16º Congresso de Inovação, Ciência e Tecnologia do IFSP - 2025

DETECÇÃO DE ANOMALIAS EM REDES INDUSTRIAIS COM ALGORITMOS DE APRENDIZADO DE MÁQUINA

MARIA EDUARDA GUEDES DOS SANTOS[1], RHAMZA MOURAD MOURAD[2], TÚLIO SANTOS DE SOUZA[3]

[1] Graduando em Bacharelado em Ciência da Computação, Câmpus Salto, e.guedes@aluno.ifsp.edu.br.
[2] Graduando em Bacharelado em Ciência da Computação, Câmpus Salto, rhamza.mourad@aluno.ifsp.edu.br.
[3] Orientador IFSP, Câmpus Salto, tulio.souza@aluno.ifsp.edu.br.

Área de conhecimento: 1.03.03.04-9 Sistemas de Informação

RESUMO: A crescente integração de sistemas na Indústria 4.0 ampliou a vulnerabilidade de redes industriais a ciberataques, tornando as abordagens de segurança tradicionais insuficientes. O objetivo deste trabalho é avaliar e comparar a eficácia de algoritmos de Aprendizado de Máquina e Aprendizado Profundo na detecção de anomalias em tráfego de redes industriais, especificamente no protocolo Modbus, para antecipar e mitigar ameaças. A metodologia adotada inclui o pré-processamento de um conjunto de dados específico, seguido pela implementação, treinamento e validação de quatro modelos: Support Vector Machine, K-Nearest Neighbors, Long Short-Term Memory e Redes Neurais Convolucionais. Os resultados preliminares indicam que, após a primeira rodada de experimentos, os modelos SVM e KNN demonstraram boa acurácia na identificação de anomalias, com taxas de detecção próximas a 99,99% para ambos. A pesquisa busca contribuir com uma solução de segurança mais robusta e adaptativa para os desafios cibernéticos do ambiente industrial, reduzindo os impactos operacionais decorrentes de ataques.

PALAVRAS-CHAVE: Segurança Cibernética, Indústria 4.0; Detecção de Anomalias; Aprendizado de Máquina; Redes Industriais.

ANOMALY DETECTION IN INDUSTRIAL NETWORKS WITH MACHINE LEARNING ALGORITHMS

ABSTRACT: The growing integration of systems in Industry 4.0 has broadened the vulnerability of industrial networks to cyberattacks, making traditional security approaches insufficient. The objective of this work is to evaluate and compare the effectiveness of Machine Learning and Deep Learning algorithms in detecting anomalies in industrial network traffic, specifically in the Modbus protocol, to anticipate and mitigate threats. The adopted methodology includes the pre-processing of a specific dataset, followed by the implementation, training, and validation of four models: Support Vector Machine, K-Nearest Neighbors, Long Short-Term Memory and Convolutional Neural Networks. Preliminary results from the first round of experiments show that SVM and KNN models achieved overall high accuracy in detecting anomalies, with detection rates of approximately 99.99% each. Ultimately, the research seeks to contribute with a more robust and adaptive security solution for the cybersecurity challenges of the industrial environment, reducing the operational impacts resulting from attacks.

KEYWORDS: Cybersecurity; Industry 4.0; Anomaly Detection; Machine Learning; Industrial Networks.

INTRODUÇÃO

Com o avanço da Indústria 4.0, o aumento da conexão de sistemas industriais à internet e a adoção da Internet Industrial das Coisas (IIoT) geraram novos desafios de segurança cibernética (Karnik et al., 2021 apud Toussaint; Krima; Panetto, 2024). As abordagens tradicionais, como os Sistemas de Detecção de Intrusão (SDI), tornaram-se insuficientes para proteger ambientes industriais complexos contra ameaças cada vez mais sofisticadas (Ahsan et al., 2022). A manufatura, em particular, tornou-se o setor mais visado por ciberataques (IBM, 2025), destacando a urgência de novas soluções de segurança. A ciência de dados está impulsionando a transformação, onde o Aprendizado de Máquina (AM), um aspecto essencial da Inteligência Artificial, pode desempenhar um papel vital na descoberta de padrões ocultos nos dados.(Ahsan et al., 2022)

Neste contexto, este trabalho propõe a aplicação de algoritmos de AM e Aprendizado Profundo (AP) para a detecção de anomalias em redes industriais. O objetivo é testar e avaliar o desempenho de modelos como *Support Vector Machines* (SVM), *K-Nearest Neighbors* (KNN), *Long Short-Term Memory* (LSTM) e *Convolution Neural Networks* (CNN) para identificar tráfego malicioso em redes que utilizam o protocolo Modbus. A pesquisa busca desenvolver uma solução de alto desempenho para fortalecer a segurança cibernética no ambiente industrial, antecipando ataques e reduzindo seus impactos operacionais.

