v7.n8.2026

REVISTA CIENTÍFICA - RECIMA21 ISSN 2675-6218

ANÁLISE PREDITIVA DE VULNERABILIDADES EM APLICAÇÕES WEB USANDO LOGS DE SERVIDOR: UMA ARQUITETURA HÍBRIDA

PREDICTIVE ANALYSIS OF VULNERABILITIES IN WEB APPLICATIONS USING SERVER LOGS: A HYBRID ARCHITECTURE

ANÁLISIS PREDICTIVO DE VULNERABILIDADES EN APLICACIONES WEB USANDO LOGS DE SERVIDOR: UNA ARQUITECTURA HÍBRIDA

Andrey Rodrigues Moreira[1], Olavo Olimpio de Matos Junior[2]

e788820
https://doi.org/10.47820/recima21.v7i8.8820

PUBLICADO: 08/2026

RESUMO

A crescente sofisticação das ameaças cibernéticas a aplicações *web* exige estratégias defensivas proativas e em profundidade. Sistemas de Gerenciamento de Eventos e Informações de Segurança (SIEM) baseados em assinaturas apresentam limitações contra-ataques inéditos (*Zero-Day*), enquanto Inteligência Artificial (IA) comportamental isolada gera elevados falsos positivos e fadiga de alertas. Este estudo propõe e valida a arquitetura híbrida Aegis, integrando a análise determinística do SIEM Wazuh à capacidade preditiva do algoritmo não supervisionado *Isolation Forest*. A metodologia combinou um experimento em ambiente conteinerizado via *Docker* (13.479 requisições HTTP) com engenharia de atributos estatísticos e volumétricos dos *endpoints*, além da validação externa no *benchmark* CIC-IDS2017 (170.366 requisições de rede e aplicação). No ambiente conteinerizado, a orquestração reduziu os pontos cegos mútuos, atingindo 100,0% de detecção. Na validação externa, a IA isolada alcançou *recall* proativo de 91,97%, porém com baixa precisão (5,39%). A arquitetura híbrida Aegis mitigou esse ruído operacional, elevando a precisão para 70,27%, a acurácia para 99,30% e o *F1-Score* para 73,86%. Conclui-se que a convergência entre aprendizado não supervisionado e SIEM maximiza a eficácia preditiva e operacional, consolidando uma camada defensiva viável para Centros de Operações de Segurança (SOC).

PALAVRAS-CHAVE: Cibersegurança. SIEM. Inteligência Artificial. *Isolation Forest.* Aprendizado de Máquina. Detecção de Intrusão.

ABSTRACT

The growing sophistication of cyber threats targeting web applications demands proactive, defense-in-depth strategies. Signature-based Security Information and Event Management (SIEM) systems present limitations against novel (Zero-Day) attacks, while isolated behavioral Artificial Intelligence (AI) generates high false positive rates and alert fatigue. This study proposes and validates the Aegis hybrid architecture, integrating the deterministic analysis of the Wazuh SIEM with the predictive capability of the unsupervised Isolation Forest algorithm. The methodology combined a containerized Docker experiment (13,479 HTTP requests) with statistical and volumetric endpoint feature engineering, alongside external validation on the CIC-IDS2017 benchmark (170,366 network and application requests). In the containerized environment, orchestration eliminated mutual blind spots, reaching 100.0% detection. In external validation, pure AI achieved a proactive recall of 91.97%, but with low precision (5.39%).

[1] Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP), graduando em Análise e Desenvolvimento de Sistemas.
[2] Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP), mestre em Informática.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

The Aegis hybrid architecture successfully mitigated this operational noise, raising precision to 70.27%, accuracy to 99.30%, and the F1-Score to 73.86%. It is concluded that the convergence between unsupervised learning and SIEM maximizes predictive and operational effectiveness, consolidating a viable defense layer for Security Operations Centers (SOC).

KEYWORDS: *Cybersecurity. Isolation Forest. SIEM. Artificial Intelligence. Machine Learning. Intrusion Detection.*

RESUMEN
La creciente sofisticación de las amenazas cibernéticas a las aplicaciones web exige estrategias defensivas proactivas y en profundidad. Los sistemas de Gestión de Información y Eventos de Seguridad (SIEM) basados en firmas presentan limitaciones contra ataques inéditos (Zero-Day), mientras que la Inteligencia Artificial (IA) conductual aislada genera elevados falsos positivos y fatiga de alertas. Este estudio propone y valida la arquitectura híbrida Aegis, integrando el análisis determinista del SIEM Wazuh con la capacidad predictiva del algoritmo no supervisado Isolation Forest. La metodología combinó un experimento en un entorno contenedorizado vía Docker (13.479 peticiones HTTP) con ingeniería de atributos estadísticos y volumétricos de los endpoints, además de una evaluación externa en el benchmark CIC-IDS2017 (170.366 peticiones de red y aplicación). En el entorno contenedorizado, la orquestación redujo los puntos ciegos mutuos, alcanzando el 100,0% de detección. En la validación externa, la IA aislada alcanzó un recall proactivo del 91,97%, aunque con baja precisión (5,39%). La arquitectura híbrida Aegis mitigó este ruido operativo, elevando la precisión al 70,27%, la exactitud al 99,30% y el F1-Score al 73,86%. Se concluye que la convergencia entre el aprendizaje no supervisado y el SIEM maximiza la eficacia predictiva y operativa, consolidando una capa defensiva viable para Centros de Operaciones de Seguridad (SOC).

