FATOR HUMANO: O ELO MAIS FRACO DA SEGURANÇA DA INFORMAÇÃO

Leonardo Carvalho Naves de Lima[1]
Carlos Eduardo de França Roland[2]

Resumo

Em decorrência dos avanços tecnológicos ocorridos nas últimas décadas, o fator humano tornou-se o elo mais frágil na segurança da informação, sendo frequentemente explorado por cibercriminosos por meio de técnicas de engenharia social. Neste artigo é apresentado um estudo, baseado em revisão bibliográfica sistemática, realizado sobre como as vulnerabilidades comportamentais, como por exemplo, a propensão ao erro, e a falta de conhecimento técnico, são amplamente exploradas em ataques de engenharia social como o *phishing*, *impersonation,* e pretextos, que utilizam princípios psicológicos como reciprocidade, autoridade e escassez para manipular as vítimas. Como resultante do estudo, são sugeridas estratégias para mitigar esses riscos, incluindo educação continuada, políticas de segurança e a implementação de controles técnicos. O estudo visou reforçar a necessidade de uma abordagem multifacetada, combinando tecnologia, conscientização e cultura organizacional, para que ocorra a transformação do fator humano de um ponto fraco em um pilar central da segurança da informação.

Palavras-chave: Cibersegurança. Engenharia social. Fator humano. Segurança da informação. Vulnerabilidade.

Abstract

Due to technological advances in recent decades, the human factor has become the weakest link in information security being often exploited by cybercriminals through social engineering techniques. This paper presents a study, based on a systematic literature review, on how behavioral vulnerabilities, such as tendency to make mistakes, and lack of technical knowledge, are widely exploited in social engineering attacks such as phishing, impersonation. Other vulnerabilities are justifications which use psychological principles such as reciprocity, authority, and scarcity to manipulate victims. As a result of the study, strategies to mitigate these risks are suggested, including continuing education, security policies, and the implementation of technical controls. The study aimed to reinforce the need for a multifaceted approach, combining technology, awareness, and organizational culture, so that the human factor can be transformed from a weak point into a central pillar of information security.

Keywords: *Cybersecurity. Social engineering Human factor. Information security.. Vulnerability.*

1 Introdução

[1] Graduando em Análise e Desenvolvimento de Sistemas pela Fatec Dr Thomaz Novelino – Franca/SP. Endereço eletrônico: leonavesdelima@yahoo.com.br.
[2] Docente em Análise e Desenvolvimento de Sistemas na Fatec Dr Thomaz Novelino – Franca/SP. Endereço eletrônico: carlos.roland@fatec.sp.gov.br.

A expansão acelerada das Tecnologias Digitais de Informação e Comunicação (TDIC) nas últimas décadas transformou fundamentalmente as interações sociais e econômicas globais. Paralelamente a essa digitalização, observa-se um crescimento exponencial das ameaças cibernéticas, com projeções indicando custos globais de US$ 10,5 trilhões até 2025 (Cybercrime Magazine, 2023), posicionando o cibercrime como a terceira maior economia mundial.

Apesar dos significativos avanços tecnológicos em segurança da informação, as vulnerabilidades humanas persistem como o principal vetor de ataques cibernéticos. Estudos internacionais demonstram que técnicas de engenharia social exploram sistematicamente falhas comportamentais, utilizando princípios psicológicos para contornar medidas de proteção técnica. Contudo, verifica-se uma lacuna científica significativa na literatura brasileira sobre a aplicação desses princípios psicológicos em ataques de engenharia social no contexto nacional.

A escassez de pesquisas nacionais sobre o fator humano na cibersegurança contrasta com a urgência crescente do tema, um estudo realizado pela DataSenado apontou que 24% dos brasileiros já foram vítimas de crimes digitais no ano de 2024 (DataSenado, 2024). Esta carência de estudos científicos compromete o desenvolvimento de estratégias de mitigação culturalmente adaptadas às características comportamentais e sociais da população brasileira.

Neste contexto, torna-se imperativo investigar como os princípios psicológicos de influência são explorados em ataques de engenharia social, com vistas ao desenvolvimento de estratégias eficazes de proteção. O presente estudo objetiva analisar as vulnerabilidades comportamentais humanas na segurança da informação, identificar os principais princípios psicológicos explorados por cibercriminosos e propor estratégias multifacetadas para transformar o fator humano de elemento vulnerável em pilar defensivo organizacional.

A relevância desta pesquisa justifica-se pela necessidade de fundamentação científica para políticas de segurança da informação que considerem as especificidades comportamentais humanas, contribuindo para a redução dos custos econômicos e sociais decorrentes de ataques cibernéticos no contexto brasileiro.

