Davi Esondem Menezes Brito

# Estudo de Correlação de Características Pré e Pós-Engenharia para Predição de Ataques DDoS

Proposta de trabalho **misto** da disciplina de Monografia em Sistemas de Informação II da UFMG

Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Ciência da Computação

Orientador: Prof.ª Dra. Michele Nogueira Lima
Coorientador: Dra. Ligia Francielle Borges

Belo Horizonte, Minas Gerais
2025

# Resumo

Os ataques de Negação de Serviço Distribuído (DDoS) são um tipo de ameaça cibernética que visa interromper a disponibilidade de um serviço por meio da utilização de uma rede de dispositivos infectados, chamada de *botnet*. Atualmente representam uma das principais ameaças aos sistemas cibernéticos e ciberfísicos, dada a sua capacidade de comprometer a disponibilidade de serviços essenciais. Para mitigar esse tipo de ameaça, têm sido desenvolvidas soluções baseadas em técnicas modernas, como o Aprendizado de Máquina, capazes de detectar e responder a ataques de forma automatizada e em tempo real. Essas soluções dependem da análise de grandes volumes de dados provenientes de múltiplas fontes, como tráfego de rede, logs de servidores e estatísticas de comportamento. Para maximizar a eficácia desses sistemas, os dados passam por processos de engenharia de características, que incluem a transformação, seleção e criação de variáveis relevantes para a identificação de padrões anômalos. No entanto, a presença de dados redundantes ou altamente correlacionados pode comprometer o desempenho das soluções. Este projeto tem como objetivo investigar o impacto da presença e da filtragem de atributos correlacionados em diferentes fases do processo de predição - antes e depois da engenharia de características - e comparar a efetividade dessas abordagens. Para tanto, foram implementadas, testadas e analisadas a técnicas de análise de correlação CFS-SU e de engenharia de características baseada na transformação em padrões ordinais de Bandt-Pompe em diferentes cenários. Os resultados demonstram que essas duas técnicas alinhadas trazem melhores resultados para o processo de classificação, com a engenharia de características sendo aplicada primeiro.

**Palavras-chave**: DDoS. Engenharia de Características. Correlação. Aprendizado de Máquina. Segurança Cibernética.

# Abstract

Distributed Denial of Service (DDoS) attacks are a type of cyber threat aimed at disrupting the availability of a service by using a network of infected devices, known as a botnet. Today, they represent one of the main threats to cyber and cyber-physical systems, given their ability to compromise the availability of essential services. To mitigate this type of threat, solutions based on modern techniques, such as Machine Learning, have been developed, capable of detecting and responding to attacks in an automated and real-time manner. These solutions rely on the analysis of large volumes of data from multiple sources, such as network traffic, server logs, and behavioral statistics. To maximize the effectiveness of these systems, the data undergo feature engineering processes, which include the transformation, selection, and creation of relevant variables for identifying anomalous patterns. However, the presence of redundant or highly correlated data can compromise the performance of these solutions. This project aims to investigate the impact of the presence and filtering of correlated attributes at different stages of the prediction process - before and after feature engineering - and to compare the effectiveness of these approaches. To this end, the correlation analysis technique CFS-SU and the feature engineering method based on Bandt-Pompe ordinal pattern transformation were implemented, tested, and analyzed across different scenarios. The results demonstrate that these two techniques, when used in combination, yield better classification performance, with feature engineering applied first.

**Keywords**: DDoS. Machine Learning. Feature Engineering. Correlation. Cybersecurity.

# Sumário

# 1 Introdução

Os ataques de Negação de Serviço Distribuído (*Distributed Denial of Service* - DDoS, na sigla em inglês) são, atualmente, uma das maiores ameaças aos sistemas cibernéticos e ciberfísicos, dada a sua capacidade de interromper serviços essenciais. Esses ataques funcionam por meio da sobrecarga de um servidor, serviço ou rede com um volume massivo de tráfego malicioso, proveniente de múltiplas fontes distribuídas - geralmente uma rede de dispositivos comprometidos, conhecida como *botnet*. A principal dificuldade em mitigar esse tipo de ataque reside na sua natureza distribuída, que dificulta a distinção entre tráfego legítimo e malicioso, além de permitir que o atacante disfarce sua origem real, tornando a defesa mais complexa. Para quantificar a magnitude dessa ameaça, no quarto trimestre de 2024, a empresa Cloudflare mitigou 6,9 milhões de ataques DDoS, representando um aumento de 83% em relação ao mesmo período do ano anterior [Yoachimik e Pacheco 2025]. Entre esses, destaca-se um ataque recorde que atingiu 5,6 terabits por segundo, considerado o maior já registrado.

Para combater esse tipo de ameaça, soluções que aplicam técnicas modernas, como Aprendizado de Máquina e Aprendizado Profundo, vêm sendo desenvolvidas para detectar e mitigar esses ataques, com a finalidade de manter a disponibilidade de serviços e evitar desastres em organizações de diferentes segmentos [Alhijawi et al. 2022]. Anteriormente, as soluções tradicionais baseavam-se em assinaturas e regras estáticas, como firewalls e sistemas de detecção de intrusões baseados em padrões conhecidos. Essas abordagens, embora eficazes contra ataques previamente identificados, apresentam limitações significativas ao lidar com ataques DDoS sofisticados e em constante evolução, que frequentemente utilizam tráfego disfarçado de legítimo ou exploram vulnerabilidades desconhecidas. As técnicas modernas de aprendizado de máquina buscam superar essas limitações ao analisar padrões de tráfego, identificar anomalias e adaptar-se a novas formas de ataques, proporcionando uma defesa mais robusta e proativa contra ameaças emergentes, a exemplo de técnicas de predição.

Essas soluções coletam, processam e analisam diversos tipos de dados para identificar possíveis preparações de ataque em andamento. Os dados provêm de fontes variadas, como: dados de tráfego de rede, capturados diretamente de equipamentos físicos na rede; dados estatísticos e comportamentais, extraídos de sistemas de monitoramento; e logs de aplicações e servidores, gerados pelos serviços e máquinas que os hospedam. A fim de extrair informações relevantes e melhorar o desempenho das soluções, os dados passam pelo processo de engenharia de características. Essa etapa busca selecionar, transformar e criar variáveis a partir dos dados brutos, incluindo tarefas como a remoção de ruídos, a normalização de valores, a extração de padrões temporais ou estatísticos e a geração de novas características que possam destacar comportamentos anômalos característicos

de ataques DDoS. A qualidade e a relevância das características construídas impactam diretamente na capacidade dos algoritmos de aprendizado em identificar padrões e tomar decisões precisas, sendo, portanto, uma etapa essencial para o sucesso na predição de ameaças cibernéticas.

## 1.1 Problema

A predição de ataques DDoS a partir de dados de tráfego de rede enfrenta um desafio importante: a presença de características redundantes ou irrelevantes nos conjuntos de dados. Isso se traduz em atributos com valores extremamente correlacionados entre si ou pouco correlacionados com a variável-alvo. Essas características não apenas desperdiçam recursos computacionais e tempo de processamento, como também podem comprometer a eficácia dos algoritmos de predição, aumentando a taxa de falsos positivos ou reduzindo a precisão dos modelos. Em muitos casos, atributos altamente correlacionados entre si ou com pouca variabilidade não contribuem de forma significativa para a diferenciação entre tráfego legítimo e malicioso.

Nesse contexto, torna-se essencial aplicar técnicas de seleção e análise de correlação de atributos, a fim de identificar e manter apenas as variáveis mais relevantes para a tarefa de predição. Essa análise pode ser realizada em diferentes momentos do processo: antes da engenharia de características (pré-processamento), com o objetivo de filtrar dados brutos; ou após a engenharia, para avaliar o impacto das transformações realizadas nas variáveis e sua relação com a variável-alvo. A escolha do momento adequado para aplicar essa análise influencia diretamente na qualidade do conjunto de dados final e, consequentemente, no desempenho do modelo de predição. No entanto, ainda não há estudos que indiquem qual seja o momento mais apropriado para sua aplicação.

Estudos demonstram que a presença de características redundantes pode degradar significativamente o desempenho de modelos de aprendizado de máquina, aumentando a complexidade computacional e reduzindo a acurácia na detecção de anomalias. Por exemplo, Yin *et al.* (2023) propuseram um método híbrido de seleção de características para detecção de intrusões em redes, que reduziu o número de características de 42 para 23 e melhorou a acurácia da classificação de 82,25% para 84,24% ao eliminar atributos redundantes.

