COMMUNICATIONS AND INNOVATIONS GAZETTE

Uma breve análise sobre phishing

Antonio Silverio Montagner[1], **Carla Merkle Westphal**[1]

[1] Departamento de Informática e Estatística
Universidade Federal de Santa Catarina (UFSC)
88.040-900 – Florianópolis – SC – Brasil

antonio.s.montagner@grad.ufsc.br, carla.merkle.westphall@ufsc.br

Abstract. *The number of phishing attacks reached, in 2021, the highest number of all time until then, being tripled since 2020. This event demonstrates that as phishing techniques are shaped considering the context of their target since its emergence in 1995, these attacks adapt over time. Thus, this problem has persisted in society for more than 25 years and continues to increase its number of victims. With that, this work presents a study about what phishing is, its concepts, its types, and the problems caused in society.*

Resumo. *O número de ataques phishing atingiu em 2021 o maior número de todos os tempos até então, sendo o triplo desde 2020. Tal evento demonstra que, desde seu surgimento em 1995, como as técnicas de phishing são moldadas considerando o contexto de seu alvo, esses ataques acabam se adaptando com o tempo e, dessa forma, tal problema persiste na sociedade há mais de 25 anos e continua aumentando seu número de vítimas. Com isso, este trabalho apresenta um estudo sobre o que é phishing, seus conceitos, seus tipos e os problemas causados na sociedade.*

1. Introdução

Após mais de 25 anos na sociedade, tendo surgido em 1995 [James 2006], o número de ataques de phishing alcançaram mais de 300 mil ataques em dezembro de 2021 [APWG 2021]. Esses ataques exploraram, entre outros cenários, o de pandemia de COVID-19 [Abroshan et al. 2021], período no qual os indivíduos ficaram mais dependentes de serviços online, como aplicativos de mensagens e de mídias sociais, aumentando o número de técnicas de abordagem para se usar no ataque.

Nesse contexto, deve-se considerar que o fator humano por trás do ataque de phishing, que é explorado por formas de engenharia social para aumentar a efetividade do ataque, é um ponto crucial. Dependendo das ações do usuário, pode existir o sucesso ou a falha do ataque [Desolda et al. 2021]. Com isso, a sensação de liberdade, privacidade e segurança acabam sendo comprometidas.

Este artigo terá mais 4 seções: na segunda seção serão apresentados os conceitos básicos sobre phishing, na terceira seção será tratado o que é o phishing e suas características, na quarta seção serão tratados os problemas causados pelo phishing e os temas de pesquisas importantes, e na quinta seção serão descritas as considerações finais sobre o trabalho.

2. Conceitos

Nesta seção serão expostos alguns conceitos importantes relacionados ao ataque de phishing.

2.1. Engenharia Social

Um ataque de engenharia social, segundo [Syafitri et al. 2022], busca manipular uma vítima atacando seu ponto mais fraco. Segundo o dicionário de Inglês de Oxford, o termo ”engenharia social” pode ter dois significados distintos, o primeiro se trata de usar um planejamento centralizado na tentativa de gerenciar uma mudança social e o segundo se trata de persuadir uma pessoa induzindo-a divulgar informações privilegiadas. Deve-se ressaltar que teve a primeira ocorrência de tal termo foi em 1842 e ainda se encontra nos dias atuais [Hatfield 2018]. Alguns tipos de engenharia social, além do phishing, segundo[Syafitri et al. 2022], são: *pretexting*, *baiting* e *ace-to-face interaction*.

Na técnica de *pretexting*, o atacante coleta uma informação pública disponível em *websites*, redes sociais e listas telefônicas, para elaboração de um ataque de comunicação bidirecional, onde o atacante pode oferecer ou pedir algo ao alvo.

A técnica de *baiting* utiliza a curiosidade de membros de uma organização sobre algum item para que conecte tal dispositivo infectado em algum aparelho da organização, assim infectando-o.

A técnica de *ace-to-face interaction* é comumente usada em ataques de engenharia social. Essa técnica explora tirar vantagem de fraquezas psicológicas da vítima, como implorando ajuda ou por acesso a algum dispositivo.

