Social Engineering

Yavor Papazov
Business Park Sofia
Building 11-B, Floor 1
Mladost 4
1766 Sofia
BULGARIA

yavor@esicenter.bg

ABSTRACT

This paper presents an attempted survey of the current state of Social Engineering – including the social context of the phenomenon, a brief history of notable social engineering attacks and their impact, a structured overview of social engineering attacks and common methods, a discussion of various defense tactics and, finally, discusses some open challenges in the topic.

1.0 INTRODUCTION

With the ever-increasing importance of and dependence on IT systems in our daily life, from smart home devices to industrial control systems (ICS) and e-government, the security of these systems is also rising in priority. Proof of these trends is omnipresent – from the increase in the sophistication of attackers' methods and tools, through to the introduction of cybersecurity as an important topic in the impending US presidential election, widely considered the most important and influential in the democratic world [1]. While awareness for cybersecurity issues has undeniably raised during recent years, efforts have often been focused on improving the technology side of the equation, while ignoring or, at least, not paying sufficient attention to the 'human factor' in cybersecurity [2]. While many reasons for this can be cited, chief among them appears to be the understanding that since IT is a technological domain, so must IT security (and cybersecurity by implication) be a technological problem. No other phenomenon illustrates the falsehood of such implications better than social engineering. Indeed, any attempt at dissecting cybersecurity issues, which ignores that humans are central to IT systems as their users, developers, administrators and maintainers, is bound to produce deeply flawed results [49].

This paper aims to present a brief overview of social engineering as an aspect of information security. While we will attempt to cover some of its psychological aspects, we will lean towards the technical aspect of the problem. To this end we examine several important topics in this domain.

Section 2 (Section 1 being this introduction) discusses the scope of the topic and several prominent definitions of the term 'Social Engineering', focusing on the subtle differences between these definitions and attempts to present a sufficiently general view of the meaning of the term that covers its use in practice.

Section 3 is an attempt at summarizing important (and highly impactful) historical applications of social engineering, including several that have no relevance to IT and information security, but have nevertheless left a considerable mark in history.

Section 4 presents a structured overview of social engineering attacks, including proposed classifications, frameworks and common methods and tactics, used in the practice of social engineering. These are key in understanding the phenomenon and how it could have such an impressive impact and remain efficient despite numerous awareness and education campaigns at different levels. A significant amount of attention is dedicated to phishing as an attack vector, as it is used in an overwhelming majority of the successful large

scale attacks. At the same time, we shall attempt to balance the exposition by discussing several other important methods and tactics, like framing, eliciting and pretexting.

Section 5 examines different approaches to defending against social engineering. It is mostly focused on the technical aspect of the problem, as deeply psychological topics are out of scope for this paper.

Finally, in Section 6 we attempt to cast light on several open challenges in the domain and the current approaches to their solution.

2.0 DEFINITION OF SOCIAL ENGINEERING

The term "Social engineering" is used at least two different contexts – information security and political science. While we will present many definitions of social engineering in the context of information security, for now we can informally state that it describes a phenomenon where people are influenced into taking a particular action, which may (and often is) be against their own best interest. In the domain of political science, the term has somewhat similar meaning, but on a larger scale – influencing large amounts ("masses") of people. We will obviously limit our discussion to the information security domain.

The most popular definition of social engineering is probably the one, provided by Harl at the Access All Areas III conference in 1997 – "Basically, social engineering is the art and science of getting people to comply to your wishes. It is not a way of mind control, it will not allow you to get people to perform tasks wildly outside of their normal behavior and it is far from foolproof" [5]. As evidence of its popularity we can point to the numerous times it is referenced, even though the original transcript of his talk is long gone. Another definition of the term belongs to Christopher Hadnagy and can be found in the very first pages of "Social Engineering: The Art of Human Hacking" [7] – "Social engineering is the act of manipulating a person to take an action that may or may not be in the target's best interest". The SANS Institute defines social engineering as "A euphemism for non-technical or low-technology means – such as lies, impersonation, tricks, bribes, blackmail, and threats – used to attack information systems" [6].

While these definitions are not generally in blatant disagreement, some important differences can be spotted: for example, the SANS Institute definition immediately narrows social engineering to negative terms and ties it to the attacker. One could argue that such a narrow interpretation is to be expected, considering that information security practitioners have had to deal mostly with the negative impact of social engineering, however Hadnagy offers some further insight into the term by noting that social engineering techniques are often used in everyday life by people with generally good intentions – for example doctors who influence patients' behavior with the ultimate goal of improving their health. In this light, he suggests that it is important to discern between the tool – social engineering – and the intention – which may often be malicious. We will accept this wider interpretation, but will nevertheless stick to addressing social engineering with 'malicious' intent, as alternative uses are hardly of any interest to information security.

An important takeout from these definitions is the complete lack of any mention of computers or even other computational devices. Indeed, if we analyze the tools of social engineering, we will find that they do not rely in any way on computer information systems and are meaningful in the context of an organization that, for example, manages its information entirely in paper form. On the other hand, if we were to specifically *exclude* IT systems from social engineering, we would most certainly miss key elements, like phishing attacks.

Another critical observation is that the *art* aspect of social engineering is mentioned as least as often as the *scientific* aspect. Indeed, outside of the academical settings, most books and resources stress that social engineering is also an art. While the meaning, conveyed by these definitions does fit in the wider definition of art, what most authors imply is that social engineering might be a talent for certain people, rather than a carefully ingrained skill. Considering the manipulative, subconscious and deceptive traits of this

phenomenon, it is certainly a valid viewpoint. Within this paper, however, we will focus more on the structured approach of examining social engineering in a scientific context.

These ambiguities have led some authors to adopt a narrower term that specifically focuses on social engineering in IT – unsurprisingly called Information Technology Social Engineering, or ITSE (suggested pronunciation - "itsy") [8]. Throughout this paper, we shall use the wider term 'social engineering', while remaining focused *mostly* on social engineering in the context of IT.

A more complete overview of different authors' definitions of social engineering can be found in Ref. [8], along with some further commentary on the main differences and incompatibilities between those definitions.

