HIKARI – Desafios Reais, Ameaças Ocultas: Uma Plataforma de Threat Hunting Gamificado

Sidnei Barbieri[1], Bruno Moreira Camargos Belo[1], Leonardo Chahud[1], Leonardo Vaz de Meneses[1], Cesar Marcondes[1], Lourenço Alves Pereira Júnior[1]

[1]Divisão de Ciência da Computação – Instituto Tecnológico de Aeronáutica (ITA)
Praça Marechal Eduardo Gomes, 50 – 12228-900 – São José dos Campos – SP – Brasil

{sidneisb,brunomoreira,chahud,leonardomeneses,cmarcondes,ljr}@ita.br

Abstract. *HIKARI is a training platform for Blue Teams operating in realistic environments, integrating progressive Capture The Flag (CTF) challenges with the ingestion and analysis of real security events through an ELK stack (Elasticsearch, Logstash e Kibana). The platform employs a Kafka bus for controlled event injection, CTFd for gamified challenge management, and Kibana as an investigation interface. This paper describes the system architecture, analyzes its differentiators compared to existing solutions, and presents experimental results from three educational deployments involving over 40 participants. The results suggest that HIKARI advances the state of the art and fosters practical skill development in cybersecurity defense operations.*

Resumo. *O HIKARI é uma plataforma para treinamento de Blue Teams em ambientes realistas, integrando desafios progressivos no formato CTF com a ingestão e análise de eventos de segurança reais em uma ELK stack (Elasticsearch, Logstash e Kibana) . A plataforma emprega um barramento Kafka para injeção controlada de eventos, o CTFd para gestão gamificada de desafios e o Kibana como interface de investigação. Este artigo descreve a arquitetura do sistema, analisa seus diferenciais em relação a soluções existentes e apresenta os resultados de experimentos aplicados em três contextos educacionais distintos, envolvendo mais de 40 participantes. Os resultados sugerem que o HIKARI contribui para a evolução do estado da arte e promove o desenvolvimento prático de competências em operações de defesa cibernética.*

1. Introdução

O aumento da frequência e sofisticação dos ataques cibernéticos tem levado organizações públicas e privadas a aprimorar continuamente suas capacidades defensivas. A digitalização de serviços essenciais e a crescente dependência tecnológica ampliaram a superfície de ataque, exigindo respostas rápidas e coordenadas por parte dos *Blue Teams*.

Apesar da importância dos *Blue Teams*, ainda existem lacunas na formação prática, especialmente na análise de logs, correlação de eventos e resposta a incidentes. O *Cybersecurity Skills Gap Report* estima que mais de 3,4 milhões de posições em cibersegurança permanecem não preenchidas globalmente, com destaque para o Brasil [Fortinet 2024]. Nesse cenário, a lacuna na capacitação torna-se crítica, impulsionada pelo aumento dos incidentes e pela evolução da maturidade regulatória.

Incidentes recentes em setores essenciais da sociedade exemplificam a necessidade de equipes capazes de atuar sob pressão e analisar rapidamente evidências técnicas. Enquanto plataformas de treinamento ofensivo tornaram-se populares, persiste uma escassez de soluções educacionais capazes de simular adequadamente as rotinas operacionais típicas de um *Security Operations Center* (SOC), com integração de ferramentas de observabilidade e desafios realistas orientados por dados.

Este artigo propõe o **HIKARI**[1], uma plataforma de treinamento gamificada desenvolvida para capacitar *Blue Teams* em ambientes de defesa cibernética realistas e orientados a dados. A plataforma integra desafios no formato *Capture The Flag* (CTF), injeção progressiva de eventos e análise de evidências em um ambiente inspirado em *Security Information and Event Management* (SIEM) – caracterizado como *SIEM-like* – utilizando a *ELK stack* para centralização e visualização de logs. A arquitetura adota uma abordagem orientada a dados, combinando pipelines de ingestão de eventos com ferramentas amplamente utilizadas na indústria e desafios baseados em cenários operacionais típicos de um SOC real.

As principais contribuições deste trabalho são: o desenvolvimento de modelos de injeção temporal progressiva de eventos para simulação realista de cenários defensivos; o emprego de uma estratégia de gamificação no estilo *Capture the Flag*, voltada à formação de *Blue Teams* — uma abordagem ainda pouco explorada na literatura; e a condução de três campanhas de treinamento, a partir das quais foram extraídas métricas de aprendizagem e experiência. Essas contribuições permitem a formulação de hipóteses verificáveis sobre a transferência de conhecimento para analistas de defesa em contextos operacionais. O artigo está estruturado da seguinte forma: a Seção 2 discute os trabalhos relacionados; a Seção 3 descreve a arquitetura do HIKARI; a Seção 4 apresenta a execução dos estudos de caso; a Seção 5 analisa os resultados obtidos; e a Seção 6 conclui o trabalho e propõe direções futuras.

2. Trabalhos Relacionados

O avanço da formação prática em defesa cibernética tem sido objeto de diversas iniciativas que buscam aproximar o ensino das demandas operacionais em nível de SOCs. A seguir, apresenta-se uma análise do estado da arte com foco nas propostas recentes voltadas à formação prática de *Blue Teams*, considerando suas contribuições e limitações.

Estudos recentes têm explorado o uso de desafios de CTF como mecanismo de desenvolvimento de competências em cibersegurança, com ênfase em ambientes acadêmicos. A análise de [Švábenskỳ et al. 2021], baseada em 52 competições, evidencia que tais desafios favorecem o treinamento de habilidades práticas classificadas segundo o framework *NICE*, porém majoritariamente centradas em técnicas ofensivas. De modo semelhante, [Savin et al. 2023] propõem a rotulagem de sessões de CTF com base no *MITRE ATT&CK*, gerando datasets úteis para análise de comportamento cibernético, mas ainda restritos a eventos discretos e descontextualizados de ambientes operacionais. Já [Leune and Petrilli 2017] avaliam impactos educacionais de CTFs em alunos de graduação, com base em autoavaliações subjetivas, sem incorporar dados técnicos ou evidências de desempenho contínuo. Por fim, [Karagiannis et al. 2021] introduzem a plataforma PocketCTF como solução portátil para exercícios gamificados, incluindo

[1] `https://hikari-edu.github.io`

ferramentas como Suricata e Wazuh; contudo, a ausência de logs reais, progressão temporal ou orquestração de eventos limita seu valor em treinamentos voltados a *Blue Teams*. Em conjunto, esses trabalhos reforçam o potencial dos CTFs como instrumento pedagógico, mas revelam lacunas críticas em realismo operacional, continuidade analítica e integração com infraestruturas de defesa cibernética como SIEMs.