2.2. Fator Humano

Considerando que estamos cercados por tecnologia da qual buscamos o uso para melhorar nossas vidas, também acaba por deixar-nos mais vulneráveis e acessíveis a sistemas enganosos e à exploração. Para considerar um sistema seguro, depende de que as pessoas responsáveis por ele ou com algum nível de relação a ele sigam as normas de segurança impostas. Dessa forma, considerando que pode ocorrer alguns fatores como distração, pressão, estresse e etc, às pessoas, criando fraquezas em um sistema, quais um invasor pode explorar [Desolda et al. 2021].

2.3. Malware

Malware pode ser classificado como um hardware, firmware ou software que, intencionalmente ou não, possam ser inseridos ou adicionados a um sistema e com capacidade de comprometer a confiabilidade, integridade ou disponibilidade dos dados, aplicativos ou do sistema operacional da vítima [Almeida et al. 2022].

2.4. Ransomware

O ransomware é um tipo de malware do qual está diretamente ligado a ataques de phishing e outros ataques de engenharia social, por serem mais direcionados aos usuários finais, com a finalidade de reter dados pessoais e liberá-los apenas mediante ao pagamento de um resgate [Almeida et al. 2022].

3. O que é Phishing

Phishing é um ataque que explora técnicas de engenharia social para realizar um roubo de informações confidenciais [Aleroud and Zhou 2017], estando presente na sociedade desde 1995 [James 2006]. O termo *phishing* se dá ao fato de que o atacante está tentando pescar, do inglês *fishing*, dados; e o "ph" é derivado de sofisticado, do inglês *sophisticated*, por conta das técnicas mais sofisticadas que tais atacantes usam para se distinguir da atividade mais simples de pescar [James 2006].

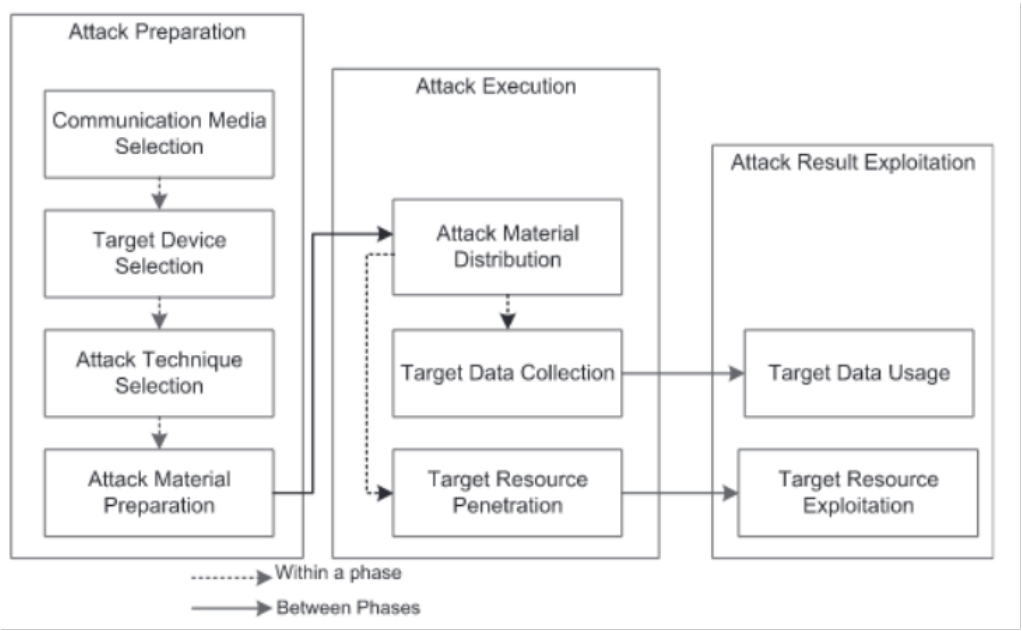

Figura 1. Fases do processo de um ataque *phishing*. [Aleroud and Zhou 2017]

3.1. Processo de elaboração

Passando por um processo, assim como da Figura 1, onde se tem as fases de preparação, execução e exploração dos resultados. Cada fase tem seus passos a serem seguidos.