MATERIAL E MÉTODOS

Para a execução deste estudo, será utilizado um computador com a linguagem de programação Python para implementar a solução e realizar as análises. O projeto será desenvolvido sobre um conjunto de dados e bibliotecas de AM e AP específicas.

1. **Conjunto de Dados**
   O estudo utilizará o CIC Modbus dataset (Kwasi Boakye-Boateng et al., 2023). Este conjunto de dados é composto por capturas de tráfego de rede (.pcap) geradas em uma plataforma de teste simulada *Wireshark*, contendo tanto comunicações benignas quanto maliciosas. Os dados de ataque simulam diversas ameaças direcionadas ao protocolo Modbus, como reconhecimento, *query flooding*, injeção de dados falsos e *baseline replay*. Os dados benignos representam o tráfego normal de um ambiente de subestação simulada.

2. **Linguagem e Bibliotecas**
   A implementação será realizada em Python, utilizando bibliotecas para AM e AP, como Scikit-learn (para os modelos SVM e KNN) e TensorFlow ou Keras (para as redes neurais CNN e LSTM).

3. **Procedimentos Metodológicos**
   O fluxo de trabalho, conforme ilustrado na Figura 1 do projeto, seguirá as seguintes etapas:
   a. Extração e Pré-processamento: As informações serão extraídas dos arquivos de captura. Em seguida, os dados passarão por um tratamento que inclui, remoção de ruídos, tratamento de valores corrompidos e codificação de variáveis. O conjunto de dados será dividido em 70% para treinamento, 15% para validação e 15% para teste, garantindo uma avaliação robusta dos modelos. As features selecionadas serão normalizadas utilizando o método *StandardScaler* para otimizar o desempenho dos algoritmos.

   b. Implementação dos Modelos: Serão implementados os quatro modelos de AM e AP propostos: SVM, KNN, LSTM e CNN. Para o SVM, será utilizado um kernel RBF com C=1.0 e gamma='scale'. O KNN será configurado com K=5 e métrica de distância euclidiana. As redes neurais LSTM e CNN terão arquiteturas otimizadas, com camadas de entrada, ocultas e de saída, e serão treinadas com o otimizador Adam e função de perda binary cross-entropy por 50 épocas, com um tamanho de batch de 32.

c. Avaliação e Comparação: O desempenho de cada modelo será avaliado com base em métricas como acurácia, *F1-score*, taxas de falsos positivos e falsos negativos, e complexidade de tempo. Por fim, os resultados serão comparados para determinar o modelo mais eficaz para a detecção de anomalias no cenário proposto.

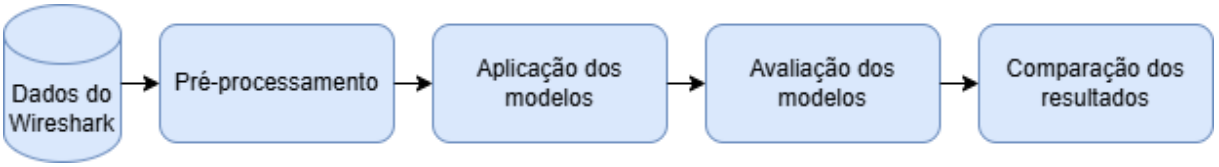

FIGURA 1. Fluxograma do modelo de desenvolvimento da solução.

RESULTADOS E DISCUSSÃO

O estudo ainda está em andamento. No entanto, são apresentados os resultados obtidos a partir da análise exploratória do conjunto de dados, seguidos pelos resultados dos modelos de Machine Learning SVM e KNN.

O trabalho parte de uma base de dados composta por 150 arquivos de captura de rede no formato .pcap. Estes arquivos são a única fonte de verdade para a análise e estão categorizados como tráfego "benigno" (rotulado como 0) ou "ataque" (rotulado como 1). Os incidentes maliciosos, por sua vez, são subclassificados em três tipos principais: compromised-ied, compromised-scada e external. A distribuição quantitativa desses arquivos, conforme detalhado na Figura 2, revela um predomínio de tráfego relacionado a ataques, o que fornece um conjunto de dados rico e desafiador para o desenvolvimento de modelos de detecção de anomalias. A ausência de uma correlação direta com fontes de log externas direciona o foco da análise para as características intrínsecas do tráfego de rede.

