Seleção de Características na Predição de Ataques DDoS com Transformação em Padrões Ordinais

Lucas Albano[1], **Ligia F. Borges**[1], **Anderson B. de Neira**[2], **Michele Nogueira**[1,2]

[1]Departamento de Ciência da Computação - Universidade Federal de Minas Gerais

[2]Departamento de Informática - Universidade Federal do Paraná

{lucasalbano,ligia.borges,michele}@dcc.ufmg.br, abneira@inf.ufpr.br

Resumo. *A predição eficaz de ataques de negação de serviço distribuído (DDoS) enfrenta desafios. Ao pré-processar o tráfego de rede em busca de sinais de preparação do ataque, conjuntos de centenas de características introduzem ruído significativo devido a fortes correlações. Selecionar características relevantes evita ruídos e melhora o desempenho da predição. Este artigo avança a literatura de predição de ataques considerando a Transformação em Padrões Ordinais, uma abordagem capaz de revelar padrões sutis no tráfego de rede ao introduzir a seleção de características como parte do método. Um dos resultados obtidos antecipou a iminência efetiva do ataque DDoS em 49 minutos e 28 segundos, reduzindo o conjunto de características em 99,16%.*

Abstract. *The effective prediction of Distributed Denial of Service (DDoS) attacks faces several challenges. Throughout the preprocessing of network traffic and searching for signs of attack preparation, datasets with hundreds of features introduce significant noise due to strong correlations. Selecting relevant features removes noise and improves prediction performance. This paper advances the literature of attack prediction using Ordinal Pattern Transformation, an approach capable of unveiling subtle patterns in network traffic, by introducing feature selection as part of the method. Results have shown to anticipate an attack with 49 minutes and 28 seconds of lead time, even reducing the feature set by 99.16%.*

1. Introdução

Os ataques de negação de serviço distribuído (DDoS) negam o acesso de usuários legítimos a determinados recursos ou serviços [Yuan et al. 2017]. Esses ataques se apoiam em dispositivos infectados, geograficamente distribuídos, para gerar um volumoso tráfego malicioso contra um ou mais alvos. Esses ataques persistem como uma das principais ameaças, afetando diversas camadas de rede, protocolos e serviços. Esses ataques exaurem recursos computacionais e de rede dos sistemas, tornando-os uma ameaça contínua e multifacetada [Yuan et al. 2017]. Uma ilustração de tal fato consiste no ataque direcionado ao Google em setembro de 2023, que registrou aproximadamente 398 milhões de requisições por segundo, ultrapassando em apenas dois minutos o tráfego total de acesso à Wikipedia no mesmo mês [Kiner and April 2023]. Além disso, em 2023, a Cloudflare reportou a mitigação de cerca de 5,2 milhões de ataques DDoS na camada de aplicação e 8,7 milhões na camada de rede [Yoachimik and Pacheco 2024].

Predizer ataques difere da detecção de um ataque já em andamento; predizer um ataque DDoS significa avaliar a probabilidade de ocorrer um ataque em um futuro próximo com base na observação de sinais extraídos do tráfego de rede. Atacantes podem deixar vários tipos de sinais da preparação do ataque na rede, como busca por portas abertas em um servidor, execução de um *malware* em dispositivos da rede ou mesmo ataques de menor intensidade como teste. Porém, os sinais de preparação dos ataques são escassos e dificilmente distinguíveis [Peng et al. 2007]. O pré-processamento dos dados revela movimentos sutis dos atacantes [Albano et al. 2023, Borges et al. 2024], mas as centenas de características dos dados de tráfego de rede geram ruído significativo devido à forte correlação dos dados. O ruído prejudica o desempenho dos modelos de predição, aumenta sua complexidade, tempo e a quantidade de recursos computacionais (e.g., armazenamento e processamento dos dados) necessários para predição [Post et al. 2016].

Assim, em [Albano et al. 2023], empregou-se a Transformação em Padrões Ordinais de Bandt-Pompe [Bandt and Pompe 2002] para predizer ataques DDoS. Demonstrou-se ser possível identificar a preparação do ataque, porém, embora com resultados promissores, a técnica proposta em [Albano et al. 2023] não considerava a presença de centenas de características, sobretudo diante da intenção de aperfeiçoar a predição com múltiplas fontes de dados. Com a introdução de várias características heterogêneas (*i.e.*, de diferentes fontes), o ruído nos dados passa a ser crítico devido à forte correlação entre as mesmas. Nesse contexto, diversas pesquisas têm abordado a aplicabilidade da seleção de características em cenários relacionados à detecção de ataques DDoS e à análise de séries temporais multi-dimensionais [Zhou et al. 2022, Kathirgamanathan and Cunningham 2020, de Araujo et al. 2021], apesar de, no melhor do nosso conhecimento, não existirem estudos considerando a seleção de características para a predição de ataques DDoS.

Este artigo apresenta uma evolução do método proposto em [Albano et al. 2023] através da introdução da seleção de características. A variação do método aprimora o desempenho da predição de ataques ao remover o ruído causado por correlações fortes, enquanto se reduz significativamente a dimensionalidade dos dados. Ao escolher as melhores características, é possível predizer um ataque com maior antecedência e aumentar a taxa de acertos dos modelos. A proposta de variação do método consiste em monitorar o tráfego de rede, extrair atributos relevantes, utilizar a Transformação em Padrões Ordinais em conjunto de descritores da Teoria da Informação para revelar os sinais de preparação do ataque, selecionar as características para reduzir o ruído e custo computacional, e por fim, treinar um modelo de classificação para identificar o tráfego da preparação de um ataque DDoS. Portanto, este artigo avança a predição de ataque DDoS via uma nova abordagem e compara diferentes algoritmos de seleção de características após a aplicação da Transformação em Padrões Ordinais.

