FELIPE SOUSA DE OLIVEIRA

USO DE IA GENERATIVA PARA ANÁLISE DE TRÁFEGO DE REDE LOCAL

Castanhal

2025

FELIPE SOUSA DE OLIVEIRA

USO DE IA GENERATIVA PARA ANÁLISE DE TRÁFEGO DE REDE LOCAL

Trabalho de Conclusão de Curso apresentado ao curso de Engenharia de Computação da Universidade Federal do Pará - Campus Castanhal, para a obtenção do título de Bacharel em 2025.

Orientador: Prof. Dr. José Jailton Júnior

Castanhal

2025

USO DE IA GENERATIVA PARA ANÁLISE DE TRÁFEGO DE REDE LOCAL

FELIPE SOUSA DE OLIVEIRA[1]

[1]Universidade Federal do Pará- (UFPA)
Av. dos Universitários, s/n - Jaderlândia, Castanhal - PA, 68746-630, Brasil
facompcastanhal@ufpa.br

Resumo. *Este trabalho investiga a aplicação de modelos de Inteligência Artificial Generativa (IAG) na análise de tráfego de rede local, com o objetivo de identificar comportamentos anormais durante o uso de serviços convencionais. A pesquisa envolve a coleta e o tratamento de dados de tráfego simulados, utilizando serviços diversos para representar cenários reais. Em seguida, os dados são analisados por diferentes modelos de IAG, como ChatGPT, DeepSeek e Gemini, com o intuito de avaliar sua eficácia na detecção de anomalias. A proposta visa explorar o uso acessível dessas ferramentas por usuários com conhecimentos básicos em redes, sem depender exclusivamente de especialistas. Os resultados buscam indicar qual modelo apresenta melhor desempenho na tarefa de identificação de tráfego malicioso, contribuindo para soluções práticas e acessíveis no campo da segurança de redes.*

Abstract. *This work investigates the application of Generative Artificial Intelligence (GAI) models in local network traffic analysis, aiming to identify abnormal behavior during the use of conventional services. The research involves collecting and processing simulated traffic data using various services to reflect real-world scenarios. The data is then analyzed using different GAI models, such as ChatGPT, DeepSeek, and Gemini, in order to assess their effectiveness in anomaly detection. The approach explores the accessibility of these tools for users with basic networking knowledge, without relying exclusively on cybersecurity specialists. The results aim to indicate which model performs best in identifying malicious traffic, contributing to practical and accessible solutions in the field of network security.*

1. Introdução

A segurança de redes é um dos desafios mais críticos no mundo da tecnologia da informação. Com o aumento exponencial das ameaças cibernéticas, a necessidade de soluções eficazes para a detecção de intrusos tornou-se fundamental. Tradicionalmente, os Sistemas de Detecção de Intrusão (IDS) baseiam-se em métodos de assinaturas ou regras predefinidas, mas essas abordagens podem enfrentar limitações na identificação de novas ameaças ou variantes de ataques existentes.

Nesse contexto, a ascensão das Inteligências Artificiais generativas trouxe novas ferramentas e abordagens para reforçar as defesas cibernéticas. Na segurança cibernética, o poder da IA generativa tem duas vertentes: É uma ferramenta poderosa para aqueles que cometem crimes cibernéticos e uma ferramenta igualmente poderosa para as equipes de segurança cibernética responsáveis pela prevenção e mitigação do risco de crimes cibernéticos (PALO ALTO NETWORKS, 2024). Utilizando aprendizado de máquina (ML) e redes neurais avançadas, essas IAs podem processar enormes quantidades de dados, identificar padrões ocultos, gerar insights e até automatizar respostas a incidentes. (CECYBER, 2024)

O Brasil ocupa o segundo lugar no ranking global de países mais afetados por ataques cibernéticos. Segundo o Panorama de Ameaças para a América Latina 2024, foram registrados mais de 700 milhões de ataques no país em apenas 12 meses, o que evidencia a urgência de soluções adaptativas (CNN BRASIL, 2024). Além disso, a especialista em cibersegurança Chelsea Jarvie alerta que a sofisticação dos ataques, impulsionada também pelo uso da IA por criminosos, aumenta o desafio de proteger redes de forma eficaz. Golpes como o vishing, por exemplo, utilizam IA generativa para simular vozes de pessoas conhecidas, ampliando o poder de persuasão dos atacantes.

2. Justificativa

A análise de tráfego em redes de computadores tradicionalmente requer um conhecimento técnico especializado, envolvendo o uso de ferramentas específicas para a captura e interpretação dos dados transmitidos. Essa prática demanda experiência para identificar possíveis indícios de tráfego malicioso, como comportamentos anômalos ou ataques de negação de serviço (DDoS), o que torna o processo inacessível à maioria dos usuários comuns.

Os ataques DDoS, por exemplo, podem ser difíceis de identificar devido a capacidade imitar problemas de serviço convencionais e são cada vez mais sofisticados (Kaspersky, 2025). No entanto, há certos sinais que podem sugerir que um sistema ou rede foi vítima desse tipo de ataque. Embora existam situações mais complexas que demandem

uma intervenção profissional, muitos ataques do tipo DDoS apresentam características detectáveis em um estágio inicial por meio da análise de volume, frequência ou repetição de pacotes. Neste contexto, a atuação proativa do próprio usuário pode ser decisiva para mitigar impactos em tempo hábil.

