v7.n10.2026

REVISTA CIENTÍFICA - RECIMA21 ISSN 2675-6218

ENGENHARIA SOCIAL NA PRÁTICA: DESENVOLVENDO UM COLETOR DE MÉTRICAS DE VULNERABILIDADE HUMANA PARA SUSCETIBILIDADE DE USUÁRIOS A ATAQUES DE PHISHING

PRACTICAL SOCIAL ENGINEERING: DEVELOPING A HUMAN VULNERABILITY METRICS COLLECTOR FOR USER SUSCEPTIBILITY TO PHISHING ATTACKS

INGENIERÍA SOCIAL EN LA PRÁCTICA: DESARROLLO DE UN RECOLECTOR DE MÉTRICAS DE VULNERABILIDAD HUMANA PARA LA SUSCEPTIBILIDAD DE USUARIOS A ATAQUES DE PHISHING

Felipe Falcai[1], Felipe Diniz Dallilo[2]

e7109193
https://doi.org/10.47820/recima21.v7i10.9193

PUBLICADO: 10/2026

RESUMO

A cibersegurança contemporânea enfrenta desafios que transcendem as barreiras tecnológicas, destacando-se a engenharia social como vetor crítico de invasões. Este trabalho objetiva desenvolver um coletor de métricas de vulnerabilidade humana via simulação *web* de *phishing*, quantificando a suscetibilidade dos usuários. A pesquisa exploratória e descritiva adotou abordagem qualiquantitativa fundamentada em um processo sistemático de engenharia de *software*. O simulador cliente-servidor foi construído em *Python* e *Flask*, ancorado em um repositório relacional robusto para garantir a integridade atômica das interações. A coleta ocorreu de forma orgânica, dispensando consentimento prévio para preservar a autenticidade comportamental das vítimas. Para assegurar a integridade ética, o sistema operou sob o princípio de minimização de dados: aplicou-se ofuscação imediata dos endereços IP via *hash* criptográfico e realizou-se um esclarecimento educativo instantâneo (*debriefing*) na própria interface. Em sete dias, a campanha alcançou 1.483 visualizações e 267 capturas válidas (18% de conversão). Os resultados atestam a eficácia da manipulação psicológica e demonstram como o vazamento do IP expõe a rede a riscos como Negação de Serviço (DDoS) e *spear phishing*. Conclui-se que mensurar empiricamente o comportamento humano é essencial para fundamentar investimentos em letramento digital e aprimorar políticas de conscientização corporativa.

PALAVRAS-CHAVE: Cibersegurança. Engenharia social. *Phishing*. Simulação. Vulnerabilidade humana.

ABSTRACT

Contemporary cybersecurity faces challenges that transcend technological barriers, with social engineering standing out as a critical invasion vector. This study aims to develop a human vulnerability metrics collector via a web-based phishing simulation, quantifying user susceptibility. The exploratory and descriptive research adopted a qualitative-quantitative approach grounded in a systematic software engineering process. The client-server simulator was built in Python and Flask, anchored in a robust relational repository to ensure the atomic integrity of interactions. Data collection occurred organically, dispensing with prior consent to preserve the behavioral authenticity of the victims. To ensure ethical integrity, the system operated under the principle of data minimization: immediate obfuscation of IP addresses via cryptographic hash was applied, and instantaneous educational clarification (debriefing) was performed directly on the interface. In seven days, the campaign reached 1,483 views and 267 valid captures (18% conversion rate).

[1] Graduando do Curso de Engenharia de Computação da Universidade de Araraquara – UNIARA. Araraquara-SP.
[2] Docente do Curso de Engenharia de Computação da Universidade de Araraquara – UNIARA. Araraquara-SP.

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 results attest to the effectiveness of psychological manipulation and demonstrate how IP leakage exposes the network to risks such as Distributed Denial of Service (DDoS) and spear phishing. It is concluded that empirically measuring human behavior is essential to justify investments in digital literacy and improve corporate awareness policies.

KEYWORDS: *Cybersecurity. Human vulnerability. Phishing. Simulation. Social engineering.*

RESUMEN

La ciberseguridad contemporánea se enfrenta a desafíos que trascienden las barreras tecnológicas, destacando la ingeniería social como un vector crítico de invasiones. Este trabajo tiene como objetivo desarrollar un recolector de métricas de vulnerabilidad humana a través de una simulación web de phishing, cuantificando la susceptibilidad de los usuarios. La investigación exploratoria y descriptiva adoptó un enfoque cualicuantitativo fundamentado en un proceso sistemático de ingeniería de software. El simulador cliente-servidor fue construido en Python y Flask, anclado en un repositorio relacional robusto para garantizar la integridad atómica de las interacciones. La recopilación ocurrió de forma orgánica, prescindiendo del consentimiento previo para preservar la autenticidad conductual de las víctimas. Para asegurar la integridad ética, el sistema operó bajo el principio de minimización de datos: se aplicó la ofuscación inmediata de las direcciones IP mediante hash criptográfico y se realizó una aclaración educativa instantánea (debriefing) en la propia interfaz. En siete días, la campaña alcanzó 1.483 visualizaciones y 267 capturas válidas (18% de conversión). Los resultados confirman la eficacia de la manipulación psicológica y demuestran cómo la filtración de la IP expone la red a riesgos como la Denegación de Servicio (DDoS) y el spear phishing. Se concluye que medir empíricamente el comportamiento humano es esencial para fundamentar inversiones en alfabetización digital y mejorar las políticas de concienciación corporativa.

PALABRAS CLAVE: *Ciberseguridad. Ingeniería social. Phishing. Simulación. Vulnerabilidad humana.*

INTRODUÇÃO

A área da cibersegurança tem assumido um papel de suma importância para empresas e usuários, os quais utilizam diariamente sistemas computacionais e a *internet*, lidando frequentemente com dados e informações confidenciais. Nesse contexto, cibercriminosos desenvolvem estratégias para atacar indivíduos com o intuito não apenas de obter acesso a esses dados, mas também de comprometer sistemas de naturezas diversas, desde redes pessoais até infraestruturas governamentais.

Contudo, para obter sucesso em uma invasão, os atacantes nem sempre necessitam de técnicas tecnológicas sofisticadas. Muitas vezes, os próprios usuários facilitam o acesso não autorizado por meio de ações simples, como o clique em um *link* suspeito recebido via *e-mail* ou outro canal de comunicação. Essa prática ilustra o *phishing*, termo utilizado na segurança da informação para designar um tipo de ataque amplamente difundido. Apesar de sua simplicidade técnica, o *phishing* apresenta alta eficácia por estar frequentemente associado à engenharia

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.

social, uma tática de manipulação psicológica que explora a confiança das pessoas e as induz a tomar atitudes que comprometem a segurança de seus sistemas.

