Cyber Social Engineering Kill Chain

Rosana Montañez Rodriguez[1] and Shouhuai Xu[2(✉)]

[1] Department of Computer Science, University of Texas at San Antonio, San Antonio, TX 78249, USA
rosana.montanezrodriguez@my.utsa.edu

[2] Department of Computer Science, University of Colorado Colorado Springs, Colorado Springs, CO 80918, USA
sxu@uccs.edu

Abstract. Cyber attacks are often initiated with a social engineering attack to penetrate a network, which we call Cyber Social Engineering (CSE) attacks. Despite many studies, our understanding of CSE attacks is inadequate in explaining why these attacks are prevalent and why humans are still the weakest link in cybersecurity. This paper aims to deepen our understanding of CSE attacks and help design effective defenses against them. Specifically, we propose a framework, dubbed *CSE Kill Chain*, for systematically modeling and characterizing CSE attacks. To demonstrate the usefulness of the framework, we perform a case study in which we apply it to analyze a real-world CSE attack.

Keywords: Cybersecurity · Cyber attacks · Cyber social engineering kill chain · Human factors · Social engineering · Cybersecurity modeling · Science of cyber security

1 Introduction

Cyber Social Engineering (CSE) attacks are cyber attacks that use humans as an attack vector. They have become prevalent as humans are often the weakest link in cybersecurity. For example, Verizon Annual Data Breach Investigation Report [6] shows that 85% of the data breach incidents occurring in the year 2020 involved a human component. Although CSE attacks are prevalent, our understanding of them remains to be improved as the prevalence of attacks suggests the lack of effective defenses.

The lack of effective defenses can be attributed to the fact that most existing defenses against CSE attacks, or CSE defenses for short, focus on detecting and filtering social engineering messages [30]. However, this approach fails to address the source of the problem: poor security decision-making resulting from errors in human judgment. As a consequence, the inadequate understanding of what makes humans vulnerable to CSE attacks results in limited approaches to mitigating this inherent vulnerability. This problem motivates the present study, which aims to propose a systematic framework for characterizing CSE attacks in

C. Su et al. (Eds.): SciSec 2022, LNCS 13580, pp. 487–504, 2022.
https://doi.org/10.1007/978-3-031-17551-0_32

a fashion similar to how we characterize technological cyber attacks. In addition to possibly leveraging the resulting framework to guide the design of effective defenses, it can also facilitate the incorporation of CSE attacks into existing technological cyber attacks frameworks.

Our Contributions. In this paper, we make three contributions. First, we establish a direct relationship between psychological constructs and CSE attacks. By doing so, we provide a "bridge" to facilitate cross-pollination between the fields of psychology, human factors, and cybersecurity. This bridge enables researchers across multiple disciplines to explore new approaches to defending against CSE attacks by leveraging the knowledge from the relevant fields.

Second, the preceding bridge guides us to propose a framework, dubbed *CSE Kill Chain*, for modeling CSE attacks. This means that the framework accommodates human psychology by relating cognitive and social psychological constructs to CSE attacks. This also means that as the term suggests, the framework is partly inspired by the Lockheed-Martin Cyber Kill Chain [21], which focuses on technological cyber attacks, not CSE attacks. Rather than offering adaptions to the Lockheed-Martin Cyber Kill Chain, the framework further incorporates elements of the MITRE ATT&CK framework [51], especially the concepts of attack *tactics* and *techniques*. Therefore, the present study can achieve the best of both worlds (i.e., these two frameworks) when coping with CSE attacks. This alignment with well-accepted frameworks is important for practitioners and researchers to adopt the CSE Kill Chain framework. All these together make the framework a foundation for studying CSE attacks in a systematic fashion.

Third, in principle, the framework can have many applications, such as modeling real-world CSE attacks and leveraging these models to prevent and/or detect ongoing CSE attacks, which requires the availability of datasets. To demonstrate the usefulness of our framework, we present a case study to provide an in-depth analysis of a real-world, new CSE attack. The case study leads to valuable insights, including: (i) the framework can guide the identification of the connection between a cyber attack and a previously known scam in the physical world; (ii) the framework can extend the knowledge of CSE attacks exploiting social media; and (iii) the framework can demonstrate how human psychological constructs can be weaponized.

Related Work. The prior studies that are closely related to the present one are the attempts at understanding CSE attacks. These attempts typically adapt social engineering attacks in the physical world to cyberspace. As a consequence, they fail to capture the attacker-user-defender interactions that are unique to CSE attacks and absent in the physical world (e.g., the success of a CSE attack may require the cooperation of a victim user, while evading the defenses that are employed). Moreover, they fail to establish connections between human psychological vulnerabilities and CSE attacks, fail to integrate CSE attacks with cyber attacks frameworks, and fail to provide insights into developing defenses against CSE attacks. Four representative prior attempts are briefly reviewed below.

The earliest study is the Social Engineering Cycle (SEC) [33] based on the experience of the authors on social engineering attacks in the physical world. It divides a social engineering attack into four phases: research, developing rapport and trust, exploiting trust, and utilizing information. However, SEC is a high-level description of social engineering attacks and does not present any tactical or operational guidance that can be actioned for cyber defense. Our framework fills this void by mapping social engineering attack processes to existing cyber attack processes, specifically the MITRE ATT&CK framework [51].

The Social Engineering Attack Framework [36] refines the aforementioned SEC and divides social engineering attacks into six phases: attack formulation (newly added), information gathering (replacing "research" in SEC), preparation (newly added), developing relationship (replacing "developing rapport and trust" in SEC), exploiting relationship (replacing "exploiting trust" in SEC), and debrief (replacing "utilizing information" in SEC). However, this framework does not explore the psychological constructs that make social engineering attacks successful.

The Attack Cycle (AC) [48] follows an approach which is similar to the aforementioned SEC. It divides a social engineering attack into four phases: information gathering, establishing relations and rapport, exploitation, and execution. Although AC is more detailed than SEC, it neither describes any attack-victim interactions nor relates psychological principles to social engineering attacks.