Com o avanço das tecnologias baseadas em Inteligência Artificial Generativa (IAG), surge a possibilidade de democratizar esse tipo de análise. Ferramentas como os grandes modelos de linguagem (LLMs), que podem ser utilizados via interfaces acessíveis, permitem interpretar padrões presentes nos dados coletados e apontar eventuais inconsistências ou anomalias no tráfego de rede. Isso possibilita que usuários com conhecimentos básicos em redes locais, munidos de ferramentas simples de captura de pacotes, possam contar com o apoio da IA para identificar ameaças e adotar medidas corretivas iniciais, mesmo sem o suporte imediato de um especialista em segurança da informação.

Diante disso, este trabalho justifica-se pela proposta de avaliar a eficácia de diferentes modelos de Inteligência Artificial Generativa na análise de tráfego de rede local para detecção de anomalias que possam indicar ataques DDoS. Ao explorar essa aplicação, espera-se contribuir para o desenvolvimento de soluções mais acessíveis e eficazes, ampliando a capacidade de defesa até mesmo em ambientes com poucos recursos técnicos.

3. Objetivo Geral

Diante do aumento dos ataques DDoS e da necessidade de soluções acessíveis para análise de tráfego em redes locais, este trabalho tem como objetivo principal analisar a aplicabilidade da Inteligência Artificial Generativa na identificação de anomalias relacionadas a ataques DDoS em redes locais.

4. Objetivos Específicos

Realizar a coleta de dados do tráfego de uma rede local, simulando o uso simultâneo de diversos serviços, e proceder com a extração e o tratamento dos dados para análise por meio de modelos de IA; Utilizar modelos de Inteligência Artificial Generativa, como ChatGPT, DeepSeek e Gemini, para analisar os dados de tráfego e identificar possíveis padrões anômalos; Avaliar a eficácia comparativa entre os diferentes modelos utilizados, com o intuito de determinar aquele que apresenta melhor desempenho na detecção de comportamentos suspeitos; Discutir os desafios técnicos e operacionais relacionados à implementação de soluções baseadas em IA generativa em ambientes de rede reais. Essa análise visa não apenas fortalecer o debate acadêmico sobre o uso da IA generativa na cibersegurança, mas também fornecer subsídios práticos para a aplicação de ferramentas

acessíveis por usuários não especialistas na identificação de tráfego malicioso em redes locais.

5. Referêncial Teórico

O referencial teórico apresenta os fundamentos sobre LLMs, com foco em seu funcionamento e aplicações. São destacadas suas capacidades na geração e interpretação de texto, além de limitações e desafios éticos. O conteúdo serve de base para compreender o uso dessas ferramentas na análise de tráfego de rede.

5.1. Large Language Models

Os Modelos de Linguagem de Grande Escala são um tipo de modelo de Inteligência Artificial criado para entender e gerar texto. Esses modelos são treinados em grandes volumes de dados da internet, aprendendo padrões sobre como as palavras e frases são comumente usadas juntas. Quando alimentado com uma nova entrada de texto, este tentará prever ou gerar a continuação mais provável desse texto com base no que aprendeu durante o treinamento. Embora esses modelos já existam há algum tempo, ganharam a mídia através do ChatGPT, interface de chat para modelos LLM GPT-3 e GPT-4 (DATA SCIENCE ACADEMY, 2023).

Os LLMs são modelos de aprendizado de máquina (Machine Learning) que usam algoritmos de aprendizado profundo (Deep Learning) para processar e entender a linguagem natural. Estes podem realizar muitos tipos de tarefas de linguagem, como tradução de idiomas, análise de sentimentos, conversas de chatbot. Podem também entender dados textuais complexos, identificar entidades e relacionamentos entre eles e gerar um novo texto coerente e gramaticalmente preciso. Eles são frequentemente usados para tarefas como responder perguntas, escrever redações, resumir documentos, gerar código em linguagem de programação e muito mais. (DATA SCIENCE ACADEMY, 2023).

Esses modelos têm visto uma série de avanços significativos nos últimos anos. Por exemplo, o GPT-3 da OpenAI, lançado em 2020, tem 175 bilhões de parâmetros e ficou famoso ao gerar texto preciso a partir de entradas feitas no ChatGPT. Outras melhorias incluem avanços na compreensão de contexto de longo alcance, a capacidade de gerar respostas mais coerentes e relevantes e a capacidade de entender e responder a uma variedade maior de entradas de texto (DATA SCIENCE ACADEMY, 2023).

No entanto, também há várias questões éticas associadas ao uso de LLMs. Por exemplo, devido à natureza do treinamento, eles podem refletir e perpetuar os preconceitos presentes nos dados de treinamento. Além disso, esses modelos podem gerar informações falsas ou enganosas, pois não têm uma compreensão do mundo real e dependem apenas dos padrões que aprenderam durante essa etapa. Finalmente, há preocupações

sobre o uso potencial de LLMs para fins mal-intencionados, como a criação de textos enganosos ou difamatórios.(DATA SCIENCE ACADEMY, 2023).

Para lidar com essas questões éticas, organizações como OpenAI, Microsoft e Google estão implementando uma série de salvaguardas. Isso inclui a introdução de diretrizes rigorosas para os revisores que revisam e ajustam a saída do modelo, o desenvolvimento de tecnologias para tornar os modelos mais controláveis e a realização de pesquisas para melhorar a compreensão e a mitigação dos preconceitos do modelo (DATA SCIENCE ACADEMY, 2023). Dentre as diversas IAs existentes atualmente, foram selecionadas 3 das mais utilizadas para a execução da atividade proposta. Dentre elas estão o ChatGPT, DeepSeek e Gemini.

