DSV Report Series No. 09-001

Securing Information Assets: Understanding, Measuring and Protecting against Social Engineering Attacks

Securing Information Assets:

Understanding, Measuring and Protecting against Social Engineering Attacks

Marcus Nohlberg

Stockholm University

UNIVERSITY OF SKÖVDE

ISSN 1101-8526
ISRN SU-KTH/DSV/R--09/1--SE
ISBN 978-91-7155-786-5

Printed in Sweden by Universitetsservice US-AB, Stockholm 2008
Distributor: Department of Computer and Systems Sciences

Abstract

Social engineering denotes, within the realm of security, a type of attack against the human element during which the assailant induces the victim to release information or perform actions they should not. Our research on social engineering is divided into three areas: understanding, measuring and protecting. Understanding deals with finding out more about what social engineering is, and how it works. This is achieved through the study of previous work in information security as well as other relevant research areas. The measuring area is about trying to find methods and approaches that put numbers on an organization's vulnerability to social engineering attacks. Protecting covers the ways an organization can use to try to prevent attacks. A common approach is to educate the users on typical attacks, assailants, and their manipulative techniques. In many cases there are no preventive techniques, dealing with the human element of security, in place.

The results show that social engineering is a technique with a high probability of success. Furthermore, defense strategies against it are complicated, and susceptibility to it is difficult to measure. Important contributions are a model describing social engineering attacks and defenses, referred to as the Cycle of Deception, together with a thorough discussion on why and how social engineering works. We also propose new ways of conducting social engineering penetration testing and outline a set of recommendations for protection. It is crucial to involve managers more, but also to train the users with practical exercises instead of theoretical education, for example, by combining measuring exercises and penetration testing with training. We also discuss the future threat of Automated Social Engineering, in which software with a simple form of artificial intelligence can be used to act as humans using social engineering techniques online, making it quite hard for Internet users to trust anyone they communicate with online.

Acknowledgements

There are four persons who were absolutely crucial in this research process, mentioned in chronological order. The first is my former boss, now friend, Kenneth Alfelt. He set the wheels in motion, and got the process started. My supervisor Benkt Wangler made it all possible through constant support and, well, supervision. It has been an honor working with you. My co-supervisor Stewart Kowalski supplied a never-ending wealth of inspiration and brilliance. Thomas Ekström and The Logic Planet AB, gave me invaluable support.

I have also had the blessing of highly supportive friends and colleagues. I would especially like to mention Rose-Mharie Åhlfeldt who not only wrote a paper together with me, but also helped me out during the processes with so many things, as did my dear friend Carl-Johan Åkerberg and Alexander Backlund. Most of the papers were written in collaboration with others: Johannes Bäckström, Kerstin Karlsson, Markus Huber and Martin Boldt. It has been a blast writing papers with you, and I would love to do it again. Vera Lindros did a fantastic job helping me with the language of the thesis.

This thesis is dedicated to my family. This is not something I just write to keep the peace (since they will all be reading this) but because I mean it. You have helped me, inspired me, fed me and put up with me when I turned massively self-occupied during the final process of this thesis (and perhaps long before that). More concretely, this is dedicated to: My parents, Arne and Ingrid Nohlberg, for their undying support and love. Carina and Julius Mikszáth, my sister in blood and my brother in spirit. My dear sister Maria Almqvist and her husband, Erik. Then we have the next generation: My niece (Sara) and my nephews (Sebastian, Olof, Oskar, Victor and Elias).

The list of names here could be very long. So for anyone who does not find his or her name in the list above, this is for you: Thank you.

Yes, this turned out to be quite sentimental. But it is sentimental to write an Acknowledgement-section after a long research project such as this thesis. Try it yourself. You should! Research rocks!

Contents

Part I: The Research Frame and Research Results

Part I of this thesis consists of six chapters and presents the background and structure of the research. It also contains a description of the results of the research, the discussion and suggestions for future research.

Chapter 1 presents the research problem, aim and objectives together with a brief overview of the research approach, delimitations, results and contributions. Related research and the thesis structure are also described.

Chapter 2 presents the research background by describing fundamental concepts in information security.

Chapter 3 presents the research design; how it was planned and executed.

Chapter 4 presents social engineering and Phishing.

Chapter 5 presents the research results.

Chapter 6 presents the concluding discussion as well as suggestions for further research.

Introduction

> “I am a lot like you. We have the same interests, are of the same age and, yes, we even think strangely alike. It is weird that we never met before! I am the kind of person you can really trust. We share so much! We even dislike the same things! If I want a small favor from you, that is really nothing to talk about, now is it? That is what friends do, is it not?”

Social engineering[1] is about creating deceptive pretexts; examples of which can be seen in the fictional text above. In social engineering the goal is to trick the victim into sharing protected information, or to make them perform certain actions. It is a part of security sometimes omitted from the big picture, even though most information security professionals seem to agree that it is of great importance. The area of this thesis is information security, which is a broader approach to security than computer security in that it does not care about the form of the data it wants to protect – thereby including humans in the big picture.

Regarding security as a broad concept has flourished in recent years. We see some interesting numbers in the Global statistics from “the Global State of Information Security in 2007” (PWC, 2007). More and more companies now employ a CSO or a CISO (60 % in 2007, 43 % in 2006). There has also been a steep increase in having an overall information security strategy (37 % in 2006 and 57 % in 2007). Technical safeguards are all the rage, 88 % of the organizations now employ firewalls, 82 % use backups and 80 % have protection against spyware (PWC, 2007). What is notable, however, is that 63 % of the organizations do not use audits or monitoring to ensure that the employees actually follow the security policy. In fact, less than half (48 %) actually try to measure and review the efficiency rate of their security policies and procedures (PWC, 2007). Among major Swedish companies it is widely thought, historically, that security is of the utmost importance, no matter the cost (Brandon, 2003). With the continued increase in the focus on information security, it is notable that much of the attention is put on increasing technical security such as firewalls, anti-virus and so on. However,

[1] The term "social engineering" is not only used within security but also in sociology where it describes the practical use of sociological knowledge, and in political science when talking about the large-scale influence on attitudes and social behavior in society by governments or private groups.

an increasing number of organizations do not know about the number and nature of security incidents (PWC, 2007). This is probably not because they do not get data from their protective systems, but because security specialists have started to realize that information security is much more than what is easily measured with the conventional solutions and that focus has to be put on the human element of information security (PWC, 2007).

In this thesis we focus on the human element of security. We do this by addressing the attack technique called “social engineering”, a major concern in information security. Social engineering was made infamous by Kevin Mitnick, partly through his actions as a hacker, and partly because of his writings and speeches on this hacking technique. It is notable that Mitnick did in no way invent the technique, since it is basically frauds used in an information security context. Mitnick managed to get access to several high security government systems, not by using high tech password crackers or obscure bugs in the systems, but by using a con man’s approach to obtaining information. By piecing information together he managed to get the access he wanted. His most frequently used tools were the telephone and a well planned out ruse. Some of these deceptive tactics are used in Phishing, where they can be combined with technological means to provide a devastatingly efficient attack. The difference between Phishing and social engineering principally lies within the high degree of personal contact within social engineering, and the very limited amount of personal contact in Phishing. Therefore the scales differ; social engineering tends to be used against a limited number of targets (that has been selected with greater care), while Phishing uses data mining, making it similar to spam, in order to attack more marks.

Social engineering attacks are quite different from many of the primarily technical attacks, due to them also having a clear, specific aim. The vast majority of attacks and threats to security have been “script kiddies”, viruses, trojans and other broad attacks without specific aims (Mitnick & Simon, 2002). Recently there are indications that organized crime is behind more and more of the attacks, and the focus has moved away from creating viruses and attacks that are mostly a nuisance to creating attacks that can generate an income (Jackson Higgins, 2008).

Since many users do not believe that anyone would ever attack them, because they are not “rich and famous”, and that hackers cannot do much damage anyway (Brostoff, Sasse & Weirich, 2002), social engineering attacks can be highly successful. This attitude is also influenced by the fact that most users do not understand how security works and therefore construct their own, often incorrect, models (Adams & Sasse, 1999). The “old” way of managing information security has led to two specific problems, according to Adams and Sasse (1999 p. 45):

(a) users' lack of security awareness, and

(b) security departments' lack of knowledge about users, producing security mechanisms and systems that are not usable. These two factors lower users' motivation to produce secure work practices. This in turn reinforces security departments' belief that users are "inherently insecure" and leads to the introduction of stricter mechanisms, which require more effort from users.

The setting is thus; we are in a world in which security spending has increased but so have security incidents. In addition, the users lack security awareness and the protective measures are not well liked, among the users, or usable. At the same time, there are attacks, based on social engineering, which are easy to learn and hard to protect against, that can be quite efficient.

The Problem

It seems that stricter technical controls may not be a viable solution to the problems associated with humans and security. In fact, many users know their behavior is not compliant with the current security policies of the organization and instead find solace in the behavior of fellow employees, and the belief that the regulations are unrealistic (Brostoff, et al. 2002).

In a study carried out by Treasury Department inspectors, one third of Internal Revenue Service (IRS) employees divulged their logins and passwords to auditors who called pretending to be computer technicians (Dalrymple, 2005). There have been other studies on the "gullibility" of users, and the extent to which they submit information while under attack from perpetrators using Phishing and social engineering techniques. The results generally indicate that users are quite susceptible to these kinds of attacks, but due to a high degree of uncertainty about the studies' results, it is hard to determine the extent of their vulnerability. Nevertheless, some studies do provide a certain shock value. For instance, when revisiting the highly publicized "Chocolate for passwords" study three years later, it was shown that 64 % would submit their office computer password in exchange for a piece of chocolate (Kelly, 2007).

Perhaps it is as in a comment from one of the interviews conducted by Björck (2005, p. 186): "It doesn't matter what technology you have - there is no technology that can protect you against human beings - forget it."