O fator humano é, indiscutivelmente, o elo mais fraco na cadeia de segurança da informação. Essa realidade exige a adoção de estratégias defensivas que ultrapassem as tradicionais barreiras tecnológicas. Afinal, apenas configurar *firewalls* e antivírus não tem se mostrado suficiente. Para mitigar os riscos reais da engenharia social, o primeiro passo é compreender e quantificar a vulnerabilidade dos usuários diante dessas táticas de manipulação. Nesse cenário, evidencia-se a necessidade do desenvolvimento de ferramentas capazes de avaliar, de forma prática e mensurável, a suscetibilidade de indivíduos a ataques de *phishing*. Assim, este trabalho propõe o desenvolvimento de um coletor de métricas de vulnerabilidade humana, com o propósito de mapear o comportamento dos usuários diante de ameaças simuladas e fornecer dados concretos para a estruturação de treinamentos e políticas de conscientização mais assertivas em cibersegurança.

O objetivo deste trabalho é desenvolver um coletor de métricas, por meio de uma aplicação *web* de simulação de *phishing*, para avaliar, quantificar e demonstrar a suscetibilidade de usuários a ataques de engenharia social. Os objetivos específicos são fundamentar os conceitos de engenharia social e *phishing*, de modo a evidenciar a manipulação psicológica como o principal vetor de exploração de falhas humanas na segurança da informação.

A partir dessa base teórica, pretende-se implementar a infraestrutura do simulador para atrair e registrar as interações dos alvos. O foco prático do trabalho é coletar e armazenar de forma segura, utilizando criptografia, os endereços IP daqueles que interagirem com a simulação. Com a obtenção desses dados, busca-se demonstrar as consequências técnicas e os riscos reais a que um indivíduo estaria exposto em um ataque autêntico, correlacionando o vazamento do IP a ameaças como rastreamento de geolocalização, ataques de negação de serviço (DDoS) e campanhas de *spear phishing*. Por fim, a pesquisa visa analisar as métricas finais obtidas pela ferramenta para comprovar a eficácia do ataque simulado e ressaltar a necessidade de políticas de conscientização corporativa e pessoal em cibersegurança.

A crescente digitalização das atividades cotidianas e a dependência tecnológica por parte de usuários, empresas e instituições governamentais ampliaram drasticamente a superfície de exposição a ataques cibernéticos. Nesse cenário, a justificativa para este projeto pauta-se na urgência de proteger essas infraestruturas contra ameaças que exploram a falha humana. Muitas vezes, invasões de proporções catastróficas não ocorrem devido a vulnerabilidades complexas de *softwares*, mas sim porque um único indivíduo, por desconhecimento ou desatenção, foi vítima de engenharia social. Assim, o tema apresenta ampla relevância e abrangência, pois a

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.

segurança da informação deixou de ser um problema focado estritamente nos departamentos de tecnologia e passou a ser uma necessidade de toda a comunidade conectada.

Sob o aspecto da importância social e corporativa, a pesquisa se justifica ao atuar diretamente na raiz do problema: a falta de letramento digital e a dificuldade em mensurar o nível real de desatenção das pessoas. É exatamente nesse ponto que se destaca a necessidade de criação do coletor de métricas de vulnerabilidade humana. Em cibersegurança, não é possível mitigar adequadamente um risco que não pode ser medido. O desenvolvimento dessa ferramenta justifica-se pela capacidade de transformar a suscetibilidade humana – um fator até então subjetivo – em dados empíricos e quantificáveis. Ao mapear o comportamento dos usuários diante de simulações e registrar a exposição de dados sensíveis, como o endereço IP, o coletor fornece insumos concretos para que as organizações substituam abordagens genéricas por treinamentos de conscientização direcionados e baseados em evidências reais.

Do ponto de vista acadêmico e científico, o trabalho inova ao materializar a ameaça, tirando o risco do campo puramente teórico. Mostrar ao usuário que o simples clique em um *link* pode resultar no rastreamento de sua geolocalização ou servir de vetor para ataques de negação de serviço (DDoS) gera um impacto cognitivo superior aos métodos tradicionais de ensino. Dessa forma, a pesquisa contribui com informações valiosas para o avanço dos estudos em segurança da informação, provando que a conscientização fundamentada na análise de métricas reais e na vivência de uma ameaça simulada é o caminho mais promissor para a adoção de boas práticas preventivas.

Apesar dos constantes avanços na implementação de barreiras tecnológicas de proteção, os incidentes de segurança da informação continuam a crescer, impulsionados majoritariamente por ataques que exploram a vulnerabilidade humana. A dificuldade das organizações reside na incapacidade de mensurar com exatidão o quão suscetíveis seus usuários estão às técnicas de manipulação psicológica antes que uma invasão real ocorra. A questão de pesquisa a ser investigada é: de que maneira o desenvolvimento e a aplicação de um coletor de métricas, baseado em simulações de *phishing*, pode evidenciar o nível de suscetibilidade dos usuários e contribuir para a mitigação dos riscos associados à engenharia social?

A hipótese é que a utilização de uma ferramenta prática de simulação permitiria não apenas quantificar a falha humana, mas também tangibilizar o risco. Parte-se da premissa de que demonstrar ao usuário a facilidade com que dados críticos – como o endereço IP criptografado coletado no simulador – podem ser capturados e correlacionados a ameaças reais (como rastreamento e ataques de negação de serviço) gera um impacto educacional superior ao

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.

das abordagens puramente teóricas. Acredita-se, portanto, que a obtenção dessas métricas empíricas fornecerá os subsídios necessários para a formulação de campanhas de conscientização altamente direcionadas, resultando na redução da vulnerabilidade humana frente a cibercriminosos.

Quanto à sua natureza, a pesquisa divide-se em exploratória e descritiva. O viés exploratório ganha destaque ao investigar a suscetibilidade humana diante de ataques de engenharia social, buscando entender as falhas cognitivas exploradas por cibercriminosos. Simultaneamente, o estudo apresenta caráter descritivo ao mapear, registrar e expor de forma métrica o comportamento dos alvos durante a simulação. A abordagem adotada é qualiquantitativa, aplicando a vertente quantitativa na análise estatística dos acessos e a qualitativa na interpretação das implicações técnicas de cada vazamento.

