DOI: 10.2478/ias-2025-0008 • JIAS • 20(3) • 2025 • 125-144

Illusion of Presence: Physical Deception and Direct Manipulation in Social Engineering

Jibran Rasheed Khan[1*], Sumaira Mustafa Qureshi [2], Farhan Ahmed Siddiqui[3] and Syed Asim Ali[4]

1 Department of Computer Science – UBIT, University of Karachi, Main University Road, Karachi - 75300, Sindh, Pakistan
2 Delta School, Riyadh-Al-Muruj District, Saudi Arabia *sumaira.mustufa@yahoo.com*
3 Department of Computer Science – UBIT, University of Karachi, Main University Road, Karachi - 75300, Sindh, Pakistan *farhan@uok.edu.pk*
4 Department of Computer Science – UBIT, University of Karachi, Main University Road, Karachi - 75300, Sindh, Pakistan *asim@uok.edu.pk*
* Corresponding author: *jibran_rasheed@hotmail.com*

Abstract: This paper proposes a novel taxonomy of physical social engineering attacks that require physical & direct interaction with the target. It addresses the classification challenge found in the literature, where approximately 35% of non-SEA cases are identified. This helps in understanding diverse methodologies employed by attackers. This paper presents a comprehensive state-of-the-art study of each attack's concept, background, methodology, and impact on individuals or organisations, to estimate the severity of the attack. Also, social and psychological factors associated with attacks, such as trust, obedience, respect, and kindness, are exploited, which requires a critical need to enhance awareness and integrate psychology with security measures. This emphasises considering the human element in cybersecurity solutions and their understanding.

Keywords: Cybersecurity, Social Engineering, Human-based Attack, Physical Attack, Human Vulnerability, Attack Taxonomy.

1 Introduction

Human-based social engineering attacks represent a small subset of social engineering attacks. Social engineering is a vast term that broadly covers a large number of social engineering attacks. Unlike other technology-assistive and modern advanced techniques, physical and human-based attack methodologies rely on the physical presence of the attacker within the target or the environment's proximity. It hinges on the attacker's smartness, understanding, actions, and ability to influence the victim in a face-to-face interaction and psychological manipulation [1, 2]. Attackers leverage innate human tendencies such as obedience [3], trust, courtesy [4], kindness, and helpfulness [5] to exploit human psychology and be trusted as an authorised person or authority [6]. This helps to bypass the security measures and achieve the goals.

Researchers estimated the annual impact of social engineering attacks would scale up from 6$ trillion [7, 8] to $10.5 trillion [9-12] by 2025, which is approximately a 15% increase compared to 2021 [8, 11]. Cybercriminals mostly (about 97%) [13] take advantage of vulnerability in human behaviour, whereas about only 3% use malware and technical vulnerabilities [13, 14]. Twitter reported in 2020 that there was a noticeable change in cybercriminal attack approach, where about 98% targeted vulnerabilities in employees instead of technical faults. This emphasises the importance of social engineering attacks; understanding is critical to devising preventive and mitigation solutions.

In today's cybersecurity landscape, even though technological defences are advanced, the human element remains a critical vulnerability due to innate factors. At the present time, technology is limited to catering to human weaknesses, and it's beyond machine controls to identify, detect, and correct human mistakes. It is essential to understand these attacks, which could lead to severe consequences

such as unauthorised access, breaches of facilities, compromised critical infrastructure, and the theft of confidential or important documents [15]. Also, it can assist in large-scale or other types of social engineering attacks or cyber espionage. This attack vector requires a comprehensive analysis and understanding, which is essential for developing an effective security measure and guideline.

The scope of this paper is limited to physical human-based social engineering attacks. It specifically focuses on the attacks that involve and require the attacker's presence or physical interaction with the victim. The primary objective of this paper is to present a detailed taxonomy and analysis of these attacks. The paper also comprehensively discusses the attack concept, background to understand its roots and nature, attack methodology, and its impact on the target organisation or a person, along with real-life examples and damage costs reported in the past.

2 Research Methodology

This work uses Critical Interpretive Synthesis (CIS) methodology developed by Dixon-Woods et al. [16, 17]. The CIS was selected over conventional SLR due to the primary objectives of the work being to critically review and integrate a diverse body of knowledge into substantial grounds. This helps in setting a strong foundational work to generate a novel theoretical and conceptual framework for SEAs classification. This approach is suited for investigating complex phenomena where definitions are contested and boundaries are blurred, as in SEA. CIS is distinguished by its capacity to deconstruct assumptions within the literature base and produce a "coherent and integrative conceptual synthesis" that transcends the findings of the original studies [18]. This aligns with our project goals, which require critical examination of distinctions and conceptual blurriness between SEA and non-SEAs, and their classification challenges.

The CIS involved an iterative and reflexive process. An initial systematic search was conducted on renowned research databases such as Google Scholar, ScienceDirect, IEEE Xplore, SpringerLink, and MDPI. Used the search strings "social engineering attack", "taxonomy", "human factors", "phishing", "classification", "categorisation", and "types" to capture conceptual richness. In CIS, the search strategy is not fixed but evolved theoretically as the synthesis developed. This allows for the inclusion of literature that is best suited to the work objectives. Initially, the literature was filtered based on its methodology, conceptual relevance, attack presence, and classification phenomena. By reciprocal translational analysis (interpreting and relating key metaphors and concepts across studies) and refutational synthesis (actively seeking and interrogating contradictory findings), the literature was synthesised to construct a new integrative logic. This process concluded in the development of a conceptual model that aims to resolve contradictions and provide a more nuanced understanding of the classification challenge in the SEA landscape.

3 Literature Review & Challenges

The term "social engineering" originally belonged to political science [19, 20] and was later used for various purposes to empower people, strengthen and structure society and in policy making before it stepped into the cybersecurity realm or for malicious purposes. Since the age of phone phreaking (1970) [21, 22], social engineering has continuously gained popularity, proving its stealth nature and scalability over time [23]. Today, social engineering is one of the most overlooked yet highly effective cybersecurity threats, demonstrating its persistent and evolving role in digital security [24]. The modern definition of social engineering blurs the line between social engineering and non-social engineering attacks. [25] states it as an art that uses human psychology, [26] defined as the practice of exploiting human characteristic(s). Whereas [13, 27] stated this as a relationship of asymmetric knowledge between the attacker and targets. [28] defined as the most creative approaches to gain access to the information system. Researchers [14, 29, 30] described a multilayered attack, which leverages social, psychological, technical, and physical aspects to bypass security. [31] refers to the art of human hacking, and [25] stated it is completely different from traditional cyberattack techniques and procedures. The authors [32] define it as the simplest method of gathering information about a target. Whereas [19] refers to technology-driven tactics, and [27, 33, 34] refer to diverting human behaviour and decision-making. Various definitions may be perceived differently with different perspectives and understanding of the social engineering definition. The definitional viability and ambiguity due to its multi-disciplinary nature and varied perspective cause misclassification, which makes it difficult to identify attacks and categorise them appropriately. For example, some new types of attacks/hypothetical attacks are presented with a lack of evidence and support, such as "Flame Wars Attack" and "Astronfing Attacks" [33].

The study found 49 SEAs, and 31 studies discussed the classification scheme. Whereas 18 of them proposed attack taxonomies, while 13 only suggested a classification but didn't classify attacks accordingly. The thorough examination of each study presented in **Appendix A - Table A1** shows a wide array of "Classification Factors" employed, ranging from "Operator, Type, Channel" to "Medium," "Computer & Human," "Behaviour," "Device," "Interaction," and "Environment". The diverse and inconsistent classification reflects the lack of understanding and lack of standardisation. This blurs the concept and makes it challenging for researchers to appropriately classify the attack. Another major issue identified is that many non-SEAs categorised under SEA taxonomies, such as SQLi [35], XSS [35, 36], and JavaScript Obfuscation [37], are a few examples, despite these taxonomies' focus on social engineering attacks. **Table 1** presents a breakdown of details, number and specific types of non-social engineering attacks identified within each taxonomy. For example, some new types of attacks/ hypothetical attacks are presented with a lack of evidence and support, such as "Flame Wars Attack" and "Astronfing Attacks". Also, many non-SEAs are misclassified such as SQLi [35], XXS [36], Session Hijacking [38], Ransomware [33, 39], Buffer Overflow [36], Rootkits [30, 38] and Secure Socket Layer (SSL) [38] Attack are few examples.

Clustering the classification schemes presented in **Table 1** into six themes: 1) No Classification, 2) Human & Computer, 3) Medium, 4) Behaviour, 5) Social & Technical, and 6) Others. The taxonomies are clustered based on their qualitative measure of classification scheme and the similar concept they adopted. For example, Human & Computer, Human & Software, Human & Technology, and Technical & Human Deception techniques are clustered as Human & Computer. **Figure 1** illustrates the frequency of each classification theme with the year of appearance, which helps in understanding the classification trend and prevalence of the issue.

No Classification: this cluster represents 41.94% (13/31) of classifications that consist of 13 literatures. These literatures either omit classification or discusses classification factors without categorising the attack accordingly. Importantly, literature in this cluster observed from 2013 to 2024, which reflects the persistence of the issue to date. This directly contributes to the lack of understanding of attacks and ambiguity in their classification.

Human and Technical: it comprises over 6/31 (19.35%) studies that used "Human & Computer", "Human & Software", "Technical & Deception" classification factor. Besides this, there are 4 more studies (i.e. [28, 31, 40], and [41]) that suggested a similar idea of classification but didn't classify attacks accordingly. Therefore, the "Human & Technical" classification studies represent an overall 32.25% (10/31). It is the most common classification factor that is frequently used in the SEAs literature. Although this cluster represented a structured distribution, a significant proportion of non-SEAs (30%) are observed, such as Ransomware, Botnets, Malware, XSS, Rootkits, and Buffer Overflow.