Na primeira, tem que selecionar a mídia a ser usada, que seria o meio a ser usado, a seleção do dispositivo alvo, que é corresponde a seleção do vetor, a técnica de abordagem que vai ser aplicada e, por fim, a preparação do material para o ataque. Após isso, tem a execução do ataque, onde se distribui o material contaminado colocando na coleção de dados do alvo ou em seus recursos. Com isso, caso o alvo usar tal dado ou explorar tal recurso e acabar fornecendo dados sensíveis ou relevantes ao atacante. Caso o ataque seja realizado em conjunto outros tipos de ataques, como o de ransomware, assim, contaminando o dispositivo.

Na Figura 2, pode-se ver alguns exemplos de meios, vetores e a técnica de abordagem para a elaboração de um ataque de *phishing*.

3.1.1. Meios

Para que um ataque de *phishing* aconteça, é necessário ter um meio para haver a interação entre o atacante e o alvo, e, para que isso aconteça, existem alguns meio mais comuns, que seria o meio da internet, *Short Messaging Service* (SMS), meios dos quais pessoas utilizam normalmente e que podem ser usados por atacantes para interagir com vitimas [Chiew et al. 2018].

3.1.2. Vetores

Existem vários vetores associados aos meios citados anteriormente dos quais servem para intermediar um ataque por um meio e uma abordagem escolhida, contudo, vetores associados ao meio de internet são os mais populares entre os ataques de *phishing* [Chiew et al. 2018]. Pode-se destacar os vetores de Email, para internet, Smishing, para SMS, e Vishing, para serviços de voz.

3.1.3. Abordagens Técnicas

Existem várias abordagens técnicas que podem ser usadas em um ou mais vetores para a implementação de um ataque de *phishing*, assim como mostrado na Figura 2. Na sequência do texto são explicados cada um dos tipos de abordagens com base na referência [Chiew et al. 2018].

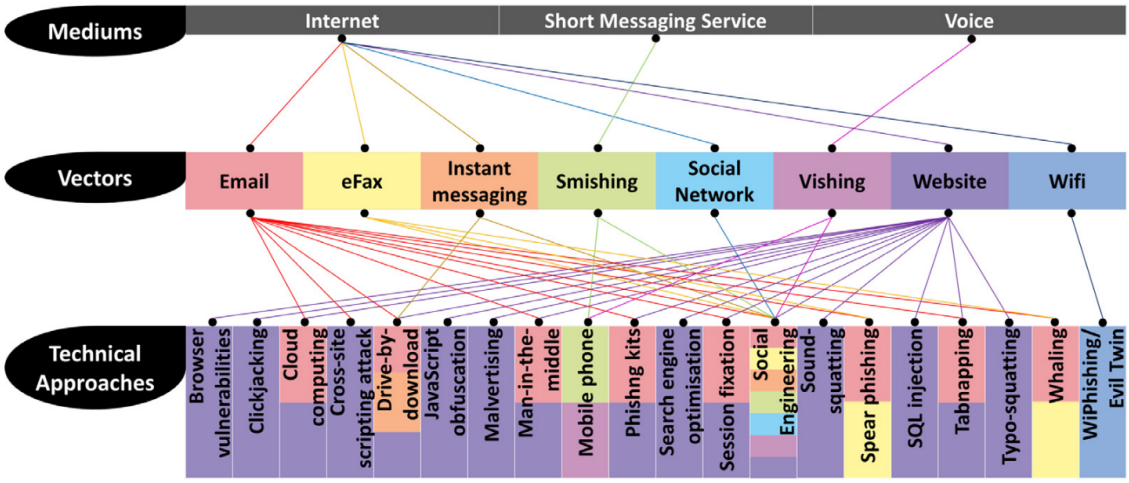

Figura 2. A interligação entre meio, vetor e abordagem das técnicas de *phishing*. [Chiew et al. 2018]

Browser vulnerabilities Trata-se de explorar vulnerabilidades de um browser para lançar um ataque de *phishing* a um usuário, por exemplo por meio de extensões e *plug-ins* de fornecedores externos.