Without attempting to introduce yet another definition of the term, we will note that social engineering is a set of mostly psychological phenomena, applied predominantly in the context of computer and information security. As such, it stands in stark contrast to the mostly technical nature of the domain and has often been misunderstood or neglected by technical specialists. As is keenly observed in Evans' "Information technology social engineering: an academic definition and study of social engineering – analyzing the human firewall" ([8]), social engineering is "always psychological and sometimes technical" and "the psychological aspect of social engineering is what makes the attack, not the technical". Indeed, one of the challenges in dealing with social engineering is precisely the non-technical root of the threat, inhibiting purely technical solutions. However, as we will see later, some technical approaches to social engineering defences do exist.

Due to the great importance and popularity of 'phishing' attacks, we will provide a definition of the term, while discussing it more in depth in Section 3 – 'phishing' refers to a form of Internet fraud, in which the attacker extracts sensitive information from the victim, by imitating communication with a trusted (by the victim) party [30]. The term is a homophone for 'fishing' – the implication being that the victim is the fish that took the bait. Typical phishing attack channels include unauthentic emails, messages in social networks (Facebook, Twitter, etc.) or instant messaging (IM) platforms, and entire counterfeit websites, mimicking the original visually, but usually having a (slightly) different domain name. Ironically enough, 'phishing' is intentionally misspelled – it is easy to mistake it for 'fishing', drawing a comparison to the relative easiness of mistaking a slightly wrong domain for the one the victim trusts and believes is currently accessing (a famous example, used in the phishing campaign against Associated Press, preceding the hijacking of their Twitter account, is hxxp://www.washinqtonpost.com – notice the q instead of g).

3.0 BRIEF HISTORY OF SOCIAL ENGINEERING

Unlike most other threats in modern IT security, social engineering has only tangential connection to technology – rather it attacks the unwitting humans, tasked with operating, maintaining, overseeing and sometimes even protecting those IT systems. As such, it has a much longer history than its current use in the domain of computer and IT security. In this section, we will examine some important milestones and events in the history of social engineering. Our motivation for doing so is mostly to demonstrate the very real threat, posed by unethical uses of social engineering and to prove the potential of its techniques. We have divided this overview in two sections – important uses of social engineering predating IT (or at least not in the context of IT) and "modern" events, concerning the current use of the word in the security domain. The list is by no means exhaustive.

3.1 Predating IT

3.1.1 The Trojan Horse

One of the first documented (albeit probably mythical) occurrences of social engineering in human history is the tale of the Trojan horse. The main source for this tale is Virgil's Aeneid, an epic poem, describing the

events around the final stages of the Greeks' siege of the city of Troy. The story goes that after spending ten years at the gates of the city, the Greeks built a large wooden horse and left it as a 'present' to the city of Troy, while sailing away. In (the tale's) reality, selected soldiers were hiding inside the horse and, after the Trojans brought the device inside the city, opened the gates, allowing the returned Greek army a quick and decisive victory [3].

The importance of this ancient tale to modern information security should not be understated, as evidenced by the term's new-found meaning – a malicious computer program, that tricks the user into allowing it to run, thereby, if not defeating, at least bypassing the machine's defences. Indeed, utilizing social engineering techniques is extremely common in the malware world.

Despite its probably mythical nature, the tale of the Trojan horse is noteworthy in several aspects – for one, it demonstrates what in hindsight appears to be extremely poor judgment on the part of the victim. This is a common trait of successful social engineering attacks – in the heat of the moment their victims feel compelled to act in compliance with the attacker, while a calm and rational examination of the situation would identify such behavior as foolish and almost unthinkable.

3.1.2 Victor Lustig

Victor Lustig was an Austro-Hungarian-born con artist, who managed to "sell" the Eiffel Tower twice for scrap, successfully persuaded Al Capone to gift him $5000 as a reward for his honesty, of all things, and has been attributed the "Ten Commandments of Con Men", which include "Be a patient listener", "Never look bored" and "Never boast. Just let your importance be quietly obvious" [11]. This is probably the first systematic "work" on social engineering.

3.1.3 ABN AMRO Bank theft

In 2007, an unidentified man, using a lost Argentinian passport managed to gain the trust of employees of the ABN AMRO Bank in Antwerp, Belgium over the course of a year. His end goal, which he ultimately achieved, was stealing a trove of diamonds, stored in the bank's safety deposit boxes, valued at about $28 million. What is particularly impressive about this incident is the utter lack of any tools, other than social engineering techniques. Quoting the manager of the Diamond High Council in Antwerp, Mr. Claes – "He used no violence. He used one weapon – and that is his charm – to gain confidence. He bought chocolates for the personnel, he was a nice guy, he charmed them, got the original of keys to make copies and got information on where the diamonds were" [10].

3.2 In the context of IT

3.2.1 Kevin Mitnick

Widely credited as the most notorious hacker, whose exploits first helped raise awareness of the power of social engineering tactics, Kevin Mitnick's hacking career started at the age of 13, when he managed to obtain, through persuasion and 'dumpster diving', documents, allowing him to ride freely the LA bus system. Among his other exploits are gaining unauthorized access to Digital Equipment Corporation (DEC) networks and Pacific Bell. During his widely publicized arrest and subsequent trial, he was painted in almost mythical proportions – it was widely circulated that he could "start a nuclear war by whistling into a pay phone" [14]. He has stated on numerous occasions that social engineering is his tactic of choice [12], [13].

3.2.2 RSA SecurID breach

On March 17, 2011, RSA Security, Inc. (at the time property of the EMC Corporation) announced that it is the victim of an attack by an Advanced Persistent Threat (APT), resulting in a "compromise and disclosure of information" [15], related to its two-factor authentication solution, the RSA SecurID. As part of the

assessment of the impact, it was stated that "this information could potentially be used to reduce the effectiveness of a current two-factor authentication implementation as part of a broader attack". The ensuing investigation uncovered that social engineering was used to gain unauthorized entry to RSA employees' systems, in particular a spear phishing email was sent to two groups of users, none of them considered high profile. One user was tricked into opening the email (from the Junk mail folder), which had a subject "2011 Recruitment Plan" and contained an attachment – "2011 Recruitment Plan.xls". Upon downloading and opening the file, it exploited a zero-day vulnerability in Adobe Flash and proceeded to install a backdoor on the PC [16], [17]. In the aftermath of the attack, RSA offered to replace SecurID tokens. While a thwarted attack on Lockheed Martin, leveraging the extracted information, ultimately no reports have been confirmed of further successful attacks. The attack's total cost to RSA was $66.3 million.