Medium and Behaviour-based: both classification schemes appeared three times and evenly constitute about 9.6% (3/31) coverage. **Figure 1** shows that these two are the emerging concepts. Medium-based distribution factor mainly focuses on digital phenomena while the same medium is shared among several attacks, such as Phishing, which all three common media (Text, Voice, Email). The non-SEAs are frequently observed in this classification, such as SQLi and XSS. Also, SE itself is represented as an attack. While authors tried to classify the SEAs based on human behaviour such as trust, emotion, cognitive level, and other factors. Although their distribution seems to be relevant, they didn't map the proposed factor to SEAs nor classify them.

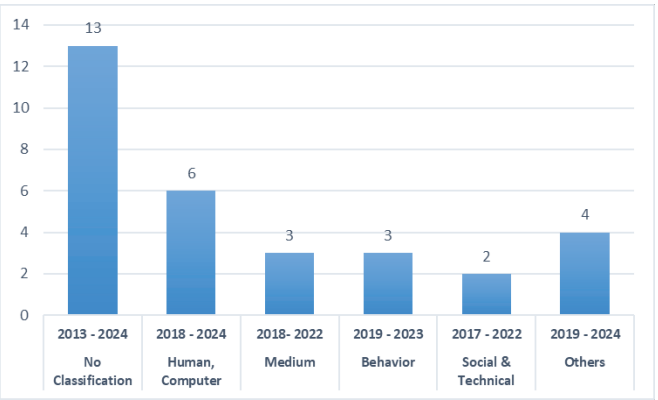

Figure 1. Classification trend and frequency of cluster themes

Social & Technical: There are only two studies that classify the SEAs based on "Social & Malware", and "Social & Technical & Mobile" schemes. It comprises only 6.45% (2/31), one in 2017 and the other in 2022. Authors tried to integrate social and technical perspectives, but notably the inclusion of non-SEAs found in their distribution, such as Ransomware, Trojan horse, Content Injection, and Key logger.

Other: four classification schemes fall into this cluster, employing a diverse and rare idea which includes "Device," "Interaction," "Environment," or "Physical & Digital space". This cluster represents the 12.9% (4/31) of coverage that appeared in recent literature from 2019 to 2024. Although they appeared only once, these studies presented unique criteria for SEA distributions, but pose similar challenges of classification and found non-SEAs like Malware, XXS, Botnet, Rootkit, and Hardware Attack (**Table 1**).

Despite the number of classification schemes, an irregular pattern was found in the classification. Moreover, 35% of attacks presented in the literature are non-SEAs, and the same attack is differently classified across the literature. These variations show a lack of understanding of the SEA landscape, which leads to classification inconsistencies and persistence of issues. This may trouble researchers to determine attack and their characteristics. Therefore, it requires the development of a clear and unified classification scheme that helps to understand SEA and strengthen the body of knowledge.

4 Spectrum of Deception Attacks (SODA)

This paper is a part of the project "Spectrum of Deception Attacks (SODA)", which aims to provide a categorical analysis of social engineering attack techniques, tricks and tactics (SEAT^3) to uncover the underlying phenomena, historical grounds, and evolution of attacks. Also, the project aims to dissect the diverse attacks and highlight humans' psychological and social mechanisms to render their drivers and measure their effectiveness and consequences. This project, structured SEAT^3 based on the attacker interaction method and level of proximity, is shown in **Figure 2**. This allows readers granular exploration from physical & direct engagement with the victim to subtle influence through unattended web-based manipulation. Each paper in this project is designated for in-depth analysis of its respective branch by exploring four parts: concept, historical evolution, methodology, and incidents reported in past.

Technology is unstoppable, so are humans as well. We cannot stop technological progression; likely, we cannot stop humans from utilising these technologies. Yet, we have to find an appropriate solution to reduce or mitigate the magnitude of SEA. As technology advances rapidly, so too does the sophistication of social engineering rise. To achieve the goals, we present a state-of-the-art systematic

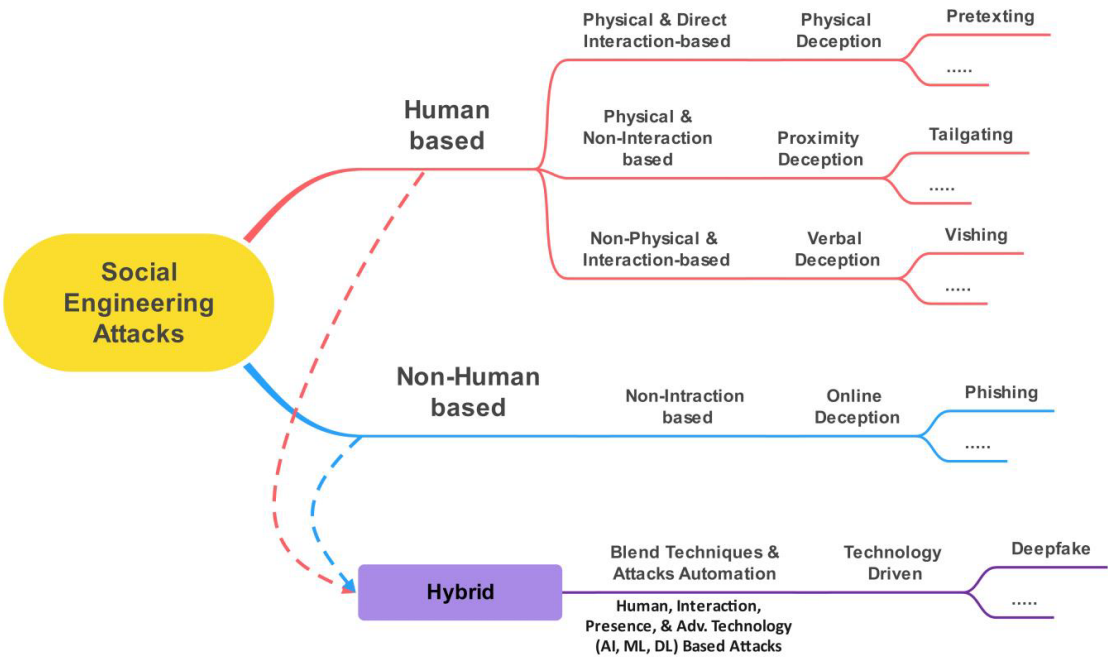

Figure 2. Taxonomy of SEA based on Attacker's Level of Proximity and Interaction

categorisation of SEA based on attacker interaction and proximity level. The proposed taxonomy is the first of its kind, which allows the reader to have a clear understanding of diverse attack strategies that target human vulnerabilities and bypass the security measures.

5 Taxonomy of Physical SE Attacks

A Introduction

This section presents a novel taxonomy of physical attacks within the sub-branch of human-based social engineering, shown in Figure 3. They proposed a sub-taxonomy of human-based attacks, which requires the attacker's physical presence and direct interaction with the victim.

1) Rationale

The proposed novel taxonomy is designed to provide a more structured form for a better understanding and clearer view of SEAs. After thorough examination of SEAs and their characteristics, we learn that 95% of SEAs rely on technology (e.g., phishing), and 70% of SEAs require human interaction and intervention either physically or digitally, with attacker presence or through media interaction [42]. We have examined different dimensions for attack classification as presented in (Section III). After careful evaluation, we found that SEAs can be categorised based on two parameters: *Level of Proximity* and *Level of Interaction* for executing attacks. It is observed that there is a varied degree of attacker presence in each SEA, completely independent of technology, medium, software, human, social, psychological behaviour and other attack vectors. This addresses the two major challenges of classification: multiclass categorisation and misclassification error. Also, it allows one to correlate one or more factors and present a broader and clearer view of attacks. Furthermore, the use of one or more attack vectors, media and execution channels does not affect the classification scheme, such as Phishing can be performed through Text, Voice, or SMS. With the advancement of technology, SEA is evolving, paving its path through one or more ways to execute an attack. Despite the number of executing vectors, we found that the base proximity phenomena remain unchanged, and apply to new emerging variants as well. Therefore, the proposed scheme does not classify attacks based on what technology or means it uses; instead, the classification factor segregates the attacks based on the degree of attacker presence and involvement with the victim. **Appendix A - Table A2** summarises existing taxonomies, their limitations, and how the proposed taxonomy addresses those limitations. This would help in building a unified and consistent framework that strengthens the body of knowledge and plays a vital role in building robust solutions.

The attacks presented in the proposed taxonomy (Figure 3) of physical attacks included only those which share common characteristics: require attacker presence, attacker interacts with the victim, attacker commonly uses impersonation, attacker directly deals with the victim, attacker identity is known (e.g., Face, body, height, etc.), high risk of execution, and highly calculated attacks. This plays a vital role in being stealthy and engaging with its target, helping to analyse their distinct methodologies and associated vulnerabilities they exploit in each attack. For example, there are two physical attacks (Physical Impersonation Attack (PIA) and Dumpster Diving Attack (DDA)) that belong to different

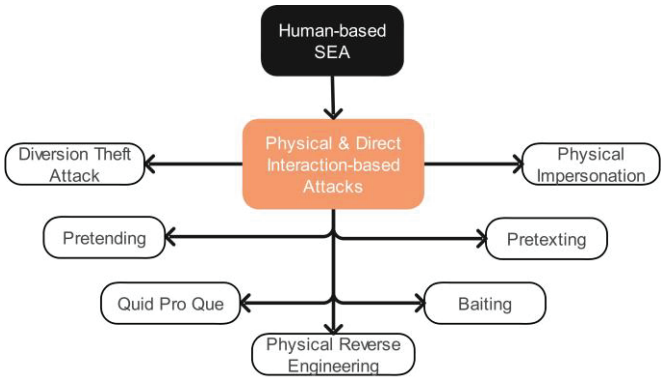

Figure 3. Taxonomy of Physical & Direct Interaction Attacks

categories with respect to the proposed classification. Whereas the PIA require an attacker to enter into the target premises and interact with the victim to gain intended information or access to the resource. DDA does not require an attacker to get in direct communication with the victim to extract the important information about the target. This allows more targeted and focused analysis to develop security measures, training programs and drills for practising. Furthermore, it allows us to analyse the risk associated with each type of attack, which helps to improve overall human behaviour, their attitude and strong cultural habits.