Para a avaliação da variante proposta, foram conduzidos experimentos com capturas de tráfego de rede do conjunto de dados *Czech Technical University dataset* (CTU-13) [Garcia et al. 2014]. Este conjunto de dados possui os cabeçalhos de pacotes durante a preparação de ataques DDoS, possibilitando a predição. O modelo de aprendizado de máquina *One Class SVM*, escolhido devido ao seu sucesso na predição [Albano et al. 2023], baixo custo computacional e por independer de dados rotulados, foi aplicado às características selecionadas por diferentes métodos de seleção,

visando estabelecer comparações. Em um dos resultados, a proposta conseguiu predizer o ataque com 49 minutos e 28 segundos de antecedência, superando a literatura. Outro resultado atingiu 95,61% de acurácia com um conjunto de somente 3 características.

Este artigo procede da seguinte forma. A Seção 2 apresenta os trabalhos relacionados. A Seção 3 detalha a variante proposta do método apresentado em [Albano et al. 2023], descrevendo a Transformação em Padrões Ordinais e os algoritmos de seleção de características utilizados. A Seção 4 apresenta a avaliação da variante. Por fim, na Seção 5, são apresentadas as conclusões.

2. Trabalhos Relacionados

Esforços na literatura predizem ataques DDoS, ou seja, identificam sinais da iminência de uma sobrecarga maliciosa contra um alvo [Jyoti and Behal 2021]. Porém, os trabalhos dessa linha de pesquisa enfrentam limitações, tais como encontrar características ideais para identificação da preparação do ataque, ao mesmo tempo em que evitam a introdução de ruído nos modelos. Em [Silva et al. 2022], métricas estatísticas auxiliam técnicas de aprendizado profundo para identificar sinais da iminência de um ataque. Similarmente, o estudo de [Brito et al. 2023] propõe um sistema baseado em *Long Short-Term Memory Autoencoder* que automaticamente configura a melhor rede neural e utiliza métricas estatísticas para predizer o ataque. Embora essas propostas atuem na predição de ataques, elas exigem recursos computacionais e tempo de treinamento significativos, devido à complexidade dos modelos.

A técnica de transformação ordinal de Bandt-Pompe [Bandt and Pompe 2002] destaca-se como uma abordagem capaz de revelar mudanças no comportamento da rede e prever ataques. Esta técnica é resistente a *outliers*, visto que desconsidera a amplitude das observações e se concentra nas relações de ordem. Em [Albano et al. 2023], essa técnica foi aplicada a diversas capturas de rede para predizer ataques DDoS de maneira eficaz, alcançando 44 minutos de antecedência em relação ao primeiro alerta do ataque, com uma acurácia de 89%. No entanto, esse estudo não considera atributos do tráfego de diferentes camadas de rede. O tráfego de rede multidimensional transporta dados heterogêneos de diferentes camadas de rede da pilha de protocolos e oferece múltiplas perspectivas para enriquecer a análise do sistema. Neste sentido, [Borges et al. 2024] propuseram a correlação de dados multifacetados para predizer ataques DDoS. A solução transforma séries temporais multivariadas em padrões ordinais para projetar características representativas. Contudo, à medida que a dimensionalidade dos dados aumenta, mais processamento e armazenamento se tornam necessários. Além disso, características redundantes podem ser inseridas e causar ruído.

A seleção de características reduz a dimensionalidade dos dados e, consequentemente, o ruído. É amplamente utilizada na literatura, inclusive no contexto de séries temporais multidimensionais [Kathirgamanathan and Cunningham 2020], e também é explorada na detecção de ataques DDoS [Osanaiye et al. 2016]. A literatura indica que a seleção de características aprimora os resultados dos modelos e reduz sua complexidade [Polat et al. 2020, de Araujo et al. 2021, Osanaiye et al. 2016]. No entanto, as pesquisas se concentram exclusivamente na seleção de características para a detecção de ataques ou focam em somente uma classe de métodos de seleção (*Filter, Wrapper* ou *Embedded*) [Bouzoubaa et al. 2021]. Este estudo avança a literatura ao introduzir a seleção

de características na técnica de predição de ataques DDoS via Transformação em Padrões Ordinais, além de avaliar o desempenho de diferentes classes de métodos de seleção.

3. Método

Esta seção detalha o método de seleção de características para a predição de ataques DDoS baseado na transformação ordinal, uma variação do método apresentado em [Albano et al. 2023]. Esta variação inclui o processo de seleção de característica visando reduzir ruídos causados pelas centenas de características presentes nos dados, tornando assim a predição de ataques DDoS mais eficiente. A escolha dos algoritmos utilizados priorizou o baixo custo computacional para permitir futuramente a implementação do método de forma *online*. Esta variação do método compreende quatro etapas: (*i*) coleta de dados do tráfego da rede; (*ii*) transformação em padrões ordinais e seleção de características; (*iii*) treinamento do modelo; (*iv*) predição do ataque.

Na Etapa 1 da Figura 1, um software *sniffer* é empregado para coletar o tráfego de rede. A Etapa 2 processa o tráfego de rede sob a perspectiva da transformação em padrões ordinais. A transformação em padrões ordinais é realizada em conjunto de descritores da Teoria da Informação [Albano et al. 2023, Borges et al. 2024]. Assim, o método proposto cria novas características para revelar as sutis ações dos atacantes. Em seguida, o método realiza a seleção de características para identificar as mais significativas. Na Etapa 3, o algoritmo de aprendizado de máquina *One-Class* SVM é treinado usando como base as características obtidas na Etapa 2. Na Etapa 4, o método usa o modelo para identificar ações dos atacantes que antecedem o lançamento do ataque e realizar a predição deles.