2 Referencial teórico e trabalhos correlatos

A engenharia social na segurança da informação constitui um conjunto de técnicas de manipulação psicológica direcionadas à obtenção não autorizada de

informações confidenciais. Mitnick e Simon (2003, p. 2) definem engenharia social como o uso de "influência e persuasão para enganar as pessoas, convencendo-as de que o engenheiro social é alguém que ele não é, ou por meio de manipulação", permitindo ao atacante "se aproveitar das pessoas para obter informações com ou sem o uso de tecnologia".

Esta conceituação fundamenta-se na exploração sistemática de vulnerabilidades comportamentais humanas, contrastando com ataques puramente tecnológicos que visam falhas em sistemas computacionais. Hadnagy (2010) complementa essa definição ao caracterizar a engenharia social como "a arte e ciência de fazer com que as pessoas façam o que você quer que façam", evidenciando a dimensão psicológica central dessas técnicas.

A eficácia da engenharia social reside na exploração de tendências comportamentais universais, independentemente do nível de conhecimento técnico da vítima. Cialdini (2007) identifica que essas técnicas aproveitam "atalhos mentais" ou heurísticas que os indivíduos utilizam para tomar decisões rápidas em situações de incerteza.

2.1 Taxonomia de Ataques de Engenharia Social

A literatura científica estabelece diversas classificações para ataques de engenharia social, baseadas em vetores de ataque, objetivos e técnicas de manipulação empregadas:

2.1.1. *Phishing* e suas variedades

Dentre os diversos tipos de ataques de engenharia social um dos mais comuns, senão o mais, é o ataque *phishing* e suas variantes. O termo *phishing* tem sua origem da palavra em inglês *fishing* que significa pesca, fazendo a alusão à ação de lançar uma vara e aguardar com que mordam a isca, para se pescar dados pessoais da vítima. O ataque *phishing* é um ataque que utiliza técnicas de engenharia social com o objetivo de enganar as vítimas e induzi-las a fornecer informações ou dados sensíveis. Normalmente os criminosos buscam credenciais de acesso, dados pessoais, números de cartões de crédito, senhas e documentos.

Esse é um dos ataques mais comuns, especialmente por meio de e-mails e redes sociais. Nesse caso os golpistas enviam mensagens que parecem ser de grandes empresas, mas que, na verdade, têm o objetivo de coletar dados.

Jakobsson e Myers (2007) classificam as variantes de *phishing* conforme especificidade do alvo e canais empregados:

Phishing: esse ataque consiste em enganar vítimas por meio de e-mails ou mensagens fraudulentas que pareçam ser originadas de fontes confiáveis. Normalmente o ataque contém links para direcionar a vítima a sites falsos, que se assemelham muito ou são idênticos aos sites originais, onde serão coletados dados confidenciais da vítima.

Spear Phishing: esse ataque é uma versão direcionada e personalizada que visa atingir públicos ou corporações específicas. Nesse tipo de ataque os criminosos realizam pesquisas para conhecerem a vítima com o intuito de criar mensagens convincentes para aumentar as chances de sucesso.

Whaling: assim como o *Spear Phishing* esse é um ataque direcionado, que tem como alvo especificamente pessoas de alto perfil dentro de organizações como diretores, gerentes e executivos. Darrel Burrel afirma que ataques do tipo *whaling* são direcionados para explorar o viés de autoridade e a confiança organizacional, tendo como alvo indivíduos de alto escalão para obter acesso a dados confidenciais (Burrel, 2024).

Clone Phishing: nesse tipo é utilizada uma réplica de um e-mail legítimo que a vítima tenha recebido anteriormente, no qual o criminoso substitui os *links* para direcionar a vítima para sites falsos ou anexos que eram legítimos por versões maliciosas. Em razão da mensagem possuir uma aparência familiar a vítima tem uma maior propensão em clicar nos endereços eletrônicos ou anexos.

Vishing: nesse tipo de *phishing*, os ataques ocorrem por meio de mensagens de voz ou ligações telefônicas.

> Nesse ataque o engenheiro social utilizará de vias telefônicas para entrar em contato com a vítima. O atacante poderá utilizar de artimanhas como alterar a identificação do número que está ligando. Esse ataque está mais relacionado para o alvo fornecer credenciais ou informações de contas bancárias, dentre outras (Santos, 2016, p. 35).

Smishing: similar ao *vishing*, esse ocorre por meio de mensagens de texto, geralmente por SMS ou outros aplicativos de mensagem, contendo links ou solicitando para que a vítima informe dados pessoais.

2.1.2. *Impersonation* e Pretextos

Dentre os ataques de engenharia social que são mais sofisticados e que mais têm apresentado crescimento, tanto em complexidade como em eficácia, são os ataques de *impersonation* e de pretextos. Estes ataques alcançaram grande crescimento em sua eficácia com o surgimento de tecnologias de inteligência artificial.

