Mapeamento dos Principais Ataques e Contramedidas de Segurança em Redes 6G Habilitadas por IA

Antonia Raiane S. Araujo Cruz[1,2], Cleitianne O. Silva[1], Joao C. da C. de Lima [1], Rossana M. C. Andrade[1]*, Emanuel B. Rodrigues[1]

[1]Universidade Federal do Ceará (UFC)

[2]Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

{raiane.santos,cleitianne,jcarloslima}@alu.ufc.br,

{rossana, emanuel}@dc.ufc.br

Abstract. *With the advancement of technology and the increasing integration of Artificial Intelligence (AI) systems across various sectors, it is crucial to understand and mitigate the risks associated with their implementation, especially in the innovative scenario proposed by the sixth generation of mobile communications (6G). This study presents a bibliographic mapping to identify and analyze the most relevant vulnerabilities found in the current literature on AI and 6G. Furthermore, it explores proposed countermeasures to address these vulnerabilities, aiming to ensure the security, reliability, and privacy of AI systems in the context of 6G.*

Resumo. *Com o avanço tecnológico e a integração crescente da Inteligência Artificial (IA) em diversos setores, compreender e mitigar os riscos associados à sua implementação torna-se crucial, especialmente diante da inovação proposta pela sexta geração de comunicações móveis (6G). Este estudo apresenta um mapeamento bibliográfico para identificar e analisar as principais vulnerabilidades descritas na literatura atual sobre IA e 6G. Além disso, investiga-se as contramedidas propostas para enfrentar tais vulnerabilidades, visando assegurar a segurança, confiabilidade e privacidade dos sistemas de IA no contexto do 6G.*

1. Introdução

Atualmente, com o advento da 5ª geração das redes móveis (5G), observa-se uma diminuição das restrições nas aplicações de dados intensivos, como as exigências por altas taxas de dados. No entanto, a sociedade está cada vez mais ávida por avanços tecnológicos, especialmente no que se refere a mecanismos inteligentes [Ji et al. 2021]. Nesse contexto, surge o 6G, destacando um direcionamento para conectar dispositivos heterogêneos e integrar a Inteligência Artificial (IA) para o desenvolvimento de aplicações capazes de coletar, analisar e interagir com os dados de forma ainda mais sofisticada.

Assim, ressalta-se a significativa importância da IA nos serviços 6G, onde as diversas tecnologias convergem para integrar seus serviços com a IA. Isso pode resultar em

*Bolsista de Produtividade Desen. Tec. e Extensão Inovadora do CNPq - Nível 1D

questões de vulnerabilidade na autenticação e no controle de acesso, bem como comportamentos maliciosos, conforme discutido por [Wang et al. 2020a], entre outros ataques mais direcionados às características específicas de cada tecnologia. Diante desse cenário, este artigo apresenta um mapeamento dos principais ataques que podem afetar a segurança e privacidade em redes 6G habilitadas por IA. Além disso, apresentamos as contramedidas que os estudos recentes indicam como soluções.

2. Trabalhos Relacionados

Esta pesquisa apresenta um mapeamento das principais ameaças à segurança e privacidade nas redes 6G habilitadas por IA. Para a busca na literatura foi utilizada a *string*: *(6g) AND ("artificial intelligence"OR "machine learning") AND ("security"OR "privacy"OR "cybersecurity")*, aplicada nas bases *IEEE Explore*, *ScienceDirect* e *Scopus*. Após, realizou-se a triagem dos artigos, nos últimos 4 anos, os quais foram anos-chave para a consolidação do 5G. Os dados foram importados para o Parsif.al[1], resultando em 235 estudos. Depois da fase de leitura, foram selecionados os trabalhos relacionados efetivamente ao tema desta pesquisa.