2) Attack Methodology

The social engineering attacks are sophisticated tactics which require a well-structured plan and careful execution. Generally, the physical attacks do not have an opportunity to refine or revise its attack methodology during attack or upon failure. They are highly risky and critical, if fails lead to immediate and serious consequences. Typically, it consists of several stages: selecting a target, reconnaissance on the target, creating a plan or credible story, executing a plan, gathering sensitive information, accessing a resource, and exiting safely, as shown in **Figure 4**. Each phase is briefly discussed in the following:

- **Selecting Target**: In this phase, the attacker picks a soft and easy target that could be manipulated or influenced for social engineering. It could be an individual or an employee of an organisation.
- **Reconnaissance on Target**: in this stage, the attacker researches the target on various sources, including social media, company websites, phone books, and industry conferences [10], [16], [23]. Attackers gather information, learn the environment, routine, employees' activities, and security protocols enforced around premises to frame all possible situations and difficulties that they may face.
- **Create Scenario**: based on the previous stages, the attacker carefully builds a proven and credible story. An attacker usually designs two strategies: entry and exit. Entry strategy, to establish his credibility, which tends to believe in the situation and act accordingly to assist him. Exit strategy, to safely leave the premises without being suspected.
- **Execute Plan:** once the attacker has built and is satisfied with the plan, step in to execute the plan with the chosen methodology, such as Pretexting. The attacker visits the target and interacts with the victim. Attacker uses different personas based on the chosen methodology and scenario, which generally include executive authority, helpdesk, IT staff or layperson personas. The authority persona commonly used to pose as the executives, manager, senior member, and auditor, where the attacker takes advantage of lower staff to obey and follow the instructions to proceed with things [23, 35]. Layperson Persona, such as delivery personnel, technicians, maintenance workers or contractors, generally use this persona to disguise employees who ask

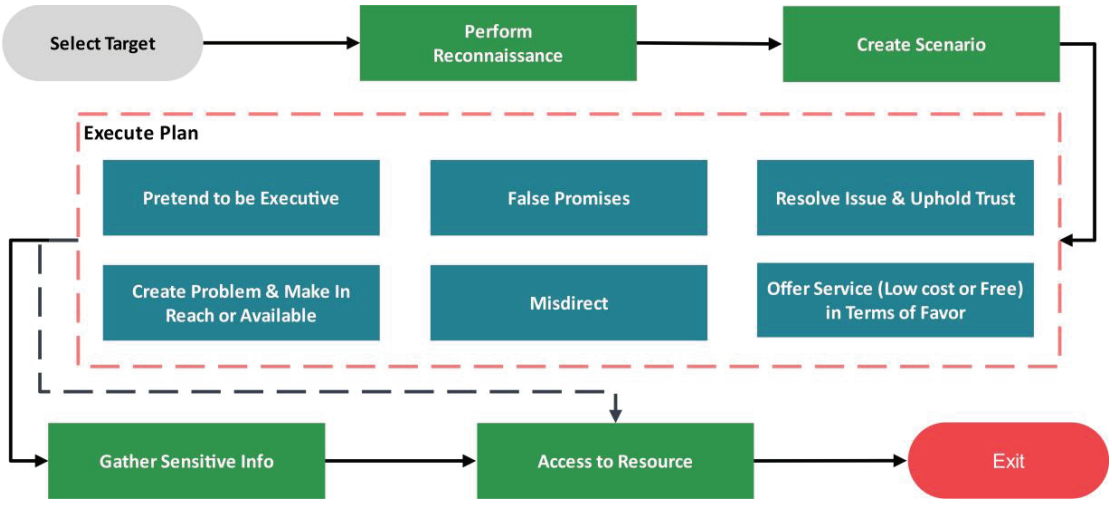

Figure 4. Generic Methodology of Human-based Physical Social Engineering Attacks

for sending parcels, visit specific areas, or to gain access to restricted areas [10]. The attacker makes sure to be easily available and in reach with the right skillset that can surely solve the problem scenario [16, 37]. If the attack fails, the consequence is typically immediate and serious. This leaves no opportunity for refinement or revision of the methodology in real-time.

- **Gather Information**: In this phase, the attacker tries to gather sensitive information under the pretext of the scenario.
- **Access to Resource**: once the attacker deceives the victim and gathers the required information, he enters into the facility or resource location, bypassing the security using the security code and performs the detrimental actions. This includes stealing documents, installing malware, disrupting business operations, or penetrating the company network.
- **Exist Strategy:** this is the crucial stage of the attack, which requires safely leaving the premises without being suspected. Safely leave the premises and make sure to make a future attempt or maintain access to the corporate network [10].

6 Physical Social Engineering Attacks

A Pretending

Pretending is an act of representing oneself as someone and falsely signifies someone who is not to deceive their target. It is a social engineering technique also known as Impersonation, which exploits the human trust factor to gain access to systems, restricted areas, or digital assets. Typically, an attacker impersonates IT staff, technicians, government agency officials, regularity compliance personnel, or an employee to manipulate the victim into divulging information or access to the resource or facility [30]. An attacker could be impersonating in a physical and digital environment (cyberspace), which may be used in various social engineering practices to trick humans. It exploits the psychological and social factors in human nature, such as politeness, kindness, and the desire to help [38, 43].

1) Impact and Costs

The impersonation technique poses a critical risk to organisations and individuals as well. Common examples of impersonation risk include Corporate Espionage to steal trade secrets or confidential data, Physical Security Breaches that lead to data theft, malware installation, damage devices to disrupt operation [38].

B Physical Impersonation

Physical Impersonation is a subtype of impersonation of a tactic which only deals with the physical interaction (face-to-face or in-person meeting) [23, 44, 45] of the attacker with the target to divulge into disclosure of information or access to a facility, such as unauthorised access to a system, network, or secure location. An attacker generally poses as a trusted individual, such as an employee, contractor, or visitor, to manipulate and gain the target's trust or seek their kindness.

1) Background

Physical Impersonation & pretending have been used in the physical domain for ages, long before the digital era. The physical deception technique, which, over time, evolved into a significant cyber threat. Historically, in the physical domain, fraudsters typically impersonated authority figures such as employees, technicians, law enforcement agents, and executives to gain access to privileged areas or rights. Over time, this has been adopted in the cybersecurity domain by the bad actors as a prevalent tactic to induce internal members into revealing sensitive information and access to the facility [43]. This is typically used to pose as an IT support or helpdesk employee [30]. This tactic is highly used in phishing and other social engineering attacks.

2) Impact and Cost

The impersonation technique poses a critical risk to organisations and individuals as well. Common examples of impersonation risk include Corporate Espionage to steal trade secrets or confidential data, Physical Security Breaches that lead to data theft, malware installation, damage to a device to disrupt operation [38].

C Pretexting

Pretexting is a social engineering attack technique that uses a prescribed, fabricated scenario known as a pretext to deceive its target into revealing information. It primly relies on the trust factor [22], used to establish credibility and win the victim's trust to execute the attack [19].

1) Background

The pretexting concept has been around for decades, but its prominence has been seen with the advancement of digital communication. Previously, an attacker impersonated bank or government officials or law enforcement agents to deceive employees into revealing information. It's a common example found in the late 1970s where a pretexting scam was performed on the telephone system [46]. Over the years, it has incorporated the advancements and methods that have evolved into a sophisticated cyber threat [5]. One of the most noticeable scams involves impersonating executives and gaining unauthorised access to the corporate systems. Technology empowering methodology by integrating AI and deepfake to enhance the credibility and enable deception of the most vigilant individuals [19, 47].

2) Impact and Cost

Pretexting emerged as a major cyber threat, which leads to data breach [38], identity theft [31, 48], financial fraud [49], physical security breach [23, 44, 47]. Attacker impersonates a corporate entity or business partner to penetrate into organisations and extract sensitive information [31, 48]. There are also some cases where attackers steal the corporate data and threaten its disclosure [38], which leads to financial loss to the organisation, or the victim unknowingly discloses or shares the account details, payment information or engages in fraudulent transactions [49]. Pretexting remains an unsolved cybersecurity problem [31, 49]; its increasing sophistication requires individuals and businesses to be more vigilant to mitigate its risks [2, 8, 45, 50] unless an effective solution is designed.

D Reverse Social Engineering

The reverse social engineering (RSE) is a social engineering attack, which is similar to the pretexting attack, where the attacker meticulously drafts the problem scenario and manipulates the victim into revealing information. RSE is a more refined or advanced version of a pretexting attack, where the attacker drafts a problem scenario, persuading the victim into initiating communication with the attacker, asking for assistance or a solution to the problem. This uses a very different approach from other social engineering tactics, which creates a plausible problem scenario and compels the victim to ask for a solution. This approach overcomes the burden of creating trust in the attacker. Psychologically, when a person reaches someone, s/he already has faith in the person whom s/he is reaching out to. Therefore, when the victim seeks out from the attacker, it reflects an indirect trust and faith in the attacker. This makes it easier for an attacker to get the desired information in correspondence to the problem solution, and the victim also feels it is feasible and comfortable to share the sensitive information without realising the consequences, including credentials or access to the system [20].