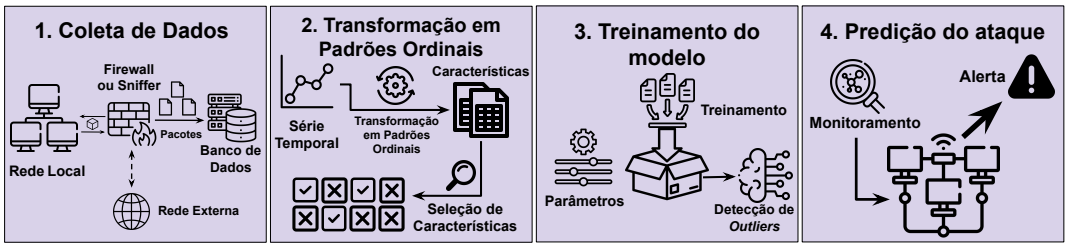

Figura 1. Visão Geral do Método

3.1. Coleta de Dados

O tráfego de rede é coletado usando uma ferramenta integrada ao *firewall* ou por meio de uma máquina equipada com um *software sniffer*, capaz de interceptar a entrada e saída dos pacotes da rede monitorada. O método proposto coleta o tráfego de rede de forma centralizada ou distribuída. Na estratégia centralizada, uma cópia de todos os pacotes (ou somente dos cabeçalhos, a depender dos objetivos e capacidade de processamento) é encaminhada para o dispositivo responsável pelas etapas seguintes. No entanto, a depender do volume de tráfego de dados da rede, esse processo exige recursos computacionais elevados. O método proposto também foi projetado para coletar o tráfego de rede distribuidamente. Nesse caso, diversos dispositivos em diferentes pontos da rede capturam o tráfego simultaneamente, enviando os pacotes para um banco de dados para processamento futuro. Embora essa estratégia evite atrasos no processamento, ela também dificulta a operacionalização do método. As informações dos pacotes, ou de seus cabeçalhos, são então armazenadas para viabilizar as etapas subsequentes.

Durante a preparação ou execução de um ataque DDoS, alguns atributos da rede sofrem alterações em seu comportamento. A identificação precisa dos atributos afetados depende do vetor de ataque. Trabalhos que visam identificar atributos relevantes para o treinamento de modelos de detecção e previsão de ataques indicam diversas possibilidades [Dayal and Srivastava 2017, Feng et al. 2018]. Exemplos de atributos de tráfego frequentemente citados incluem o número total de pacotes enviados e recebidos, a quantidade de diferentes endereços IP de origem e destino, a entropia das portas de origem e destino e métricas estatísticas do tamanho dos pacotes.

O método proposto coleta 45 atributos do tráfego de rede (Tabela 1). O objetivo de definir esses atributos foi proporcionar a proposta várias fontes de informação para maximizar a predição dos ataques DDoS por meio da Etapa 2 (transformação em padrões ordinais e seleção das características). Mesmo com a definição desses atributos, os usuários podem usar outros seguindo suas necessidades. O método proposto extrai os atributos da Tabela 1 por intervalos de tempo e os organiza sequencialmente pela ordem de captura. Por padrão, os intervalos de tempo possuem um segundo. Contudo, os intervalos podem ser ajustados pelo usuário. Assim, cada atributo do tráfego de rede coletado representa uma série temporal manipulada pelo método proposto.

Tabela 1. Atributos do tráfego de rede coletados

Atributos	
Total de Pacotes no Intervalo	Quantidade de flags “reserved” no intervalo
Porcentagem de pacotes TCP/UDP/outros no intervalo	Quantidade de flags “cwr” no intervalo
Endereços IP de origem únicos no intervalo	Quantidade de flags “urgent” no intervalo
Endereços IP de destino únicos no intervalo	Quantidade de flags “acknowledgment” no intervalo
Endereços MAC de origem únicos no intervalo	Quantidade de flags “push” no intervalo
Endereços MAC de destino únicos no intervalo	Quantidade de flags “reset” no intervalo
Média/Mediana/Desvio Padrão/Variância/Mínimo/Máximo/ Primeiro e Terceiro quartil de tamanho dos pacotes no intervalo	Quantidade de flags “synchronize” no intervalo
Média/Mediana/Desvio Padrão/Variância/Mínimo/Máximo/ Primeiro e Terceiro quartil do tempo entre pacotes no intervalo	Quantidade de flags “finish” no intervalo
Média/Mediana/Desvio Padrão/Variância/Mínimo/Máximo/ Primeiro e Terceiro quartil do TTL dos pacotes no intervalo	Quantidade de flags “ae” no intervalo
Entropia das portas de destino no intervalo	Quantidade de flags “ece” no intervalo
Entropia das portas de origem no intervalo	Entropia do número da sequência TCP no intervalo

3.2. Transformação em Padrões Ordinais

A segunda etapa do método proposto define a transformação em padrões ordinais e a seleção de características. A transformação em padrões ordinais consiste na aplicação da Transformação Ordinal de Bandt-Pompe [Bandt and Pompe 2002] sobre as séries temporais do tráfego de rede coletado na etapa anterior. Deste modo, ao fim da transformação em padrões ordinais, o método proposto cria novas características baseadas nos descritores da Teoria da Informação sem a influência de *outliers* no tráfego de rede [Rosso et al. 2012]. A hipótese que embasa a transformação em padrões ordinais é a possibilidade dela revelar os movimentos sutis dos atacantes durante a preparação do ataque. Essa combinação tem demonstrado eficácia na caracterização da dinâmica de séries temporais, tanto para a detecção de bots quanto para a predição de ataques DDoS [Borges et al. 2022, Chagas et al. 2022, Albano et al. 2023, Borges et al. 2024]. A transformação de Bandt-Pompe consiste em representar as relações de ordem entre pontos consecutivos de uma série temporal como padrões. Para obter o conjunto de padrões

ordinais das séries temporais capturadas são necessários dois processos: (*i*) dividir a série temporal em subconjuntos e (*ii*) obter o padrão ordinal de cada subconjunto através do índice da permutação dos pontos em ordem crescente. A Figura 2 ilustra esse processo. Considerando um subconjunto de dimensão 3, a técnica avança ao longo da série temporal caracterizando os padrões formados pelos índices da permutação em ordem crescente entre as observações. Ao fim da transformação, é obtida uma nova representação relativa aos padrões ordinais encontrados ao percorrer a série temporal.