The Cycle of Deception framework [38], which is based on criminal manipulations (i.e., grooming), divides social engineering attacks into five phases: goal and plan, map and bond, execution, recruiting and cloaking, and evolving/regressing. However, it does not describe how to apply each phase in practice. For example, it is unclear how to execute the recruiting and cloaking phase, which requires a continuous relationship with the victim, but this relationship is not always present in social engineering attacks.

Putting the present study into a broader context, it is important to realize that understanding, characterizing, and incorporating CSE attacks is integral to the emerging Science of Cyber Security since CSE attacks are integral to most, if not all, cyber attacks. For example, the Cybersecurity Dynamics framework aims to explicitly model human factors from a holistic perspective [63–65,70]. Moreover, the specific family of Cybersecurity Dynamics models known as preventive and reactive defense dynamics models [17,27–29,61,62,67–69,71] have explicitly modelled *pull-based* cyber attacks, which include CSE attacks (e.g., phishing and drive-by download).

Paper Outline. The remainder of this paper is organized as follows. Section 2 presents the CSE Kill Chain framework. Section 3 describes a case study on applying the CSE Kill Chain framework to analyze a real-world CSE attack. Section 4 discusses the limitations of the present study. Section 5 concludes the paper with open problems for future research.

2 The Cyber Social Engineering (CSE) Kill Chain

The premise of the CSE Kill Chain framework is human psychology. From a psychological perspective, a successful CSE attack is when the attacker's external sensory input, combined with the victim's internal cognitive processes, leads the victim to behave in the way desired by the attacker (e.g., clicking a link in a malicious email). In what follows, we start with a discussion on the cognitive and psychological foundation of the CSE Kill Chain framework, including the following psychological constructs: information processing, trust, framing, persuasion, deception, risk acceptance, and attention. These constructs collectively lead to the desired behavior.

2.1 Cognitive and Psychological Foundation of CSE Kill Chain

The success of a CSE attack depends on the attacker's ability to influence a victim individual's response. This is because, in order for an attacker to accomplish this goal, the attacker must influence the victim's internal cognitive processes by manipulating the elements in the attack (e.g., tailored message in spear-phishing). The cognitive processes pertinent to CSE attacks are the ones that are involved in human information processing, which is naturally affected by psychological constructs such as trust, attention, and risk. Table 1 summarizes the relationships between the constructs that are discussed in the present subsection and the techniques that are described later in Sect. 2.2. The relationships are elaborated below.

Table 1. Mapping between psychological constructs and CSE attack techniques

Psychological Constructs	CSE Kill Chain Technique	
Trust	Trusted Relationship (CSE-1201)	Affection Trust (CSE-T1204)
Framing	Personalization (CSE-T1110)	Contextualization (CSE-T1101)
Persuasion	Persuasion (CSE-T1108)	Foot-in-the-Door (FITD) (CSE-T1205)
Deception	Impersonation (CSE-T1102)	Visual Deception (CSE-T1103)
	Website or App. Spoofing (CSE-T1104)	URL Spoofing (CSE-T1105)
	Rev. Social Engineering (CSE-T1202)	Pretexting (CSE-T1203)
	Scamming Techniques (CSE-T1207)	
Risk Acceptance	Incentives & Motivators (CSE-T1109)	Quid-Pro-Quo (CSE-T1206)
Attention	Urgency Cues (CSE-T1106)	Attention Grabber (CSE-T1107)

Human information processing combines internal cognitive processes with external stimuli and information. The output of this process is reflected by an individual's response to the external input. Some researchers believe that there is a dual-processing system involving intuitive (heuristic) and analytic thinking [23]. Intuitive thinking is automatic and fast as it is based on general rules and recognizable patterns; analytical thinking is often slow because it is effortful and involves deliberations and deep analyses [23]. Several models attempt to explain the interactions between these two systems [7,41,42]. We leverage the three-stage dual-processing model described in [41] as it provides the insights relevant to the present study, supported by observations of several social engineering experiments. Specifically, this model states that conflict monitoring mediates the switching from intuitive to analytic processing, where a conflict is a deviation from the expected rules or known patterns. Once the analytical processing is engaged, there are two possible outcomes: the individual can generate a different response to resolve the conflict (i.e., decouple) or justify (i.e., rationalize) the initial response to resolve the conflict. Consequently, a failed social engineering attack is when message inconsistencies are detected, triggering suspicion [14, 53,58]. A successful social engineering attack is when conflict is not detected, or when conflict is detected, the individual still rationalizes their decision to comply. The later case helps explain why individuals can knowingly fall victims to phishing scams [26]. Using the preceding model as a point of departure, we can dissect the elements of a successful social engineering attack into two scenarios, which will guide us in designing the CSE Kill Chain framework.

In the first scenario (i.e., conflict not detected), the attacker avoids raising conflict by projecting trust, which is the belief that the other party is acting in good faith. Trust counters suspicion in both the real world [11,24] and social engineering [35]. For example, a social engineering message that projects trust is indistinguishable from a benign message and thus encourages the recipient to comply. The psychological constructs relevant to our research that encourage compliance are framing, persuasion and deception. The effect of framing and persuasion on compliance has been the subject of extensive research in consumer psychology, and marketing [10,23,42]. The effect of deception on compliance has been studied in research on the psychology of consumer scams [25,26,50]. Our previous studies [30,34,35] show that these psychological constructs are relevant to cyber social engineering attacks. At a high level, framing is a technique that can be used to increase the appeal of a social engineering message by manipulating its interpretation to evoke a particular emotion [23]; framing can be used in social engineering through personalization and contextualization. Persuasion is the use of arguments to encourage a specific behavior [10]; the use of persuasion techniques (e.g., foot-in-the-door) in CSE attacks is discussed in [35]. Deception is the use of arguments to encourage a false belief [5]. Deception techniques are widely used in CSE attacks, and may involve using visual elements associated with a brand, mirroring the "look-and-feel" of legitimate websites or applications, mimicking legitimate website addresses, or encouraging communication with the attacker [34].

In the second scenario (i.e., rationalization), we observe that social engineering attack can succeed because the individual justifies their initial response and accepts the risk of the action. Risk acceptance is a process through which an individual tolerates the prospect of a potential loss. Rationalization may lead an individual to believe that the perceived reward of complying with the request outweighs the perceived potential loss [19,26,30]. Rewards commonly used in CSE attacks are monetary rewards in return of a favor, free goods, or the opportunity to perform a good deed.

