Detecção Proativa de Intrusos em Redes Wi-Fi Utilizando CSI e Machine Learning

Felipe Silveira de Almeida, Eduardo Fabrício Gomes Trindade e Lourenço Alves Pereira Junior
Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos/SP - Brasil

Resumo – Sistemas de Detecção de Intrusão em Redes (NIDS) têm sido amplamente desenvolvidos. No entanto, ataques sofisticados demandam camadas de proteção adicionais. Este estudo propõe utilizar dados do Channel State Information (CSI) Wi-Fi com técnicas de aprendizado de máquina para detecção nas camadas física e de enlace, visando identificar tentativas de acesso não autorizado a uma rede Wi-Fi. Foram analisadas 70.513 instâncias coletadas em dois ambientes e identificadas atividades maliciosas com 94,24% de precisão, utilizando o Random Forest. Esta nova abordagem explora o CSI, incorporando características das camadas 1 e 2 ao contexto de NIDS e permite uma detecção mais precisa e eficaz de atividades maliciosas. O artigo avança o estado da arte ao empregar dispositivos de baixa capacidade computacional, como o ESP32, coletando dados de CSI, demonstrando a viabilidade dessa técnica em ambientes de Internet das Coisas (IoT). A pesquisa apresenta dados rotulados com ações maliciosas e neutras, além de avaliar o desempenho de classificadores. Os resultados mostram que o Random Forest obteve a melhor precisão, destacando-se pela eficácia e velocidade. Este trabalho traz contribuições significativas ao campo de NIDS, propondo uma camada de segurança complementar para redes Wi-Fi, capaz de detectar proativamente ações maliciosas antes que comprometam a rede.

I. INTRODUÇÃO

No âmbito das Operações Militares, o Espaço Cibernético de Interesse exige medidas de defesa interna, englobando ações para eliminar ameaças e mitigar seus efeitos [1-3]. Nesse sentido e considerando o estudo [4], mostrando que o uso de dispositivos móveis é cada vez mais frequente, torna-se evidente a necessidade da utilização de ferramentas que possam monitorar o ambiente de rede sem fio e detectar atividades suspeitas. Assim, os Sistemas de Detecção de Intrusão (IDS) têm sido atualizados, para detectar antecipadamente ameaças modernas. Contudo, a maioria das soluções conhecidas monitoram o ambiente a partir da camada de rede, em outras palavras, tradicionalmente o monitoramento ocorre no momento em que o atacante encontra-se dentro da rede explorada, exigindo uma ação reativa.

Nossa proposta é utilizar dados do *Channel State Information (CSI)* do Wi-Fi capturados por dispositivos com reduzida capacidade computacional e analisar padrões de interferências ocorridas na transmissão do sinal para monitorar ativamente as camadas física e de enlace na área de cobertura de uma *Basic Service Set (BSS)*, permitindo o incremento de uma camada adicional de proteção contra dispositivos não autorizados, complementando os IDSs atuais e trazendo a possibilidade de detecção antes que os atacantes se associarem.

II. CONTRIBUIÇÕES

1. Dataset rotulado com ações maliciosas e neutras.
2. Proposta de sistema complementar para IDS com *features* de camada física e de enlace, voltado à redes militares em Centros de Coordenação de Operações.
3. Análise de padrões de conexão de dispositivos Android e IOS.
4. Avaliação de desempenho de algoritmos *Support Vector Machine (SVM)*, *Random Forest*, *K-Nearest Neighbors (KNN)*, *Random Tree* e *Naive Bayes* para identificação de ações maliciosas.

III. DESENVOLVIMENTO

A Fig. 1 ilustra que o CSI possui propriedades que representam a propagação multipercurso do sinal em amplitude e fase para cada subportadora dentro do domínio de frequência (2,4 GHz – 5 GHz). Essas propriedades trazem características do canal de rádio do sinal Wi-Fi à medida que ele percorre caminhos diretos (LOS) e indiretos (NLOS) entre pares de antenas, como visto na Fig. 2. Consequentemente, esses atributos facilitam a observação de efeitos como atraso de tempo, mudanças de amplitude e de fase no sinal recebido que ficam registrado na Matriz CSI conforme relatado por [5].