5.2. ChatGPT

O ChatGPT é um modelo avançado de inteligência artificial desenvolvido para gerar respostas em linguagem natural, com o objetivo de simular interações humanas com alta precisão e flexibilidade. Seu funcionamento baseia-se no uso de aprendizado por reforço com feedback humano (RLHF), técnica que permite ao sistema reconhecer padrões linguísticos e contextos a partir de grandes volumes de texto, produzindo respostas coerentes e contextualizadas (CHATGPT BRASIL, 2025). A **FIGURA 1** ilustra a tela inicial da interface do ChatGPT, destacando seu ambiente de interação textual.

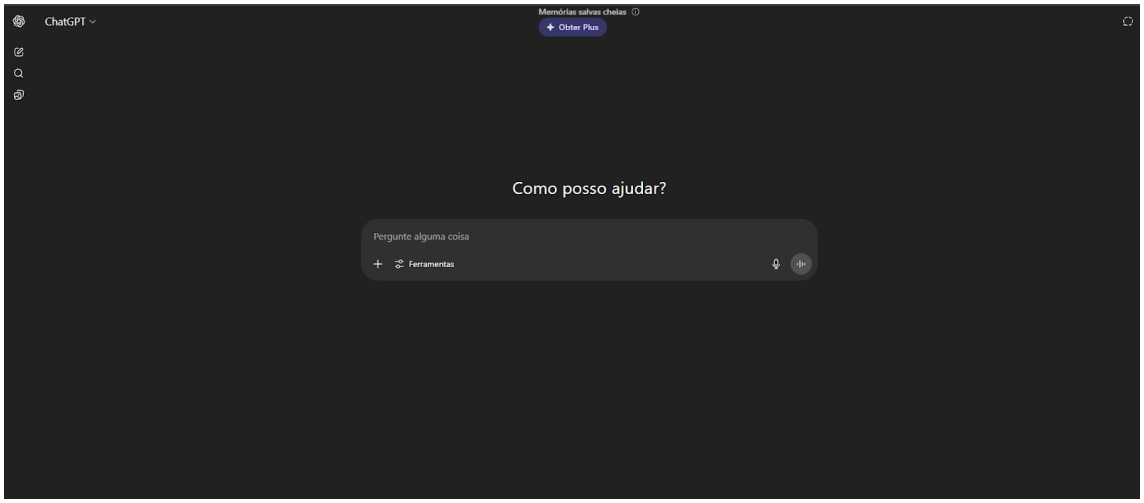

Figura 1: Tela inicial do ChatGPT

A fluidez e naturalidade das respostas do ChatGPT decorrem do seu treinamento em textos humanos de alta qualidade, o que contribui para a construção de diálogos que se assemelham à comunicação real entre pessoas. No entanto, o modelo pode ocasionalmente gerar respostas imprecisas ou irrelevantes, sendo fundamental que as informações fornecidas sejam verificadas antes de seu uso em contextos críticos ou sensíveis (CHATGPT BRASIL, 2025).

Com relação à privacidade, os dados são armazenados em conformidade com a Política de Privacidade e os Termos de Uso da plataforma. Algumas conversas podem ser analisadas para fins de aprimoramento do sistema e conformidade com políticas de segurança. Os usuários têm a opção de desativar o uso de suas conversas para fins de treinamento, por meio do portal de privacidade. Além disso, é possível solicitar a exclusão de dados conforme os procedimentos estabelecidos (CHATGPT BRASIL, 2025).

O sistema permite o acesso ao histórico de conversas, possibilitando a continuidade do diálogo em momentos distintos. Ressalta-se que, embora o modelo demonstre alto desempenho linguístico, não possui compreensão real do conteúdo gerado, o que pode levar a inconsistências ou erros factuais.

Por fim, o ChatGPT pode ser integrado a outros sistemas e plataformas através de uma interface de programação de aplicações (API), permitindo a utilização de seus recursos em diferentes contextos computacionais (CHATGPT BRASIL, 2025). Ainda assim, é essencial que os usuários mantenham uma postura crítica diante das respostas fornecidas, colaborando com feedbacks para o aprimoramento contínuo do sistema.

5.3. Gemini

O Google Gemini é uma IA Generativa que fornece respostas e realiza tarefas complexas a partir da interação do usuário. Essa interação pode ser em formato de texto, imagem ou áudio, sendo feita em prompts (comandos) que podem ser realizados no chat da ferramenta. O Google Gemini faz parte de uma família de LLMs multimodais formada também pelo Gemini Ultra, Gemini Pro, Gemini Flash e Gemini Nano. Ele é responsável por alimentar a IA Generativa de mesmo nome e foi anunciado em dezembro de 2023, posicionando-se como concorrente do GPT-4 da OpenAI (TECNOBLOG, 2024). A **FIGURA 2** apresenta a interface do Google Gemini, evidenciando sua funcionalidade multimodal.

Figura 2: Tela inicial do Gemini

O Google Gemini é usado com frequência para pesquisar sobre determinado assunto e obter uma resposta abrangente, mas a IA Generativa pode ser útil para outras tarefas como redação e programação, vendas, resumos, brainstorming e entre outros. São utilizados trilhões de parâmetros para processar vários tipos de dados simultaneamente, incluindo texto, imagem, áudio, vídeo e código de programação. O chatbot aprende padrões automaticamente através do ajuste de parâmetros durante o treinamento para interpretar as informações (TECNOBLOG, 2024).