The following quote is Gartner's (2002a) comments about the risks:

"Malicious individuals have always known that the best way around any security system is to manipulate a human target into giving them what

> they want – what we call social engineering. It remains the single greatest security threat to enterprises."

It is apparent that there are great risks associated with humans and security, and especially with regard to attacks aimed at human weaknesses. We, in the information security world, often know little about what it is that makes these attacks succeed, how we can measure them and, perhaps more importantly, how we can prevent them. If we do not know the extent of the risks, we can hardly realistically judge whether any protective measures applied are actually increasing security or if they are just for show. From a cybernetics perspective, this amounts to Ashby's "Law of Requisite Variety" (Heylighen & Joslyn, 2000). If we want to secure a system, we need to know what to secure it against. Furthermore, if we ignore attacks aimed at humans, we almost ensure that those attacks are the ones to which we will be vulnerable. Security must be holistic.

Early in the research process, in attempting to identify what social engineering consists of, we created a mind-map inspired by Näckros (2005), which included different, and possible, influencing factors. This mind-map, Figure 1, which is in no way a complete description of the topic, provides an understanding of the complex issues that in some way *can* be argued to be connected to or influence social engineering and the human element of security. This mind-map was created as a tool with which to visualize the research area and is based on the findings from the literature study. The purpose of the mind-map is to illustrate the broad selection of research areas associated with the human element of security and should not be regarded as a complete analysis of the interconnecting areas.

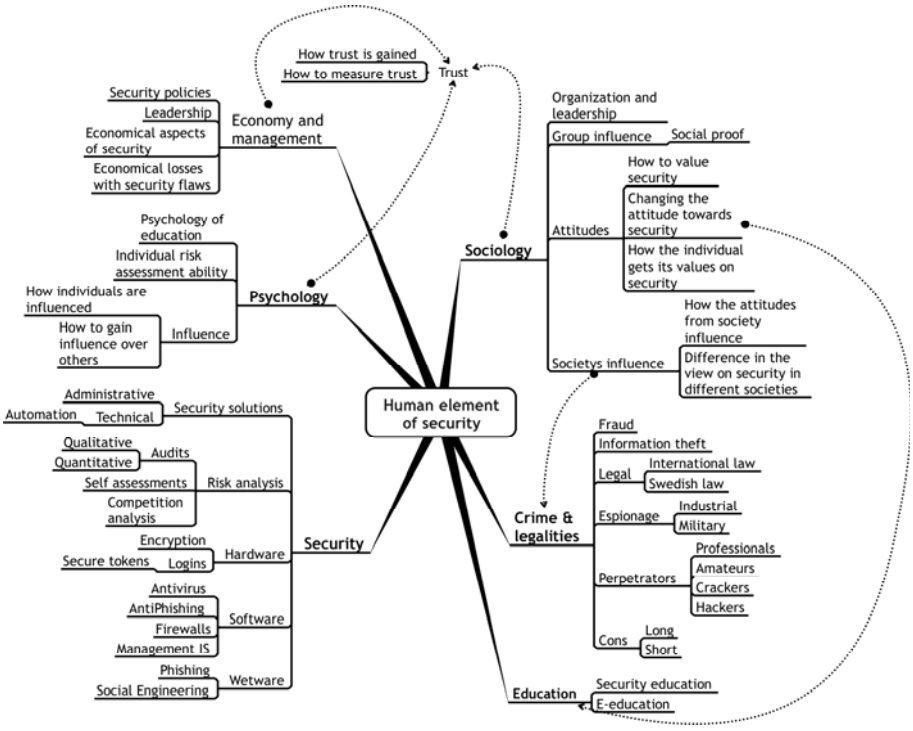


Figure 1. An illustration of research areas and subjects connected to the human element of security/social engineering.

In order to organize the research work, as well as its presentation, this thesis is structured like a temple, because as Sun Tzu states in “The Art of War” (Sun Tzu, translation from 1910): “Now the general who wins a battle makes many calculations in his temple ere the battle is fought.”

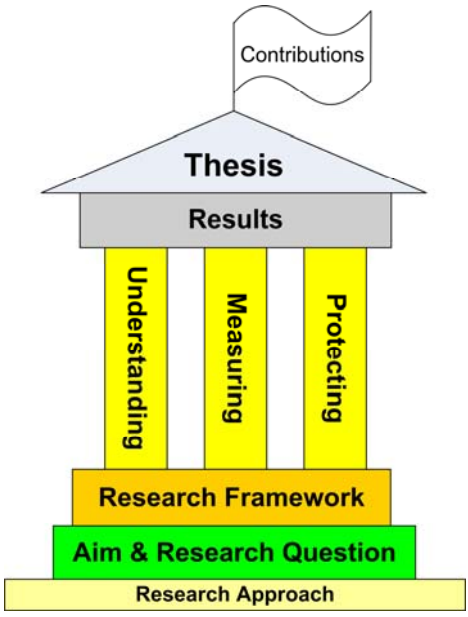


Figure 2. The research temple.

The temple, as illustrated in Figure 2, consists of a sturdy foundation, the research approach, on which the aim and the research question described in this section have been added. Additional sections of the thesis build the structure that culminates in the flag of contributions; the final part of the calculations made in the temple that comprises this research project.

The Aim

The aim of this work is to increase overall information security readiness by addressing some of the most efficient attacks that exist – attacks against the human element. This is achieved by gaining a working knowledge of the threat named social engineering, an often neglected, but, by many considered, crucial area of information security. In extending the knowledge of what makes humans susceptible to attacks, as well as learning about current attacks, their countermeasures and methods of testing organizations and individuals' vulnerability to them, we can make new and important research contributions to academia, and provide results that are useful for professionals working within information security.

The Research Question and Objectives

To put it succinctly, the research question of this work is: "What is social engineering and how can we best protect ourselves against attackers who use it?"

With regard to the research question, we believe there are three aspects of the area that are the most important ones to study. The selection of these three areas is based on previous work, experience, ongoing literature studies, as well as informal discussions with a number of information security professionals. In addition, each area represents an objective.

- Objective 1: To learn more about what constitutes social engineering. Also, what the underlying mechanisms are, how humans are influenced, and what techniques are being used by attackers? Furthermore, to create a model that better describes the concept, in order to understand the term. To improve knowledge of the area it is important to review the literature on the subject, and learn from other areas of research. This objective covers the area of understanding.
- Objective 2: In order to even begin protecting ourselves against social engineering attacks, we need a means of measuring the current level of readiness to see if the protective measures used have any effect. We cannot control what we cannot measure. Thus, it is necessary to study the methods of measuring an organization's vulnerability to social engineering. Although a few methods of penetration testing are in use today, many of these have ethical or practical problems. An effort to assess the efficiency of preventive approaches in this area is necessary. By testing several ways of measuring, a set of recommendations is created. This objective covers the area of measuring.
- Objective 3: To study how social engineering attacks can be prevented. In order to ensure security, protection against nefarious attacks is of the utmost importance. The current means of protection are studied, and novel defense approaches based on knowledge from other areas are tested. This objective covers the area of protecting.

The research approach aims to gain as much knowledge as possible from literature studies, mostly to obtain background knowledge on the subject. In order to acquire a deeper understanding and to achieve novel results from the organizations examined, both qualitative and quantitative studies are used.

The Research Approach

The research began with an early phase, which developed from an idea of examining the human element while interviews with systems administrators were being conducted in a study on overall security readiness in health care (Åhlfeldt & Nohlberg, 2005), in combination with an ongoing literature study. As this literature study and the interviews with systems administrators indicated a need to examine social engineering further, the subject was divided into three sub-areas, as discussed in 1.3. The subsections below discuss these three areas in more detail.

In order to address the research question and the objectives, a research plan was designed. The goal was to work in each of the three objective areas, and to try to maintain an industrial focus with some practical applicability of the results. During the research process, literature was continuously studied in order to obtain the necessary background information, as well as updates on new related research. The plan is illustrated in Figure 3, and described further in Chapter 3. Protecting is covered in most of the papers, but Paper 2 specifically focuses on that area.

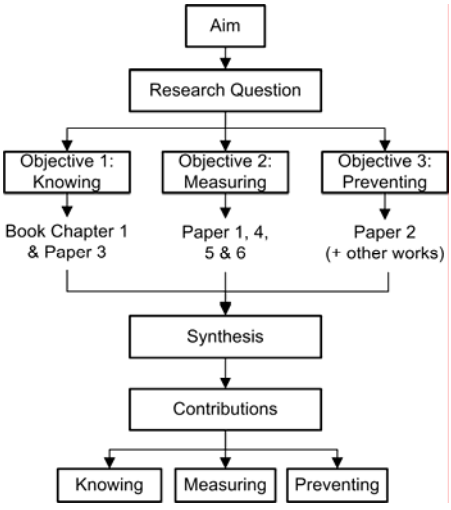


Figure 3. The research plan concerning aim, research questions and objectives.

The published results from the studies are found in Part II, and summaries of the results and contributions are presented in "Concluding Discussions" on page 73.

Understanding

In order to learn more about social engineering in general, we studied the materials that could be found in the literature, written and online. A broad range of subjects were studied, ranging from sociology, psychology, and criminology, as well as information security and more obscure literature, both academic and popular. One of the conclusions we could draw was that noticeable amounts of texts are derived from a small selection of imperative works that we then chose to focus on. The most relevant parts of the knowledge about social engineering and manipulative techniques were summarized in a book chapter (Book Chapter 1, Part II). As we found some flaws in the common description of what constitutes social engineering, as well as the attack cycle in general, we developed a new model describing social engineering. The cycle of deception was created using the results from the literature study, as well as those gained from semi-structured interviews, as sug-

gested by May (2001), conducted with the lead criminal investigator in a relevant case. This cycle was subsequently further validated by discussing its usefulness (1) with security experts and (2) with a group of social workers who are often subject to deception.