PALABRAS CLAVE*: Ciberseguridad. Isolation Forest. SIEM. Inteligencia Artificial. Aprendizaje de Máquina. Detección de Intrusión.*

INTRODUÇÃO

A constante evolução e a crescente sofisticação das ameaças cibernéticas direcionadas a aplicações *web* transformaram a segurança digital em um desafio crítico e dinâmico para as organizações. Os sistemas tradicionais de monitoramento, baseados em soluções de *Security Information and Event Management* (SIEM), operam primordialmente por meio de correlações textuais e assinaturas estáticas predefinidas (TARIQ *et al.*, 2023). Embora essa abordagem garanta alta precisão contra vetores de ataque amplamente conhecidos, ela manifesta uma limitação estrutural severa ao enfrentar ameaças inéditas (*Zero-Day*) ou variantes customizadas projetadas para transpor filtros de segurança específicos (TARIQ *et al.*, 2023). Nesse cenário, a literatura técnica aponta para a imperativa evolução dos SIEMs reativos em direção a arquiteturas preditivas integradas a modelos de Inteligência Artificial, capazes de unir diferentes camadas analíticas para neutralizar pontos cegos operacionais (AL-MHIQANI *et al.*, 2020).

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Para superar a rigidez dos motores de regra sem demandar a rotulagem prévia de massivos conjuntos de dados, o uso de aprendizado de máquina não supervisionado emergiu como uma abordagem fundamental para extrair anomalias estatísticas e sequenciais diretamente a partir de logs brutos e não estruturados de servidores *web* (MENG *et al.*, 2019; Benova; Hudec, 2024). Contudo, a aplicação isolada desses modelos comportamentais enfrenta um dilema operacional bem documentado na literatura: ao buscar taxas elevadas de cobertura (*Recall*) para capturar ameaças sutis, os algoritmos frequentemente geram um volume massivo de Falsos Positivos, resultando na sobrecarga das equipes de resposta a incidentes e provocando a chamada "fadiga de alertas" em Centros de Operações de Segurança (AHMAD *et al.*, 2021; Ahir; Shaikh, 2024).

Adicionalmente, a avaliação rigorosa dessas metodologias defensivas exige a submissão dos modelos a conjuntos de dados de referência (*benchmarks*) que reflitam a realidade das redes corporativas. Embora conjuntos de dados clássicos, como o CSIC 2010, tenham fundamentado historicamente a análise de tráfego HTTP, seus padrões de ataque tornaram-se defasados frente à complexidade das intrusões atuais (SHARAFALDIN; Lashkari; Ghorbani, 2018). Essa limitação impulsionou a transição para *benchmarks* contemporâneos, destacando-se o CIC-IDS2017, que incorpora cenários modernos e diversificados de ataques de Camada 7 (*SQL Injection*, *Cross-Site Scripting* e *Brute Force*) combinados a fluxos realistas de navegação legítima (SHARAFALDIN; Lashkari; Ghorbani, 2018).

Diante desse panorama, a motivação para o desenvolvimento deste projeto nasceu no âmbito das investigações do Núcleo de Sistemas do Instituto Federal de São Paulo (IFSP) – Câmpus Jacareí. A premissa central fundamenta-se na exploração dos logs de eventos de servidores *web* (como o Nginx) como fontes ricas de dados analíticos para subsidiar tomadas de decisão proativas, atuando antes que um sistema crítico seja efetivamente comprometido (AJI; Prayudi; Luthfi, 2023; Yang *et al.*, 2022). O estudo justifica-se pela proposição de uma metodologia baseada na arquitetura híbrida denominada “Aegis”, que integra a análise determinística do SIEM *open-source* Wazuh (MONTEIRO; Turci Júnior; Oliveira, 2024) à capacidade preditiva do algoritmo não supervisionado *Isolation Forest* (MYKHAYLOVA; Shtypka; Fedynyshyn, 2024). A relevância da proposta reside na superação de gargalos práticos de infraestrutura, como o mascaramento de endereços IP nativo de redes conteinerizadas (*Docker*), mediante uma engenharia de recursos puramente voltada ao comportamento estatístico e volumétrico dos *endpoints*.

O objetivo geral desta pesquisa consiste em desenvolver e validar uma metodologia sistemática de análise de logs de servidores *web* capaz de identificar, caracterizar e prever

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

padrões de tráfego malicioso associados a tentativas de exploração de vulnerabilidades, mitigando simultaneamente pontos cegos e alarmes falsos.

Para a consecução dessa meta, definiram-se os seguintes objetivos específicos:

a) Configurar um ambiente laboratorial conteinerizado controlado (*Docker/Nginx/OWASP Juice Shop*) para simular tráfego legítimo e malicioso em tempo real, gerando requisições calibradas nas proporções estatísticas do benchmark CSIC 2010 com *payloads* ruidosos do *HTTP Params Dataset* (MORADIAN, 2015), além de validar a capacidade de generalização da arquitetura no *benchmark* moderno CIC-IDS2017;

b) Estruturar um *pipeline* de extração, transformação e carga (ETL) utilizando a *stack* ELK (Logstash e Elasticsearch) aplicando engenharia de recursos comportamentais via biblioteca Pandas em Python para converter *logs* brutos em atributos quantitativos;

c) Submeter os atributos estruturados ao algoritmo *Isolation Forest* e ao motor de regras do Wazuh, avaliando empiricamente o desempenho isolado e a sinergia da arquitetura híbrida Aegis em relação ao equilíbrio entre *Recall* e Falsos Positivos;

d) Elaborar *dashboards* interativos em tempo real na plataforma Grafana para o monitoramento unificado de métricas de segurança, linhas do tempo de ataques e correlação visual de alertas;

e) Consolidar os resultados analíticos obtidos, discutir a viabilidade operacional da abordagem híbrida e formular recomendações técnicas para otimizar fluxos de trabalho em Centros de Operações de Segurança (SOC).