## 1.2 Objetivos

### 1.2.1 Objetivo Geral

Este projeto, portanto, propõe analisar o impacto de dados altamente correlacionados e aferir os resultados da solução nos dois momentos possíveis: anteriormente à etapa de engenharia de características, onde há somente os atributos brutos, ou após o processamento

realizado na etapa de engenharia de características, com os dados tratados. O principal objetivo é comparar as métricas e trazer as vantagens e desvantagens de cada aplicação para determinado objetivo e contexto de predição de ataques DDoS.

### 1.2.2 Objetivos Específicos

Os objetivos específicos deste trabalho incluem: i) aplicar técnicas de engenharia de características para transformar e preparar o conjunto de dados selecionado, criando novas variáveis; ii) estabelecer e conduzir análises de correlação nos atributos e características, com o intuito de avaliar a relevância e redundância dos atributos em cada cenário; iii) treinar e avaliar modelos de predição de ataques DDoS com base nos diferentes conjuntos de dados resultantes dessas análises, utilizando métricas apropriadas de desempenho e eficiência; e iv) comparar os resultados obtidos a fim de identificar as vantagens e limitações de cada abordagem no contexto da predição de ataques DDoS, fornecendo diretrizes para aplicações em ambientes diversos.

## 1.3 Organização do trabalho

Este trabalho está organizado da seguinte forma. A Seção 2 apresenta os fundamentos e os trabalhos relacionados à análise de correlação e predição de ataques DDoS. A Seção 3 detalha a proposta estabelecida e sua metodologia. A Seção 4 apresenta os resultados e a discussão. E, por fim, a Seção 5 conclue o trabalho.

# 2 Conceitos e Trabalhos Relacionados

## 2.1 Ataques DDoS

Um ataque de Negação de Serviço Distribuído ou *Distributed Denial of Service* (DDoS) consiste na utilização de múltiplos dispositivos comprometidos, que estejam conectados à Internet, para atacar um alvo e causar a indisponibilidade de um serviço para os usuários do alvo afetado. O alvo pode ser um servidor ou outro recurso de rede. Para realização do ataque, uma enxurrada de mensagens, solicitações de conexão ou pacotes malformados, são enviados para o sistema de destino, o forçando a aumentar o consumo de seus recursos computacionais de maneira a reduzir o seu tempo de resposta ou até mesmo travar e desligar, negando assim um serviço a usuários ou sistemas legítimos.

O DDoS é uma das ameaças mais prejudiciais na Internet de hoje. Sua capacidade de interromper sistemas essenciais é visada e é comumente ofertada como um serviço por criminosos no mercado negro, para pessoas ou organizações que desejam sabotar outra instituições, sejam elas públicas ou privadas. De acordo com [Brodsky 2021], as motivações dos atacantes e clientes desse tipo de serviço são diversas e partem desde a obtenção de benefícios financeiros ou econômicos, vingança, crenças ideológicas, guerra cibernética e até mesmo pela simples satisfação pessoal.

### 2.1.1 Tipos de Ataques DDoS

Os ataques de Negação de Serviço Distribuído (DDoS) podem ser classificados em três categorias principais, conforme a camada do modelo OSI que visam: ataques volumétricos, ataques de protocolo e ataques na camada de aplicação [Mirkovic e Reiher 2004]. Cada tipo explora diferentes vulnerabilidades para sobrecarregar recursos e interromper serviços legítimos. Ataques Volumétricos têm como objetivo saturar a largura de banda do alvo, enviando um volume massivo de tráfego para impedir que usuários legítimos acessem os serviços. Um exemplo comum é o UDP Flood, no qual o atacante envia pacotes UDP a portas aleatórias, fazendo com que o sistema alvo responda com mensagens ICMP de "Destino Inalcançável", sobrecarregando a rede. Outro exemplo é o ataque de amplificação DNS, que explora servidores DNS abertos para enviar pequenas solicitações que resultam em grandes respostas direcionadas ao alvo, amplificando o tráfego malicioso. Esses ataques são frequentemente medidos em gigabits por segundo (Gbps) e podem ser mitigados com filtragem de tráfego e soluções de proteção DDoS baseadas em nuvem.

Ataques de Protocolo exploram vulnerabilidades em protocolos da camada de rede para consumir recursos do servidor ou de dispositivos intermediários, como firewalls. O SYN Flood é um exemplo clássico, onde o atacante envia múltiplas solicitações de conexão

TCP sem concluir o processo de handshake, esgotando a capacidade do servidor de manter novas conexões. O Ping of Death envolve o envio de pacotes ICMP fragmentados que, ao serem remontados, excedem o tamanho máximo permitido, podendo travar o sistema alvo. Já o Smurf Attack utiliza pacotes ICMP com endereços de origem falsificados enviados a endereços de broadcast, fazendo com que múltiplos dispositivos respondam simultaneamente ao alvo, inundando-o com tráfego.

Ataques na Camada de Aplicação são mais sofisticados e visam exaurir recursos específicos de aplicações, como servidores web, simulando tráfego legítimo. O HTTP Flood consiste no envio massivo de requisições HTTP GET ou POST, sobrecarregando o servidor com processamento de solicitações aparentemente válidas. O Slowloris mantém múltiplas conexões abertas enviando cabeçalhos HTTP incompletos, impedindo que o servidor feche essas conexões e abrindo espaço para novas.

### 2.1.2 Predição

A crescente sofisticação e frequência dos ataques de DDoS têm exigido o desenvolvimento de abordagens proativas que vão além da simples detecção. Nesse cenário, técnicas de predição surgem como uma alternativa promissora, permitindo a identificação de sinais indicativos de ataques antes que eles comprometam a disponibilidade dos serviços [de Neira, Kantarci e Nogueira 2023]. A pesquisa em predição de ataques DDoS visa analisar padrões de tráfego, comportamentos históricos e variações sutis nos dados de rede para antecipar ações maliciosas com o máximo de antecedência possível. Diferente das soluções tradicionais de detecção e mitigação, que agem durante e após a detecção do ataque, os sistemas preditivos têm como objetivo atuar de forma preventiva, desencadeando medidas de defesa automaticamente ou alertando os administradores com base em indícios confiáveis. Essa abordagem é especialmente relevante em ambientes críticos, onde segundos de indisponibilidade podem gerar perdas significativas.

Contudo, prever ataques DDoS é uma tarefa desafiadora. A variabilidade dos padrões de ataque, a presença de tráfego legítimo que pode se assemelhar ao malicioso, e a necessidade de detecção em tempo real exigem algoritmos robustos e eficientes. A escolha de atributos relevantes, o tratamento adequado dos dados e a capacidade dos modelos de generalizar para novas situações são fatores determinantes para o sucesso de qualquer solução preditiva. Com base nesse mesmo raciocínio, [Kebede et al. 2022] estabeleceram um sistema de predição inteligente de ataques DDoS em redes, utilizando técnicas de mineração de dados como JRip, J48 e k-NN para detectar e prevenir ataques DDoS. Esses algoritmos foram implementados e avaliados individualmente para validar seu desempenho na detecção de ataques. Em [Silva, Neira e Nogueira 2022], os autores criaram uma solução baseada em rede neural não supervisionada com arquitetura *Long Short Term Memory* (LSTM) *Autoencoder* para reconstruir os tráfegos de rede e avaliar o custo obtido, transformando-os em sinais precoces de alerta de ataque.

## 2.2 Engenharia de Características e Análise de Correlação

Engenharia de características é o processo de transformar dados brutos em atributos que melhor representem os padrões subjacentes de um problema, com o objetivo de maximizar o desempenho de algoritmos de aprendizado de máquina. Essa etapa é especialmente crucial em sistemas de predição de ataques DDoS, onde os dados coletados de tráfego de rede, logs de aplicação ou sensores podem conter milhares de variáveis - muitas das quais irrelevantes, redundantes ou ruidosas.