Os ataques de *pretexting,* ou pretexto, são aqueles onde o criminoso cria uma situação ou cenário fabricado para conseguir convencer a vítima a fornecer dados confidenciais. Neste tipo de ataque é comum que o criminoso finja ser parte do setor de TI da companhia ou de uma prestadora de serviços para solicitar dados como senhas, alegando alguma necessidade para solucionar um problema reportado.

Já o ataque de *impersonation* ocorre quando o atacante tenta se passar por uma pessoa específica como a figura de uma autoridade, um executivo ou alguém conhecido pela vítima, por exemplo, para explorar pilares psicológicos dela.

Um dos ataques que ocorreu em 2020, e que ganhou grande repercussão midiática, foi ao X, antiga rede social Twitter, no qual o cibercriminoso Joseph James O'Connor, conhecido como *PlugWalkJoe*, conseguiu invadir cerca de 130 contas de perfis verificados da rede social. Dentre as contas invadidas estavam as de personalidades como Joe Biden, Barack Obama, e Elon Musk (BIASE, 2023). A invasão ocorreu através de um ataque de engenharia social, onde O'Connor por meio de técnicas de pretexto e *impersonation* conseguiu enganar funcionários do X para fornecer a ele acesso de administrador às ferramentas de gerenciamento.

Os ataques de *impersonation* e pretextos tiveram crescimento em sua eficácia e em sua complexidade nos últimos anos, em razão da popularização das inteligências artificiais. Neste sentido a Verizon apresentou em seu Relatório de Investigações de Violação de Dados de 2024, que nos 2 anos anteriores, 25% dos ataques que tiveram como motivação obter vantagem financeira, envolveram técnicas de pretexto (Verizon, 2024).

2.2. Pilares psicológicos

É conveniente que sejam esclarecidos o que são os pilares psicológicos. São os conceitos fundamentais que explicam como a mente humana pode ser manipulada ou influenciada em determinadas situações. Estes pilares são utilizados por diversas áreas assim como pelos cibercriminosos.

Cialdini (2007) propõe em sua obra que existem 7 princípios de influência psicológica que são popularmente explorados nos contextos de ataques de engenharia social.

2.2.1 Reciprocidade

Este princípio se baseia em uma tendência natural que os seres humanos possuem de retribuir favores, presentes ou concessões recebidas. Este comportamento humano é tão comum que é encontrado em todas as sociedades, e é quase considerado uma regra não escrita, sendo essencial para a convivência e cooperação na sociedade humana. Porém a reciprocidade também pode ser uma ferramenta poderosa para manipular ou convencer alguém a realizar uma determinada ação.

O criminoso oferece à vítima algo de valor, como uma vantagem inicial ou um presente, de tal forma que ela tenha a sensação de obrigação moral de retribuir. Um exemplo desse tipo de ação ocorre em ataques de *phishing*, onde os criminosos costumam enviar e-mails semelhantes aos legítimos oferecendo um presente, como um curso ou uma ferramenta útil, de forma a atrair a atenção da vítima. Porém, ao clicar para receber o presente a vítima acaba fornecendo dados sensíveis ou até mesmo sendo infectado por *malwares* sem perceber.

2.2.2 Compromisso e consistência

O princípio do compromisso e consistência esta correlacionado à necessidade humana de ser coerente com aquilo que já foi dito ou já foi realizado. Ou seja, após tomar uma determinada decisão ou de proferir uma determinada declaração, mesmo que sejam pequenas, tem-se a tendência de seguir a mesma linha de ação, ou do que foi declarado, com o intuito de manter a imagem pessoal consistente.

Organizações utilizam deste princípio de forma positiva, com o intuito de gerar engajamento entre os seus colaboradores por meio de tarefas, como por exemplo a participação em treinamentos, ou outras atividades diárias com o foco em criar um ciclo de consistência.

Porém no campo da cibersegurança, este princípio é comumente utilizado de forma extremamente sutil. Por exemplo, é comum que em um processo de ataque a uma empresa um cibercriminoso inicie sua abordagem solicitando confirmações simples como por um e-mail, ou executando um *link,* ou de dados cadastrais. Ocorre

que após realizar a primeira solicitação, a vítima pode ser levada a realizar compromissos maiores como por exemplo compartilhar informações confidenciais acreditando estar agindo de forma consistente com o cibercriminoso.

Nos ataques que envolvem o princípio do compromisso e consistência é comum também que ocorram abordagens em datas diferentes, com o intuito de gerar uma tendencia maior ao êxito.