3.2.3 Target and eBay breaches

On December 19, 2013, Target Corporation announced that a massive data breach has taken place, impacting about 40 million credit and debit cards of their clients [21]. While verifiable official information about the aftermath of the incident is scarce, the security journalist Brian Krebs reported that the suspected attack vector was email phishing that targeted a Target supplier – Fazio Mechanical [20] – information that was partially confirmed by Fazio, who announced that they are the "the victim of a sophisticated cyber-attack operation". The estimated number of affected customers was later increased to 70 million, making it one of the largest data breaches in history at the time [22].

A somewhat similar incident followed with another large retailer – the e-commerce platform eBay [23]. On May 21, 2014, eBay Inc. published a news article on its corporate website, admitting it has been breached, that attackers have stolen a database, containing customers' "name, encrypted password, email address, physical address, phone number and date of birth", and urging its customers to change their passwords. The breach, however, did not affect financial data, which was stored separately, in an encrypted format. The article also pointed out that a "small number of employee log-in credentials" have been compromised, causing many security professionals and companies to speculate that phishing was involved as the initial attack vector [18], [19].

3.2.4 'Carbanak' Group

The 'Carbanak' cybercrime group is responsible for targeting banks in multiple countries. Initial reports of more than $300 million stolen were quickly revised to the staggering $1 billion. This would make the operation one of the largest bank thefts in human history. Although reports on the APT are somewhat conflicting [24], [26], particularly about the APT's targets, all sources agree that spear phishing emails are used as the initial attack vector [25].

3.2.5 "CEO frauds"

This is rather a general phishing scheme, rather than a particular attack, though noteworthy cases are plentiful – for example, Matell, Inc. lost $3 million in 2015 to that type of scam; Ubiquiti parted ways with the impressive $46.7 million (2015) and The Scoular Co. lost $17.2 million [27]. The fraud is exceedingly simple – with a well-crafted spear phishing email from the company CEO (hence the name), that requests that an employee (with the appropriate privileges) perform a wire transfer to an account, specified by his superior. At the time when it becomes clear that the email was not authentic, the wired sum has long been withdrawn by the attackers.

4.0 SOCIAL ENGINEERING ATTACKS

The following section represents an attempt at an organized, but not exhaustive, enumeration of common tools, tactics and methods, related to social engineering. It should be noted that an overwhelming majority of

the available resources on the topic are not academical in nature. Therefore, while our overview attempts to present a purely academic discussion, some degree of informality may be inevitable.

4.1 The 'Attack Cycle' (Mitnick)

One of the first systematic description of the process of exploiting social engineering as an attack vector has been given by Kevin Mitnick in Ref. [12]. The process he proposes has four phases:

1) Research – refers to the process of collecting as much information as possible about the target. This information is utilized subsequent phases and is of critical importance for the quality of the final result in cases of targeted attacks.
2) Developing Rapport and Trust – this phase consists of using various techniques (discussed later in this section) for ensuring the victim trusts the attacker. Information, gathered in the previous phase, is often used for that purpose.
3) Exploiting Trust – the actual 'exploitation' phase of a social engineering attack. At this phase a measurable gain in information or privileges is achieved by the attacker.
4) Utilize information – the final phase of the cycle refers to 'cashing in' on the gains from the previous phase. As hinted by the term 'Attack cycle', this phase can transition again into another research phase, thereby completing the cyclical nature of the process. This transition is the social engineering equivalent of the 'pivoting' process, described in many penetration testing methodologies.

Published in 2002, this model has not aged particularly well, as evidenced, for example, by its failure to fully model attacks like email (non-spear) phishing, where information gathering is optional and may not be performed at all. This has led some researchers to suggest an extension to this model that better reflects the diverse nature of social engineering attacks [28], which is discussed immediately below.

4.2 The "Social Engineering Attack Framework" and attack classification (Mouton et al.)

One of the most complete models of a social engineering attack can be found in the article "Social Engineering Attack Framework" by Mouton, Malan, Leenen and Venter [28]. The article has been published in 2014 and contains, among other things a summary of the authors' previous work on the topic – an ontological model of social engineering attacks [29]. We will first present their ontological model.

A Social engineering attack has:

- One "Social engineer", which may either be an *Individual* or a *Group of Individuals.*
- One "Target", which may either be an *Individual* or an *Organization.*
- One or more "Compliance Principles", which can be any of the following:
  - *Friendship or Liking* – the victim will be more willing to comply with the attacker if a predisposition is present or the latter is considered a friend.
  - *Commitment or Consistency* – people have a tendency to follow through commitments and to act in consistency with their position.
  - *Scarcity* – a psychological tendency is to value scarce "resources" higher, regardless of their actual value.
  - *Reciprocity* – informally 'returning the favor'; at a subconscious level people are predisposed towards complying with others, who treated them well.
  - *Social Validation* – utilizing social pressure; people tend to behave in ways they consider socially acceptable and to avoid socially unacceptable behaviour.
  - *Authority* – victims have harder time denying requests that come from people with perceived authority over them.

- One or more "Techniques", which can be any of the following (discussed later):
  - *Phishing*
  - *Pretexting*
  - *Baiting*
  - *Quid Pro Quo*
- One "Medium", which can be:
  - *Email*
  - *Face to Face*
  - *Telephone*
  - *SMS*
  - *Paper Mail*
  - *Storage Media*
  - *Webpage*
  - *Pamphlets*
- One "Goal", which can be *Financial gain, Unauthorized Access or Service Disruption*.

Additionally, the attack itself can either utilize *Direct Communication*, further subdivided into *Bidirectional Communication* (e.g. using telephone or face-to-face) and *Unidirectional Communication* (e.g. Paper mail or SMS); or *Indirect Communication* (e.g. USB flash drives).

In their follow-up research, the authors propose an extended attack framework, consisting of six phases, which on their own require more than one steps. The extended framework is also cyclical at more than one level – it may involve repeating some steps and phases without completing the full cycle. The phases are as follows:

1) Attack Formulation – in this phase, the attacker defines his goals for the attack and subsequently identifies the target of the attack. This phase is not present in the original model, but is of great importance, as it helps explain attacks vectors like phishing, which do not need information gathering.

2) Information Gathering – this phase consists of 3 steps, which are however cyclical in nature, that is they may be repeated multiple times within the phase. These steps are:
   1) Identify potential sources.
   2) Gather information from sources (which can be public or private).
   3) Assess gathered information and depending on its relevance continue with next phase or go back to step 1.

