection. *Cluster Computing*, 24:3779–3795.
- Shao, X., Ma, X., Chen, F., Song, S., Pan, X., and You, K. (2020). A random parameters ordered probit analysis of injury severity in truck involved rear-end collisions. *International Journal of Environmental Research and Public Health*, 17:395.
- Tajrrian, M., Rahman, A., Kabir, M. A., and Islam, M. R. (2023). A review of methodologies for fake news analysis. *IEEE Access*.
- Torky, M., Nabil, E., and Said, W. (2019). Proof of credibility: A blockchain approach for detecting and blocking fake news in social networks. *IJACSA*, 10(12).
- Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology innovation management review*, 9(11).
- World Economic Forum (2024). Global risks 2024: Disinformation tops global risks 2024 as environmental threats intensify. <https://www.weforum.org/press/2024/01/global-risks-report-2024-press-release/>. Accessed: [June, 2024].
- Yu, P., Xia, Z., Fei, J., and Lu, Y. (2021). A survey on deepfake video detection. *Iet Biometrics*, 10(6):607–624.