Clickjacking Também conhecido como *user interface (UI) redressing attack*, trata-se da manipulação da UI de uma página web, onde possibilita uma ação externa ao utilizar a página.

Cloud computing Trata-se de um serviço online que está aumentando a popularidade e que podem ser atacados de 6 formas, atacando os as relações entre os três componentes que compõem tal serviço, que são os usuários, o serviço e o provedor da nuvem. Relações mostradas na Figura 3.

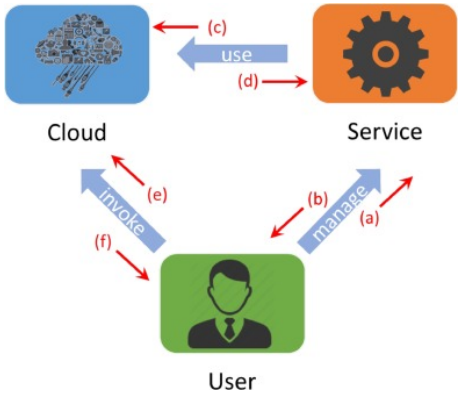

Figura 3. A relação entre os três componentes da computação em nuvem. [Chiew et al. 2018]

Cross-site scripting (XSS) attack Trata-se da exploração de uma vulnerabilidade de *websites* onde aceitam que o atacante injete um código malicioso dentro de algum campo de dados do site e ele interprete tal código podendo permitir que acesse informações pessoais como login e credenciais.

Drive-by-download Trata-se de injetar um malware, vírus ou código em uma máquina apenas visitando um *websites*, vendo um HTML ou recebendo um email.

javascript obfuscation Trata-se de utilizar JavaScript para mascarar a barra de endereços, barra de *status*, barra de ferramentas ou a área de *menu*, assim conseguindo falsificar os endereços ou *status* de tais áreas.

Malvertising Trata-se de utilizar o serviço de hospedagem online de anúncios como meio de distribuir malware, onde, ao clicar no anúncio, um malware dinâmico infecta a máquina da vítima e explorar suas vulnerabilidades com o objetivo de roubar informações pessoais.

Man-in-the-middle (MITM) Trata-se de o atacante se colocar no meio da comunicação entre a vítima e uma aplicação *web*, assim como na Figura 4, onde o atacante é capaz de controlar as informações submetidas da vítima à aplicação *web*, possibilitando a captura

de credenciais de autenticação. Tal captura pode ser feita por meio da utilização de um proxy transparente, onde o atacante tem acesso a essa rede e ao tráfego dela, demonstrado na Figura 5.

Figura 4. Exemplo de um ataque Man-in-the-Middle. [Chiew et al. 2018]

Figura 5. Exemplo de utilização de um proxy transparente em um ataque Man-in-the-Middle. [Chiew et al. 2018]

Mobile phone Trata-se de distribuir aplicativos maliciosos para aparelhos de telefones móveis, onde buscam controlar a transferência de dados entre as aplicações do dispositivo, possibilitando a captura de informações pessoais.

Phishing kits Trata-se de ferramentas que possibilitam que o atacante crie um *websites*, emails e *scripts* para obter dados inseridos por um usuário sem ser necessário conhecimento de programação para isso.

Search engine optimisation Trata-se de otimizar a entrega *websites* de *phishing* para potenciais vítimas usando técnicas de otimização para ferramentas de busca.

Session fixation Trata-se de roubar identificadores de sessões, como cookies, gerados quando um usuário faz login em um site com esse tipo de falha de segurança, assim, o atacante é capaz de usar a seção do usuário para efetuar atividades maliciosas como transferência de dinheiro.

Social engineering Trata-se da utilização de técnicas de engenharia social (seção 2.1) para obter vantagem sobre a vítima.