Measuring

In the attempt to devise how to best measure susceptibility to social engineering attacks, a selection of methods was used. In the first study, a quantitative approach was used, in which hundreds of subjects were deceived into believing that they were answering questions related to “micro efficiency”. This false concept was used rather than openly informing that the questions related to information security (Paper 1, Part II). The same quantitative approach was used in a highly sophisticated spear-phishing attack against students in a study to determine whether security education actually made students behave with more security awareness. It did not, apparently (Paper 6, Part II). In a much softer approach, and in an attempt to find novel aspects of security, qualitative research was used in the next study. The subjects of this study were interviewed about security in general and social engineering in particular. They were asked to think about possible flaws after being given short introductions to the area. This approach exposed a selection of weaknesses and perhaps also increased the subjects’ interest in security afterwards. The study thereby bordered on action research, although this was not a primary goal. With regard to readiness at management level, we used qualitative interviews structured on the deception cycle as developed in Paper 3, Part II. By interviewing managers and senior employees at large corporations, we did obtain an overall image of the readiness for a highly probable kind of future attack, which is automated social engineering.

Protecting

One of the methods useful for preventing successful attacks is to carry out the measuring as discussed above. During those tests, a number of suggestions for protection also emerged and were discussed. We did realize that one of the crucial aspects of preventing these kinds of attacks was to make it possible for managers to understand the security risks in general and social engineering in particular. One way this was done was to study how senior security professionals inform managers about security (Nohlberg & Bäckström, 2007). We also realized that in order to facilitate ongoing information and awareness at a higher management level, a software solution was necessary. To address the problems associated with informing managers, who often have little interest in security nor time to learn a new piece of software, we developed a prototype management information system for information security (Paper 2, Part II). This was created using a user-

centered approach to security and software development. The interface was tested by the target audience and regarded as good. It was considered crucial by the studied organization that the interface design had a high degree of usability. If human weaknesses are to be addressed in a software solution such as the one proposed in Paper II, it is important to design a solution that is easy to use and adapted to the needs of the target audience. Usability is, in fact, a crucial part of security in general, and human related security in particular. If the users do not understand how to use the products correctly, there is a risk that they will find another way of using it, as unsafe as that might be. We need to pay attention to ensuring that the users understand the programs and emphasizing the importance of secure behavior when using computers and computer networks (Whitten & Tygar, 1998; Flechais & Sasse in Cranor & Garfinkel 2005; Dourish & Redmiles, 2002).

Research Delimitations

The human element of security can cover both unintentional human mistakes, as well as deliberate attacks by perpetrators. In this research, the focus lies on intended attacks that may exploit unintentional mistakes, such as divulging information to a stranger that asks for it, but not unintentional accidents such as pouring coffee on a laptop. This is because controls against mistakes are not useful in preventing intentional attacks, but controls against intentional acts can provide protection against both intentional and unintentional attacks.

There is a distinction between what is technical, and what we consider human related security in general and social engineering in particular. For example, it could be claimed that a virus exploiting bad code in an operating system is exploiting a human weakness; poor programming. However, in this thesis, social engineering is considered from an aspect of attackers intentionally using manipulative techniques as the main method of attack. Consequently, phishing is included to some degree, and regarded as a subset of social engineering, but we do not consider that viruses exploiting human weaknesses as a means of spreading, such as the well known e-mail viruses, belong to social engineering in this context, even if they use some of its techniques.

There is a distinct possibility that gender issues can play a part in social engineering, but they have not been a focus of this study. While the statistics are collected in some of the studies, we draw no specific conclusions based on gender, although this an interesting area of study in the future.

Furthermore, although the legal implications of social engineering are an interesting area, they are beyond the scope of this thesis.

Results and Contributions

This research has used a rather broad approach which has resulted in useful and new knowledge, suitable for a wide audience in both academia and among security professionals. The approach can also be used by future researchers, who can initiate further research based on this thesis. The results, in the form of papers, have all been published or presented in peer-reviewed academic conferences, journals or books.

Our work contributes with a merger of the aspects that enable social engineering, both from a social psychological, and a descriptive perspective that uses a model to describe the actions of the victim, the attacker and the defender. We also recommend how to carry out social engineering penetration testing, as well as suggest methods of protection against social engineering attacks. In addition, a novel future threat, in the form of Automated Social Engineering is described. The contributions are described further on page 76.

Related Research

This section presents related work in the area of information security.

There are a number of specific problems associated with humans and security in general and social engineering especially. In this chapter, the three main areas of interest for this research are discussed. The research has been divided into these areas on the basis of previous work, experience, ongoing literature studies, as well as informal discussions with a number of information security professionals.

One of the problems associated with this area of research is that there has been little interest in the area in the past. Björck and Yngström's (2001) study, for example, attempted to classify research in information security.

In their study, they classified the papers accepted by the "IFIP World Computer Congress" (SEC 2000) and placed their contribution and research area in a matrix. The Y-axis deals with whether or not the contribution is primarily focused on being *empirical*, or *theoretical*, and the X-axis consists of three areas. The *technical* area, for example, deals with computer hardware and software, communication protocols, as well as cryptographic algorithms and technical evaluation methodologies. In the *formal* area, they place research dealing with procedures to formalize human behavior in the information system. Examples include information security policy, the legal system, and so on. In the *informal* area, there is research about informal human behavior, for example, social relations, ethics, and security implications of intrapersonal communication. The results are illustrated in Figure 4. The dots within the dotted line represent research aiming to move from one level of abstraction to another, for instance, implementing a theory in the empirical world.

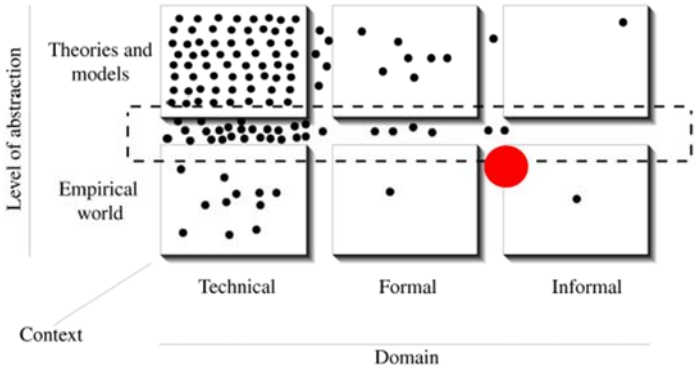


Figure 4. The classification of the 125 papers from the SEC 2000 proceedings (Björck and Yngström, 2001). The added large dot marks the research position in this thesis.

The research conducted in this thesis has mainly dealt with the informal area, what Björck and Yngström (2001) call "security implications of intrapersonal communication". A large part of our research deals with the empirical world, as we believe there is a need for a practical understanding of the risks and procedures, but at the same time there is also a need for models that can explain them. Our intended research position is indicated by the large red dot in the model above.

This research position is further motivated by the conclusions of Björck and Yngström (2001), in which they argue that the human element of security is one of the most important. They maintain that while 80 % of all information security research is currently being carried out in the technical domain, resources are perhaps better spent in the formal and informal domains where critical problems can be found.

In recent years there has been an increased interest in this research area, with books written on the subject and special conferences that focus on the human element, exemplified by the International Symposium on Human Aspects of Information Security & Assurance (HAISA).

The three research areas, understanding, measuring and protecting are presented below. They are discussed and compared to some of the state of the art research today, together with a conclusion of each area.

Understanding

The research framework of this thesis describes social engineering and Phishing, two of the major areas of what is often attributed to be the human element of security. They are examples of attacks aimed with deliberation at primarily humans and human weaknesses. There are, of course, security

flaws related to humans that are without deliberation, for example, those not carried out by attackers, but instead by unconcerned and unwitting users. Examples of such flaws can be a user that mistakenly destroys the wrong back-up CD, deletes the wrong file, and so on. In this research the focus is, however, on the deliberate attacks against humans, not the unintentional mistakes users can make.

Previous Work in the Research Field

Much of the material presented in this thesis describes parts of what constitutes the human element of security, the major impact on social engineering. There is ample material to be found regarding Phishing and social engineering; the problem is that most of the material is not of adequate academic standard. Most articles, and so on, are from contemporary magazines, web pages, and other such sources. Those of an academic standard tend to use the same references as the more contemporary ones, meaning that the factual basis on which the field is grounded is quite narrow. It is difficult to overlook the tremendous impact that Mitnick and Simon (2002) have had on the field, and little is written that is not to some extent covered or mentioned in their book, even though the field nowadays spans several hundreds, perhaps thousands, of information sources. This small set of foundational references does not automatically mean that the quality is low, but it can potentially be a problem since the research area is narrow and few researchers are working in it. It is important to continuously remember to check sources and to be critical when reading, especially since a lot of the sources are highly anecdotal web pages.

The typical perpetrators of social engineering attacks are described in this thesis on page 32. In the field of criminology, much work has been carried out to determine the types of individuals that become criminals. Within the study of deviance, there are ample theories explaining why people turn to crime. Most relevant to this area of research is perhaps the Differential Association theory developed by Edwin H. Sunderland (DeMelo, 2007). In Sutherland's differential association theory, the view is that criminal behavior, both techniques and values, is learned from social interaction with others. Once a potential perpetrator has learned the techniques, be they simple or complex, the values supporting the crime can be learned from just about anyone (DeMelo, 2007).

It is unavoidable that a research area such as this one describes the techniques actually used for social engineering attacks, which can then be used by the aspiring criminal, in cooperation with a social network supporting criminal actions, to become a criminal. This is, however, not anything unusual for this particular field, but a dilemma shared with much of informa-

tion security research. It is better that we all know and understand the flaws, rather than only the attackers.