Além dos CTFs tradicionais, outras propostas educacionais têm buscado explorar abordagens gamificadas alternativas para fomentar competências em cibersegurança, sobretudo em públicos diversos e contextos de ensino mais amplo. O estudo de [Russo et al. 2023] descreve um *Cyber Defense Exercise* estruturado com simulações de rede, ataques automatizados com CVEs reais e tráfego legítimo simulado, compondo um cenário realista de defesa. No entanto, a execução foi individual e não incorporou mecanismos de ingestão contínua de eventos nem instrumentação SIEM-like, comprometendo a reprodutibilidade de análises forenses automatizadas. De forma análoga, [Gough et al. 2024] propõem o uso de brinquedos IoT hackeáveis como instrumentos pedagógicos acessíveis. Embora incentivem o engajamento prático com dispositivos embarcados, as atividades permanecem isoladas, sem logs operacionais, sem encadeamento de evidências nem suporte à investigação progressiva. Já [DeCusatis et al. 2022] apresentam uma sala de escape digital baseada na plataforma *Unity 3D*, com ênfase em mecânicas de jogo voltadas à retenção conceitual. Ainda que envolvente, a proposta carece de continuidade analítica e não simula fluxos típicos de resposta a incidentes. Complementarmente, o trabalho de [Zhong et al. 2024] aborda práticas de gamificação inclusiva, com foco em diversidade de perfis e acessibilidade, mas sem proposição de arquiteturas técnicas ou instrumentação realista. Assim, embora ampliem a base de participação e inovem em formato, essas iniciativas ainda não satisfazem os requisitos técnicos mínimos para treinamento progressivo e investigativo de *Blue Teams* em ambientes análogos a SOCs.

No domínio de ambientes industriais, observam-se iniciativas baseadas em *Cyber Ranges* que buscam aproximar o treinamento técnico das particularidades de sistemas ciberfísicos. A plataforma *BUTCA*, descrita por [Kuchar et al. 2024], emprega desafios gamificados e tráfego industrial simulado para treinamento em *Industrial Control Systems*, mas não inclui mecanismos de ingestão contínua de eventos nem suporte a investigações forenses encadeadas. Já [Vasilakis et al. 2024] desenvolvem uma arquitetura dockerizada com múltiplas HMIs e desafios integrados via CTFd e Naumachia, permitindo a exploração técnica de falhas específicas (e.g., *SQLi*, *Snort*, *MITM*); ainda assim, não há modelagem FSM-like nem integração com ferramentas de análise progressiva, dificultando simulações investigativas alinhadas a fluxos de SOCs. Em perspectiva complementar, o modelo sócio-técnico de [Kianpour et al. 2019] enfatiza o papel das interações organizacionais e perfis de usuário no design pedagógico de *Cyber Ranges*, mas carece de instrumentação operacional com dados reais ou controle temporal sobre o encadeamento de eventos. Por fim, o trabalho de [Diakoumakos et al. 2021] introduz um modelo federado de gamificação, atribuindo pesos a categorias de habilidades e consolidando pontuações entre ambientes distintos; no entanto, o modelo ignora ingestão de logs e reconstrução de incidentes em tempo real. Em conjunto, esses estudos ampliam o escopo das *Cyber Ranges* para além da dimensão técnica, incorporando aspectos organizacionais e estruturais, mas ainda não satisfazem integralmente os requisitos de simulação progressiva, análise forense contínua e adaptação dinâmica aos

estados de defesa típicos de ambientes *SOC-like*.

Adicionalmente, identificam-se esforços voltados à modelagem semântica de comportamentos adversariais e à compreensão das limitações práticas do *threat hunting* em ambientes reais. O trabalho de [Chetwyn et al. 2024] propõe uma ontologia baseada em eventos do Sysmon e emulações com o *MITRE Caldera*, permitindo inferir cadeias de comportamento adversário por meio de regras semânticas (SWRL) e consultas RDF. A abordagem possibilita a abstração de indicadores técnicos (como comandos *PowerShell* ou acessos administrativos) em *Indicators of Behaviour* mais genéricos, expandindo a capacidade de raciocínio contextualizado sobre ações hostis. Contudo, o sistema foca na construção de conhecimento e inferência semântica, sem demonstrar sua aplicação em treinamentos realistas com ingestão contínua de eventos ou integração com fluxos operacionais de SOCs. Em complemento, o estudo de [Badva et al. 2024] adota uma abordagem qualitativa para mapear os desafios enfrentados por analistas durante atividades de *threat hunting*, revelando barreiras cognitivas, limitações de *tooling* e dificuldades na formulação e validação de hipóteses diante de dados brutos. Apesar de relevante para entender os fatores humanos e processuais envolvidos, o estudo não propõe arquiteturas técnicas nem ambientes simulados para o desenvolvimento prático dessas competências. Assim, ambos os trabalhos contribuem para aprofundar a compreensão dos elementos conceituais e humanos do *threat hunting*, mas permanecem aquém dos requisitos técnicos e instrumentais exigidos para a formação progressiva de *Blue Teams*.

Os trabalhos analisados não integram plenamente a combinação de realismo, progressão controlada e integração com ferramentas *SIEM-like*. Em geral, priorizam o desenvolvimento de habilidades ofensivas, apresentam pouca integração com ferramentas reais de observabilidade ou requerem infraestruturas excessivamente complexas. A partir dessa análise, identificam-se três lacunas recorrentes: (i) ausência de ingestão contínua de eventos realistas, como logs técnicos de sistemas operacionais e aplicações legítimas; (ii) falta de mecanismos de orquestração temporal, que permitam a progressão encadeada de evidências e *flags*; e (iii) baixa integração com ferramentas reais de observabilidade, como a pilha ELK. Tais limitações constituíram os requisitos fundacionais considerados no desenho arquitetural do HIKARI.

Portanto, o HIKARI foi concebido como uma plataforma educacional que integra práticas recorrentes na literatura — como a utilização de logs reais, ambientes controlados e desafios progressivos — com requisitos operacionais de realismo, acessibilidade e controle de eventos. A Tabela 1 sintetiza essa avaliação, estruturando sete critérios técnicos extraídos da literatura recente e mapeando as lacunas às capacidades previstas na plataforma.

No contexto da formação prática em defesa cibernética, observam-se plataformas que adotam ambientes gamificados e laboratórios simulados. Tais soluções, embora eficazes na introdução de técnicas ofensivas, como *penetration testing* e *reverse engineering*, apresentam deficiências recorrentes em cenários de defesa cibernética realista. As lacunas mais frequentes incluem: ausência de ingestão contínua de logs reais; inexistência de mecanismos explícitos de orquestração temporal de eventos; e baixa integração com ferramentas amplamente utilizadas em SOCs, como a *ELK stack*. Esses aspectos inviabilizam a simulação de investigações encadeadas com evidências técnicas progressivas, limitando a aplicabilidade dessas plataformas em treinamentos defensivos.

Tabela 1. Comparação entre o HIKARI e outros trabalhos da literatura

Trabalho	Foco em Blue Team	Ingestão de Logs Reais	Integração com SIEM	Orquestração Temporal	IaC	Gamificação	Avaliação Empírica
[Švábenskỳ et al. 2021]	◐	◐	○	○	○	●	●
[Savin et al. 2023]	●	◐	○	○	○	●	●
[Russo et al. 2023]	●	○	◐	○	○	●	◐
[Chetwyn et al. 2024]	●	◐	●	○	○	○	●
[Badva et al. 2024]	●	○	●	◐	○	○	●
[Kuchar et al. 2024]	●	○	◐	○	○	●	●
[Gough et al. 2024]	●	○	○	○	○	●	●
[Vasilakis et al. 2024]	●	○	◐	○	○	●	◐
[Diakoumakos et al. 2021]	◐	○	◐	●	○	●	●
[Kianpour et al. 2019]	◐	○	○	○	○	●	◐
[Zhong et al. 2024]	●	○	○	○	○	●	●
[Leune and Petrilli 2017]	○	○	○	○	○	●	◐
[DeCusatis et al. 2022]	○	○	○	○	○	●	◐
[Karagiannis et al. 2021]	●	○	●	◐	○	◐	◐
HIKARI	✓	✓	✓	✓	✓	✓	✓

Nota: ● presença funcional; ○ ausência funcional; ◐ presença limitada ou parcial.