O estudo de [Porambage et al. 2021b] tem como foco realizar um detalhamento das principais tecnologias em relação à segurança e IA: segurança de IA/Machine Learning (ML) distribuída e escalável; utilização de *Distributed Ledger Technology* (DLT) para proteger a integridade dos dados de IA através de registros imutáveis e confiança distribuída entre diferentes partes; segurança quântica e algoritmos de ML para aprimorar o aprendizado supervisionado e não supervisionado; e rede 6G com métodos de *Physical Layer Security* (PLS) para adicionar uma camada de proteção adaptável, incorporando novas técnicas como *TeraHertz (THz) Communication*, *Visible Light Communication* (VLC) e *Molecular Communication* (MC). Além disso, os autores discutem potenciais ameaças e possíveis soluções. Contudo, as conclusões carecem de direcionamentos futuros claros e apresentam mais questões do que soluções específicas.

Em [Wang et al. 2020b], os autores discutem quatro aspectos-chave das redes 6G: computação de borda inteligente, IA distribuída, rádio e intercomunicadores 3D. Os autores também mencionam tecnologias emergentes nessas áreas, como IA na detecção de anomalias de rede e a MC para promover o desenvolvimento de novos mecanismos de autenticação. Embora o estudo apresente tecnologias 6G e aprofunde as áreas-chave da rede 6G para identificar vulnerabilidades de segurança conectadas a tecnologias futuras, os autores apresentam apenas prováveis aplicações que a rede 6G suportará, carecendo de direcionamentos mais específicos como estratégias para mitigação das vulnerabilidades.

Os autores [Nguyen et al. 2021] propuseram uma investigação sistemática das questões de segurança e privacidade nas camadas física, rede e aplicações. De modo que, os ataques como *Denial of Service* (DoS), *Distributed Denial of Service* (DDoS) e esgotamento de energia continuarão a ser ameaças à arquitetura de segurança 6G. Os pesquisadores afirmam que novas aplicações frequentemente introduzem ameaças à segurança, exigindo melhorias e ajustes nos protocolos para reduzir vulnerabilidades. Por fim, os autores afirmam que além das ameaças à segurança comuns, como ataques de vírus, *malware*, DDoS, *deep fake*, também serão preocupantes os ataques com poder de aprendizagem e violação maciça de dados. Desse modo, será necessário assegurar as tecnologias

[1]https://parsif.al/

de alto impacto para 6G e melhorar as tecnologias 5G que serão suporte para o 6G.

Enquanto diversos estudos focam em aspectos gerais da tecnologia ou em análises de segurança, nossa pesquisa se concentra em identificar e compreender os principais ataques e contramedidas específicas que a IA enfrentará no contexto do 6G.

3. Mapeamento dos Principais Ataques e Contramedidas para IA em 6G

Nesta seção é realizado um mapeamento dos ataques e contramedidas voltados para IA. Esse mapeamento desempenha um papel crucial na garantia da segurança e privacidade das redes de comunicação móvel de próxima geração, como o 6G. A identificação antecipada dos ataques emergentes possibilita a implementação de contramedidas indicadas pelas pesquisas atuais, como garantia de segurança da informação em redes 6G, em face das ameaças que também estão em constante evolução.

IA é significativa na evolução das redes 6G e dos diversos casos de uso 6G [Silva et al. 2023]. Os autores também destacam os requisitos de segurança para os casos de uso. Desta forma, os problemas de segurança e privacidade também serão explorados nesta tecnologia. A IA poderá ser utilizada como mecanismo para o lançamento de ataques inteligentes, como ilustrado na Figura 1. São destacados ataques nas seguintes categorias: (i) ataques adversariais, (ii) ataques de inversão de modelos, (iii) ataques de extração de modelos, (iv) ataques de inferência de atributos e (v) roubo de modelos.

Os **ataques adversariais** visam direcionar entradas diversas nas fases de treino e teste dos dados e modelo de aprendizagem, gerando aprendizado de forma incorreta. O envenenamento dos dados busca inserir dados incorretamente rotulados/amostras maliciosas ou adulterar objetos de entrada para induzir os algoritmos de aprendizagem automática ao erro. Pode ocorrer envenenamento de algoritmos para influenciar o processo de aprendizagem distribuído através do carregamento de pesos manipulados em modelos de aprendizagem locais e, envenenamento de modelos para substituir o modelo implementado por um modelo malicioso [Nguyen et al. 2021].