A description about what makes humans susceptible to influence from others is given in the Book Chapter of part II below. There is a wealth of information on these subjects, although most often not from an information security perspective, but on marketing or other areas instead. Prominent authors in the field are Cialdini (2001) and Levine (2003). These authors, and many others, write about social psychological aspects of deception and the ways we influence each other. From the literature on deception in general, we learn of several ways deception can occur. In the Book Chapter of part II, we use that knowledge in an information security and social engineering setting, by describing the deceptive methods used in typical social engineering attacks. Finding research in the area with an information security setting is rare, but in one paper (Jordan & Goudey, 2005), an interesting taxonomy of twelve categories of social psychological vulnerabilities is revealed. This taxonomy is used to describe a selection of current attacks by malicious code and the social engineering areas they exploit. With regard to more contemporary sources, we discover Harl's (1997) influential presentation that describes how social engineering can be used by an attacker.

With regard to deception, and techniques that educate about deception, there is a surprisingly large amount of literature that seems to be unknown to most researchers in the field of social engineering. Grazioli (2004) writes about different theories that describe deception, and focuses on the "Theory of Deception", ToD.

> "the Theory of Deception describes the information processing involved in both deceiving and detecting deception, […] the Theory of Deception states that individuals detect deception by noticing and interpreting anomalies in their environment in light of the goals and capability for action that they ascribe to others with whom they interact. The interpretation process is triggered when individuals notice inconsistencies between their experience and their expectations about their experience."*(Grazioli, 2004, p. 151).*

This theory is interesting, as are other, conflicting theories of deception, such as the Interpersonal Deception Theory, IDT. The difference between the two theories, according to Grazioli (2004), is that ToD would be more suitable to use in a context with more personal contact (therefore social engineering), while IDT is more aimed at communication with little personal contact (therefore Phishing).

Conclusion on Understanding

We have found notably little research carried out on the human elements of information security. There is, however, interesting and relevant research in other fields than information security. Furthermore, there is also much potential for learning from psychology, social psychology and sociology.

Measuring

Growing up, this author was raised with his father's favorite quote from Lord Kelvin: "To measure is to know". Perhaps it can be argued that it is the only way to fully understand the impact, and the relevance, of the research area It is rather easy to measure, and thus to understand, the impact, for example, that the deployment of an anti-virus software has on an organization. One of the most obvious effects is, hopefully, the disappearance of viruses, and logs that probably display significant numbers of thwarted attacks, and updates, as well as successful recoveries carried out by the software. Something concrete has a value that is easy to grasp, and easier to market. With regard to humans, it is harder to measure, both the inefficiency and efficiency of security measures. Another fact that can be troubling is that while technical attacks tend to be on a large scale, for example, viruses or attacks against firewalls, the attacks aimed at humans are often on a smaller scale, with a greater focus on individuals. This makes collecting relevant statistics difficult, thus making it hard to fully grasp the scale of the problem.

Previous Work in the Research Field

One approach to measuring is sending out fake Spear Phishing e-mails to the organizations' own users. This has been done by both the State of New York (Bank, 2005), and the US military school, West Point (Dodge & Ferguson, 2006). In the West Point case, students were sent an e-mail from a person claiming to be a Colonel, ordering them to click on an attached link to verify their grades. This approach received 80 % compliance among the students. In the case of the State of New York, 15 % of the employees tried to enter their passwords into a special online "password checker" after receiving an e-mail from the "Office of Cyber Security and Critical Infrastructure Coordination", urging them to do so. This was after they had received educational materials on security. A follow-up study several months later, using a similar approach, received a lower compliance rate (8 %).

This approach is interesting, but it creates a new set of problems, both ethical and practical. There is a possibility that the trust between the organization and the employees can be affected, and there are also other ethical questions. Still, it may be a very efficient method, not only for diagnosing a level of insecurity, but also for educating the users. If they do submit information,

and are criticized for it, they may become inoculated against further, real attacks.

There have, of course, also been other academic studies on Phishing reviews. A highly publicized and interesting study was done by Jagatic, Johnson, Jakobsson and Menczer (2007) in which a highly specialized and targeted Phishing attack was attempted against university students. The experiment was a stunning success, if seen from the perspective of a potential attacker. The test using a classic Phishing attack was 16 % successful, but the more advanced attack was 72 % successful.

While the Phishing study by Jagatic, et al. (2007) in itself is highly interesting, the debate that followed with its highly vocal complaints and articles in the media criticizing the study after its publication, as well as the ethical and emotional dilemmas, are also interesting. This once again demonstrates the necessity of a strictly ethical approach while conducting these kinds of studies, as there is an inherent problem that with large scale deceiving of users they may feel violated by the test. Judging from the reactions in many of the cases where users have been deceived, there are often strong opinions against using these kinds of tests, based on the feelings of the individual subjects. The strong interest in conducting Phishing research in that particular study, and the experiences that were gained, led to the publication of an excellent paper on how to perform fraud experiments (Jakobsson, Finn, & Johnson 2008). However, the article was published late in our research process.

While large scale attacks using Phishing techniques to measure a level of insecurity are quite manageable because it does not take much longer to send 10,000 e-mails than it does to send 10, with social engineering attacks it is different. It is, obviously, not feasible to carry out a social engineering review on every single employee. One reason is that employees would probably notice if they all suddenly started to get friends who wanted them to reveal information. Another reason is the fact that it would take a tremendous amount of time for the penetration tester to properly social engineer a large number of people individually. The ethical complications would be even greater than with those for Phishing attacks, as a social engineer should try to develop a relationship with the mark, preferably over a long period of time. Therefore, large-scale social engineering reviews are probably unfeasible for most, if not all, organizations.

The ethical problems connected to social engineering reviews are also discussed at length in Hasle, Kristiansen, Kintel and Snekkenes (2005), and especially in Jakobsson, et al. (2008). It is important to consider that the subjects in these tests are humans and not machines. The impact of a review on the individual must be considered and minimized, and anonymity ensured. A novel proposal to avoid the dilemmas associated with reviewing individuals,

that are also discussed in length, is suggested by Vroom and von Solms (2004). They actually propose focusing on reviewing the organizational culture rather than the individuals. While this is an interesting approach in theory, we find it hard to develop a practical deployment using that approach for reviewing. Nevertheless, the discussion and arguments against individual reviewing are relevant and interesting.

In our early research, (Paper 1, Part II), we tried a slightly different approach to this problem. That study tried to test users' awareness and degree of susceptibility to common social engineering attacks, and if a quantitative approach to penetration testing of social engineering could be used. By conducting a quantitative study using the false cover of studying "micro efficiency", an organization with above average skilled users was surveyed on three classic social engineering cons. The results indicate that the approach could be useful as a part of, or a separate reviewing technique. The human element was not merely vulnerable, but vulnerable to the extent that it shadows most other security areas.

By using a web based study and false pretences, the people assessed (highly qualified IT-consultants) were asked a set of questions in a different context than security. The results can perhaps be significant with regard to which extent the organization is vulnerable to social engineering. This approach shares some of the dilemmas discussed with Bank (2005) above, but it is at least a practically feasible method of conducting social engineering reviews on a large scale in organizations.

The dilemmas associated with penetration testing and social engineering are also discussed by Barrett (2003), who concludes that it is preferable to use a review style which has results and objectives that are clear and can be accepted by both subjects and company. Furthermore, while Barret (2003) argues that reviews should not lead to discipline or dismissal for the individuals, nothing more concrete than that is discussed.

Another academic approach to social engineering review was taken by Hasle, et al. (2005), whose approach to social engineering penetration tried to test a larger population. They performed two tests. The first was a survey where the users were asked to submit their login information to authenticate if they had won a prize; the second test was an e-mail which triggered a login box. According to their findings, approximately one quarter of the users could be tricked into submitting their passwords. A more recent study (Bakhshi, Papadaki & Furnell, 2008) produced similar results; about one quarter of the users were easily deceived by a Phishing attack.

A traditional approach to social engineering reviewing is argued by Jones (2003), who advises the reviewer to actually conduct social engineering attacks on the users. A similar approach is used by Orgill, Romney, Bailey, Orgill

(2004) who actually have a person trying to manipulate his way to gaining information from the employees of the tested organization. This is done in two parts. The first is to let the person wander around submitting employees to a written questionnaire with questions on security, logins, and so on, and in the second part the person tries to gain physical access to the perimeters. Both approaches are disturbingly efficient; 81 % of the subjects asked gave their login names, and 59 % also revealed their passwords. Very few employees asked for identification or questioned the reviewer. The auditor also managed to obtain unrestricted, physical access to the building.

Dalrymple (2005) describes the highly successful internal review on social engineering conducted by the IRS, where a select number of users were called, under some pretext, and asked to reveal their passwords, which 35 % of the employees complied with.

The classic approach, as used by Ogrill, et al. (2004), definitely has its uses, but the flaws are that it is costly (since it takes a lot of time to perform), and the employees can perceive it as being more ethically questionable than a more indirect form of deceptive study. It is also possible that the subset of users who are tricked will not tell their colleagues about it. If everyone in the organization is told about the study, it is possible that the users who were not reviewed will then keep to the "lie detection" bias (Marett, Biros, Knode, 2004), feeling that they, themselves, would not fall for "tricks like that".

Information Systems Audit and Control Association, ISACA (2004) provides a list of areas that should be tested when doing a social engineering audit. They suggest the four areas to test are:

- Test of Controls – a general overview of the organization, can give basic knowledge usable in further tests.
- Telephone Access – to use a set of well known attacks to test the organizations' resistance to attacks over the telephone.
- Garbage Viewing – to see if there is any sensitive information being thrown away (dumpster diving).
- Desktop Review – Check the user's workplace. Merge the data from the social engineering audits with other audits.