A principal motivação por trás da engenharia de características é melhorar a capacidade dos modelos de identificar padrões úteis para uma tarefa específica (como classificar tráfego como legítimo ou malicioso). Isso pode ser feito por meio de diversas estratégias, como a geração de atributos derivados a partir dos existentes, como taxas, agregações ou estatísticas temporais, a transformação de dados, que envolve aplicação de técnicas como normalização, discretização ou codificação categórica, e redução de dimensionalidade (técnicas como Análise de Componentes Principais, Autoencoders ou seleção de atributos). A engenharia de caracterísiticas pode reduzir a complexidade computacional, aumentar as métricas da predição, evitar overfitting, e permitir uma operação mais eficiente em tempo real.

Vários trabalhos recentes têm explorado diferentes formas de engenharia de características no contexto de segurança de rede. Em [Borges et al. 2024], os autores aplicam a simbolização de Bandt-Pompe para transformar séries temporais multivariadas de tráfego e logs em padrões ordinais, que são então representados por grafos de transição, facilitando a distinção entre comportamentos legítimos e maliciosos. Já em [Albano et al. 2024], é proposto um método que integra transformações ordinais com descritores da Teoria da Informação e múltiplos algoritmos de seleção de atributos, como MultiSURF e LASSO, focando na predição antecipada de ataques e baixo custo computacional.

### 2.2.1 Análise de Correlação

Complementar à engenharia de características está a análise de correlação, cuja função é identificar relacionamentos estatísticos entre os atributos. A presença de correlações elevadas pode indicar redundância - ou seja, quando dois ou mais atributos carregam informações similares [Gregorutti, Michel e Saint-Pierre 2016]. Manter todos esses atributos no conjunto de dados pode tornar o modelo mais complexo do que o necessário, aumentar o tempo de treinamento e prejudicar a capacidade de generalização do modelo (levando ao *overfitting*). Por essa razão, técnicas de seleção de atributos baseadas em correlação buscam identificar subconjuntos de atributos que tenham alta relevância em relação à variável-alvo (por exemplo, um rótulo que indica ataque ou tráfego normal) e, ao mesmo tempo, baixa redundância entre si.

Uma abordagem comum é a aplicação de filtros estatísticos como o *Correlation-*

*Based Feature Selection* (CFS), que considera tanto a relevância quanto a redundância ao selecionar subconjuntos de atributos. No entanto, o CFS é baseado no coeficiente de correlação de Pearson, o que o torna limitado a correlações lineares. Para lidar com relações não-lineares, técnicas complementares como a *Symmetrical Uncertainty* (SU) são frequentemente utilizadas. Essa métrica é derivada da informação mútua e mede a dependência entre variáveis, sendo adequada para cenários em que os dados apresentam complexidade estrutural mais elevada.

A combinação dessas estratégias pode ser aplicada em uma ou mais etapas do *pipeline* de desenvolvimento do modelo, seja antes da engenharia de características (trabalhando diretamente com atributos brutos) ou depois (considerando os atributos já transformados). A decisão sobre o momento mais adequado para aplicar essas técnicas depende do objetivo do sistema, da complexidade dos dados e das restrições computacionais envolvidas.

Diversos trabalhos da literatura têm empregado essas abordagens. Em [Dasari K.B. 2022], os autores buscaram aumentar a eficiência dos modelos de detecção de ataques DDoS eliminando atributos redundantes por meio da análise de correlação e da técnica de limite de variação. Os resultados demonstraram que atributos altamente correlacionados aumentam o tempo de treinamento e contribuem para o risco de overfitting, prejudicando a generalização do modelo.

No estudo de [Michalak e Kwasnicka 2006], foi proposta uma abordagem híbrida de seleção de atributos baseada em medidas de dependência estatística. Inicialmente, foi utilizado o CFS para identificar subconjuntos com alta relevância e baixa redundância, seguido pela aplicação do SU para refinar o conjunto e eliminar atributos irrelevantes ou redundantes de maneira não-linear. O processo foi realizado com estratégias de forward selection e backward elimination, resultando na redução de 42 para apenas 4 atributos no conjunto de dados NSL-KDD, sem comprometer o desempenho do sistema.

## 2.3 Aprendizado de Máquina

O aprendizado de máquina (*Machine Learning* - ML, em inglês) é uma subárea da inteligência artificial que permite que sistemas aprendam automaticamente a partir de dados, sem serem explicitamente programados para executar tarefas específicas. Por meio da identificação de padrões e regularidades nos dados, algoritmos de ML são capazes de construir modelos que fazem previsões ou tomam decisões com base em novas entradas.

Os métodos de aprendizado de máquina podem ser classificados, de maneira geral, em três grandes categorias: aprendizado supervisionado, não supervisionado e por reforço. No aprendizado supervisionado, o modelo é treinado em um conjunto de dados rotulado, aprendendo a mapear entradas para saídas conhecidas, como em problemas de classificação ou regressão. Já no aprendizado não supervisionado, o modelo busca descobrir padrões ocultos ou agrupamentos nos dados sem o auxílio de rótulos, sendo amplamente utilizado em

tarefas de segmentação ou detecção de anomalias. Além disso, há o chamado aprendizado semi-supervisionado, que mescla as duas técnicas anteriores com o objetivo de classificar dados não-rotulados com o uso de dados rotulados. O aprendizado por reforço, por sua vez, envolve a tomada de decisões sequenciais, onde um agente aprende a maximizar uma recompensa cumulativa ao interagir com um ambiente dinâmico [N e Gupta 2020].

### 2.3.1 Aprendizado Profundo

O desenvolvimento do aprendizado profundo (*Deep Learning* — DL), uma subárea do aprendizado de máquina inspirada em redes neurais artificiais, revolucionou a capacidade dos modelos em capturar representações complexas e hierárquicas dos dados. O aprendizado profundo funciona organizando camadas sucessivas de neurônios artificiais, onde cada camada extrai características de maior nível de abstração em relação à anterior, permitindo a construção automática de representações ricas a partir de dados brutos [LeCun, Bengio e Hinton 2015]. Diferentemente dos métodos tradicionais de aprendizado de máquina, que dependem fortemente da engenharia manual de características, o aprendizado profundo é capaz de aprender essas representações de forma automática, reduzindo a necessidade de intervenção humana e alcançando desempenhos superiores em tarefas complexas. Redes neurais profundas, como as arquiteturas convolucionais (CNNs) e recorrentes (RNNs), permitiram avanços significativos em áreas como visão computacional, processamento de linguagem natural e segurança cibernética.

### 2.3.2 Arquiteturas para Predição de Ataques DDoS

No contexto da predição de ataques DDoS, técnicas de aprendizado de máquina e aprendizado profundo desempenham um papel central. Algoritmos de classificação supervisionada como *k-Nearest Neighbors* (k-NN), *Random Forests*, *Decision Trees* e *Support Vector Machines* (SVMs) têm sido amplamente empregados para identificar padrões de tráfego anômalo associados a ataques. Esses modelos são treinados em conjuntos de dados históricos de tráfego de rede, nos quais as instâncias são rotuladas como benignas ou maliciosas, e, posteriormente, usados para detectar comportamentos suspeitos em tempo real. O k-NN, por exemplo, classifica novas amostras com base nas classes predominantes entre seus vizinhos mais próximos no espaço de características, sendo especialmente útil em problemas de detecção onde as classes possuem fronteiras complexas. *Decision Trees* funcionam criando uma sequência hierárquica de decisões com base em características dos dados, permitindo interpretações claras sobre o processo de classificação. *Random Forests* combinam diversas árvores de decisão treinadas em subconjuntos dos dados, melhorando a precisão e reduzindo o risco de sobreajuste. Já as SVMs buscam encontrar o hiperplano ótimo que separa as diferentes classes no espaço de características, sendo eficazes em conjuntos de dados de alta dimensão [Sen, Hajra e Ghosh 2020].

Além dos métodos clássicos, abordagens baseadas em DL vêm ganhando destaque na predição de DDoS. Modelos como *Autoencoders*, Redes Neurais Convolucionais (CNNs) e *Long Short-Term Memory* (LSTM) têm demonstrado eficiência na modelagem de padrões temporais do tráfego de rede.

# 3 Metodologia

A metodologia deste trabalho será conduzida por meio da avaliação e uma análise detalhada da correlação entre diferentes dados aplicados em um algoritmo de predição de ataques DDoS. Foram utilizados três experimentos para comparação das métricas entre os dados e seus resultados. Cada experimento trata de um conjunto de dados e tipo de ataque diferente. A abordagem é estruturada em etapas distintas, que abrangem a seleção dos dados, a análise de correlação, a engenharia de características, a predição dos ataques e, por fim, a análise dos resultados. A seguir, cada uma das etapas é detalhada.