An attacker can manipulate attention in both scenarios to make the message noticeable. Attention is a key component of human information processing as it moderates the prioritizing of information [57]. Attention is directed to the most salient information. To increase CSE message saliency, an attacker can use emotional triggers like urgency cues or graphical and auditory elements to redirect attention, such as attention grabber.

2.2 Framework Overview

Figure 1 presents a high level view of the framework, which is based on the CSE attacks lifecycle of four phases: pre-stage; resource development; execution; and exploitation. The lifecycle is inspired by the Lockheed Martin Cyber Kill Chain [21], where each phase builds on the success of the previous one. Inspired by the MITRE ATT&CK framework [51], each phase is decomposed into *tactics* and *techniques*, where a *tactic* defines the short-term objective of an attack and a *technique* is the activities that support a tactic. Table 2 decomposes each phase into tactics and techniques. The phases and their associated tactics and techniques are elaborated below.

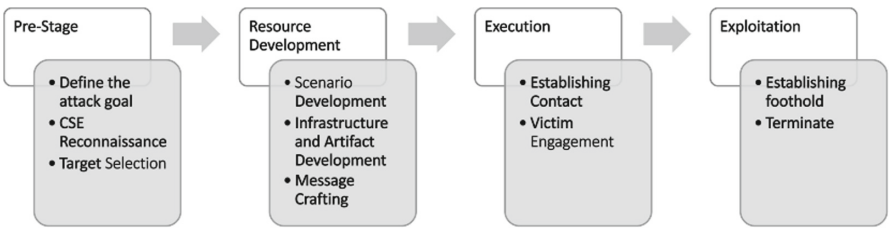

Fig. 1. The four phases and their associated tactics of the CSE Kill Chain framework.

2.3 The Pre-Stage Phase

In this phase, the attacker identifies opportunities to increase their success. This phase consist of three tactics: *Define the Attack Goal*, *CSE Reconnaissance*, and *Target Selection*.

Table 2. CSE kill chain matrix (inspired by the MITRE ATT&CK matrix). Note that ATT&CK techniques (e.g., T1957) may be leveraged to instantiate some CSE techniques (e.g., CSE-T1003) as discussed in the text.

PHASE	Pre-Stage			Resource Development			Execution		Exploitation	
TACTIC	Define the goal (CSE-TA01)	CSE Reconnaissance (CSE-TA02)	Target Selection (CSE-TA03)	Scenario Development (CSE-TA04)	Infrastructure and Artifact Development (CSE-TA05)	Message Crafting (CSE-TA06)	Initial Contact (CSE-TA07)	Victim Engagement (CSE-TA08)	Establish Foothold (CSE-TA09)	Terminate (CSE-TA10)
TECHNIQUES	Steal Proprietary Information	Passive Surveillance (CSE-T1001)	Analyze Vulnerabilities (CSE-T1004)	Contextualize. (CSE-T1101)	Compromise Infrastructure (T1584)	Urgency Cues (CSE-T1106)	Replication through Removable Media (T1091)	Pretexting (CSE-T1203)	Command and Scripting Interpreter (T1059)	Funds Transfer (CSE-T1301)
	Financial Fraud	Dumpster Diving (CSE-T1002)	Identify Available Resources (CSE-T1005)	Impersonation (CSE-T1102)	Develop Capabilities (T1587)	Attention Grabber (CSE-T1107)	Phishing (T1566)	Affection Trust (CSE-T1204)	Exploitation for Client Execution (T1203)	Information Transfer (CSE-1302)
	Steal Personal Information	Open Source Reccon. (CSE-T1003)	Identify Env. Limitations (CSE-T1006)		Obtain Capabilities (T1588)	Persuasion (CSE-T1108)	Drive-By Compromise (T1198)	FITD (CSE-T1205)	Native API (T1106)	
	Acquire Information		Determine Impact (CSE-T1007)		Stage Capabilities (T1608)	Incentives & Motivators (CSE-T1109)	Trusted Relationships (CSE-T1201)	Quid-Pro-Quo (CSE-T1206)	Scheduled Task/Job (T1053)	
	Gain Access to a Network		Select Attack Conduit (CSE-T1008)		Visual Deception (CSE-T1103)	Personalization (CSE-T1110)	Rev. Social Engineering (CSE-T1202)	Scamming Techniques (CSE-T1207)	Software Deployment Tools (T1072)	
			Determine Exploitability (Exposure) (CSE-T1009)		Website or Application Spoofing (CSE-T1104)				System Services (T1569)	
					URL Spoofing (CSE-T1105)				User execution (T1204)	

FITD = Foot-in-the-Door

The "Define the Attack Goal" Tactic (denoted by CSE-TA01). This tactic identifies the desired outcome of a CSE attack to maximize success. Possible CSE attack goals include: *steal proprietary information*, *financial fraud*, *steal personal information*, *acquire information*, and *gain access to a network*, which are self-explaining. Understanding attackers' goals can help design effective defense to prevent, detect, and attribute CSE attacks.

The "CSE Reconnaissance" Tactic (denoted by CSE-TA02). This tactics gathers information to identify vulnerabilities and opportunities to conceal a CSE attack by mimicking legitimate activities, interactions, expected behaviors and norms in the environment and/or exploiting the widely accepted biases and beliefs that are unique to the environment. This tactic can be supported by the following CSE techniques: (i) *Passive Surveillance* (denoted by CSE-T1001), which is the close monitoring of a targeted location to identify activity patterns, entry and exit points, and to assess security [16]; (ii) *Dumpster Diving* (denoted by CSE-T1002), which is the act of searching through the trash for useful information, such as corporate phone books, corporate letterheads, calendar of events, or system manuals [47]; and (iii) *Open-Source Reconnaissance* (denoted by CSE-T1003), which is the process of gathering information available online and in social media [39], and can be instantiated as the following MITRE ATT&CK techniques [51]: "*gather victim identity information*" (denoted by T1589 in [51]), "*gather victim organization information*" (denoted by T1591 in [51]), "*search closed sources*" (denoted by T1597 in [51]), "*search open technical database*" (denoted by T1593 in [51]), "*search open websites/domain*" (denoted by T1596 in [51]), and "*search victim-owned websites*" (denoted by T1594 in [51]).