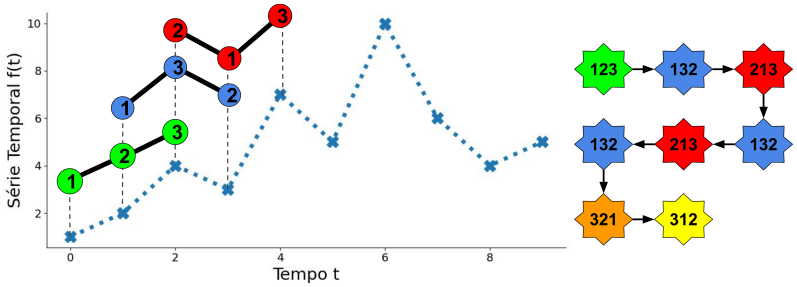

Figura 2. Transformação de Série Temporal para Padrões Ordinais

Essa transformação possibilita a realização de dois tipos de análises: (i) a distribuição de probabilidade de cada padrão e (ii) as transições entre os padrões ao longo da série. O cálculo da distribuição de probabilidade de padrões ordinais pode ser conduzido por meio do histograma de frequência. Nesse sentido, é necessário extrair a distribuição de probabilidade do conjunto de padrões ordinais resultante da transformação da série temporal [Chagas et al. 2022]. A segunda análise do conjunto de padrões ordinais é baseada na transformação da série temporal em grafos de transição. Para representar um conjunto de padrões ordinais como um grafo orientado que ilustra as transições entre dois padrões ordinais consecutivos ao longo do tempo, a nova representação atribui a cada vértice, um padrão, enquanto as arestas denotam as transições entre eles [Borges et al. 2022].

A partir da análise do histograma de frequência e do grafo de transição gerados pela transformação em padrões ordinais, são derivadas novas características. Para isso, são empregados três quantificadores da Teoria da Informação: entropia de permutação, complexidade estatística e informação de Fisher, os quais já foram comprovados na literatura como ferramentas eficazes para descrever o comportamento dinâmico observado em séries temporais [Ribeiro et al. 2017]. Com base nisso, são obtidas as seguintes características representativas relacionadas aos atributos do tráfego de rede: (i) entropia de permutação normalizada; (ii) complexidade estatística; (iii) medida de informação de Fisher; (iv) número de vértices no grafo de transição; (v) entropia de permutação normalizada da distribuição de pesos das arestas; (vi) complexidade estatística da distribuição de pesos das arestas; (vii) medida de informação de Fisher da distribuição de pesos das arestas; e $(viii)$ probabilidade de autotransição entre os vértices.

A entropia de permutação é utilizada para avaliar a complexidade e a imprevisibilidade da sequência de padrões ordinais. Quanto maior a entropia de permutação, maior a diversidade e imprevisibilidade dos padrões na sequência temporal, enquanto valores menores indicam uma série temporal mais determinística [Bandt and Pompe 2002]. A complexidade estatística identifica padrões e anomalias nos dados das séries temporais

capturadas. A complexidade estatística dos padrões ordinais pode ser avaliada usando a Entropia de Permutação, que mede o desequilíbrio em relação a uma distribuição esperada. A informação de Fisher é empregada para caracterizar a complexidade estatística dos padrões presentes na série temporal, fornecendo percepções sobre a complexidade e desequilíbrio desses padrões. Para isso, a métrica considera as diferenças entre as probabilidades consecutivas na distribuição. Por fim, a probabilidade de autotransição no grafo (i.e., padrões ordinais consecutivos) é utilizada para compreender a correlação temporal das séries [Borges et al. 2022]. A probabilidade de autotransição é calculada como a soma dos pesos das arestas na diagonal principal do grafo de transição. As equações referentes a essas métricas são descritas em [Albano et al. 2023].

3.2.1. Seleção de Características

Após a obtenção das características resultantes da transformação em padrões ordinais, o método proposto realiza a seleção destas. A escolha criteriosa de características desempenha um papel fundamental na otimização do desempenho dos modelos preditivos, contribuindo para a eficiência, interpretabilidade e redução da complexidade computacional do método [de Araujo et al. 2021]. Vale ressaltar que algumas características são fundamentais para a predição dos ataques e outras podem ter pouca relevância. Em alguns casos até mesmo piorar o modelo com a introdução de ruído [Hasan et al. 2016, Zhou et al. 2022]. Portanto, para alcançar o melhor desempenho e tempo de predição em cada cenário, é crucial selecionar o melhor subconjunto de características.

A literatura apresenta diversos métodos para a seleção de características. Esses métodos podem ser categorizados em três grupos: *Filters*, *Wrappers* e *Embedded*. Os métodos *filter* univariados analisam cada característica individualmente por meio de cálculos como variância, dependência e ganho de informação dos dados. Os métodos *filters* são, geralmente, mais rápidos e independentes dos modelos de classificação. No entanto, podem selecionar características redundantes ao ignorar as relações entre estas [Bouzoubaa et al. 2021], para evitar isto métodos *filters* multivariados podem ser utilizados [Urbanowicz et al. 2018]. Os métodos *wrapper*, por sua vez, empregam o desempenho de algoritmos de aprendizado de máquina como critério de pontuação para cada subconjunto de características. Em métodos *wrapper*, um novo modelo deve ser treinado para cada subconjunto de características, o que possibilita encontrar o conjunto com melhor desempenho para este modelo; no entanto, isso é computacionalmente custoso [de Araujo et al. 2021].

Os métodos *embedded* realizam a seleção de variáveis em conjunto com o treinamento do modelo. Geralmente são mais eficientes que os métodos *wrapper*, porém seus resultados são dependentes do algoritmo de aprendizado de máquina utilizado [Guyon and Elisseeff 2003]. Para a escolha dos métodos a serem avaliados, considerou-se o grande número de atributos de rede coletados. Assim, para manter o baixo custo computacional do método proposto, este trabalho não usou métodos de seleção do tipo *wrapper*. Os algoritmos de seleção de características selecionados para os experimentos são detalhados a seguir.