## 3.1 Conjunto de Dados

Para a realização dos experimentos deste trabalho, foram utilizados três subconjuntos de dados distintos provenientes de dois conjuntos amplamente reconhecidos na literatura: o CTU-13 e o CIC-DDoS2019. Esses conjuntos contêm dados sobre tráfego de rede real e simulado, além de ataques DDoS, e foram escolhidos por sua relevância em pesquisas de predição desse ataques. Os dados abrangem diferentes tipos de tráfego, incluindo tráfego legítimo, malicioso e de preparação ao ataque. Esse último tipo é crucial para as soluções de predição, visto que será a base usada para treinar e avaliar o modelo, uma vez que antecede o ataque propriamente dito. Neste caso, foram extraídos atributos brutos da camada de rede e transporte.

### 3.1.1 CTU-13

O conjunto de dados CTU-13 [Garcia et al. 2014] foi capturado na rede de uma universidade e é dividido em treze cenários de diferentes amostras de *botnets* e tipos de ataque. O *dataset* captura, simultaneamente, tráfego legítimo e tráfego de *botnet*. Para os experimentos conduzidos, foram utilizadas as capturas de número 51 e 52 por conterem tráfego da *botnet* anterior ao ataque, o que pode ser considerado como uma preparação ao ataque.

A captura 51 possui 8803 segundos, 10 bots, 41 GB, 46.997.342 pacotes e ataques do tipo inundação com os protocolos Internet Control Message Protocol (ICMP) e UDP. Nesse cenário os bots da rede foram infectados no segundo 2643 e lançaram os ataques a partir do segundo 5632 da captura. Esse tempo da infecção até o início da ataque será considerado como a preparação dele.

Já a captura 52 contém um tamanho consideravelmente menor. O cenário possui 972 segundos de tráfego na rede, com 6 milhões de pacotes, um ataque de DDoS do

tipo inundação ICMP e três bots. A preparação se inicia no segundo 527 e o ataque, efetivamente, no segundo 797.

### 3.1.2 CIC-DDoS2019

Também foi utilizado o *dataset* CIC-DDoS2019 [Sharafaldin et al. 2019], um ambiente simulado no qual a rede vítima está conectada ao atacante pela Internet. Essa captura possui 19 ataques DDoS lançados pelos pesquisadores em dois dias. O conjunto de dados possui 61.407.883 de pacotes e 27 GB de dados referente aos ataques e tráfego comum. A experimentação se concentrou na predição do primeiro ataque DDoS realizado. O ataque começa no segundo 1484 da captura e possui duração de 540 segundos.

## 3.2 Engenharia de Características

A etapa de engenharia de características, como é conhecido o processo de extração de informações a partir dos atributos brutos coletados da rede, é vantajosa às soluções de predição de ataques DDoS. Por meio dela é possível criar novas representações dos dados existentes, de modo a tornar mais explícito ao modelo de predição o comportamento da rede e o efeito das anomalias que representam as preparações e, também, o ataque DDoS. Desse modo, a definição de um processo de engenharia de características adequado ao contexto dos dados pode trazer significativas melhorias aos resultados.

### 3.2.1 Transformação em Padrões Ordinais

Para este trabalho, foi adotada a técnica de transformação dos dados da rede em padrões ordinais. Ela utiliza a Transformação Ordinal de Bandt-Pompe [Bandt e Pompe 2002] e transforma as séries temporais obtidas a partir do tráfego de rede em representações simbólicas que descrevem a ordem relativa entre valores consecutivos da série. Diferentemente de métodos sensíveis a *outliers*, a transformação ordinal é robusta a ruídos extremos, permitindo capturar variações estruturais mais sutis no comportamento do tráfego. A hipótese que sustenta a utilização dessa técnica é que pequenas flutuações provocadas por atividades maliciosas, mesmo antes de um ataque ser iniciado de fato, podem alterar o padrão dinâmico da série, tornando essas alterações detectáveis por meio da análise dos padrões de ordem.

A Figura 1 mostra o funcionamento geral dessa etapa. A técnica é composta por duas etapas principais: a primeira consiste em segmentar a série temporal em subconjuntos de tamanho fixo. Cada subconjunto representa uma sequência de observações consecutivas do tráfego. Em seguida, é calculado o padrão ordinal de cada subconjunto, que corresponde à permutação dos índices dos valores ordenados de forma crescente.

A partir da sequência completa de padrões ordinais extraída da série, duas representações são construídas para possibilitar a geração de novas características: a distribuição de probabilidade dos padrões e o grafo de transições entre padrões. A distribuição de probabilidade é obtida por meio de um histograma de frequências, contabilizando a ocorrência de cada padrão ordinal na série. Já o grafo de transições é um grafo orientado em que cada vértice representa um padrão ordinal, e as arestas indicam a transição de um padrão para outro ao longo do tempo. O peso de cada aresta corresponde à frequência com que tal transição ocorre.

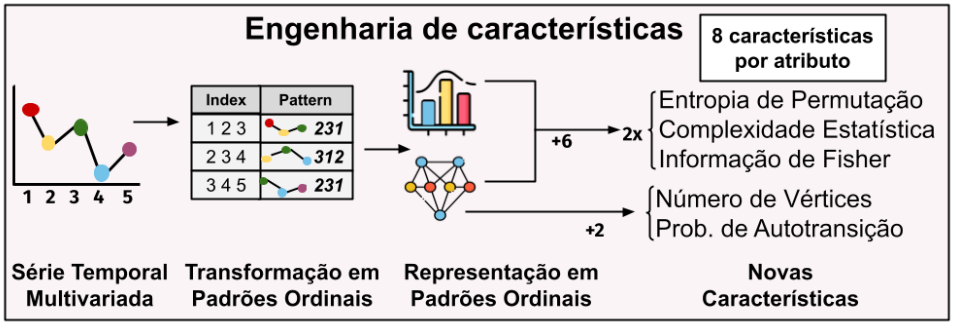


Figura 1 – Processo de engenharia de características

A partir dessas duas representações, são extraídas oito características representativas com base em quantificadores da Teoria da Informação. São elas: (i) a entropia de permutação normalizada, que mede a diversidade e a imprevisibilidade dos padrões ordinais, sendo maior em séries caóticas e menor em séries determinísticas; (ii) a complexidade estatística, que combina o grau de desordem da série com o afastamento da distribuição em relação ao caso uniforme, permitindo identificar estruturas internas não triviais; (iii) a informação de Fisher, que mensura a sensibilidade a variações locais na distribuição de padrões; (iv) o número de vértices no grafo de transições, representando a diversidade de padrões observados; (v) a entropia de permutação da distribuição de pesos das arestas do grafo, refletindo a aleatoriedade das transições entre padrões; (vi) a complexidade estatística da distribuição de pesos, revelando estruturas recorrentes nas transições; (vii) a informação de Fisher aplicada à distribuição de pesos, que avalia descontinuidades nas transições entre padrões; e (viii) a probabilidade de autotransição, que indica a frequência com que um padrão ordinal é seguido por ele mesmo, refletindo a autocorrelação da série.

Essas oito métricas geradas a partir da Transformação Ordinal de Bandt-Pompe constituem um novo conjunto de atributos derivados dos dados originais de tráfego de rede. Ao representar de forma compacta e robusta a dinâmica temporal dos dados, esses atributos têm se mostrado eficazes na caracterização de comportamentos maliciosos e na predição de ataques DDoS [Borges et al. 2024].