A condução prática desta pesquisa fundamentou-se em um modelo de coleta orgânica e não controlada, caracterizando-se pela ausência absoluta de identificação dos participantes. O universo de pesquisa abrangeu usuários da internet em geral. Para a composição da amostra, adotou-se o método não probabilístico por conveniência e a técnica de bola de neve, disponibilizando o *link* da simulação de forma pública. Essa característica metodológica impossibilitou o conhecimento prévio ou posterior sobre os indivíduos que visualizaram ou interagiram com o artefato, garantindo o anonimato estrutural dos sujeitos.

A dispensa de um termo de consentimento prévio justificou-se pela própria natureza do fenômeno investigado. A eficácia da engenharia social baseia-se na desatenção e na confiança do usuário em seu ambiente digital. O aviso antecipado sobre a realização de um teste de *phishing* alteraria artificialmente o estado de alerta do indivíduo, invalidando a medição empírica da vulnerabilidade. Para capturar o comportamento autêntico, a interação ocorreu em um cenário de normalidade, eliminando o viés de observação.

Para assegurar a integridade moral e mitigar riscos de exposição dos envolvidos, a arquitetura do simulador operou sob o princípio da minimização de dados. O sistema não solicitou ou armazenou informações pessoais, nomes ou credenciais. O único metadado interceptado para fins estatísticos — o endereço IP — sofreu ofuscação imediata por meio de funções de *hash* criptográfico unidirecional diretamente no *back-end*. O banco de dados consolidou apenas sequências alfanuméricas irreversíveis, tornando tecnicamente impossível o rastreamento ou a individualização das vítimas.

O procedimento de esclarecimento e mitigação de danos ocorreu de forma instantânea e automatizada. O usuário que clicava no *link* simulado era redirecionado imediatamente para uma página educativa contida na própria aplicação. Este ambiente revelava a natureza

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.

inofensiva do ataque, explicava o conceito de *phishing* e assegurava que nenhum dado sensível havia sido comprometido. Essa abordagem substituiu o contato posterior direto — inviável devido ao anonimato das capturas — e cumpriu o papel ético de letramento digital no exato momento da falha, sem gerar constrangimento ou exposição perante terceiros.

1. REFERENCIAL TEÓRICO

A sofisticação dos ataques cibernéticos transformou a segurança digital em uma urgência social e corporativa. Hoje, as ameaças que exploram o fator humano representam o maior desafio para a proteção de dados. Para embasar o desenvolvimento deste simulador, a pesquisa foi estruturada em três pilares teóricos principais: engenharia social, *phishing* e o peso da vulnerabilidade humana na cibersegurança.

1.1. Engenharia social

No cenário tecnológico, a engenharia social atua como uma técnica de manipulação psicológica. O objetivo do invasor é claro: induzir o indivíduo a cometer erros. Isso inclui burlar protocolos de segurança ou entregar informações sigilosas de forma voluntária. Na prática, o atacante explora a ingenuidade e a confiança das pessoas, transformando a mente do usuário no alvo mais acessível de um sistema interligado.

1.1.1. A engenharia social no contexto de aplicações *web*

A engenharia social moderna vai muito além das tradicionais ligações telefônicas fraudulentas. Ela se apoia fortemente em infraestruturas digitais, utilizando interfaces e páginas forjadas para viabilizar os golpes. Stuttard e Pinto (2011) deixam claro que as defesas convencionais de rede perdem a força quando o vetor de ataque é a própria aplicação *web* que o indivíduo está utilizando. A interface tecnológica vira, então, a ponte por onde a falha humana é tecnicamente explorada:

> Uma segunda tendência proeminente tem sido uma mudança gradual de atenção dos ataques com o lado do servidor da aplicação para aqueles que têm como alvo os usuários da aplicação. Esse último tipo de ataque ainda explora defeitos na própria aplicação, mas geralmente envolve algum tipo de interação com outro usuário para comprometer as relações desse usuário com a aplicação vulnerável. (STUTTARD; PINTO, 2011, p. 14, tradução nossa).

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.

1.2. *Phishing*

Entre as táticas de engenharia social, o *phishing* se destaca pela sua escala e periculosidade. Tonezer et al. (2024) definem esse vetor de ataque pelo uso primordial de *e-mails*, SMS e páginas *web* fraudulentas criadas especificamente para enganar as vítimas. O sucesso dessa tática depende intimamente da capacidade do atacante de simular um ambiente de confiança. Sob o disfarce de uma comunicação legítima, os criminosos camuflam intenções maliciosas, como o roubo de credenciais ou o rastreamento silencioso de endereços IP.

1.3. Vulnerabilidade humana na segurança da informação

Entender como os atacantes operam exige, obrigatoriamente, compreender por que os usuários caem nesses golpes. Historicamente, o mercado focou seus esforços na defesa exclusiva de *hardware* e *software*. Sêmola (2014) rompe com essa visão ao argumentar que as vulnerabilidades de um sistema não são puramente técnicas. Elas envolvem aspectos físicos, naturais e, sobretudo, humanos. Quando somamos a falta de letramento digital à baixíssima percepção de risco imediato, o usuário acaba sendo posicionado no epicentro dos incidentes cibernéticos.

1.3.1. Conscientização e simulação de ameaças

Para combater essas falhas comportamentais, as corporações precisam de políticas de conscientização eficientes. Pena *et al.* (2025) reforçam que a melhor forma de prevenir danos críticos é avaliar as equipes na prática. Isso significa expor os funcionários a estímulos maliciosos artificiais, como simulações controladas de *phishing*. Fica evidente, portanto, a lacuna nítida entre a proteção perimetral instalada e o comportamento real do usuário. É justamente essa lacuna que valida a necessidade de se desenvolver um coletor de métricas capaz de transformar a suscetibilidade humana em dados tangíveis.

1.4. Endereçamento IP e a exposição do usuário

Quando um alvo clica em um *link* de *phishing*, ele imediatamente expõe uma informação crucial: o seu endereço IP (*Internet Protocol*). Kurose e Ross (2014) definem o IP como o identificador lógico exclusivo atribuído a cada dispositivo conectado. Sendo um elemento de 32 bits na arquitetura IPv4, é ele quem garante a comunicação de ponta a ponta na *internet* global.

Na perspectiva da engenharia social, capturar o IP da vítima transforma um simples erro de julgamento em uma vulnerabilidade técnica grave. Esse dado abre caminho para consequências reais, permitindo o rastreamento da geolocalização do usuário, o mapeamento

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.

de falhas na sua rede local e a execução de ataques direcionados, como a Negação de Serviço (DDoS).