Por exemplo, um invasor que busca explorar o acesso físico de um ambiente corporativo pode inicialmente se apresentar como um visitante ou até mesmo um prestador de serviços solicitando algo simples como a autorização para aguardar na recepção ou usar o banheiro. Em outra ocasião ele poderá ir tentando aumentar o acesso como ao uso de uma tomada, ponto de rede ou WIFI com o intuito de coletar dados necessários para um possível ataque.

Esta sequência de pedidos gradualmente mais significativos tende a funcionar, uma vez que como o próprio princípio explica, após atender o primeiro pedido a tendência natural é a de manter a linha comportamental, principalmente se o invasor demonstrar confiança ou familiaridade.

2.2.3 Prova social

A prova social também conhecida como efeito de manada é um comportamento natural que leva as pessoas a seguirem as ações que outras ao seu redor estão realizando. Este comportamento ocorre especialmente quando não há certeza sobre a melhor decisão a tomar. É comum acreditar que, se muitas pessoas estão realizando uma determinada ação, ela deva ser a escolha correta a ser tomada. Este princípio está diretamente relacionado à necessidade de validação e do sentimento de pertencimento.

Um exemplo clássico deste princípio ocorre em campanhas de arrecadação (*crowdfunding*). Quando um projeto arrecadou uma quantia significativa e possui um número significativo de apoiadores, é comum que as pessoas se sintam mais confiantes e confortáveis em contribuir e apoiar o projeto, uma vez que conquistou a validação de muitas pessoas.

Já no contexto da engenharia social o princípio da prova social é amplamente explorado principalmente no que tange a sites falsos. Um exemplo é o de um site falso que exibe inúmeros depoimentos fictícios, um número inflacionado de usuários com

avaliações positivas com o intuito de gerar legitimidade para que usuários forneçam dados pessoais e informações de pagamento.

Um outro exemplo comum ocorre em redes sociais onde criminosos utilizam contas falsas para comentar e reagir em postagens de *phishing*, de forma a simular um grande volume de interações orgânicas, gerando uma falsa validação coletiva para atrair vítimas a clicarem no *link* ou seguir determinadas instruções. Isto ocorre em razão de que a vítima se sente segura uma vez que muitas pessoas supostamente já fizeram aquilo.

2.2.4 Simpatia

Dentre todos os princípios pode-se considerar que a simpatia é um dos mecanismos mais poderosos de persuasão e influência. É inegável que se tenha uma tendência maior a não só confiar em pessoas quando se gosta dela ou sentir que há alguma forma de conexão, mas também se tem a tendência de aceitar as ideias ou os pedidos feitos, com uma maior facilidade. Estas conexões podem ser construídas de diversas formas, por exemplo, a partir de características como atratividade, similaridade, senso de humor ou até mesmo por meio de elogios.

O princípio da simpatia é amplamente utilizado por diversas áreas como o marketing e a comunicação. No marketing é utilizado em campanhas publicitárias pelo emprego da imagem de celebridades, influenciadores ou personalidades carismáticas, uma vez que, conforme demonstrado pelo estudo da agência de branding Matter (2023), 69% dos consumidores tendem a confiar na recomendação de influenciadores, ou seja, as campanhas que possuem figuras carismáticas ou familiares ao público possuem uma taxa de sucesso significativamente maior do que as outras.

Na engenharia social o princípio da simpatia é amplamente utilizado nos ataques de *impersonation* e pretexto, onde é comum que o atacante utilize de perfis falsos em redes sociais para se conectar à vítima, utilizando fotos atrativas, interesses semelhantes e conversas de forma a criar um vínculo de afinidade para que posteriormente possa vir a pedir favores ou explorar a vítima em eventuais crimes.

Casos semelhantes ao do exemplo são comuns e muito fáceis de se encontrar na internet. Com uma busca rápida é possível pesquisar inúmeras reportagens relatando situações do que foram batizados de golpes românticos. Esses golpes consistem em criminosos que constroem um relacionamento de confiança com a

vítima para que estas voluntariamente forneçam acesso à suas contas bancárias ou realizem transferência de dinheiro ao criminoso. Conforme a CNBC (2024), esse tipo de crime tem se tornado cada vez mais frequente, com métodos que exploram as vulnerabilidades emocionais das vítimas. Segundo o relatório da FTC (2024), o prejuízo relacionado a este tipo de golpe chegou a US$ 1,14 bilhões em 2023, evidenciando o impacto deste tipo de crime.

2.2.5 Escassez

Este princípio baseia-se na ideia de que se tem a tendência de atribuir um maior valor a algo quando se percebe que aquilo é raro ou está se esgotando. Em razão disso o comportamento humano pode ser impulsionado pelo medo da perda de uma oportunidade, uma vez que, psicologicamente, seja gerada uma pressão para se agir rapidamente.