Plataformas comerciais de *Cyber Range*, como o Cyberbit Range[2] e o IBM X-Force Cyber Range[3], exemplificam essa tendência. Embora ofereçam simulações ricas em telemetria e incidentes complexos, operam sob arquiteturas proprietárias não auditáveis, com forte dependência de dados sintéticos e ausência de interfaces abertas para ingestão ou visualização externa de eventos. Além disso, os custos elevados e o modelo fechado de licenciamento dificultam sua adoção em contextos educacionais.

Soluções de treinamento técnico como o Hack The Box[4], Immersive Labs[5] e RangeForce[6] também apresentam limitações estruturais sob os critérios avaliados. No Hack The Box, os desafios CTF são majoritariamente isolados, desprovidos de fluxo temporal ou integração com fontes contínuas de telemetria. A Immersive Labs oferece laboratórios guiados voltados à resposta a incidentes e *threat hunting*, mas os dados são pré-processados e não permitem exploração investigativa com rastreabilidade contínua. O RangeForce, por sua vez, provê exercícios *hands-on* estáticos com foco técnico, sem progressão baseada em eventos e sem flexibilidade para inserção de logs reais.

No domínio de soluções de código aberto, o Security Onion[7] destaca-se por sua instrumentação precisa para coleta e análise de eventos, integrando ferramentas como Suricata, Zeek e Kibana. No entanto, sua adoção em treinamentos é limitada por fatores pedagógicos e operacionais: a ausência de gamificação, a elevada curva de aprendizado e a complexidade de implantação dificultam seu uso em simulações progressivas para formação de novos analistas.

Além dos aspectos técnicos e da gamificação, as soluções analisadas divergem

[2] https://www.cyberbit.com/platform/cyber-range/
[3] https://www.ibm.com/services/xforce-cyber-range
[4] https://www.hackthebox.com
[5] https://www.immersivelabs.com
[6] https://www.rangeforce.com
[7] https://securityonionsolutions.com/software

em critérios fundamentais para a capacitação prática de analistas em segurança, tais como regionalização de conteúdos, capacidade de simular *threat hunting*, realismo de ambientes simulados, suporte à gestão de competições estruturadas e modelo de precificação. Embora algumas plataformas possam utilizar automação interna para provisionamento de ambientes, nenhuma das soluções analisadas expõe sua infraestrutura como código – *Infrastructure as Code* (IaC). Nenhuma plataforma analisada reproduz integralmente o conjunto de requisitos operacionais implementados no HIKARI.

A seguir, a Tabela 2 analisa as principais plataformas que integram gamificação e ambientes simulados voltados à capacitação defensiva em cibersegurança. Ela compara as principais soluções existentes a partir de cinco critérios estratégicos para formação prática em defesa cibernética. A regionalização refere-se à capacidade de adaptar conteúdos, desafios e interfaces a contextos culturais e linguísticos. A caça de ameaças (*threat hunting*) avalia se a plataforma permite investigações proativas. O realismo do ambiente simulado refere-se ao uso de dados reais ou simulações representativas. O suporte à gestão de competições considera a existência de mecanismos de orquestração, progressão e ranqueamento de desafios. Por fim, o critério de precificação examina o modelo de acesso — seja código aberto, assinatura comercial ou licenciamento fechado. Os critérios adotados foram extraídos a partir das funcionalidades reportadas nas plataformas revisadas, considerando requisitos operacionais relevantes à formação técnica de analistas defensivos.

Tabela 2. Comparação entre plataformas de capacitação em defesa cibernética

Soluções	Regionalização	Caça de ameaças	Ambiente Simulado Realista	Gerência de competições	Direcionada para Educação	Precificação
Cyber Range Solutions	●	●	●	●	●	Soluções fechadas
Hack The Box	○	●	●	●	○	Assinatura
Immersive Labs	○	●	●	●	●	Assinatura
RangeForce	○	●	●	●	●	Assinatura
Security Onion	○	●	●	○	○	Código aberto
HIKARI	●	●	●	●	●	Código aberto

Nota: ● indica presença funcional; ○ indica ausência funcional.

Com efeito, as soluções revisadas abordam aspectos isolados da formação de *Blue Teams*, mas raramente integram ingestão contínua de dados realistas, progressão temporal de evidências e ferramentas *SIEM-like*. O HIKARI diferencia-se ao combinar eventos reais anonimizados, injetados progressivamente via Kafka [8], com desafios gamificados orquestrados por meio do CTFd [9], em uma arquitetura aberta baseada na *ELK stack* [10]. Essa abordagem permite simular incidentes com progressão controlada, análise investigativa com evidências reais e suporte à formação prática de analistas em ambientes próximos ao de um SOC real.

A Tabela 3 sintetiza o mapeamento entre as lacunas identificadas na literatura e os requisitos técnicos assumidos no HIKARI. Esses critérios foram extraídos da análise crítica dos trabalhos discutidos e validados pela estrutura técnica descrita nas seções 3 e 4.

[8] https://kafka.apache.org
[9] https://ctfd.io/
[10] https://www.elastic.co/elastic-stack

Tabela 3. Síntese crítica: lacunas recorrentes e requisitos assumidos no HIKARI

Lacunas recorrentes na literatura	Requisitos técnicos assumidos no HIKARI
Ausência de ingestão contínua de eventos realistas (e.g., logs de SO, rede e aplicação)	Ingestão de eventos reais anonimizados com progressão temporal controlada via Kafka e Filebeat
Falta de orquestração encadeada de desafios e evidências ao longo do tempo	Composição modular de flags e evidências via CTFd com suporte à progressão e investigação
Baixa integração com ferramentas reais de observabilidade (SIEM-like)	Arquitetura baseada na ELK stack com visualização contínua em tempo real (Kibana, Elasticsearch)
Foco excessivo em técnicas ofensivas e desafios isolados	Ênfase em investigações orientadas à defesa, com foco em *Blue Teams* e contexto operacional de SOC
Arquiteturas fechadas ou pouco reprodutíveis	Modelo aberto, baseado em *Infrastructure as Code*, com suporte à reprodutibilidade e extensão

3. Arquitetura da Plataforma HIKARI

A arquitetura do HIKARI foi concebida para atender aos requisitos técnicos derivados das limitações identificadas na literatura (Tabela 3), visando simular, de forma reproduzível e escalável, a dinâmica de resposta a incidentes por equipes defensivas. O sistema adota uma arquitetura orientada por dados, com ingestão contínua de eventos, orquestração gamificada de desafios e ferramentas reais de análise em tempo real. A Figura 1 apresenta uma visão geral da infraestrutura técnica, composta por componentes de gamificação (CTFd), transporte de eventos (Kafka e scripts auxiliares) e análise centralizada por meio da *ELK stack*.

