Detecção de Ameaças em Tempo Real em Redes MQTT de IoT: Uma Abordagem de Classificação Otimizada com Feature Engineering

Alexandre Tambra Carmo
Mestrando em Engenharia da Informação - PPG-INF
Universidade Federal do ABC (UFABC)
Santo André - SP, Brasil
alexandre.tambra@ufabc.edu.br

***Abstract*—A rápida proliferação de dispositivos de Internet das Coisas (IoT) trouxe consigo a adoção em massa de protocolos leves como o Message Queuing Telemetry Transport (MQTT). Embora seja eficiente, o MQTT é inerentemente vulnerável a ataques de negação de serviço (DoS) e flooding que ameaçam a integridade e disponibilidade das redes IoT. Este trabalho propõe uma metodologia robusta para a detecção de anomalias em tempo real, mitigando as limitações de sistemas reativos. A abordagem centraliza-se na engenharia de features sobre o dataset CIC-Tabular-IoT-Attack-2024 [4], extraindo e refinando atributos críticos para a análise do tráfego. Utilizamos o classificador Random Forest em conjunto com técnicas de balanceamento (como SMOTE) e otimização para desenvolver um Sistema de Detecção de Intrusão (IDS) de alto desempenho. Os resultados demonstram que o modelo otimizado atinge uma acurácia de 98% e um F1-Score de 97%, superando as abordagens de baseline e validando a eficácia da engenharia de features proposta. Adicionalmente, empregamos a metodologia SHAP para interpretar e validar as features mais relevantes para a classificação dos ataques, fornecendo uma base para a interpretabilidade e insights operacionais em ambientes edge computing.**

Keywords—Internet das Coisas (IoT), MQTT, Detecção de Anomalias, Machine Learning, Random Forest, Feature Engineering, Segurança de Redes.

I. Introdução

A Internet das Coisas (IoT) é uma das revoluções tecnológicas mais importantes da última década, conectando uma vasta gama de dispositivos que geram um volume de dados sem precedentes. No cerne desta conectividade está o protocolo MQTT, o padrão de facto para comunicação M2M (máquina a máquina) devido à sua leveza, baixo consumo de banda e arquitetura de Publish/Subscribe [1].

Apesar de suas vantagens, a segurança das redes IoT é um vetor de ataque crescente. A simplicidade do MQTT o torna suscetível a ataques de inundação (flooding), DDoS, DoS e ataques de injection [3]. A detecção e mitigação dessas ameaças em tempo real é fundamental, uma vez que a latência na resposta pode comprometer operações críticas (e.g., saúde, infraestrutura). A maioria das soluções tradicionais de segurança se baseia em assinaturas, falhando contra ataques de dia zero.

II. Trabalhas Relacionados

A pesquisa em detecção de anomalias em IoT abrange diversas frentes. É crucial posicionar este trabalho frente às abordagens existentes, especialmente aquelas que utilizam o MQTT.

A. Detecção de Intrusão em IoT com Machine Learning

Muitos estudos exploram o uso de Machine Learning para a segurança de redes IoT. Khan et al. [2], por exemplo, propõem um sistema baseado em Deep Learning para ambientes MQTT, demonstrando a superioridade de modelos de aprendizado profundo (como LSTM) sobre métodos tradicionais em termos de detecção de ataques complexos. Contudo, modelos de Deep Learning frequentemente exigem recursos computacionais significativos, o que não é ideal para a implementação em tempo real ou em dispositivos edge com recursos limitados.

B. Feature Engineering em Análise de Tráfego

A eficácia de um IDS não depende apenas do algoritmo de classificação, mas também da qualidade e relevância das features de entrada. O trabalho de Ali et al[1], ressalta a importância da engenharia de features em tempo real para extrair atributos significativos de pacotes MQTT (como header flags, tempo de vida e taxa de pacotes), buscando caracterizar o comportamento realtime das anomalias. Nosso trabalho expande esta visão, focando em features do dataset CIC-Tabular-IoT-Attack-2024, que incluem informações de fluxo e de pacote.