Cada parâmetro da IA Generativa utiliza “pesos” para determinar a relação entre a entrada (prompt, comando) e a saída (resposta). A utilização de um peso é “aprendida” por meio do treinamento da IA, que é realizado através de Redes Neurais Artificiais (RNAs). As RNAs foram inspiradas no nosso cérebro e utilizam deep learning para criar várias camadas que são ajustadas por técnicas de machine learning. Criar múltiplas camadas serve para imitar a forma como o cérebro humano processa as percepções, criando, assim, uma linguagem natural (TECNOBLOG, 2024).

O Google Gemini coleta informações fornecidas pelos usuários, incluindo o uso de produtos relacionados, conversas, localização, entre outros. O objetivo é melhorar o serviço e desenvolver novos produtos para o Google, segundo a política de privacidade da empresa disponibilizada durante o aceite do uso do serviço (TECNOBLOG, 2024).

5.4. DeepSeek

O DeepSeek é um modelo de linguagem de grande escala (LLM) desenvolvido pela empresa chinesa DeepSeek.ai. Assim como outros LLMs, como o ChatGPT, o DeepSeek é capaz de gerar textos originais e criativos, traduzir idiomas, escrever diferentes tipos de conteúdo e responder a perguntas de forma informativa (DIVIA, 2024). A **FIGURA**

3 exibe a tela principal da plataforma DeepSeek, com foco na sua área de inserção de comandos.

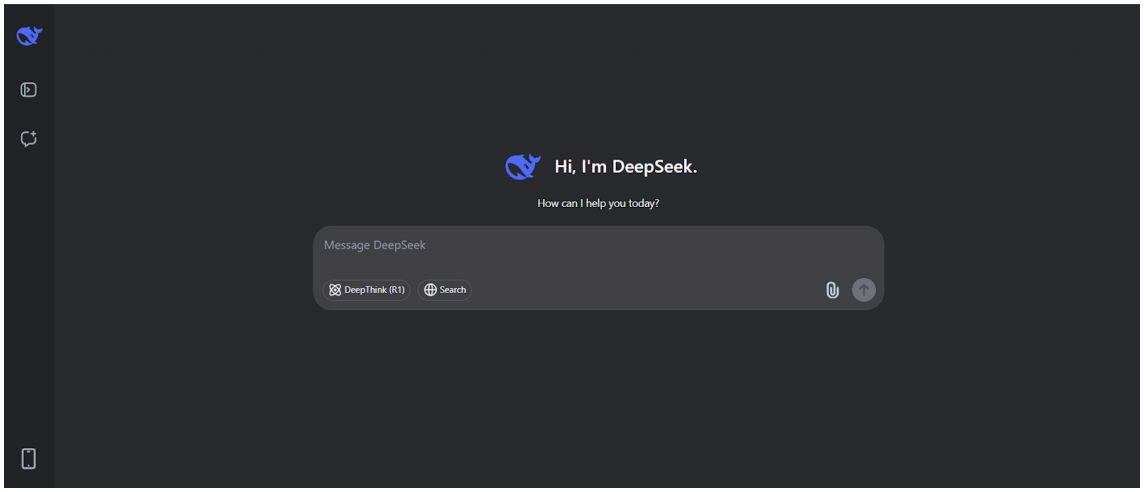

Figura 3: Tela inicial do DeepSeek

O diferencial do DeepSeek reside na sua arquitetura e filosofia. Ele utiliza uma abordagem de ”Mistura de Especialistas”(MoE), o que significa que o modelo é composto por vários ”especialistas”em diferentes áreas do conhecimento. Quando você faz uma pergunta ao DeepSeek, o sistema direciona sua solicitação para o ”especialista”mais adequado, garantindo respostas mais precisas e eficientes (DIVIA, 2024)].

A arquitetura MoE é um projeto de rede neural que melhora a eficiência e o desempenho ativando dinamicamente um subconjunto de redes especializadas, chamadas especialistas, para cada entrada. Uma rede de gating determina quais especialistas ativar, resultando em ativação esparsa e custo computacional reduzido. Essa arquitetura consiste em dois componentes essenciais: a rede de gating e os especialistas (EXXACT CORPORATION, 2024).

Em sua essência, a arquitetura MoE funciona como um sistema de tráfego eficiente, direcionando cada veículo – ou, neste caso, os dados – para a melhor rota com base nas condições em tempo real e no destino desejado. Cada tarefa é encaminhada ao especialista mais adequado, ou submodelo, especializado em lidar com aquela tarefa específica. Esse roteamento dinâmico garante que os recursos mais capacitados sejam empregados para cada tarefa, aumentando a eficiência e a eficácia geral do modelo (EXXACT CORPORATION, 2024).

Alguns especulam que o nome ”DeepSeek”pode ser uma alusão ao conceito de ”busca profunda”em áreas como mineração de dados e aprendizado de máquina, onde algoritmos complexos são usados para descobrir padrões e insights ocultos em grandes conjuntos de dados (DIVIA, 2024).

6. Metodologia

Para conduzir o estudo, foram definidas etapas bem estruturadas que envolveram a coleta, o pré-processamento e a análise dos dados de tráfego de rede. A abordagem adotada buscou simular situações reais de uso, combinando tráfego legítimo com padrões potencialmente suspeitos. Também foram aplicados prompts padronizados a diferentes ferramentas de inteligência artificial, com o objetivo de comparar o desempenho e a coerência das respostas. Todo o processo foi planejado para garantir consistência e permitir futuras reproduções do experimento.