The "Target Selection" Tactic (denoted by CSE-TA03). This tactic selects a target based on its exploitability and how much the exploitation contributes to achieving the attacker's goal. Factors that affect exploitability are environmental limitations, attacker's available resources, weakness exposure [4], and attack conduit. This tactic can be supported by the following CSE techniques: (i) *Analyze Vulnerabilities* (denoted by CSE-T1004); (ii) *Identify Available Resources* (denoted by CSE-T1005); (iii) *Identify Environmental Limitations* (denoted by CSE-T1006); (iv) *Determine Impact* (denoted by CSE-T1007); (v) *Select Attack Conduit* (denoted by CSE-T1008); and (vi) *Determine Exploitability* (denoted by CSE-T1009).

2.4 The Resource Development Phase

In this phase, the attacker crafts the resources that are necessary to launch a CSE attack. This phase has three tactics: *Scenario Development*, *Infrastructure and Artifact Development*, and *Message Crafting*.

The "Scenario Development" Tactic (denoted by CSE-TA04). This tactic develops a scenario to provide a credible reason for contacting the victim. A scenario consists of two parts: backstory and persona. A backstory sets the stage for the attacker to approach the victim. A persona is a fake identity that aligns with the backstory and adds credibility. A persona can also increase the victim's cooperation. This tactic is supported by the following techniques: *contextualization* (denoted by CSE-T1101), which incorporation of details in the backstory to project the appearance of belonging to a group [35]; and *impersonation* (denoted by CSE-T1102), which creates a persona to increase cooperation by leveraging societal norms and expectations [2,15,16,46].

The "Infrastructure and Artifacts Development" Tactic (denoted by CSE-TA05). This tactic creates the artifacts and infrastructure that make a scenario believable [56]. Artifacts serve as evidence of the credibility of a CSE message and increase a victim's trust perception of the CSE message. For artifact hosting, an attacker can build the infrastructure or use a compromised third-party infrastructure. This tactic is supported by the following CSE techniques: (i) *Compromise Infrastructure* (T1584, inherited from the ATT&CK framework [51]), which entails the process of gaining control of a third-party infrastructure to reuse it in a cyber attack; (ii) *Develop Capabilities* (T1587, inherited from ATT&CK [51]), which involves the building of capabilities by the attacker; (iii) *Obtain Capabilities* (T1588, inherited from ATT&CK [51]), which involves the purchase or stealing of capabilities by the attacker; (iv) *stage capabilities* (T1608, inherited from ATT&CK [51]), which involves the action of preposition capabilities for an attack; (v) *Visual Deception* (denoted by CSE-T1103), which is the manipulation of known artifacts to project credibility, such as high-quality artifacts (e.g., logos, images, videos, or name spoofing [12]) and proper grammar, implicit third-party endorsement (SSL padlocks, use of HTTPS, antifraud statements, corporate protection), and content presentation [52]; (vi) *Website or Application Spoofing* (denoted by CSE-T1104), which is the mimicking of a legitimate website or application; and (vii) *URL Spoofing* (denoted by CSE-T1105), which is the replacement of visually similar characters in a legitimate URL to create a malicious URL (e.g., replacing "w" with "vv" in a URL [12]).

The "Message Crafting" Tactic (denoted by CSE-TA06). This tactic creates a CSE message to persuade a victim, by lowering the victim's risk perception, to perform the action that is desired by the CSE attacker [26]. This tactic is supported by the following techniques, which may be used individually or in combination. (i) *Urgency Cues* (denoted by CSE-T1106), which encourages an automatic response from a victim based a perceived sense of urgency [54]; (ii) *Attention Grabbers* [37] (denoted by CSE-T1107), which is the use of visual and auditory elements to force an individual to switch their attention to a message [30]; (iii) *Persuasion* [10] (denoted by CSE-T1108), which involves the use of persuasion principles in the message to encourage a behavior like clicking a link or providing personal information [35]; (iv) *Incentives and Motivators* (denoted by

CSE-T1109) [46], which are the rewards that encourage a desired victim behavior [15,49]; and (v) *Personalization* (denoted by CSE-T1110), which involves implying a connection with the victim, by expressing personal knowledge of a victim, such as addressing a victim by name, using appropriate pronouns, and aligning a message with the recipient characteristics (e.g., interests, biases and beliefs, and personal information) [20,35,55].

2.5 The Execution Phase

This phase convinces a victim to comply with the attacker's request and has two tactics: *Initial Contact* and *Victim Engagement*.

The "Initial Contact" Tactic (denoted by CSE-TA07). This tactic initiates a contact with a victim. This tactic is supported by the following techniques: (i) *Replication through Removable Media* (T1091, inherited from ATT&CK [51]), which is the use of removable media to copy malware into a system; (ii) *Phishing* (T1566, inherited from ATT&CK [51]), which involves the use of emails to deliver malicious content to gain access to a system; (iii) *Drive-By Compromise* (T1198, inherited from ATT&CK [51]), which is about gaining access to a system by alluring a victim to visit a compromised website; (iv) *Trusted Relationship* (denoted by CSE-T1201), which is adapted from the ATT&CK Trusted Relationship technique (T-1199 [51]) to social engineering attacks and involves leveraging relations or connections to other agents, or exploiting third parties (e.g., social media sites, dating sites, etc.) to increase perceived credibility [1]; and (v) *Reverse Social Engineering* (denoted by CSE T-1202), which triggers the victim to initiate a communication with the attacker.