Os métodos *Principal Component Analysis* (PCA) e *Analysis of Variance* (ANOVA) foram escolhidos como base para a comparação. Estes são algoritmos *filter*

univariados, amplamente utilizados em análises estatísticas e de aprendizado de máquina para selecionar características relevantes em conjuntos de dados. A seleção de características com **PCA** é feita considerando os componentes principais que explicam a maioria da variabilidade nos dados. Já o **ANOVA** compara as médias de diferentes grupos para determinar se há diferenças estatisticamente significativas entre eles em relação a uma variável dependente [Guyon and Elisseeff 2003].

O **MultiSURF** (*Multiple Surrogate ReliefF*) é um método de seleção de características *filter* multivariado que avalia a importância das características considerando as interações entre elas. Ele utiliza vizinhos próximos como referência para determinar o quanto uma característica influencia na classificação ou na regressão, sendo especialmente útil em conjuntos de dados com características altamente correlacionadas [Urbanowicz et al. 2018].

O **LASSO** (*Least Absolute Shrinkage and Selection Operator*) é um método de regularização usado para seleção de características e redução de dimensionalidade em modelos de regressão [Tibshirani 1996]. Seguindo a abordagem *embedded*, ele penaliza os coeficientes dos preditores menos importantes, eliminando-os ou tornando-os muito pequenos. Isso evita *overfitting* e auxilia na identificação das características mais relevantes para a predição. O **XGBoost** (*Extreme Gradient Boosting*) é um algoritmo de aprendizado de máquina baseado em árvores de decisão. Ele utiliza um processo de *ensemble* para melhorar a precisão preditiva, ajustando sequencialmente os modelos para corrigir os erros dos anteriores e pode ser utilizado como um método de seleção de características do tipo *embedded* [Chen and Guestrin 2016].

3.3. Treinamento do Modelo

Na Etapa 3, emprega-se o algoritmo de aprendizado de máquina *One-Class* SVM para identificação de *outliers*. Este algoritmo opera de maneira não supervisionada, necessitando apenas de dados normais para o treinamento [Amer et al. 2013]. O *One-Class* SVM determina um limite de decisão que separa os dados normais dos atípicos, sendo este limite definido por um hiperplano que maximiza a margem entre os dados. Este hiperplano, por sua vez, categoriza novos pontos de dados como normais ou atípicos com base em seu posicionamento em relação a ele.

A eficácia do *One-Class* SVM na classificação de *outliers* está condicionada a alguns parâmetros. O parâmetro *nu* atua como um limiar para a taxa de erros durante o treinamento e como um limite inferior para a proporção de vetores de suporte. Ao ajustar o valor do parâmetro *nu*, é possível controlar a quantidade de pontos de dados considerados *outliers* durante o treinamento, variando entre [0, 1] para representar a porcentagem de pontos classificados como *outliers*. Além disso, o parâmetro *Kernel* determina como os dados são separados no espaço, oferecendo várias opções, como *linear*, *poly* e *sigmoid*. Estes parâmetros se adaptam aos dados, ajustando-se conforme as necessidades específicas do problema em questão.

3.4. Predição do Ataque

Após o treinamento, o modelo identifica *outliers* (*i.e.*, alterações no tráfego de rede) que podem não ser perceptíveis normalmente. A transformação em padrões ordinais realizada na Etapa 2 permite a correlação entre os diversos padrões de dados identificados nas séries

temporais, auxiliando a determinar se o tráfego monitorado foi impactado pelas diversas possíveis ações que simbolizam a preparação de ataques DDoS. O método proposto pode ser implementado ao nível de rede para notificar a equipe responsável sobre a preparação de ataques DDoS. Após a notificação, as equipes de segurança podem tomar ações para evitar que o ataque DDoS interrompa a disponibilidade do serviço e cause prejuízos.

4. Avaliação

Essa seção descreve os experimentos conduzidos para avaliar o desempenho do método equipado com as diferentes técnicas da seleção de características. Para avaliar a predição dos ataques, os conjuntos de dados devem possuir tráfego de rede prévio ao ataque, a identificação das atividades maliciosas e do início do ataque. A predição ocorre quando o método proposto identifica corretamente sinais da preparação dos ataques causados pelo tráfego maligno antes do início do ataque. Testes de ataques, busca por portas abertas ou infecção de dispositivos na rede são exemplos de ações de preparação dos atacantes [Griffioen et al. 2021]. O método proposto não utiliza dados rotulados e manipula somente o tráfego anterior ao ataque para seleção de características, treinamento e teste dos modelos. Assim, o método fica independente do vetor de ataque e pode avaliar diferentes métodos de seleção de características aliados com a predição dos ataques.

4.1. Definição dos Experimentos

O **Experimento 1** emprega a captura 51 do *dataset* CTU-13 [Garcia et al. 2014]. Esta captura de tráfego de rede local compreende 8803 segundos, contendo 46 milhões de pacotes, possuindo ataques do tipo *Internet Control Message Protocol* (ICMP) e *User Datagram Protocol* (UDP) *flood*, executados por dez *bots*. Os pesquisadores iniciam a preparação no segundo 2643 e lançam os ataques no segundo 5614 da captura, possibilitando assim a predição. Portanto, o conjunto de dados utilizado se restringe ao tráfego prévio ao início do ataque, dividido em um terço para seleção de características e treinamento, e dois terços para teste.

O **Experimento 2**, por sua vez, utiliza a captura 52 do CTU-13. A captura compreende 972 segundos, contém 6 milhões de pacotes, um ataque do tipo ICMP *flood* e três *bots*. A preparação se iniciou com 527 segundos e o ataque foi conduzido pelos pesquisadores no segundo 797 da captura. É importante ressaltar que a predição do ataque nesta captura é desafiadora devido ao seu tamanho reduzido, limitando o conjunto de treino e o tempo disponível para a predição do ataque. Da mesma forma que no experimento anterior, uma divisão é feita no conjunto de dados, reservando um terço para seleção de características e treinamento do modelo, e o restante para teste até o início do ataque.