1.5. *Hash* criptográfico e o armazenamento seguro de dados

Coletar endereços IP para fins de pesquisa exige responsabilidade e mecanismos rígidos de proteção. Para garantir a privacidade absoluta dos usuários testados na simulação, a adoção do *hash* criptográfico é indispensável. Conforme explicam Tanenbaum e Wetherall (2011), uma função de *hash* é um algoritmo matemático que recebe um dado de qualquer tamanho e o converte em uma saída de tamanho fixo, conhecida como resumo da mensagem.

A força essencial dessa técnica é a sua via de mão única: é computacionalmente inviável fazer o caminho inverso e descobrir o dado original a partir da saída gerada (TANENBAUM; WETHERALL, 2011). Ao aplicar essas funções criptográficas no *back-end* do coletor, assegura-se que os endereços IP dos participantes sejam contabilizados estatisticamente de forma completamente ilegível e anônima, respeitando com rigor a ética na pesquisa científica.

1.6. A escalada do ataque: ISPs, DDoS e *spear phishing*

A captura de um endereço IP isolado desencadeia uma cascata de vulnerabilidades secundárias. No âmbito da infraestrutura, os Provedores de Serviços de *Internet* (ISPs - *Internet Service Providers*) desempenham um papel passivo nesse rastreamento. Ao obter o IP, o cibercriminoso pode consultar os blocos de roteamento do ISP e determinar a localização geográfica aproximada do alvo, transformando a rede da vítima em um ponto estático e localizável na *internet*.

Com o alvo mapeado, o invasor ganha a capacidade de orquestrar ataques de Negação de Serviço Distribuída (DDoS). Conforme aponta Stallings (2014), a negação de serviço tem como objetivo inibir ou impedir o uso e o gerenciamento normal das instalações de comunicação. Na prática, o cibercriminoso utiliza o endereço IP descoberto para perturbar a rede da vítima, sobrecarregando-a com um volume massivo de mensagens e tráfego malicioso a fim de degradar e paralisar o seu desempenho.

Paralelamente à ameaça estrutural contra a rede, o fator humano volta a ser explorado de forma agravada através do *spear phishing*. Hadnagy (2011) detalha que a eficácia desse tipo de engenharia social exige um processo minucioso de coleta de informações (*information gathering*) sobre os alvos. Em vez de disparar mensagens genéricas, o cibercriminoso utiliza os dados obtidos na primeira interação — como a confirmação do IP, o perfil de navegação e a geolocalização — para construir pretextos altamente personalizados e convincentes. O objetivo

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.

dessa manipulação direcionada é forçar a vítima a executar ações prejudiciais, como a instalação silenciosa de *malwares*, garantindo o acesso e a escalada de privilégios na infraestrutura comprometida sem levantar suspeitas.

1.7. Engenharia de *software* e metodologia

A construção técnica do simulador exigiu diretrizes bem definidas para assegurar a estabilidade do sistema. Pressman e Maxim (2021) defendem que a adoção de um processo de software estabelece um roteiro sistemático, retirando o desenvolvimento do campo do improviso empírico e garantindo a organização das tarefas. O ciclo deste projeto seguiu essa premissa. A equipe especificou os requisitos de captura, modelou a arquitetura cliente-servidor e executou a validação funcional dos pacotes de segurança locais antes de homologar a implantação definitiva na nuvem.

1.8. Linguagem e arquitetura de rede

A aplicação back-end operou com base na linguagem Python. A escolha técnica justificou-se pela alta eficiência dessa ferramenta no processamento de expressões lógicas e numéricas. Perkovic (2016) detalha que a sintaxe simplificada e a avaliação dinâmica de variáveis no Python permitem a criação de rotinas complexas de forma legível e com alto desempenho. O código-fonte resolveu a geração das funções de hash e a triagem algébrica das requisições *web* em milissegundos, evitando qualquer gargalo no servidor.

No controle do tráfego *web* e na resposta aos usuários, o simulador adotou a arquitetura cliente-servidor por meio do framework Flask hospedado em nuvem. Coulouris *et al*. (2013) definem esse modelo arquitetônico pela separação direta de funções, em que processos clientes interagem com servidores remotos para acessar recursos compartilhados e centralizados. Aplicando esse conceito à engenharia social, o navegador da vítima opera como o cliente que solicita a visualização da página, enquanto o servidor gerencia a lógica de captura nos bastidores e devolve a interface falsa. Essa divisão de responsabilidades atendeu ao modelo teórico e garantiu o registro dos acessos de forma totalmente oculta ao alvo.

1.9. Arquitetura de banco de dados

A retenção segura das métricas de conversão demandou um repositório relacional robusto. Silberschatz, Korth e Sudarshan (2020) explicam que sistemas gerenciadores de banco de dados solucionam problemas críticos de atomicidade e previnem anomalias de acesso concorrente presentes em modelos de gravação comuns. A aplicação dessas propriedades na

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.

arquitetura do simulador garantiu o processamento exato dos registros durante o pico de interações simultâneas, ocorrido nas primeiras 48 horas da campanha. Essa estrutura relacional evitou inconsistências e manteve a base de dados íntegra para a avaliação estatística.