Os ataques de inversão de modelos e extração de modelos objetivam invadir a privacidade. Em algumas situações o atacante pode utilizar de previsões de modelos para expor a privacidade de registos sensíveis que foram utilizados como parte do conjunto de treinamento [Siriwardhana et al. 2021]. Esses ataques são chamados de **ataques de inversão de modelos** e podem comprometer o modelo de IA e a divulgação de dados sensíveis [Sandeepa et al. 2022]. Já os **ataques de extração de modelos** centram-se na informação do modelo, com a finalidade de roubar os parâmetros/hiperparâmetros do modelo ou duplicá-los. Podem comprometer a confidencialidade do algoritmo de IA e a propriedade intelectual do aprendente [Sandeepa et al. 2022] e [Liu et al. 2020].

Os ataques de inferência têm a finalidade de inferir se um determinado dado está incluso nos dados de treino do modelo, logo os atacantes têm como alvo os dados que violam a privacidade, podendo afetar diretamente os geradores dos dados. Segundo [Sandeepa et al. 2022], quando o sistema é comprometido por um **ataque de inferência de atributos**, o atacante se utiliza disso para tentar identificar os atributos privados de um utilizador, como localização, gênero, entre outros.

Já no ataque **roubo de modelos**, um adversário tenta replicar a funcionalidade de um modelo de IA/ML vítima, explorando consultas, como entradas e saídas. Esse tipo

de ataque é essencialmente uma forma de envenenamento de previsões e podem induzir os usuários finais a tomar decisões relacionadas à privacidade com base em informações incorretas.

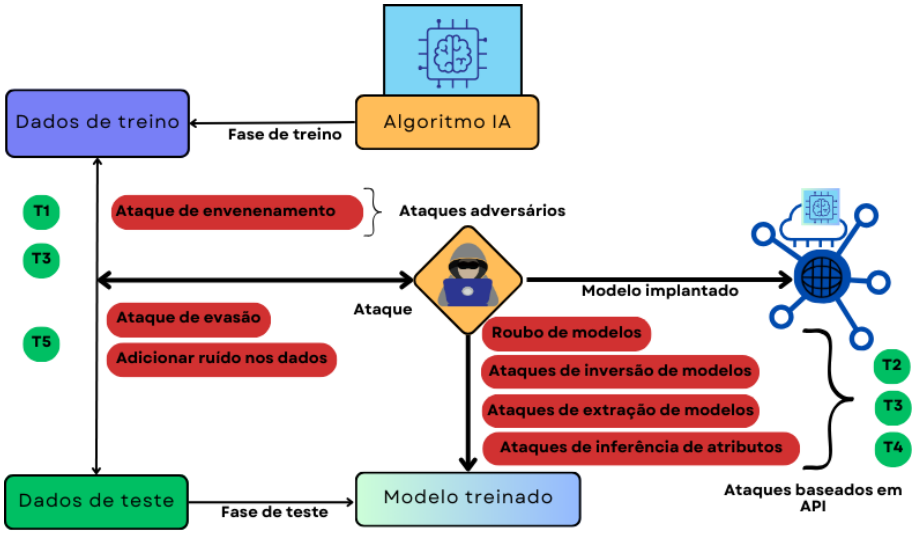

Figura 1. Principais ataques de IA em 6G.
Fonte: Adaptado de [Sandeepa et al. 2022].

Buscamos documentações técnicas da indústria para encontrar as tecnologias habilitadoras como contramedidas de segurança propostas pelo mercado. Essas tecnologias não apenas têm o potencial de contribuir significativamente para atender aos requisitos de segurança no contexto da IA, mas também de enfrentar alguns dos desafios previstos para a próxima geração de redes móveis, o 6G. A Figura 1 ilustra as contramedidas T1 como solução para os ataques de envenenamento, T2, T3 e T4 para os ataques baseados em Interface de Programação de Aplicação (API) e T5 para os ataques que visam a exploração dos dados na fase de teste.