1) Background

Reverse social engineering is widely recognised due to its unique deceptive nature and effective strategy. Unlike traditional social engineering approaches, where an attacker initiates the communication with the target and tries to deceive them into taking the intended action. As by its name, it uses the reverse tactics; instead of the attacker initiating communication, the attacker creates a well-prescribed problem scenario which compels the victim to reach out to the attacker for its solution. This remarkable attack is used in both physical and cyber worlds. It first appeared during the late 1970s or in the early 1980s, in the early age of computer networking and the era of telecommunication fraud commonly known as the "phone phreaking" age. The earliest presence of reverse engineering appeared in academic research and the hacker community, where it was stated that the victim reaching out to the attacker for help unknowingly exposed the system to the attacker. In 1996, Kevin Mitnick published a notable discussion in his book "The Art of Deception" on the reverse social engineering technique. He described how attackers use deception techniques to gain access to critical information. Since the early 2000s, beyond phone scams, it has evolved into powerful, unique strategic tactics empowered by the indirect or reverse psychology principle to exploit the "trust" factor. This inverse psychological exploitation increases the chances of deception and success ratio [30, 51].

2) Impact and Costs

Reverse social engineering attacks have been observed in both the physical world and the cyber world. In both cases, the attack methodology remains the same and effective [52]. But, in cyberspace, it is also used to share malware as a fake solution to a problem. It has a significant impact and cost to an individual or organisation, not limited to data breach, reputational damage, and financial loss [48, 53]. In 2013, the Target Data Breach 3 [54] incident was a very prominent incident; an HVAC vendor provided access credentials to an attacker posing as IT support personnel. This attack resulted in the theft of more than 40 million customer credit card numbers and over 162 million dollars' financial loss in settlements and security upgrades. The Twitter Celebrity Hack in 2020 [55], where an organisation employee reset the password and granted the attacker privileged access on the attacker's request as an IT staff member. The attack ended in reputational damage that affected individuals, resulting in more than $100,000 in losses, along with facing legal scrutiny and security revamps to Twitter.

E Baiting Attack

The baiting attack is a type of social engineering attack, also known as "road apples". This technique is used to lure the victim with enticing offers and deceive them into compromising their security. The most common and well-known form of baiting attack is a USB dropping attack [4]. It commonly uses attractive offers in both physical or digital assets or services as bait, such as a USB drive, Free software, watching an ad, or sending something among people. This offer usually aims to install the malware in victims' computers or steal their credentials or personal information [19]. It also sometimes refers to a similar act as the Trojan horse, due to presenting fake, attractive offers, software or services as legitimate and free to lure victims [1]. It is a malware-based social engineering attack that exploits the psychological factors: curiosity, greed, and [1].

1) Background

The baiting attack dated back to two decades ago, with its first documented case reported in the early 2000s. The concept of this technique dates back to its existence in the cybersecurity realm. Long before the digital era, traditional tactics bribe or bait the person to disclose the secret information, betray friends, group or company, reveal secret code and confidential information, and the victim gets paid or rewarded with a highly attractive package and incentive. However, one of the earliest baiting attacks reported in 2006, when researchers at Sunbelt Software found an infected MP3 file disguised as a pirated music track was a Trojan malware, which swiftly spread upon playing the music track file.

2) Impact and Cost

The consequences of baiting attacks have affected both individuals and organisations. It primly focus on; Data Theft such as confidential files, personal information, and corporate data [31, 38], Malware Infections such as keyloggers, spyware, or ransomware to disrupt operations or compromise security [38], Financial Losses such as fraudulent or unauthorized transactions, or demand ransoms against encrypted or stolen data [49], and Reputational Damage, such as losing customer trust and credibility due to data breaches [48]. Significant incidents of baiting attacks were determined by the U.S. Department of Defence (DoD) USB drive experiment. In 2008 [56-58], an Experiment was conducted by dropping a USB drive in a parking lot within a military facility. The study observed that about 60% picked up and inserted into a computer which was infected with the "Agent.BTZ" worm. This demonstrated the high probability and effectiveness of psychological drivers leading to success. As a consequence, the US military has banned the use of USB drives within the facility. Another experiment conducted in 2016, Google researchers also experimented [59] to measure the human behaviour against the USB drop attack, aka the Baiting attack, by placing the 297 USBs at various locations within the university campus. The result was so high, as 135/297 (45%) picked up and plugged in USB drives into their system, 290/297 (98%) USB drives were not found in their location that were picked up by people, but neither plugged in nor not. Collectively, experiments conclude with a 98% success ratio, showing that baiting attack exploits the curiosity, while 45% showed exploitation of another factor, "greed", with a very high success rate. Another significant attack occurred in 2019, the BMW and Mercedes Cyber Espionage Attack [60], where an attacker deliberately dropped a USB in the workplace disguised as job applications or supplier files. An employee of curiosity plugged in a USB drive and unknowingly injected the malware into the system. Consequently, this led to financial and reputational loss, which was estimated in millions of dollars.

F Quid Pro Quo

The Quid Pro Quo is derived from the Latin word, which means "something for something" or "this for that". It is a social engineering technique where an attacker offers enticing deals, such as free service, in exchange for desired information [48]. An attacker impersonates IT staff or a support representative to persuade the victim into providing confidential information, including login credentials and WiFi password, to assist him [48, 53]. It uses the psychological principle of reciprocity and the trust factor to exploit the human element in the security chain and execute an attack.

1) Background

The Quid Pro Quo attack has existed since the early 2000s, when the attacker started impersonating IT staff to gain system access [13]. One of the earliest documented attacks was recorded in 2013, where the attacker posed as help desk staff and offered free system updates [3]. The incident occurred when many employees trusted the attacker and gave their credentials, and the attacker gained unauthorised access to the network [43, 44].

2) Impact and Cost

The consequences of this attack can lead to Unauthorised Access [14, 48], Data Breaches [5, 38], Financial Losses [31], regulatory penalties and reputational damage. Some significant incidents in the recent past are: in 2011, an RSA Security Breach occurred when an attacker pretended to be an IT support and offered to fix the issue. The attacker deceived the employee into installing malware and stealing the confidential data. In 2015, the US healthcare industry scammed an attacker posing as a HIPAA compliance auditor offering free service. Unknowingly, an employee shared the patient data, resulting in a privacy violation penalty. Another attack was noticed in 2018 [26], where an attacker compromised a Cryptocurrency Exchange. The attacker impersonated customer support and convinced the user to verify the account. The employee was deceived into sharing login credentials, resulting in the loss of millions of dollars. The FBI & CISA Fake IT Support Scam occurred in 2021, where an attacker pretending to be federal agency staff instructed to install security updates, which included the remote access Trojans (RATs) [28]. This ended up with several private and government sector networks compromised.

G Diversion Theft Attack

The diversion theft attack is a concept of misdirecting a courier service or package delivery to a false location or manipulating the delivery of a package by installing malware or forging secret documents [19]. The core aim of this attack is to mislead the courier service by an internal or external source and to exploit the logistics, supply chain, and its delivery system [3]. Also, aims to gain access to the package or information, steal the goods or replace the goods with infected ones [2]. This attack is also known as "Corner Game" or "Round the Corner Game" [30]. This usually happens without the sender's and receiver's knowledge. The attacker also misguides the courier company, unknowingly installs the malware into the computer system as preloaded software that the company commonly use. Once these infected system enters the organisation's facility, the attacker easily has access to the corporate network [13].

7 Future Work

This work discussed the physical deception techniques which form the foundational understanding of social engineering attacks. It highlights the vulnerabilities that attackers exploit in direct and physical interaction with the victim to deceive and delve into revealing information. Besides the essential understanding, there are several critical avenues that remain to be investigated in future to expand knowledge and modern emerging threats. Some of the important avenues that researchers should consider;

- Challenges and opportunities in targeting technology to deal with physical deception techniques. This could be an interesting topic to explore how attacker leverage advanced technology measures to enhance their physical deception tactics.
- Evolving social dynamic paradigm and cultural norms require research into the psychological underpinnings of physical deception. Researchers delve into psychological factors that influence physical deception, especially in diverse cultural contexts. It is important to learn how these factors work across different cultures and their severity of impact on them.

- AI-driven surveillance system potential should be explored in future to enhance security against physical deception techniques by integrating human behavioural analysis and measuring its effectiveness.

In the evolving landscape of social engineering attacks, the enduring relevance of physical deception techniques requires ongoing research in this domain to effectively encounter these attacks and minimise their magnitude.

8 Limitations

The study has two primary limitations. First, due to financial constraints, the literature review was limited to the open-source and freely available works. Second, the proposed taxonomy is inherently interpretative and based on our synthesis of existing literature. These constraints might omit the potential and relevant studies and introduce selection bias.

We tried to minimise these limitations by using institutional access where possible, searching preprint repositories, and using citation chaining (snowballing technique) to identify seminal works. To enhance objectivity, we grounded our categories in consistently observed patterns across multiple sources and explicitly defined the criteria for attack inclusion (e.g., the requirement for direct physical interaction).

However, it is possible that a more comprehensive resource pool could reveal additional nuances or rare attack variants of physical human-based SEA. This presents a clear opportunity for future research to validate and expand upon this taxonomy with a broader literature base.

9 Conclusion

This work presented a comprehensive survey of physical human-based social engineering attacks, covered tactics, methodology, and impact that exploit innate vulnerabilities by direct physical engagement or with no engagement with the attacker's proximity. The paper provides a clear understanding of the diverse strategies that attackers employed to deceive their victims. Through detailed analysis, the study highlighted the critical role of social and psychological factors enabling these attacks to deceive humans. From impersonation technique to dumpster diving, and to eavesdropping attack demonstrates effectiveness, and the impact of physical presence that influences bypassing the security measures. The findings underscore the importance of human-centric security solutions along with technological defence measures. Technological advancement plays an important role in protecting against cyberattacks, while the human element remains the weakest and significant vulnerability in the cybersecurity chain. Over time, technology and social dynamics are evolving, which directly contribute to the social engineering landscape that requires ongoing research and adaptation for developing effective countermeasures. This work serves as a foundation for further investigation in this domain and other dimensions. By understanding these increasingly prevalent and sophisticated attacks, we can enhance our defence mechanism to mitigate the risk posed by them.