6.1. Coleta e pré-processamento de dados

A etapa de coleta de dados foi conduzida em um ambiente de rede controlado, utilizando um host com sistema operacional Microsoft Windows 10, conectado a uma rede local (LAN) com acesso à internet. O cenário foi estruturado de forma a simular um ambiente típico de uso residencial, visando capturar tanto tráfego legítimo quanto possíveis variações anômalas. Durante o período de coleta, foram executadas atividades comuns de navegação e utilização de serviços de rede, bem como simulações de tráfego anormal para representar possíveis vetores de ataque. O objetivo foi assegurar a diversidade e a representatividade dos dados, preservando a consistência do ambiente para evitar interferências externas que comprometesse a qualidade das amostras capturadas

6.2. Ferramentas utilizadas

A principal ferramenta empregada para a captura de pacotes foi o Wireshark, um analisador de protocolos de rede amplamente reconhecido por sua eficácia na inspeção detalhada de tráfego em tempo real. O Wireshark permitiu a coleta e exportação dos dados em formato .csv, possibilitando a posterior análise fora da ferramenta. Conforme a **FIGURA 4** abaixo temos o ambiente inicial da ferramenta para realizar a coleta de dados.

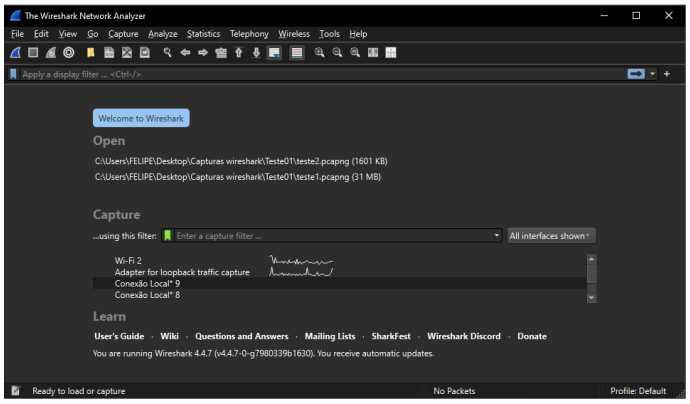

Figura 4: Tela inicial do Wireshark

Para a etapa de análise dos dados capturados, optou-se pelo uso de modelos de Inteligência Artificial Generativa, com o intuito de avaliar sua capacidade interpretativa e diagnóstico sobre o tráfego de rede. Foram utilizadas três IAs distintas: ChatGPT, DeepSeek e Gemini. Cada uma das ferramentas foi empregada com sua respectiva abordagem de análise textual e contextual, recebendo o mesmo conjunto de dados brutos extraídos do Wireshark. Embora apresentem arquiteturas diferentes, as três IAs convergiram em relação às conclusões obtidas, demonstrando consistência nos diagnósticos realizados a partir dos padrões observados no tráfego de rede.

6.3. Dados coletados

Os dados coletados consistem em pacotes de tráfego de rede capturados em ambiente real, contendo atributos relevantes para identificação de padrões e possíveis atividades anômalas. Entre os campos registrados estão os dados conforme indicados na **TABELA 1**, que apresenta os principais atributos extraídos durante a captura de tráfego.

Tabela 1: Dados coletados

Campo	Exemplo do Arquivo	Descrição
Endereço IP de Origem	35.215.245.50, 66.22.200.55	IPs que iniciam a comunicação (ex.: enviam pacotes UDP/RTCP).
Endereço IP de Destino	192.168.128.191	IP alvo da comunicação (rede interna).
Porta de Comunicação	50005 > 52941, 50006 > 49834	Portas de origem e destino (UDP). Portas altas indicam serviços específicos.
Protocolo de Transporte	UDP, RTCP, TCP, HTTP/JSON	Protocolos identificados (predomínio de UDP).
Carimbo Temporal	0.000000, 0.005581, 0.006871	Tempo em segundos desde o início da captura.
Quantidade de Pacotes	620 pacotes UDP (65.5% do total)	Volume por protocolo (ex.: UDP dominante).
Volume de Dados	Len=147, Len=400	Tamanho dos pacotes em bytes (variação entre 8 e 500 bytes).
Flags TCP	[SYN], [ACK], [FIN, ACK] (linhas 559–561)	Flags observadas em conexões TCP (ex.: handshake e encerramento).

Para viabilizar a análise pelas IAs Generativas, realizou-se unicamente a seleção de um subconjunto representativo dos dados, com o intuito de evitar sobrecarga computacional durante o processo interpretativo. Não foram aplicadas rotinas de pré-processamento, limpeza ou transformação dos dados; os arquivos foram mantidos em seu formato original, com apenas a quantidade de linhas ajustada para compatibilidade com os limites de entrada textual das ferramentas de IA utilizadas.

6.4. Procedimento de análise com LLMs

Para a realização da análise com LLMs, foram definidos procedimentos padronizados de envio e avaliação de prompts em diferentes ferramentas de IA generativa. O processo incluiu a preparação dos dados capturados, adaptação do conteúdo técnico ao formato de entrada textual e interpretação comparativa das respostas geradas.

6.4.1. Exemplo dos prompts utilizados

Durante a análise do tráfego de rede, foram utilizados dois prompts padronizados enviados às três ferramentas de Inteligência Artificial Generativa selecionadas (ChatGPT, DeepSeek e Gemini), sendo eles: "1. Esse padrão de tráfego representa um ataque DDoS ou uma conexão legítima?" e "2. Você consegue descrever os tipos de dados capturados?".

