



## **Social Engineering: I-E based Model of Human Weakness to Investigate Attack and Defense**

**Wenjun Fan<sup>1</sup>, Kevin Lwakatare<sup>2</sup> and Rong Rong<sup>3</sup>**

<sup>1</sup>Department of Telematics Engineering, ETSI Telecommunication  
Technical University of Madrid, Madrid, Spain

Email: [efan@dit.upm.es](mailto:efan@dit.upm.es)

<sup>2</sup>Department of Computer Science, TUT Centre for Digital Forensics and Cyber Security  
Tallinn University of Technology, Tallinn, Estonia

Email: [kevinlwaks@gmail.com](mailto:kevinlwaks@gmail.com)

<sup>3</sup>IE Business School, Madrid, Spain

Email: [rrong.mm2013@alumni.ie.edu](mailto:rrong.mm2013@alumni.ie.edu)

### **Abstract**

Social engineering is the attack aimed to manipulate dupe to divulge sensitive information or perform actions to help the adversary bypassing the secure perimeter in front of the information-related resources in order to complete attacking goals. Though there are a number of security tools, such as firewalls and intrusion detection systems, which can be used to protect the machines from being attacked, there is a lack of widely accepted mechanism to prevent dupe from fraud. However, the human element is often the weakest link in an information security chain, particularly, in a human-centered environment. In this paper, we reveal that the human psychological weaknesses result in the main vulnerabilities that can be exploited by social engineering attacks, and also, we capture two essential levels, internal

characteristics of human nature and external circumstance influences, to discover the root cause of the human weaknesses. We unveil that the internal characteristics of human nature can be converted into weaknesses by external circumstance influences. So, we propose the I-E based model of human weakness for social engineering investigation. Based on this model, we analyzed the vulnerabilities exploited by different techniques of social engineering, and also, we conclude several defense approaches to strengthen the human weaknesses. This work can help the security researchers to gain insights into social engineering from a different perspective, and especially, enhance research for the current and future social engineering defense mechanisms.

**Keywords:** Social Engineering, Semantic Attacks, Information Security, Data Privacy, Hacking Techniques, Human Weaknesses

## I. INTRODUCTION

Information security and privacy are very important to personal assets, corporate properties, and even state secrets, which across the globe are faced with various hacking threats. In modern society, people use various digital equipments, such as cell phones, laptops, tablet pads and personal computers, connecting by the Internet to communicate with each other and share information. Hence, modern information security become increasingly interconnected and dependent on IT security [1][2]. The IT security includes not only protecting the organization systems from being attacked but also preventing the system-related human or users from being tricked in order to avoid leaking valuable information.

Due to the intelligence of blackhat community, on one hand, there are many hacking techniques, such as buffer overflow, SQL injection, and cross-site scripting (XSS), can be used to attack the computer systems for accessing the sensitive information [3]. These attacks depend on exploiting the vulnerabilities of the software systems, which can be addressed through timely system update, and supplementing the production system by security tools such as firewall and intrusion detection system (IDS). On the other hand, nevertheless, some hackers pioneered the art of human hacking (also called phreakers in some earlier articles [4]) known as social engineering (SE) attacks to deceive the dupe in order to get the valuable information, such as account names, ID numbers and even passwords, which can be used to

bypass the access control and evade intrusion detection. Hence, unlike the traditional technical vulnerabilities, social engineering attacks focus on human element's psychological vulnerabilities.

The SE attacks are much more difficult for system administrators to defend against. At present, the great percent of information security depends on the human rather than the technical security measures. According to the Verizon 2015 DBIR report [5], humans account for 90\% of security incidents. A recent research report from Ponemon sponsored by Wombat Security Technologies [6] also concluded that the average 10,000-employee company spends \$3.7 million a year dealing with phishing attacks. Symantec 2015 Internet Security Threat Report (ISTR) [7] also shows that five out of six large (2,500 + employees) companies were targeted with spear-phishing attacks during 2014 - a 40 percent increase over the previous year; small and medium-sized businesses also saw an uptick, with attacks increasing 26 percent and 30 percent, respectively. Thus, protection of sensitive information is vitally important to governments and organizations. Although the effectiveness of protecting information is increasing, human element still remains susceptible to manipulation and is the weakest link.

In this paper, therefore, the objective is to make an effort to gain an insight into the social engineering research area. The contributions of this paper can be summarized as follows:

- 1) We capture two essential levels - the internal characteristics of human nature and the external circumstance influences - shaping human psychological states, and propose a novel I-E based model of human weakness.
- 2) We apply the I-E based model of human weakness to analyze the typical social engineering attack techniques in order to get insight into the social engineering attacks.
- 3) According to the model, we also suggest some social engineering defense measures to strengthen the human weaknesses for facilitating information security and privacy.

The remainder of this paper is organized as follows: section 2 reviews the related work; section 3 provides an overview of social engineering to identify the importance of the human weakness; section 4 proposes a novel I-E based model of human weakness; section 5 analyzes the SE attack techniques in terms of the I-E based model; section 6 suggests some SE defense measures; section 7 makes a conclusion.

## II. RELATED WORKS

### A. Social engineering taxonomies

At present, there are a plenty of materials [8][9], introducing social engineering attacks where security researchers can learn the concepts, the attacking techniques, the interesting real cases, etc. Studying the dedicated taxonomies is another way to know well over a study field. To our knowledge, a decade ago, though there were a number of taxonomies of network attacks [10][11], there is little taxonomy that is specially designed for social engineering attacks. Thereafter, some succeeding taxonomies of network attacks began to consider the classifications about social engineering. For example, Simmons et al. [12] proposed a taxonomy called AVOIDIT, which classifies cyber attacks into six categories: attack vector, operational impact, defense, informational impact, and attack target. The attack vector is a vulnerability or path used to compromise a system, such as misconfiguration, buffer overflow, insufficient authentication validation, etc., and one of the subcategory of the attack vector is social engineering. Another taxonomy [13] proposed by Van Heerden et al. consists of twelve classes and each class containing multiple subclasses. The social engineering is one of the subclasses of the class Attack Mechanism. Hence, both of AVOIDIT and Van Heerden's taxonomy simply treated the social engineering as one of the attacking methods but did not unveil the technique details about social engineering attacks.