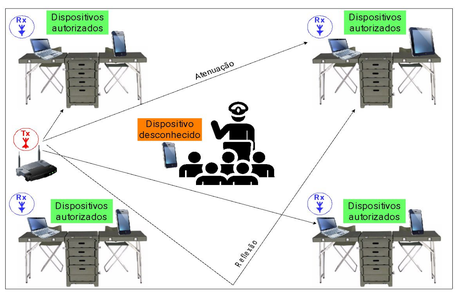

Fig. 1 - Multipercurso do sinal

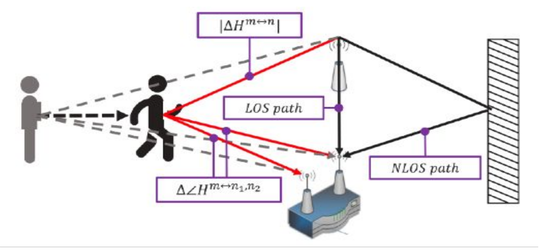

Fig. 2 - Caminhos diretos e indiretos

Recentemente [6] e [7] mostraram que tanto os humanos quanto os dispositivos, possuem características únicas que os diferenciam uns dos outros, incluindo informações de amplitude média e fase, correlação entre subportadoras e interferências no multipercurso. Essas características fornecem uma assinatura distinta para cada dispositivo, como visto em nossos experimentos e exemplificados pelas tentativas de conexão realizadas por equipamentos distintos, nas Fig 3 e 4.

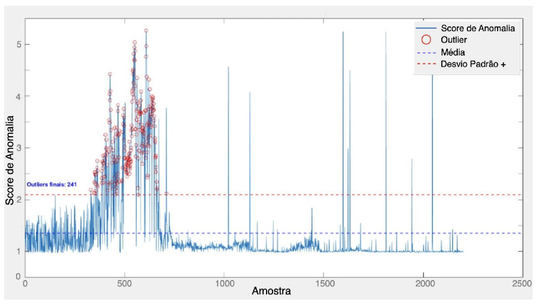

Fig. 3 - Tentativas de conexão do Iphone

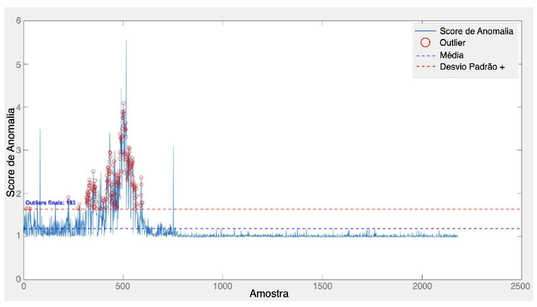

Fig.4 - Tentativas de conexão do Android

No contexto de IoT e pensando em ambientes cada vez mais conectados, aplicamos nossa proposta utilizando um dispositivo ESP32 (Fig. 5), ostentando capacidades para conectividade WiFi e registro de dados CSI. Além disso, eles são programáveis, podendo integrar uma variedade de sensores e funcionar com baterias, garantindo tanto portabilidade quanto flexibilidade.

Em termos gerais, o processo de atividade ou reconhecimento de dispositivos é segmentado em três fases: coleta de dados, extração de recursos e classificação [8]. Por fim, a capacidade de reconhecimento da nossa proposta é medida pela acurácia média obtida por algoritmos de classificação. A partir de uma abordagem supervisionada, representada pela Fig. 6, é possível identificar dispositivos desconhecidos antes mesmo que se associem à rede da BSS monitorada.

Fig. 5 - ESP32

Fig. 6 - Modelo proposto

IV. RESULTADOS

Em nossos experimentos, observou-se que as capturas são influenciadas diretamente pelo ambiente monitorado, variando a correlação dos atributos de acordo com o local. Assim, a partir de uma calibração da ferramenta de captura e de uma seleção de atributos, é possível mapear como cada dispositivo afeta as subportadoras no canal Wi-Fi. Nosso modelo foi treinado com validação cruzada em 10 etapas, trazendo os resultados da Tabela 1. Além disso, verificou-se também que o comportamento dos dispositivos durante as tentativas de conexão são bem distintos, conforme elucidado pelas CDF da Fig. 7.