A escolha por perguntas diretas e objetivas visou avaliar a capacidade dos modelos em interpretar dados técnicos e emitir diagnósticos precisos com base em registros reais extraídos via Wireshark. Além disso, esses prompts permitiram uma avaliação qualitativa da coerência e consistência entre os modelos, pois todos receberam exatamente as mesmas informações de entrada.

O uso desse tipo de abordagem evidencia o potencial das IAs generativas convencionais como ferramentas acessíveis e eficazes para a análise de tráfego de rede, viabilizando que um indivíduo com conhecimentos básicos em redes e tráfego de pacotes possa interpretar e compreender os resultados gerados, mesmo sem a necessidade de desenvolver algoritmos próprios ou dominar técnicas avançadas de aprendizado de máquina. Assim, além de promover agilidade, essa metodologia reduz significativamente as barreiras técnicas para aplicações iniciais em cibersegurança.

7. Resultados e discussões

Nesta seção, são apresentados os resultados obtidos com a análise do tráfego de rede e as interpretações feitas pelas ferramentas de inteligência artificial utilizadas. O objetivo é mostrar como cada uma respondeu aos prompts e lidou com os dados capturados. A seguir, são destacados os principais padrões encontrados e como cada IA interpretou essas informações.

7.1. Principais padrões identificados por cada ferramenta

A análise do arquivo revelou padrões de tráfego distintos, destacando-se o domínio de pacotes UDP, que representaram 65,5% do total de tráfego capturado. Dentro desse protocolo, observou-se um alto volume de comunicações originadas de dois endereços IP

específicos: 35.215.245.50 e 66.22.200.55. O primeiro enviou 210 pacotes UDP marcados como HiPerConTracer, um protocolo não padrão que apresentou características incomuns, como valores fixos de TTL (Time to Live) (87 e 224) e um campo ”Round”com incrementos sequenciais (ex.: Round=39, 40). Esses pacotes variavam em tamanho, sendo que 47% tinham comprimento fixo de 47 bytes, enquanto outros 53% ultrapassavam 200 bytes, sugerindo possíveis testes de desempenho ou sondagem de rede.

Além disso, o IP 66.22.200.55 enviou 185 pacotes UDP para a porta 49834, com tamanhos entre 300 e 500 bytes, em intervalos curtos (média de um pacote a cada 0,1 segundo). Esse comportamento pode estar associado a streaming de mídia em tempo real (como VoIP ou videoconferência), mas também é compatível com tentativas de inundação UDP (DDoS) devido ao alto volume e tamanho dos pacotes.

Outro padrão relevante foi a presença de tráfego RTCP (Real-Time Control Protocol), utilizado para controle de fluxo em aplicações multimídia. Foram identificados 12 pacotes ”Sender Report”e 10 pacotes ”Receiver Report”, além de um pacote malformado (linha 347), que pode indicar erro na implementação ou manipulação maliciosa. Por fim, o tráfego legítimo—como HTTP/JSON (0,8%) e TCP/TLS (11,8%)—foi minoritário, concentrando-se em comunicações web seguras (ex.: solicitações POST em JSON) e conexões criptografadas (TLSv1.2).

7.1.1. HiperContracer como protocolo não padrão

A detecção do protocolo HiPerConTracer, classificado como desconhecido, com parâmetros repetitivos como Time-To-Live (TTL) fixo e campos ”Round”padronizados, sugere uma atividade anômala dentro do tráfego analisado. Em ambientes corporativos, a presença de protocolos não documentados publicamente pode indicar o uso de ferramentas de diagnóstico não autorizadas ou até sondagens maliciosas destinadas à exploração de vulnerabilidades, como mapeamentos de rede baseados em TTL. A ausência de documentação pública sobre o HiPerConTracer reforça a suspeita de que sua utilização não se alinha a práticas legítimas comuns em redes empresariais.

7.1.2. Alto volume de UDP com padrão de inundação

O endereço IP 66.22.200.55 apresentou um padrão de envio de pacotes UDP de tamanho elevado, chegando a 500 bytes, em frequência intensiva, o que caracteriza um comportamento típico de ataques de amplificação UDP, nos quais serviços como DNS e NTP podem ser utilizados como vetores de reflexão. Embora esse padrão possa ocasionalmente refletir tráfego legítimo, como transmissões de streaming, o volume observado está fora do

esperado para redes convencionais. A ausência de correlação com protocolos amplamente utilizados, como o RTP (empregado em VoIP), reforça a necessidade de aprofundamento na investigação sobre a origem e finalidade desse tráfego.

7.1.3. Pacotes RTCP mal formados

A detecção de pacotes RTCP com estrutura malformada indica a ocorrência de possíveis anomalias, que podem ser resultado de corrupção de dados durante a transmissão ou, em um cenário mais crítico, de tentativas intencionais de manipulação do fluxo multimídia. Esse comportamento pode estar relacionado a ataques que visam comprometer a qualidade e a estabilidade de serviços de comunicação, como chamadas VoIP, representando um risco relevante à integridade dos sistemas de mídia em tempo real.