3) Preparation – in this phase, the attacker evaluates the sufficiency of the gathered information and prepares an attack vector:
   1) Combination and analysis of gathered information – in this step, the 'bigger picture' is assembled and a believable pretext is constructed. We will discuss pretexts further.
   2) Development of an attack vector – this attack vector should specify the following attributes of the attack – Goal, Target, Social engineer, Medium, Compliance Principles and Techniques. As an example of the result of that phase, the authors present a typical phishing attack: composing an email with similar content to the ones, internal to the organization, and attaches or links a malicious file, which will be used for the attack, following the exploitation of the social engineering vector. As another example, an attacker may prepare a much more generic (yet sufficiently specific to succeed) phishing email, targeting hiring managers and containing a link

to a ransomware binary. Indeed, this is currently one of the most popular variations of phishing attacks [35]. An important note is due – should the attacker consider his attack vector as insufficiently satisfactory in the context of the end goal of the attack, he can instead return to further information gathering.

4) Develop a relationship – this is the phase where the attacker ensures the predisposition of the victim. It is further broken down into two steps:
    1) Establishment of communication – in case a pretext is present, it is used at the very beginning of the communication.
    2) Rapport building – probably the trickiest step in the whole attack process. A positive relationship (in the mind of the victim) must be established – e.g. the attacker must endear himself or herself to the victim.

5) Exploit the relationship – this phase is the actual implementation of the attack:
    1) Priming the target – manipulating the victim to achieve the desired emotional (and, more generally, mental) state. At the end of this step, the *actual* gain in trust has already happened, even if it is "consumed" at the next step, so in certain regards it is analogous to privilege escalation in technical exploitation.
    2) Elicitation – extracting the outlined information or 'favor' from the victim. This step is the ultimate test for the success of the attack vector – at this point the defined vector will either fail or produce the expected result. A typical example of information requested is the victim's credentials (their password in particular).

6) Debrief – the final phase of the attack cycle:
    1) Maintenance – this step involves (once again) manipulating the victim to achieve the desired emotional state, however in this case the desired state is actually calm and relaxed. This is important, as it decreases the likelihood that the victim will realize that an attack took place – on the other hand, if the victim is stressed after the end of the attack, they will reflect on the event for much longer and with much greater intensity – stressful events lead to intensive memories. The authors additionally cite an example where the psychological stress of being successfully exploited has led to extreme feelings of inadequacy, depression and, ultimately, suicide. While concerns about the well-being of the victim are humane, it is hardly expected that an attacker will share them.
    2) Transition – the final step, at this point the attacker must decide whether the set goal is achieved, or the cycle must be restarted in order to gain further access and/or information.

It should be noted that there are alternative classifications, for example "The Cycle", as defined in Ref. [32], probably referencing the "Social Engineering Framework" by Hadnagy et al. [7], which attempts to provide systematic learning resources on the topic.

4.3 Some Social Engineering Primitives

This list is by no means exhaustive, however it contains some of the more popular social engineering methods. Among the topics that are not covered are: obtaining used/refurbished/recycled HDDs and other storage devices and performing forensic analysis on them; using intimidation and/or coercion instead of a rapport (a much harder task, as people generally react negatively when given negative stimuli); elicitation; advanced psychological tactics, such as microexpressions and neuro-linguistic programming (NLP). It should be noted that there is considerable debate in scientific circles (in particular in psychology) pertaining to the actual validity and scientific value of the last two approaches. Successful social engineers have consistently claimed success in using these tools and have advertised them.

4.3.1 "Dumpster Diving"

The practice of examining an organization's trash for any information leaks. This method is the main reason for the adoption of document disposal (a.k.a. security shredding) policies in many organizations. Widely popularized by Kevin Mitnick.

4.3.2 "Shoulder Surfing"

An extremely simple method – simply looking over a victim's shoulder to observe and hijack important information, being typed by the victim, e.g. a password. More advanced versions may include remote surveillance, using binoculars or other equipment. [42]

4.3.3 Collecting OSINT

Open-source intelligence (OSINT) is a term, denoting information, collected from publicly available sources. Such information has many upsides for an attacker, chief of which is the possibility to remain anonymous and untraceable (given that the appropriate technical measures are taken). Therefore, such collecting such information carries practically no risk of detection. Typical sources of OSINT are social networks (applicable mostly for individuals); search engines; publicly-available information for organizations such as DNS WHOIS records, government-mandated information, such as SEC filings and others.

4.3.4 Pretexting

Pretexting is the process of 'inventing' a believable and sound (to the victim) pretext, and more generally context, that gives the attacker a legitimate-sounding reason to request sensitive information or a favor from the victim. Classical examples of general pretexts in social engineering include participation in a survey, 'research projects' or similar situations, where the victim would be more predisposed to give up valuable information or perform an action on behalf of the social engineer, in this particular case due to the implied (or even explicit) guarantee of anonymity to participants in such initiatives. Another traditional example of a pretext is a situation in which the attacker claims to work for a utility company, wears a uniform and is present at the location to 'take care' or 'fix' some problem. Due to the subconscious desire to not inhibit other people's work and waste their time, the victim is often ready to give physical access to the attacker without further verification, due to the implicit 'authority' of the uniform, even if this authority does not give 'social power' over the victim.

Pretexting is one of the aspects of social engineering that is almost entirely psychological in nature and as such we will not discuss it any further.

4.3.5 Phishing

We have already introduced the concept of phishing, however due to the pandemic proportions of attacks, utilizing this vector, it deserves a more detailed look in our examination of attack methods. While phishing as a method is not theoretically related to email, that is almost always the case in practice. We will limit our discussion to email-based phishing. According to Verizon's Data Breach Investigation Report [4] for 2016, phishing was involved in 9576 data breaches (916 of them having resulted in confirmed data disclosure) out of about 100000, or roughly 10% of the considered data breaches were caused by phishing. Phishing is often used as an initial attack vector and usually more elaborate technical methods are used once an initial PC inside the organization's network has been infected by the attackers. Another interesting metric is open rate for phishing emails – per the same report, 13% of users clicked on the link or attachment in a phishing email. Combining a phishing campaign with a zero-day vulnerability (as is commonly the case with APTs, e.g. the SecurID breach) means that all the aforementioned users will have their PCs infected. To put that in an even clearer perspective – any organization with at least 5 users with public email addresses is statistically expected (i.e. will happen with probability > 50%, assuming a Bernoulli distribution with success rate 13%

for individual user opening a phishing email) be infiltrated by phishing-borne malware if targeted by such a campaign. Furthermore, research indicates that the 'best' phishing campaigns have click rates of 45% [34].