The "Victim Engagement" Tactic (denoted by CSE-TA08). This tactic establishes continuous communications with the victim to increase compliance. This tactic is often employed when an attack involves a high-risk request (e.g., money wire transfer). This tactic is supported by the following techniques: (i) *Pretexting* (denoted by CSE-T1203), which is the use of pretenses to entice a victim into willingly disclosing information [18]; (ii) *Affection Trust* (denoted by CSE-T1204), which involves establishing an emotional connection through continuous communications to increase a victim's risk tolerance and compliance [31]; (iii) *Foot-in-the-Door* (denoted by CSE-T1205), which manipulates a victim to comply with a large request by making several smaller requests over time [13]; (iv) *Quid-Pro-Quo* (denoted by CSE-T1206), which promotes a victim's compliance by promising a future reward for the victim's help [19]; and (v) *Scamming* [49] (denoted by CSE-T1207), which is the use of deceptive elements to encourage a false belief [5].

2.6 The Exploitation Phase

This phase makes an attack that achieves its goal. It has two tactics: *Establishing a Foothold* and *Terminate*.

The "Establish Foothold" Tactic (denoted by CSE-TA09). This tactic leverages a victim's compliance with the attack's request to penetrate into a network. For example, clicking on a link could trigger the installation of a malicious software, which enables the attacker to gain access to a network or launch a ransomware attack [3]. This tactic can be supported by the following ATT&CK techniques: (i) *Command and Scripting Interpreter* (T1059, inherited from ATT&CK [51]), which is the use of commands, scripts or binaries to compromise a system; (ii) *Exploitation for Client Execution* (T1203, inherited from ATT&CK [51]), which is the exploitation of a vulnerability in the software installed in a system; (iii) *Native API* (T1106, inherited from ATT&CK [51]), which is the use of native Operating System programming interfaces to compromise a systems; (iv) *Scheduled Task/Job* (T1053, inherited from ATT&CK [51]), which is the use of the task scheduling functionality to compromise a system; (v) *Software Deployment Tools* (T1072, inherited from ATT&CK [51]), which is the use of a third-party enterprise applications to conduct lateral movement in a network; (vi) *System Services* (T1569, inherited from ATT&CK [51]), which is the abuse of system services to execute commands and programs; and (vii) *User Execution* (T1204, inherited from ATT&CK [51]), which is execution of an action by a victim to help the attacker gain access into a system.

The "Terminate" Tactic (denoted by CSE-TA10). Once the victim performs the desired action, the attack terminates. This tactic can be supported by the following techniques: (i) *Financial Transfer* (denoted by CSE-T1301), which may be conducted by an employee to transfer funds to a foreign bank account [22] or by an individual to transfer money to a romantic [45] or venture partner [19]; and (ii) *Information Transfer* (denoted by CSE-T1302), which involves the movement of information to an asset that is under the attacker's control.

3 Case Study

This section presents a case study to demonstrate how the CSE Kill Chain can be applied to analyze a real-world CSE attack described in [1]. This attack is a cyberspace variant of the recruiting scam in the physical world described in [50]. The attack involves the four phases of the CSE Kill Chain. The attack uses a social media platform to exploit the victims. To demonstrate the usefulness of the framework, we leverage the findings of the case study to generate an attack pattern that describes the implementation of each technique (Table 3). These two products synthesize details of a CSE attack to assist characterization.

3.1 The Pre-stage Phase of the Attack

The CSE attack in this case study mirrors the Recruitment Scam described in [50], meaning that the attacker just has to recreate the scam elements in the digital space. At this phase, the CSE attacker uses three tactics. (i) Corresponding to the "Define the Attack Goal" tactic (CSE-TA01), the attack attempts

Table 3. Illustration of generating attack patterns from the case study [1], where attack patterns can help identify distinct elements in an attack to define a signature for the attacker.

Tactic	Technique	Implementation
SE Recon. (CSE-TA02)	Open Source Recon.	Online research on potential targets and methods of exploitation
Target Selection (CSE-TA03)	Analyze Vul.	Features that allow a attacker to operate undetected
	Identify Available Resources	Platform recruiting tools, job seeker analysis tools, and job adverting tools
	Identify Env. Limitations	Violations of use of the platform services
	Determine Impact	Attract job seeker with financial means
Scenario Develop. (CSE-TA04)	Contextualiz.	Construction Company hiring project managers
	Impersonation	Hiring Manager Persona
Inf. & Artifact Development (CSE-TA05)	Visual Deception	High quality images and brand logos
	Website/App. Spoofing	Carbon copy website from legitimate construction company
Message Crafting (CSE-TA06)	Persuasion	∘ Authority: Hiring manager demands in-person interview ∘ Social Compliance: Interview would determine if candidate is a good fit for their culture
	Incentives & Motivators	High paying salary
Initial Contact (CSE-TA07)	Rev. SE	Job seekers contact the attacker
	Trusted Relationships	Using LinkedIn, a site the job seekers trust, for job posting
Victim Engagement (CSE-TA08)	Pretexting	Invitation for an in-person interview used for request personal information and money
	FITD	Gradual increase of request: ∘ 1st request: resume ∘ 2nd request: personal information ∘ 3rd request: money
	Scamming Tech.	Distraction: job seeker fails to notice inconsistent, e.g., paying for travel cost for an interview
Terminate (CSE-TA10)	Funds Transfer	Job seeker transfer money to attacker's bank account
	Information Transfer	Job seeker send copy of passport and identification

to steal a victim's personal information and money (i.e., financial fraud). (ii) As part of the "CSE Reconnaissance" tactic (CSE-TA02), the attacker identifies the social media platforms that would better support its goal by using technique CSE-T1003 (Open Source Reconnaissance) and, more specifically, the ATT&CK technique T1596 (Search Open Website/Domains). Since the attack is a digital version of the Recruitment Scam, the attacker would gather information on available employment sites, the users, the cost of corporate membership, and membership benefits. (iii) The "Target Selection" tactic (CSE-TA003) focuses

on selecting an employment platform to wage the attack. The employment platform would (indirectly) recommend potential victims. This can be achieved by using the technique CSE-T1004 (Analyze Vulnerabilities), whereby the attacker identifies platform features that can benefit the attacker in evading detection. Furthermore, by using the technique CSE T-1005 (Identify Available Resources), the attacker evaluates the platform features to recommend victims and increase the visibility of the (fake) job posting. By using the technique CSE T-1006 (Identify Environmental Limitations), the attacker evaluates the employment platforms based on the conditions that violate the user agreement. As described in CSE-T1007 (Determine Impact), the attacker must attract individuals with financial means. The attacker may focus on white-collar professionals, especially project managers.