Cialdini (2007, p. 179) afirma que "as oportunidades parecem mais valiosas quando sua disponibilidade é limitada". Este conceito é explorado por diversas áreas em anúncios de promoções. É comum serem encontrados anúncios com frases como: Últimas unidades; Oferta limitada; Edição limitada. Essas propagandas têm o intuito de criar um senso de urgência de modo a fazer com que os consumidores ajam impulsivamente para obter o produto antes que ele acabe.

No âmbito da cibersegurança não é diferente, este princípio é amplamente explorado, principalmente no que tange a fraudes envolvendo sites de venda on-line e ataques de engenharia social. Golpes por meio de SMS ou mensagens em aplicativos de comunicação com frases: Clique aqui para resgatar seu prêmio; Sua CNH está prestes a ser bloqueada, clique aqui para consultar mais informações; ou até mesmo mensagens de alerta de supostas transações ou compras suspeitas que foram realizadas em sua conta bancária ou cartão de crédito, indicando para que a vítima clique em um *link* ou ligue em um suposto canal de suporte ao cliente, estão cada dia mais populares.

2.2.6 Autoridade

O princípio da autoridade é um conceito muito estudado no campo da psicologia social. Define-se este princípio como sendo a tendência humana de obedecer ou seguir ordens ou orientações de figuras que são percebidas como autoridades ou especialistas (Cialdini, 2007, p. 163). Esta inclinação pode ser facilmente percebida

na sociedade atual, uma vez que não são raras situações em que se veem pessoas sendo influenciadas por alguém com prestígio, ou por superior hierárquico, no cotidiano.

A autoridade é comumente associada a títulos acadêmicos, uniformes profissionais, credenciais específicas ou cargos de gestão e liderança em uma organização. Normalmente estes títulos ou cargos estão relacionados à influência do indivíduo que conferem legitimidade e poder a este indivíduo. Estes símbolos ainda podem estar relacionados com indicadores de conhecimento ou competência em determinadas áreas. Em razão disto, a autoridade é frequentemente percebida e legitimada por meio de símbolos externos, que, por sua vez, moldam a maneira como os indivíduos respondem a figuras de poder. Essa dinâmica será mais explorada adiante por meio dos estudos de Milgram e de Zimbardo (Cialdini, 2007).

Um dos estudos mais emblemáticos sobre o princípio da autoridade foi o experimento realizado pelo psicólogo Stanley Milgram na década de 1960. Esse experimento buscava investigar até que ponto indivíduos comuns seriam capazes de obedecer a ordens de uma figura de autoridade, mesmo quando tais ordens implicassem em infringir sofrimento a terceiros (Cialdini, 2007, p. 157).

Para o experimento, Milgram recrutou 40 homens com idades entre 20 e 50 anos de diferentes níveis educacionais e com diferentes ocupações. Os participantes foram apresentados a um voluntário que na realidade era um ator contratado para o experimento. No experimento os participantes teriam o papel de professores, enquanto o ator sempre desempenhava o papel de aluno.

O professor recebia a tarefa de ler pares de palavras ao aluno e, posteriormente, testar sua memória. Para cada resposta errada, o professor deveria aplicar choques elétricos, aumentando progressivamente a carga a cada erro subsequente. Os choques começavam em 15V e poderiam chegar até 450V. Os choques não eram reais, porém o ator que fazia o papel de aluno simulava as reações de dor, incluindo gritos e súplicas para interromper o experimento. Durante todo o experimento os voluntários que faziam o papel de professor eram supervisionados por uma figura de autoridade vestida com um jaleco de laboratório.

Segundo Milgram (1963), cerca de 65% dos participantes continuaram a aplicar choques até o nível máximo de 450V, mesmo diante das súplicas do ator que simulava sofrimento, demonstrando que pessoas comuns podem agir contra seus valores sob ordens de uma autoridade legítima (Blass, 1999).

Em 1971 ocorreu outro estudo que foi notável sobre o assunto, Philip Zimbardo conduziu o Experimento da Prisão de Stanford, onde vários voluntários foram designados aleatoriamente a papéis de guardas ou prisioneiros em uma prisão simulada. No experimento rapidamente os guardas adotaram comportamentos autoritários e, em alguns casos, abusivos, enquanto os prisioneiros mostraram sinais de estresse e de submissão. Como resultado do estudo Zimbardo (1971) evidenciou como a atribuição de autoridade e papéis podem influenciar comportamentos de maneiras profunda e de forma negativa.

No que tange à cibersegurança, o princípio da autoridade costuma ser explorado em diversos tipos de golpes de engenharia social. Normalmente os atacantes utilizam de técnicas de *impersonation* e pretextos para se passarem por figuras de autoridades como executivos de alto escalão ou representantes de determinadas instituições, com o intuito de manipular as vítimas e obter informações confidenciais ou acessos não autorizados a um sistema.