7.1.4. Disparidade entre tráfego legítimo e suspeito

A análise do tráfego revela uma discrepância significativa entre o volume de dados associados a protocolos legítimos e suspeitos: apenas 12,6% do tráfego registrado corresponde a HTTP/HTTPS, enquanto pacotes UDP representam 65,5% do total. Essa assimetria é preocupante, principalmente em redes onde se espera predominância de tráfego web tradicional, e pode ser um indicativo de atividades não usuais que requerem investigação, como ataques ou uso de aplicações não autorizadas.

7.2. Comparativo entre as respostas obtidas (em Conteúdo e formato)

As respostas geradas pelas três ferramentas de IA analisadas apresentaram semelhanças em alguns pontos, mas também revelaram diferenças importantes na forma como interpretaram os dados. Cada modelo adotou uma abordagem própria, variando no nível de detalhe, na linguagem utilizada e na clareza das recomendações. A seguir, é feito um comparativo entre os conteúdos e formatos das respostas, destacando os pontos fortes e limitações de cada uma.

7.2.1. Análise de DDoS vs Conexão legítima

As ferramentas de análise utilizadas compartilham pontos em comum, como a identificação do tráfego UDP como atípico e a menção recorrente aos IPs 35.215.245.50 e 66.22.200.55 como fontes potencialmente suspeitas. No entanto, há divergências interpretativas importantes: o ChatGPT é a única ferramenta que conclui de forma categórica tratar-se de um ataque DDoS, enquanto DeepSeek e Gemini adotam uma postura mais cautelosa, sugerindo que o protocolo HiPerConTracer pode, eventualmente, ter finalidade

legítima como ferramenta de diagnóstico ou monitoramento. A **TABELA 2** sintetiza essas interpretações comparativas entre as ferramentas analisadas.

Tabela 2: Análise dos dados

IA	Conclusão Principal	Argumentos-Chave	Diferenciais
DeepSeek	"Não é um ataque DDoS típico, mas há tráfego suspeito."	- HTTP/JSON e TLS são legítimos. - UDP/RTCP pode ser legítimo (streaming) ou suspeito (varredura).	Foco em métricas (65.5% UDP) e análise de IPs específicos (35.215.245.50).
ChatGPT	"Fortes indícios de DDoS."	- Pacotes UDP rápidos e constantes. - Mesmo IP de origem atacando repetidamente.	Menos nuances; interpreta HiPerConTracer como malicioso sem explorar alternativas.
Gemini	"Difícil afirmar, mas provavelmente legítimo (serviço HiPerConTracer)."	- Baixa diversidade de fontes (não típico de DDoS). - Serviço específico pode justificar o tráfego.	Ênfase no contexto ausente e análise de tamanho de pacotes/variedade.

7.2.2. Descrição dos dados capturados

No que se refere à abordagem descritiva dos dados, o DeepSeek destaca-se pela abrangência na análise de protocolos e pela inclusão de representações visuais, proporcionando uma visão mais detalhada e acessível. O ChatGPT, por outro lado, prioriza a aplicação prática dos dados, oferecendo scripts e recursos para automação da análise. Já o Gemini apresenta foco técnico nas camadas inferiores do modelo OSI, como Ethernet e IPv4, porém carece de maior concretude na interpretação contextual dos dados capturados. A **TABELA 3** apresenta um resumo das características principais observadas em cada ferramenta quanto à descrição dos dados capturados.

Tabela 3: Descrição dos dados capturados

IA	Protocolos Identificados	Destaques	Formato de Resposta
DeepSeek	UDP, TCP/TLS, HTTP/JSON, ARP, RTCP, SSDP, ICMPv6.	Detalha porcentagens (ex: 65.5% UDP) e padrões como "HiPerConTracer".	Lista organizada com métricas e gráficos sugeridos.
ChatGPT	UDP, HiPerConTracer (ênfase).	Foca em campos do CSV (Time, Source, Protocol) e padrões de ataque.	Descrição textual simples + código Python para análise.
Gemini	UDP, HiPerConTracer, cabeçalhos Ethernet/IPv4.	Analisa tamanho de pacotes (8 bytes a 400 bytes) e fluxos de comunicação (portas 50005/52941).	Descrição conceitual detalhada, sem métricas precisas.

7.2.3. Comparativo geral

As recomendações das ferramentas de inteligência artificial divergem em escopo e profundidade. O DeepSeek propõe ações como a investigação de IPs suspeitos, o monitoramento de portas elevadas (entre 50000 e 51000) e o bloqueio do HiPerConTracer caso este não seja essencial ao ambiente. O ChatGPT, de forma implícita, sugere o bloqueio direto do tráfego proveniente dos IPs 35.215.245.50 e 66.22.200.55. Já o Gemini recomenda a obtenção de maior contexto sobre o protocolo HiPerConTracer, além da análise de janelas de tráfego mais extensas. No comparativo geral, o DeepSeek foi o mais completo tecnicamente, apresentando métricas específicas, suspeitas fundamentadas e elementos visuais. O ChatGPT se mostrou mais ágil para alertas genéricos, enquanto o Gemini trouxe uma análise contextual relevante, porém incompleta, evidenciando a necessidade de coleta adicional de dados. A **TABELA 4** consolida as recomendações finais propostas por cada ferramenta de inteligência artificial.