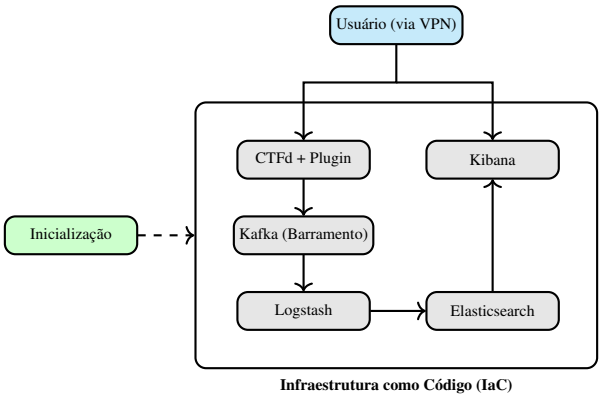

Figura 1. Arquitetura da plataforma HIKARI

A infraestrutura do HIKARI é implementada como IaC, permitindo a inicialização completa dos serviços a partir de um ponto único de orquestração (Inicialização), responsável pelo provisionamento da infraestrutura e pela configuração operacional básica dos serviços. O CTFd gerencia os desafios e dispara o processo de ingestão de eventos, representados por arquivos JSON, que são processados, indexados e persistidos por meio do Kafka, Logstash e Elasticsearch, tornando-se posteriormente disponíveis para visualização na interface do Kibana. Cada usuário ou equipe participante da competição ou treinamento interage exclusivamente com o CTFd e o Kibana mediante conexão VPN (Virtual Private Network), conduzindo atividades típicas de *Blue Teams* sobre os dados disponibilizados, localizando as *flags* necessárias para superar os desafios e progredir

na competição. Todos os serviços executam-se em contêineres Docker, assegurando isolamento, paralelismo e reprodutibilidade operacional.

O CTFd gerencia a apresentação dos desafios, a submissão de *flags* e o controle de progressão dos participantes. O Kafka, o Logstash e o Elasticsearch, conforme descrito anteriormente, orquestram o fluxo de eventos estruturados, da ingestão à disponibilização para análise no Kibana. A análise interativa é conduzida no Kibana, que disponibiliza dashboards e mecanismos de consulta para a investigação de evidências pelos participantes.

Cada desafio é descrito no CTFd e pode ou não ser associado a um novo conjunto de eventos (logs) em formato JSON. Caso associado, a liberação do desafio implica na injeção do respectivo conjunto de logs em tópicos Kafka. O participante que primeiro avançar na competição aciona a publicação dos eventos, que são consumidos por pipelines Logstash, parseados e indexados no Elasticsearch, ficando disponíveis para análise no Kibana, conforme ilustrado na Figura 2.

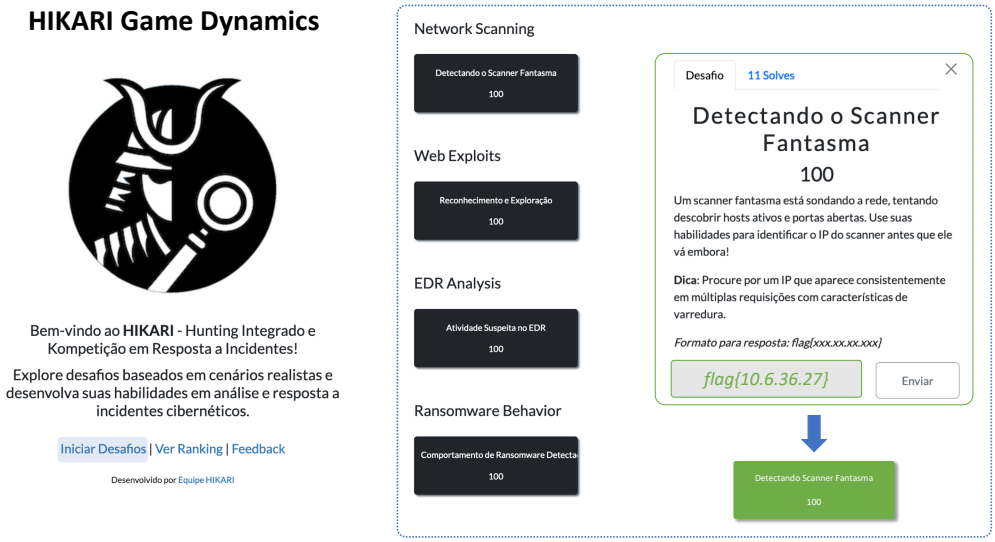

Figura 2. Dinâmica de liberação de desafios no HIKARI.

Esse fluxo simula o funcionamento de um SOC, no qual alertas e evidências chegam continuamente, exigindo correlação temporal e análise contextual pelos analistas. Todo o gerenciamento interno da plataforma é automatizado e transparente para os competidores. As competições contemplam cenários comuns a ambientes corporativos, como varreduras de rede, exploração de serviços vulneráveis, movimentação lateral e exfiltração de dados.

Para coordenar a progressão dos desafios, foi desenvolvido um plugin customizado para o CTFd, como parte integrante da plataforma HIKARI. Esse componente automatiza a associação entre desafios e conjuntos de logs técnicos em formato JSON, acionando a publicação dos dados em tópicos Kafka a cada estágio vencido. Essa funcionalidade garante a injeção controlada e progressiva de evidências, viabilizando simulações realistas com controle temporal e coordenação por equipe. O Kafka, Logstash e Elasticsearch orquestram o fluxo estruturado de eventos, que são

instantaneamente disponibilizados para consultas no Kibana a partir da ativação do desafio correspondente.

A Figura 3 apresenta a sequência de execução da plataforma HIKARI, desde o registro da competição até a resolução dos desafios e a injeção progressiva de logs para investigação. O diagrama evidencia como o plugin do CTFd coordena a publicação de logs a cada estágio vencido, enquanto o pipeline ELK processa, indexa e disponibiliza as evidências para análise interativa em tempo real.

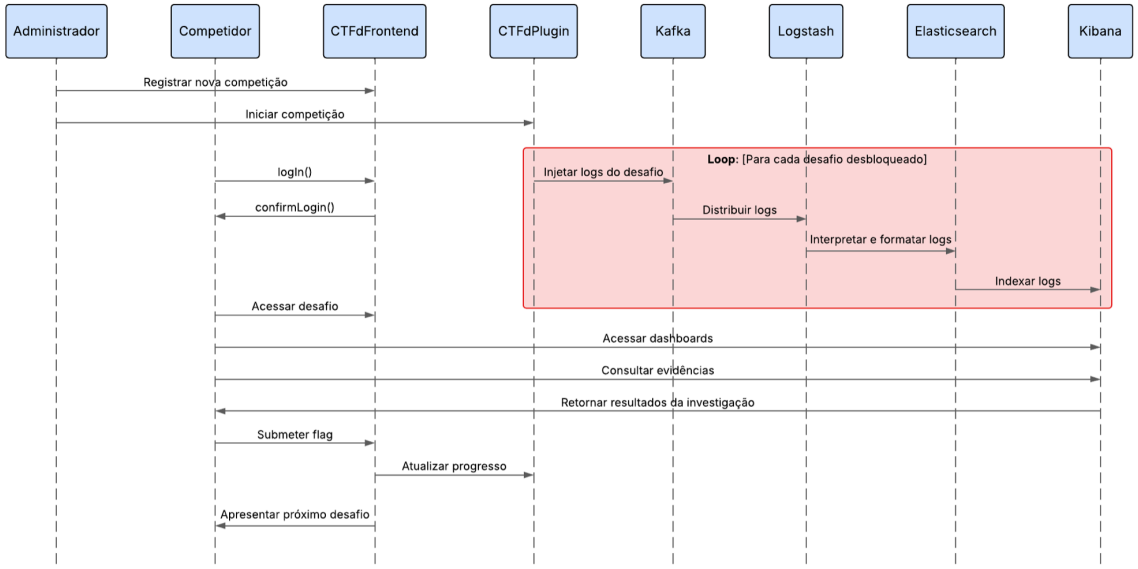

Figura 3. Diagrama de sequência: cadastro, injeção de logs e resolução de desafios na plataforma HIKARI.