Sound-squatting Trata-se de registrar domínio de sites com nomes similares aos de sites legítimos, assim, o atacante tira vantagem da confusão do usuário que é redirecionado a uma versão de *phishing* do *website* do qual pretendia acessar.

Spear phishing Trata-se de um ataque direcionado a um indivíduo, um grupo ou organização, onde se desenvolve um email com conteúdo relevante e que a vítima conhece o remetente, assim, evitando alguma suspeita da vítima e possibilitando efetuar solicitações de detalhes de login ou rodar algum conteúdo com malware.

SQL injection Trata-se da exploração de uma vulnerabilidade do banco de dados onde é possível injetar comandos e capturar dados da tabela de dados do usuário.

Tabnapping Trata-se do roubo de abas do navegador, onde o atacante envia um link em um email, do qual é aberto no navegador da vítima e, tal site, possui um código JavaScript do qual monitora a atividade do navegador da vítima e também carrega uma tela de login conhecida pela vítima, como do Gmail, para ela achar q a seção foi fechada e que é necessário conectar-se novamente.

Typo-squatting Trata-se de ataque que o atacante registra nomes de domínio com possíveis erros de digitação que o usuário possa fazer, tal como ”www.mybank.br” e ”wwwmybank.br”, assim possibilitando o acesso acidental ao site malicioso e descarregando um malware no dispositivo da vítima.

Whaling Trata-se de um tipo de *spear phishing* do qual tem como alvo pessoas de alto nível executivo e com altos privilégios de acesso na organização, onde, por meio de um malware, o atacante tem acesso a uma porta de acesso ao sistema da organização.

Wiphishing / Evil Twin Trata-se de um ataque que usa redes sem fio, onde o atacante se coloca entre o usuário e a verdadeira rede sem fio, possibilitando que o atacante seja capaz de espionar os dados enviados e recebidos pelo usuário.

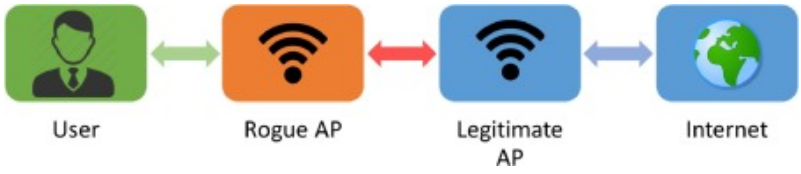

Figura 6. Ataque WiPhishing/Evil Twins. [Chiew et al. 2018]

4. Phishing durante os anos de 2020 e 2021

Problemas com *phishing* persistem na sociedade há mais de 25 anos [James 2006]. Além da persistência, o número desse tipo de fraude vem crescendo nos últimos anos, atingindo o maior número de ataques de todos os tempos em Dezembro de 2021 [APWG 2021]. Com isso, nesta seção serão expostos alguns problemas, sobre *phishing*, encontrados na literatura entre os anos de 2020 e 2021.

4.1. Acontecimentos Práticos

Com a pandemia de COVID-19 em 2020, despertou um novo nível de ansiedade, medo e *stress* na sociedade, e, com isso, pessoas mal intencionadas buscam tirar proveito de pessoas que possuem tais problemas, facilitando os ataques de *phishing*.

Nesse contexto de pandemia, também houve o aumento no número de indivíduos dependentes de serviços online, como compras, videochamadas entre outros, com isso, deixando as pessoas mais vulneráveis a fraudes online, tal como envio de malwares, dos quais 94% é enviado por e mail e, desses, 32% envolve *phishing*.

Com tudo isso, invasores cibernéticos aproveitaram para tirar vantagens de possíveis vítimas ao usar informações e palavras chaves relacionadas ao COVID-19, para realizar ataques de *phishing* [Abroshan et al. 2021].

A Figura 7 demonstra algumas variáveis que influenciaram na probabilidade de acontecer um ataque de phishing durante a pandemia de COVID-19. Nota-se que, com mais vetores hipotéticos (siglas H1, H2 e H3 da figura), maior a possibilidade de acontecer um *phishing*, já que aumenta o número de abordagens, como tratado na seção 3.1.