O processamento das capturas segue os passos delineados na Seção 3. Um intervalo de 1 segundo foi escolhido para segmentar as capturas devido aos tamanhos destas e por ser adequado para implementações em ambientes reais. Uma janela deslizante com tamanho relativo a 5% do total de cada captura foi utilizada para a aplicação da técnica. Isto é necessário, pois a transformação é realizada em uma série temporal completa e o objetivo é avaliar os dados ao longo da série.

O modelo de classificação escolhido, o *One-Class* SVM, não requer dados rotulados para treinamento. No entanto, durante a extração de atributos, os intervalos são rotulados como malignos sempre que houver atividade de algum bot nesse intervalo. Embora

essa rotulagem não seja viável em ambientes reais, ela é essencial para avaliar o desempenho e o tempo de predição dos modelos, além de permitir a comparação de métodos de seleção de características que dependam de dados rotulados. Dessa forma, o modelo *One-Class* SVM é treinado e testado aumentando progressivamente o conjunto de características, utilizando uma lista ordenada por importância fornecida por cada um dos métodos descritos na Subseção 3.2.1. Os parâmetros do *One-Class* SVM foram mantidos constantes em todos os experimentos, seguindo os padrões da biblioteca Scikit-Learn[1], com exceção do parâmetro *nu*, definido como 0.05, e o *kernel*, definido como *poly*. Estes parâmetros foram definidos manualmente para maximizar a eficácia do modelo nos experimentos.

A avaliação dos resultados utiliza a acurácia, precisão, *recall*, *F1-score* e o tempo de antecedência do alerta do ataque. A acurácia representa a proporção de amostras classificadas corretamente em relação ao total de amostras. A precisão indica a proporção de observações corretamente classificadas como um tipo específico em relação ao total de observações classificadas como esse tipo. O *recall* apresenta a proporção de todas as observações esperadas de um tipo específico que foram corretamente classificadas pela técnica. O *F1-score* é uma medida que combina precisão e *recall*, calculada como a média harmônica dessas duas métricas. Devido ao desbalanceamento inerente ao ataque DDoS e à variação na quantidade de amostras nas classes, é necessário considerar a precisão, *recall* e *F1-score* ponderadamente. Neste trabalho, utiliza-se a média ponderada da precisão, *recall* e *F1-score* considerando a quantidade de amostras em cada classe. Isso permite uma avaliação mais equilibrada e representativa. O tempo de predição corresponde ao tempo entre o primeiro alerta correto produzido pelo modelo e o início efetivo do ataque DDoS.

4.2. Resultados

Esta subseção apresenta os resultados obtidos a partir da aplicação dos algoritmos de seleção de características para a predição de ataque DDoS. Estes experimentos exploram a relação entre a acurácia dos modelos, o tempo de antecedência dos alertas e a quantidade de características, utilizando a transformação do tráfego de rede em padrões ordinais. Os principais resultados são discutidos a seguir e todos eles estão disponíveis online[2].

No **Experimento 1**, ao aplicar a transformação em padrões ordinais, o conjunto de características aumenta de 45 para 360. A Tabela 2 apresenta os principais resultados obtidos durante a avaliação da seleção de características. O método proposto equipado com o método MultiSURF foi a versão do método que obteve o melhor equilíbrio entre acurácia e tempo de predição. Neste caso, a transformação ordinal aliada com a seleção de características proporcionou a predição do ataque DDoS com acurácia de 88,89%. Também é oportuno ressaltar que esse resultado foi obtido utilizando 5 características. As características selecionadas são: Complexidade do TTL mínimo, Informação de Fisher dos Ips de Origem Únicos, Entropia do TTL mínimo, Informação de Fisher do TTL mínimo e Probabilidade de autotransição da mediana do TTL dos pacotes. Por fim, o XGBoost e o PCA obtiveram o maior tempo de predição dos ataques. Nestes casos, o método proposto predisse o ataque com 49 minutos e 28 segundos antes do lançamento do ataque.

[1]https://scikit-learn.org/stable/modules/generated/sklearn.svm.OneClassSVM.html
[2]github.com/EoSingle/WGRS24

Tabela 2. Principais Resultados Obtidos no Experimento 1

Algoritmo	Características	Acurácia	Precisão	Recall	F1-Score	Predição
PCA	3	75,48%	85,24%	75,48%	79,96%	**49m 28s**
PCA	303	86,76%	85,8%	86,76%	86,27%	33m 48s
ANOVA	7	85,23%	87,32%	85,23%	86,23%	48m 9s
MultiSURF	5	88,89%	**87,63%**	88,89%	88,24%	48m 39s
LASSO	3	**92,38%**	86,49%	**92,38%**	**89,05%**	45s
LASSO	4	80,41%	86,28%	80,41%	83,15%	45m 41s
XGBoost	6	71,53 %	85%	71,53%	77,51%	**49m 28s**
XGBoost	168	86,09%	86,34%	86,09%	86,22%	49m 24s

No **Experimento 2,** devido ao tamanho reduzido da amostra, a predição dos ataques torna-se desafiadora. Com a transformação em padrões ordinais, o método proposto realiza a seleção de características em um conjunto de 360 características. A Tabela 3 apresenta os principais resultados obtidos neste experimento. O XGBoost selecionou três características (Informação de Fisher do número de flags ACK, Complexidade dos endereços MAC de origem únicos e Entropia do grafo de transição do número de flags PSH) que proporcionou a predição do ataque DDoS com 95,61% de acurácia. O maior tempo de predição foi obtido pelo MultiSURF a partir das 32 características selecionadas. Neste caso, a predição ocorreu 4 minutos e 9 segundos antes do lançamento do ataque.