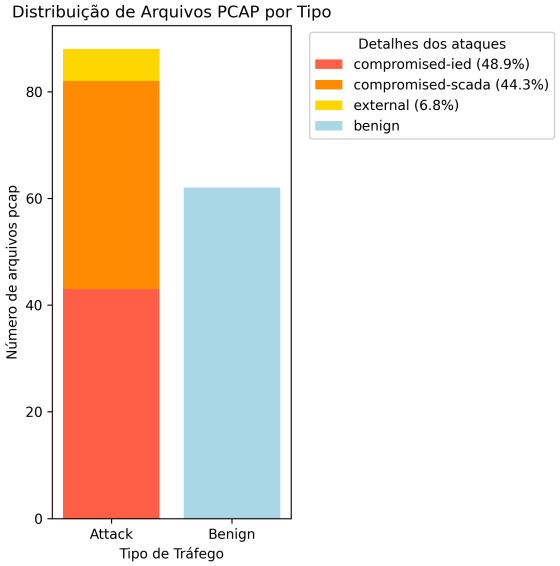

Figura 2 – Distribuição dos arquivos de captura de rede (.pcap) por tipo de tráfego. Elaborado pelos Autores.

Diante das características do tráfego e dos ataques, a fase subsequente concentrou-se em no processo de engenharia e seleção de features. Foram selecionadas sete features iniciais com base em sua relevância teórica para a detecção de anomalias: tamanho do pacote, atraso entre pacotes, tempo de processamento, comprimento do quadro Modbus, código de função, overhead do protocolo e os dados da payload. A validade dessa escolha foi confirmada por meio de uma Análise Exploratória de Dados (AED), que mostra assinaturas estatísticas e padrões visuais distintos que diferenciam claramente as operações normais das anômalas, estabelecendo uma base sólida para o alto potencial preditivo do

conjunto de features. Para enriquecer ainda mais os modelos, foram criadas novas features derivadas, como a classificação binária de pacotes (*attack*), variáveis temporais (hora, minuto, tempo em segundos, dia da semana) com suas respectivas representações em seno e cosseno para capturar a ciclicidade, e a entropia da *payload*, uma métrica importante para quantificar a aleatoriedade dos dados. O pré-processamento foi finalizado com a remoção de *flags* TCP redundantes e da *payload* Modbus, seguida pelo *encoding* das flags TCP restantes (utilizando *multi-hot encoding*, já que um pacote pode ter múltiplas *flags* ativas) e do código de função (*one-hot encoding*).

Com um conjunto de dados robusto e bem preparado, a metodologia de treinamento e otimização dos modelos foi implementada. A pipeline de AM iniciou-se com a normalização das variáveis numéricas utilizando *StandardScaler*, essencial para o bom desempenho de muitos algoritmos de ML. Em seguida, o *dataframe* foi separado para balancear a proporção entre pacotes de ataque e benignos antes da divisão em conjuntos de treino e teste. A seleção de atributos foi refinada com o método *SelectKBest*, que emprega testes estatísticos univariados para identificar as variáveis mais informativas e reduzir a dimensionalidade. Foram então aplicados e otimizados dois algoritmos supervisionados, KNN e SVM, utilizando uma busca em grade (*Grid Search*) com validação cruzada para garantir a generalização dos modelos. A capacidade de generalização e a estabilidade do treinamento foram verificadas visualmente através da análise das Curvas de Aprendizado, presentes na Figura 3. Para ambos os algoritmos, as curvas de treino e validação apresentam uma convergência estável, indicando que os modelos não sofrem de alto viés ou alta variância e se beneficiam do aumento no volume de dados de treinamento.