Acknowledgment

This work was carried out independently, and no external grants or funding were received for this work.

References

[1] Z. Wang, L. Sun, and H. Zhu, "Defining Social Engineering in Cybersecurity," *IEEE Access*, vol. 8, pp. 85094–85115, 2020, doi: 10.1109/ACCESS.2020.2992807.

[2] A. Ferreira, L. Coventry, and G. Lenzini, "Principles of persuasion in social engineering and their use in phishing," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9190, no. May 2017, pp. 36–47, 2015, doi: 10.1007/978-3-319-20376-8_4.

[3] M. A. Siddiqi, W. Pak, and M. A. Siddiqi, "A Study on the Psychology of Social Engineering-Based Cyberattacks and Existing Countermeasures," *Appl. Sci.*, vol. 12, no. 12, 2022, doi: 10.3390/app12126042.

[4] G. Sarkar and S. K. Shukla, "Behavioral analysis of cybercrime: Paving the way for effective policing strategies," *J. Econ. Criminol.*, vol. 2, no. August, p. 100034, 2023, doi: 10.1016/j.jeconc.2023.100034.

[5] H. Taherdoost, "Analyzing Influential Psychological Factors in Social Engineering; Human Psyche and Cybersecurity," *Psychomachina*, vol. 1, pp. 1–7, Feb. 2024, doi: 10.59388/pm00374.

[6] S. Kuraku, D. Kalla, N. Smith, and F. Samaah, "Exploring How User Behavior Shapes Cybersecurity Awareness in the Face of Phishing Attacks," *Int. J. Comput. Trends Technol.*, vol. 71, no. 11, pp. 74–79, 2023, [Online]. Available: https://doi.org/10.14445/22312803/IJCTT-V71I11P111

[7] N. F. Khan, N. Ikram, S. Saleem, and S. Zafar, *Cyber-security and risky behaviors in a developing country context: a Pakistani perspective*, vol. 36, no. 2. Palgrave Macmillan UK, 2023. doi: 10.1057/s41284-022-00343-4.

[8] Z. Alkhalil, C. Hewage, L. Nawaf, and I. Khan, "Phishing Attacks: A Recent Comprehensive Study and a New Anatomy," *Front. Comput. Sci.*, vol. 3, no. March, pp. 1–23, 2021, doi: 10.3389/fcomp.2021.563060.

[9] A. H. Washo, "An interdisciplinary view of social engineering: A call to action for research," *Comput. Hum. Behav. Reports*, vol. 4, p. 100126, 2021, doi: 10.1016/j.chbr.2021.100126.

[10] A. Abzakh and A. Althunibat, "A Review: Human Factor and Cybersecurity," *IEEE Xplore*, 2023, doi: 10.1109/ICIT58056.2023.10225828.

[11] G. Desolda, L. S. Ferro, A. Marrella, T. Catarci, and M. F. Costabile, "Human Factors in Phishing Attacks: A Systematic Literature Review," *ACM Comput. Surv.*, vol. 54, no. 8, 2022, doi: 10.1145/3469886.

[12] M. F. Arroyabe, C. F. A. Arranz, I. F. De Arroyabe, and J. C. F. de Arroyabe, "Revealing the realities of cybercrime in small and medium enterprises: Understanding fear and taxonomic perspectives," *Comput. Secur.*, vol. 141, no. February, p. 103826, 2024, doi: 10.1016/j.cose.2024.103826.

[13] W. Syafitri, Z. Shukur, U. A. Mokhtar, R. Sulaiman, and M. A. Ibrahim, "Social Engineering Attacks Prevention: A Systematic Literature Review," *IEEE Access*, vol. 10, pp. 39325–39343, 2022, doi: 10.1109/ACCESS.2022.3162594.

[14] B. S. Almutairi and A. Alghamdi, "The Role of Social Engineering in Cybersecurity and Its Impact," *J. Inf. Secur.*, vol. 13, no. 04, pp. 363–379, 2022, doi: 10.4236/jis.2022.134020.

[15] R. F. Abu Hweidi and D. Eleyan, "Social Engineering Attack concepts, frameworks, and Awareness: A Systematic Literature Review," *Int. J. Comput. Digit. Syst.*, vol. 13, no. 1, pp. 691–700, 2023, doi: 10.12785/ijcds/130155.

[16] M. Dixon-Woods *et al.*, "Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups," *BMC Med. Res. Methodol.*, vol. 6, no. 1, p. 35, Dec. 2006, doi: 10.1186/1471-2288-6-35.

[17] M. Dixon-Woods *et al.*, "How can systematic reviews incorporate qualitative research? A critical perspective," *Qual. Res.*, vol. 6, no. 1, pp. 27–44, Feb. 2006, doi: 10.1177/1468794106058867.

[18] E. Barnett-Page and J. Thomas, "Methods for the synthesis of qualitative research: a critical review," *BMC Med. Res. Methodol.*, vol. 9, no. 1, p. 59, Dec. 2009, doi: 10.1186/1471-2288-9-59.

[19] F. Salahdine and N. Kaabouch, "Social engineering attacks: A survey," *Futur. Internet*, vol. 11, no. 4, 2019, doi: 10.3390/FI11040089.

[20] J. M. Hatfield, "Social engineering in cybersecurity: The evolution of a concept," *Comput. Secur.*, vol. 73, pp. 102–113, 2018, doi: 10.1016/j.cose.2017.10.008.

[21] B. B. Gupta, A. Tewari, A. K. Jain, and D. P. Agrawal, "Fighting against phishing attacks: state of the art and future challenges," *Neural Comput. Appl.*, vol. 28, no. 12, pp. 3629–3654, 2017, doi: 10.1007/s00521-016-2275-y.

[22] A. Smith, M. Papadaki, and S. M. Furnell, "Improving Awareness of Social Engineering Attacks," in *IFIP International Federation for Information Processing 2013*, Springer, 2013, pp. 249–256.

[23] Z. Wang, H. Zhu, and L. Sun, "Social Engineering in Cybersecurity: Effect Mechanisms, Human Vulnerabilities and Attack Methods," *IEEE Access*, vol. 9, pp. 11895–11910, 2021, doi: 10.1109/ACCESS.2021.3051633.

[24] B. Banire, D. Al Thani, and Y. Yang, "Investigating the experience of social engineering victims: Exploratory and user testing study," *Electron.*, vol. 10, no. 21, 2021, doi: 10.3390/electronics10212709.

[25] A. A. Dziwa, C. CISA, and C. CEH, "How Social Engineering Bypasses Technical Controls," *ISACA J.*, vol. 5, pp. 23–26, 2022, [Online]. Available: https://www.isaca.org/resources/isaca-journal/issues/2022/volume-5/how-social-engineering-bypasses-technical-controls

[26] W. Fuertes *et al.*, "Impact of Social Engineering Attacks: A Literature Review," *Smart Innov. Syst. Technol.*, vol. 255, no. September, pp. 25–35, 2022, doi: 10.1007/978-981-16-4884-7_3.

[27] P. Burda, L. Allodi, and N. Zannone, "Cognition in Social Engineering Empirical Research: A Systematic Literature Review," *ACM Trans. Comput. Interact.*, vol. 31, no. 2, 2024, doi: 10.1145/3635149.

[28] A. Maraj and W. Butler, "Taxonomy of Social Engineering Attacks: A Survey of Trends and Future Directions," in *International Conference on Cyber Warfare and Security*, Mar. 2022, pp. 185–193. doi: 10.34190/iccws.17.1.40.

[29] H. Aldawood and G. Skinner, "Challenges of implementing training and awareness programs targeting cyber security social engineering," in *Proceedings - 2019 Cybersecurity and Cyberforensics Conference, CCC 2019*, 2019, pp. 111–117. doi: 10.1109/CCC.2019.00004.

[30] H. Aldawood and G. Skinner, "An Advanced Taxonomy for Social Engineering Attacks," *Int. J. Comput. Appl.*, vol. 177, no. 30, pp. 1–11, 2020, doi: 10.5120/ijca2020919744.

[31] N. Y. Conteh and P. J. Schmick, "Cybersecurity: risks, vulnerabilities and countermeasures to prevent social engineering attacks," *Int. J. Adv. Comput. Res.*, vol. 6, no. 23, pp. 31–38, Feb. 2016, doi: 10.19101/IJACR.2016.623006.

[32] V. Ambika, N. Shashank, T. Kayargo, and P. Pruthvi, "A Review Article on Impact of Social Engineering Attacks against Security of IoT," *Turkish Online J. Qual. Inq.*, vol. 12, no. 7, pp. 13552–13560, 2021, [Online]. Available: https://www.tojqi.net/index.php/journal/article/view/6882

[33] A. Naz, M. Sarwar, M. Kaleem, M. A. Mushtaq, and S. Rashid, "A comprehensive survey on social engineering-based attacks on social networks," *Int. J. Adv. Appl. Sci.*, vol. 11, no. 4, pp. 139–154, 2024, doi: 10.21833/ijaas.2024.04.016.

[34] A. A. Nugraha *et al.*, "Social Engineering Awareness : A Social Science Approach to Cybersecurity Education," in *The 3rd International Conference on Education Innovation and Social Science*, 2024, pp. 376–386.