We will consider two particular variants of phishing:

- Spear phishing – a spear phishing attack is similar to normal phishing in its technical principles. The difference is, however, in the degree of information gathering done before the attack. To be more precise, spear phishing is much more targeted – it involves selecting a target and tailoring the messages to the particular victim – e.g. using the actual names of employees in the "From:" field, following similar structure and tone to in-company communication, etc. The name is a wordplay on 'spear fishing', a fishing technique, where the fisherman targets the fish he wants to catch (with a spear), instead of waiting for it to catch the bait. Another possible analogy is sharpness – while standard phishing campaigns operate by using strength in numbers and sending millions of messages, spear phishing sends only a tiny fraction of the traffic and is, therefore, much harder to notice as a pattern on a large scale.
- Soft targeting – this type of phishing attack sits somewhere between standard phishing and spear phishing: the email messages are targeted, but not towards a certain individual or an organization, but rather towards a *profession* or *occupation*. This allows the attackers to use the 'strength in numbers' paradigm by sending millions of messages, while still having a higher success rate than completely generic phishing messages. Soft targeting is particularly popular at the moment of writing of this paper, as it is used as the main distribution vector for ransomware – the malware 'trend' of 2016 [35].

4.3.8 "Baiting"

A technique in which the attacker places a 'bait' for the victim to take on their own initiative – the typical example being leaving one or more USB flash drives, containing a malicious executable, in a spot, where the victim is likely to notice them. Then, motivated by curiosity or greed, the victim may take the 'bait' and unwittingly help the attacker's payload cross a trust or security boundary that the attacker himself/herself cannot – for example a physical barrier with adequate access control.

4.3.9 Reverse Social Engineering

Reverse social engineering is present when the attacker creates a situation that influences the victim into initiating the contact, usually asking for a solution to a technical issue. The main advantage of reverse social engineering is that the users very rarely doubt the authenticity of communication that *they* have initiated, due to the present illusion of control (i.e. the assumption that it was them who actually initiated the connection).

4.3.10 "Piggybacking"

When "piggybacking", an attacker attempts to pressure the victim into allowing him past a security checkpoint. The canonical example is something functionally equivalent to asking the victim "Can you hold the door? I forgot my access card upstairs". Due to the natural inclination to help, many people would let them through, without realizing that they are, in fact, bypassing a physical access control mechanism and often even inhibiting any log entries, that would otherwise be generated by the system, which could allow easier detection of the attack in hindsight.

5.0 DEFENDING AGAINST SOCIAL ENGINEERING

5.1 Host and Network Technical Defences

Most, if not all, traditional host and network technical defense methods in information security, including firewalls, antivirus solutions, intrusion detection/prevention systems (IDS/IPS), generally consider social

engineering out of scope. Therefore, they do not attempt to address the 'human factor' in the security equation, instead focusing only on technical threats.

While these methods have very little positive impact on preventing social engineering attacks, they may be useful for thwarting further escalation by technical means. In practice, such a defense mechanism is of little efficiency, as any competent attacker would ensure that a backdoor he intends to use is not detected by AV and other defensive products. This renders such tools almost completely inefficient against anything but the most amateurish attacks of this type.

5.2 Email Security Solutions

As we already noted, email phishing is the most common attack vector for social engineering attacks. Given that phishing emails are almost exclusively spoofed, solutions, providing email authentication, are effective at preventing phishing. Of course, such solutions are only effective against email phishing attacks and do not address the larger and more general concerns of social engineering.

There are at least three technical solutions to providing email communication authentication – the OpenPGP standard [36] for encrypting and (more important in this context) signing messages, the S/MIME protocol [37], which is functionally similar to OpenPGP with the notable exception of public key handling, and the SPF/DKIM/DMARC set of protocols, which provide a mostly DNS-based solution [39].

The first one provides support end-to-end encrypted and/or signed email messages. While technically sound in most regards, the OpenPGP standard unfortunately does not provide a definitive and user-friendly solution for distributing public keys, mostly due to the inherent trust issues – placing trust in a central authority that handles PKI goes against the core ideas of the standard. Therefore, the burden of verifying the authenticity of the senders or receiver's public key lies entirely on the user, which is a significant barrier to Internet-level scaling [38]. Furthermore, the protocol is notably geared towards individuals, rather than organizations.

The S/MIME protocol is somewhat similar in design and goals [37] to the OpenPGP standard, with the notable difference that it utilizes the same centralized trust model of the X.509 certificate system, used to secure (among other things) TLS connections. This has the obvious upside that it integrates much better into typical PKI operations and, in general, S/MIME is much more widely deployed in enterprises than OpenPGP.

The last solution, the set of protocols SPF, DKIM and DMARC are complementary methods in assuring authentication that all work on a domain basis using the DNS protocol, making them much easier to deploy in enterprises or large organizations. The Sender Policy Framework (SPF) defines a format for DNS TXT records for the domain that define a list of IP addresses, allowed to send emails on behalf of the domain. Any conformant receiving email server must check the presence of DNS SPF records for the sender domain and reject emails that are sent from unauthorized addresses. The DomainKeys Identified Mail protocol inserts a header in each email message, containing a cryptographic signature of the contents of the email. The receiving party can then retrieve the corresponding public key through DNS TXT records and verify the signature. Finally, the Domain-based Message Authentication, Reporting and Conformance (DMARC) protocol, allows the sending domain to announce its policy for sending email, in particular whether SPF and DKIM are implemented and supported [39], [40].

A common issue of all three approaches is the opt-in nature, i.e. it is necessary for an organization to convince *other* organizations it has regular and expected email communication with to adopt one or more of these solutions. Other practical issues that is not addressed by any of these solutions are the 'display name' attack and visually similar sender domain names, which can trick users into believing the email originates from a 'trusted' domain. However, assuming a full and correct implementation of these technologies, the purely technical aspect of phishing can be mitigated. Unfortunately, as we have previously noted, social

engineering, and by implication phishing, is first and foremost psychological and not technical in nature. A good overview of 'best practices' in providing email authentication can be found in Refs. [40] and [41].