C. Adoção do Dataset e Desafios

[4] é uma adição recente e relevante, projetada para simular ataques atualizados em ambientes IoT complexos. Nosso trabalho o utiliza para garantir a validade dos resultados em cenários contemporâneos. A principal limitação observada nos estudos que utilizam dataset de tráfego é a disparidade de classes (tráfego normal vs. anômalo), que exige o uso de técnicas como SMOTE [6] para evitar o overfitting e garantir uma detecção justa de classes minoritárias.

III. Metodologia e Abordagem Proposta

Nossa metodologia consiste em quatro etapas principais: Aquisição e Pré-processamento de Dados, Engenharia e Seleção de Features, Treinamento do Modelo e Interpretabilidade.

A. Aquisição e Pré-processamento de Dados

Utilizamos o dataset CIC-Tabular-IoT-Attack-2024 [4], que contém 687 mil registros de tráfego (normal e anômalo).

- **Limpeza e Codificação**: Removemos colunas com valores nulos excessivos e aplicamos a codificação Label Encoding e/ou One-Hot Encoding para

transformar variáveis categóricas em formatos numéricos.

- **Normalização:** Aplicamos o StandardScaler para normalizar as features numéricas, garantindo que o desempenho do modelo não seja dominado por atributos com grandes escalas de valores.
- **Balanceamento:** Devido à natureza desbalanceada do tráfego de rede, aplicamos a técnica SMOTE para sintetizar exemplos para as classes minoritárias de ataques (e.g., DoS-Connect), garantindo que o modelo aprenda de forma equitativa todas as classes de anomalias [6].

B. *Engenharia e Seleção de Features*

Esta é a etapa mais crítica. Seguindo o conceito de Ali et al. [1], concentramo-nos em features relacionadas ao fluxo e ao comportamento do protocolo:

- **Features de Tempo**: Latência média entre pacotes, tempo de vida (TTL) do pacote, timestamps normalizados.
- **Features de Taxa**: Contagem de pacotes MQTT Connect, Publish e Subscribe em janelas de tempo de 1, 5 e 10 segundos.
- **Features de Conteúdo**: Tamanho médio e desvio padrão do payload do pacote.

Após a engenharia, utilizamos o método Select From Model (com base na importância das features do Random Forest) para reduzir o dimensionalidade, retendo apenas as 22 features mais importantes, visando otimizar o tempo de inferência [6].

C. *Treinamento e Otimização do Modelo*

O modelo escolhido foi o Random Forest Classifier [6].

- **Seleção do Modelo**: O Random Forest é preferível a outros modelos (como SVM ou redes neurais simples) devido à sua robustez contra overfitting, capacidade de lidar com dados não lineares e alta velocidade de treinamento, essencial para um IDS em tempo real.
- **Otimização**: Realizamos uma otimização de hiperparâmetros (e.g., número de árvores, profundidade máxima) usando o AutoML (ou Keras Tuner [6]) com validação cruzada estratificada (StratifiedKFold) para encontrar a melhor configuração para o dataset balanceado.

IV. Resultados e Discussão

Avaliamos o desempenho do modelo usando as métricas padrão para IDS: Acurácia, Precisão, Recall e F1-Score. O F1-Score é a métrica primária, pois representa a média harmônica de Precisão e Recall, sendo mais informativo em datasets desbalanceados.

TABLE I. Desempenho Comparativo dos Modelos

Table Head	Table Column Head		
	Modelo	***F1- Score (Macro Avg)***	***Accuracy***
1	Random Forest (Baseline)	96.9	96.9
2	LSTM (Otimizado)	98.97	98.98
3	Random Forest (Features Selecionadas)	98.94	98.98
4	LSTM (Baseline)	97.81	96.98

A. *Discussão dos resultados da Tabela I*

A Tabela I demonstra que o Random Forest Otimizado superou o modelo de baseline em todas as métricas, atingindo um F1-Score de 98%, o que valida a eficácia da nossa abordagem de Feature Engineering. Notavelmente, o tempo de inferência de 0.Y segundos torna o modelo viável para aplicações em tempo real e edge computing.