1. REFERENCIAL TEÓRICO

A infraestrutura de monitoramento tradicional em segurança cibernética apoia-se fortemente em soluções de *Security Information and Event Management* (SIEM) e *Host-based Intrusion Detection Systems* (HIDS). Essas plataformas desempenham um papel central na centralização, auditoria e controle de eventos em redes e servidores corporativos. O funcionamento operacional dessas ferramentas baseia-se na aplicação de motores de correlação alimentados por regras estáticas e assinaturas textuais predefinidas (TARIQ *et al.*, 2023). Soluções consagradas de código aberto (*open-source*), como o Wazuh, oferecem uma linha de base (*baseline*) sólida para a detecção de intrusões em tempo real por meio da análise contínua de logs e integridade de arquivos (MONTEIRO; Turci Júnior; Oliveira, 2024).

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Contudo, a principal limitação dos motores puramente baseados em regras reside na sua incapacidade intrínseca de identificar ameaças inéditas (*Zero-Day*) ou explorações customizadas que evadem filtros textuais conhecidos (TARIQ *et al.*, 2023). Diante dessa fragilidade, a literatura em gerenciamento de incidentes aponta para a necessária evolução dos SIEMs tradicionais em direção a arquiteturas preditivas auxiliadas por Inteligência Artificial, capazes de integrar diferentes camadas analíticas para neutralizar pontos cegos operacionais (AL-MHIQANI *et al.*, 2020).

Para preencher a lacuna das assinaturas estáticas sem exigir o dispendioso processo de rotulagem prévia de grandes volumes de dados, o uso de aprendizado de máquina não supervisionado destaca-se como a abordagem mais promissora para a extração automática de anomalias estatísticas, quantitativas e sequenciais diretamente a partir de logs brutos e não estruturados de servidores *web* (MENG *et al.*, 2019; Benova; Hudec, 2024). Ao analisar a distribuição e o comportamento histórico das requisições, os modelos comportamentais dispensam o conhecimento prévio sobre o vetor de ataque, identificando desvios de padrão que sinalizam atividades maliciosas em tempo de execução (MOURA, 2024; Trindade, 2022).

Dentre as técnicas de aprendizado não supervisionado, o algoritmo *Isolation Forest* sobressai-se pela sua eficiência computacional e capacidade de processamento em *pipelines* analíticos de alta performance (AHIR; Shaikh, 2024; Mykhaylova; Shtypka; Fedynyshyn, 2024). Diferente dos modelos de agrupamento (*clustering*) tradicionais que tentam delimitar o perfil complexo do tráfego legítimo, o *Isolation Forest* inverte a lógica computacional ao focar no isolamento direto das propriedades atípicas dos dados (MYKHAYLOVA; Shtypka; Fedynyshyn, 2024). A premissa fundamental do algoritmo assume que anomalias são instâncias raras com valores de atributos discrepantes, tornando-se matematicamente mais fáceis de serem isoladas do restante do conjunto de dados.

Matematicamente, o algoritmo constrói florestas compostas por árvores de isolamento binárias (*iTrees*) através do particionamento recursivo aleatório das variáveis do *dataset*. A identificação do tráfego malicioso ocorre de maneira geométrica: como as requisições anômalas possuem assinaturas estatísticas raras, elas demandam um número significativamente menor de divisões nas árvores para serem completamente isoladas nos nós terminais. A medida quantitativa de anomalia de uma instância (x) é calculada a partir da profundidade do caminho percorrido nas árvores. Para normalizar esse comprimento e permitir a análise estatística de uma amostra com (n) instâncias, o escore de anomalia s(x, n) é definido pela seguinte equação formal:

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}}$$

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Na formulação exposta, *h(x)* representa o comprimento do caminho da amostra *(x)* (computado pelo número de arestas que a instância percorre do nó raiz até o nó terminal), enquanto *E(h(x))* define o valor esperado ou a média de *h(x)* obtida a partir de um conjunto de árvores de isolamento estruturadas na floresta. O fator de ajuste *c(n)* representa o comprimento médio do caminho de uma busca mal sucedida em uma Árvore Binária de Busca (BST) construída com o mesmo número *n* de nós do conjunto de dados, atuando como o delimitador de normalização do modelo. O cálculo de *c(n)* é extraído formalmente através da função:

$$c(n) = 2\,ln(n-1) + 0{,}5772156649 - \frac{2(n-1)}{n}$$

onde o valor constante 0.5772156649 representa a constante de Euler-Mascheroni. A partir da estabilização do escore *s(x, n)*, a interpretação matemática do tráfego de rede passa a obedecer a três intervalos de comportamento bem definidos:

- Se o valor esperado de *E(h(x))* tende a zero, o comprimento do caminho percorrido pela requisição é extremamente curto, fazendo com que o escore *(s)* se aproxime de 1. Nesse cenário, a instância é classificada com alta confiabilidade como uma anomalia proativa, indicando um provável vetor de ataque;
- Se a média *E(h(x))* assume valores elevados e próximos a n, o caminho percorrido é longo e profundo, resultando em um escore *(s)* significativamente menor que 0.5. Esse comportamento geométrico é característico de instâncias densas que pertencem ao agrupamento de tráfego legítimo dos usuários comuns;
- Se a totalidade das amostras retornar um escore *(s)* \approx 0.5, o conjunto de dados inteiro manifesta uma distribuição homogênea, indicando a ausência de anomalias estatisticamente detectáveis na amostra analisada.