## 3.3 Análise de Correlação

Para a etapa de análise de correlação e filtragem de características, adotou-se a metodologia proposta por [Shahbaz et al. 2016], a qual combina medidas de dependência linear e não-linear com o objetivo de reduzir a dimensionalidade dos dados sem comprometer, e por vezes até melhorando, o desempenho dos classificadores no contexto de sistemas de detecção de intrusão. A técnica baseia-se em uma abordagem em duas etapas que utiliza as métricas *Correlation-Based Feature Selection* (CFS) e *Symmetrical Uncertainty* (SU). Inicialmente, aplica-se o CFS para selecionar um subconjunto de atributos que apresente alta relevância com relação à variável-alvo e, simultaneamente, baixa redundância entre os próprios atributos. Ao contrário de métodos que avaliam atributos de forma isolada, o CFS considera o valor informativo de conjuntos de atributos, utilizando a fórmula:

$$\text{CFS} = \frac{m \cdot \overline{rcf}}{\sqrt{m + m(m-1)\overline{r_{ff}}}} \tag{3.1}$$

em que $m$ representa o número de atributos no subconjunto, $\overline{r_{cf}}$ é a média da correlação entre os atributos e a variável-alvo, e $\overline{r_{ff}}$ corresponde à média das correlações entre os atributos. Essas correlações são calculadas por meio do coeficiente de Pearson [Pearson 1895], dado por:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{3.2}$$

Embora eficiente, o CFS considera apenas correlações lineares, o que pode ser limitante em contextos com relações não-lineares entre os dados [Michalak e Kwasnicka 2006]. Para mitigar essa limitação, é realizada uma segunda etapa com o uso da métrica *Symmetrical Uncertainty* (SU), que complementa a análise ao considerar dependências não-lineares. A SU é uma métrica baseada em informação mútua e entropia, expressa por:

$$SU(X, Y) = \frac{2 \cdot I(X; Y)}{H(X) + H(Y)} \tag{3.3}$$

onde I(X; Y) representa a informação mútua entre os atributos X e Y, e H(X) é a entropia de X. Como métrica normalizada, a SU varia entre 0 e 1, sendo que valores próximos de 1 indicam maior redundância entre os atributos.

O processo de seleção se inicia com a remoção de atributos que não possuem variância, seguido da geração de subconjuntos de atributos por meio de busca sequencial progressiva (*forward selection*). Para cada subconjunto, calcula-se a métrica CFS, e aquele com maior valor é selecionado. Em seguida, é computada a SU para cada par de atributos do subconjunto escolhido, organizando os resultados em ordem decrescente de SU. A partir disso, aplica-se uma estratégia de eliminação regressiva (*backward elimination*), removendo iterativamente os atributos menos colaborativos— ou seja, com menor SU — desde que sua remoção não comprometa significativamente a acurácia do modelo. A seleção final

é determinada com base na estabilidade da acurácia observada no próprio conjunto de treinamento, evitando a remoção de atributos relevantes.

## 3.4 Predição dos Ataques DDoS

Após a seleção dos atributos, extração de características e filtragem por meio da análise de correlação, os dados resultantes são utilizados para treinamento e teste de um modelo de predição de ataques DDoS. A partir do teste, é possível avaliar as métricas obtidas para o conjunto de dados, os processos utilizados e sua ordem, e o modelo utilizado. Para este trabalho, o foco não foi direcionado à definição de um algoritmo complexo de predição e obtenção de resultados sem precedentes, mas sim à variação deles conforme a transição dos dados.

Para esse propósito, foi empregado o algoritmo de aprendizado de máquina One-Class Support Vector Machine (One-Class SVM) com o objetivo de identificar anomalias no tráfego de rede, tratadas neste contexto como potenciais sinais de ataques DDoS. Esse algoritmo atua de forma não supervisionada, ou seja, requer apenas dados rotulados como normais para o treinamento do modelo. A partir desses dados, o One-Class SVM aprende a delimitar uma região de normalidade no espaço de características, construindo um hiperplano que separa o conjunto normal do espaço restante, onde eventuais anomalias podem ocorrer [Lima et al. 2023].

O funcionamento do One-Class SVM baseia-se na construção de um limite de decisão que maximiza a margem entre os dados normais e a origem, criando uma fronteira capaz de distinguir entre padrões esperados e comportamentos atípicos. Após o treinamento, o modelo é capaz de classificar novas observações como normais ou anômalas com base em sua posição relativa ao hiperplano aprendido. Dessa forma, qualquer instância que ultrapasse essa fronteira é interpretada como um possível *outlier*, que seria, neste caso, uma alteração suspeita no padrão de tráfego que pode indicar o início de um ataque.

A eficácia do modelo depende de dois parâmetros principais. O primeiro é o parâmetro $\nu$, que atua como um controle sobre a proporção de exemplos considerados anômalos durante o treinamento. Esse valor varia entre 0 e 1, sendo que valores mais altos aumentam a sensibilidade do modelo, permitindo que ele identifique uma maior quantidade de *outliers*, ao custo de um possível aumento nos falsos positivos. O segundo parâmetro é o kernel, que define o tipo de transformação utilizada para separar os dados no espaço de características. Os kernels mais comuns são o linear, o polinomial e o sigmoide, sendo a escolha do kernel ajustada de acordo com a natureza e a complexidade dos dados analisados.

## 3.5 Comparação de Resultados

Por fim, a última etapa consiste em uma análise dos resultados obtidos nos experimentos feitos. Essa análise envolve a avaliação de eficácia do conjunto de modelo e dados utilizados por meio de métricas definidas. Os impactos da presença de dados correlacionados também foram analisados, assim como da variação da ordem em que as etapas de engenharia de características e análise de correlação são aplicadas. Os resultados foram interpretados de forma a oferecer esclarecimentos sobre a relevância da análise de correlação em diferentes fases do processamento de dados, explicitando suas vantagens e desvantagens. Adicionalmente, propõe-se um fluxo otimizado para o tratamento de dados em sistemas de predição de ataques DDoS, integrando a seleção de variáveis por correlação, a engenharia de características e os algoritmos de predição de forma eficiente.

# 4 Avaliação e Resultados

A avaliação dos processos está dividida em três experimentos, nos quais há uma diferença dos conjuntos de dados utilizados e dos tipos de ataque DDoS presentes. Os experimentos 1 e 2 utilizarão o *dataset* CTU-13, com as capturas 51 e 52, respectivamente. Já o Experimento 3 considera o *dataset* CIC-DDoS2019. A presença de diferentes conjuntos de dados auxilia a reforçar as análises e os padrões percebidos nos resultados.

A princípio, foram comparados os resultados do teste do modelo de predição utilizando quatro subconjuntos de dados distintos. O primeiro é dos atributos brutos, extraídos diretamente do arquivo de captura de rede disponibilizados pelo *dataset*. O segundo consiste no primeiro subconjunto filtrado por meio da etapa de análise de correlação, restando somente atributos com baixa correlação entre si e alta correlação com o rótulo. Já o terceiro e quarto subconjunto são compostos das características obtidas a partir dos atributos brutos, possuindo como diferença, também, a filtragem por correlação no subconjunto quatro. Por fim, também foi feita a comparação dos resultados entre diferentes ordens de aplicação das etapas de engenharia de características e de análise de correlação. Na primeira ordem, é executada a etapa de análise de correlação antes, e na segunda ordem, a etapa de engenharia de características. Todos os resultados estão disponíveis online[1]

O modelo One-Class SVM foi configurado de modo a extrair resultados nos conjuntos de testes de forma padronizada. Para isso, foi utilizada a biblioteca Scikit-Learn[2] e foram configurados alguns hiperparâmetros do modelo, como o tipo do kernel, seu coeficiente (denotado por $\gamma$) e a sensibilidade $\nu$. O tipo do kernel selecionado foi o polinomial, com o coeficiente fixado em "auto", que consiste na razão entre 1 e o número de variáveis, e, por fim, a sensibilidade $\nu$ com o valor de 0.3, representando um equilíbrio na classificação do tráfego da rede.

As variações no desempenho dos algoritmos utilizados para a predição de ataques DDoS foram mensuradas utilizando quatro métricas: Acurácia, Precisão, *Recall* e F1-Score. Essas métricas foram calculadas a partir de classificações binárias feitas pelo algoritmo, as quais são dadas como corretas se houver ao menos um pacote proveniente da *botnet* no instante de tempo analisado ou não, no caso oposto. Para isso, foram definidas as métricas de avaliação:

- Verdadeiro Positivo (VP): número de tráfego malicioso corretamente classificado como malicioso.

[1] `https://github.com/daviembrito/MSI`
[2] `https://scikit-learn.org/stable/modules/generated/sklearn.svm.OneClassSVM.html`

- Verdadeiro Negativo (VN): número de tráfego normal corretamente classificado como normal.
- Falso Positivo (FP): número de tráfego normal erroneamente classificado como malicioso.
- Falso Negativo (FN): número de tráfego malicioso erroneamente classificado como normal.