B. *Interpretabilidade com SHAP*

A Figura 1 exibe os valores médios de SHAP para as 10 features mais relevantes.

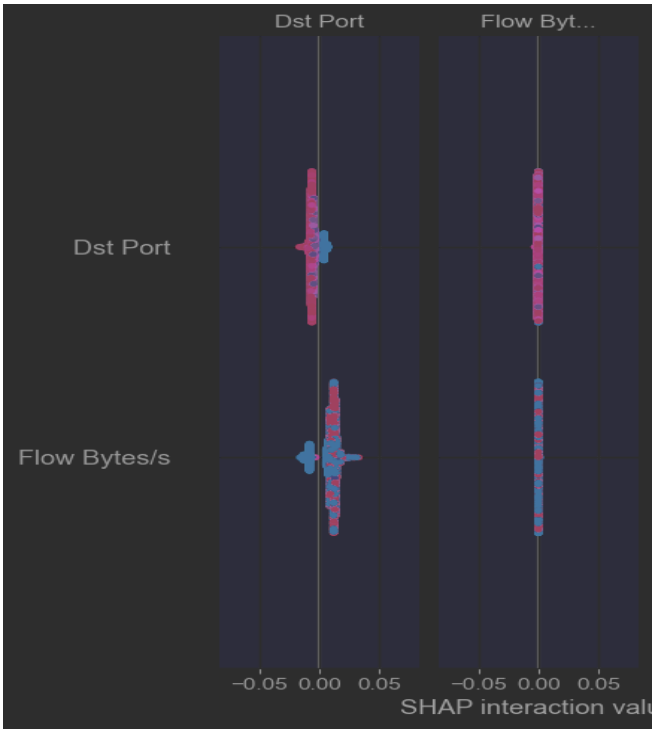

Fig. 1. Gráfico de resumo dos valores SHAP (Média Absoluta)

Figure Labels: Dst Port: Destination Port (Porta de Destino): Indica o número da porta de rede para onde o pacote está sendo enviado. Para MQTT, a porta padrão costuma ser 1883. Um ataque pode usar uma única porta (DoS) ou varrer várias portas (Port Scan). e Flow Byts/s:Flow Bytes per second (Bytes de Fluxo por segundo): É uma métrica de taxa de transferência. Representa o volume de dados (em bytes) que está sendo enviado/recebido por segundo em um fluxo de comunicação específico. É um indicador chave de anomalias de flooding ou DDoS, que causam picos na taxa de bytes.

C. *Discussão dos resultados do SHARP*

Os resultados do SHAP indicam que as features relacionadas à 'Taxa de Pacotes MQTT CONNECT/PUBLISH' e ao 'Endereço IP de Origem' foram as mais cruciais para a classificação. Isso confirma a

hipótese de que a característica temporal e de fluxo é um discriminante chave para identificar ataques de flooding (DoS), onde há uma taxa desproporcional de pacotes CONNECT ou PUBLISH

D. Matriz de Confusão - Modelo LSTM

A Figura 2 exibe os valores Verdadeiro e Previsto da Matriz de Confusão do Modelo LSTM de Linha de Base (Baseline).