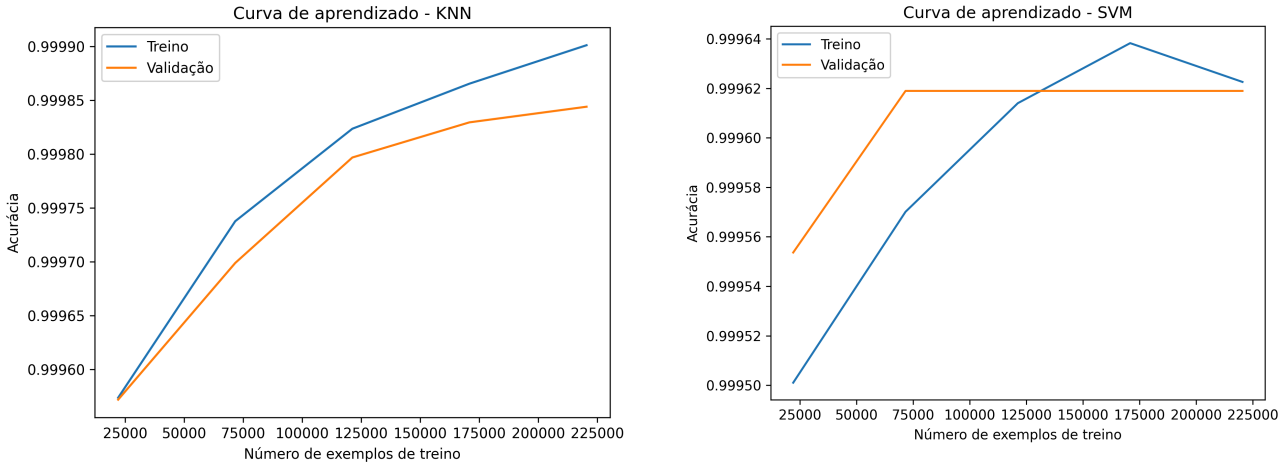

Figura 3 – Curva de aprendizado das AM KNN e SVM. Elaborado pelos Autores.

A avaliação de desempenho dos modelos foi abrangente, e os resultados quantitativos, particularmente para o modelo KNN, foram excepcionais, alcançando uma acurácia geral de 0.999915. Para complementar esses números e oferecer uma visão qualitativa dos erros de classificação, as Matrizes de Confusão para ambos os modelos foram geradas. A análise das matrizes na Figura 4 valida a alta eficácia dos classificadores: o modelo KNN apresentou apenas 2 falsos positivos e 3 falsos negativos, enquanto o SVM, embora também muito preciso, registrou 14 falsos negativos e nenhum falso positivo, permitindo uma visualização clara do comportamento de cada modelo na distinção entre tráfego benigno e de ataque. Uma análise comparativa final, ilustrada na Figura 5, mostra que o KNN obteve um desempenho ligeiramente superior em acurácia e *F1-score*, apesar de um tempo de treinamento consideravelmente maior.

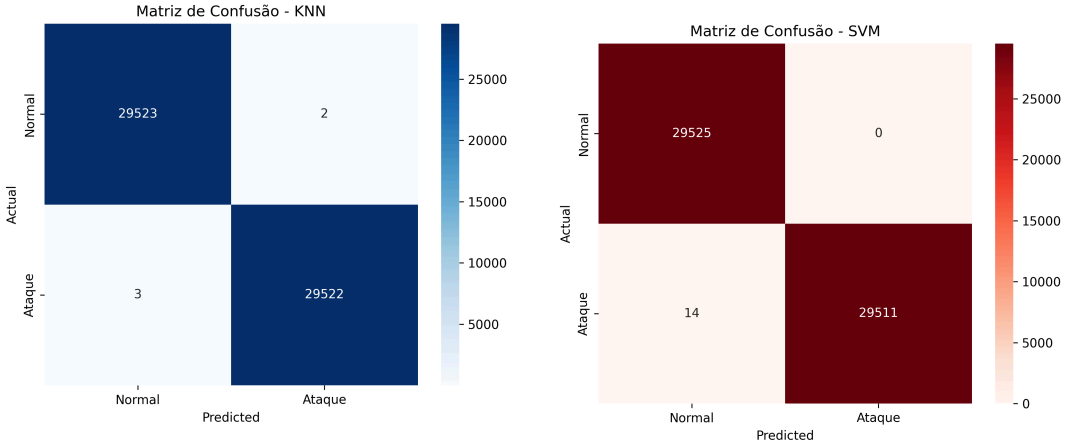

Figura 4 – Matriz de confusão das AM KNN e SVM. Elaborado pelos Autores.

Ainda, foi realizada uma análise comparativa do desempenho dos modelos KNN e SVM, por meio de um gráfico comparativo, baseando-se nas métricas propostas pelo trabalho. Conforme ilustrado na Figura 5, embora ambos os modelos apresentem acurácia e *F1-score* elevados, observa-se uma leve diferença entre eles, com o KNN obtendo desempenho ligeiramente superior nessas métricas. Quanto às taxas de falsos positivos e falsos negativos, o KNN apresenta desempenho inferior no primeiro, mas melhor no segundo. Por fim, a análise do tempo de treinamento evidencia que o SVM, utilizando o classificador linear, apresenta maior eficiência computacional (medida por tempo de treinamento) em relação ao KNN.