In recent years, several novel taxonomies focusing on social engineering attacks have been proposed, which can help us to reveal more detail. In 2015, Krombholz et al. proposed a novel taxonomy [14] aimed to classify social engineering attacks. This taxonomy proposed three main categories for dissecting social engineering, and they are: channel, operator and type. The channel means the medium where the SE attacks conduct. The channel consists of e-mail, instant messaging, telephone, VoIP, social network, cloud and website. The Operator indicates the actor who launches the SE attacks, which can be human or software. The type refers to the approach that the SE attacks carry out. The taxonomy includes four approaches: physical, technical, social and socio-technical. Furthermore, the author summarized seven representative SE attack vectors (or scenarios): phishing, dumpster diving, shoulder surfing, reverse social engineering, waterholing, advanced persistent threat, and baiting, nevertheless the author mentioned the fact that the individual SE attack scenario have not been technically exhausted. In order to verify the taxonomy, the author applied it to these representative attack scenarios, which proves the taxonomy works well in analyzing these typical SE attack vectors. Indeed, it is a scenario-driven taxonomy, which draws out the attacking characteristics from

the actual attacking scenario and then categories the characteristics into taxonomy. This taxonomy is designed mainly from the attack point of view, however, there is a lack of the main cause of social engineering attacks.

Another recent novel taxonomy of SE attacks was proposed by Heartfield and Loukas [15]. It adopts three distinct control stages, orchestration, exploitation, and execution, as the basic categories of the taxonomy. For each stage, it poses questions that can help to develop the technical protection mechanisms. The answers to these questions compose the corresponding categories, which consequently establish the whole taxonomy. The Orchestration consists of target type (target of choice or opportunity), attacking mode (manual or automated), and attacking approach (software, hardware without software or hardware with software). The Exploitation includes the deception vector (cosmetic, behavior, or hybrid) and the manipulation interface (user interface or programmatic interface). The Execution is comprised of execution steps (single or multiple) and attack persistence (one-off or continual). Furthermore, this taxonomy depicts several mutual-exclusive sub-categories whose characteristics should be considered for developing the technical protection mechanisms. The taxonomy is not exhaustive and can be expanded based on the three main categories. Also, it was evaluated by applying to 30 different attacks observed in the wild, which is aimed to help in developing the technical protection mechanisms. However, the taxonomy adopted the definition of the three distinct control stages of orchestration, exploitation, and execution suggested by CESG [16], which aimed to describe common cyber attacks instead of social engineering attacks. Hence, the categories of this taxonomy are more related to common cyber attacks than to social engineering attacks, which have some specific concerns that should be taken into account.

In addition, Mouton et al. [17] proposed an ontological model to define the social engineering domain and offered important insights into the various SE attack methods. The set of the categories provided by this social engineering ontology can be considered as a taxonomy as well. This work is also on the side of attacks to analyze the social engineering. There is little consideration and analysis on the weaknesses from human element. Also, some categories concluded by the authors are also very general for common cyber attacks. For example, the values of the class "target" and class "goal" are also valid for the common cybercrimes.

## **B. Social engineering conceptual models**

A taxonomy is also a conceptual model. In this subsection, several dedicated conceptual models of social engineering will be presented.

In the book [4], Mitnick ever proposed a conceptual model, from a perspective of attackers, to describe the social engineering attack cycle (SEAC). But the SEAC is explained too briefly and lacks many details. Based on that, Nohlberg and Kowalski [18] proposed a new model to describe the cycle of deception, which merges attacker, defender and victim. In the model each cycle has five steps. If an attacker is not able to meet the requirement in one step, his attack will fail. Similarly, if one of the steps in the defense cycle can stop the attacker, the attack will fail as well. Otherwise, the attacker will be successful and even is going to be able to do it again. This model can be used to build defenses or to map and describe an attack. Mouton et al. [19] proposed another social engineering attack framework combining their previously proposed SE ontological model [20] and extended Mitnick's social engineering attack cycle through specifying attack steps. It provides full details in every attacking step and can map historical SE attacks to a standardized format.

A system archetype is also a good way to conceptualize a warfare framework of social engineering through describing the relations between the system, the countermeasures and the intruder. Gonzalez et al. [21] use system archetypes as idealized patterns to describe the main modes of social engineering attacks. From both the attack and defense perspectives, the presented system archetypes unveil the two feedback loops, called controlling balancing (B) and reinforcing (R) loops, whose four basic combinations can be used to describe the intended consequence (IC) of the social engineering attack and the unintended consequence (UC) as organizational defense. The UC is the result of the organizational reaction to the SE attacks. However, SE attackers also have solution loop (SOL) to deal with the reaction of the organizational reaction, and always seek the ways to outsmart the single-loop defense lines. So, the paper suggests designing organizational security controls that can provide multi-layer feedback against the combined action of SE attacker's IC and SOL.