3.2 The Resource Development Phase of the Attack

From the previous phase, the attacker selects LinkedIn to connect with job seekers. At this phase, the CSE attacker uses three tactics. (i) As part of the "Scenario Development" tactic (CSE-TA04), the attacker uses the technique CSE-T1101 (Contextualization) to create a fictional company and job postings for project managers. The attacker also uses technique CSE-T1102 (Impersonation) to assume the persona of a Hiring Manager. (ii) In support of the "Infrastructure and Artifact Development" tactic (CSE-TA05), the attacker uses technique CSE-T1104 (Website or Application Spoofing) to create a copy of a legitimate company website and modify it to align with the scenario in question. The attacker also uses technique CSE-T1103 (Visual Deception) by displaying legitimate logos and professional graphics on the website to increase its credibility. Using technique T1608 (Stage Capability), the attacker activates the website in a web hosting platform. (iii) Corresponding to the "Message Crafting" tactic (CSE-TA06), the attacker uses techniques CSE-T1109 (Incentives) to offer a high salary range of $105K-$160K. By using CSE-T1108 (Persuasion Techniques) which exploits the principle of authority [10] or social compliance [49], the attacker demands an in-person interview for a candidate to demonstrate that the candidate is a proper fit to the chief project manager of a company. Summarizing the preceding discussion, we draw:

Insight 1. *Understanding how human perception is formed is key to detecting CSE attacks.*

3.3 The Execution Phase of the Attack

At this phase, a CSE attacker uses two tactics. (i) Corresponding to the "Initial Contact" tactic (CSE-TA07), the attacker uses technique CSE-T1202 (Reverse Social Engineering) by posting the job announcement on LinkedIn, which automatically advertises the position to suitable candidates. A victim contacts the attacker by applying for the job and providing their curriculum and personal

information. Using LinkedIn also leverages technique CSE-T1201 (Trusted Relationship). (ii) Corresponding to tactic CSE-TA08 (Victim Engagement), the attacker uses CSE technique CSE-T1203 (Pretexting) to request the victim's personal information (VISA/travel information, passport, identification, etc.) with the pretext of arranging for an in-person job interview. Using technique CSE-T1205 (FITD), the attacker obtains a victim's compliance by gradually increasing the size of the request (i.e., the attacker initially requests a victim's resume, then personal information, and finally, a money transfer). Using technique CSE-T1207 (Scamming Technique), the attacker uses the prospect of a well-paying job as a distraction to prevent a victim from noticing the inconsistencies in their interaction. Examples of inconsistency include: providing personal information without an interview or job offer and covering travel costs upfront for a job interview. Summarizing the preceding discussion, we draw:

Insight 2. *Social media enables a CSE attacker to increase the credibility and trustworthiness of CSE attacks.*

3.4 The Exploitation Phase of the Attack

At this phase, and as described in the "Terminate" tactic (CSE-TA10), the attacker uses techniques CSE-T1302 (Information Transfer) and CSE-T1301 (Funds Transfer) to receive a victim's personal information and money transfer into a designated bank account. Once the funds are received, the attacker terminates communications with the victim.

4 Limitations

The present study has several limitations, which need to be overcome by future studies. First, because it focuses on attacks, the CSE Kill Chain does not accommodate the victim attributes that contribute to the success of CSE attacks. For example, it does not capture victims' temporal psychological attributes (e.g., stress and workload), which can contribute to social engineering victimization by reducing victims' attention and vigilance. Second, the CSE Kill Chain does not account for the environment where the attacker-victim interaction occurs (e.g., workplace, home office), which can affect a victim's risk acceptance. For example, a user may be more willing to click on a link in a CSE message when receiving a message on a network which employs multiple defense layers. Third, more case studies are needed to evaluate the effectiveness of the CSE Kill Chain.

5 Conclusion

We have presented the CSE Kill Chain framework, which describes CSE attacks in four phases: pre-stage, resource development, execution, and exploitation. We established connections between the CSE Kill Chain and the MITRE ATT&CK framework, which does not consider CSE attacks. Therefore, it is hopeful that the

CSE Kill Chain can be adopted to accompany the MITRE ATT&CK framework to accommodate CSE attacks.

We hope this work will inspire many more studies on defending against CSE attacks. There are several outstanding open problems for future research. In addition to addressing the limitations of the present study as described above, we mention the following: How can we apply the CSE Kill Chain to make existing warning models (e.g., the Communication-Human Information Processing model [57]) more effective in communicating security threats to users? How can we apply the CSE Kill Chain to identify an attacker's psychological signature so as to enhance attack attributions? How can we leverage the CSE Kill Chain to characterize the relationship between suspicion and cognitive engagement so as to defend from CSE attacks? How can we transform the CSE Kill Chain framework into mathematical models for quantitative analysis purposes? This would require to defining pertinent cybersecurity metrics [8,9,32,40,66] and conducting quantitative studies to characterize CSE attacks [43,44,59,60].

Acknowledgement. We thank the anonymous reviewers for their comments that helped us in improving the paper. Approved for Public Release; Distribution Unlimited. Public Release Case Number 21–1635.The first author is also affiliated with The MITRE Corporation, which is provided for identification purposes only and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions, or viewpoints expressed by the authors. This work was supported in part by ARO Grant #W911NF-17–1-0566, NSF Grants #2122631 and #2115134, and Colorado State Bill 18–086.