Tabela 4: Comparativo geral

Critério	DeepSeek	ChatGPT	Gemini
Tom da Resposta	Analítico, com conclusões cautelosas e recomendações de investigação.	Direto, com forte indicação de atividade maliciosa (DDoS).	Equilibrado, destacando a necessidade de mais contexto para conclusões.
Abordagem	Quantitativa (estatísticas, métricas) e qualitativa. .	Qualitativa, focada em padrões suspeitos.	Qualitativa, com ênfase em hipóteses alternativas (serviço legítimo).
Visualizações	Inclui gráficos sugeridos (barras, pizza, heatmap) e métricas estruturadas.	Gera código Python para análise automatizada e visualização.	Descreve visualizações conceituais, mas não as implementa.
Profundidade Técnica	Alta (detalha protocolos, portas, TTL, volumetria).	Média (foca em SYN/ACK, IPs suspeitos).	Alta (analisa cabeçalhos, padrões temporais, tamanho de pacotes).

A análise das respostas das três IAs mostra pontos fortes e fracos em cada abordagem. O ChatGPT erra ao tratar o tráfego HiPerConTracer como claramente malicioso sem considerar explicações alternativas, como ser um serviço legítimo de rede. O Gemini tem uma visão mais equilibrada, mas falha por não apresentar dados concretos ou ferramentas para comprovar suas suposições. Já o DeepSeek oferece a análise mais completa, com métricas detalhadas e recomendações técnicas, porém sua complexidade pode dificultar o entendimento para usuários menos experientes. No geral, isso revela um desafio impor-

tante no uso de IAs para análise de segurança: como equilibrar precisão técnica, clareza e consideração de diferentes possibilidades ao avaliar tráfego de rede.

8. Conclusão

Através deste estudo, foi possível examinar como os Modelos de Linguagem de Grande Escala, a exemplo do ChatGPT, DeepSeek e Gemini, operam e são empregados na avaliação do tráfego de dados em uma rede local. Fortalecidos por progressos no aprendizado profundo e em estruturas como o Transformer, eles exibem um notável desempenho na compreensão de informações textuais complicadas, na criação de respostas contextualizadas e na descoberta de padrões importantes.

A utilização efetiva de IAs generativas na avaliação do fluxo de dados da rede, principalmente na identificação de investidas como o DDoS, revelou-se encorajadora. As avaliações realizadas apontaram que esses modelos são capazes de ajudar na identificação de atitudes fora do comum, oferecendo percepções valiosas e acelerando o processo de análise. Isso pode simbolizar uma alteração considerável nas formas tradicionais de acompanhamento, concedendo mais rapidez e exatidão aos grupos de segurança da informação.

Entretanto, também foi possível identificar limitações importantes. O uso de LLMs depende da qualidade e diversidade dos dados com os quais foram treinados, o que pode introduzir vieses ou limitações em determinados contextos técnicos. Além disso, como não possuem uma compreensão real do mundo, essas ferramentas podem apresentar falhas em interpretações específicas ou gerar respostas imprecisas. Questões éticas, como privacidade e uso indevido das respostas geradas, também precisam ser consideradas com cautela.

Conclui-se, portanto, que o uso de IAs generativas na área de redes e segurança apresenta uma alternativa viável e inovadora, com capacidade real de complementar análises técnicas. Contudo, sua adoção deve ser acompanhada de responsabilidade, validação humana e contínuo aprimoramento, garantindo que sua aplicação seja segura, eficaz e eticamente adequada.

Referências

CECYBER. *A importância das IAs generativas na análise de vulnerabilidades e resposta a incidentes*. 2024. ⟨https://cecyber.com/blog/a-importancia-das-ias-generativas-na-analise-de-vulnerabilidades-e-resposta-a-incidentes/⟩. Acesso em: 01 jul. 2025.

CHATGPT BRASIL. *FAQ - O que é o ChatGPT?* 2025. ⟨https://chatgpt.com.br/faq/#o-que-e-chatgpt⟩. Acesso em: 15 jun. 2025.

CNN BRASIL. *Brasil é vice-campeão em ataques cibernéticos, com 1.379 golpes por minuto, aponta estudo*. 2024. ⟨https://www.cnnbrasil.com.br/economia/negocios/brasil-e-vice-campeao-em-ataques-ciberneticos-com-1-379-golpes-por-minuto-aponta-estudo⟩. Acesso em: 29 abr. 2025.

DATA SCIENCE ACADEMY. *O que são Large Language Models (LLMs)?* 2023. ⟨https://blog.dsacademy.com.br/o-que-sao-large-language-models-llms/⟩. Acesso em: 15 jun. 2025.

DIVIA. *DeepSeek: o que é, como funciona e como usar*. 2024. ⟨https://www.divia.com.br/deepseek-o-que-e-como-funciona-e-como-usar⟩. Acesso em: 20 jun. 2025.

EXXACT CORPORATION. *Why New LLMs Use MoE: Mixture of Experts Architecture*. 2024. ⟨https://www.exxactcorp.com/blog/deep-learning/why-new-llms-use-moe-mixture-of-experts-architecture⟩. Acesso em: 20 jun. 2025.

Kaspersky. *O que é um ataque DDoS?* 2025. ⟨https://www.kaspersky.com.br/resource-center/threats/ddos-attacks⟩. Acesso em: 01 jul. 2025.

PALO ALTO NETWORKS. *Generative AI in Cybersecurity*. 2024. ⟨https://www.paloaltonetworks.com.br/cyberpedia/generative-ai-in-cybersecurity⟩. Acesso em: 01 jul. 2025.

TECNOBLOG. *O que é o Google Gemini: entenda para que serve e como funciona a IA do Google*. 2024. ⟨https://tecnoblog.net/responde/o-que-e-o-google-gemini-entenda-para-que-serve-e-como-funciona-a-ia-do-google/⟩. Acesso em: 18 jun. 2025.