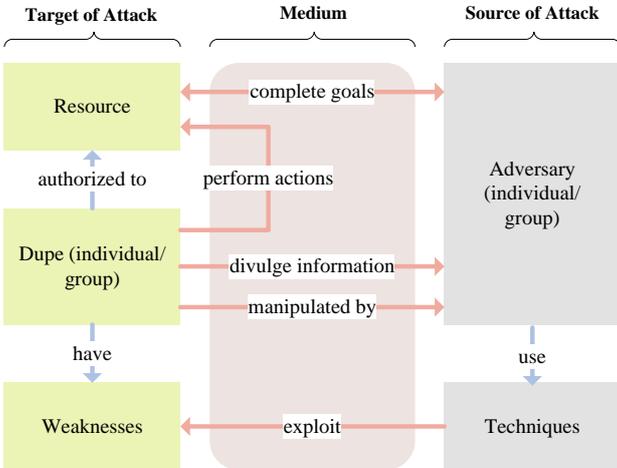
The system archetype approach is good at conceptualizing the SE to a high level of abstraction. However, the power of its analysis still remains questionable in terms of clarifying the techniques in detail. Tetri and Vuorinen [22] proposed a conceptualization of SE consisting different dimensions of SE can be used to exam the techniques of social engineering. Through reviewing the techniques used in actualizing the attacks, the paper extracted three different dimensions of SE techniques: persuasion, fabrication, and data gathering. After that, it proposed an abstract SE framework: intruder-techniques-dupe. The authors emphasized that in real scenario the SE attacker would use multidimensional

approaches to attack organization, which prove, in particular case, the information security policy is the weakest link rather than the human element.

Besides, in particular, Abraham and Smith developed a framework [23] that shows the steps social engineering malware executes to be successful. Indeed, this paper reveals some malware using social engineering channels to be activated, which includes psychological and technical ploys. The psychological techniques include some persuasive tactics as well, such as using the victim’s curiosity, empathy, excitement, fear and greed. The authors claimed that, although it is important for organizations to build comprehensive information security program, the SE malware cannot be mitigated by organizations alone, the shared responsibility of governments, ISPs, end users, and international bodies is needed to combat SE malware.

**III. AN OVERVIEW OF SOCIAL ENGINEERING ATTACK**

A common network intrusion can be divided into five steps: reconnaissance, scan, exploit, gain access and maintain access. For social engineering attacks, Kevin Mitnick's model [4] proposed another five steps: research, developing trust, exploiting trust and utilize information. The main difference is the human element that is the weak link in SE attacks and is exploited by the attackers. Hence, we mainly focus on the difference between the SE attacks and the common cyber attacks, and present an overview of SE attack. The elements and workflow of SE attack are shown in Figure 1.



**Figure 1. Elements and workflow of social engineering attack**

A social engineering attack is aimed to manipulate human into performing actions or divulging confidential information in order to access the target resource to complete goals.

The goal could be financial gain, unauthorized access and service disruption [17]. In order to achieve the goal, an adversary has to trick someone, who is authorized to access the resource, into providing sensitive information or breaking normal security procedures according to his manipulation. The source of attack is an individual or a group of adversaries who often use different social engineering techniques. The adversary known as social engineer is often entrenched in techniques of both IT technology and social psychology [14][15]. Those social engineering techniques often rely on some medium to attack the human weaknesses. The medium could be direct interactions such face-to-face interview or communication over the telephones, or indirect interaction by letters, emails, websites or even unidirectional interaction, e.g. leaving an USB on the ground to wait the dupe to pick it up [14][17]. Also, the dupe could be an individual or a group of victims, since SE threatens not only individuals, but also companies, organizations, and governments. After an successful exploitation on the weakness, the dupe will be manipulated by the adversary. The dupe will divulge sensitive information for accessing the resource or even take place of the adversary to perform malicious behavior so that the adversary can complete the malicious goal. The whole process of doing that is known as social engineering attack.

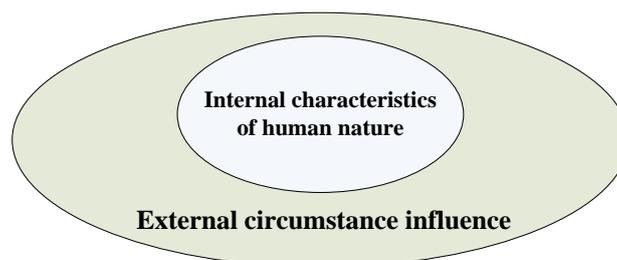
Though security measures have made some fraudulent activities more difficult, the smart and skilled social engineer can also excavate new opportunities to overcome them. Hence, the knowledge of both sides, i.e. attack and defense, is needed to perform research in social engineering. From the workflow of SE attack, we can discover that the human weakness is a link of strategic importance to both the attack and defense sides. In order to frustrate the exploitation, it is necessary to strengthen the human weakness. In contrast, social engineer needs to exploit the weakness to complete the goal. Thus, it is reasonable to consider the human weakness as the foundation for investigating the SE attack and defense. In the next section, we will propose the dissection of the root cause of the human weakness in order to provide our insights into the social engineering.

#### **IV. I-E BASED MODEL OF HUMAN WEAKNESS**

The topic of human weaknesses is a big subject, which includes not only the psychology but also biology and even some principles related to sociology, economics, etc. For example, from the perspective of biology in Richard Dawkins's book – "The Selfish Gene" [24], it explains a lot of altruistic behaviors in the nature, especially the relationship between the

relatives: when an organism in order to fight to protect the safety of relatives, although taking big risks, but because of the relatives have a lot of the same genes in the body, so that is good for gene spreading. This can be used to infer that a man often prefers to trust his family and relatives, but it does not mean those people will never deceive the man. Another case in point is the Dale Carnegie's best-selling book –"How to win friends and influence people" [25], which combines age-old truisms with the emerging field of psychology to give an instruction in handling people, winning friends, bringing people to your way of thinking, being a great leader, and even navigating home life successfully. Carnegie presents the use of other's egotistical tendencies to one's advantage to get success of building trust. Though this handbook is not used for social engineering, the way used to develop trust between human is the same.