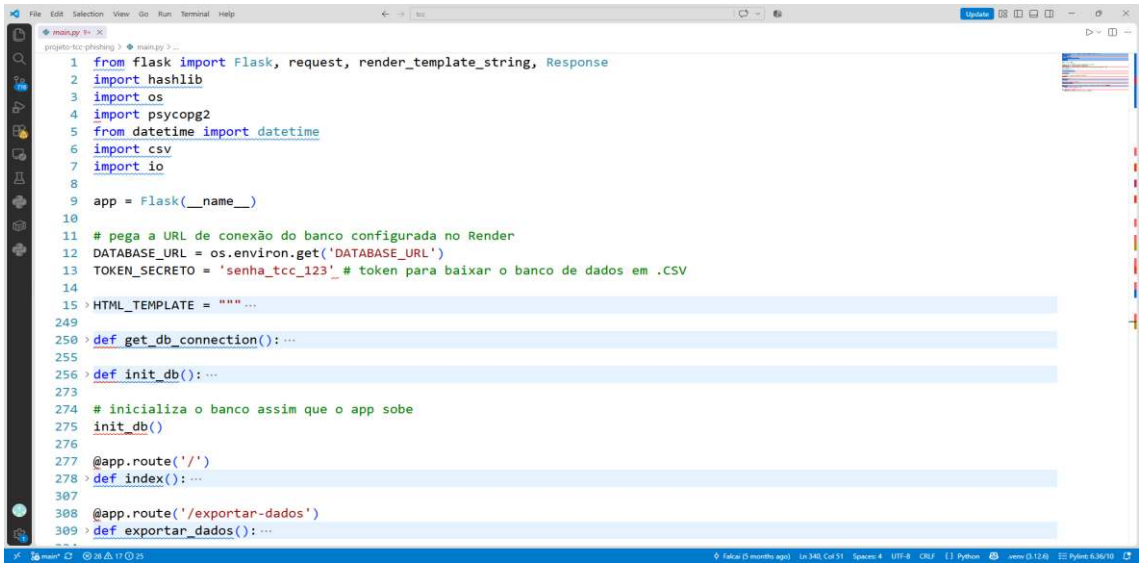
2. METODOLOGIA

Esta seção detalha a construção, a implementação e a hospedagem da aplicação *web* focada na simulação de ataques de *phishing*. O objetivo é evidenciar as escolhas tecnológicas e a arquitetura do simulador, demonstrando na prática como os mecanismos de coleta e armazenamento anônimo de dados funcionaram.

2.1. Ambiente de desenvolvimento e codificação

Para construir a lógica da aplicação e o servidor *back-end*, a escolha recaiu sobre o *Visual Studio Code* (*VSCode*). A IDE ofereceu a flexibilidade necessária para integrar *scripts* em *Python* com o *framework Flask*. Na prática, o *VSCode* centralizou a estruturação do projeto: desde a codificação das rotas de rede até a implementação das funções criptográficas de *hash* e a construção das páginas estáticas. A Figura 1 ilustra a interface do ambiente de desenvolvimento durante a fase de codificação do projeto.

Figura 1. Ambiente de desenvolvimento *Visual Studio Code (VSCode)*.



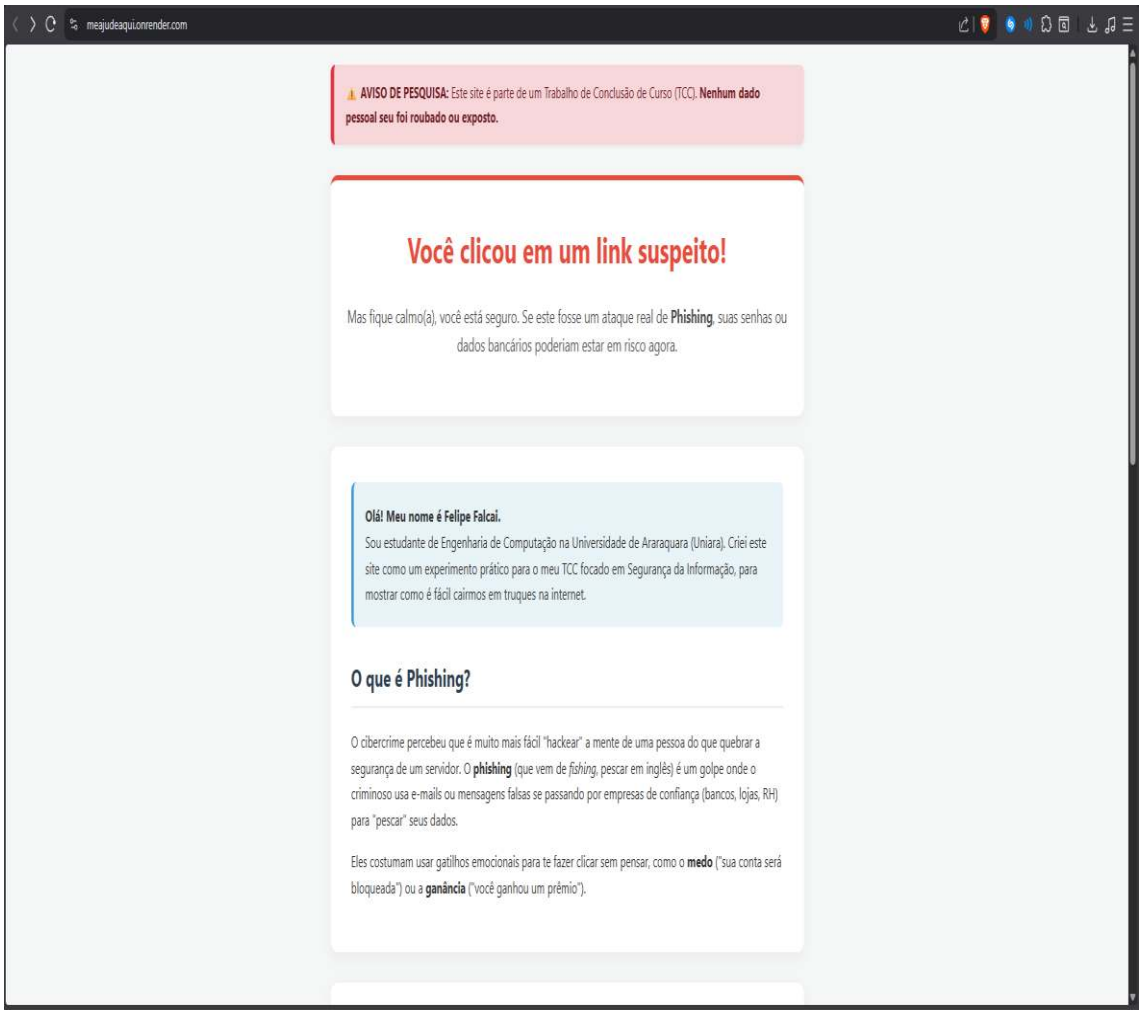
Fonte: Os autores, 2026.

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.

2.2. Interface da aplicação simuladora

O sucesso de qualquer ataque de engenharia social depende inteiramente de quão real o cenário aparenta ser. Por isso, projetou-se a interface do usuário (*front-end*) para emular um ambiente virtual autêntico. O objetivo era fazer o alvo interagir naturalmente com a página. É exatamente no momento dessa interação que o sistema age nos bastidores, capturando os metadados da conexão da vítima para posterior criptografia. A Figura 2 demonstra o simulador em funcionamento, evidenciando a tela com a qual os usuários testados interagiram.