Um caso relatado em 2024, de um golpe milionário, ocorreu na China. Segundo a South China Morning Post (2024) uma companhia multinacional sofreu um golpe de 25,6 milhões de dólares, no qual os atacantes utilizaram de *deepfake* (técnica que utiliza inteligência artificial para criar imagens, sons ou vídeos falsos) para recriar digitalmente o diretor financeiro da empresa, ordenando a realização de transferências de dinheiro durante uma videoconferência. Todos que estavam na reunião, exceto a vítima que realizou as transferências, eram representações falsas de pessoas reais que foram criadas por meio do *deepfake*.

2.3.7 Unidade

Cialdini (2007) apresenta o sétimo princípio, que é a unidade, que foi considerada como um amplificador dos demais princípios, porém com o decorrer do tempo e com a evolução das interações humanas em uma sociedade ultra conectada digitalmente, o autor passou a perceber que a unidade possui características para possuir força independentemente de qualquer outro princípio.

A essência deste princípio reside na ideia de que as pessoas são mais propensas a sofrer influência daqueles com quem compartilham uma identidade ou pertencimento comum. Isto pode ocorrer em diversos contextos, como nacionalidade religião, cultura, causas sociais, ou até por comunidades digitais.

Este princípio tornou-se extremamente relevante em razão da evolução da internet e das redes sociais. Empresas e influenciadores utilizam a unidade para criar comunidades engajadas, de forma a promover um senso de pertencimento entre os participantes.

Assim, tendo por base os marcos conceituais apresentados, foram realizadas pesquisas complementares para identificar suas aplicações, por cibercriminosos, em ataques de segurança da informação.

3 Ferramentas e métodos

Esta seção detalha as ferramentas e os métodos empregados para a coleta e análise das informações referentes ao fator humano na segurança da informação, seguindo protocolos reconhecidos para revisão sistemática da literatura.

A pesquisa foi realizada por meio de uma revisão sistemática, estruturada conforme as diretrizes do protocolo PRISMA, com o objetivo de garantir transparência e rigor metodológico. Inicialmente, definiu-se a pergunta de pesquisa focada nas vulnerabilidades humanas em segurança da informação e nas estratégias de mitigação.

Em seguida, foram estabelecidos critérios claros de inclusão e exclusão: foram selecionados artigos científicos publicados entre 2015 e 2024, disponíveis em português e inglês, que abordassem engenharia social, fator humano e segurança da informação. Foram excluídos trabalhos sem revisão por pares e fontes não acadêmicas, exceto livros e posts em sites reconhecidos do setor, que complementaram a fundamentação teórica.

As buscas foram realizadas nas bases Google Scholar e IEEE Xplore, utilizando palavras-chave como “engenharia social”, “fator humano”, “segurança da informação”, “vulnerabilidades humanas” e “mitigação de riscos”. A seleção dos documentos seguiu etapas de triagem por título e resumo, seguida da leitura integral dos textos que atendiam aos critérios, conforme o fluxo PRISMA.

Além dos artigos, foram incluídos livros relevantes que aprofundam a compreensão das técnicas de manipulação e dos fundamentos psicológicos envolvidos, bem como postagens e relatórios de organizações em portais especializados, que forneceram dados estatísticos atualizados e exemplos práticos. Destacam-se, por exemplo, os relatórios da Verizon (2024) e da Cybercrime Magazine

(2023), além de casos reais como o ataque ao Twitter em 2020 e o golpe de *deepfake* na China em 2024.

A análise dos dados consistiu na categorização das técnicas de engenharia social identificadas, tais como *phishing*, *impersonation* e ataques baseados em princípios psicológicos, permitindo uma análise estruturada e comparativa das abordagens criminosas. Essa metodologia integrada, que combina fontes acadêmicas e do setor, visa garantir a validade e a aplicabilidade dos resultados, contribuindo para um melhor entendimento do papel do fator humano na segurança da informação.

4 Resultados e discussão

Esta seção tem como objetivo apresentar os principais resultados obtidos a partir da análise das fontes pesquisadas, destacando os principais riscos quanto ao fator humano na segurança da informação bem como apresentar estratégias para mitigá-las. Além disso, são discutidas as implicações dessas descobertas para organizações e indivíduos, com foco nas medidas preventivas contra os ataques de engenharia social.

4.1 Principais Resultados

Apesar dos avanços tecnológicos em segurança da informação, o comportamento humano continua sendo o elo mais fraco no que tange à segurança de dados. Estatísticas do relatório da Verizon (2024) indicam que 25% dos ataques financeiros realizados envolvem uma das técnicas de engenharia social, como ataques *impersonation* e de pretexto. Bem como que 68% de todas as violações de dados envolveram um elemento humano.