Therefore, we can discover that the success of manipulating human is often happened when some characteristics of human nature are triggered by some external influences and converted into weaknesses and they are exploited by the SE techniques. So, based on this discovery, we capture two essential levels (or elements) shaping human psychological states - the internal characteristics of human nature and the external circumstance influences - which cause the human weaknesses. We name it I-E based model of human weakness Figure 2.



**Figure 2. I-E based model of human weakness**

We will describe the internal characteristics of human nature and the external circumstance influence respectively in the next subsections.

#### **A. Internal characteristics of human nature**

There are many types of characteristics of human nature. From the point of view of psychology, it can be roughly divided into two big categories: positive and negative. Also, we cite the seven virtues and sins in the Catholic catechism to make up the cardinal characteristics of human nature in both psychology categories respectively. The positive and negative psychology characteristics are described as follows:

- a) *Positive characteristics refer to the bright-side of personality traits.*

- **Chastity:** discretion of sexual conduct according to one's state in life; the practice of courtly love; cleanliness by cultivated good health and hygiene, and maintained by refraining from intoxicants.
- **Temperance:** constant mindfulness of others and one's surroundings; practicing self-control, restraint, abstinence, moderation and deferred gratification.
- **Charity:** generosity and helpfulness especially toward the needy or suffering; aid given or voluntary giving of help to those in need; benevolent goodwill toward or love of humanity.
- **Diligence:** a zealous and careful nature in one's actions; decisive work ethic, steadfastness in belief, fortitude, and the capability of not giving up.
- **Patience:** building a sense of peaceful stability and harmony rather than conflict, hostility, and antagonism; resolving issues and arguments respectfully, as opposed to resorting to anger and fighting.
- **Kindness:** compassion and friendship for its own sake; empathy and trust without prejudice or resentment; unselfish love and voluntary kindness without bias or spite; having positive outlooks and cheerful demeanor; to inspire kindness in others.
- **Humility:** a spirit of self-examination; a hermeneutic of suspicion toward yourself and charity toward people you disagree with; Modest behavior, selflessness, and the giving of respect; the courage of the heart necessary to undertake tasks that are difficult, tedious or unglamorous, and to graciously accept the sacrifices involved.

*b) Negative characteristics indicate the dark-side of personality traits.*

- **Lust:** it is usually thought of as intense or unbridled sexual desire, which leads to fornication, adultery, rape, bestiality, and other immoral sexual acts.
- **Gluttony:** it is the overindulgence and overconsumption of anything to the point of waste.
- **Greed:** also known as avarice, cupidity or covetousness, is, like lust and gluttony, a sin of desire. However, greed is applied to an artificial, rapacious desire and pursuit of material possessions.

- **Sloth:** it refers to a peculiar jumble of notions, dating from antiquity and including mental, spiritual, pathological, and physical states. It may be defined as absence of interest or habitual disinclination to exertion.
- **Wrath:** it can be defined as uncontrolled feelings of anger, rage, and even hatred, often revealing itself in the wish to seek vengeance. In its purest form, it presents with injury, violence, and hate that may provoke feuds that can go on for centuries.
- **Envy:** like greed and lust, it is characterized by an insatiable desire. It can be described as a sad or resentful covetousness towards the traits or possessions of someone else.
- **Hubris:** the negative version of pride is considered; it describes a personality quality of extreme or foolish pride or dangerous over-confidence; a feeling of deep pleasure or satisfaction derived from one's own achievements, the achievements of one's close associates, or from qualities or possessions that are widely admired.

We maintain that all other personality traits can be attributed to those fourteen characteristics. For example, curiosity is often originated from some desires, such as Lust, Greed and Envy; one's kind-hearted is often based on the Kindness and the politeness relies on the Humility; the incaution arises from the mental Sloth. Also, we believe the fact that any common people have those fourteen characteristics of human nature. Usually, those characteristics are implicit, but under some circumstance influences, they will become increasingly explicit and convert into human weaknesses used by the social engineer. The next subsection we will focus on describing those external circumstance influences.

## **B. External circumstance influences**

As stated, social engineering is a social exercise, and the attackers usually exploit the victims' weaknesses or psychological vulnerabilities, to get the attack success. An external circumstance influence is often an intensive impact from the environment where the dupe locates. The external circumstance influence can stimulate or trigger the psychological characteristics and convert them into human psychological weaknesses, which latterly can be considered as the targets of SE attack. So, in this subsection we describe the related external circumstance influences for social engineering attacks, and we summaries them as follows:

- **Strong affect:** it is an impact using a heightened emotion as a powerful distraction, such feeling a strong sense of surprise, anticipation or even anger, to interfere with the victim's ability to evaluate and think logically when arguments are being presented. This can stimulate one's characteristics such as greedy, lust, gluttony, envy, etc.

- **Overloading:** it refers to the victim has too much information to process but does not have enough time to evaluate it. Hence, this is an influence to reduce the victim's ability to process and scrutinize the argument so that the target is more willing to accept arguments that should have been challenged. This can impact one's sloth, wrath, envy, etc.
- **Reciprocation:** this influence indicates the social interaction rule; if someone give us something or promise us something, we should return the favor. The reasoning follows that people are more willing to comply with a request if the requester has treated them favorably in the past. So, this can trigger one's charity, humility, kindness, etc.
- **Deceptive Relationships:** this influence indicates that the attacker builds the fabricated relationship with the dupe in order to increase the chance the dupe divulges private information to the attacker. The reason is people are more willing to comply with requests from friends or people they like and perform activity under a legitimated and trustworthy relationship. So, one way of doing this is sharing information through discussing a common enemy. Another example is the attacker appears to the target as if they are very much alike, e.g., have the same interests or desire the same things out of life. This can influence almost all the cardinal characteristics.
- **Diffusion of Responsibility and Moral Duty:** this influence means victims are more willing to accept requests or perform actions when they feel it is none of their business or they will not be held solely responsible for their actions. Hence, this can trigger the characteristics such as sloth, charity, humility, kindness, etc.
- **Authority:** it indicates that people are easily to response the requests given by the people with more authority than they have. This can influence one's humility, patience, etc.
- **Integrity and consistency:** this influence refers to people have a tendency to follow commitments and comply the requests that consistent with them, even though commitments may not very wise at the first place. This influence can tigger one's sloth, greed, lust, gluttony, envy, etc.
- **Social validation:** this influence means victims are more easily to comply to requests if they are seen as the socially correct thing to do. This can influence one's humility, charity, kindness, etc.
- **Scarcity:** this influence presents people are more likely to comply with a request that is scarce or decreasing in availability. The reason hiding behind is people subconsciously