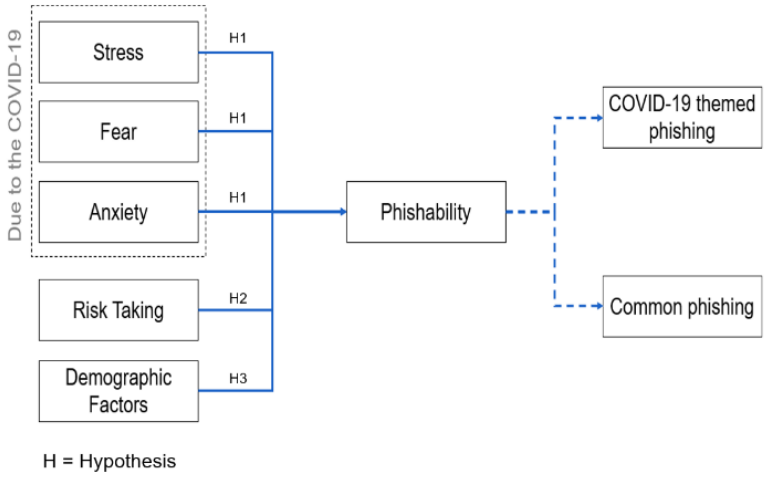

Figura 7. Grafo do processo de efetivação de um ataque *phishing*. [Abroshan et al. 2021]

Com tudo isso, e considerando o ano de 2021 e também a continuação da pandemia de COVID-19, dados demonstram o crescente número de ataques de *phishing* no decorrer do ano (Figura 8), até atingir o maior valor histórico até então, em dezembro, assim corroborando o efeito dos problemas supracitados.

No Brasil, durante esse período, houve uma leve queda no número de ataques, como pode-se ver na Figura 9. No entanto, os números ainda são preocupantes por conta de serem em um período de 3 meses. Dentre esses números, a quantidade de *phishing* contra empresas de SaaS e Webmail totalizam 18% dos ataques no terceiro trimestre e contra empresas de comércio eletrônico foram 37,5% [APWG 2021].

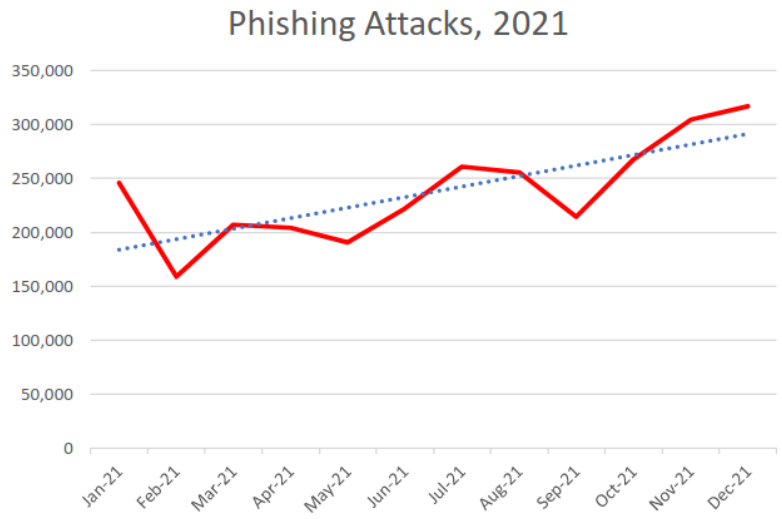

Figura 8. Números de ataques *phishing* em 2021. [APWG 2021]

Figura 9. Números de ataques *phishing* no Brasil em 2021. [APWG 2021]

4.2. Trabalhos relacionados

Sobre pesquisas na área de detecção de *phishing*, destacam-se algumas técnicas, tais como heurísticas e *Machine Learning*.

Um exemplo de técnica heurística foi desenvolvida em [Almeida et al. 2022], trabalho desenvolvido com base em *web crawling*, que é um processo usado pelos motores de busca para coletar páginas da *web* [Almeida et al. 2022]. Neste trabalho, também é possível monitorar o comportamento do usuário, detectando suas ações e determinando o momento em que o usuário acessou o conteúdo malicioso.