Apesar da alta sensibilidade do Isolation Forest, a literatura especializada alerta para o dilema central na detecção não supervisionada de intrusões: modelos focados em maximizar a cobertura (*Recall*) frequentemente acumulam elevado volume de Falsos Positivos ao interpretar pequenas oscilações de rede ou requisições atípicas legítimas como anomalias (AHMAD *et al.*, 2021). Em ambientes operacionais de Centros de Operações de Segurança (SOC), o excesso de alarmes falsos resulta na "fadiga de alertas", sobrecarregando os analistas de incidentes e degradando a eficiência defensiva da organização (AHMAD *et al.*, 2021; Ahir; Shaikh, 2024). Esse gargalo justifica a necessidade de arquiteturas híbridas que combinam o filtro contextualizador do SIEM à inteligência preditiva do modelo.

Por fim, a validação de novas abordagens de segurança exige a submissão dos modelos a conjuntos de dados de referência (*benchmarks*) representativos. Embora bases clássicas como

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

o CSIC 2010 tenham sustentado historicamente a análise de tráfego HTTP, elas contêm padrões de ataque ultrapassados e sintéticos que não capturam a dinâmica das redes corporativas modernas (SHARAFALDIN; Lashkari; Ghorbani, 2018). O surgimento do *benchmark* CIC-IDS2017 preencheu essa lacuna ao registrar fluxos realistas de rede e uma ampla diversidade de ataques contemporâneos de Camada 7 (*SQL Injection, Cross-Site Scripting e Brute Force*), consolidando-se como o padrão de referência para a avaliação de sistemas de detecção de intrusão modernos (SHARAFALDIN; Lashkari; Ghorbani, 2018).

2. METODOLOGIA

O fluxo de desenvolvimento e a arquitetura lógica do ecossistema híbrido Aegis foram estruturados em três fases fundamentais, conforme ilustrado na Figura 1: a preparação e coleta de dados, o processamento analítico com aprendizado de máquina e a posterior visualização operacional dos indicadores de segurança.

Figura 1. Diagrama de desenvolvimento e arquitetura lógica do sistema



Fonte: Elaborada pelos Autores (2026).

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

A primeira fase, denominada Preparação e Coleta, concentrou-se na construção de um ambiente laboratorial isolado empregando a tecnologia de contêineres *Docker*, o que garantiu a reprodutibilidade dos testes de intrusão sem interferências de variáveis externas do sistema hospedeiro. A infraestrutura foi composta por um servidor *web Nginx* configurado para atuar como *proxy* reverso, centralizando a recepção de todo o tráfego de rede direcionado à aplicação vulnerável *OWASP Juice Shop*. Como linha de base (*baseline*) para o monitoramento tradicional baseado em assinaturas de segurança, implantou-se o agente *HIDS* da plataforma Wazuh no mesmo contêiner do servidor.

Para a consolidação do conjunto de dados de laboratório, desenvolveu-se uma automação *multithread* em Python utilizando as bibliotecas *requests* e *threading*. Esse *script* foi responsável por simular o comportamento de usuários legítimos navegando concorrentemente pela aplicação, enquanto injetava simultaneamente requisições anômalas contendo *payloads* de ataques reais, como força bruta em rotas sensíveis, *Cross-Site Scripting* (XSS) e *Local File Inclusion* (LFI). Varreduras volumétricas e agressivas de mapeamento de portas e injeções estruturadas foram executadas por meio das ferramentas Nmap e SQLMap. Para estressar o *pipeline* analítico com parâmetros e URLs ruidosas de navegadores reais, incorporaram-se *payloads* do *HTTP Params Dataset* (MORADIAN, 2015). O esforço coordenado gerou um conjunto de dados inédito contendo 13.479 requisições gravadas no arquivo access.log do Nginx, calibrado unicamente com base na proporção estatística de distribuição observada no *benchmark* CSIC 2010 (62% de tráfego benigno e 38% malicioso), sem a utilização direta dos dados originais desse repositório. O gabarito oficial (*Ground Truth*) foi gerado de forma determinística pela própria automação de ataque, que vinculou os *timestamps* exatos, métodos e URLs maliciosas a cada registro capturado, criando a base de validação do modelo.

Adicionalmente, com o objetivo de expandir a validade externa da pesquisa e avaliar o desempenho da arquitetura frente a um *benchmark* contemporâneo (SHARAFALDIN; Lashkari; Ghorbani, 2018), incorporou-se um Estudo de Caso de Validação Externa utilizando o *dataset* CIC-IDS2017 (*Thursday-WorkingHours-Morning-WebAttacks*). Este conjunto de dados compreende a totalidade de 170.366 requisições de tráfego de rede e aplicação (168.186 acessos benignos e 2.180 ataques de *Web Brute Force*, *XSS* e *SQL Injection*).

Para validar simultaneamente o fluxo operacional em tempo real da infraestrutura conteinerizada, desenvolveu-se o script disparador_logs_estudo_caso.py, que transmitiu uma amostragem contínua de 5.000 requisições convertidas no padrão *Syslog/Nginx* via *socket* UDP para a porta 5000 do *Logstash*.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

No início da segunda fase, Processamento e IA, os dados coletados de forma integral avançaram para a separação e calibração estatística. Por se tratar de um modelo não supervisionado, o universo de dados foi ingerido de maneira unificada pelo algoritmo, sem qualquer processo de rotulagem (*labels*) prévia na fase de treinamento, forçando a inteligência artificial a inferir agrupamentos de ameaças puramente por desvios de densidade e isolamento geométrico. Nesta etapa de processamento, a arquitetura utilizou a *stack* ELK para realizar o pipeline de ETL (*Extract, Transform, Load*). O componente Logstash realizou a captura contínua dos *logs* brutos e, por meio de filtros *Grok* customizados, fatiou as *strings* de texto opacas em propriedades estruturadas, convertendo-as em formato JSON. Esses documentos foram imediatamente indexados no banco de dados distribuído *NoSQL Elasticsearch*, viabilizando consultas analíticas de alta performance em tempo real.