A associação de desafios a diferentes arquivos JSON permite a construção de desafios progressivos, possibilitando ajustes finos de dificuldade e foco temático em cada etapa da competição ou treinamento. Os desafios do HIKARI podem ser estruturados em múltiplos estágios, representando fases típicas de incidentes cibernéticos, como reconhecimento, exploração, persistência e impacto (ex.: ransomware). A liberação progressiva dos eventos é coordenada pelo CTFd, que publica novos conjuntos de logs em intervalos controlados ou sincronizados com o avanço dos participantes. A base Elasticsearch é compartilhada entre todos os competidores, aqueles que progridem mais rapidamente interagem inicialmente com quantidade reduzida de eventos, enquanto novos dados são injetados dinamicamente e ficam visíveis para todas as equipes. Esse modelo reforça a simulação da dinâmica de investigação sob pressão, típica de centros reais de operação de segurança, ao mesmo tempo em que recompensa a agilidade dos participantes.

Os desafios atualmente disponíveis no HIKARI foram construídos com base em logs reais anonimizados, extraídos do setor financeiro para fins de simulação realista. Entretanto, a arquitetura é flexível e permite a substituição ou complementação desses desafios. Também podem ser utilizados logs sintéticos, gerados artificialmente para simular ataques específicos ou padrões de comportamento, permitindo a criação de cenários controlados conforme os objetivos pedagógicos.

Os logs utilizados nos desafios da plataforma HIKARI foram extraídos de

ambientes reais do setor financeiro, com autorização institucional, tratados sob rígidos critérios de anonimização e formatados para permitir sua utilização nos desafios. O pipeline completo abrangeu desde a coleta e limpeza dos dados até a estruturação em JSON e a ingestão na *ELK stack*. O objetivo foi preservar a veracidade técnica dos eventos, eliminando, ao mesmo tempo, quaisquer vestígios de informações sensíveis. A coleta foi orientada por quatro domínios estratégicos de observabilidade: atividades em endpoints, comportamento de ransomware, eventos de autenticação e anomalias de rede. As fontes principais foram:

(1) EDR (Endpoint Detection and Response): registros de sensores instalados em estações Windows corporativas, fornecendo informações granulares sobre a cadeia de processos e execução de binários. Os eventos incluem artefatos maliciosos como `GetDockVer32W.exe`, execuções suspeitas por `svchost.exe` fora do contexto legítimo, técnicas como DLL Sideloading e o uso de ferramentas como `Ammy.exe`. A execução de comandos capturada pelos campos `Command Line`, `Parent Command Line` e `Grandparent Command Line`, juntamente com os campos Tactic e Technique mapeados no *MITRE ATT&CK*, permite reconstruir a linha temporal da infecção e identificar táticas de evasão e persistência com base em análise comportamental e de hash.

(2) Logs de ransomware (RBMWS): eventos que fazem parte das fases preparatórias de ataques de ransomware, incluindo exclusão de backups via `wbadmin.exe`, execução remota com `PsExec` e sabotagem de restauração mediante exclusão de `Volume Shadow Copies`. Esses registros foram empregados para treinar a identificação de estágios de ataque e avaliar a resposta a incidentes em tempo real, simulando fases avançadas de impacto (MITRE: Inhibit System Recovery).

(3) Eventos de autenticação (ALS-D): registros extraídos de dispositivos de borda e firewalls, focados em tentativas de login, mudanças de contexto de segurança e padrões anômalos de acesso. Esses dados fornecem visibilidade de movimentação lateral e *footprinting*, forçando os participantes a correlacionarem acessos originados de máquinas comprometidas.

(4) Eventos de rede (EAP): registros provenientes de firewalls de aplicações, documentando requisições HTTP anômalas, como *Server-Side Request Forgery (SSRF)*, enumeração de recursos e manipulação de cabeçalhos. Esses eventos são utilizados para treinar a análise de tráfego e a identificação de padrões de exfiltração, técnicas de disfarce (como spoofing de user agents) e evasão por meio de proxies.

O processamento inicial após a extração dos logs da ferramenta de SIEM foi realizado por um script em Python desenvolvido para converter arquivos CSV brutos para o formato JSON. Essa conversão possibilitou a ingestão estruturada via pipelines do Logstash, utilizando parsing baseado em *grok* e enriquecimento seletivo de campos. Cada linha do CSV foi transformada em um dicionário JSON, mantendo os nomes das colunas como chaves, o que é indispensável para indexação semântica no Elasticsearch e para a construção de dashboards legíveis no Kibana. Cada linha representa um evento independente, ideal para simulação cronológica no ambiente dos desafios.

Para garantir a integridade dos dados sensíveis e o cumprimento das diretrizes de privacidade exigidas tanto por normas internas da organização parceira quanto pela

legislação vigente (como a Lei Geral de Proteção de Dados Pessoais – LGPD), são aplicados critérios rigorosos de anonimização e consistência antes da ingestão dos logs na plataforma HIKARI com finalidade não apenas proteger informações identificáveis, mas também assegurar que os dados mantenham sua veracidade técnica e valor analítico, permitindo a simulação realista de ameaças em ambiente seguro e ético.

A anonimização foi operacionalizada por meio de uma automação customizada desenvolvida em Python, com o objetivo específico de varrer os *datasets* de logs com base em uma lista de palavras-chave sensíveis. Cada termo identificado foi substituído por uma nomenclatura padronizada da plataforma HIKARI (ex: `USR_GEN_001`, `DOMAIN_HKR`, `IP_PRIV_HKR`), isto garante que o conteúdo semântico do evento seja preservado, mas desvinculado de qualquer informação real capaz de comprometer a confidencialidade organizacional ou pessoal. Campos como contas de usuários nominais, contas privilegiadas e administrativas, de termos institucionais e produtos proprietários, domínios e URLs foram mapeados para rótulos fictícios sem alterar o tipo de comportamento associado no log. Campos como *timestamp*, tipo de evento, portas de rede, método HTTP, códigos de status e outros metadados operacionais foram mantidos inalterados para garantir validade técnica nos desafios permitindo que os logs anonimizados mantenham sua estrutura analítica íntegra sem exposição de ativos ou informações reais da organização parceira.