Tab. 1 - Resultados

Classificador	Precisão (%)	F1-Score	Tempo de Construção (s)
SVM	94.43	0.969	360.79
RandomForest	94.24	0.968	28.67
KNN	92.76	0.959	0.04
RandomTree	90.93	0.948	0.60
NaiveBayes	86.64	0.928	0.13

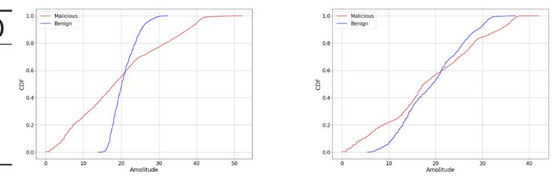

Fig. 7 - CDF

V. CONCLUSÃO

Os resultados mostram que a partir da combinação de aprendizado de máquina e algoritmos de classificação, é possível detectar ações maliciosas. O estudo introduz uma camada de segurança complementar para redes Wi-Fi, capaz de detectar proativamente ações maliciosas por meio de características (features) das camadas física e de enlace. Assim, é possível aplicar a proposta em Centros de Coordenação de Operações ou em qualquer outro ambiente com fluxo de informações sensíveis. A análise dos padrões de conexão de cada dispositivo rotulado viabiliza a distinção proativa entre aqueles autorizados e outros desconhecidos, elevando o nível de proteção contra ações de exploração, ataques do oponente e ainda atividade de evasão de dados. Por fim, a acurácia de 94.24% detecção do modelo com o algoritmo Random Forest, apresenta-se como uma alternativa na prevenção de ações de espionagem e sabotagem, como uma ação de contra inteligência.

REFERÊNCIAS

1. Ministério da Defesa, "Doutrina Militar de Defesa Cibernética," 2ª Edição, Estado-Maior Conjunto das Forças Armadas, 2023. Disponível em: https://mdlegis.defesa.gov.br/pesquisar_normas/.
2. Ministério da Defesa, "Estratégia de Inteligência de Defesa," 1ª Edição, Estado-Maior Conjunto das Forças Armadas, 2023. Disponível em: https://mdlegis.defesa.gov.br/pesquisar_normas/.
3. Ministério da Defesa, "Política de Inteligência de Defesa," 1ª Edição, Estado-Maior Conjunto das Forças Armadas, 2023. Disponível em: https://mdlegis.defesa.gov.br/pesquisar_normas/.
4. Di Gao, Hao Lin, Zhenhua Li, Feng Qian, Qi Alfred Chen, Zhiyun Qian, Wei Liu, Liangyi Gong, e Yunhao Liu, "A Nationwide Census on Wifi Security Threats: Prevalence, Riskiness, and the Economics," (*MobiCom '21)*, DOI: 10.1145/3447993.3448620.
5. Yongsen Ma, Gang Zhou, e Shuangquan Wang, "WiFi Sensing with Channel State Information: A Survey," *ACM Computing Surveys*, vol. 52, no. 3, Art. no. 46, pp. 1-36, Jun. 2020. DOI: 10.1145/3310194.
6. Yi Zhang, Yue Zheng, Guidong Zhang, Kun Qian, Chen Qian, e Zheng Yang, "GaitID: Robust Wi-Fi Based Gait Recognition," em *Wireless Algorithms, Systems, and Applications: 15th International Conference, WASA 2020, Part I 15*, pp. 730–742, Springer, 2020.
7. Ying He, Yan Chen, Yang Hu, e Bing Zeng, "WiFi Vision: Sensing, Recognition, and Detection With Commodity MIMO-OFDM WiFi," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8296-8317, 2020. DOI: 10.1109/JIOT.2020.2989426.
8. Dan Wu, Daqing Zhang, Chenren Xu, Hao Wang, e Xiang Li, "Device-Free WiFi Human Sensing: From Pattern-Based to Model-Based Approaches," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 91-97, 2017. DOI: 10.1109/MCOM.2017.1700143.