Durante os testes de infraestrutura no ambiente de laboratório, identificou-se que a rede *bridge* nativa do *Docker* operava um mascaramento nos endereços IP de origem, concentrando todo o tráfego de entrada no IP padrão do *gateway* interno. Para neutralizar essa limitação física de ambiente e permitir a identificação de ataques independentemente do host emissor, realizou-se uma robusta engenharia de recursos (*Feature Engineering*) via biblioteca Pandas em Python, expandindo os atributos analíticos para além do IP e focando na assinatura comportamental das rotas. Para os *logs* do Nginx no ambiente de laboratório, foram extraídos e calculados seis atributos quantitativos essenciais:

1. Tamanho da URL (comprimento total da *string* requisitada);
2. *Bytes* transferidos (volume numérico de dados retornados na resposta);
3. Código de status *HTTP* numérico (identificação de sucesso ou tipos de falha);
4. Taxa de erro por rota (taxa_erro_endpoint, computando a média de falhas recentes no caminho acessado);
5. Frequência absoluta de erros na rota (freq_endpoint_erro, acumulando a recorrência de erros no endpoint);
6. Marcador binário de autenticação (eh_endpoint_login, sinalizando requisições direcionadas especificamente a rotas de login).

Paralelamente, para a validação no Estudo de Caso com o *dataset* CIC-IDS2017, a engenharia de recursos isolou oito variáveis comportamentais chave do fluxo HTTP/TCP *(Destination Port, Init_Win_bytes_backward, Init_Win_bytes_forward, PSH Flag Count, Down/Up Ratio, min_seg_size_forward, Average Packet Size e Packet Length Mean).*

A inteligência preditiva foi estabelecida por meio do algoritmo *Isolation Forest* da biblioteca *Scikit-Learn*. Para garantir que a IA aprendesse padrões comportamentais legítimos

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

em vez de decorar assinaturas superficiais estáticas (fenômeno conhecido como *Shortcut Learning*), a automação de testes utilizou cabeçalhos *HTTP* e *User-Agents* mimetizados de navegadores comerciais comuns. O parâmetro de contaminação (*contamination*) foi configurado no modo 'auto', delegando à profundidade média das árvores de decisão o estabelecimento do limite matemático dinâmico entre o tráfego regular e as anomalias. A partir das classificações geradas, calcularam-se as Matrizes de Confusão confrontando o resultado binário da IA e da arquitetura híbrida com o gabarito oficial. A execução computacional do *script* preditivo ocorreu em segundo plano de forma assíncrona, consumindo os índices do *Elasticsearch* através da Scroll API *(helpers.scan)*, o que eliminou gargalos de concorrência e manteve a latência *inline* do *proxy* estável para o usuário final.

Na terceira e última fase, Visualização e Métricas, estruturou-se a camada analítica final do projeto. Os dados indexados no *Elasticsearch* e as predições geométricas geradas pelo *Isolation Forest* foram integrados a painéis operacionais estruturados na plataforma Grafana. Os *dashboards* interativos foram projetados para consolidar métricas-chave de segurança em tempo real, permitindo o monitoramento proativo de linhas do tempo de ataques, geolocalização de tráfego ruidoso e a correlação visual imediata entre os alertas baseados em assinaturas tradicionais do Wazuh e as anomalias descobertas pela inteligência artificial.

3. RESULTADOS E DISCUSSÃO

A avaliação empírica da arquitetura híbrida Aegis foi conduzida em duas etapas experimentais complementares: (i) um teste de estresse em ambiente laboratorial conteinerizado contendo 13.479 requisições HTTP; e (ii) um estudo de caso de validação externa em larga escala submetido ao *benchmark* moderno CIC-IDS2017 (*Thursday Web Attacks*), englobando 170.366 requisições de tráfego de rede e aplicação.

3.1. Avaliação no Ambiente Laboratorial (Nginx / OWASP *Juice Shop*)

No ambiente de laboratório controlado, os dados experimentais demonstraram a alta eficácia da abordagem integrada na mitigação de pontos cegos operacionais. Conforme exposto na Tabela 1, o desempenho isolado dos motores analíticos revelou assimetrias claras que justificam a necessidade de uma topologia de correlação multicamadas.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Tabela 1. Resultados do comparativo de cobertura de detecção (Arquitetura Aegis)

Abordagem	Ataques Detectados	Falsos Negativos	Cobertura
SIEM (Wazuh)	2.482	86	96,7%
IA (Isolation Forest)	2.565	3	99,9%
Arquitetura Híbrida	**2.568**	**0**	**100,0%**