Figura 2. Interface *web* da aplicação simuladora de *phishing*



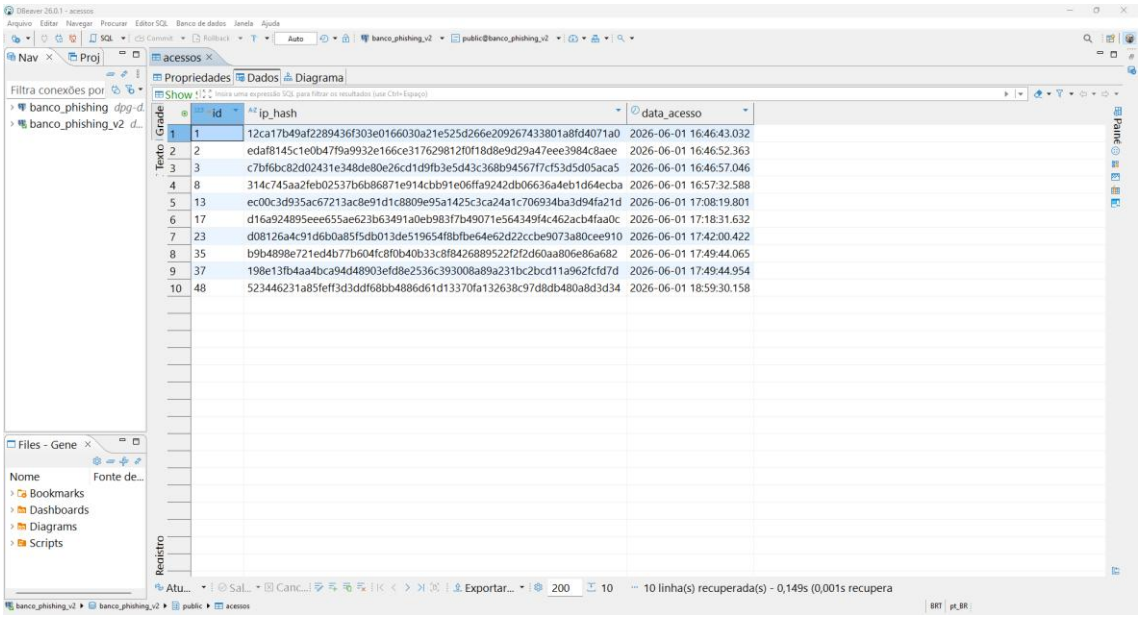
Fonte: Os autores, 2026.

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.

2.3. Gerenciamento do banco de dados

Mensurar a falha humana exige um repositório seguro e ágil. O *software* DBeaver cumpriu esse papel como administrador do banco de dados relacional do projeto. A ferramenta facilitou a modelagem das tabelas, a execução de consultas SQL e a validação visual do processo de anonimização. Isso garantiu que todos os endereços IP capturados passassem pelas funções de *hash* antes de serem persistidos, respeitando os critérios éticos da pesquisa. A Figura 3 exibe a interface do DBeaver operando a base de dados do simulador.

Figura 3. Gerenciamento do banco de dados da aplicação por meio da ferramenta DBeaver



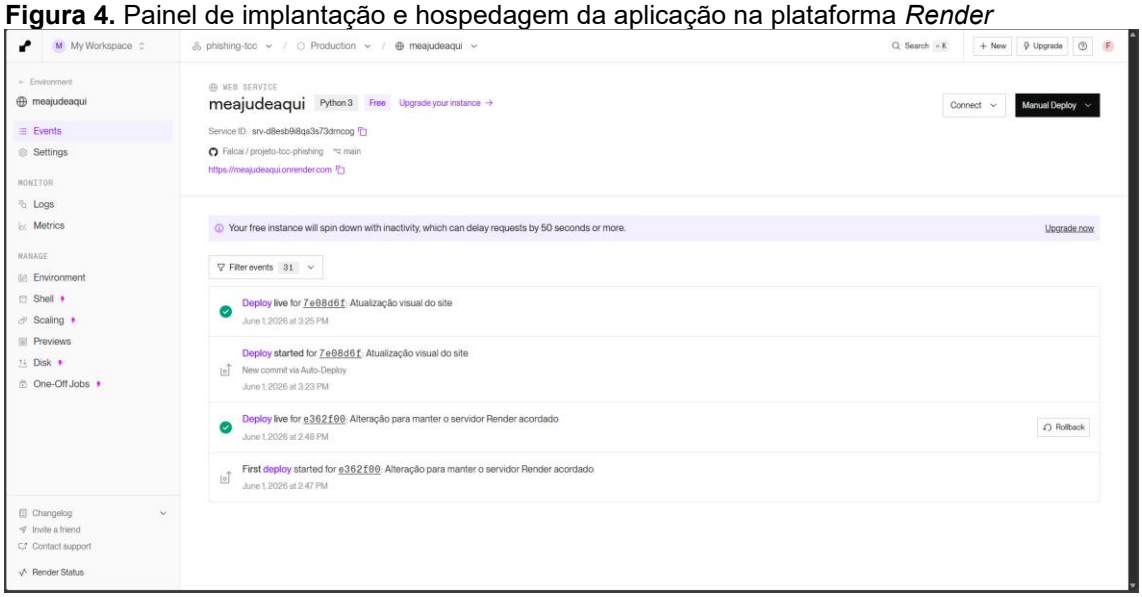
Fonte: Os autores, 2026.

2.4. Implantação e hospedagem em nuvem

A metodologia da pesquisa exigia um simulador público e acessível, viabilizando a amostragem por conveniência na *internet*. Para colocar a aplicação no ar, a plataforma *Render* serviu como provedora de Nuvem na modalidade Plataforma como Serviço (PaaS). A escolha permitiu manter o servidor *Flask* operando de forma contínua e distribuir o *link* de forma segura aos alvos. A Figura 4 apresenta o painel de controle do Render, validando o estado ativo da aplicação na *web*.

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 4. Painel de implantação e hospedagem da aplicação na plataforma *Render*



Fonte: Os autores, 2026.

3. RESULTADOS E DISCUSSÃO

A fase de testes práticos validou a arquitetura do simulador e gerou a base de dados necessária para avaliar a suscetibilidade humana. Durante os sete dias de operação (de 09 a 16 de setembro de 2026), o sistema rodou de forma ininterrupta na nuvem, coletando as interações orgânicas dos usuários expostos à ameaça simulada.