approve the fact that objects are valued because of their rarity. So, this can impact one's greed, gluttony, lust, envy, diligence, etc.

These external circumstance influences could be objectively existed around the dupe or subjectively constructed by the social engineer. If the dupe submerges into these scenarios, the probability of being exploited by the adversary will be very high. Additionally, there could be many other external circumstance influences, we can not enumerate all of them here. In the next section, we will analyze the SE attack techniques based on the I-E based human weakness.

## V. ATTACK TECHNIQUES

### A. Descriptions of techniques

The SE attack techniques (also known as attack vectors in the paper [14]) represent the approaches used to exploit the human weaknesses. We classify them into four categories: physical, technical, social and hybrid.

a) *Physical approaches refer to those where the adversary performs some physical activities to gather information.*

- **Dumpster diving:** it represents the action of digging through trash at corporations in search sensitive data.
- **Shoulder surfing:** indicates the observation techniques, such as looking over someone's shoulder, for the sake of getting security information.

b) *Technical approaches refer to the technical actions mainly carried out over the Internet to gather sensitive information.*

- **Phishing:** it is the attempt to acquire sensitive information, such as username, passwords, credit card details, etc. or to make someone to act in a desired way by masquerading as a trustworthy entity in an electronic communication. The general phishing will attack a group of targets of opportunity. However, the spear-phishing focus on attacking some specific individuals or cooperators, thus it requires the adversary to gather information on the intended target in advance. So, the spear-phishing needs more effort but also has a higher success rate than the general phishing attacks.

- **Waterholing:** it refers to the adversary compromise websites which are often browsed or are likely to be of interest to the targets of choice, and infect them with malware, and then waits the target victims getting infected at the waterhole.

- **Baiting:** it is like the real-world Trojan Horse that exploits the victims' greedy and curiosity to access the malware infected temptation, which could be physical media or software and online item. The baiting attack is very similar to the phishing attack, while the baiting is more like a gift or a good left in somewhere can be found by the victims.

c) *Social approaches rely on socio-psychology to manipulate the victims in order to get sensitive information.*

- **Persuasion:** it is aimed to get a victim to comply with an inappropriate request to make them perform some illicit action based on some psychological weaknesses, such as purported authority. One representative persuasion is diversion theft, which is also known as the "Corner Game" or "Round the Corner Game". It is a ``Con" exercised by professional thieves, normally against a transport or courier company. The objective of diversion theft is to persuade the persons responsible for a legitimate delivery that the consignment is requested elsewhere - hence, "round the corner".

- **Pretexting:** it refers to the art of creating and using a fabricated scenario (the pretext) that can be used to increase the chance the dupe divulge information or perform actions that would be unlikely in ordinary circumstances. In comparison with the persuasion, pretexting stands for deceiving the dupe though using some of the techniques, such as impersonation, name-dropping and using false ID, etc.

- **Quid pro quo:** it means "something for something" or "this for that" in Latin, refers to the SE attacks promise a benefit in exchange for information. This benefit usually assumes the form of a service, whereas baiting frequently takes the form of a good.

- **Reverse social engineering:** it is a type of attacks calling back when the victim needs help from someone who claimed can solve the problem, which relies on the established trust between the attacker and the victim, so that the attacker will be allowed to gain the privileged information.

d) *Hybrid approaches refer to the exploiting techniques consisting of multiple different single approaches described above.*

- **Tailgating:** also known as "piggybacking", it refers to that type of attacks involve the adversary who lacks the proper authentication seek entry to a restricted area through following a person who has legitimate access. These restricted areas, i.e. organizations and corporations, are often secured by unattended and electronic access control such as RFID based entrance guard card. For instance, the adversary impersonates a delivery driver and waits outside a building. When an employee appears to gain security's approval and open the door, the adversary will hold the door open, or the employee may hold the door open for the trailing adversary following common courtesy, or the attacker may even ask the employee to hold the door open while the legitimate employee may fail to ask for identification for any of several reasons, such as accept a fabricated assertion that the attacker has forgotten or lost the appropriate identity token.
- **Vishing:** known as phone phishing, it is the act of using the telephone in an attempt to scam the dupe into surrendering private information that will be used for identity theft. The scammer usually pretends to be a legitimate business, and fools the victim into thinking he or she will profit.

## **B. Analysis of techniques**

In this subsection, we analyze the SE attack techniques by applying the proposed I-E based human weaknesses. TABLE I. shows the comparison of those SE attack techniques in terms of I-E based model.