A Acurácia é a proporção de acertos (VP e VN) do algoritmo em relação ao número total. É dada pela seguinte fórmula:

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \tag{4.1}$$

A Precisão mede a proporção de classificações positivas corretas em relação ao total de classificações positivas.

$$Precisão = \frac{VP}{VP + FP} \tag{4.2}$$

O *Recall* mede a quantidade de classificações positivas e corretas do modelo, em relação ao total de positivos existentes.

$$Recall = \frac{VP}{VP + FN} \tag{4.3}$$

O F1-Score é a média harmônica da Precisão e *Recall*, sendo utilizado para balancear a importância dessas duas métricas.

$$F1\text{-}Score = 2 \times \frac{Precisão \times Recall}{Precisão + Recall} \tag{4.4}$$

Além disso, também foi considerado o tempo que a solução antecipou até a ocorrência de ataque, que é denominada neste trabalho como "tempo de predição". Quanto maior esse tempo, mais capaz é a solução de evitar danos com a presença de um ataque na rede. Por fim, a quantidade de variáveis utilizadas para teste também foi avaliada, a fim de mensurar a dimensionalidade necessária dos dados para a obtenção dos respectivos resultados.

## 4.1 Experimento 1

Para o Experimento 1, foi utilizada a captura 51 do *dataset* CTU-13. Nesse cenário, o ataque começa após 1:33:34 do início da captura do tráfego. A Tabela 1 contém os atributos brutos extraídos a partir do arquivo de captura de rede. Ao total, foram utilizados 21 atributos para esse experimento. Na etapa de engenharia de características foram geradas oito novas variáveis a partir de cada atributo. Contudo, com a remoção de algumas variáveis com variância nula, o conjunto de características utilizado teve 164

características. Para o treinamento e teste da solução, foram utilizados 5614 segundos da captura, correspondentes ao tráfego anterior ao ataque. Dessa porção, os conjuntos de treinamento e teste foram segmentados de forma aleatória. O conjunto de treinamento possui 40% do tamanho total, enquanto o conjunto de teste possui 60%.

| **Atributos** | |
|---|---|
| Total de pacotes | Quantidade de flags “reserved” |
| Total de pacotes TCP | Quantidade de flags “cwr” |
| Total de pacotes UDP | Quantidade de flags “urgent” |
| Total de pacotes de outros protocolos | Quantidade de flags “acknowledgment” |
| Endereços IP de origem únicos | Quantidade de flags “push” |
| Endereços IP de destino únicos | Quantidade de flags “reset” |
| Endereços MAC de origem únicos | Quantidade de flags “finish” |
| Endereços MAC de destino únicos | |
| Mínimo/Máximo/Amplitude dos tamanhos dos pacotes | |
| Mínimo/Máximo/Mediana do TTL dos pacotes | |

Tabela 1 – Atributos brutos extraídos no Experimento 1

Na Tabela 2, observa-se que os dados baseados em características filtradas por possuírem baixa correlação entre si apresentaram os maiores valores em todas as métricas de avaliação: Acurácia (76.73%), *recall* Ponderado (76.23%), F1-Score Ponderado (80.17%) e uma Precisão Ponderada de 87.48%, com tempo de predição de 1:33:34, ou seja, no primeiro segundo da preparação do ataque. Além disso, essa configuração utilizou apenas cinco variáveis como entrada, em contraste com as 164 características disponíveis antes da filtragem. As características selecionadas nesse cenário foram: (*i*) entropia dos pacotes de outros protocolos; (*ii*) informação de Fisher dos pacotes de outros protocolos; (*iii*) probabilidade de autotransição dos pacotes com a flag "ack"do protocolo TCP; (*iv*) probabilidade de autotransição do total de pacotes; e (*v*) probabilidade de autotransição do TTL mínimo dos pacotes.

Já as abordagens baseadas apenas nos atributos brutos, com ou sem filtragem, apresentaram desempenho inferior, exceto na precisão ponderada. Além disso, mantiveram o tempo de predição inalterado (1:33:32) e quantidades de variáveis significativamente menores (21 e 10, respectivamente). O uso direto das características (sem filtragem) também apresentou bons resultados, com uma acurácia de 71.83% e F1-Score ponderado de 75.83%, porém com o maior número de atributos entre todos os cenários. Por fim, os resultados também se mostraram melhores nas métricas da etapa de engenharia de características precedendo a etapa de análise de correlação, conforme a Tabela 3.

| **Dados Usados** | **Acurácia (%)** | **Precisão Ponderada (%)** | ***Recall* Ponderado (%)** | **F1-Score Ponderado (%)** | **Tempo de predição** | **Qtd. variáveis** |
|---|---|---|---|---|---|---|
| Atributos com Correlação | 65.63 | 88.95 | 65.63 | 75.09 | 1:33:32 | 21 |
| Atributos sem Correlação | 65.84 | **88.97** | 65.84 | 75.24 | 1:33:32 | 10 |
| Características com Correlação | 71.83 | 82.45 | 71.83 | 75.83 | **1:33:34** | 164 |
| Características sem Correlação | **76.73** | 87.48 | **76.23** | **80.17** | **1:33:34** | **5** |

Tabela 2 – Predição de ataque no Experimento 1

| **Etapa Aplicada Primeiro** | **Acurácia (%)** | **Precisão Ponderada (%)** | ***Recall* Ponderado (%)** | **F1-Score Ponderado (%)** | **Tempo de predição** | **Qtd. variáveis** |
|---|---|---|---|---|---|---|
| Análise de correlação | 65.57 | 79.38 | 65.57 | 70.87 | **1:33:34** | 80 |
| Engenharia de características | **76.73** | **87.48** | **76.73** | **80.17** | **1:33:34** | **5** |

Tabela 3 – Resultados das ordens de aplicação das etapas no Experimento 1

## 4.2 Experimento 2

O Experimento 2 aplicou a captura 52, do *dataset* CTU-13, a qual possui uma quantidade de dados e de tempo significativamente menor em relação ao experimento anterior. Essa captura possui 972 segundos, sendo que o ataque começa no segundo 778. Os dados de treinamento foram limitados a 33% da preparação de ataque por conta do tamanho reduzido. Para esse experimento, o conjunto de atributos brutos extraídos teve tamanho 11, especificado na Tabela 4.

| **Atributos** | |
|---|---|
| Total de pacotes | Endereços IP de origem únicos |
| Total de pacotes TCP | Endereços IP de destino únicos |
| Total de pacotes UDP | Porta de origem mais frequente |
| Total de pacotes ICMP | Porta de destino mais frequente |
| Mínimo/Máximo/Soma dos tamanhos dos pacotes | |

Tabela 4 – Atributos brutos extraídos no Experimento 2 e 3

Neste experimento, o melhor desempenho geral foi obtido novamente pelo cenário com características filtradas durante a etapa de análise de correlação, que atingiu uma acurácia de 74.69%, *recall* ponderado de 74.69% e o maior F1-Score ponderado entre todos os cenários (80.74%). O tempo de predição nesse caso foi de 12 minutos e 52 segundos, e a quantidade de variáveis utilizadas foi reduzida para apenas 9, o que demonstra um ganho considerável de eficiência em relação ao uso de todas as características (88 variáveis).

O cenário com uso direto dos atributos apresentou desempenho intermediário, com acurácia de 67.96% e F1-Score de 75.93%. Já os cenários baseados nos atributos com correlação e sem correlação apresentaram desempenho semelhante, com destaque para a maior precisão ponderada alcançada ao usar os atributos com correlação (95.40%). Apesar disso, ambos apresentaram menor *recall* (67.55% e 67.14%, respectivamente), o que impactou no F1-Score final. Neste experimento, a filtragem dos atributos baseou-se na remoção de atributos altamente correlacionados, resultando na seleção de apenas duas variáveis: o total de pacotes UDP e de pacotes TCP.