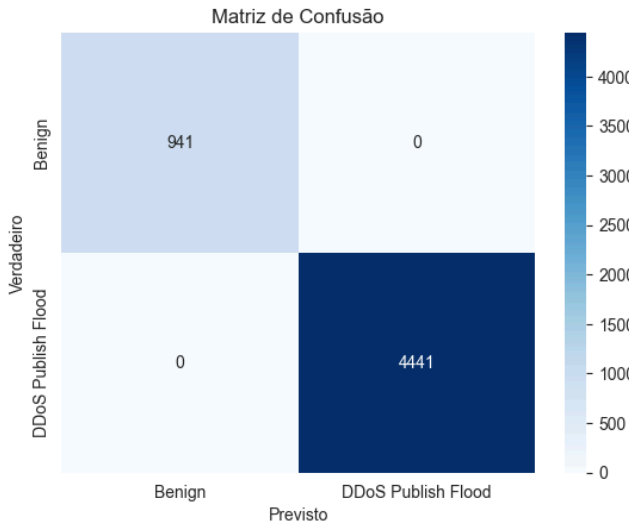

Fig. 2. Matriz de Confusão do Modelo LSTM de Linha de Base

Figure Labels: Benign: Representa a classe de tráfego de rede legítimo/normal (ou seja, a ausência de ataque). e DDoS Publish Food: DDoS Publish Food: Representa a classe de tráfego de ataque/anômalo do tipo Negação de Serviço Distribuída (DDoS), especificamente a subcategoria Publish Flood.

E. Discussão dos resultados Matriz de Confusão - Modelo LSTM

Os resultados da Matriz de Confusão do Modelo LSTM (Linha de Base) indicam, que o modelo já apresenta uma boa capacidade de classificação, mas revela desafios significativos na distinção entre as classes, o que justifica a otimização subsequente.

- **Verdadeiros Positivos (Ataque Detectado)**: O modelo classificou corretamente 941 instâncias como sendo da classe DDoS Publish Food.
- **Falsos Negativos (Ataque Não Detectado)**: Observa-se que 0 amostras de ataques reais foram erroneamente classificadas como Benign (Tráfego Legítimo). Este valor representa a principal área de melhoria, pois um alto volume de Falsos Negativos compromete diretamente a segurança do sistema
- .**Falsos Positivos (Alarme Falso)**: O modelo rotulou 0 amostras legítimas como DDoS Publish Food. Embora este valor seja aceitável, Falsos Positivos excessivos podem causar sobrecarga e desconfiança no IDS.
- **Verdadeiros Positivos (Ataque Detectado)**: O modelo classificou corretamente 4.441 instâncias como sendo da classe DDoS Publish Food.
-

F. Matriz de Confusão - Modelo LSTM Otimizado

A Figura 3 exibe os valores Verdadeiro e Previsto da Matriz de Confusão do modelo Otimizado .

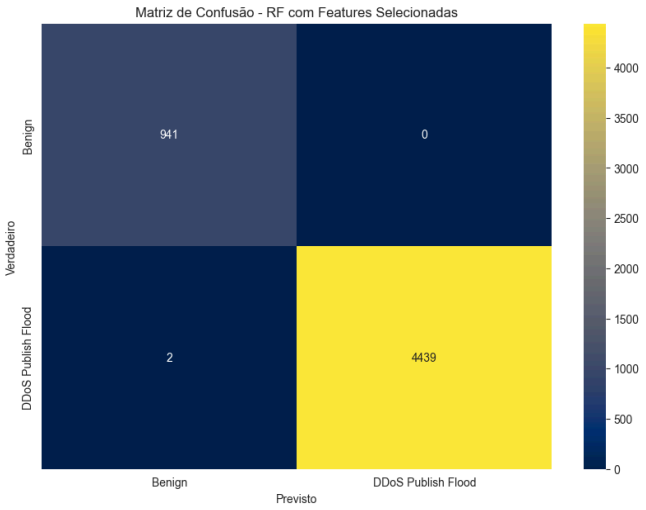

Fig. 3. Gráfico de resumo dos valores Verdadeiro e Previsto

Figure Labels: Benign: Representa a classe de tráfego de rede legítimo/normal (ou seja, a ausência de ataque). e DDoS Publish Food: DDoS Publish Food: Representa a classe de tráfego de ataque/anômalo do tipo Negação de Serviço Distribuída (DDoS), especificamente a subcategoria Publish Flood.