[35] K. L. Chiew, K. S. C. Yong, and C. L. Tan, "A survey of phishing attacks: Their types, vectors and technical approaches," *Expert Syst. Appl.*, vol. 106, pp. 1–20, 2018, doi: 10.1016/j.eswa.2018.03.050.

[36] H. Aldawood and G. Skinner, "A Taxonomy for Social Engineering Attacks via Personal Devices," *Int. J. Comput. Appl.*, vol. 178, no. 50, pp. 19–26, 2019, doi: 10.5120/ijca2019919411.

[37] R. Alabdan, "Phishing attacks survey: Types, vectors, and technical approaches," *Futur. Internet*, vol. 12, no. 10, pp. 1–39, 2020, doi: 10.3390/fi12100168.

[38] S. K. Birthriya, P. Ahlawat, and A. K. Jain, "A Comprehensive Survey of Social Engineering Attacks: Taxonomy of Attacks, Prevention, and Mitigation Strategies," *J. Appl. Secur. Res.*, vol. 0, no. 0, pp. 1–49, Jul. 2024, doi: 10.1080/19361610.2024.2372986.

[39] A. K. Jain and B. B. Gupta, "A survey of phishing attack techniques, defence mechanisms and open research challenges," *Enterp. Inf. Syst.*, vol. 16, no. 4, pp. 527–565, 2022, doi: 10.1080/17517575.2021.1896786.

[40] B. Atkins and W. Huang, "A Study of Social Engineering in Online Frauds," *Open J. Soc. Sci.*, vol. 01, no. 03, pp. 23–32, 2013, doi: 10.4236/jss.2013.13004.

[41] D. Alharthi and A. Regan, "A Literature Survey and Analysis on Social Engineering Defense Mechanisms and Infosec Policies," *Int. J. Netw. Secur. Its Appl.*, vol. 13, no. 2, pp. 41–61, Mar. 2021, doi: 10.5121/ijnsa.2021.13204.

[42] K. Krombholz, H. Hobel, M. Huber, and E. Weippl, "Advanced social engineering attacks," *J. Inf. Secur. Appl.*, vol. 22, pp. 113–122, Jun. 2015, doi: 10.1016/j.jisa.2014.09.005.

[43] C. Sekhar Bhusal, "Systematic Review on Social Engineering: Hacking by Manipulating Humans," *J. Inf. Secur.*, vol. 12, no. 01, pp. 104–114, 2021, doi: 10.4236/jis.2021.121005.

[44] Z. Wang, H. Zhu, P. Liu, and L. Sun, "Social engineering in cybersecurity: a domain ontology and knowledge graph application examples," *Cybersecurity*, vol. 4, no. 1, 2021, doi: 10.1186/s42400-021-00094-6.

[45] J.-W. Bullée and M. Junger, "Social Engineering," in *The Palgrave Handbook of International Cybercrime and Cyberdeviance*, no. October, Cham: Springer International Publishing, 2019, pp. 1–28. doi: 10.1007/978-3-319-90307-1_38-1.

[46] B. Azizbek Zoxid ugli, "The Impact of Social Engineering on Cybercrime: Psychological Manipulation and Prevention Methods," *Int. J. Cyber Law*, p. 5, 2023, [Online]. Available: https://irshadjournals.com/index.php/ijcl/article/view/52/38

[47] K. Chetioui, B. Bah, A. O. Alami, and A. Bahnasse, "Overview of Social Engineering Attacks on Social Networks," *Procedia Comput. Sci.*, vol. 198, no. 2021, pp. 656–661, 2021, doi: 10.1016/j.procs.2021.12.302.

[48] M. Zaoui, B. Yousra, S. Yassine, M. Yassine, and O. Karim, "A Comprehensive Taxonomy of Social Engineering Attacks and Defense Mechanisms: Toward Effective Mitigation Strategies," *IEEE Access*, vol. 12, no. May, pp. 72224–72241, 2024, doi: 10.1109/ACCESS.2024.3403197.

[49] T. Rathod, N. K. Jadav, S. Tanwar, A. Alabdulatif, D. Garg, and A. Singh, "A comprehensive survey on social engineering attacks, countermeasures, case study, and research challenges," *Inf. Process. Manag.*, vol. 62, no. 1, p. 103928, 2025, doi: 10.1016/j.ipm.2024.103928.

[50] T. Grassegger and D. Nedbal, "The role of employees' information security awareness on the intention to resist social engineering," *Procedia Comput. Sci.*, vol. 181, no. 2019, pp. 59–66, 2021, doi: 10.1016/j.procs.2021.01.103.

[51] N. Duarte, N. Coelho, and T. Guarda, "Social Engineering: The Art of Attacks," *Commun. Comput. Inf. Sci.*, vol. 1485 CCIS, pp. 474–483, 2021, doi: 10.1007/978-3-030-90241-4_36.

[52] R. Montañez, A. Atyabi, and S. Xu, "Social engineering attacks and defenses in the physical world vs. cyberspace: A contrast study," in *Cybersecurity and Cognitive Science*, Elsevier, 2022, pp. 3–41. doi: 10.1016/B978-0-323-90570-1.00012-7.

[53] A. Lopes, H. S. Mamede, L. Reis, and A. Santos, "Common Techniques, Success Attack Factors and Obstacles to Social Engineering: A Systematic Literature Review," *Emerg. Sci. J.*, vol. 8, no. 2, pp. 761–794, 2024, doi: 10.28991/ESJ-2024-08-02-025.

[54] M. Hasan, "Common Cybersecurity Vulnerabilities: Software Bugs , Weak Passwords , Misconfigurations , Social Engineering," *Glob. Mainstream J. Innov. Eng. Emerg. Technol.*, vol. 03, no. 04, pp. 42–57, 2024, doi: 10.62304/jieet.v3i4.193.

[55] Y. Kano and T. Nakajima, "Trust factors of social engineering attacks on social networking services," *LifeTech 2021 - 2021 IEEE 3rd Glob. Conf. Life Sci. Technol.*, no. LifeTech, pp. 25–28, 2021, doi: 10.1109/LifeTech52111.2021.9391929.

[56] B. Stilwell, "The worst cyber attack in DoD history came from a USB drive found in a parking lot," wearethemighty. Accessed: Feb. 20, 2025. [Online]. Available: https://www.wearethemighty.com/mighty-history/worst-cyber-attack-usb/

[57] N. Shachtman, "Under Worm Assault, Military Bans Disks, USB Drives," Wired.com. Accessed: Feb. 20, 2025. [Online]. Available: https://www.wired.com/2008/11/army-bans-usb-d/

[58] A. Moscaritolo, "Military's ban of USB thumb drives highlights security risks," scworld.com. Accessed: Feb. 18, 2025. [Online]. Available: https://www.scworld.com/news/militarys-ban-of-usb-thumb-drives-highlights-security-risks

[59] M. Tischer *et al.*, "Users Really Do Plug in USB Drives They Find," in *2016 IEEE Symposium on Security and Privacy (SP)*, IEEE, May 2016, pp. 306–319. doi: 10.1109/SP.2016.26.

[60] Robert Abel, "Luxury hackers crack Bimmers and Benzes, a tale of BMW's bugs and Mercedes-Benz thugs," scworld.com. Accessed: Feb. 15, 2025. [Online]. Available: https://www.scworld.com/news/luxury-hackers-crack-bimmers-and-benzes-a-tale-of-bmws-bugs-and-mercedes-benz-thugs

[61] I. A. M. Abass, "Social Engineering Threat and Defense: A Literature Survey," *J. Inf. Secur.*, vol. 09, no. 04, pp. 257–264, 2018, doi: 10.4236/jis.2018.94018.

[62] T. Mashiane and E. Kritzinger, "Cybersecurity Behaviour: A Conceptual Taxonomy," in *12th IFIP International Conference on Information Security Theory and Practice (WISTP)*, Brussels, Belgium, 2019, pp. 147–156. doi: 10.1007/978-3-030-20074-9_11.

[63] M. R. Arabia-Obedoza, G. Rodriguez, A. Johnston, F. Salahdine, and N. Kaabouch, "Social Engineering Attacks A Reconnaissance Synthesis Analysis," in *2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, IEEE, Oct. 2020, pp. 0843–0848. doi: 10.1109/UEMCON51285.2020.9298100.

[64] M. Hijji and G. Alam, "A Multivocal Literature Review on Growing Social Engineering Based Cyber-Attacks/Threats During the COVID-19 Pandemic: Challenges and Prospective Solutions," *IEEE Access*, vol. 9, pp. 7152–7169, 2021, doi: 10.1109/ACCESS.2020.3048839.

[65] N. A. Odeh, D. Eleyan, and A. Eleyan, "A Survey of Social Engineering Attacks :," *J. Theor. Appl. Inf. Technol.*, vol. 99, no. 18, 2021.

[66] N. Mashtalyar, U. N. Ntaganzwa, T. Santos, S. Hakak, and S. Ray, "Social Engineering Attacks: Recent Advances and Challenges," in *HCI for Cybersecurity, Privacy and Trust (HCII 2021)*, 2021, pp. 417–431. doi: 10.1007/978-3-030-77392-2_27.