A Tabela 6 demonstra a diferença entre a ordem da execução das duas etapas de processamento dos dados. É possível perceber um ganho considerável nas métricas ao aplicar a etapa de engenharia de características primeiro. Além disso, houve uma redução de dimensionalidade em relação à etapa de análise de correlação precedendo, apesar de uma pequena queda no tempo de predição.

| **Dados Usados** | **Acurácia (%)** | **Precisão Ponderada (%)** | ***Recall* Ponderado (%)** | **F1-Score Ponderado (%)** | **Tempo de predição** | **Qtd. variáveis** |
|---|---|---|---|---|---|---|
| Atributos com Correlação | 67.55 | **95.40** | 67.55 | 78.40 | **0:12:57** | 11 |
| Atributos sem Correlação | 67.14 | 95.09 | 67.14 | 78.13 | **0:12:57** | **2** |
| Características com Correlação | 67.96 | 90.37 | 67.96 | 75.93 | **0:12:57** | 88 |
| Características sem Correlação | **74.69** | 90.92 | **74.69** | **80.74** | 0:12:52 | 9 |

Tabela 5 – Predição de ataque no Experimento 2

| **Etapa Aplicada Primeiro** | **Acurácia (%)** | **Precisão Ponderada (%)** | ***Recall* Ponderado (%)** | **F1-Score Ponderado (%)** | **Tempo de predição** | **Qtd. variáveis** |
|---|---|---|---|---|---|---|
| Análise de correlação | 66.73 | 87.13 | 66.73 | 74.91 | **0:12:57** | 16 |
| Engenharia de características | **74.69** | **90.92** | **74.69** | **80.74** | 0:12:52 | **9** |

Tabela 6 – Resultados das ordens de aplicação das etapas no Experimento 2

## 4.3 Experimento 3

Por fim, o Experimento 3 faz uso do *dataset* CIC-DDoS2019, que tem o seu ataque iniciado no segundo 1483. Para essa captura, o conjunto de dados de treinamento possui 33% dos dados de preparação de ataque. O conjunto de atributos utilizado é o mesmo do Experimento 2, presente na Tabela 4.

A Tabela 7 apresenta os resultados do terceiro experimento. Novamente, foram avaliadas quatro abordagens distintas: atributos brutos com correlação, atributos filtrados sem

correlação, características extraídas com correlação e características filtradas sem correlação após engenharia. Dentre os cenários testados, o melhor desempenho foi alcançado com o uso de características sem correlação, que resultaram em acurácia de 56.85%, precisão ponderada de 55.62%, *recall* ponderado de 56.85% e F1-Score ponderado de 56.18%. Este cenário também se destacou por utilizar apenas 3 variáveis, mantendo o mesmo tempo de predição (0:24:44) dos demais. As características selecionadas após o processo de filtragem por análise de correlação foram: (*i*) complexidade estatística dos pesos das arestas do total de IPs de destino; (*ii*) entropia de permutação da distribuição de pesos das arestas do total de IPs de destino; e (*iii*) entropia de permutação da distribuição de pesos das arestas do total de pacotes UDP.

As abordagens baseadas em características completas (sem filtragem) também apresentaram desempenho considerável, com acurácia de 54.69% e F1-Score de 54.84%, embora utilizando um conjunto significativamente maior de atributos (88 variáveis). Em contraste, os cenários baseados apenas em atributos brutos, com ou sem filtragem, apresentaram desempenho inferior. O uso direto dos atributos resultou em acurácia de 41.42% e F1-Score de 41.27%, enquanto a filtragem dos atributos (mantendo apenas duas variáveis) reduziu ainda mais esses valores, indicando possível perda de informações relevantes nesse caso.

Os resultados apresentados na Tabela 8 demonstram, novamente, um ganho maior obtido ao se aplicar a etapa de engenharia de características antes da filtragem por análise de correlação. Apesar da ligeira vantagem nas métricas de classificação, essa sequência utilizou apenas 3 variáveis para obter os resultados, contra 16 variáveis da sequência inversa.

| **Dados Usados** | **Acurácia (%)** | **Precisão Ponderada (%)** | ***Recall* Ponderado (%)** | **F1-Score Ponderado (%)** | **Tempo de predição** | **Qtd. variáveis** |
|---|---|---|---|---|---|---|
| Atributos com Correlação | 41.42 | 41.12 | 41.42 | 41.27 | **0:24:44** | 11 |
| Atributos sem Correlação | 38.30 | 39.26 | 38.30 | 38.77 | 0:24:43 | **2** |
| Características com Correlação | 54.69 | 55.00 | 54.69 | 54.84 | **0:24:44** | 88 |
| Características sem Correlação | **56.85** | **55.62** | **56.85** | **56.18** | **0:24:44** | 3 |

Tabela 7 – Predição de ataque no Experimento 3 (CIC-DDoS2019)

| **Etapa Aplicada Primeiro** | **Acurácia (%)** | **Precisão Ponderada (%)** | ***Recall* Ponderado (%)** | **F1-Score Ponderado (%)** | **Tempo de predição** | **Qtd. variáveis** |
|---|---|---|---|---|---|---|
| Análise de correlação | 55.88 | 55.61 | 55.88 | 55.74 | 0:24:40 | 16 |
| Engenharia de características | **56.85** | **55.62** | **56.85** | **56.18** | **0:24:44** | **3** |

Tabela 8 – Resultados das ordens de aplicação das etapas no Experimento 3

## 4.4 Discussão dos resultados

A análise comparativa dos três experimentos evidencia padrões consistentes quanto à eficácia das diferentes abordagens aplicadas à predição de ataques DDoS. Em todos os cenários avaliados, a combinação entre a engenharia de características com base na Transformação de Bandt-Pompe e a posterior filtragem por correlação foi responsável pelos melhores desempenhos obtidos, tanto em termos de métricas de avaliação quanto na redução da dimensionalidade.

No Experimento 1, observou-se uma melhoria significativa de desempenho quando se utilizaram características extraídas e sem correlação: o F1-Score ponderado aumentou para 80.17%, e a acurácia chegou a 76.73%. A filtragem reduziu o conjunto de entrada de 164 para apenas 5 variáveis, mantendo o tempo de predição praticamente inalterado. Isso sugere que grande parte das informações redundantes ou ruidosas foi eliminada sem prejuízo, e sim ganho à performance do modelo.

O Experimento 2 reforçou essa tendência: características filtradas sem correlação produziram a maior acurácia (74.69%) e o maior F1-Score (80.74%), mesmo utilizando apenas 9 variáveis. Embora o modelo treinado com atributos brutos tenha apresentado a maior precisão (95.40%), seu baixo Recall (67.55%) indica que muitos ataques deixaram de ser identificados. Isso mostra que uma alta precisão isolada não garante a efetividade do sistema de predição, especialmente em cenários onde a sensibilidade à detecção de eventos maliciosos é crítica.

No Experimento 3, os ganhos trazidos pela engenharia de características tornam-se ainda mais evidentes. Enquanto os modelos baseados em atributos apresentaram desempenho abaixo de 42% em acurácia e F1-Score, a aplicação de características extraídas elevou essas métricas para até 56.85%, mesmo com um conjunto reduzido a apenas 3 variáveis após filtragem. Dado o maior tempo de predição neste experimento (cerca de 25 minutos), essa economia de recursos computacionais também é relevante, especialmente em contextos de operação em tempo real.

Outro ponto relevante diz respeito à eficácia da filtragem e da engenharia de características. Nos experimentos, o uso da engenharia de características baseada na transformação em padrões ordinais foi capaz de capturar dinâmicas sutis no tráfego de rede, refletindo variações comportamentais que precedem ataques DDoS que auxiliam na classificação dos dados. Da mesma forma, a seleção de variáveis por análise de correlação demonstrou a capacidade de manter e, até mesmo, melhorar os resultados, principalmente quando aplicada nos conjunto de características. Isso indica que a filtragem por correlação é mais eficaz quando utilizada em conjunto com uma técnica de engenharia de característica robusta.

# 5 Conclusão

Este trabalho teve como objetivo investigar o impacto da correlação entre atributos e da engenharia de características na predição de ataques DDoS a partir de tráfego de rede. Para isso, foram conduzidos três experimentos distintos utilizando diferentes conjuntos de dados e cenários de avaliação, comparando o desempenho de modelos treinados com dados brutos, atributos filtrados por correlação, características extraídas com base na Teoria da Informação e características filtradas após essa engenharia. Os resultados obtidos evidenciam que o uso de características derivadas a partir da transformação ordinal de Bandt-Pompe, combinada com descritores da Teoria da Informação, proporciona uma representação mais rica e discriminativa do tráfego de rede. Em todos os experimentos, essa abordagem superou, em desempenho, os modelos baseados apenas em atributos brutos, especialmente após a aplicação de técnicas de seleção de características.

