# Detecção de Ataques Cibernéticos em Ambiente HTTPS Utilizando Aprendizado Profundo

**Edson B. de Souza[1], Ronaldo Ribeiro Goldschmidt[1], Paulo Ivson N. Santos[2], Paulo César Pellanda[1], Ronaldo Moreira Salles[1,3]**

[1]Departamento de Engenharia de Defesa – Instituto Militar de Engenharia (IME)
Praça Gen. Tibúrcio, 80 - Urca, Rio de Janeiro - RJ

[2]Departamento de Informática – PUC-Rio
Rio de Janeiro, RJ.

[3]CIICESI, ESTG, Instituto Politécnico do Porto, Portugal

{edsonbsouza,ronaldo.rgold,pellanda,salles}@ime.eb.br,
pivson@inf.puc-rio.br

## 1. Introdução

A escolha do *dataset* impacta diretamente a validação de modelos de aprendizado profundo e sua possível aplicabilidade em cenários reais de detecção de intrusão. Surpreendentemente, muitos *datasets* consagrados para tarefas de classificação em tráfego criptografado têm predominância de tráfego em claro, como o ISCXVPN2016 e o USTC-TFC2016, com proporções de tráfego não criptografado chegando a 98,9% e 94,7%, respectivamente, de acordo com [Wickramasinghe et al. 2025]. Mesmo *datasets* mais recentes, como o CICIDS-2017 e o CICIDS-2018, apresentam sérias falhas em sua construção, como rótulos incorretos e problemas nas *flags* dos pacotes capturados e, inclusive, pouco tráfego HTTPS, o que pode comprometer a validade de diversos trabalhos científicos.

Este trabalho aborda uma lacuna identificada na literatura referente à detecção de ataques cibernéticos em tráfego criptografado, por meio de uma arquitetura que assegura a preservação da privacidade do usuário, sem recorrer, por exemplo, à inspeção profunda de pacotes, e um protocolo de validação rígido, com o objetivo de mitigar riscos de *overfitting* e *data leakage* e testada com os *datasets* públicos HIKARI-2021 e CIRA-CIC-DoHBrw-2020. O primeiro é focado em ataques de camada de aplicação, como *probing* e *brute-force*, e o segundo *dataset* especializado na detecção de tunelamento DNS sobre HTTPS.

## 2. Metodologia e Arquitetura Proposta

A normalização com `MinMaxScaler` foi ajustada exclusivamente sobre o conjunto de treino e aplicada às demais partições. Utilizou-se a divisão estratificada em treino/validação/teste (70/15/15), mantendo as proporções de classe e o balanceamento com SMOTE foi aplicado somente ao treino para evitar *data leakage*.

A arquitetura híbrida proposta inspirada em [Liu et al. 2025] combina redes neurais convolucionais com mecanismos de atenção ECANet (*Efficient Channel Attention Network*) e módulos *Transformer Encoder*. O modelo contém três blocos CNN (32/64/128 filtros, kernel $3\times3$, ReLU e *max–pooling* $2\times2$), intercalados com módulos ECANet para recalibração por canal com baixo custo computacional. Em seguida, um *Transformer Encoder* com 4 cabeças de atenção, *dropout* de 0,3, *feed–forward* de 256

unidades e conexões residuais captura dependências globais entre as representações. Camadas densas com regularização L2 ($\lambda = 0{,}01$) e *dropout* de 0,3 antecedem a saída sigmoide. O treinamento utilizou Adam com taxa de $10^{-3}$, 100 épocas e *batch* de 64, além de *EarlyStopping* e *ReduceLROnPlateau* para estabilizar a otimização e mitigar *overfitting*.

Uma diferença crucial em relação a [Liu et al. 2025] está no foco dos dados e no na tarefa de classificação. Enquanto o presente trabalho utiliza *datasets* voltados especificamente para a detecção de ataques cibernéticos em tráfego encriptado, [Liu et al. 2025] emprega o *dataset* ISCX VPN-nonVPN, para diferenciação entre VPN e não VPN. Por fim, o presente trabalho remove identificadores fortes como IP, portas e protocolos, enquanto [Liu et al. 2025] os mantém, o que pode comprometer a generalização e tornar o modelo mais dependente de elementos específicos dos dados, conforme recomendado em [Lotfollahi et al. 2020, Wickramasinghe et al. 2025].

## 3. Resultados e Discussão

A fim de efetuar uma comparação experimental mais justa com o trabalho de [Liu et al. 2025], foram utilizados os mesmos *datasets* para ambos. No conjunto de teste, a arquitetura proposta atingiu uma AUC de 92,7%, *recall* de 99,5%, precisão de 87,6%, *F1-score* de 93,2% e acurácia de 92,7% e o de [Liu et al. 2025] 89,7%, 94,4%, 82,5%, 88,1% e 89,7%, respectivamente, no *dataset* HIKARI-2021. Já no *dataset* CIRA-CIC-DoHBrw-2020, o modelo proposto também superou o de [Liu et al. 2025] em todas as métricas avaliadas: AUC (97,9% vs. 94,3%), *Recall* (98,9% vs. 93,6%), Precisão (97,1% vs. 91,8%), *F1-score* (97,9% vs. 92,6%) e Acurácia (97,9% vs. 94,3%). Tais métricas demonstram a capacidade do modelo em capturar eficientemente os padrões de tráfego malicioso consolidando a arquitetura proposta como mais promissora para aplicação em cenário real.

A precisão menor obtida no HIKARI-2021, em comparação com o outro *dataset*, indica a presença de falsos positivos que poderiam ser mitigados com técnicas de XAI (*eXplainable Artificial Intelligence*). Poder-se-ia, também, avaliar futuramente a robustez do modelo contra ataques adversariais. Por fim, a análise de sensibilidade dos hiperparâmetros poderia testar a estabilidade e a robustez do modelo de modo a otimizar o classificador e melhorar sua capacidade de generalização.

## Referências

Liu, Z., Xie, Y., Luo, Y., Wang, Y., and Ji, X. (2025). Transeca-net: A transformer-based model for encrypted traffic classification. *Applied Sciences*, 15(6):2977.

Lotfollahi, M., Jafari Siavoshani, M., Shirali Hossein Zade, R., and Saberian, M. (2020). Deep packet: A novel approach for encrypted traffic classification using deep learning. *Soft Computing*, 24(3):1999–2012.

Wickramasinghe, N., Shaghaghi, A., Tsudik, G., and Jha, S. (2025). Sok: Decoding the enigma of encrypted network traffic classifiers. In *2025 IEEE Symposium on Security and Privacy (SP)*, pages 1825–1843, San Francisco, CA, USA. IEEE.