The guidelines given by ISACA (2004) present a basis for testing that could be perceived as ethical, at least by the organization, but the attacks suggested and the general set-up seems, in our opinion, to provide little data that can actually be useful, and the approach, while nicely structured, is slightly shallow and incomplete.

A related study was done by Grazioli (2004) who studied the impact of deception on MBA students trying to evaluate whether to trust a web-site or

not. This study was mostly centered on deception, but proposed testing against deception cues in order to discover to what extent it was possible to influence the students by deceptive tactics. The findings were that as a group, the students were unable to discriminate between deceptive web pages and genuine ones. This approach could probably be adapted for testing if users are able to discriminate between genuine requests for help or assistance and malignant ones.

Conclusion on Measuring

One of the reasons for the focus on technical solutions to the security problem is perhaps that it is easy to see the benefits of using a product against a measurable threat. We believe that there is a need for similar methods of presenting the risks associated with humans. There are a couple of different approaches that can be used. The first is to select what one wants to test. If the test should be a broad approach covering a large number of subjects, a Phishing attack would be the most suitable. If fewer subjects should be tested somewhat more thoroughly, then social engineering would be better. Phishing is, in our opinion, basically using social engineering techniques against a wider audience, by using technical means, with less precision, but greater coverage. It is hard to choose the preferred number of the subject group. While a study on a smaller subset may give useful statistics, the fact is that it is still enough with just a single vulnerable employee for the organization to be vulnerable. Furthermore, the ethical implications of conducting extensive tests that attempt to deceive the employees can also be difficult to handle.

Protecting

Learning about, and measuring, a problem is interesting, but it is important to also try to find possible solutions to the problem. While the extent of the vulnerability of social engineering is not precisely known, and may never be conclusively proven, there are few arguments against its existence. Consequently, there is a need for protection against attacks on the human element of information security. Currently, the typical recommendation for protection is education, which, for example, Mitnick and Simon (2002) argue for.

Previous Work in the Research Field

While conducting a more general case study on the status of information security in the healthcare domain by interviewing persons responsible for information security, it was obvious that education in the field was lacking in most of the subjects' organizations. One organization had not provided any

security education in the last 10 years, and in none of them was there now an active education program for the users (Åhlfeldt & Nohlberg, 2005).

While education is an important tool to use, it is important not to lose focus on the psychological aspects of the field. A defense against social engineering attacks must take psychology and persuasion into account, and develop that in order to understand, and counter, the persuasive attack (Gragg, 2002).

In the Background Chapter of this thesis, the current state of the art techniques of social engineering are presented, including a thorough description of the interesting "A multi-layered defense against social engineering" by Gragg (2002). The chapter includes a description of the specific educational needs, as well as general guidelines from other sources. In addition, the chapter includes a presentation of protection against Phishing, which describes practical end user measures as well as somewhat more organizational aspects.

There is also an interesting Masters Thesis aimed at the area of education for protection against social engineering attacks, "Fighting Social Engineering - Increasing information security in organizations by combining scenario based learning and psychological factors of persuasion", by Hermansson and Ravne (2005). In their thesis, the authors test, with some success, scenario based learning on psychological factors of persuasion, and create a software prototype for this. Subsequently, this scenario method was assessed as being more efficient than using ordinary lectures.

While their approach is interesting, we believe there is a risk in focusing too closely on certain methods of manipulation, since the typical characteristic of the social engineer is the adaptability and the flexibility of the attack. Therefore, it is hard to know whether such a strict and controlled model for education would be successful in real life, even though it is successful in the evaluation done by Hermansson and Ravne (2005). Many measures have been used in attempts to prevent Phishing attacks, both novel ones, such as using cartons (Srikwan & Jakobsson, 2008), and variants on Phishing attacks followed by targeted education of the users who "fell" for them, referred to as embedded training (Kumaraguru, Rhee, Sheng, Hasan, Acquisti, Cranor, & Hong, 2007).

Thomson and von Solms (1998) present a novel set of guidelines for information security awareness training that could easily be used for education on social engineering. They actually apply the same manipulative techniques on their students that a social engineer uses on a target in order to educate more efficiently. In this way they gently try to persuade the students into changing their security behavior.

Once again, turning to the field of deception, a couple of interesting studies have been done on educating users in detecting deception. These studies

dealt with an interesting piece of software named Agent99, developed to train military personnel in detecting deception. This software uses a multimedia approach, and is, according to the studies, an efficient way of training users in detecting deception (Cao, Lin, Deokar, Burgoon, Crews, & Adkins (2004), Biros (2005)). Marett, et al. (2004) also evaluate deception training in a military context, and suggest the field should be studied further, as it is promising.

General Aspects of Security and Education

Lee and Harley (2002) provide insights into the problems associated with security education, and express some views that education is hopeless, because users do not want to be educated. With this in mind, security education must be "maintained as strongly and vigorously as the technological aspects of the wider policy" in order to be useful (Lee & Harley, 2002, p. 81). However, they do argue that while security education does work if done well enough, it cannot be relied on as a complete solution to the problem. In the experience of Lee and Harley (2002), education works best on lower-grade staff, such as secretaries and administrators, but often fails with engineers and managers.

In an article about education regarding security, Adams and Sasse (1999) provide a set of guidelines for how efficient security education should be performed. Their general idea is to inform and empower the users by guiding them into the right actions.

Conti, Ahamad and Stasko (2005) give another view on how to educate users, which is more aimed at security awareness. Their idea is to train users to:

- Be alert for manipulation.
- Be aware of their own personal weaknesses.
- Take maximum advantage of the abilities in the system to counter these weaknesses.

It is believed that this approach makes the users more protected and resistant to attacks.

Björck (2005) also has some suggestions on the optimal way to educate the users. One of the recommendations is to use examples of previous security breaches. However, it is important that the users understand the rationale behind the security rules, and that top managers act in accordance to the same rules as ordinary employees. One of the conclusions made by Björck (2005, p. 238) is that more focus should be put on information security education, as well as other informal areas of research, such as ethics, awareness and policies.

One novel approach to protecting against social engineering attacks is to educate the users in transactional analysis, and how it can be used to identify "attacker" and “victim” communication patterns. Transactional analysis is based on the works of Eric Berne, and can be used to analyze communication. It is based on every person having three “ego states”, Parent, Adult and Child. In any communication and at any time, one of these is dominant. In any kinds of communications with others, the ego state the communicators are performing in influences the outcome of the communication and reflects on the individuals (Berne, 1996). There is a set of common counterproductive social interactions, from which the most interesting one in this area is the third degree interaction, in which one, or both, of the communicators can get hurt. This can be used to analyze the language patterns of social engineering attacks, and perhaps to train employees to be more resilient to them.

Conclusion on Protection

In the area of security education, there is a lot of material, both with regard to traditional education, which requires more time and resources from the end user, and security awareness, something that is quite useful in this context. With regard to assessing what the best approach would be, it would probably be necessary to actually test the efficiency of the approaches, and in order to do that, a metric is needed, leading back to problems discussed in the Measuring section above.

The other methods of protection could also be tested once a metric is in place, but since some of them, like Graggs (2002) “A multi-layered defense against social engineering”, are costly and perhaps overly complicated, only a dedicated organization would be able to employ it.

The smaller, organizational changes that can be made to increase protection are perhaps best employed when educating the responsible personnel, making education the natural first step in building defenses.

Thesis Structure

This thesis consists of two parts. Part I is an introduction to the research area and the research questions, how the research has been conducted, as well as a summary of contributions and discussion of the work. The second part contains the published materials; one book chapter and six papers. These are in Part II. A description is found in Table 1.

Table 1. Structure of the thesis.

Part I	**Chapter 1** Introduction
	Chapter 2 Background
	Chapter 3 The Research Design
	Chapter 4 Social Engineering and Phishing
	Chapter 5 Research Results
	Chapter 6 Concluding Discussion
Part II	**Paper 1:** "Social Engineering Audits Using Anonymous Surveys – Conning the Users in Order to Know if They Can Be Conned" Published in *Proceedings of the 4th Security Conference*, Las Vegas, USA, March 2005. ISBN 0-9729562-5-5.
	Paper 2: "User-centered security applied to the development of a management information system." Published in *Information Management and Computer Security* vol. 15, issue 5. ISBN: 978-1-84663-696-7.
	Book Chapter 1: "Why Humans are the Weakest Link" Published in Gupta, M. and Sharman, R. *Social and Human Elements in Information Security: Emerging Trends and Countermeasures*, IGI Global, Hershey, PA, USA. ISBN: 978-1-60566-036-3.
	Paper 3: "The cycle of deception - a model of social engineering attacks, defenses and victims." Published in *Proceedings of the Second International Symposium on Human Aspects of Information Security and Assurance (HAISA 2008)*, Plymouth, UK, July 2008. ISBN: 978-1-84102-189-8.
	Paper 4: "Non-Invasive Social Engineering Penetration Testing in a Medical Environment." Published in *Proceedings of the 7th Security Conference*, Las Vegas, USA, June 2008. ISBN: 978-1-935160-01-4.
	Paper 5: "Measuring Readiness for Automated Social Engineering" Published in *Proceedings of the 7th Security Conference*, Las Vegas, USA, June 2008. ISBN: 978-1-935160-01-4.
	Paper 6: "Phishing with Gifts as Bait: Measurement and Analysis of Phishing Attacks within a University Environment" Submitted to the *International Journal of Information Security*.

Background

This chapter presents an overview of fundamental information security concepts significant to this thesis. This is useful knowledge, especially for the reader not well versed in security.

Information Security

In order to provide a basic framework for security in computing, a presentation of *basic terminology and concepts* follows, dealing with the classic views on what constitutes information security compared to other security views. There is also a brief presentation of typical groups of *perpetrators*.

Basic Terminology and Concepts

While the field of information security is a rapidly evolving field of research, the basic concepts do not tend to change as quickly. This section is useful for the reader with limited insight into security, since it describes some of the classic models of what constitutes security. It also provides an argument for our selection of security model for this thesis.