Por meio da análise do estudo realizado por Cialdini (2007) verificou-se que os sete princípios da influência (reciprocidade, compromisso e consistência, prova social, simpatia, escassez, autoridade e unidade) são constante e amplamente explorados por cibercriminosos para manipular suas vítimas.

Quanto às técnicas de engenharia social mais utilizadas o *phishing* e suas variantes (*spear phishing, whaling e smishing*) lideram como os ataques mais frequentes, representando mais de 80% de todos os incidentes relatados conforme a Cybercrime Magazine (2023).

Já os ataques que envolvem *impersonation* e pretexto estão em ascensão devido ao uso de *deepfake* e IA, como analisado no caso da empresa chinesa que

perdeu US$ 25,6 milhões em um golpe por videoconferência falsa (South China Morning Post, 2024).

Como resultado a Cybercrime Magazine (2023) apresenta a projeção de que o cibercrime custará US$ 10,5 trilhões no ano de 2025, posicionando-se como a terceira maior economia global. Destaca-se que os prejuízos ocasionados em razão dos ataques vão muito além do financeiro uma vez que também causam danos reputacionais além das interrupções operacionais.

4.2 Estratégias de Mitigação

Diante dos riscos e evidências que foram apresentados, torna-se imprescindível a adoção de estratégias multifacetadas para mitigar as vulnerabilidades humanas na segurança da informação. Mitnick e Simon (2003) afirmam que a engenharia social explora as falhas comportamentais, o que exige o emprego de contramedidas que combinem tanto aspectos técnicos quanto aspectos organizacionais e psicológicos.

Como uma das principais formas de defesa contra-ataques baseados em engenharia social tem-se a educação continuada. Esse método consiste em “um processo de aprendizado contínuo cujo propósito é atualizar ou aperfeiçoar as habilidades e os conhecimentos de um indivíduo” (UniCesumar, 2025).

Complementando a educação continuada, é necessária a implementação de programas de conscientização em segurança da informação, com treinamentos periódicos junto de simulações de ataques como uma estratégia para fortalecer a capacidade individual dos usuários em identificar e responder a ameaças.

Além disso é essencial que outras estratégias sejam implementadas. A implementação de políticas de segurança é fundamental para qualquer organização, com diretrizes claras sobre boas práticas como o uso de senhas robustas, uso de autenticação multifator, bem como a verificação cuidadosa de fontes antes de realizar qualquer compartilhamento de informações sensíveis. Essas políticas devem ser divulgadas amplamente e de forma regular com o intuito de garantir a adesão consistente dos colaboradores.

O fortalecimento de controles técnicos desempenha um papel fundamental na segurança da informação. A utilização de ferramentas como filtros de e-mail avançado, sistemas de detecção de intrusão bem como o uso de autenticação em

dois fatores são básicos para a proteção com objetivo de mitigar golpes que visam o roubo de credenciais, uma vez que adicionam camadas extras de proteção.

Outro aspecto relevante para mitigar os riscos mostrados é a promoção de uma cultura organizacional de segurança, onde a proteção da informação seja valorizada em todos os níveis hierárquicos da corporação. Essa cultura deve promover uma comunicação aberta sobre incidentes e reconhecer comportamentos seguros, criando um ambiente em que os colaboradores se sintam responsáveis pela segurança coletiva.

Considerações finais

A revisão sistemática da literatura realizada permitiu mapear o estado atual do conhecimento sobre vulnerabilidades comportamentais humanas na segurança da informação, evidenciando a centralidade do fator humano como vetor de ataques cibernéticos contemporâneos.

Os achados da literatura confirmam que 68% de todas as violações de dados envolvem elementos humanos não maliciosos. A análise de casos emblemáticos, particularmente o ataque ao Twitter (2020) e o golpe *deepfake* em Hong Kong (2024), evidencia a sofisticação crescente das técnicas empregadas e a necessidade de abordagens defensivas multifacetadas.

A convergência identificada entre princípios psicológicos clássicos e técnicas contemporâneas de ataque sugere que estratégias de mitigação eficazes devem integrar educação continuada, controles técnicos e desenvolvimento de cultura organizacional de segurança. A literatura revisada indica que abordagens unidimensionais apresentam eficácia limitada, enquanto estratégias integradas demonstram resultados superiores.

Contudo, a análise crítica revela limitações significativas nos estudos disponíveis. Poucos estudos sobre *phishing* focam efetivamente em usuários através de metodologias rigorosas, evidenciando lacuna metodológica substancial. A evolução acelerada das técnicas de ataque, especialmente com incorporação de inteligência artificial, torna estudos mais antigos potencialmente obsoletos.