**TABLE I. COMPARISON OF SE ATTACK TECHNIQUES IN TERMS OF I-E BASED HUMAN WEAKNESS**

	Internal Characteristics											External Influences											
	Chastity	Temperance	Charity	Diligence	Patience	Kindness	Humility	Lust	Gluttony	Greed	Sloth	Wrath	Envy	Hubris	Strong affect	Overloading	Reciprocation	Deceptive Relationships	Diffusion of Responsibility	Authority	Integrity and consistency	Social validation	Scarcity
<b>Physical</b>																							
Dumpster diving	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Shoulder surfing	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
<b>Technical</b>																							
Phishing	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Waterholing	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Baiting	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
<b>Social</b>																							
Persuasion	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Pretexting	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Quid pro quo	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Reverse SE	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
<b>Hybrid</b>																							
Tailgating	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Vishing	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o

The comparison presents that different attack techniques have different emphasis on exploiting human weaknesses. First, the physical approaches often rely on the human element's incaution under social validations. For example, the company employees are often simply drop the unused materials into the trashes, which complies the normal social scenarios. However, the undestroyed materials leave the opportunity to the adversary to perform the dumpster diving. Also, the regular working scenario will let down one's guard to other colleagues' shoulder surfing. Second, to the technical approaches, we find the human element's greed is the main cause of the weakness. It is not strange that the dupe accesses some contents through emails or websites, which raises his interest due to the greed characteristics such as lust and gluttony. For example, some phishing emails or websites often use porn information to lure the dupe to divulge personal information. Third, considering the social approaches, we can discover the fact that both of the positive and negative characteristics under certain circumstance influences can become human weakness. The dupe's charity under some social validation can allow the adversary to build trust easily. Also, it is very normal that someone will trust the other who ever helped him previously. General employees in a company often comply the boss's request and even order as well. Fourth, for the hybrid approaches, on one hand, the tailgating uses the dupe's charity, kindness and

humility under some social validation, and for politeness, the dupe perform the action under the adversary's manipulation, which helps the adversary to access the target resource; on the other hand, similar to phishing, the vishing approach is aimed to exploit the dupe's greed to complete the goal.

We discover that the most vulnerable characteristic of human nature is greed that includes lust, gluttony, avarice, etc, which can be easily exploited by SE attack techniques under external circumstance influences. It could be attributed to the theory of "selfish-gene" [24] as Dawkins presented. Also, the sloth is a very important human element's vulnerability. Many social engineers actually exploit the poor dupe's lazy personality to gain complete goals. So, there is an old Chinese saying that "the poor person must have detestable place." In summary, the most weaknesses are exploited by the technical and social approaches, and those negative characteristics are much more vulnerable than the positive ones.

## **VI. DEFENSE MEASURES**

Since threats cannot be eliminated thoroughly, but can be reduced by using security measures, in this section, we present some defense measures to strengthen the human weaknesses to reduce the risk of social engineering attacks. Now that the proposed I-E based model includes two levels, we consider two corresponding categories of defense measures, subjective and objective, to cope with the weaknesses. The next two subsections describe these two categories of defense measures in detail.

### **A. Objective defense measures**

An objective defense measure is aimed to provide some objective conditions to avoid or reduce the impact of the external circumstance influences over the internal characteristics of human nature.

#### ***1) Using standard security policies***

First, well defined and documented security policy is the foundation for defending SE attacks. Using the standard security policy is an effective way to help the organizations to train the employees and control security risk. Organizations often use information security management system (ISMS) to provide a framework for information security risk management. ISMS consists of sets of security policies to define, construct, develop and maintain the computer system (including hardware and software resources) based security

within companies. At present, there are several security standards for IT Governance which leads to information security, and the big five of ISMS standards are: ISO/IEC 27001, BS 7799, COBIT, PCI DSS, ITIL & ISO 2000. These policies dictate the way where computer resources can be used. However, most security standards and policies are defined to address general information security risks, such as malware, hackers and phishers, which threaten organizations. Hence, these general security policies are ineffective, owing to a failure to acknowledge all that is actually required to cope with SE attacks.

For defending SE attacks, the set of policies provided by the security standards should cover not only the computer-based risks but also the human-based risks. ISO/IEC 27032, extended from ISO/IEC 27001, is a completely new international standard published by ISO that covers the baseline security practices for all stakeholders in cyberspace. In particular, it provides technical guidance for addressing SE attacks. Thus, the organization concerning information security can choose ISO/IEC 27032 to implement the cyber security framework to prevent SE attacks. However, this novel security standard still needs to be validated how it will turn out in practice and how widely it will be accepted. Using the security standards makes an ease of security measurement.

## 2) *Updating facilities*

If the corporation has a good financial position, it is suggested to update the office facilities. For example, in order to prevent the dumpster diving, the organization should equip the paper shredder to avoid the sensitive information leaving in the trash. Furthermore, using the fingerprinting based authentication approach replace the password typing based access control to avoid shoulder surfing. Another case in point is to hire the security guards at the entrance of building and some restricted areas as the supplement of the electronic access control. All persons entering the building are required to swipe the ID card. The one no ID entry has to register his information and pass the security check by the security guards.

Besides, it is necessary to apply monitoring facilities to record social activities, which can increase the difficulty of performing social engineering attacks. The phone-call recording is used to capture the social techniques happened in the call center environment, such as the e-bank system where the call center agents will directly communicate with the social engineer. Indeed, the phone-call recording mechanism has been applied widely in many bank call centers. Commonly, the user will be asked at the beginning of the conversation with the call center agent that this talk will be recorded. Therefore, it can be imagined that even the social engineer can disguise perfectly bypass the detection layer, his malicious behavior has been

captured. These data can be used to track the social engineer and even applied as the evidence of crime. Also, the digital surveillance is aimed to capture the physical malicious behavior, such as dumpster diving, shoulder surfing, and even tailgating. The surveillance camera is the widely used device to facilitate the monitoring task. It can be seen that many public places, e.g. super markets, hospitals, banks, etc., using digital surveillance. Definitely, many enterprises also apply the surveillance camera as a security approach. Note that the camera should be equipped not only at the entrance but also should be fixed inside of the enterprises to monitor the potential insider social engineering behavior. The captured data by the camera can be used to track the social engineer and even applied as the evidence of crime as well.