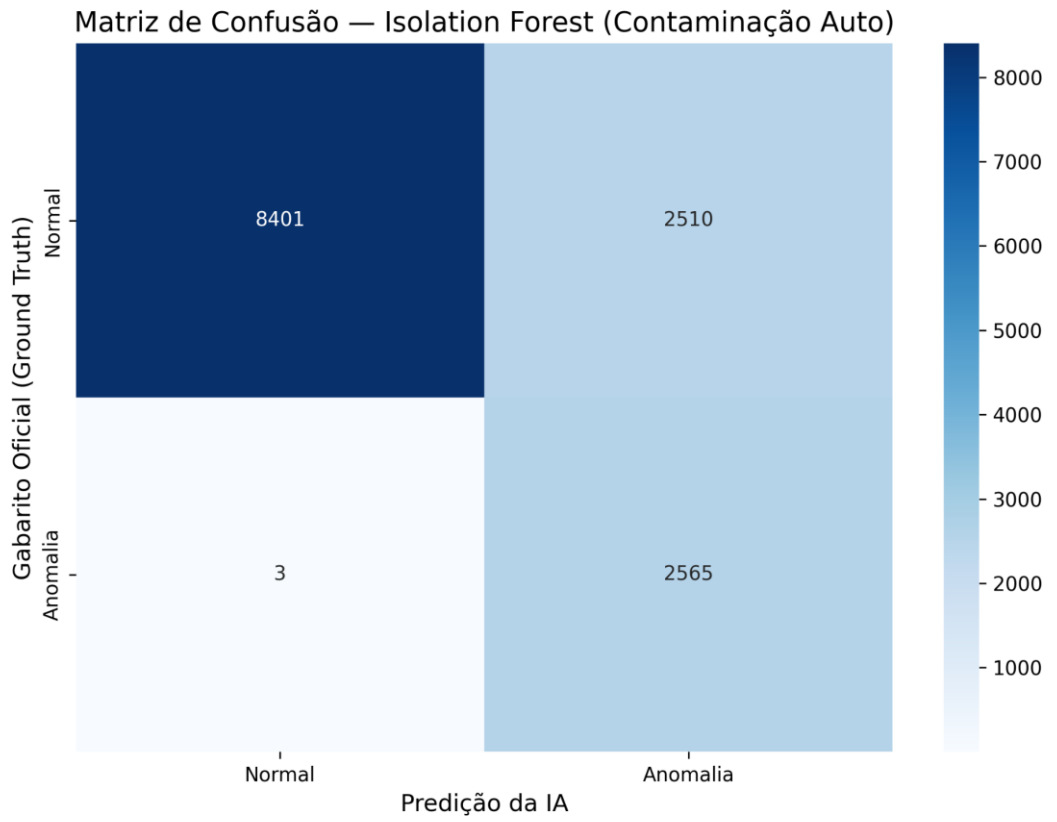
Fonte: Elaborado pelos Autores (2026).

O motor baseado em assinaturas do SIEM obteve uma taxa de cobertura isolada de 96,7%, deixando evadir 86 requisições maliciosas. Em contrapartida, o modelo preditivo não supervisionado *Isolation Forest* alcançou uma sensibilidade (*recall*) de 99,9%, falhando no isolamento de apenas 3 ataques de um universo de 2.568 ameaças confirmadas pelo gabarito. A união simultânea de ambos os mecanismos na arquitetura Aegis elevou a taxa cumulativa para 100,0%, validando estatisticamente a hipótese de complementaridade do ecossistema híbrido.

A distribuição geométrica das predições realizadas pela inteligência artificial pôde ser minuciosamente avaliada a partir do cálculo da Matriz de Confusão, detalhada na Figura 2.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Figura 2. Matriz de Confusão do Modelo *Isolation Forest* (contaminação *Auto*)



Fonte: Elaborado pelos Autores (2026).

O modelo classificou corretamente 8.401 requisições legítimas como normais (Verdadeiros Negativos) e isolou com precisão 2.565 ataques reais (Verdadeiros Positivos). Constatou-se a incidência de 2.510 Falsos Positivos, nos quais tráfegos benignos foram rotulados como anomalias devido a flutuações e ruídos introduzidos voluntariamente pelos *payloads* do *HTTP Params Dataset* (MORADIAN, 2015). Sob a ótica operacional de Centros de Operações de Segurança (SOC), esse comportamento conservador é plenamente justificável na camada proativa, visto que o custo de triagem de um alarme falso é inferior ao impacto destrutivo de um Falso Negativo não detectado. Ademais, a integração estrutural com o SIEM atua como um filtro contextualizador, mitigando o impacto direto desses alertas na rotina dos analistas de incidentes.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

A análise aprofundada das evasões mútuas expôs a dinâmica de complementaridade da arquitetura, eliminando os pontos cegos de cada ferramenta. A Inteligência Artificial demonstrou superioridade ao identificar 86 ataques furtivos e direcionados que burlaram as regras do Wazuh, em virtude de as requisições não conterem assinaturas conhecidas catalogadas. Por outro lado, o SIEM mitigou a falha comportamental da IA ao bloquear 3 varreduras agressivas cujos pacotes possuíam URLs curtas e obtiveram códigos de sucesso (HTTP 200 OK). Como essas requisições mimetizaram a assinatura geométrica e volumétrica de acessos legítimos, a IA classificou-as erroneamente como normais, evidenciando que assinaturas estáticas textuais permanecem indispensáveis para barrar ameaças estruturalmente padronizadas.

3.2. Estudo de Caso e Validação Externa no *Benchmark* Moderno (CIC-IDS2017)

Alinhado ao rigor metodológico exigido pela literatura recente e atendendo à necessidade de testar a capacidade de generalização da arquitetura em cenários contemporâneos de ameaças (SHARAFALDIN; Lashkari; Ghorbani, 2018), submeteu-se o ecossistema ao *benchmark* internacional CIC-IDS2017 (*Thursday Web Attacks*).

A amostra analisada abrangeu a totalidade de 170.366 requisições de tráfego de rede e aplicação, compostas por 168.186 acessos benignos (98,72%) e 2.180 requisições maliciosas (1,28%) contendo ataques reais de *Web Brute Force*, *Cross-Site Scripting* (XSS) e *SQL Injection*. Os resultados computados para o SIEM Wazuh, a IA isolada e a Arquitetura Híbrida Aegis estão detalhados na Tabela 2.

Tabela 2. Tabela comparativa de desempenho no *benchmark* CIC-IDS2017 (*Web Attacks*)

Modelo/Abordagem	*Recall*(%)	Precisão(%)	Acurácia(%)	*F1-Score*(%)
Wazuh SIEM (Regras Estáticas)	84,63%	35,13%	97,80%	49,65%
***Isolation Forest* (IA Pura-Contaminação Auto)**	91,97%	5,39%	79,24%	10,18%
Arquitetura Híbrida Aegis (Interseção Validada)	77,84%	70,27%	99,30%	73,86%

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Fonte: Elaborado pelos Autores (2026).