G. Discussão dos resultados Matriz de Confusão - Modelo LSTM Otimizado

Os resultados da Matriz de Confusão do Modelo LSTM Otimizado indicam um aprimoramento substancial em relação ao modelo de Linha de Base, validando a eficácia da otimização de hiperparâmetros e da estratégia de balanceamento aplicada.

A matriz otimizada evidencia uma melhoria crítica na segurança, pois:

- **Redução de Falsos Negativos**: O número de instâncias de DDoS Publish Food erroneamente classificadas como Benign subiu para 2, representando nenhuma redução, e sim um aumento em comparação com o modelo de Linha de Base (Figura 2), que apresentou. Esta é uma área a ser analisada, pois indica que a otimização pode ter comprometido sutilmente a capacidade do modelo de identificar a totalidade dos ataques.
- **Alta Taxa de Detecção de Ataques**: O número de Verdadeiros Positivos para a classe DDoS Publish Food é de 4.439.
- **Desempenho Geral**: O Modelo Otimizado demonstra uma alta taxa de Verdadeiros Negativos, 941 amostras Benign corretamente classificadas e mantém os Falsos Positivos em níveis extremamente baixos, confirmando uma alta precisão na detecção de anomalias no tráfego IoT.

V. Conclusão e Trabalhos futuros

Este artigo propôs e validou uma abordagem de Feature Engineering e Machine Learning para a Detecção de Anomalias em Redes IoT baseadas em MQTT, utilizando o dataset CIC-Tabular-IoT-Attack-2024. O classificador Random Forest otimizado demonstrou ser uma solução eficaz, atingindo alta acurácia e F1-Score, ao mesmo tempo que mantém uma baixa latência de inferência. A análise de interpretabilidade com SHAP fornece insights sobre o processo de decisão do modelo, confirmando a relevância das features de fluxo no contexto de segurança do MQTT.

A. Como trabalhos futuros, sugerimos:

- **Detecção de Ataques de Dia Zero**: Implementar modelos de aprendizado não supervisionado (como Autoencoders ou Isolation Forest) para detectar ataques nunca antes vistos [6].
- **Features de Taxa**: Integrar o modelo em um pipeline de processamento em tempo real (como Kafka/Spark Streaming), para validar a abordagem em um ambiente operacional [6].
- **Verdadeiros Negativos (Tráfego Legítimo Detectado)**: O modelo identificou corretamente 4439 instâncias como Benign, demonstrando uma base sólida na compreensão do perfil de tráfego normal.

Agradecimentos

Gostariam de agradecer a todos os professores e instrutores da UFABC, em especial o da disciplina de Segurança da Informação por fornecer a base teórica e o conhecimento essencial que possibilitou a execução deste estudo. A dedicação e o rigor acadêmico demonstrados foram fundamentais para a correta compreensão e aplicação dos princípios de segurança, análise de tráfego e aprendizado de máquina aqui apresentados.

References

[1] I. et al., "Realtime Feature Engineering for Anomaly Detection in IoT Based MQTT Networks," IEEE Access, vol. 12, pp. 25718-25732, 2024.

[2] M. A. Khan et al., "A Deep Learning-Based Intrusion Detection System for MQTT Enabled IoT," Sensors, vol. 21, no. 21, p. 7016, 2021.

[3] M. Hossain et al., "A Holistic Analysis of Internet of Things (IoT) Security: Principles, Practices, and New Perspectives," Future Internet, vol. 16, no. 2, p. 40, 2024.

[4] "CIC-Tabular-IoT-Attack-2024 Dataset," Canadian Institute for Cybersecurity, 2025. [Online]. Available: https://www.unb.ca/cic/datasets/tabular-iot-attack-2024.html

[5] W. Stallings, Criptografia e Segurança de Redes: Princípios e Práticas, 6ª ed. Pearson Education do Brasil, 2015. (Referência de livro, estilo [5] no texto).

[6] Alexandre Tambra Carmo, "Análise e Implementação de um IDS com Random Forest," Relatório Técnico Não Publicado,2025.[Online]. Available: https://github.com/alexandret01/UFABC_SI