Every part of a system needs well-balanced security. It is only after all the parts of the system have reasonable protection that it can be said to be secure. Historically there have been several terms for security, such as computer security, network security and IT security. Nowadays, the most common term when discussing security in a broader context is information security, as the focus is on the information, which has the higher value, rather than the technology. There are three fundamental terms used when talking about information security (SIS, 2003):

- *Confidentiality*. Only those individuals who are entitled to access a resource can access it. Access may also include printing, and knowing that an object exists.
- *Integrity*. Only authorized individuals should have the possibility to modify the asset. This also includes writing, changing, status changing, deleting, and creating.

- *Availability*. Assets should be available to those who need them when they need them. If someone has access rights to a resource, that individual should be able to access it.

These three terms are often referred to as the "CIA-triad" which is the basis of all security modeling. There have been additions to these terms over the years, of which one of the most common is accountability. Organizations want to be able to audit what decisions a user has made, in such a way that the user cannot deny having made a decision (SIS, 2003).

A different perspective on the CIA-triad and its extensions is the Parkerian Hexad, as proposed by Parker (1998). It adds three new attributes to the CIA-triad:

- Possession or Control, which is for situations when the data might be encrypted, but the confidentiality has not been broken. An example of this is the loss of an encrypted USB-memory.
- Authenticity, which concerns the correct labeling of information and who is attributed to it.
- Utility deals with usefulness; for instance, if all the files are encrypted but the encryption key is lost, the data would not breach any of the other traits in the Parkerian Hexad, but the files would not be useful.

It is important that a secure system incorporates all these aspects, and that the aspects often, but not always, overlap.

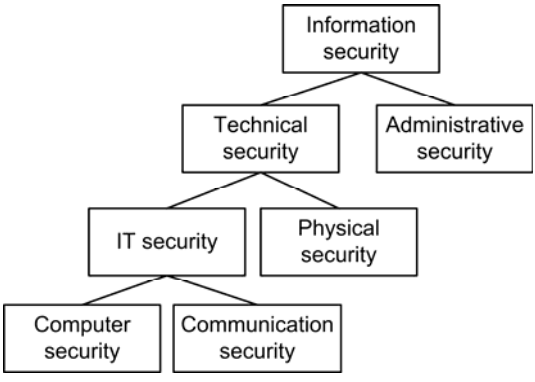


Figure 5. The SIS illustration of information security (SIS, 2003).

Another model for describing information security is the one used by the Swedish Standardization of Information Technology (SIS, 2003). This authority maintains that information security is the protection of information assets, which is achieved by maintaining secrecy, integrity, availability and accountability of information. SIS (2003) illustrates these terms in a hierarchical figure, where the terms are also classified in a ranking order, as illustrated in Figure 5. We argue that this model is suboptimal, especially in the

areas concerning “Administrative security”, which is poorly described. Considering the SIS model of security, our social engineering research would probably primarily belong to “Administrative security”, but would also need to address most other parts of the model. The boundaries are fuzzy as it is apparent that this model was not created with an intention to cover social engineering. The SIS model generally seems to be one more concerned with the hierarchical structure of an information security organization, rather than a useful description of the term information security. In fact, the efficiency of social engineering attacks against an organization that has closely modeled its security according to the SIS-model would probably be high, due the attacks falling “between the cracks” in the structure. One way of addressing this would be to introduce a new section in the model named “Cultural security”, which could include cultural, psychological and social aspects of security. This is, however, outside of the scope of this thesis.

A further problem with the SIS-model is the absence of the attacker in the model. One of the existing models of security that includes the role of the attacker is the ISO/IEC 15408-1:1999(E) Common Criteria model, as seen in Figure 6. Another positive aspect of the Common Criteria model is the inclusion of ownership. The model also includes such fundamental areas as risk, assets, threats and vulnerabilities. It is a very good model to describe information security as a whole, and also has the added benefit of being more useful, in our opinion, than the SIS-model described above.

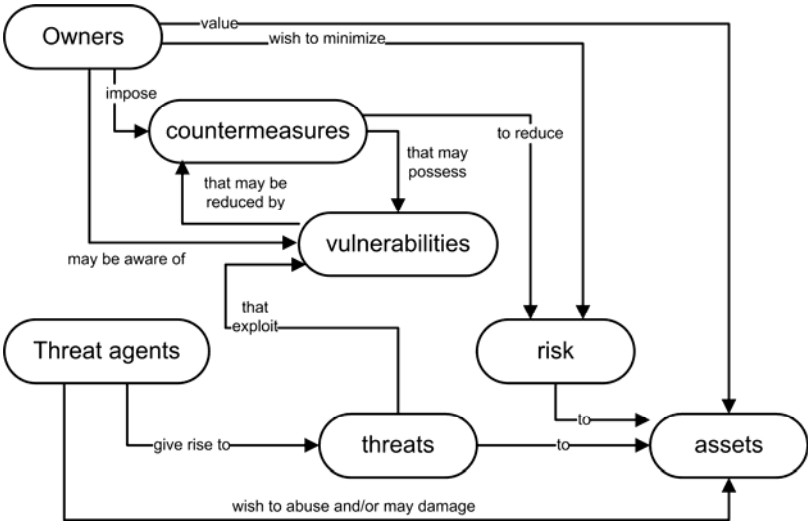


Figure 6. Security concepts and relationships, from ISO/IEC 15408-1:1999(E) (ISO/IEC, 1999).

The problem with using the Common Criteria model in this research is that it has a different focus than what we need. It provides a good general description of attacks and information security, but it is not well adapted to describ-

ing social engineering attacks due to the models broad scope. One example is that owners, in fact, are a risk themselves from a social engineering perspective, which is not described in the model, where they are reduced to simply trying to impose countermeasures. Nevertheless, the Common Criteria model was successfully used as a foundation during the creation of the conceptual models describing key concepts of social engineering that can be seen in this chapter below.

A security model that has a more holistic approach and is thus more useful in this research context is the SBC model proposed by Kowalski (1994), which provides a description of security that focuses on the perspective of the organization and assets, as seen in Figure 7.

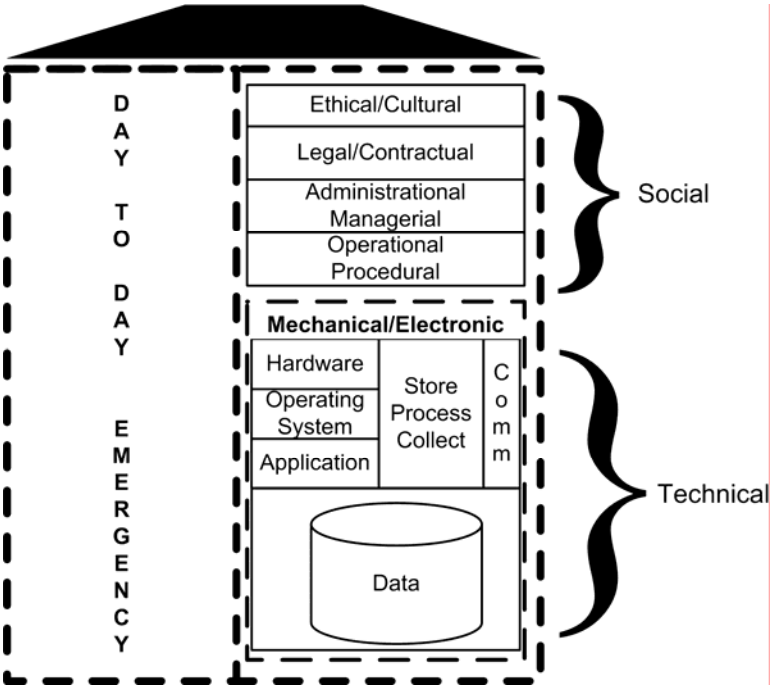


Figure 7. SBC Model, from Kowalski (1994, p. 19).

In the SBC (Security By Consensus) model a greater emphasis is put on a holistic approach, thus including the social aspects completely lacking in the SIS model above, as well as in the common criteria model. In the SBC model the owner or user of a system is perceived to create opportunities to become a victim by not protecting the systems they use or own. It is notable here that the perpetrators are not included in the model, due to the fact that collecting enough data on the perpetrators to enable a crime prevention program for IT crime is regarded as almost impossible (Kowalski, 1994).

The Systemic-Holistic Model (Yngström, 1996) was developed in order to address the problem of how to structure and present security knowledge at

an academic level, but to also have a use as a general description of security in security informatics. The Systemic-Holistic Model builds on the same framework as the SBC-model, but has a different scope. It is based on General Systems Theory, Cybernetics and General Living Systems Theory (Yngström, 1996). The idea is to be able to view a system both from small details and from a larger whole aspect with the same model; the overview can be seen in Figure 8. The framework is organized into level of abstraction (physical constructions, theories/models and designs and architecture), the context organization (geographical, space and time bound) and the content subject areas (technical and non-technical areas) (Yngström, 1996). The systemic module is the epistemological part of the model, the methodology of how to understand and use the framework.

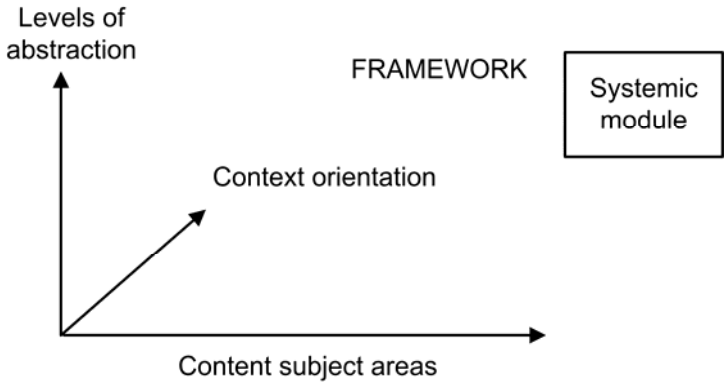


Figure 8. Overview of the Systemic-Holistic Model, from Yngström (1996, p. 19).