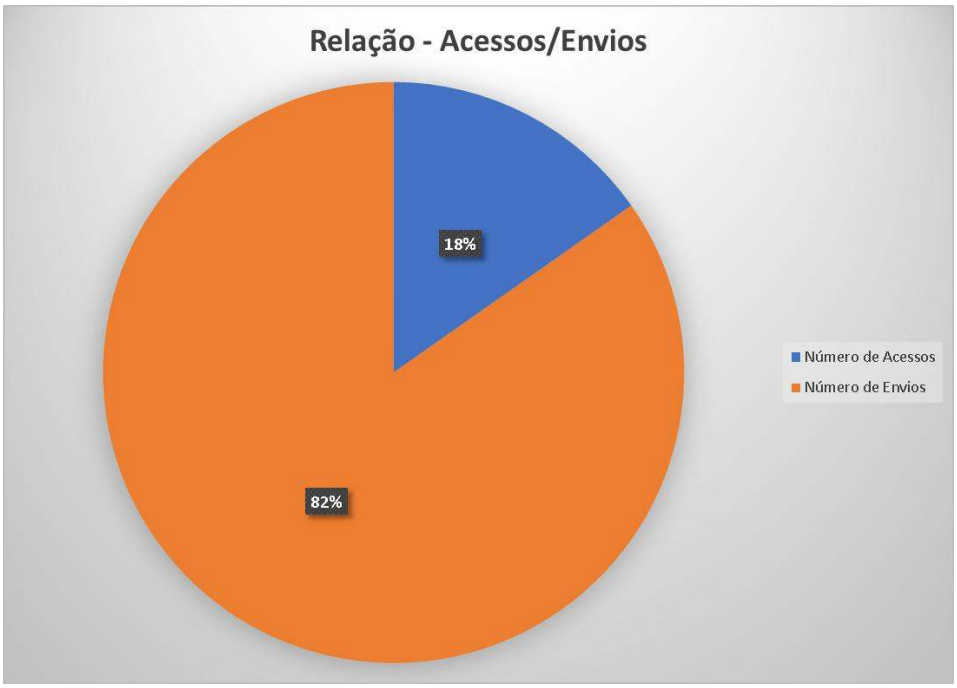
3.1. Consolidação das métricas de acesso

A disseminação do *link* forjado ocorreu exclusivamente de forma orgânica, atingindo um alcance estimado de 1.483 visualizações em aplicativos de mensagens e redes sociais. No *back-end* da aplicação, o algoritmo de triagem trabalhou para filtrar requisições incompletas e varreduras automatizadas feitas por bots. O saldo final no banco de dados contabilizou 267 interações reais e únicas.

Cruzando o volume de acessos com os cliques registrados, a taxa de conversão (*Click-Through Rate* – CTR) atingiu 18%. Traduzindo o indicador para a realidade do teste: praticamente um em cada cinco usuários baixou a guarda, ignorou o senso crítico e entregou seus metadados de rede ao servidor.

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 5. Relação entre acessos válidos e visualizações da campanha simulada



Fonte: Os autores, 2026.

3.2. Comportamento da amostra ao longo do tempo

A análise cronológica dos registros, extraída diretamente das tabelas do *DBeaver*, revelou um padrão comportamental nítido. O maior volume de vítimas caiu na armadilha nas primeiras 48 horas da campanha. Esse pico inicial prova na prática a eficácia do "senso de urgência". Quando a engenharia social simula uma situação limite ou uma oferta imperdível, o alvo reduz o tempo de raciocínio lógico e clica por instinto.

Outro detalhe técnico importante surgiu na numeração do banco de dados. Os identificadores primários (coluna id) saltaram organicamente e ultrapassaram o número 310 para entregar os 267 IPs válidos. Essa quebra na sequência numérica atesta o bom funcionamento do filtro da aplicação, que descartou tentativas de cliques repetidos da mesma vítima antes de consolidar a métrica final.

A Tabela 1 apresenta um extrato dos registros validados pelo banco de dados, evidenciando o formato de captura anonimizada e a progressão temporal dos acessos.

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. Extrato dos dados coletados durante a simulação

id	ip_hash	data_acesso
1	0001a2c4b6d8e0f1a3c5b7d9e1f3a5c7b9d9b2d4e6f8a0c1b3d5e7c3b5d7e9f04f8a	2026-09-09 08:14:22.304
2	0002a2c4b6d8e0f1a3c5b7d9e1f3a5c7b9d9b2d4e6f8a0c1b3d5e7c3b5d7e9f0b1c2	2026-09-09 08:21:05.112
3	0003a2c4b6d8e0f1a3c5b7d9e1f3a5c7b9d9b2d4e6f8a0c1b3d5e7c3b5d7e9f0d3e4	2026-09-09 08:45:57.046
5	0005a2c4b6d8e0f1a3c5b7d9e1f3a5c7b9d9b2d4e6f8a0c1b3d5e7c3b5d7e9f0f5a6	2026-09-09 09:12:10.120
...	...	...
313	0139a2c4b6d8e0f1a3c5b7d9e1f3a5c7b9d9b2d4e6f8a0c1b3d5e7c3b5d7e9f09c1d	2026-09-16 18:30:15.796
314	013aa2c4b6d8e0f1a3c5b7d9e1f3a5c7b9d9b2d4e6f8a0c1b3d5e7c3b5d7e9f01e2f	2026-09-16 18:55:50.897

Fonte: Os autores, 2026.

3.3. Discussão das implicações técnicas

Registrar 267 endereços IP válidos em apenas uma semana projeta um cenário de altíssimo risco cibernético. Um endereço IP isolado não exibe o CPF ou o nome civil da vítima. No entanto, ele entrega a chave para ataques em cadeia muito mais severos. Quando um

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.

cibercriminoso obtém essa identificação de rede de um usuário desatento, o cruzamento do IP com bases públicas de ISPs permite mapear a região geográfica do alvo.

A partir dessa localização, o ataque de Negação de Serviço consiste em inundar intencionalmente a largura de banda do alvo com tráfego malicioso, sobrecarregando a rede com mensagens a fim de prejudicar e paralisar o seu desempenho (STALLINGS, 2014). Paralelamente à ameaça estrutural, o fator humano volta a ser explorado de forma agravada através do *spear phishing*. Hadnagy (2011) explica que a eficácia desse ataque exige o cruzamento prévio de informações sobre os alvos. Em vez de disparar mensagens genéricas, o cibercriminoso utiliza os dados obtidos na primeira interação para construir pretextos altamente personalizados e convincentes, forçando a vítima a executar ações maliciosas sem levantar suspeitas.

3.4. Validação empírica da ameaça

O experimento tirou o risco humano do campo teórico. Os dados atestam que as empresas precisam de métricas exatas e palpáveis para justificar investimentos em letramento digital. As barreiras tecnológicas tradicionais, como antivírus e *firewalls*, falham no exato milissegundo em que o próprio usuário decide clicar na ameaça. O coletor cumpriu seu papel ao tangibilizar esse comportamento, transformando uma vulnerabilidade psicológica abstrata em estatística comprovada.