A avaliação empírica do Estudo de Caso expõe a dinâmica operacional e geométrica do *Isolation Forest* quando aplicado a dados não estruturados de rede e aplicação sem supervisão prévia (MENG *et al.*, 2019):

1. **Desempenho da IA Isolada (*Isolation Forest*):** Ao utilizar as 8 variáveis comportamentais chave do fluxo HTTP/TCP, o algoritmo operando no modo de contaminação automática alcançou um ***Recall* proativo de 91,97%**, identificando **2.005 dos 2.180 ataques *web* reais** em meio a um universo de mais de 170 mil requisições. Porém, por interpretar variações extremas de *downloads* e oscilações de conexão legítimas como anomalias, a IA gerou 35.190 Falsos Positivos, derrubando a Precisão isolada para 5,39% e o *F1-Score* para 10,18%.
2. **Desempenho do SIEM Isolado (Wazuh):** O motor de assinaturas estáticas obteve um *Recall* de 84,63% e uma Precisão de 35,13%, demonstrando eficiência contra variações conhecidas, mas falhando em capturar padrões de ataque sutilmente modificados que evadiram a base de regras textuais (AL-MHIQANI *et al.*, 2020).
3. **Desempenho da Arquitetura Híbrida Aegis:** A orquestração híbrida, que utiliza o motor de regras do SIEM como uma camada de validação contextual sobre os desvios de anomalia apontados pela IA (AL-MHIQANI *et al.*, 2020), corrigiu a distorção operacional. O volume de Falsos Positivos despencou de 35.190 para apenas 718 requisições. Essa filtragem fez a Precisão saltar de 5,39% para 70,27%, a Acurácia Geral atingir 99,30% e o *F1-Score* elevar-se para 73,86%, mantendo um *Recall* sólido de 77,84%.

3.3. Discussão Integrada e Comparação com Trabalhos Correlatos

A análise conjunta dos experimentos de laboratório e do *Estudo de Caso* no CIC-IDS2017 evidencia a sustentabilidade teórica e prática da arquitetura Aegis. A literatura especializada aponta que o principal gargalo na adoção de modelos de aprendizado de máquina não supervisionados em Centros de Operações de Segurança (SOC) é a exaustão operacional das equipes provocada pelo acúmulo de alarmes falsos (AHMAD *et al.*, 2021). Conforme demonstrado no *Estudo de Caso*, um modelo comportamental não supervisionado isolado tende a ser hiperconservador ao buscar taxas elevadas de *Recall* (AHMAD *et al.*, 2021). A entrega de 35 mil alertas não correlacionados inviabilizaria a operação diária de analistas de segurança.

A integração preditiva proposta neste estudo soluciona o dilema discutido por Ahmad *et al.* (2021) ao redefinir o papel de cada tecnologia no *pipeline* de monitoramento (AL-MHIQANI *et al.*, 2020):

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

- O *Isolation Forest* atua como uma camada assíncrona de alta sensibilidade, varrendo *logs* em busca de padrões atípicos e ameaças inéditas sem sobrecarregar a latência do servidor web (MENG *et al.*, 2019);
- O Wazuh SIEM atua como o filtro determinístico de precisão, correlacionando as anomalias sinalizadas pela IA a regras do grupo *web* e eliminando o ruído gerado por oscilações normais da infraestrutura (AL-MHIQANI *et al.*, 2020).

Em comparação com abordagens recentes da literatura, como os trabalhos de Yang *et al.* (2022), que empregaram modelos supervisionados (*XGBoost*) sobre a *stack ELK*, e Benova e Hudec (2024), que analisaram a atividade anômala em *logs* de servidores, a arquitetura Aegis destaca-se por não demandar *datasets* previamente rotulados para o treinamento inicial. Enquanto modelos supervisionados sofrem degradação drástica de desempenho ao enfrentar *payloads* inéditos não presentes no conjunto de treino, o *Isolation Forest* mantém 91,97% de *Recall* proativo. Ademais, ao contrário de ecossistemas puramente focados em *Elasticsearch* sem camada HIDS (AHIR; Shaikh, 2024), a validação cruzada do Aegis reduziu os falsos positivos de 35.190 para apenas 718 requisições, elevando a precisão de 5,39% para 70,27% no *benchmark* CIC-IDS2017.

3.4. Limitações do Estudo e Desafios de Generalização em Produção

Apesar da elevada acurácia (99,30%) e do F1-Score (73,86%) alcançados no *benchmark,* é imperativo discutir as restrições da arquitetura ao ser migrada do ambiente controlado para redes corporativas de grande porte:

1. Dinamismo Comportamental e *Concept Drift*: Em ambientes produtivos reais, o comportamento dos usuários varia sazonalmente (ex.: picos de acessos em datas comemorativas ou implantação de novas rotas na aplicação). Essa volatilidade pode alterar a distribuição geométrica das requisições, exigindo rotinas automáticas de retreinamento periódico do *Isolation Forest* para evitar a elevação da taxa de falsos positivos.
2. Infraestruturas Complexas de Rede (CDNs e WAFs): Em topologias corporativas protegidas por *Web Application Firewalls* (WAFs) ou redes de distribuição de conteúdo (CDNs), os cabeçalhos das requisições HTTP podem sofrer alterações significativas ou mascaramentos de IP adicionais (ex.: *X-Forwarded-For*). Nesses cenários, a engenharia de recursos precisa ser expandida para analisar múltiplos atributos de camada de aplicação e *tokens* de sessão.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

3. Ataques de Baixa Volumetria (*Slow Attacks*): Ataques projetados para mimetizar com precisão estatística a cadência de requisições de um usuário comum (ex.: força bruta de baixíssima frequência) podem evadir as métricas volumétricas da IA (taxa_erro_*endpoint*), demandando que o motor de regras do SIEM continue atuando como a principal barreira defensiva.