A escassez de estudos sobre adaptação cultural dos princípios psicológicos em contextos de cibersegurança constitui lacuna científica significativa identificada nesta

revisão. Esta limitação compromete o desenvolvimento de estratégias de mitigação culturalmente adaptadas

Os resultados desta revisão contribuem para o corpus teórico da área ao sistematizar conhecimentos sobre o fator humano na cibersegurança baseados em evidências científicas verificáveis. Do ponto de vista prático, os achados fundamentam o desenvolvimento de políticas organizacionais de segurança da informação baseadas em pesquisas acadêmicas rigorosas.

A transformação do fator humano de elemento vulnerável em pilar defensivo da segurança da informação representa desafio complexo que demanda investimento sustentado em pesquisa, educação e desenvolvimento de soluções tecnológicas complementares. A efetividade dessa transformação depende fundamentalmente da integração sinérgica entre avanços tecnológicos, compreensão aprofundada de mecanismos psicológicos e desenvolvimento de culturas organizacionais orientadas à segurança da informação.

Referências

BIASE, Nicholas. **UK citizen sentenced to five years in prison for cybercrime offenses.** 2023. Disponível em: https://www.justice.gov/usao-sdny/pr/uk-citizen-sentenced-five-years-prison-cybercrime-offenses. Acesso em: 11.out.2024.

BLASS, T. **The Milgram paradigm after 35 years: some things we now know about obedience to authority.** Journal of Applied Social Psychology, v. 29, n. 5, p. 955–978, 1999. Disponível em: https://doi.org/10.1111/j.1559-1816.1999.tb00134.x. Acesso em: 24.jun.2025.

BURRELL, Darrell Norman. **Exploring the cyberpsychology and criminal psychology of whaling and spear fishing on-line attacks.** RAIS Conference Proceedings 2022-2024, n. 0465, 2024.

CIALDINI, R. B. **Influence: the psychology of persuasion.** New York: Collins, 2007.

CNBC. **Here's how to avoid romance scams, which cost consumers $1.14 billion last year.** 2024. Disponível em: https://www.cnbc.com/2024/07/03/heres-how-to-avoid-romance-scams-which-cost-consumers-1point14-billion-last-year.html. Acesso em: 12.jan.2025.

CYBERCRIME MAGAZINE. **Cybercrime damage is predicted to cost the world $9.5 trillion in 2024.** 2023. Disponível em: https://www.youtube.com/watch?v=MjHO7ghELNs. Acesso em: 18.set.2024.

FTC. FEDERAL TRADE COMMISSION. **Love stinks when a scammer is involved.** 2024. Disponível em: https://www.ftc.gov/business-guidance/blog/2024/02/love-stinks-when-scammer-involved. Acesso em: 12.jan.2025.

HADNAGY, C. **Social engineering: the art of human hacking.** Indianapolis: Wiley, 2010.

JAKOBSSON, M.; MYERS, S. (Ed.). **Phishing and countermeasures: understanding the increasing problem of electronic identity theft.** Hoboken: Wiley-Interscience, 2007.

MATTER. **Consumers continue to seek influencers who keep it real.** 2023. Disponível em: https://www.matternow.com/blog/consumers-seek-influencers-who-keep-it-real/. Acesso em: 12.jan.2025.
MITNICK, Kevin; SIMON, William. **The art of deception: controlling the human element of security.** New York: Wiley, 2003.

SANTOS, Daniel Pitanga dos. **A engenharia social no Brasil e seus riscos.** 2016. 121 f. Trabalho de Conclusão de Curso (Especialização em Gestão de Tecnologia da Informação e Comunicação) – Universidade Tecnológica Federal do Paraná, Curitiba, 2016.

SOUTH CHINA MORNING POST. **'Everyone looked real': multinational firm's Hong Kong office loses HK$200 million after scammers stage Zoom call with deepfake CFO.** 2024. Disponível em: https://www.scmp.com/news/hong-kong/law-and-crime/article/3250851/everyone-looked-real-multinational-firms-hong-kong-office-loses-hk200-million-after-scammers-stage. Acesso em: 12.jan.2025.

UNICESUMAR. **Educação continuada: o que é e quais os formatos?** 2025. Disponível em: https://www.unicesumar.edu.br/blog/educacao-continuada/. Acesso em: 20.abr.2025.

VERIZON. **2024 Data Breach Investigations Report.** 2024. Disponível em: https://www.verizon.com/business/resources/reports/2024-dbir-data-breach-investigations-report.pdf. Acesso em: 18.abr.2025.

ZIMBARDO, Philip. **Stanford Prison Experiment.** Simply Psychology, 2023. Disponível em: https://www.simplypsychology.org/zimbardo.html. Acesso em: 19.jan.2025.