Looking at the details of the framework in Figure 9, we see the similarity to the SBC-model above.

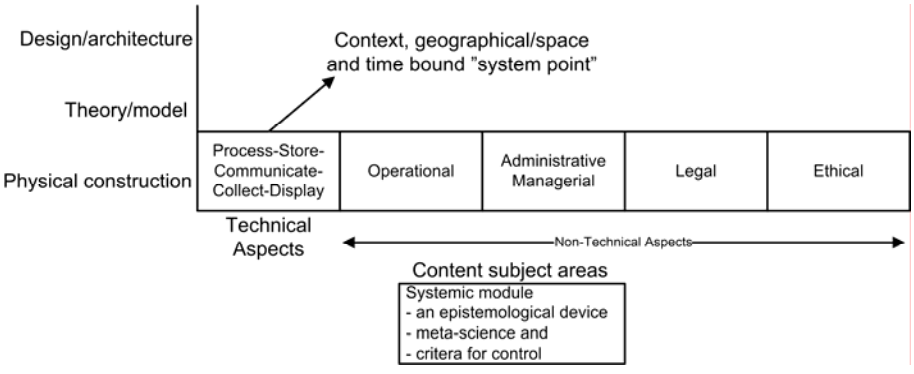


Figure 9. The Systemic-Holistic Model, details, from Yngström (1996, p. 20).

One of the differences between the SBC-model and the Systemic-Holistic Model is within the three levels of abstractions visible in the latter model. Each knowledge/subject area is supported by reality or a physical construc-

tion. The second dimension deals with the level of abstraction, that is, how detailed the view of the knowledge/subject area is. The third dimension, context, brings meaning to that particular subject area (Yngström, 1996).

The SBC-model and the Systemic-Holistic Model both not only illustrate that there is a need for a holistic and multidisciplinary approach to security, but also the importance of focusing on more than technical aspects of security.

Perpetrators

Most people have a very specific idea of who the computer criminals are. Accordingly, they are pale, socially awkward teenagers with high IQ's, low EQ's and a desire for destruction. They are extremely skilled at what they do, and their competence often surpasses even the most experienced professional. At least that is the way they are perceived in the movies, and, perhaps, how they were at the beginning of the computer era. However, the world has moved on. The motives for the early hackers were to gain access to computer resources, something that had a high value in those days. The goal of the next phase was the gathering of information, and the goal of the current phase is financial gain. This change of goals has also meant a change of perpetrators. Rogers (2000), updated in Wilson (2007), describes eight categories of hackers:

1. The Novice: Often referred to as script kiddies. Limited skills and often uses software developed by someone else.
2. The cyber punk: Young, often male, with higher skills. Often attacks high profile targets. No stranger to vandalism.
3. The Internal: Insiders who use their access either for financial gain or for revenge if they are disgruntled.
4. The Petty Thief: Perpetrators who start as regular thieves, but learn to use technology to increase their earning potential and lower the risks. Often not highly skilled in the beginning, but can acquire skills in the long run.
5. The Old Guard: Regards hacking as a challenge for the mind, and are quite curious. Often very skilled and often also lacking criminal intent. Will share their findings.
6. The Virus Writer: Mostly young males, often motivated by revenge or curiosity, but this is a group Rogers has yet to define.
7. The Professional Criminal: Highly trained, perhaps ex-intelligence operatives, use their skills for financial gain. Seldom caught, and work for organized, criminal groups.

8. The Information Warrior: Motivated by patriotism, they use their skills to disrupt an enemy country.

There are subgroups being developed, but these are the basic types of hackers.

In order to further explain the concept of the perpetrator in a social engineering context, we created the model in Figure 10. This conceptual model, as well as those in further sections of the thesis, was created as a single, large model that provides a conceptual overview of social engineering. The models were verified to the best knowledge of security experts to be valid. From this large model, sections have been used to illustrate key concepts in this thesis. These conceptual models are related in spirit to the ISO/IEC 15408-model that is illustrated in Figure 6, however, they have been adapted to the field of social engineering, as that model's scope of security is too broad. One example of change is that our term "mark", could be seen as a subclass of "owners", specifically the humans, who are affected in an attack. Correspondingly, we use the term "perpetrator" to refer to the "threat agents" that actually conduct the attack. In the same manner "threats" are referred to as "attacks" in our model, as our scope is far less general and instead specifically focused on social engineering attacks. The term "Assets" is the same in our conceptual models and in the ISO/IEC 15408-model. We chose to use the terminology common in social engineering rather than the more general terminology used within the ISO/IEC 15408-model.

One further difference is the level of detail. As our conceptual models are created specifically for social engineering, they cover far more details than those offered in the ISO/IEC 15408-model. For instance, the model in Figure 10 covers the Perpetrator in great detail, and also includes such aspects as the criminal organization and classic criminological traits. Our conceptual models do, however, chiefly match the flow of the ISO/IEC 15408-model, even if specific social engineering terms and concepts are used.

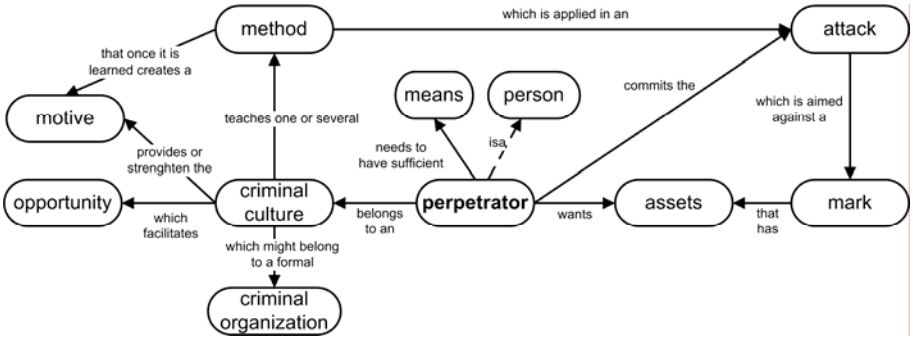


Figure 10. Conceptual model of the perpetrator.

An aspiring computer criminal must possess three qualities (Pfleeger & Pfleeger, 2003):

- *Method*. He or she must have the skills and tools and other necessary resources to perpetrate an attack.
- *Opportunity*. A perpetrator must have the time and the access to perform and succeed with an attack.
- *Motive*. There must be a reason for a perpetrator to perform an attack on the system.

If any one of these factors is not available to the criminal, the attack will never occur. The problem is that knowledge about systems and methods of attacks are easily obtainable, and since most systems today have Internet access, attackers often have an opportunity. Motives are diverse. Some carry out attacks to steal money, or specific data, while others do it for the challenge and the fun. There are also those that do it because of revenge (Pfleeger & Pfleeger, 2003).

The criminal culture, as discussed by Ferrell (1995), can be seen as the major factor determining crime. In fact, one of the flaws of traditional criminological reasoning is that the contemporary culture is sometimes neglected in the consideration of criminological analysis. The criminal subculture spans more than simply proximity, something that is available almost anywhere in a connected world, it also concerns motives, drives, rationalizations and attitudes, as well as certain appearances, group specific language and self presentation, and style (Ferrell, 1995).

In order to be able to perform attacks, the perpetrator must have knowledge about what kinds of attacks are possible, the method. Some attacks are obvious, and require no great cunning or planning, while others require certain skills or knowledge as well. There are three basic ways of acquiring this knowledge. The method may be known in advance, one can search for it with the specific intention of using it for attacks, or it may be found by chance. The perpetrator can discover an attack method that works well, on the first try, or a book or text describing attacks without having any prior intention of carrying out an attack. It is notable here that Sunderland's Differential Association Theory (DeMelo, 2007) states that once a potential Perpetrator learns the methods required, he or she can easily get the required motive from just about anyone. Thus, by learning the necessary methods, it is probable that the perpetrator will also pick up the motives needed.

While the potential motives for attacks are many, the majority of them seem to fit into certain categories. They are either carried out for financial gain, or information, or for revenge against someone or something. They may also be done just for the fun of it, entertainment, or for sabotage purposes. An in-

creasing trend in recent years is attacks that support certain values, a kind of political hacking.

When the perpetrator knows both how and why something should be attacked, an opportunity is needed. We have made the following distinctions among these:

Opening: An opening can either be known or assumed. This is a weakness or a specific opportunity that the perpetrator knows about. For instance, the fact that it can be assumed a credit card company knows about credit card numbers. However, the perpetrator can also have prior knowledge, either from his/her own experience or from others, about certain weaknesses. An example would be a perpetrator who in his/her previous career has learned about a certain weakness.

Random: A perpetrator can choose to carry out random attacks systematically, "trial & error", for instance, by calling a set of phone numbers while searching for a specific person. A perpetrator can also just conduct an attack based on a (mostly) sudden random impulse.

If a perpetrator has prior knowledge about an organization, perhaps from a previous attack (or knows someone who has), it is easier to use a "step n+1" attack, where prior knowledge is used. An organization may, for example, have a reputation of being easy to attack, or having things of great value.

When the perpetrator has all of the above, the only thing missing is the means. The first of these needed is the skill to carry out the attack. Skills can comprise influence techniques, programming knowledge, or simply the desired language. No matter how skilled the perpetrator is, if he/she does not know the language of the target, an attack is very difficult. Purely technical means are the communication channel used for the attack. The basic set includes the telephone, e-mail, www or other network communication, such as instant messaging. Uniforms that are used by the perpetrator also constitute means. In addition, time to perform and prepare is also necessary. All of these means can, however, be bought using the means of funding.