Appendix A

Table A1. Summary of issues in 31 examined classification schemes and 18 taxonomies

Ref	Year	Classification Factor	Attack(s) Discussed	Observation / Non-SEA	Attacks	Non-SEA Err (%)
[40]	2013	Computer & Human	**Trojan** e-mail and phishing messages, advance-fee fraud, impersonation, persuasion, bribery, shoulder surfing, and dumpster diving.	Persuasion (not attack), bribery (not attack). Suggested but didn't classify into proposed classification scheme	8	37.50
[42]	2015	Operator, Type, Channel	Phishing, Shoulder surfing, Dumpster diving, RSE, Waterholing, APT, Baiting, ASE, FakeProfile, Sybil, Spear-phishing,	Advanced persistent threat, Sybil attacks. But didn't classify attack into suggested categories.	10	20.00
[31]	2016	Human & Technical	Phishing, Pretexting, Baiting, Quid pro quo, Tailgating, dumpster diving,	Suggested but didn't classify into proposed classification scheme	6	0.00
[21]	2017	SE & Malware	Website phishing, Email phishing, Spear phishing, Key loggers (Hardware & Software), screen loggers, Session hijacking, DNS phishing, Host file poisoning, System reconfiguration attacks, Pharming, Proxy attack, Content injection, Phishing through search engines, Spandexing	Key loggers (Hardware & Software), screen loggers, Session hijacking, DNS poisoning, Host file poisoning, System reconfiguration attacks, Proxy attack, Content injection, Spandexing	16	75.00
[35]	2018	Medium	drive-by-download, MITM, XSS, tabnapping, spear phishing, whaling, SEO Phishing, session fixation, malvertising, social engineering, JavaScript obfuscation, browser vulnerabilities, mobile phone, cloud computing and WiPhishing or Evil Twins, SQL injection, typo-squatting, and sound-squatting, ClickJacking, CAPTCHA attack, Trojan, Bot Attack (BotNet), APT, Tabnapping,	MITM, XSS, session fixation, malvertising, social engineering, browser vulnerabilities (generic term), mobile phone (generic term), cloud computing (generic term), SQL injection, CAPTCHA attack, Trojan, Bot Attack (BotNet), APT, Tabnapping,	24	45.83
[61]	2018	Computer & Human	temptation to have something rare, Show the similarities with the target, Pay back the favour, style of flattery, Move with the flow, Pretexting, Reverse Social Engineering, Phishing, Spear phishing, Watering hole attack	Pay back the favour, show the similarities, style of flattery, temptation to have something rare. These are too general terms and do not reflect a specific attack	10	50.00
[36]	2019	Device	phishing, CSRF, malware, email, and popups. Session hijacking attacks, connection-oriented protocol attacks, SQLi, SEO poisoning, Vishing, botnets, rootkits, buffer overflow, worms, viruses, Trojan horses, bot attacks, and spywares.	CSRF, malware, email (generic term), session hijacking attacks, connection-oriented protocol attacks (generic term), SQLi, botnets, rootkits, buffer overflow, worms, viruses, Trojan horses, bot attacks, and spywares.	18	83.33
[62]	2019	Behavior	-	Not discussed any attack, but suggested classification based on behavioural factors.	0	0.00

[45]	2019	Communication Type	Vishing, Phishing, Physical (F2F), Smishing,	Face-to-Face (Too generic term). Suggested but didn't classify.	4	0.00
[19]	2019	None	phishing, impersonation on help desk calls, shoulder surfing, dumpster diving, stealing important documents, diversion theft, fake software, baiting, quid pro qou, pretexting, tailgating, Pop-Up windows, Robocalls, ransomware, online social engineering, RSE, fake software, SMSishing, physical access, phishing, spear-phishing, whaling, vishing, BEC, interactive voice response phishing, Pharming attacks	Ransomware, Online Social Engineering (general term), Phone Social Engineering Stealing (generic term), Important Documents (General action), Whitelisting flow. Didn't classify the attack or proposed classification scheme.	26	11.54
[30]	2020	Human & Computer	Impersonation, Pretexting, Tailgating, Quid Pro Quo, Diversion Theft, Phishing, Spear-Phishing, Whaling, Clone-Phishing, Pharming, Website Phishing, Pop-Up Windows, XSS, Spyware, Malware, Worms, Rootkits, Botnets, Ransomware, Baiting, Trojans, KeyLoggers, Screen Loggers, Malicious Links, Fake Groups, Digital Impersonation, Fake Profiles, DNS Poisoning, SSL Attack, MIMT, Compromised Web Server Attack, Host File Poisoning, Session Hijacking, SEO Phishing, SMSishing, Mobile Apps, Vishing, Similarity Attacks, Forwarding Attacks, Background Attacks, Notification Attack, Floating Attacks, (VoIP) Phishing	Ransomware, Botnets, Cross-Site Scripting (XSS), Rootkits, Malware, Spyware, Worms, Trojans, Key Loggers, Screen Loggers, Session Hijacking, Host File Poisoning, Compromised Web Server Attack, Man-in-the-Middle Attack, Secure Socket Layer (SSL) Attack, DNS Poisoning, Background Attacks, Floating Attacks,	40	55.00
[37]	2020	Medium	Phishing (Email), Vishing, Smishing, Whaling, Spear Phishing, BEC, QRishing, Wiphishing, Social media phishing, Typo Squatting, Sound Squatting, Social Engineering, XSS (stored XSS and reflected XSS), SQLi, APT, browser sniffing, DNS cache poisoning, CAPTCHA Attack, Drive-by Download, Malvertizing, Session hijacking, Botnets, Browser Vulnerabilities, Tab-Napping, Skill squatting, 404 Error Manipulation, Click Jacking, Malicious Browsing Extensions Browser, MIMT, Application phishing, GUI-Squatting, Session Fixation, JavaScript Obfuscation,	Email (generic term), IM (Too generic term), Cross-Site Scripting (XSS), CAPTCHA Attack, Social Engineering (generic term), Browser Vulnerabilities, Tab-Napping, 404 Error Manipulation, Malicious Browsing Extensions, MIMT, Mobile Phones (Generic Term), GUI-Squatting, Session Fixation, JavaScript Obfuscation, Wiphishing,	35	60.00
[63]	2020	Operator, Type, Channel	online hoaxes or advance-fee fraud, credit card, high-tech disaster and identity theft fraud, web and e-mail spoofing, instant messaging fraud, and spamming, phishing	credit card, high-tech disaster (generic term). Not discussed any attack, but suggested classification type based on Operator, Type, and Channel.	8	25.00
[41]	2021	Computer & Human	BEC, Vishing, Smishing, Pharming,	Discussed the classification scheme, but didn't classify	4	0.00
[64]	2021	Social, Technical, Socio-Technical, and Physical	Trojan (20 variants), SQLi, Espionage, DDoS, DoS, Brute-Force, Impersonation, Typo-squatting, Bots (7 Variants), Spoofing, Cyber-sabotage, e-skimming, Bit-Squatting, BEC, Mal-URLs, CryptoMinors, Malware (6, Other-13 variants, Ransomware-31 variants), Worms.	No attack discussed. Only the classification scheme was discussed, but it was not classified.	91	95.60
[32]	2021	Generic & Targeted	Phishing, Spare-Phishing	Discussed the classification scheme, but didn't classify	2	0.00

[65]	2021	Operator, Methods, Nature	Impersonation, Shoulder surfing, Dumpster diving, Phishing, Spear-phishing, Whaling, Vishing, Angler, Baiting, Pretexting, Tailgating, Ransomware, Pop-up windows, Scareware, Smishing, Quid pro quo,	Ransomware, Phone Scam (generic term). Suggested classification based on Operator, Method and Nature. But, didn't classify the attack.	17	11.76
[8]	2021	Technical & Deception	Phishing, Spare-phishing, Whaling, Vishing and Smishing, Spoofed Website, Soshing, Malware, Key Loggers, Screen Loggers, Viruses, Worms, Spyware, Adware, Ransomware, RootKit, Session Hijacking, Web Trojans, Hosts File Poisoning, System Reconfiguration Attack, MIMT, Content Injection Phishing, Pharming, DNS Piosoning, Data Theft, SEO Phishing, URL Obfuscation Attack, ARP Poisoning, DNS spoofing,	Ransomware, Malware, Key Loggers, Screen Loggers, Viruses, Worms, Data Theft (output not attack), RootKit, Session Hijacking, Web Trojans, Hosts File Poisoning, System Reconfiguration Attack, MIMT, Content Injection Phishing, ARP Poisoning, DNS spoofing	29	68.97
[66]	2021	Computer & Human	Persuasion, Impersonation, Tailgating, Piggybacking, Shoulder Surfing, Dumpster Diving, Phishing, Vishing, Watering Hole, Bot Attacks, Brand Theft, Typosquatting, Baiting,	Bot Attacks, Brand Theft (Generic term)	13	15.38
[44]	2021	Computer & Human	Impersonation, Shoulder Surfing, Pigggybacking, Trailing, Pretexting, Phishing, Web-Phishing, Smishing, Whaling, WiFi-Phishing, Trojan, Baiting, Water Hole, ATP. Ransomware, XSS, CSRF, Pop-up Window, F2F, MIMT, Dumpster Diving, Vishing, RSE, Spear-phishing, Honey trap,	Manipulating Conversation (generic term), Trojan attack, and Honey Trap.	25	40.00
[23]	2021	None	Impersonation, Shoulder Surfing, Trailing, Pretexting, Phishing, Smishing, Whaling, Trojan, Water Hole, Vishing, RSE, Spare-phishing, Honey trap,	Honey Trap, Manipulating conversation (too generic term, not an attack) No classification scheme, and didn't classify	13	7.69
[43]	2021	Interaction	Impersonation, Shoulder surfing, Dumpster diving, Eavesdropping, Vishing, Tailgating, Quid pro quo, Phishing, Baiting, Pretexting, Water holing, Pop-up window	Found misclassification such as Vishing, Shoulder surfing and dumpster diving categorised in direct interaction, and phishing & pretexting in the indirect interaction category.	12	0.00
[39]	2022	Social, Technical, Mobile, Others	Website spoofing, Email spoofing, spear-phishing, Ransomware, Trojan, Content Injection, Keylogger, Screen Logger, MIMT, Smishing, Vishing, WiFi-Phishing, MalApp, Compromised Web-Server, Botnet, DNS Poisoning, Soshing	Website spoofing, Email Spoofing, Trojan horse, Content injection, Key logger, screen logger, man-in-the-middle attack, DNS poisoning, ransomware, compromised web server,	17	41.18
[11]	2022	Medium	drive-by-download, MITM, XSS, tabnapping, spear phishing, whaling, SEO Phishing, session fixation, malvertising, social engineering, JavaScript obfuscation, browser vulnerabilities, mobile phone, cloud computing and WiPhishing or Evil Twins, SQL injection, typo-squatting, and sound-squatting, ClickJacking, Pharming, Trojan, Tabnapping,	Browser Vulnerability, Clickjacking, XSS, Javascript Obfuscation, MIMT (Man-in-the-Middle), DNS Poisoning, Session Fixation, SQLi, Tabnapping, Cloud Computing, Mobile-Based (too general terms)	23	34.78