Com efeito, a fim de refletir cenários enfrentados por equipes de *Blue Team* durante as competições no ambiente HIKARI, esses quatro conjuntos de eventos foram utilizados na elaboração dos seguintes desafios:

(1) Detectando o Scanner Fantasma: (Categoria: Network Scanning) Um scanner fantasma está sondando a rede, tentando descobrir hosts ativos e portas abertas. Os participantes devem identificar o IP do scanner antes que ele interrompa suas atividades. Dica: procurar um IP que aparece consistentemente em múltiplas requisições com características de varredura.

(2) Reconhecimento e Exploração: (Categoria: Web Exploits) Alertas de segurança indicam que um atacante realizou reconhecimento seguido da exploração de vulnerabilidades. A tarefa consiste em descobrir o IP do atacante para capturar a flag. Dica: verificar IPs envolvidos em atividades de reconhecimento e subsequente exploração.

(3) Atividade Suspeita no EDR: (Categoria: EDR Analysis) Uma atividade suspeita foi detectada por sensores EDR em múltiplos sistemas. O desafio é localizar o IP do sistema comprometido para ajudar na mitigação da ameaça. Dica: concentrar-se em IPs associados a eventos críticos que ocorrem repetidamente e indicam possível comprometimento.

(4) Comportamento de Ransomware Detectado: (Categoria: Ransomware Behavior) Sinais de comportamento de ransomware foram detectados em um dos sistemas monitorados. O objetivo é encontrar o IP do sistema afetado e agir rapidamente para impedir o ataque. Dica: buscar atividades indicativas de criptografia em massa ou comportamentos anômalos de arquivos em sistemas específicos. Esses quatro desafios foram utilizados na execução dos três experimentos descritos na seção a seguir.

4. Avaliação Experimental

A plataforma HIKARI foi aplicada em três eventos educacionais independentes mas com perfis complementares. O objetivo central foi avaliar a usabilidade da solução, o realismo dos cenários simulados e a efetividade da interface de análise baseada em eventos reais. Ao todo, participaram mais de 40 estudantes e profissionais, organizados em múltiplos times e distribuídos entre os seguintes contextos: um exercício interno no Laboratório de Comando e Controle de Defesa Cibernética (Lab C2-DC) do ITA; uma competição técnica com pós-graduandos da PUC Minas; e a execução no programa Hackers do Bem, promovido pela RNP. Todos os eventos seguiram o mesmo padrão arquitetural: provisionamento automatizado de instâncias por equipe, desafios sequenciais controlados via CTFd, e análise centralizada em dashboards Kibana. A liberação de eventos foi temporizada por meio de scripts que publicavam logs nos tópicos do Kafka, permitindo simulação incremental da progressão de um incidente real.

A metodologia adotada baseou-se em estudos de caso de natureza exploratória, com observação indireta. Cada instância executou desafios multiestágio, utilizando logs realistas, em ambientes controlados, utilizando contêineres Docker. A avaliação da plataforma foi conduzida em três etapas principais: validação funcional da infraestrutura; execução dos desafios com participantes reais; e coleta e análise dos dados de interação. Cada equipe atuou em uma instância isolada da plataforma, acessando os desafios por meio do CTFd e investigando evidências em dashboards personalizados no Kibana. A coleta de dados foi realizada de forma automatizada, por meio da instrumentação nativa dos serviços utilizados: logs de submissão e progresso no CTFd (incluindo tempo, número de acertos e ordem de resolução); e formulários estruturados compostos por questões objetivas e campos abertos para feedback qualitativo. Toda a infraestrutura foi projetada para ser replicável, permitindo sua reutilização em novos estudos, seja com os mesmos desafios ou com instâncias adaptadas a diferentes perfis de participantes.

A liberação dos desafios seguiu uma ordem progressiva, sincronizada com a injeção controlada de eventos via Kafka. A visibilidade dos logs por estágio foi coordenada por scripts automatizados, evitando exposição antecipada das evidências. Os participantes não tiveram acesso direto à infraestrutura da plataforma, assegurando isonomia e controle experimental. **Limitações.** Como não houve grupo controle, randomização ou acompanhamento longitudinal, os resultados devem ser interpretados como evidência exploratória de viabilidade e aplicabilidade educacional. As análises quantitativas e qualitativas, entretanto, sugerem coerência entre os contextos e indicam o potencial formativo da plataforma.

Participaram da avaliação três grupos distintos, distribuídos em diferentes momentos e contextos institucionais: Estudantes de graduação em Engenharia de Computação do ITA, durante sessão prática no laboratório C2-DC; Estudantes (profissionais da área de Tecnologia da Informação) de um curso de pós-graduação em Cibersegurança da Pontifícia Universidade Católica de Minas Gerais (PUC Minas); Participantes do programa Hackers do Bem, promovido pela RNP. Os testes foram realizados em ambientes controlados, com autenticação individual, isolamento por equipe e sincronização por cronômetro global. O provisionamento completo foi automatizado via contêineres Docker, com um orquestrador central responsável por subir as instâncias de CTFd, Logstash, Elasticsearch e Kibana para cada grupo. Todos os times receberam

acesso simultâneo aos desafios e às ferramentas de investigação analítica.

Durante os exercícios, foram extraídas métricas específicas que contribuíram para a análise da percepção dos participantes em relação a utilização do Hikari. Para avaliação quantitativa e qualitativa, as questões objetivas e os comentários abertos do formulário estruturado foram analisados por agrupamento temático, com foco em aspectos como usabilidade da interface, clareza dos desafios, realismo dos logs e aplicabilidade prática dos conteúdos. Os resultados dessa análise são discutidos na Seção 5.

5. Resultados

O HIKARI foi concebido e aplicado como uma plataforma de capacitação prática voltada à formação de competências em defesa cibernética. Cada edição seguiu uma dinâmica estruturada em desafios do tipo CTF, com atividades destinadas especificamente ao desenvolvimento de habilidades de *Blue Team*. As ações foram simuladas em um ambiente que utilizou o Kibana para análise de logs, detecção de incidentes, correlação de eventos e resposta cibernética. A mecânica da plataforma incorporou o conceito de gamificação progressiva e multiestágio, de modo que novos desafios eram liberados somente após a solução dos anteriores, estimulando a construção gradual do conhecimento técnico. Ao término das atividades, todos os participantes foram convidados a responder uma avaliação da experiência prática da plataforma, a fim de levantar informações que pudessem medir a aderência da plataforma a seus objetivos formativos, bem como identificar melhorias e evoluções pertinentes para sua evolução.

A metodologia de análise dos dados coletados foi pautada na avaliação de temas específicos, como o perfil do usuário, sua experiência com a plataforma, o impacto na aprendizagem e desenvolvimento, a aplicabilidade e utilidade do sistema, a relevância e o realismo dos desafios e o potencial de negócio da solução. A avaliação foi realizada nas três edições: em novembro de 2024, com estudantes de graduação em Engenharia de Computação do ITA durante sessões práticas no laboratório C2-DC; em março de 2025, com estudantes (profissionais da área de Tecnologia da Informação) de um curso de pós-graduação em Cibersegurança da Pontifícia Universidade Católica de Minas Gerais (PUC Minas); e em abril de 2025, com participantes do programa Hackers do Bem promovido pela RNP. Ao todo, foram obtidas 46 respostas únicas.