5.3 Education and Awareness Trainings

A commonly suggested defensive tactic against social engineering attacks is to ensure all employees receive (mandatory) training in recognizing and dealing with social engineering attacks.

Additionally, special attention should be given to social networks/media, since their use is extremely prevalent, yet users often do not realize the amount of data they have publicly shared. Since this data probably identifies them as an employee of the organization, it indirectly raises their own risk of becoming a social engineering attack victim. One interesting advice is to suggest to employees that they do not indicate their workplace in social media, so that they are not targeted as such.

It is important to note that employees at different positions should receive a different version of the training [42]. Indeed, the chief officers are probably the highest priority target group in the organization for a social engineer, therefore, they must be much more aware of the challenges they may face and the consequences of a successful social engineering attack for the company (e.g. the "CEO frauds").

Some authors, however, argue that awareness campaigns are bound to fail and instead suggest combining them with phishing simulations [43].

5.4 Phishing Simulations

Currently, a very popular approach to mitigating the threat of phishing is to use a "phishing simulation". The approach can be concisely described as performing a phishing attack towards the target group in a controlled environment and delivering feedback to the participants, based on personal performance.

The methodology varies between providers, but often includes providing an initial 'basic' training on the subject and most vendors recommend continuous efforts in this direction – in particular repeating the phishing simulations on a regular basis [44].

One vendor survey claimed that after only 4 sessions, user susceptibility to phishing decreased with 97.14% [44].

5.5 Policy-Level Defences

Many literature sources and general information security 'best practices' demand that an organization have a formal and sound policy for personnel. Suggestions from different sources include:

- Implementing a 'least privilege' policy [46] and **ensuring that users understand** the reasoning behind it and that it is designed to protect them, too – a perceived arbitrary imposition of management over other employees will only generate dissatisfaction and could probably increase the risk, associated with the 'human factor', due to e.g. disgruntled employees.
- Implementing the usage of 'scripts' for each workflow when employees are communicating over phone/email.
- Having a well-established contact point and escalation path for reporting suspected social engineering attempts. This will help employees become proactive in defending the organization from such attacks, if combined with the proper organizational culture and training. Of course, unlike with IT systems, should the designated contact fail to act on received information, this would have an adverse effect – users will assume that such events 'are not a big deal' or that management is not sufficiently interested and will quickly become demotivated to report such occurrences.

- Deploying two-factor authentication wherever possible. This has the obvious upside of preventing direct damage from employee credential theft.
- Create and deploy an information classification – in particular formally denote sensitive information as such and ensure that handling procedures for sensitive information impose increased security requirements.
- Albeit somewhat out of scope for a policy, creating an organization culture of general awareness of social engineering. As an example, explaining why tailgating/piggybacking is an issue and communicating clearly that it is each employee's personal responsibility to prevent such occurrences can help immensely.

As a more structured approach to the issue, the COBIT 5 framework "also considers human factors and behavior as key enablers (albeit often underestimated) for designing a holistic approach for GEIT" (Governance of Enterprise IT) [45].

5.6 Audits and Penetration Tests

Another common line of defense is performing regular external audits and penetration tests. Such verification and validation of security controls in place of the organizations is almost always a positive force with regards to the overall information security of the organization. Of course, any respectable penetration test and most comprehensive security audits include testing social engineering attack vectors. The results of these tests and audits could greatly benefit the organization by providing an overview of the current issues, which could in turn allow management to implement a plan for improvement and, at the end, to achieve an acceptable level of security. We should once again point out that this should *not* be a one-time occasion, but rather a regular process that provides feedback on set intervals, as, like any other process, security tends to erode when left unchecked.

5.7 Technological Anti-Phishing solutions

In addition to all other discussed methods, there are several relatively new technologies that attempt to detect phishing attempts and warn the user of it [47]:

- **C**arnegie Mellon **Anti**-phishing and **N**etwork **A**nalysis Tool (CANTINA) – in this approach, a combination of DOM analysis and search engine results are combined to detect potential phishing sites. The approach boasts some impressive technical results – it successfully detects 95% of phishing sites.
- CodeShield uses a Personalized Application Whitelist (PAW) to automatically block any attempted phishing website by virtue of them not being in the whitelist.
- Google's Password Alert is a browser extension that focuses on early warning and mitigation, rather than prevention – it detects whenever a user enters their Google account credentials on another website and warns them to immediately change their Google password.
- AuntieTuna – a browser plugin that uses personalization in conjunction with detection algorithms to decide whether or not a page is a phishing attempt or not.

6.0 OPEN CHALLENGES AND RESEARCH DIRECTIONS

6.1 Passwords as Authentication Factors

Passwords are the cornerstone of current security in IT systems. However, passwords are also quite bad as authentication factors in the context that we currently use them.

In particular, passwords have non-trivial requirements in the context of Web security. A particularly striking issue is the necessity for using different passwords for different service providers. It is proven that keeping secrets generates stress in humans and passwords are authentication factors of the 'knowledge' variety, which makes them secrets by design. However, in modern life it is often common for an average user to have accounts in tens, if not hundreds, different service providers – and remembering such an amount of passwords is an impractical, if not outright impossible task. From a psychological point of view, such a burden is essentially too heavy to place on an individual for the reward of basic secure browsing, which is exactly why many users choose to (sometimes wilfully) disregard best security practices. While not yet a generally shared opinion, it is the author's belief that passwords have mostly outlived their usefulness as an authentication mechanism and that view is slowly becoming more and more prevalent, both in academia and in industry. Many researchers are proposing alternatives to using passwords that would make the problem more manageable for people in general. It is worth noting that while passwords are not *directly* related to social engineering, their theft is often the sole purpose of such an attack and replacing them with an alternative mechanism that is not prone to such theft would automatically alleviate the pressure on humans to preserve the confidentiality of their authentication factors.

While research in this direction is active, not many viable (and sufficiently mature) alternatives have been presented. Using purely biometric authentication factors has many nontrivial issues, for example some biometric factors may change for an individual and therefore cause rejection of an authorized user, and that most third-party services would not allow integrating such authentication mechanism. A practical alternative that is currently used is password managers. While they improve the usability concerns, they do not substantially improve security and sometimes even reduce it – a compromised machine will still result in credential theft and, in most cases, theft of *all* credentials, stored in the manager.