O **Aprendizado Federado - *T1***, uma técnica de aprendizado de máquina, destaca-se na documentação técnica da Aliança NGMN [NGMN Alliance 2023] sobre os casos de uso do 6G. Essa abordagem permite treinar modelos localmente, agregando apenas os parâmetros centralmente, o que preserva a privacidade dos dados dos usuários. No contexto do 6G, o Aprendizado Federado pode ser crucial para a segurança, já que permite atualizar modelos com os dados dos usuários sem comprometer a confidencialidade desses dados. Isso é essencial em um ambiente onde a proteção de dados é uma prioridade crescente.

Também considerada uma tecnologia habilitadora por vários autores, [Porambage et al. 2021a], [Ylianttila et al. 2020], [Siriwardhana et al. 2021] para resolver questões de privacidade no contexto do 6G, a **Privacidade Diferencial (DP) - *T2*** é uma técnica de proteção de dados que visa preservar a privacidade dos indivíduos enquanto permite a análise dos dados para obter *insights* valiosos. A DP pode desempenhar um papel crucial na segurança do 6G, garantindo que as informações dos usuários sejam protegidas enquanto ainda são utilizadas para análises e melhorias nos sistemas.

Outra tecnologia relevante no contexto de IA citada por [Siriwardhana et al. 2021]

e [Ylianttila et al. 2020] é o **Aprendizado Federado Baseado em Borda - *T3***. Essa abordagem preserva a privacidade dos dados ao manter um controle físico mais próximo do usuário, mantendo os dados localmente em dispositivos de borda.

Também recomendada como uma tecnologia habilitadora por por [Siriwardhana et al. 2021], o **Homomorphic Encryption - *T4*** é uma técnica de criptografia que permite realizar operações diretamente em dados criptografados, sem a necessidade de descriptografá-los primeiro. Essa abordagem possibilita o processamento seguro e privado de dados sensíveis, garantindo a confidencialidade das informações mesmo durante as operações. Essa técnica pode ser essencial para a segurança do 6G, protegendo os dados enquanto são processados, o que é importante em um contexto onde a privacidade e a segurança dos dados são prioridades.

Por fim, também é apresentado por [Siriwardhana et al. 2021] o **Aprendizado de Máquina Adversarial - *T5***, que se concentra no desenvolvimento de técnicas e estratégias para garantir que os modelos de aprendizado de máquina sejam robustos e resilientes a ataques e manipulações. No contexto do 6G, onde o uso de algoritmos de aprendizado de máquina será generalizado em diversas aplicações, é primordial garantir a segurança desses modelos. Essa abordagem pode desempenhar um papel fundamental na proteção contra potenciais ameaças e ataques, assegurando a confiabilidade das operações e dos sistemas no ambiente do 6G.

4. Desafios futuros

As aplicações 6G futuras exigirão mais recursos de rede do que as redes 5G atuais. É fundamental avançar em abordagens e mecanismos de encriptação, autenticação, integridade e privacidade dos dados. Diversas questões de segurança persistem; por exemplo, ataques DoS e *Man-In-The-Middle (MITM)*, que são comuns.

Outros desafios de segurança incluem o processamento de tráfego, abrangendo detecção de ataques, pipeline de IA/ML, análise de tráfego e criptografia onipresente. Com a necessidade de lidar com o tráfego de forma local e dinâmica em diferentes partes da rede, soluções de segurança distribuídas podem mitigar esses problemas. No entanto, os sistemas de ML, em particular, estão vulneráveis a ataques como envenenamento, injeção, manipulação e corrupção de dados, além da extração e inversão de modelo.