### 3) *Detecting malicious data*

As stated, the two effective ways to exploit human weaknesses are technical and social techniques. For example, the phishing is often based on the unidirectional communication media and dialog-based attacks are often based on the bidirectional communication media. In order to detect the attack pattern spreading through the digital media, such as email, instant messaging, website, etc., it is probable to use automated security program. For example, Bhakta and Harris [26] presented a novel approach based on a pre-defined Topic Blacklist (TBL) to detecting SE attacks by verifying whether the discussion topics of each line of text generated by the potential attacker matches the topic in the TBL. The topic blacklist (TBL) is proposed to check if the sender requests sensitive information or not. The TBL is a list of statement topics, which describe a sensitive operation associating to a sensitive data. So, if the request message hits the TBL, the system will make an alert to remind the dupe to raise vigilance.

## **B. Subjective defense measures**

A subjective defense measure is used to improve human element's subjective will to overcome the impact of the external circumstance influences over the internal characteristics of human nature.

### 1) *Training human awareness*

Once the foundation of a security policy has been established and approved, all employees should be trained in security awareness. Though the organizations apply appropriate security standards, they still need to train the employees' awareness to defend the SE attacks.

This task can be done by defining the awareness needs of various audience groups within the organization (executives, line managers, users, etc.); determining the most effective

awareness methods for each audience group (i.e., briefings, messages, courses); and developing and disseminating awareness materials (presentations, posters, mailings, etc.) regarding the need for adherence to the policy. The awareness function also includes efforts to integrate up-to-date policy compliance and enforcement feedback as well as current threat information to make awareness information as topical and realistic as possible. For example, Mataracioglu et al. [27] proposed a qualitative method called security lifecycle model against SE attacks (SLM-SEA). Although this approach still mainly focuses on enhancing the individuals' awareness to prevent social engineering, it proposed a comprehensive model consisting of user training, testing, measuring, and result feedback.

However, the conventional human being involved awareness training methods, such as educational courses, routine remind, interview, awareness quiz and survey, etc., are labor intensive, repeated and even perhaps tedious. At present, there are some automated tools can be used to train and promote user awareness by simulating real world SE attacks. For example, the King Phisher (<https://n0where.net/phishing-campaign-toolkit-king-phisher/>) is an open-source tool for automatically training the users' awareness to prevent phishing attacks. It can be used to run campaigns ranging from simple awareness training to more complicated scenarios in which user aware content is served for harvesting credentials.

Nevertheless, the previous mentioned methods are passive solutions that enforce the employees to be aware of the sensitive information protected by the security policy. Indeed, awareness training does not simply require the employees to keep secret of the sensitive information, but desire them to know how to identify confidential information and understand their responsibility to protect it. Thus, a positive method is to combine the employee's profit, which could be the bonus, reward or merit pay, with the sensitive information security. Thereafter, all employees will actively improve their awareness because the information security has associated to their own financial benefit.

## 2) *Detecting human emotion*

Changes in emotional state will have an influence over the individual's cognitive functioning. Hence, the employee's emotional state can affect his awareness of the sensitive information. For example, people often perform abnormal activity under some extreme emotions, such as wrath. However, it is not an easy task to determine one's emotional state, and it is even an impossible task for an individual to adjust his or her own emotional state, because individuals have their own perception of emotional state and even some individuals are unable to perform this kind of task in a rational way when their emotions are irrationally challenged. Besides,

some basic concepts of emotional state should be taken into account. First, one's emotional state is something that can stay constant for a long time unless the individual experience great discomforting incident, such as economic crisis, health issue, loved ones die, etc., which have intense effect on the cognitive function. Second, an individual's emotional state can be impacted in a short time when the individual is under attack by the SE attacker. Mathews [28] presented that the experiencing severe stress will have an influence over the individual's cognitive function. Hence, it is desired to use some emotion detection model for automatically performing this task.

Bezuidenhout et al. [29] proposed a SE attack detection model (SEADM), which can be used by the workers to detect SE attacks from the requesters in a call centre environment. The authors claimed that the social engineers often use psychological weaknesses to influence the victim's emotional state and cognitive abilities for getting objective information. In order to enhance the individual's awareness to the social engineering requests, the paper proposed an automated self-evaluation electronic questionnaire. If the individual is detected too emotional, the call or the email request will be elevated to another individual. However, this strategy could initiate the work responsibility shift and even promote further frustration with all individuals involved. The detection of one's emotional state of the first SEADM is subjective, and it is impossible to make instantaneous decision whilst working under pressure. Thus, Mouton et al. [30] improved the SEADM by proposing and incorporating a cognitive functioning psychological measure to determine the emotional state and decision-making ability. Nevertheless, the two previous papers related to SEADM only focus on the call centre environment. Mouton et al. [31] therefore proposed a revised version of SE attack detection model, namely SEADMv2, through extending the model to be able to cope with much more different SE scenarios.

## **VII. CONCLUSIONS**

Social engineering attack is an open and big challenge in the area of cyber crime at the modern human-centered environment. In this paper, we revisited the overview of social engineering attack and identified the root problem, i.e. the human weakness. We captured two essential levels - internal characteristics of human nature and external circumstance influences - that shape the human weakness for social engineering. Therefore, we proposed a novel I-E based model of human weakness and defined these two levels's terminology. We classified

the characteristics into two categories - positive and negative, and cited the seven virtues and sins in the Catholic catechism to make up the cardinal characteristics. Also, we presented nine common circumstance influences. Using this new I-E based model we analyzed a number of typical social engineering attack techniques. We discovered that the human negative characteristics, such as greed and sloth, are much more vulnerable than those positive ones, which can be exploited by SE techniques. Finally, we presented a number of defense measures to strengthen the human weaknesses. These defense measures are categorized into objective approaches and subjective approaches in terms of the I-E based model. In summary, this work provided another perspective to investigate social engineering, and we hope it can help the related security researchers to get insights into the social engineering and enhance the future research.