6.2 Usability and Security

Another issue that consistently has plagued information and IT security is the lack of concern for usability when designing security systems, controls and mechanisms. This is somewhat due to the perceived 'technical' nature of the problem – developers, architects and designers, tasked with security, often consider the human a 'passive' element that must comply with their chosen design. Such treatment is one of the root causes of social engineering attacks. A designer's assumption of his users should not be that of 'a critical unpatchable weakness' that are there to ruin the security of the system – quoting Josephine Wolff, "It's an attitude strangely reminiscent of a certain type of hostile librarian who gives the impression that she would much prefer you not touch, or even breathe on, any of the precious books in her care" [49].

What many authors consider as a better approach is to instead focus on improving *usability* of security, as far as end users are concerned. This includes up to a certain degree the previous point – most of the concerns raised against passwords can be summarized as usability issues. Another example is OpenPGP – while mostly technologically sound, the technology has been widely criticized as almost unusable [38]. Among the suggested possibilities for solutions to this issue are: including usability as a requirement in particular when related to security, defining metrics for security usability and performing usability tests that identify usability issues during the development process itself [50].

7.0 REFERENCES

[1] The role of cybersecurity in this year's US presidential election, National Cybersecurity Institute, September 13, 2016, accessed October 9, 2016, http://www.nationalcybersecurityinstitute.org/general-public-interests/the-role-of-cybersecurity-in-this-years-us-presidential-election/.

[2] Siegel, Bari Faye, *Internet Creations Security Expert Warns of Costly Effects of Ignoring Cyber Threats*, Internet Creations Blog, March 31, 2016, accessed October 9, 2016, https://blog.internetcreations.com/2016/03/one-wrong-click/#.WAdbCfl97IV.

[3] Goodchild, Joan, *History's infamous social engineers*, Network World, Inc., January 4, 2012, accessed October 9, 2016, http://www.networkworld.com/article/2287427/network-security/history-s-infamous-social-engineers.html#slide6.

[4] *Verizon's 2016 Data Breach Investigations Report*, Verizon Enterprise, downloaded on October 9, 2016 from http://www.verizonenterprise.com/verizon-insights-lab/dbir/2016/.

[5] Harl, *People Hacking: The Psychology of Social Engineering*, Talk at Access All Areas III, May 7, 1997, accessed on October 9, 2016 through the Internet Wayback Machine, https://web.archive.org/web/20060306051447/http://noblit.com/docs/people-hacking.pdf, http://www.isoc.org/inet99/proceedings/3g/3g_2.htm.

[6] *Glossary of Security Terms*, SANS Institute, accessed on October 9, 2016, https://www.sans.org/security-resources/glossary-of-terms/.

[7] Hadnagy, Christopher, *Social Engineering: The Art of Human Hacking, 1st Edition*, ISBN-13: 978-0470639535.

[8] Evans, Nathaniel Joseph, *"Information technology social engineering: an academic definition and study of social engineering – analyzing the human firewall",* 2009. Graduate Theses and Dissertations. Paper 10709.

[9] Atkins, B. and Huang, W., *A Study of Social Engineering in Online Frauds. Open Journal of Social Sciences*, 2013, 1, 23-32. doi: 10.4236/jss.2013.13004.

[10] Schneier, B., *Social Engineering Diamond Theft*, "Schneier on Security" blog, March 19, 2007, accessed on October 9, 2016, https://www.schneier.com/blog/archives/2007/03/social_engineer_3.html.

[11] Usher, S., *10 Commandments for Con Men*, "Lists of Note", February 29, 2012, accessed on October 9, 2016, http://www.listsofnote.com/2012/02/10-commandments-for-con-men.html.

[12] Mitnick, D., and Simon, W., *The Art of Deception: Controlling the Human Element of Security*, 2003, ISBN-13: 978-0764542800.

[13] Gelner, B., *Super-hacker Kevin Mitnick takes a plea*, "Computer Fraud & Security", Volume 1999, Issue 5, 1999, doi:10.1016/S1361-3723(99)90141-0.

[14] Ilyin, Y., *Can we beat social engineering?*, Kaspersky Lab Business, August 11, 2014, accessed October 9, 2016, https://business.kaspersky.com/can-we-beat-social-engineering/2363/.

[15] Coviello, A., *Open Letter to Customers*, RSA Security, Inc., June 6, 2011, hosted on "Network Computing Architects, Inc.", accessed on October 9, 2016, http://www.ncanet.com/resources/press-releases/91-2011-06-08-art-coviello-rsa-open-letter-customers.html.

[16] *Anatomy of an attack*, RSA Blogs, April 1, 2011, accessed on October 9, 2016, http://blogs.rsa.com/anatomy-of-an-attack/.

[17] Jarmoc, J., *RSA compromise: Impacts on SecurID*, SecureWorks, Inc., March 17, 2011, accessed on October 9, 2016, https://www.secureworks.com/research/rsacompromise.

[18] *Social engineering attacks on the rise, part 1: eBay breach*, Trend Micro, June 26, 2014, accessed on October 9, 2016, http://blog.trendmicro.com/social-engineering-attacks-rise-part-1-ebay-breach/.

[19] Hunt, T., *The eBay breach: answers to the questions that will inevitably be asked*, May 21, 2014, accessed on October 9, 2016, https://www.troyhunt.com/the-ebay-breach-answers-to-questions/.

[20] Krebs, B., *Email Attack on Vendor Set Up Breach at Target*, February 12, 14, accessed on October 9, 2016, https://krebsonsecurity.com/2014/02/email-attack-on-vendor-set-up-breach-at-target/.

[21] *Target Confirms Unauthorized Access to Payment Card Data in U.S. Stores*, Target Brands, Inc. Corporate, December 19, accessed on October 9, 2016, https://corporate.target.com/press/releases/2013/12/target-confirms-unauthorized-access-to-payment-car.

[22] *Target Provides Update on Data Breach and Financial Performance*, Target Brands, Inc. Corporate, January 10, 2014, accessed on October 9, 2016, https://corporate.target.com/press/releases/2014/01/target-provides-update-on-data-breach-and-financia.

[23] *eBay Inc. To Ask eBay Users To Change Passwords*, eBay Inc. Staff, May 21, 2014, accessed on October 9, 2016, https://www.ebayinc.com/stories/news/ebay-inc-ask-ebay-users-change-passwords/.