Tabela 3. Principais Resultados Obtidos no Experimento 2

Algoritmo	Características	Acurácia	Precisão	Recall	F1-Score	Predição
PCA	297	88,24%	94,48%	88,24%	91,09%	2m 43s
ANOVA	1	93,82%	**95,48%**	93,82%	94,57%	3m 38s
MultiSURF	14	89,44%	95,31%	89,44%	92%	3m 41s
MultiSURF	64	86,45%	95,40%	86,45%	90,26%	**4m 9s**
LASSO	8	89,24%	93,74%	89,24%	91,41%	2m 10s
XGBoost	3	**95,61%**	94,27%	**95,61%**	**94,9%**	2m 10s
XGBoost	14	89,64%	93,76%	89,64%	91,63%	4m 6s

4.3. Discussão

No Experimento 1, a seleção de características que otimizou o tempo de predição foi a versão do método proposto, equipada com XGBoost, que selecionou 6 características, ou com PCA, com três características. Em ambas as configurações, o tempo de predição foi de 49 minutos e 28 segundos. Apesar deste tempo de predição, as acurácias obtidas variaram entre 75,48% e 71,53% representando as menores acurácias conforme observado na Tabela 2. O MultiSURF selecionou cinco características que proporcionam o melhor equilíbrio das demais métricas, atingindo 88,89% de acurácia. Neste caso, a predição ocorreu apenas 49 segundos mais tarde do que na melhor predição. Outro importante resultado em termos de acurácia foi obtida com o LASSO com três características. Neste caso, a acurácia atingiu 92,38%. Contudo, seu tempo de predição caiu para apenas 45 segundos. Assim, neste experimento, o método MultiSURF se destaca como aquele que alcança o melhor equilíbrio entre tempo de predição e desempenho.

No Experimento 2, ao utilizar o conjunto de uma característica selecionada pelo método ANOVA, foi possível identificar a preparação do ataque com 3 minutos e 38 segundos de antecedência com uma acurácia 93,82%. Embora este resultado não seja o ótimo em termos de precisão ou tempo de predição, é notável por usar exclusivamente a característica relacionada ao número de nós no grafo de transição, considerando o atributo endereços MAC de origem. O melhor tempo de predição foi alcançado pelo MultiSURF a partir de 32 características (o máximo de acurácia foi obtido com 64) identificando o ataque 4 minutos e 9 segundos antes de seu início. A melhor acurácia, *recall* e *F1-score* foram obtidos pelo XGBoost com três características. No entanto, a predição ocorre com somente 2 minutos e 10 segundos de antecedência. É importante destacar que a predição nesta captura é desafiadora, pois esta possui apenas alguns minutos, limitando o conjunto de treino assim como o tempo disponível para a predição.

Com base nos resultados obtidos, é possível concluir que a seleção de características, juntamente com a transformação em padrões ordinais, apresenta resultados promissores para a predição de ataques DDoS. Devido ao ineditismo desta abordagem na predição de ataques, é difícil realizar a comparação com outros trabalhos, pois estes são escassos e muitos não empregam as mesmas métricas para avaliação. Dessa forma, os resultados superam o tempo de predição de trabalhos com o mesmo conjunto de dados [Albano et al. 2023, Brito et al. 2023, Rahal et al. 2020, de Neira et al. 2023] (Tabela 4). Com a eliminação de fortes correlações entre as características, o ruído é reduzido junto ao custo computacional enquanto a eficácia do modelo é aprimorada.

Tabela 4. Comparação dos resultados

Experimento	Conjunto de Dados	Acurácia	Precisão	Recall	F1-Score	Predição
Experimento 1	CTU-13 (Captura 51)	88,89%	87,63%	88,89%	**88,24%**	48m 39s
Experimento 1	CTU-13 (Captura 51)	81,25%	86,26%	81,25%	83,61%	**49m 28s**
[Albano et al. 2023]	CTU-13 (Captura 51)	89,88%	84,94%	89,88%	87%	44m 41s
[Brito et al. 2023]	CTU-13 (Captura 51)	**97,89%**	**97,4%**	**97,9%**	N/A	29m e 51s
[Rahal et al. 2020]	CTU-13 (Captura 51)	N/A	N/A	N/A	N/A	5m e 41s
Experimento 2	CTU-13 (Captura 52)	89,64%	93,76%	89,64%	91,63%	**4m 6s**
Experimento 2	CTU-13 (Captura 52)	95,61%	**94,27%**	**95,61%**	**94,9%**	2m 10s
[Albano et al. 2023]	CTU-13 (Captura 52)	88,77%	90,13%	88,17%	89%	1m 50s
[de Neira et al. 2023]	CTU-13 (Captura 52)	**98,87%**	N/A	N/A	N/A	3m 49s

Em ambientes onde o custo de processamento não é uma limitação, algoritmos mais complexos, como redes neurais, podem alcançar acurácias ainda maiores. Mesmo nesses casos, a remoção de características não representativas contribui para melhorar a eficácia dos modelos. Além disso, é necessário diferentes capturas de tráfego de rede que incluam o tráfego de preparação de ataques para análises complementares. A rotulação do conjunto de dados é crucial para avaliar o desempenho dos modelos. Contudo, as abordagens não devem depender dos dados rotulados, já que isso pode ser inviável em cenários reais. Vários outros experimentos foram conduzidos com esses conjuntos de dados em diferentes configurações de conjuntos de treinamento e teste, utilizando outros métodos de seleção de características. Por questões de brevidade, apenas os mais relevantes foram apresentados neste trabalho. No entanto, todos os resultados estão disponíveis online[3].

[3] github.com/EoSingle/WGRS24

5. Conclusão

Ao adotar uma nova abordagem, este trabalho avança na predição de ataques DDoS. A estratégia combina a Transformação em Padrões Ordinais com a seleção de características, visando aprimorar a eficácia dos modelos de predição de ataques. A introdução de numerosos atributos do tráfego de rede resulta em um ruído significativo devido às fortes correlações entre as características. Portanto, este trabalho demonstra que a seleção das características mais relevantes remove o ruído, diminui o custo computacional e melhora o tempo de predição. Os resultados indicam que a proposta conseguiu antecipar o ataque com 49 minutos e 28 segundos de antecedência. Além disso, obteve-se uma precisão de 95,61%, ao mesmo tempo, em que o conjunto de características foi reduzido em 99,16%.

Referências

Albano, L., Borges, L., Neira, A., and Nogueira, M. (2023). Predição de ataques DDoS pela correlação de séries temporais via padrões ordinais. In *SBSeg*, pages 69–82, Porto Alegre, RS, Brasil. SBC.