Um exemplo de técnica de machine learning, é descrita em [Sahingoz et al. 2019], que considera o uso de diferentes algoritmos de aprendizado de máquina, como *Decision Tree* que se trata de um algoritmo que usa um modelo de decisões em forma de árvore e suas possíveis consequências[1]. O trabalho implementa uma forma de detecção em tempo real de páginas *web* de *phishing* investigando sua URL. Tal trabalho, tem como foco não apenas executar os algoritmos de aprendizado de máquina, mas também a extração de características importantes de tais dados para compor seu próprio conjunto de dados.

Outras técnicas de inteligência artificial, como *deep learning* e *Hybrid learning* são apresentadas em [Basit 2021], para implementação de formas de detecção de *phishing*.

5. Considerações finais

Neste trabalho são discutidas as principais características de um ataque de *phishing*. Também são mostrados dados sobre o problema causado na sociedade por tal tipo de ataque. Com isso, considera-se que *phishing* é um problema persistente na sociedade há anos e com muitas possibilidades de evolução quanto ao combate e mitigação.

Foi explicitada a estrutura do *phishing*, como ela é preparada, considerando todo um processo de seleção de meio, vetor e técnica de abordagem, seguido da execução do ataque mediante a técnica escolhida e, por fim, a exploração dos resultados obtidos. Em relação às pesquisas relacionadas, nota-se que existem pesquisas bem atuais, utilizando as mais recentes formas computacionais e que sempre há espaço para aperfeiçoamentos.

Por fim, não deve-se desconsiderar o *phishing* e deixar de pesquisar meios de detecção e mitigação pois tal ataque persiste há mais de 25 anos e deverá persistir por ainda mais tempo, já que existem sempre novas maneiras e técnicas de ataques a serem combatidas. Dessa forma, este trabalho surge como uma fonte de fácil compreensão para introduzir o tema proposto, proporcionando a abertura de um leque de oportunidades para a aprofundamento no conteúdo.

[1] https://en.wikipedia.org/wiki/Decision_tree

Referências

Abroshan, H., Devos, J., Poels, G., and Laermans, E. (2021). Covid-19 and phishing: Effects of human emotions, behavior, and demographics on the success of phishing attempts during the pandemic. *IEEE Access*, 9:121916–121929.

Aleroud, A. and Zhou, L. (2017). Phishing environments, techniques, and countermeasures: A survey. *Computers Security*, 68:160–196.

Almeida, R. A. O. C. B. d. et al. (2022). Heuristic phishing detection based on web crawling and user behaviour monitoring with a deterministic approach for cybersecurity.

APWG (2021). Phishing activity trends report 4th quarter 2021. In *PHISHING ACTIVITY TRENDS REPORT 4th Quarter 2021*. apwg.org.

Basit, A. e. a. (2021). A comprehensive survey of ai-enabled phishing attacks detection techniques. *Telecommunication Systems*.

Chiew, K. L., Yong, K. S. C., and Tan, C. L. (2018). A survey of phishing attacks: Their types, vectors and technical approaches. *Expert Systems with Applications*, 106:1–20.

Desolda, G., Ferro, L. S., Marrella, A., Catarci, T., and Costabile, M. F. (2021). Human factors in phishing attacks: A systematic literature review. 54(8).

Hatfield, J. M. (2018). Social engineering in cybersecurity: The evolution of a concept. *Computers Security*, 73:102–113.

James, L. (2006). Chapter 1 - banking on phishing. In James, L., editor, *Phishing Exposed*, pages 1–35. Syngress, Burlington.

Sahingoz, O. K., Buber, E., Demir, O., and Diri, B. (2019). Machine learning based phishing detection from urls. *Expert Systems with Applications*, 117:345–357.

Syafitri, W., Shukur, Z., Mokhtar, U. A., Sulaiman, R., and Ibrahim, M. A. (2022). Social engineering attacks prevention: A systematic literature review. *IEEE Access*, 10:39325–39343.