[24] *The Great Bank Robbery: the Carbanak APT*, AO Kaspersky Lab, user GReAT, February 16, 2015, accessed on October 9, 2016, https://securelist.com/blog/research/68732/the-great-bank-robbery-the-carbanak-apt/.

[25] Paganini, P., *Carbanak Cybergang Swipes Over $1 Billion from Banks*, InfoSec Institute, February 18, 2015, accessed on October 9, 2016, http://resources.infosecinstitute.com/carbanak-cybergang-swipes-1-billion-banks/.

[26] *Anunak: APT against financial institutions*, Group-IB and Fox-IT, December 2014, accessed on October 9, 2016, https://www.fox-it.com/en/files/2014/12/Anunak_APT-against-financial-institutions2.pdf.

[27] Krebs, B., *FBI: $2.3 Billion Lost to CEO Email Scams*, "Krebs on Security", April 07, 2016, accessed on October 9, 2016, https://krebsonsecurity.com/2016/04/fbi-2-3-billion-lost-to-ceo-email-scams/.

[28] Mouton, F. et al., *Social Engineering Attack Framework*, Information Security for South Africa, 2014, DOI: 10.1109/ISSA.2014.6950510.

[29] Mouton, F. et al., *Towards an Ontological Model Defining the Social Engineering Domain*, IFIP Advances in Information and Communication Technology 431, July 2014, DOI: 10.1007/978-3-662-44208-1_22.

[30] Higbee, A., *Phishing and Ransomware Threats Soared in Q1 2016*, PhishMe, June 9, 2016, accessed on October 9, 2016, http://phishme.com/phishing-ransomware-threats-soared-q1-2016/.

[31] Ivaturi, Koteswara and Janczewski, Lech, *A Typology Of Social Engineering Attacks – An Information Science Perspective*, 2012, PACIS 2012 Proceedings, Paper 145, http://aisel.aisnet.org/pacis2012/145.

[32] Allen, M., *Social Engineering: A Means To Violate A Computer System*, SANS Institute, June 2006, accessed October 9, 2016, https://www.sans.org/reading-room/whitepapers/engineering/social-engineering-means-violate-computer-system-529.

[33] Social Engineer, Inc., *Security through Education*, accessed on October 9, 2016, http://www.social-engineer.org/.

[34] Bursztein, E., Benko, B., Margolis, D, Pietraszek, et al., *Handcrafted Fraud and Extortion: Manual Account Hijacking in the Wild*, IMC '14 Proceedings of the 2014 Conference on Internet Measurement Conference, ACM, 1600 Amphitheatre Parkway, pp. 347-358.

[35] Sjouwerman, S., *93% of phishing attacks now have ransomware payloads*, KnowBe4 Inc., June 4, 2016, accessed on October 9, 2016, https://blog.knowbe4.com/alert-93-of-phishing-attacks-now-have-ransomware-payloads.

[36] Callas, J., Donnerhacke, et al., *OpenPGP Message Format*, RFC 4880, IETF, November 2007, accessed on October 9, 2016, https://tools.ietf.org/html/rfc4880.

[37] Ramsdell, B. et al., *Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.2 Message Specification*, RFC 5751, IETF, January 2010, accessed on October 9, 2016, https://tools.ietf.org/html/rfc5751.

[38] Green, M., *What's the matter with PGP?*, A Few Thoughts on Cryptographic Engineering, August 13, 2014, accessed on October 9, 2016, https://blog.cryptographyengineering.com/2014/08/13/whats-matter-with-pgp/.

[39] Kucherawy, M., *Domain-based Message Authentication, Reporting, and Conformance (DMARC)*, RFC 7489, IETF, March 2015, accessed on October 9, 2016, https://tools.ietf.org/html/rfc7489.

[40] Steve Lynch, *Domain-Based Message Authentication Reporting and Conformance*, InfoSec Institute, January 4, 2016, accessed on October 9, 2016, http://resources.infosecinstitute.com/domain-based-message-authentication-reporting-and-conformance/.

[41] Crocker, D., Zink, T., *M^{3}AAWG Trust in Email Begins with Authentication*, M^{3}AAWG, June 2008, Updated February 2015, accessed on October 9, 2015, https://www.m3aawg.org/sites/default/files/document/M3AAWG_Email_Authentication_Update-2015.pdf.

[42] Kee, J., *Social Engineering: Manipulating the Source*, SANS Institute, April 28, 2008, accessed on October 9, 2016, https://www.sans.org/reading-room/whitepapers/engineering/social-engineering-manipulating-source-32914.

[43] Randell, C., *Why security awareness campaigns fail*, MWR InfoSecurity, August 8, 2013, accessed on October 9, 2016, https://www.mwrinfosecurity.com/our-thinking/why-security-awareness-campaigns-fail/.

[44] *Enterprise Phishing Susceptibility Report*, PhishMe, 2015, downloaded from https://phishme.com/project/enterprise-phishing-susceptibility-report/.

[45] Puricelli, R., *The Underestimated Social Engineering Threat in IT Security Governance and Management*, ISACA Journal Volume 3, 2015.

[46] Emekauwa, U., *The Human Layer of Information Security Defense*, date not present, accessed on October 9, 2016, http://www.infosecwriters.com/text_resources/pdf/UEmekauwa_Social_Engineering.pdf.

[47] Ardi, C., Heidemann, J., *AuntieTuna: Personalized Content-based Phishing Detection*, NDSS Usable Security Workshop 2016, February 21, 2016, accessed on October 9, 2016, https://www.isi.edu/~calvin/papers/Ardi16a.pdf.

[48] Snyder, C., *Handling Human Hacking Creating a Comprehensive Defensive Strategy Against Modern Social Engineering*, Liberty University, Spring 2015, accessed on October 9, 2016, http://digitalcommons.liberty.edu/cgi/viewcontent.cgi?article=1510&context=honors.

[49] Wolff, J., *Calling Humans the "Weakest Link" in Computer Security Is Dangerous and Unhelpful*, Slate, January 22, 2016, accessed on October 9, 2016, http://www.slate.com/blogs/future_tense/2016/01/22/calling_humans_the_weakest_link_in_computer_security_is_dangerous.html.

[50] Jøsang, A., AlFayyadh, B. et al., *Security Usability Principles for Vulnerability Analysis and Risk Assessment*, Computer Security Applications Conference, 2007. ACSAC 2007. Twenty-Third Annual, DOI: 10.1109/ACSAC.2007.14.