Amer, M., Goldstein, M., and Abdennadher, S. (2013). Enhancing one-class support vector machines for unsupervised anomaly detection. *ACM SIGKDD*, pages 8–15.

Bandt, C. and Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. *Physical review letters*, 88(17):174102.

Borges, J. B., Medeiros, J. P., Barbosa, L. P., Ramos, H. S., and Loureiro, A. A. (2022). IoT botnet detection based on anomalies of multiscale time series dynamics. *IEEE TKDE*.

Borges, L. F., de Neira, A. B., Albano, L., and Nogueira, M. (2024). Multifaceted DDoS attack prediction by multivariate time series and ordinal patterns (to appear). In *ICC*, pages 1–6. IEEE.

Bouzoubaa, K., Taher, Y., and Nsiri, B. (2021). Predicting DOS-DDOS attacks: Review and evaluation study of feature selection methods based on wrapper process. *Int. J. Adv. Comput. Sci. Appl*, 12(5):132–145.

Brito, D., Neira, A., Borges, L., Araújo, A., and Nogueira, M. (2023). Um sistema autônomo para a predição de ataques de DDoS em redes locais e Internet. In *WGRS*, pages 29–42, Porto Alegre, RS, Brasil. SBC.

Chagas, E. T., Borges, J. B., and Ramos, H. S. (2022). Uso de padrões ordinais na caracterização e análise de ataques de botnets em Internet das coisas (IoT). In *WebMedia*, pages 133–137. SBC.

Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *KDD '16*, pages 785–794.

Dayal, N. and Srivastava, S. (2017). Analyzing behavior of DDoS attacks to identify DDoS detection features in SDN. In *COMSNETS*, pages 274–281.

de Araujo, P. H. H. N., Silva, A., Junior, N. F., Cabrini, F., Santiago, A., Guelfi, A., and Kofuji, S. (2021). Impact of feature selection methods on the classification of DDoS attacks using XGBoost. In *JCIS*, pages 200–214.

de Neira, A. B., de Araujo, A. M., and Nogueira, M. (2023). An intelligent system for DDoS attack prediction based on early warning signals. *IEEE TNSM*, 20(2):1–13.

Feng, Y., Akiyama, H., Lu, L., and Sakurai, K. (2018). Feature selection for machine learning-based early detection of distributed cyber attacks. In *DASC/PiCom/DataCom/CyberSciTech*, pages 173–180.

Garcia, S., Grill, M., Stiborek, J., and Zunino, A. (2014). An empirical comparison of botnet detection methods. *C&S*, 45:100–123.

Griffioen, H., Oosthoek, K., van der Knaap, P., and Doerr, C. (2021). Scan, test, execute: Adversarial tactics in amplification DDoS attacks. In *ACM SIGSAC*, pages 940–954.

Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182.

Hasan, M. A. M., Nasser, M., Ahmad, S., and Molla, K. I. (2016). Feature selection for intrusion detection using random forest. *JIS*, 7(3):129–140.

Jyoti, N. and Behal, S. (2021). A meta-evaluation of machine learning techniques for detection of DDoS attacks. In *INDIACom*, pages 522–526, India. IEEE.

Kathirgamanathan, B. and Cunningham, P. (2020). A feature selection method for multi-dimension time-series data. In *AALTD*, pages 220–231, Cham. Springer International Publishing.

Kiner, E. and April, T. (2023). Google mitigated the largest DDoS attack to date, peaking above 398 million rps. *Google Cloud Blog*.

Osanaiye, O., Cai, H., Choo, K.-K. R., Dehghantanha, A., Xu, Z., and Dlodlo, M. (2016). Ensemble-based multi-filter feature selection method for DDoS detection in cloud computing. *EURASIP*, 2016:1–10.

Peng, T., Leckie, C., and Ramamohanarao, K. (2007). Survey of network-based defense mechanisms countering the DoS and DDoS problems. *CSUR*, 39(1):3–es.

Polat, H., Polat, O., and Cetin, A. (2020). Detecting DDoS attacks in software-defined networks through feature selection methods and machine learning models. *Sustainability*, 12(3):1035.

Post, M. J., Van Der Putten, P., and Van Rijn, J. N. (2016). Does feature selection improve classification? A large scale experiment in OpenML. In *IDA 2016*, pages 158–170. Springer.

Rahal, B. M., Santos, A., and Nogueira, M. (2020). A distributed architecture for DDoS prediction and bot detection. *IEEE Access*, 8:159756–159772.

Ribeiro, H. V., Jauregui, M., Zunino, L., and Lenzi, E. K. (2017). Characterizing time series via complexity-entropy curves. *Physical Review E*, 95(6):062106.

Rosso, O. A., Carpi, L. C., Saco, P. M., Gómez Ravetti, M., Plastino, A., and Larrondo, H. A. (2012). Causality and the entropy–complexity plane: Robustness and missing ordinal patterns. *Physica A: Statistical Mechanics and its Applications*, 391(1):42–55.

Silva, G. L. F. E., de Neira, A. B., and Nogueira, M. (2022). A deep learning-based system for DDoS attack anticipation. In *LATINCOM*, pages 1–6. IEEE.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *JRSSSB*, 58(1):267–288.

Urbanowicz, R. J., Olson, R. S., Schmitt, P., Meeker, M., and Moore, J. H. (2018). Benchmarking relief-based feature selection methods for bioinformatics data mining. *Journal of biomedical informatics*, 85:168–188.

Yoachimik, O. and Pacheco, J. (2024). DDoS threat report for 2023 Q4. *Cloudflare Blog*.

Yuan, X., Li, C., and Li, X. (2017). Deepdefense: Identifying DDoS attack via deep learning. In *SMARTCOMP*, pages 1–8.

Zhou, L., Zhu, Y., Zong, T., and Xiang, Y. (2022). A feature selection-based method for DDoS attack flow classification. *Future Generation Computer Systems*, 132:67–79.