4. CONSIDERAÇÕES FINAIS

A pesquisa atingiu o seu objetivo principal ao desenvolver e implementar um coletor de métricas de vulnerabilidade humana por meio de uma aplicação *web* de simulação de *phishing*. A campanha de testes, que alcançou 1.483 visualizações e registrou 267 interações válidas, consolidou uma taxa de conversão de 18%. Esse resultado responde de forma direta à questão central da pesquisa: a aplicação de um simulador evidencia o alto nível de suscetibilidade dos usuários de forma empírica e contribui para a mitigação de riscos associados à engenharia social ao transformar o comportamento humano — uma variável até então subjetiva — em dados métricos exatos, permitindo intervenções corretivas direcionadas.

O trabalho demonstra sua relevância no contexto atual da cibersegurança ao comprovar que defesas puramente tecnológicas de perímetro são ineficazes quando o atacante manipula os vieses cognitivos do alvo. Em comparação com as abordagens predominantes na literatura existente — que frequentemente avaliam o risco do fator humano de maneira puramente teórica

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.

ou por meio de questionários autoavaliativos —, este projeto apresenta a vantagem de mensurar o comportamento real e instintivo do usuário diante de uma ameaça ativa, garantindo, simultaneamente, a privacidade dos dados coletados por meio de funções de *hash* criptográfico. Como desvantagem, aponta-se a limitação inerente ao método de amostragem por bola de neve e ao curto período de coleta (sete dias), o que restringe a generalização absoluta dos dados para perfis demográficos ou corporativos específicos, diferentemente de estudos de longo prazo conduzidos em ambientes controlados.

As aplicações práticas desta pesquisa são imediatas para o mercado de tecnologia. A arquitetura do simulador construído pode ser adotada por departamentos de segurança da informação e equipes de *Red Team* para auditar o nível de letramento digital de colaboradores. A demonstração de como um simples clique resulta no vazamento do endereço IP conscientiza as corporações sobre a porta de entrada para ataques subsequentes, como o rastreamento, a negação de serviço (DDoS) e o *spear phishing*.

Por fim, para o desenvolvimento de pesquisas futuras, recomenda-se a expansão das capacidades do simulador para testar a inserção de credenciais fictícias em páginas de autenticação clonadas, elevando a complexidade da métrica além da simples coleta de IP. Adicionalmente, o desenvolvimento de um painel de controle (*dashboard*) dinâmico para a análise comportamental em tempo real, cruzando dados como o sistema operacional da vítima e o tempo de permanência na página forjada, agregará ainda mais valor analítico e preditivo à ferramenta.

REFERÊNCIAS

COULOURIS, George; DOLLIMORE, Jean; KINDBERG, Tim; BLAIR, Gordon. Modelos de Sistema. In: COULOURIS, George *et al*. **Sistemas distribuídos**: conceitos e projeto. 5. ed. Porto Alegre: Bookman, 2013. cap. 2, p. 37-79.

HADNAGY, Christopher. Information Gathering. In: HADNAGY, Christopher. **Social Engineering**: The Art of Human Hacking. Indianapolis: Wiley Publishing, 2011. cap. 2, p. 39-57.

KUROSE, James F.; ROSS, Keith W. A Camada de Rede. In: KUROSE, James F.; ROSS, Keith W. **Redes de computadores e a internet**: uma abordagem top-down. 6. ed. São Paulo: Pearson Education do Brasil, 2014. cap. 4, p. 224-320.

PENA, S. A.; GONÇALVES FILHO, C.; GALIZA, J. P. L. de M.; MARQUES, J. S. Estratégias de políticas públicas no combate aos crimes cibernéticos: uma análise crítica. **Revista do Instituto Brasileiro de Segurança Pública (RIBSP)**, São Paulo, v. 7, n. 18, p. 27-42, 2025. Disponível em: https://revista.ibsp.org.br. Acesso em: 1 abr. 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.

PERKOVIC, Ljubomir. Expressões, Variáveis e Atribuições. In: PERKOVIC, Ljubomir. **Introdução à computação usando Python**: um foco no desenvolvimento de aplicações. Rio de Janeiro: LTC, 2016. cap. 2, p. 24-51.

PRESSMAN, Roger S.; MAXIM, Bruce R. Processos de Software. In: PRESSMAN, Roger S.; MAXIM, Bruce R. **Engenharia de software**: uma abordagem profissional. 9. ed. Porto Alegre: AMGH, 2021. cap. 2, p. 38-63.

SÊMOLA, Marcos. Segurança da Informação. In: SÊMOLA, Marcos. **Gestão da Segurança da Informação**: uma visão executiva. 2. ed. Rio de Janeiro: Campus, 2014. cap. 4, p. 41-73.

SILBERSCHATZ, Abraham; KORTH, Henry F.; SUDARSHAN, S. Introdução. In: SILBERSCHATZ, Abraham; KORTH, Henry F.; SUDARSHAN, S. **Sistema de banco de dados**. 7. ed. Rio de Janeiro: LTC, 2020. cap. 1, p. 1-32.

STALLINGS, William. Introdução. In: STALLINGS, William. **Criptografia e segurança de redes**: princípios e práticas. 6. ed. São Paulo: Pearson Education do Brasil, 2014. cap. 1, p. 5-19.

STUTTARD, Dafydd; PINTO, Marcus. Web Application (In)security. In: STUTTARD, Dafydd; PINTO, Marcus. **The Web Application Hacker's Handbook**: Finding and Exploiting Security Flaws. 2. ed. Indianapolis: John Wiley & Sons, 2011. cap. 1, p. 1-15.

TANENBAUM, Andrew S.; WETHERALL, David. Segurança de redes. In: TANENBAUM, Andrew S.; WETHERALL, David. **Redes de Computadores**. 5. ed. São Paulo: Pearson Prentice Hall, 2011. cap. 8, p. 479-548.

TONEZER, L. N.; SILVA, A. C. M.; ALMEIDA, A. H. de; NEVES, J. E. D. A. Simulações multiagentes e phishing: explorando a segurança em ambientes de nuvem. **Revista de Tecnologia da Fatec de Americana (RTec)**, Americana, v. 11, n. 2, p. 1-17, 2024. Disponível em: https://fatec.edu.br/revista/index.php/RTecFatecAM. Acesso em: 1 abr. 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.