## REFERENCES

- [1] Hossein Bidgoli. Handbook of Information Security, Information Warfare, Social, Legal, and International Issues and Security Foundations (Handbook of Information Security). John Wiley & Sons, Inc., New York, NY, USA, 2006.
- [2] Ji-Xuan Feng and Janet Hughes. Analyzing privacy and security issues in the information age - an ethical perspective. WSEAS Trans. Info. Sci. and App., 6(1):126–135, January 2009.
- [3] RC Joshi and Anjali Sardana. Honeypots: a new paradigm to information security. CRC Press, 2011.
- [4] Kevin D Mitnick and William L Simon. The art of deception: Controlling the human element of security. John Wiley & Sons, 2011.
- [5] Verizon RISK Team. 2015 data breach investigations report. 2015.
- [6] Ponemon Institute. The cost of phishing and value of employee training, Aug 2015.
- [7] Symantec Enterprise. Internet security threat report 2015, 2015.
- [8] Christopher Hadnagy. Social engineering: The art of human hacking. John Wiley & Sons, 2010.
- [9] Ian Mann. Hacking the human: social engineering techniques and security countermeasures. Gower Publishing, Ltd., 2012.
- [10] A. Avizienis, J. C. Laprie, B. Randell, and C. Landwehr. Basic concepts and taxonomy of dependable and secure computing. IEEE Transactions on Dependable and Secure Computing, 1(1):11–33, Jan 2004.

- [11] Simon Hansman and Ray Hunt. A taxonomy of network and computer attacks. *Computers & Security*, 24(1):31 – 43, 2005.
- [12] Chris Simmons, Charles Ellis, Sajjan Shiva, Dipankar Dasgupta, and Qishi Wu. Avoidit: a cyber attack taxonomy. Technical Report CS-09-003, University of Memphis, Aug 2009.
- [13] RP Van Heerden, Barry Irwin, ID Burke, and L Leenen. A computer network attack taxonomy and ontology. *International Journal of Cyber Warfare and Terrorism (IJCWT)*, 2(3):12–25, 2012.
- [14] Katharina Krombholz, Heidelinde Hobel, Markus Huber, and Edgar Weippl. Advanced social engineering attacks. *Journal of Information Security and Applications*, 22(C):113–122, June 2015.
- [15] Ryan Heartfield and George Loukas. A taxonomy of attacks and a survey of defence mechanisms for semantic social engineering attacks. *ACM Comput. Surv.*, 48(3):37:1–37:39, December 2015.
- [16] CERT-UK. Common cyber attacks: Reducing the impact, 2015.
- [17] Francois Mouton, Louise Leenen, and H.S. Venter. Social engineering attack examples, templates and scenarios. *Computers & Security*, 59:186 – 209, 2016.
- [18] Marcus Nohlberg and Stewart Kowalski. The cycle of deception: a model of social engineering attacks, defenses and victims. In *Second International Symposium on Human Aspects of Information Security and Assurance (HAISA 2008)*, pages 1–11, Plymouth, UK, July 8-9 2008.
- [19] F. Mouton, M. M. Malan, L. Leenen, and H. S. Venter. Social engineering attack framework. In *2014 Information Security for South Africa*, pages 1–9, Aug 2014.
- [20] Francois Mouton, Louise Leenen, Mercia M. Malan, and H. S. Venter. Towards an ontological model defining the social engineering domain. In *11th IFIP TC 9 International Conference on Human Choice and Computers (HCC11 2014)*, pages 266–279, Turku, Finland, July 31 – August 1, 2014.
- [21] Jose J. Gonzalez, Jose M. Sarriegi, and Alazne Gurrutxaga. A framework for conceptualizing social engineering attacks, pages 79–90. Samos, Greece, August 31 - September 1, 2006.
- [22] Pekka Tetri and Jukka Vuorinen. Dissecting social engineering. *Behaviour & Information Technology*, 32(10):1014–1023, 2013.
- [23] Sherly Abraham and InduShobha Chengalur-Smith. An overview of social engineering malware: trends, tactics, and implications. *Technology in Society*, 32(3):183 – 196, 2010.

- [24] Richard Dawkins. *The selfish gene*. 1976.
- [25] Dale Carnegie. *How to win friends and influence people*. Simon and Schuster, 2010.
- [26] R. Bhakta and I. G. Harris. Semantic analysis of dialogs to detect social engineering attacks. In *Semantic Computing (ICSC), 2015 IEEE International Conference on*, pages 424–427, Feb 2015.
- [27] Tolga Mataracioglu, SevgiÖzkan, and Ray Hackney. Towards a security lifecycle model against social engineering attacks: SLM-SEA. *CoRR*, abs/1507.02458, 2015.
- [28] Andrew Mathews. Why worry? the cognitive function of anxiety. *Behaviour Research and Therapy*, 28(6):455 – 468, 1990.
- [29] M. Bezuidenhout, F. Mouton, and H. S. Venter. Social engineering attack detection model: SEADM. In *2010 Information Security for South Africa*, pages 1–8, Aug 2010.
- [30] Francois Mouton, Mercia M Malan, and Hein S Venter. Development of cognitive functioning psychological measures for the seadm. In *HAISA*, pages 40–51, 2012.
- [31] F. Mouton, L. Leenen, and H. S. Venter. Social engineering attack detection model: Seadm2. In *2015 International Conference on Cyberworlds (CW)*, pages 216–223, Oct. 2015.