References

1. Allodi, L., Chotza, T., Panina, E., Zannone, N.: The need for new antiphishing measures against spear-phishing attacks. IEEE Secur. Priv. **18**(2), 23–34 (2019)
2. Anderson, R.: Security Engineering: a Guide to Building Dependable Distributed Systems. John Wiley & Sons (2020)
3. Applegate, S.D.: Social engineering: hacking the wetware! Inf. Secur. J. a Glob. Perspect. **18**(1), 40–46 (2009)
4. Barrett, N.: Penetration testing and social engineering-hacking the weakest link. Inf. Secur. Tech. Rep. **4**(8), 56–64 (2003)
5. Buller, D.B., Burgoon, J.K.: Interpersonal deception theory. Commun. Theory **6**(3), 203–242 (1996)
6. Center, V.T.R.A.: 2021 data breach investigation report. Tech. rep, Verizon Threat Research Advisory Center (2021)
7. Chen, S., Chaiken, S.: The heuristic-systematic model in its broader context. In: Dual-process theories in social psychology, pp. 73–96. The Guilford Press (1999)
8. Cho, J., Hurley, P., Xu, S.: Metrics and measurement of trustworthy systems. In: Proceedings IEEE MILCOM (2016)
9. Cho, J., Xu, S., Hurley, P., Mackay, M., Benjamin, T., Beaumont, M.: STRAM: measuring the trustworthiness of computer-based systems. ACM Comput. Surv. **51**(6), 1–47 (2019)

10. Cialdini, R.B., Cialdini, R.B.: Influence: the psychology of persuasion, vol. 55. Collins New York (2007)
11. Deutsch, M.: Trust and suspicion. J. Conflict Resolut. **2**(4), 265–279 (1958)
12. Dhamija, R., Tygar, J.D., Hearst, M.: Why phishing works. In: Proceedings of the SIGCHI conference on Human Factors in computing systems, pp. 581–590 (2006)
13. Freedman, J.L., Fraser, S.C.: Compliance without pressure: the foot-in-the-door technique. J. Pers. Soc. Psychol. **4**(2), 195 (1966)
14. Gavett, B.E., Zhao, R., John, S.E., Bussell, C.A., Roberts, J.R., Yue, C.: Phishing suspiciousness in older and younger adults: the role of executive functioning. PLoS ONE **12**(2), e0171620 (2017)
15. Gragg, D.: A multi-level defense against social engineering. SANS Reading Room **13**, 1–21 (2003)
16. Greenlees, C.: An intruder's tale - [it security]. Engineering & Technology, pp. 55–57 (2009)
17. Han, Y., Lu, W., Xu, S.: Preventive and reactive cyber defense dynamics with ergodic time-dependent parameters is globally attractive. IEEE TNSE **8**(3), 2517–2532 (2021)
18. Hechler Baer, M.: Corporate policing and corporate governance: what can we learn from Hewlett-Packard's pretexting scandal. Univ. Cincinnati Law Rev. **77**, 523 (2008)
19. Herley, C.: Why do Nigerian scammers say they are from Nigeria? In: WEIS (2012)
20. Hirsh, J.B., Kang, S.K., Bodenhausen, G.V.: Personalized persuasion: tailoring persuasive appeals to recipients' personality traits. Psychol. Sci. **23**(6), 578–581 (2012)
21. Hutchins, E.M., Cloppert, M.J., Amin, R.M., et al.: Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. Leading Issues Inf. Warfare Secur. Res. **1**(1), 80 (2011)
22. Junger, M., Wang, V., Schlömer, M.: Fraud against businesses both online and offline: crime scripts, business characteristics, efforts, and benefits. Crime Sci. **9**(1), 1–15 (2020). https://doi.org/10.1186/s40163-020-00119-4
23. Kahneman, D.: Thinking, fast and slow. Macmillan (2011)
24. Kirmani, A., Zhu, R.: Vigilant against manipulation: the effect of regulatory focus on the use of persuasion knowledge. J. Mark. Res. **44**(4), 688–701 (2007)
25. Langenderfer, J., Shimp, T.A.: Consumer vulnerability to scams, swindles, and fraud: a new theory of visceral influences on persuasion. Psychol. Marketin. **18**(7), 763–783 (2001)
26. Lea, S.E., Fischer, P., Evans, K.M.: The Psychology of Scams: Provoking and Committing Errors of Judgement. Tech. rep, Office of Fair Trading (2009)
27. Li, X., Parker, P., Xu, S.: Towards quantifying the (in) security of networked systems. In: 21st IEEE International Conference on Advanced Information Networking and Applications (AINA2007), pp. 420–427 (2007)
28. Li, X., Parker, P., Xu, S.: A stochastic model for quantitative security analyses of networked systems. IEEE TDSC **8**(1), 28–43 (2011)
29. Lin, Z., Lu, W., Xu, S.: Unified preventive and reactive cyber defense dynamics is still globally convergent. IEEE/ACM ToN **27**(3), 1098–1111 (2019)
30. Longtchi, T., Rodriguez, R.M., Al-Shawaf, L., Atyabi, A., Xu, S.: Internet-based social engineering attacks, defenses and psychology: a survey. arXiv preprint arXiv:2203.08302 (2022)
31. McAllister, D.J.: Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. Acad. Manag. J. **38**(1), 24–59 (1995)