Além disso, foi possível observar que a filtragem criteriosa de características, com base em correlação, contribui significativamente para a redução da dimensionalidade sem comprometer - e, em diversos casos, melhorando - o desempenho dos modelos. Essa combinação resultou em modelos mais leves, com menos variáveis de entrada e tempos de predição semelhantes ou até superiores, o que reforça a viabilidade de sua aplicação em cenários reais com restrições de tempo e recursos computacionais. Em especial, os melhores resultados de F1-Score foram obtidos nos cenários em que as características extraídas foram posteriormente filtradas, indicando que a extração de padrões ordinais a partir da estrutura temporal dos dados, seguida de uma eliminação de redundâncias, é uma estratégia eficaz para detectar comportamentos maliciosos sutis e predizer ataques DDoS.

# Referências

[Albano et al. 2024]ALBANO, L. et al. Seleção de características na predição de ataques ddos com transformação em padrões ordinais. In: *Anais do XXIX Workshop de Gerência e Operação de Redes e Serviços*. Porto Alegre, RS, Brasil: SBC, 2024. p. 126–139. ISSN 2595-2722. Disponível em: <https://sol.sbc.org.br/index.php/wgrs/article/view/30092>.

[Alhijawi et al. 2022]ALHIJAWI, B. et al. A survey on dos/ddos mitigation techniques in sdns: Classification, comparison, solutions, testing tools and datasets. *Computers and Electrical Engineering*, v. 99, p. 107706, 2022. ISSN 0045-7906. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0045790622000234>.

[Bandt e Pompe 2002]BANDT, C.; POMPE, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.*, American Physical Society, v. 88, p. 174102, Apr 2002. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.88.174102>.

[Borges et al. 2024]BORGES, L. et al. Multifaceted ddos attack prediction by multivariate time series and ordinal patterns. In: . [S.l.: s.n.], 2024.

[Borges et al. 2024]BORGES, L. F. et al. Multifaceted ddos attack prediction by multivariate time series and ordinal patterns. In: *2024 IEEE International Conference on Communications Workshops (ICC Workshops)*. [S.l.: s.n.], 2024. p. 427–432.

[Brodsky 2021]BRODSKY, Z. *The Psychology Behind DDoS: Motivations and Methods*. [S.l.], 2021. Disponível em: <https://www.perimeter81.com/blog/network/the-psychology-behind-ddos-attacks>. Acesso em: 26 jun. 2021.

[Dasari K.B. 2022]DASARI K.B., D. N. Tcp/udp-based exploitation ddos attacks detection using ai classification algorithms with common uncorrelated feature subset selected by pearson, spearman and kendall correlation methods. *Revue d'Intelligence Artificielle*, v. 36, n. 1, p. 61–71, 2022.

[de Neira, Kantarci e Nogueira 2023]de Neira, A. B.; KANTARCI, B.; NOGUEIRA, M. Distributed denial of service attack prediction: Challenges, open issues and opportunities. *Computer Networks*, v. 222, p. 109553, 2023. ISSN 1389-1286. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1389128622005874>.

[Garcia et al. 2014]GARCIA, S. et al. An empirical comparison of botnet detection methods. *Computers & Security*, v. 45, p. 100–123, 2014. ISSN 0167-4048.

[Gregorutti, Michel e Saint-Pierre 2016]GREGORUTTI, B.; MICHEL, B.; SAINT-PIERRE, P. Correlation and variable importance in random forests. *Statistics and*

*Computing*, Springer Science and Business Media LLC, v. 27, n. 3, p. 659–678, mar. 2016. ISSN 1573-1375. Disponível em: <http://dx.doi.org/10.1007/s11222-016-9646-1>.

[Kebede et al. 2022]KEBEDE, S. D. et al. Predictive machine learning-based integrated approach for DDoS detection and prevention. *Multimedia Tools and Applications*, v. 81, n. 3, p. 4185–4211, jan 2022. ISSN 1573-7721. Disponível em: <https://doi.org/10.1007/s11042-021-11740-z>.

[Klymash et al. 2020]KLYMASH, M. et al. Concept of intelligent detection of ddos attacks in sdn networks using machine learning. In: *2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC ST)*. [S.l.: s.n.], 2020. p. 609–612.

[LeCun, Bengio e Hinton 2015]LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, v. 521, n. 7553, p. 436–444, may 2015. ISSN 1476-4687. Disponível em: <https://doi.org/10.1038/nature14539>.

[Lima et al. 2023]LIMA, M. et al. Predição não-supervisionada de ataques ddos por sinais precoces e one-class svm. In: *Anais do XXIII Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais*. Porto Alegre, RS, Brasil: SBC, 2023. p. 403–416. ISSN 0000-0000. Disponível em: <https://sol.sbc.org.br/index.php/sbseg/article/view/27222>.

[Mao et al. 2006]MAO, Z. M. et al. Analyzing large ddos attacks using multiple data sources. In: *Proceedings of the 2006 SIGCOMM Workshop on Large-Scale Attack Defense*. New York, NY, USA: Association for Computing Machinery, 2006. (LSAD '06), p. 161–168. ISBN 1595935711. Disponível em: <https://doi.org/10.1145/1162666.1162675>.

[Mary E. Sabaridha e Rani 2014]MARY E. SABARIDHA, A. S. S. A. L.; RANI, M. U. An empirical study on dos attacks and ddos defense mechanism. In: _____. [S.l.]: Advances in Natural and Applied Sciences, American-Eurasian Network for Scientific Information, 2014. v. 8.

[Michalak e Kwasnicka 2006]MICHALAK, K.; KWASNICKA, H. Correlation-based feature selection strategy in classification problems. *International Journal of Applied Mathematics and Computer Science*, v. 16, p. 503–511, 01 2006.

[Mirkovic e Reiher 2004]MIRKOVIC, J.; REIHER, P. A taxonomy of ddos attack and ddos defense mechanisms. *SIGCOMM Comput. Commun. Rev.*, Association for Computing Machinery, New York, NY, USA, v. 34, n. 2, p. 39–53, abr. 2004. ISSN 0146-4833. Disponível em: <https://doi.org/10.1145/997150.997156>.

[N e Gupta 2020]N, T. R.; GUPTA, R. A survey on machine learning approaches and its techniques:. In: *2020 IEEE International Students' Conference on Electrical,Electronics and Computer Science (SCEECS)*. [S.l.: s.n.], 2020. p. 1–6.

[Pearson 1895]PEARSON, K. Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, The Royal Society, v. 58, p. 240–242, 1895. ISSN 03701662. Disponível em: <http://www.jstor.org/stable/115794>.

[Sen, Hajra e Ghosh 2020]SEN, P. C.; HAJRA, M.; GHOSH, M. Supervised classification algorithms in machine learning: A survey and review. In: MANDAL, J. K.; BHATTACHARYA, D. (Ed.). *Emerging Technology in Modelling and Graphics*. Singapore: Springer Singapore, 2020. p. 99–111. ISBN 978-981-13-7403-6.

[Shahbaz et al. 2016]SHAHBAZ, M. B. et al. On efficiency enhancement of the correlation-based feature selection for intrusion detection systems. In: *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. [S.l.: s.n.], 2016. p. 1–7.

[Sharafaldin et al. 2019]SHARAFALDIN, I. et al. Developing realistic distributed denial of service (ddos) attack dataset and taxonomy. In: *2019 International Carnahan Conference on Security Technology (ICCST)*. [S.l.: s.n.], 2019. p. 1–8.

[Silva, Neira e Nogueira 2022]SILVA, G. L. F. M. E.; NEIRA, A. B. de; NOGUEIRA, M. A deep learning-based system for ddos attack anticipation. In: *2022 IEEE Latin-American Conference on Communications (LATINCOM)*. [S.l.: s.n.], 2022. p. 1–6.

[Yin et al. 2023]YIN, Y. et al. Igrf-rfe: a hybrid feature selection method for mlp-based network intrusion detection on unsw-nb15 dataset. *Journal of Big Data*, v. 10, n. 1, p. 15, feb 2023. ISSN 2196-1115. Disponível em: <https://doi.org/10.1186/s40537-023-00694-8>.

[Yoachimik e Pacheco 2025]YOACHIMIK, O.; PACHECO, J. *Record-breaking 5.6 Tbps DDoS attack and global DDoS trends for 2024 Q4*. [S.l.], 2025. Disponível em: <https://blog.cloudflare.com/ddos-threat-report-for-2024-q4/>. Acesso em: 20 abr. 2025.