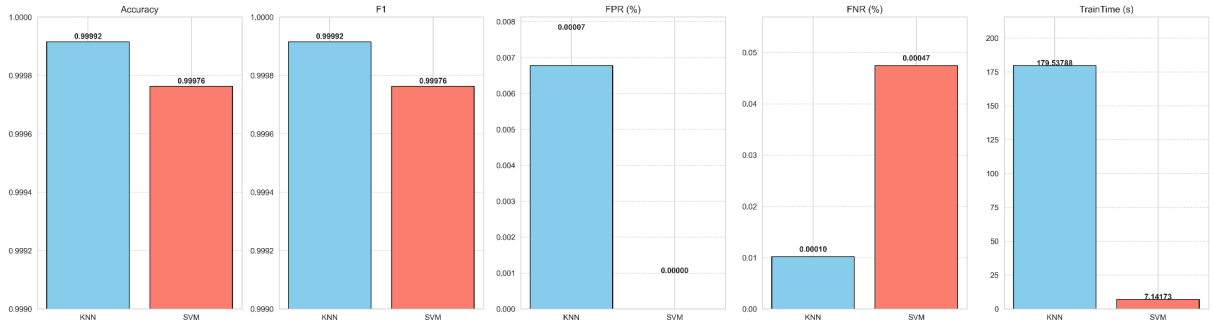

Figura 5 – Análise comparativa das métricas e do tempo de treinamento dos modelos de AM KNN e SVM com gráfico de barras. Elaborado pelos Autores.

CONCLUSÕES

Este trabalho apresentou as etapas iniciais focando na análise de dados e engenharia de features. A análise exploratória validou a seleção de sete características de rede ao demonstrar que o tráfego anômalo possui assinaturas estatísticas distintas do tráfego legítimo, estabelecendo uma base sólida para a modelagem preditiva. Ainda, foi possível treinar os modelos tradicionais de AM propostos, SVM e KNN, com uma amostra representativa dos dados de captura de rede. Como próximos passos, propõe-se a criação e treinamento das arquiteturas CNN e LSTM de AP, e, posteriormente, o aumento da amostra para verificação mais precisa dos modelos.

CONTRIBUIÇÕES DOS AUTORES

Maria Eduarda Guedes dos Santos e Rhamza Mourad Mourad contribuíram com a modelagem, desenvolvimento e aplicação dos estudos e experimentos apresentados. Túlio de Santos Souza contribuiu com a organização e orientação do projeto. Todos os autores contribuíram para o desenvolvimento do trabalho.

AGRADECIMENTOS

Agradecemos ao Instituto Federal por apoiar este estudo.

REFERÊNCIAS

AHSAN, Mostofa; NYGARD, Kendall E.; GOMES, Rahul; CHOWDHURY, Md Minhaz; RIFAT, Nafiz; CONNOLLY, Jayden F. Cybersecurity Threats and Their Mitigation Approaches Using Machine Learning—A Review. Journal of Cybersecurity and Privacy, [S. l.], v. 2, n. 3, p. 527–555, 2022. DOI: https://doi.org/10.3390/jcp2030027.

GHOSH, T. et al. Anomaly Detection for Modbus over TCP in Control Systems Using Entropy and Classification-Based Analysis. Journal of Cybersecurity and Privacy, v. 3, n. 4, p. 895–913, 1 dez. 2023. DOI: https://doi.org/10.3390/jcp3040041

IBM. IBM X-Force 2025 Threat Intelligence Index. 2025. Disponível em: https://www.ibm.com/thought-leadership/institute-business-value/report/2025-threat-intelligence-index.

KARNIK, Niharika; BORA, Urvi; BHADRI, Karan; KADAMBI, Prasanna; DHATRAK, Pankaj. A comprehensive study on current and future trends towards the characteristics and enablers of industry 4.0. Journal of Industrial Information Integration, [S. l.], v. 27, p. 100294, 2021. DOI: https://doi.org/10.1016/j.jii.2021.100294.

KWASI BOAKYE-BOATENG; GHORBANI, Ali A.; ARASH HABIBI LASHKARI. Securing Substations with Trust, Risk Posture, and Multi-Agent Systems: A Comprehensive Approach. 20th International Conference on Privacy, Security and Trust (PST), [S. l.], p. 1–12, 2023. DOI: https://doi.org/10.1109/pst58708.2023.10320154.