32. Mireles, J., Ficke, E., Cho, J., Hurley, P., Xu, S.: Metrics towards measuring cyber agility. IEEE T-IFS **14**(12), 3217–3232 (2019)
33. Mitnick, K.D., Simon, W.L.: The art of deception: controlling the human element of security. John Wiley & Sons (2003)
34. Montañez, R., Atyabi, A., Xu, S.: Social engineering attacks and defenses in the physical world vs. cyberspace: a contrast study. In: Cybersecurity and Cognitive Science, pp. 3–41. Elsevier (2022)
35. Montañez, R., Golob, E., Xu, S.: Human cognition through the lens of social engineering cyberattacks. Front. Psychol. **11**, 1755 (2020)
36. Mouton, F., Malan, M.M., Leenen, L., Venter, H.S.: Social engineering attack framework. In: 2014 Information Security for South Africa, pp. 1–9. IEEE (2014)
37. Nelms, T., Perdisci, R., Antonakakis, M., Ahamad, M.: Towards measuring and mitigating social engineering software download attacks. In: 25th {USENIX} Security Symposium ({USENIX} Security 16), pp. 773–789 (2016)
38. Nohlberg, M., Kowalski, S.: The cycle of deception - a model of social engineering attacks, defenses and victims. In: HAISA (2008)
39. Pastor-Galindo, J., Nespoli, P., Gómez Mármol, F., Martínez Pérez, G.: The not yet exploited goldmine of OSINT: opportunities, open challenges and future trends. IEEE Access **8**, 10282–10304 (2020). https://doi.org/10.1109/ACCESS.2020.2965257
40. Pendleton, M., Garcia-Lebron, R., Cho, J., Xu, S.: A survey on systems security metrics. ACM Comput. Surv. **49**(4), 1–35 (2016)
41. Pennycook, G., Fugelsang, J.A., Koehler, D.J.: What makes us think? a three-stage dual-process model of analytic engagement. Cogn. Psychol. **80**, 34–72 (2015)
42. Petty, R.E., Cacioppo, J.T.: The elaboration likelihood model of persuasion. In: Communication and Persuasion: Central and Peripheral Routes to Attitude Change, pp. 1–24. Springer, New York (1986). https://doi.org/10.1007/978-1-4612-4964-1_1
43. Pritom, M., Schweitzer, K., Bateman, R., Xu, M., Xu, S.: Characterizing the landscape of COVID-19 themed cyberattacks and defenses. In: IEEE ISI'2020 (2020)
44. Pritom, M., Schweitzer, K., Bateman, R., Xu, M., Xu, S.: Data-driven characterization and detection of COVID-19 themed malicious websites. In: IEEE ISI'2020 (2020)
45. Rege, A.: What's love got to do with it? exploring online dating scams and identity fraud. Int. J. Cyber Criminol. **3**(2) (2009)
46. Riegelsberger, J., Sasse, M.A., McCarthy, J.D.: The researcher's dilemma: evaluating trust in computer-mediated communication. Int. J. Hum.-Comput. Stud. **58**(6) (2003)
47. Robinson, S.W.: Corporate espionage 101. https://www.giac.org/paper/gsec/1587/corporate-espionage-101/102941 (2003). Accessed 19 Jun 2021
48. Social Engineer, L.: The attack cycle. https://www.social-engineer.org/framework/attack-vectors/attack-cycle/. Accessed 22 June 2021
49. Stajano, F., Wilson, P.: Understanding Scam Victims: Seven Principles For Systems Security. University of Cambridge, Computer Laboratory, Tech. rep. (2009)
50. Stajano, F., Wilson, P.: Understanding scam victims: seven principles for systems security. Commun. ACM **54**(3), 70–75 (2011)
51. Strom, B.E., Applebaum, A., Miller, D.P., Nickels, K.C., Pennington, A.G., Thomas, C.B.: Mitre att&ck design and philosophy. Tech. rep., MITRE (2020). Accessed 25 June 2021
52. Van Der Heijden, A., Allodi, L.: Cognitive triaging of phishing attacks. In: 28th {USENIX} Security Symposium ({USENIX} Security 19), pp. 1309–1326 (2019)

53. Vishwanath, A., Harrison, B., Ng, Y.J.: Suspicion, cognition, and automaticity model of phishing susceptibility. Commun. Res. **45**(8), 1146–1166 (2018)
54. Vishwanath, A., Herath, T., Chen, R., Wang, J., Rao, H.R.: Why do people get phished? testing individual differences in phishing vulnerability within an integrated, information processing model. Decis. Support Syst. **51**(3), 576–586 (2011)
55. Wang, J., Chen, R., Herath, T., Rao, H.R.: An exploration of the design features of phishing attacks. Inf. Assur. Secur. Priv. Serv. **4**(29), 178–199 (2009)
56. Wathen, C.N., Burkell, J.: Believe it or not: factors influencing credibility on the web. J. Am. Soc. Inform. Sci. Technol. **53**(2), 134–144 (2002)
57. Wogalter, M.S.: Communication-human information processing (c-hip) model. In: Forensic Human Factors and Ergonomics, pp. 33–49. CRC Press (2018)
58. Wright, R.T., Marett, K.: The influence of experiential and dispositional factors in phishing: An empirical investigation of the deceived. J. Manag. Inf. Syst. **27**(1) (2010)
59. Xu, L., Zhan, Z., Xu, S., Ye, K.: An evasion and counter-evasion study in malicious websites detection. In: IEEE CNS, pp. 265–273 (2014)
60. Xu, L., Zhan, Z., Xu, S., Ye, K.: Cross-layer detection of malicious websites. In: Third ACM Conference on Data and Application Security and Privacy (CODASPY'13), pp. 141–152 (2013)
61. Xu, M., Da, G., Xu, S.: Cyber epidemic models with dependences. Internet Math. **11**(1), 62–92 (2015)
62. Xu, M., Xu, S.: An extended stochastic model for quantitative security analysis of networked systems. Internet Math. **8**(3), 288–320 (2012)
63. Xu, S.: Emergent behavior in cybersecurity. In: HotSoS 2014: Proceedings of the 2014 Symposium and Bootcamp on the Science of Security, pp. 1–2 (2014)
64. Xu, S.: Cybersecurity dynamics: a foundation for the science of cybersecurity. In: In: Wang, C., Lu, Z. (eds.) Proactive and Dynamic Network Defense. AIS, vol 74. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-10597-6_1
65. Xu, S.: The cybersecurity dynamics way of thinking and landscape (invited paper). In: ACM Workshop on Moving Target Defense (2020)
66. Xu, S.: SARR: a cybersecurity metrics and quantification framework (keynote). In: Proceedings of the 2021 International Conference on Science of Cyber Security (2021)
67. Xu, S., Lu, W., Xu, L.: Push- and pull-based epidemic spreading in networks: thresholds and deeper insights. ACM Trans. Auton. Adapt. Syst. **7**(3), 1–26 (2012)
68. Xu, S., Lu, W., Xu, L., Zhan, Z.: Adaptive epidemic dynamics in networks: thresholds and control. ACM Trans. Auton. Adapt. Syst. **8**(4), 1–19 (2014)
69. Xu, S., Lu, W., Zhan, Z.: A stochastic model of multivirus dynamics. IEEE Trans. Dependable Secure Comput. **9**(1), 30–45 (2012)
70. Xu, S.: Cybersecurity dynamics. In: Proc. HotSoS'14, pp. 1–2 (2014)
71. Zheng, R., Lu, W., Xu, S.: Preventive and reactive cyber defense dynamics is globally stable. IEEE TNSE **5**(2), 156–170 (2018)