[52]	2022	Masquerading, Physical, Digital	pretexting, impersonation, physical reverse engineering (physical RE), and tailgating, passive surveillance, dumpster diving, open source reconnaissance, Vishing, Email Scam, Phishing, Smishing, Spear-phishing, Whaling, Soshing, Catphishing, Baiting, Shoulder Surfing, Web-RSE, Social-RSE, angler phishing,	Email scams (generic term), Web-RSE, and Social-RSE do not justify. Open source reconnaissance is an activity.	19	15.79
[28]	2022	Human & Technical	phishing	Discussed the classification scheme, but didn't classify	1	0.00
[5]	2023	Behaviour, Emotion, Social, Cognitive	Phishing, Pretexting, Impersonation, Quid pro quo, Tailgating exploits	Multiclass categorisation issue	5	0.00
[4]	2023	Behavior	Dumpster Diving, fake profiles, Phishing, Vishing, Smishing, Fake Websites, Hosting Fake Online Advertisements, Malvertising, USB Dropping, Juice Jacking, Email Bombing, QRshing, SEO-phishing, Befriending, Sextortion:,	Web Application Hosting, Email Bombing, Juice Jacking,	14	14.29
[38]	2024	Human & Software	Impersonation, Pretexting, Quid pro quo, Dumpster diving, Shoulder surfing, Phishing, Spamming, Spear phishing, Whaling, Pharming, Website phishing, Pop-up windows, XSS, Malware, Virus, Worms, Spyware, Trojans, Botnets, Adware, Rootkits, Ransomware, Wiper, Mobile malware, Fileless malwar, Keyloggers, Screen Logger, Fake groups, Fake profiles, Malicious links, Digital impersonation, Baiting, Tailgating, watering hole, Sybil, Vicinity, Plug-in, Session hijacking, Host file poisoning, Graph-based, MIMT, SEO-phishing, Compromised web server, SSL attack, RSE, DNS poisoning, Smishing, Malicious apps, Vishing, Phishing,	Malware, Virus, Worms, Spyware, Trojans, Botnets, Rootkits, Ransomware, Wiper, Fileless malware, Key loggers, Social network (generic term), Cross-site scripting (XSS), Pop-up, Host file poisoning, Compromised web server, Man-in-the-Middle (MitM), SSL attack, DNS poisoning, Malicious apps (generic term), Plug-in attacks, Session hijacking, Graph-based attack, Spamming, Sybil attack	49	48.98
[33]	2024	No Classification	Email Phishing, Phishing, Spear Phishing, Whaling, Smishing, VIsshing, Angler, Romance Scam, Online Dating Scam, Cat Phishing, Military Romance, Investment Scam, Social Media Scam, Baiting, USB Drop, Fake Job Posting, Fake Software Downloading, Free WiFi, Online Surveys, RSE Pretexting, Tailgating, Impersonation, Thread Hijacking, Flamewares, Trolling, Spamming, Astronfing, RSE, Scareware Fake Anti Virus, Ransomware, Tech Support	Hijacking (generic term), Flamewares, Scareware, Ransomware, Astronfing, Online Survey (generic term), Social Media Scam (generic term). Didn't classify attacks, and there was a lack of description and evidence in many attacks	30	46.67
[48]	2024	Environment	Phishing, Spear phishing, Smishing, Vishing, Whaling, Baiting, quid pro quo, watering hole, Dumpster diving, Pretexting, shoulder surfing, piggybacking, Scareware, RSE, keyloggers, spyware,	key loggers, spyware,	16	12.5

Table A2. Comparative analysis of existing classification methodologies with the proposed taxonomy.

Ref	Year	Non-SEA	Misclassification	Multiclass	Issue/Limitation
[40]	2013	✓	✓	—	Doesn't classify as discussed, Non-SEAs included, e.g. Trojan, leads to classification ambiguity.
[42]	2015	✓	✓	✓	Partial attacks classified, Non-SEAs included, Multi-class categorisation, e.g. Phishing in more than 1 branch (Ref Lit. Table 1). Leads classification ambiguity
[31]	2016	×	—	—	Doesn't classify as discussed, leads to Classification Ambiguity.
[21]	2017	✓	✓	—	The classification scheme doesn't cover physical attacks. Non-SEAs are included, e.g. Session Hijacking, Misclassification. Leads to Classification Ambiguity
[35]	2018	✓	✓	✓	The classification scheme covers only internet attacks; SE itself is listed as an attack, Non-SEAs are included, e.g. XSS, Misclassification error. Multiclass error, e.g. Literature Fig.2, leads to Classification Ambiguity.
[61]	2018	✓	✓	—	Non-SEAs, e.g. "style of flattery", lead to Classification Ambiguity.
[36]	2019	✓	✓	✓	Multi-class categorisation error, Non-SEAs included, e.g. CSRF, SQLi, Multiclass error found, e.g. literature Fig.1, covers only internet-based attacks. Leads to Classification Ambiguity
[62]	2019	×	—	—	Conceptualised classification doesn't classify as discussed, leading to Classification ambiguity.
[45]	2019	×	×	—	Doesn't classify as discussed, doesn't cover all attacks, e.g. Dumspter Diving, leads to Classification Ambiguity.
[19]	2019	✓	✓	—	Does not classify as a scheme. Non-SEAs, e.g. Ransomware, lead to classification ambiguity.
[30]	2020	✓	✓	×	Non-SEAs included, e.g. Rootkit & SSL. Leads to classification ambiguity
[37]	2020	✓	✓	—	Doesn't cover physical attack, Non-SEAs included, e.g. APT, XSS. SE is also listed as an Attack. Misclassification error leads to classification ambiguity.
[63]	2020	✓	✓	—	Doesn't classify as discussed, not cover physical attacks, Non-SEAs included, e.g. Spamming. Leads to classification ambiguity.
[41]	2021	×	—	—	Doesn't classify the attack as discussed.
[64]	2021	✓	✓	—	Doesn't classify as discussed, Non-SEA included, e.g. Brute Force, Misclassification error. Leads to classification ambiguity
[32]	2021	×	—	—	Doesn't classify as discussed; only two variants of Phishing attacks are discussed.
[65]	2021	✓	✓	—	Doesn't classify as discussed, Non-SEA included, e.g. Ransomware. Leads to classification ambiguity
[8]	2021	✓	✓	×	Doesn't cover physical attacks, Non-SEAs included, e.g. Virus & Worms. Leads to classification ambiguity

[66]	2021	✓	✓	×	Non-SEAs included, e.g. Bot Attack. Persuasion is a principal & technique used in an attack, not an attack itself. Misclassification error. Leads to classification ambiguity
[44]	2021	✓	✓	×	Non-SEAs included, e.g. APT, CSRF. Influence & deception are techniques used in an attack, not an attack itself. Misclassification error. Leads to classification ambiguity.
[23]	2021	✓	✓	×	Non-SEAs included, e.g. Trojan. Misclassification error. Leads to classification ambiguity.
[43]	2021	×	✓	×	Ambiguous distribution, e.g. vishing categorises as a direct attack, phishing and Pretexting as indirect, which contradict the classification definition. Misclassification
[39]	2022	✓	✓	✓	SE itself is listed as a phishing subtype, Non-SEAs included, e.g. Ransomware, Trojan, BotNet.
[11]	2022	✓	✓	✓	Non-SEA included, e.g. SQLi. Cloud Computing and Mobile Apps are not attacks, e.g. Ref Lit. Figure.2.
[52]	2022	✓	✓	✓	Non-SEA included, e.g. passive surveillance is not an attack. Multiclass error, e.g. Shoulder Surfing as Physical Access-based, and Tailgating as Physical world-based.
[28]	2022	×	—	—	Doesn't classify the attack as discussed.
[5]	2023	×	×	✓	Only a few are classified, observed, and shared behavioural/psychological aspects among attacks. This leads to classification ambiguity.
[4]	2023	✓	✓	—	Doesn't classify as discussed, Non-SEAs included, e.g. Email Bombing. Leads to classification Ambiguity
[38]	2024	✓	✓	×	Non-SEAs included, e.g. Wiper & SSL. Malicious App is not an attack, but an attack vector. Misclassification error. Leads to classification ambiguity.
[33]	2024	✓	✓	✓	Non-SEAs included, e.g. Spamming & Ransomware. Multiclass error, e.g. RSE, is present in two categories. Leads to classification ambiguity
[48]	2024	×	×	✓	Multiclass error, e.g. Phishing, Pretexting, Tailgating, classified in more than 1 class. This leads to classification ambiguity.
This Study	*2025*	×	×	×	*Overcome the classification ambiguity, clear distribution, independent of factors that are shared among different attacks, such as technology, communication type, behaviour, environment, and device. This will help in understanding attacker psychology and approaches within the target proximity and at a distance.*

✓ **Error Exist** × ***Error Not Exist*** — **Undetermined (due to absence of taxonomy or no classification)**