Análise comparativa de métricas de avaliação em modelos de aprendizado de máquina na detecção de ataques de spoofing de GPS

Ana Carla Rodrigues[1], **Jessica Fileto**[1]

[1] Centro de Matemática, Computação e Cognição
Universidade Federal do ABC (UFABC)
Av. dos Estados, 5001 – 09210-580 – Santo André – SP – Brasil

{carla.rodrigues,jessica.fileto}@ufabc.edu.br

Resumo. *O Veículo Aéreo Não Tripulado (VANT), comumente conhecido como drone, é uma plataforma que está revolucionando indústrias ao redor do mundo, oferecendo soluções de alta flexibilidade e riscos reduzidos. O Global Positioning System (GPS) fornece ao VANT navegação precisa, rastreamento e vôo autônomo. Os VANTs podem sofrer ataques de spoofing nas coordenadas de GPS, comprometendo a segurança da informação e demonstrando que a precisão na análise dos dados são cruciais na tomada de decisões. A inteligência artificial pode ser uma aliada na detecção destes ataques através de técnicas sofisticadas, reduzindo a quantidade de falsos alarmes e fornecendo informações confiáveis por meio do aprendizado de máquina que exerce papel fundamental na classificação destes ataques. A proposta deste projeto é realizar uma análise comparativa das métricas de avaliação dos seguintes modelos de aprendizado de máquina: Random Forest, Naive Bayes, K-NN, Gaussian Naive Bayes e Extremely Randomized Trees, visando identificar aquele que melhor equilibre a detecção correta de spoofing e viabilidade de implementação do modelo.*

1. Introdução

Os veículos aéreos não tripulados (VANTS), também conhecidos como drone, são um tipo de tecnologia fundamental tanto em áreas civis quanto militares. Seu uso reduz a exposição humana a tarefas repetitivas, de longa duração e em alguns casos, de alto risco [Brasil 2025]. Desempenham papel importante nas seguintes áreas: gerenciamento de desastres, vigilância aérea, fotografia aérea de rastreamento, pesquisa e resgate, monitoramento de gado, dentre outros. [Titouna e Naït-Abdesselam 2021]

Por serem operados de maneira autônoma e remota suas coordenadas são essenciais para o sucesso de suas missões, para isso são usados sinais de Sistema Global de Navegação por Satélite (GNSS), sendo o mais conhecido o *Global Positioning System* (GPS), esses sinais podem ser criptografados ou não [A. Faria1 et al. 2018]. Sendo os não criptografados mais suscetíveis a ataques conhecidos como *Spoofing* de GPS [Srinivasan S e Sathyadevan 2023].

Para garantir um vôo seguro do VANT é essencial ter medidas para detecção dos ataques de *Spoofing de GPS*. A identificação precisa desses ataques é extremamente importante, pois técnicas avançadas de *spoofing* podem causar disrupção no funcionamento do VANT, portanto são necessários métodos robustos e adaptativos [İşleyen e Bahtiyar

2024]. Nesse contexto algoritmos de aprendizado de máquina tem se mostrado promissores, pois são capazes de analisar grandes volumetrias de dados, identificando padrões e anomalias. Tais abordagens oferecem uma resposta robusta e adaptativa a essas ameaças contribuindo para as operações dos VANTs [İşleyen e Bahtiyar 2024].

2. Trabalhos relacionados

Diversos estudos recentes tem utilizado essa abordagem para detecção de *spoofing* de GPS. Podemos mencionar algumas abordagens utilizadas:

- **Redes neurais profundas (DNN):** Extração de padrões complexos em séries temporais dos sinais GNSS. [İşleyen e Bahtiyar 2024]
- **Máquinas de vetores de suporte (SVM):**
  Abordagem teve como objetivo a identificação de discrepâncias entre as posições determinadas pelos sinais de GPS e aquelas medidas pelo sistema de navegação inercial (INS). Especificamente, o modelo SVM foi treinado através dos erros obtidos das diferenças posicionais, permitindo que o sistema detecte inconsistências que poderiam indicar um ataque de *spoofing* de GPS. [Panice et al. 2017]
- **Redes neurais convolucionais (CNN):** Em comparação aos modelos tradicionais de aprendizagem de máquina, os de aprendizagem profunda obtiveram alta acurácia na detecção dos ataques de *spoofing*. A grande vantagem desses modelos baseados em aprendizagem profunda, é pelo fato de aprenderem automaticamente a extraírem as *features* sem precisarem de intervenção humana. Além de se adaptarem melhor a *datasets* mais complexos. [Sun et al. 2023]
- **Aprendizagem supervisionada (Baseado em Árvores):** Aissou et al. (2021) utilizou os seguintes modelos baseados em árvores: Random Forest, Gradient Boost, XGBoost e LightGBM para fazer um comparativo de qual seria melhor na detecção dos ataques de *spoofing* de GPS. Sendo que o XGBoost obteve a melhor acurácia (95,52%).
- **IA Generativa:** Abordagem que em comparação com outros modelos de aprendizagem de máquina, se destacou pela eficácia na deteção dos ataques de *spoofing* e *jamming*. [El alami e Rawat 2024]

3. Delimitação e Escopo

O *dataset* que será utilizado como base de dados é o *Mendeley Data* [Aissou 2022], esses dados foram gerados de sinais GPS autênticos e contém aproximadamente 500,000 dados, no artigo escolhido de [Aissou et al. 2021] não possui uma citação do *dataset* usado para o treinamento dos modelos, então para as análises decorrentes deste artigo será utilizada essa fonte de dados simplificada e disponibilizada pelos autores do artigo. De acordo com o artigo de [Aissou et al. 2021], os ataques do tipo *spoofing* podem ser classificados em: simples, intermediários e sofisticados.