A análise quantitativa consistiu no tratamento dos dados coletados por meio de questões estruturadas que abordaram, de maneira categórica, o perfil do usuário, a experiência com a plataforma HIKARI, o impacto no aprendizado técnico, a aplicabilidade da ferramenta em cenários reais de treinamento, a relevância dos desafios simulados e o potencial de expansão da solução. Para as avaliações de interface, eficácia de aprendizado, utilidade prática e realismo dos desafios, foi utilizada uma escala ordinal de 1 a 5, em que 5 representava a melhor avaliação. A análise buscou identificar correlações entre o nível de experiência em segurança cibernética dos participantes, sua familiaridade prévia com desafios CTF e técnicas de defesa *(Blue Team)*, e sua percepção sobre a dificuldade e efetividade dos desafios propostos. Além disso, foram analisadas respostas relacionadas à gamificação da plataforma, ao engajamento dos usuários, à aplicabilidade dos desafios para a formação de equipes de *Blue Team* e à recomendação da plataforma como instrumento de treinamento técnico. De forma complementar, as respostas abertas foram submetidas a uma análise temática qualitativa, categorizando

comentários sobre aspectos como usabilidade da interface, dificuldade de navegação e utilização de filtros no Kibana, percepção de realismo dos cenários simulados, e sugestões práticas de aprimoramento da experiência e da dinâmica da plataforma.

Em relação ao perfil dos participantes, constatou-se que a maioria apresentava nível intermediário ou avançado em segurança cibernética. Ainda que grande parte já tivesse experiência prévia em CTFs, as edições mais recentes indicaram uma maior presença de iniciantes. A familiaridade dos participantes com técnicas de defesa cibernética, por sua vez, revelou-se predominantemente moderada a alta. Essa heterogeneidade reforça a versatilidade do HIKARI para públicos de diferentes níveis técnicos, sendo que a abordagem progressiva adotada pela plataforma favoreceu a inclusão de usuários iniciantes sem comprometer a complexidade e a profundidade necessárias para usuários mais experientes.

A experiência dos usuários com a plataforma foi, em geral, bastante positiva. A interface recebeu avaliação elevada, com 85,4% das notas atribuídas variando entre 4 e 5. A dificuldade dos desafios foi considerada adequada, sobretudo por perfis que já possuíam algum grau de familiaridade com atividades de *Blue Team*. A gamificação utilizada foi aprovada por mais da metade dos participantes, contribuindo significativamente para o aumento do engajamento durante as atividades. Quanto ao realismo dos desafios, cerca de 87% dos respondentes os classificaram como realistas ou muito realistas, enquanto mais de 90% consideraram o HIKARI útil para treinamento de equipes de segurança. A eficácia da ferramenta no processo de aprendizado foi igualmente bem avaliada, com uma nota média superior a 4,2 em uma escala de 1 a 5. Entre os problemas técnicos reportados, destacaram-se a interrupção do acesso à plataforma (como no caso do Kibana inoperante) e dificuldades relacionadas à falta de familiaridade com a ferramenta, especialmente quanto ao uso de filtros e queries, indicando uma necessidade de melhoria na curva de aprendizado inicial.

Observou-se um impacto positivo na aprendizagem técnica dos participantes. O HIKARI foi considerado eficaz ou muito eficaz para o aprendizado em defesa cibernética. Observou-se que o maior impacto foi entre os participantes com perfil intermediário, em forte sinergia com o modelo multinível progressivo proposto pela plataforma. As competências mais desenvolvidas pelos usuários envolveram a análise de logs, citada por 92% dos respondentes, seguida pela detecção de intrusão, que variou entre 45% e 85%, análise de ameaças, entre 45% e 58%, e resposta a incidentes, que ficou entre 23% e 25%. Além dessas competências centrais, os participantes também destacaram a aquisição de conhecimentos sobre sintaxe KQL (Kibana Query Language), o manuseio do Kibana e o desenvolvimento de raciocínio analítico orientado à evidência. A estrutura progressiva dos desafios foi fundamental para estimular a aplicação prática contínua e o fortalecimento da capacidade de análise dos participantes.

Quanto à aplicabilidade e utilidade, a plataforma foi considerada muito útil para treinamento prático de *Blue Teams* por quase todos os participantes. Entre as sugestões de evolução mais recorrentes destacam-se a implementação de um modo assistido com dicas progressivas, para apoiar os iniciantes, e a inclusão da integração do Elastic diretamente no HIKARI para otimizar o fluxo de trabalho. A percepção geral reforça a relevância da plataforma não apenas para formação acadêmica, mas também para treinamentos internos em organizações e programas educacionais focados em segurança defensiva.

No que tange à relevância e realismo dos desafios, a maioria dos usuários considerou que os cenários simulados reproduziram com fidelidade a realidade enfrentada por equipes de *Blue Team* em operações reais. Entre os pontos destacados positivamente estão a simulação de múltiplas etapas de ataques, como reconhecimento, exploração e persistência. Algumas sugestões para melhoria incluíram a adição de diversidade de fontes de logs, a simulação de logs em tempo real e a criação de cenários mais desastrosos, ampliando assim a complexidade e variedade das simulações. Em termos de avaliação quantitativa, entre 85% e 92% dos respondentes consideraram os desafios realistas ou parcialmente realistas, enquanto 85% a 90% reconheceram que o HIKARI contribui para a preparação prática de analistas de segurança para incidentes cibernéticos reais.

Finalmente, no que diz respeito ao potencial de negócio, todos os respondentes recomendaram o HIKARI como ferramenta para treinamento de equipes de segurança, e mais de 90% acreditam que a plataforma possui grande potencial para se destacar no mercado, ainda que cerca de 60% a 70% ressaltem a necessidade de melhorias incrementais para alcançar excelência. As sugestões mais recorrentes para a evolução do produto incluíram o desenvolvimento de modos assistidos, a criação de rankings transparentes, a inclusão de dicas progressivas e a possibilidade de trabalhar com logs gerados em tempo real. De maneira geral, a plataforma demonstra elevado potencial estratégico para ser escalada e replicada em múltiplos setores da indústria e educação.

5.1. Critérios de Sucesso

A avaliação experimental da plataforma HIKARI considerou como critério de sucesso sua capacidade de cumprir os objetivos formativos definidos, que a diferenciam dos ambientes discutidos na Seção 2, com foco em realismo, progressão controlada e integração com ferramentas de análise compatíveis com SIEM. Participaram 46 usuários, que responderam a formulários com questões objetivas e abertas. Os resultados indicaram mais de 85% de concordância quanto à utilidade e ao realismo dos desafios, com nota média superior a 4,2 (em uma escala de 1 a 5) sobre aprendizado técnico. Comentários abertos destacaram a sequência lógica dos desafios, a clareza das evidências e a dinâmica de progressão. Esses aspectos correspondem diretamente às funcionalidades previstas na arquitetura da plataforma.