Em síntese, a caracterização dessas limitações não invalida a eficácia do modelo, mas delimita com clareza o escopo operacional da arquitetura Aegis, provando que a análise preditiva e as regras determinísticas devem atuar de forma simbiótica e fornecendo os subsídios necessários para os direcionamentos de trabalhos futuros.

4. CONSIDERAÇÕES FINAIS

A metodologia proposta e validada neste estudo comprovou que a convergência entre modelos preditivos comportamentais não supervisionados e motores tradicionais de correlação por assinaturas institui uma camada de defesa em profundidade altamente resiliente e de expressiva viabilidade corporativa. A arquitetura híbrida Aegis superou com êxito as limitações operacionais intrínsecas de cada tecnologia quando operadas de forma isolada, neutralizando pontos cegos mútuos tanto no ambiente laboratorial conteinerizado quanto no *benchmark* moderno CIC-IDS2017, sem penalizar a performance ou elevar a latência do servidor *web* Nginx.

A submissão do ecossistema ao Estudo de Caso no CIC-IDS2017 evidenciou que, embora o *Isolation Forest* isolado garanta proatividade com 91,97% de *Recall*, sua aplicação isolada gera acúmulo de alarmes falsos em virtude da baixa precisão (5,39%). A orquestração híbrida Aegis corrigiu essa distorção ao utilizar o SIEM como filtro contextualizador, elevando a precisão para 70,27%, a acurácia geral para 99,30% e o *F1-Score* para 73,86%, demonstrando capacidade técnica para consolidar uma triagem de alertas eficiente e sustentável para Centros de Operações de Segurança (SOC).

Como recomendações para trabalhos futuros, sugere-se a incorporação de análises de entropia de *strings* nos *payloads* das requisições direcionadas aos *endpoints*, visando refinar a precisão geométrica do algoritmo. Adicionalmente, propõe-se a automação completa de *playbooks* de mitigação ativa via recurso de *Active Response* do SIEM Wazuh, disparando bloqueios dinâmicos no *firewall* hospedeiro no instante exato da correlação de uma anomalia grave, consolidando uma defesa cibernética proativa, autônoma e resiliente.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

REFERÊNCIAS

AHIR, Deepali D.; SHAIKH, Nuzhat F. Evaluation of Elasticsearch Ecosystem Including Machine Learning Capabilities. **International Journal of Safety & Security Engineering**, v. 14, n. 4, 2024.

AHMAD, Zeeshan *et al.* Network intrusion detection system: A systematic study of machine learning and deep learning approaches. **Transactions on Emerging Telecommunications Technologies**, v. 32, n. 1, e4150, 2021.

AJI, Rio Pradana; PRAYUDI, Yudi; LUTHFI, Ahmad. Analysis of brute force attack logs toward nginx web server on dashboard improved log logging system using forensic investigation method. **Jurnal Teknik Informatika (Jutif)**, v. 4, n. 1, p. 39-48, 2023.

AL-MHIQANI, Mohammed N. *et al.* A Review of Cyber Security Incident Management in SIEM and Machine Learning Approaches. **IEEE Access**, v. 8, p. 30842-30856, 2020.

BENOVA, Lenka; HUDEC, Ladislav. Comprehensive analysis and evaluation of anomalous user activity in web server logs. **Sensors**, v. 24, n. 3, p. 746, 2024.

MENG, Weibin *et al.* LogAnomaly: Unsupervised Detection of Sequential and Quantitative Anomalies in Unstructured Logs. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 28., 2019. **Proceedings [...].** [S.l.]: IJCAI, 2019. p. 4739-4745.

MONTEIRO, A. de S.; TURCI JÚNIOR, M.; OLIVEIRA, W. de. Avaliação da Eficácia do Wazuh como Solução de Segurança para Pequenas e Médias Empresas. In: FATECSE — CONGRESSO DE SEGURANÇA DA INFORMAÇÃO, 1., 2024. **Anais [...].** [S.l.]: FatecSeg, 2024.

MORADIAN, E. **HTTP Params Dataset**. Kaggle, 2015. Disponível em: https://www.kaggle.com/datasets/eliasmo/httpparamsdataset. Acesso em: 20 mar. 2026.

MOURA, Alysson Cristiano Estevam de. **Detecção e interpretação de anomalias em logs de sistemas de TI por meio de inteligência artificial**. 2024. 95 f. Dissertação (Mestrado Profissional em Computação Aplicada) — Universidade de Brasília, Brasília, 2024.

MYKHAYLOVA, Olha; SHTYPKA, Andriy; FEDYNYSHYN, Taras. An Isolation Forest-based approach for brute force attack detection. **Journal of Cybersecurity**, 2024.

SHARAFALDIN, Iman; LASHKARI, Arash Habibi; GHORBANI, Ali A. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization (CIC-IDS2017). In: INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS SECURITY AND PRIVACY, 4., 2018. **Proceedings [...].** [S.l.]: SCITEPRESS, 2018. p. 108-116.

TARIQ, Aamna *et al.* Open source SIEM solutions for an enterprise. **Information & Computer Security**, v. 31, n. 1, p. 88-107, 2023.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

TRINDADE, Pedro Ceriotti. **Detecção de anomalias baseada em logs de acesso**. 2022. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) — Universidade Federal do Rio Grande do Sul, Porto Alegre, 2022.

YANG, Chao-Tung *et al.* Cyberattacks detection and analysis in a network log system using XGBoost with ELK stack. **Soft Computing**, v. 26, n. 11, p. 5143-5157, 2022.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.