Nas próximas seções se estabelecerá a execução do trabalho. Na seção 4, faremos nossa proposta de mudança ao trabalho de [Aissou et al. 2021], discutindo a metodologia e modelos de aprendizado de máquina a serem implementados. Logo, os resultados e discussões serão apresentados na seção 5. Finalmente, nossas conclusões sobre o trabalho serão expostos na seção 6.

4. Metodologia (Proposta de mudança)

A metodologia adotada e os passos seguidos foram inspirados no artigo de [Aissou et al. 2021], que utilizou a implementações de algoritmos como: *Random Forest*, *Gradient Boost*, *XGBoost*, e *LightGBM* com o objetivo de explorar os dados obtidos.

A proposta de mudança deste artigo é avaliar outros modelos para a classificação de possíveis ataques de *spoofing* a sinais de GPS, com enfoque em algoritmos que não são baseados em árvore, sendo estes: (i) *Naive Bayes*, (ii) *K-NN* e (iii) *Gaussian Naive Bayes*.

Além da reimplementação do algoritmo (iv) *Random Forest* que foi utilizado no artigo [Aissou et al. 2021] e obteve o pior desempenho no comparativo. Com base no desempenho do *Random Forest*, será feita a implementação do algoritmo (v) *Extremely randomized trees*, um modelo melhorado do *Random Forest*, que se diferem pelo fato de no *Extremely randomized trees* não existir a fase de *bagging* e no momento da separação dos nós, esta escolha é feita aleatoriamente. **geurtsExtremelyRandomizedTrees2006**

Serão analisados a acurácia, precisão e a quantidade de falsos negativos. Para isso a base será dividida em porções diferentes para testes e o treinamento do modelo, sendo 70% para treino e 30% para testes como no artigo de [Aissou et al. 2021].

Será aplicada a técnica de pré processamento *Principal Component Analysis* (PCA) para melhorar a precisão dos resultados finais, ela consiste em reduzir as dimensões do dataset para deixar apenas as informações mais importantes, removendo assim as redundâncias [IBM 2025]. Esse tipo de processamento ajuda algoritmos de aprendizagem de máquina, simplificando o processo de reconhecimento dos dados, essa técnica é diferente da usada no artigo de [Aissou et al. 2021] que se chama *Spearman Correlation Coefficient*.

Referências

A. Faria1, Lester de et al. (2018). "GPS Jamming Signals Propagation in Free-Space, Urban and Suburban Environments". Em: *Journal Aerospacial Tecnology Managment*. DOI: 10.5028/jatm.v10.870.

Aissou, Ghilas (2022). *A DATASET for GPS Spoofing Detection on Unmanned Aerial System*. Mendeley Data, V3. DOI: 10.17632/z7dj3yyzt8.3.

Aissou, Ghilas et al. (2021). "Tree-based Supervised Machine Learning Models For Detecting GPS Spoofing Attacks on UAS". Em: *2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pp. 0649–0653. DOI: 10.1109/UEMCON53757.2021.9666744.

Brasil, Dialogos Uniao Europeia (2025). *Estudo sobre a Indústria Brasileira e Europeia de Veículos aéreos não tripulados*. Rel. técn. Dialogos Uniao Europeia Brasil. DOI: https://www.gov.br/mdic/pt-br/images/publicacaoa_DRONES-20161130-20012017-web.pdf.

El alami, Hassan e Danda B. Rawat (2024). "DroneDefGANt: A Generative AI-Based Approach for Detecting UAS Attacks and Faults". Em: *ICC 2024 - IEEE International Conference on Communications*, pp. 1933–1938. DOI: 10.1109/ICC51166.2024.10622524. (Acesso em 23/07/2025).

IBM (2025). *What is principal component analysis (PCA)?* URL: https://www.ibm.com/think/topics/principal-component-analysis (acesso em 31/07/2025).

İşleyen, Emre e Şerif Bahtiyar (2024). "GPS Spoofing Detection on Autonomous Vehicles with XGBoost". Em: *2024 9th International Conference on Computer Science and Engineering (UBMK)*, pp. 500–505. DOI: 10.1109/UBMK63289.2024.10773593. (Acesso em 23/07/2025).

Panice, G. et al. (2017). "A SVM-based detection approach for GPS spoofing attacks to UAV". Em: *2017 23rd International Conference on Automation and Computing (ICAC)*, pp. 1–11. DOI: 10.23919/IConAC.2017.8081999.

Srinivasan S, Prasanna e Shiju Sathyadevan (2023). "GPS Spoofing Detection in UAV Using Motion Processing Unit". Em: *2023 11th International Symposium on Digital Forensics and Security (ISDFS)*, pp. 1–4. DOI: 10.1109/ISDFS58141.2023.10131729. (Acesso em 23/07/2025).

Sun, Yichen et al. (2023). "A Deep-Learning-Based GPS Signal Spoofing Detection Method for Small UAVs". Em: *Drones* 7.6. ISSN: 2504-446X. DOI: 10.3390/drones7060370. URL: https://www.mdpi.com/2504-446X/7/6/370.

Titouna, Chafiq e Farid Naït-Abdesselam (2021). "A Lightweight Security Technique For Unmanned Aerial Vehicles Against GPS Spoofing Attack". Em: *2021 International Wireless Communications and Mobile Computing (IWCMC)*, pp. 819–824. DOI: 10.1109/IWCMC51323.2021.9498734. (Acesso em 23/07/2025).