6. Conclusão e Trabalhos Futuros

Este artigo apresentou a arquitetura, execução e avaliação da plataforma HIKARI, concebida para o treinamento técnico de equipes *Blue Team* em ambientes realistas. A solução integra desafios gamificados no formato CTF, geração controlada de eventos, análise de evidências com a *ELK stack* e orquestração temporal via Kafka. A plataforma foi testada em três contextos educacionais distintos — ITA, PUC Minas e Hackers do Bem — com mais de 40 participantes no total, demonstrando sua viabilidade técnica, escalabilidade e aplicabilidade prática em formações orientadas à defesa cibernética. Os resultados indicam que o HIKARI é capaz de simular cenários realistas de resposta a incidentes com progressão controlada e correlação temporal de eventos, oferecendo uma alternativa robusta às abordagens centradas exclusivamente em técnicas ofensivas. A integração com ferramentas amplamente adotadas no mercado, a infraestrutura baseada em contêineres e a ausência de configurações manuais por parte dos usuários favorecem sua adoção ampla em contextos acadêmicos, corporativos e governamentais.

Como direções futuras, pretende-se expandir o conjunto de desafios disponíveis, incorporar mecanismos de geração dinâmica de logs e oferecer módulos de orientação adaptados ao perfil dos participantes. Além disso, sereá introduzida uma análise quantitativa baseada número de tentativas por desafio; tempo médio de resolução por fase; distribuição de flags corretas e incorretas (taxa de acertos); e um índice *"Log-Query Efficiency"* derivado de pesquisas KQL corretas/total. Desta maneira, através de informações sobre a navegação e consultas realizadas no Kibana será possível inferir padrões de exploração das evidências. Também está em avaliação a integração de funcionalidades baseadas em Modelos de Linguagem de Grande Escala (LLMs), com o objetivo de apoiar a análise semântica de logs e sugerir queries de investigação. A adoção de notebooks Jupyter para fins didáticos e o uso de Kubernetes para viabilizar escalabilidade horizontal também estão previstos. Essas funcionalidades ainda não integram a implementação atual e são consideradas extensões planejadas para versões futuras da plataforma.

Neste contexto, a literatura recente discute a possibilidade de agentes baseados em LLMs conduzirem experimentos científicos de forma autônoma, como exemplificado pelo AI Scientist [Lu et al. 2024]. Apesar do potencial desses sistemas para acelerar tarefas analíticas e gerar hipóteses, ainda persistem limitações quanto à confiabilidade, compreensão contextual e capacidade de operar em cenários realistas de segurança cibernética. O HIKARI propõe uma abordagem complementar, na qual analistas humanos interagem com logs reais, ferramentas industriais e desafios progressivos em ambientes gamificados, preservando o raciocínio interpretativo e o aprendizado incremental. Futuramente, a integração de LLMs poderá atuar como suporte auxiliar — por exemplo, na sugestão de queries ou explicações — sem substituir a experiência imersiva e formativa centrada no usuário.

Agradecimentos

Os autores agradecem ao programa Hackers do Bem, à Rede Nacional de Ensino e Pesquisa (RNP) e ao Instituto Tecnológico de Aeronáutica (ITA) pelo suporte institucional, técnico e logístico na concepção, implementação e validação da plataforma HIKARI. Agradecem também aos participantes e avaliadores pelos feedbacks fornecidos durante a execução dos testes realizados em 2024 e no início de 2025.

Referências

Badva, P., Ramokapane, K. M., Pantano, E., and Rashid, A. (2024). Unveiling the Hunter-Gatherers: Exploring threat hunting practices and challenges in cyber defense. In *33rd USENIX Security Symposium (USENIX Security 24)*, pages 3313–3330, Philadelphia, PA. USENIX Association.

Chetwyn, R. A., Eian, M., and Jøsang, A. (2024). Modelling indicators of behaviour for cyber threat hunting via sysmon. In *Proceedings of the 2024 European Interdisciplinary Cybersecurity Conference*, EICC '24, page 95–104, New York, NY, USA. Association for Computing Machinery.

DeCusatis, C., Alvarico, E., and Dirahoui, O. (2022). Gamification of cybersecurity training. In *Proceedings of the 1st International Workshop on Gamification of Software Development, Verification, and Validation*, Gamify 2022, page 10–13, New York, NY, USA. Association for Computing Machinery.

Diakoumakos, J., Chaskos, E., Kolokotronis, N., and Lepouras, G. (2021). Cyber-range federation and cyber-security games: A gamification scoring model. In *2021 IEEE International Conference on Cyber Security and Resilience (CSR)*, pages 186–191.

Fortinet (2024). 2024 cybersecurity skills gap global research report. Available at: `https://www.fortinet.com/content/dam/fortinet/assets/reports/2024-cybersecurity-skills-gap-report.pdf` (Acceso em: Abril 2025).

Gough, C., Mann, C., Ficke, C., Namukasa, M., Carroll, M., and OConnor, T. (2024). Remote controlled cyber: Toward engaging and educating a diverse cybersecurity workforce. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, SIGCSE 2024, page 394–400, New York, NY, USA. Association for Computing Machinery.

Karagiannis, S., Ntantogian, C., Magkos, E., Ribeiro, L. L., and Campos, L. (2021). Pocketctf: A fully featured approach for hosting portable attack and defense cybersecurity exercises. *Information*, 12(8).

Kianpour, M., Kowalski, S., Zoto, E., Frantz, C., and Øverby, H. (2019). Designing serious games for cyber ranges: A socio-technical approach. In *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pages 85–93.

Kuchar, K., Blazek, P., and Fujdiak, R. (2024). From playground to battleground: Cyber range training for industrial cybersecurity education. In *Proceedings of the 2023 13th International Conference on Communication and Network Security*, ICCNS '23, page 209–214, New York, NY, USA. Association for Computing Machinery.

Leune, K. and Petrilli, S. J. (2017). Using capture-the-flag to enhance the effectiveness of cybersecurity education. In *Proceedings of the 18th Annual Conference on Information Technology Education*, SIGITE '17, page 47–52, New York, NY, USA. Association for Computing Machinery.

Lu, C., Lu, C., Lange, R. T., Foerster, J., Clune, J., and Ha, D. (2024). The ai scientist: Towards fully automated open-ended scientific discovery. *arXiv preprint arXiv:2408.06292*.

Russo, E., Ribaudo, M., Orlich, A., Longo, G., and Armando, A. (2023). Cyber range and cyber defense exercises: Gamification meets university students. In *Proceedings of the 2nd International Workshop on Gamification in Software Development, Verification, and Validation*, Gamify 2023, page 29–37, New York, NY, USA. Association for Computing Machinery.

Savin, G. M., Asseri, A., Dykstra, J., Goohs, J., Melaragno, A., and Casey, W. (2023). Battle ground: Data collection and labeling of ctf games to understand human cyber operators. In *Proceedings of the 16th Cyber Security Experimentation and Test Workshop*, CSET '23, page 32–40, New York, NY, USA. Association for Computing Machinery.

Švábenskỳ, V., Čeleda, P., Vykopal, J., and Brišáková, S. (2021). Cybersecurity knowledge and skills taught in capture the flag challenges. *Computers & Security*, 102:102154.

Vasilakis, M., Karampidis, K., Tampouratzis, M., Malamos, A., Panagiotakis, S., and Papadourakis, G. (2024). Enhancing industry 4.0 cybersecurity training through cyber range platform. In *2024 5th International Conference in Electronic Engineering, Information Technology & Education (EEITE)*, pages 1–6.

Zhong, C., Kim, J. B. J. B., and Liu, H. (2024). The art of inclusive gamification in cybersecurity training. *IEEE Security & Privacy*, 22(5):40–51.