5. Considerações Finais e Agradecimentos

Neste trabalho, buscamos apresentar os principais ataques em redes 6G habilitadas por IA, trazendo contramedidas de segurança e privacidade. Além disso, a segurança e privacidade nas redes móveis deve ser ponto chave, considerando a heterogeneidade, o número massivo de conexões e o poder de aprendizagem empregado. Citamos alguns artigos relevantes nessa área. Com esses artigos, notamos que a IA é recorrente nos estudos sobre 6G. Também pudemos observar as pesquisas latentes e os desafios da área.

Como trabalho futuro, propomos realizar uma prova de conceito para abordar os desafios emergentes que as aplicações 6G enfrentarão. Essa prova de conceito será uma oportunidade para testar a eficácia das tecnologias habilitadoras mencionadas neste estudo e avaliar seus impactos sobre os requisitos e desafios de segurança do 6G, potencialmente identificando novos desafios de pesquisa promissores. Ademais, é válido explorar

soluções distribuídas de segurança para mitigar problemas como ataques DoS, MITM e outros desafios relacionados ao processamento de tráfego e à integridade dos sistemas de ML/IA.

Por fim, agradecemos pelo apoio recebido do projeto INCT INES 2.0 (Instituto Nacional de Ciência e Tecnologia para Engenharia de Software) para a condução desta pesquisa.

Referências

Ji, B., Wang, Y., Song, K., Li, C., Wen, H., Menon, V. G., and Mumtaz, S. (2021). A survey of computational intelligence for 6g: Key technologies, applications and trends. *IEEE Transactions on Industrial Informatics*, 17(10):7145–7154.

Liu, X., Xie, L., Wang, Y., Zou, J., Xiong, J., Ying, Z., and Vasilakos, A. V. (2020). Privacy and security issues in deep learning: A survey. *IEEE Access*, 9:4566–4593.

NGMN Alliance (2023). 6g use cases and analysis.

Nguyen, V.-L., Lin, P.-C., Cheng, B.-C., Hwang, R.-H., and Lin, Y.-D. (2021). Security and privacy for 6g: A survey on prospective technologies and challenges. *IEEE Communications Surveys & Tutorials*, 23(4):2384–2428.

Porambage, P., Gür, G., Osorio, D. P. M., Liyanage, M., Gurtov, A., and Ylianttila, M. (2021a). The roadmap to 6g security and privacy. *IEEE Open Journal of the Communications Society*, 2:1094–1122.

Porambage, P., Gür, G., Moya Osorio, D. P., Livanage, M., and Ylianttila, M. (2021b). 6g security challenges and potential solutions. In *2021 Joint European Conference on Networks and Communications 6G Summit (EuCNC/6G Summit)*, pages 622–627.

Sandeepa, C., Siniarski, B., Kourtellis, N., Wang, S., and Liyanage, M. (2022). A survey on privacy for b5g/6g: New privacy challenges, and research directions. *Journal of Industrial Information Integration*, 30:100405.

Silva, C., Cruz, A., Andrade, R., and Rodrigues, E. (2023). Requisitos e desafios de segurança e privacidade em redes 6g. In *Anais do III Workshop de Redes 6G*, pages 19–24, Porto Alegre, RS, Brasil. SBC.

Siriwardhana, Y., Porambage, P., Liyanage, M., and Ylianttila, M. (2021). Ai and 6g security: Opportunities and challenges. In *2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, pages 616–621. IEEE.

Wang, M., Zhu, T., Zhang, T., Zhang, J., Yu, S., and Zhou, W. (2020a). Security and privacy in 6g networks: New areas and new challenges. *Digital Communications and Networks*, 6(3):281–291.

Wang, M., Zhu, T., Zhang, T., Zhang, J., Yu, S., and Zhou, W. (2020b). Security and privacy in 6g networks: New areas and new challenges. *Digital Communications and Networks*, 6(3):281–291.

Ylianttila, M., Kantola, R., Gurtov, A., Mucchi, L., Oppermann, I., Yan, Z., Nguyen, T. H., Liu, F., Hewa, T., Liyanage, M., et al. (2020). 6g white paper: Research challenges for trust, security and privacy. *arXiv preprint arXiv:2004.11665*.