There is increased suspicion that many of the perpetrators have ties to organized crime, thereby making it easier for them to acquire means, method and opportunity (Hansell, 2004).

The Research Design

This chapter describes how the research was planned, conducted and which methods were used.

The Research Strategy

The aim of this research, as well as the previous work, is to try to cover the research area as completely as possible, while still maintaining a focus on information systems. It is easy to become lost in details not quite relevant to the research area. In order to maintain a focus during the process, certain delimitations must be made.

The intention with this research was not to base it on certain case studies or a single organization, but to try to achieve general knowledge, and to cover as wide an assortment of organizations as possible. This gives a broader understanding but also exposes the research to the vulnerability of being too broad to actually offer useful results, due to the variance of the studied organizations. This was addressed by a careful selection of the organizations that were studied, as well as contrasting research where possible.

The focus was on small to medium sized organizations. This focus is due to the different situations facing smaller organizations compared to major ones. It is also a delimitation which was made because of the problems associated with getting major organizations involved in these kinds of studies. This delimitation mainly affected the areas dealing with prevention and measuring, as the area about knowing remains basically the same regardless of the size of the organizations that are studied. We also tried to study major organizations to provide information and a broader understanding of the field.

There was little possibility of covering the whole field of the human element of information security, since this includes several research disciplines. It is unavoidable that there is a need to learn from other fields of research, such as sociology, psychology, and so on, while still maintaining the focus on information systems.

One problem is when can the area be said to be sufficiently studied to be able to draw any final conclusions. The easy answer is probably never, due to the complexity of the field, but by conducting studies that are sufficiently broad from several viewpoints, a valid contribution can be made.

Data Collection Techniques

In order to gather data from the real world, a selection of techniques was used. We mostly conducted interviews and surveys, but also observations and prototype testing. The interviews were all semi-structured (May, 2001) with a predetermined set of questions, but we also asked follow-up questions where and when needed. This is a common approach in qualitative studies. In most cases, the interviews were recorded digitally and then transcribed, but in some cases notes were taken during the interviews instead. There are advantages to both approaches. When taking notes during the interview, there is the risk of missing pieces of information for the coding process. In addition, it is hard for one interviewer to both ask questions and take notes at the same time. However, taking notes during the interview does reduce the workload associated with transcribing recorded interviews.

The surveys used in this study were both traditional ones, in which we asked people their opinions, and some that were conducted using false pretexts. This caused some ethical concerns, but in all the studies and tests the anonymity of the subjects has been a major consideration. The surveys were all done digitally over the Internet in order to create cost and time efficient studies that reached as many recipients as possible. The digital surveys conducted in this thesis have worked satisfactorily.

The observations were carried out during the development of the management information system. In this case we actually watched, and recorded, the subjects trying to use the interface in its various stages. Notes were taken in a task log and then used to improve the interface in the next step.

The penetration tests were mostly carried out either with interviews or surveys, but one test consisted of an actual penetration test using Phishing. When this test was created, the highest possible levels of anonymity were assured for the subjects, and any data they were tricked into submitting was guaranteed not to fall into the wrong hands. This is described with great detail in Paper 6, Part II.

Research Process

In Figure 11, a broad plan can be seen, mostly focused on the order in which the material was completed. The plan also provides an historical view of the order of completion of the previous material. The work has mainly been conducted in parallel, and knowledge has been collected in all areas throughout the project.

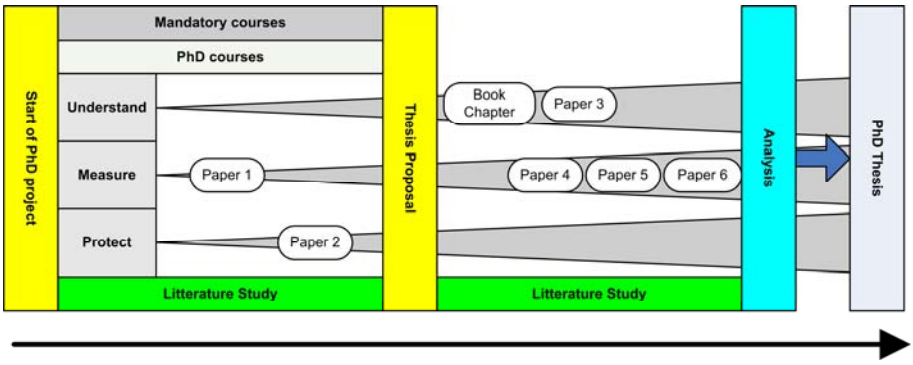


Figure 11. The research process.

This research was conducted in parallel with working at an information security company, The Logic Planet AB. The research process started with a number of mandatory courses required for the PhD, as well as several optional courses. These were within risk management, information security, law, ethics and criminology, and so on. All in all they provided a good background and understanding of the field of research. Together with input and considerations acquired at conferences, symposiums, workshops, and so on, as well as the early papers, this constitutes the background of the thesis.

A literature analysis not only consists of reviewing the literature, but also an evaluation of the knowledge gained from seminars, conferences, and so on, as well as the daily work as a security specialist. This has been carried out in parallel to the writing of papers.

Using the classification presented in Chapter 1.4, the planned and the completed papers have been added in accordance to the position we considered appropriate. Some minor position changes have been made in order to make them legible. Most of the research has been done in, or near, the informal research area with an empirical focus, as shown in Figure 12.

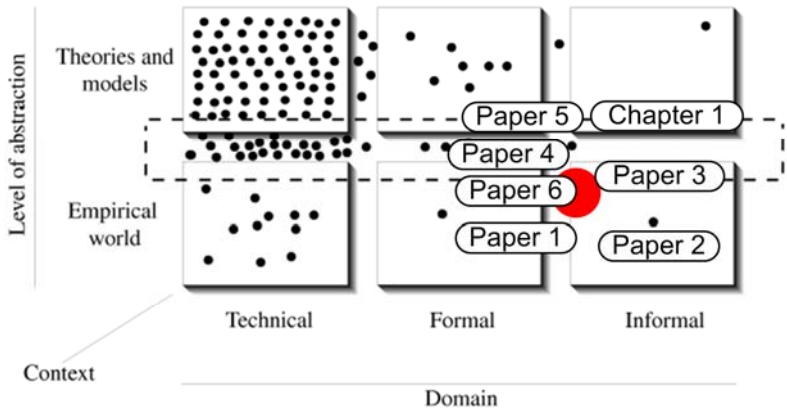


Figure 12. The classification of the 125 papers from the SEC 2000 proceedings (Björck and Yngström, 2001) with our contributions added.

When the contributions are matched against the three research objectives, we obtain an image showing the coverage of each, as can be seen in Figure 13.

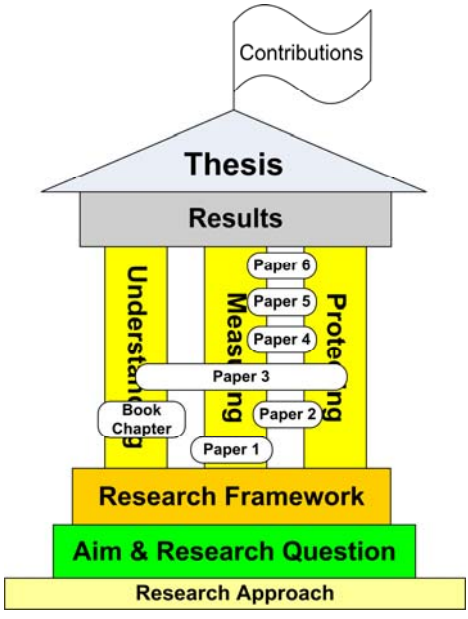


Figure 13. The contributions matched against the research objectives.

While some of the contributions are positioned in one single area, most of them actually cover two, or in one case, three areas. Exactly how the division has been made can be argued; the focus here has been on which area the main contribution is in.

As the cycle of deception is one of the major contributions of this work, some effort has been made to validate this further than what is presented in the included papers (Papers 3 and 5, Part II). Two different approaches were used. The first was interviews with three security experts, one of whom has a PhD in security, another is a researcher with twenty years experience in military information security, and the third is a researcher and teacher of information security. The subjects were selected due to their interest in the human aspects of information security. The interviews were conducted on the telephone, after the subjects had been given a document describing the cycle of deception. They were then asked to answer questions about the model: did they consider it to be useful, could they see any other uses for it, was the model complete, should some aspect not be included in the model, and could it be used to improve security in general. The complaints the three experts raised mostly concerned issues related to the presentation of the model (it was *unclear where it started*) as well as on *how it was connected to the "big picture"* of security. There were different views on the extent of the generality of the model; two of the subjects found it to *be useful for a far broader set of crimes*, while one found it *useful specifically for social engineering* due to the wording in it. The subjects all felt that it could be useful for *teaching about social engineering*, but also to *prepare defenses and to investigate attacks*. In general, the subjects found the model to be *useful*, *relevant* and that it *contributed to the overall knowledge in security*.

As the model was created with information security in mind and mainly security professionals have discussed it, we wanted a group of professionals with a good understanding of deception, but with little knowledge of information security to examine it. There are some professional groups that work extensively with other people in relationships where one part can benefit from deceiving the other. Examples of this include the police force, lawyers, medical doctors and social workers. We chose to study social workers since they work with both adults and younger persons, in different kinds of situations, and they have a good knowledge of psychology, criminology as well as social psychology from their education. The social workers that took part in the seminar are specialists in questions related to younger children and are regarded as some of the best in their field. The model was presented verbally for the group, consisting of four experienced and trained social workers. The model was presented to them and they were asked whether they felt it was useful and reasonable according to their experience. The general consensus was that the model did explain what happened when someone was trying to deceive them, but also that it gave a reasonable description according to their experience. They did note, however, that, in a short presentation, it was somewhat difficult to understand the model and that many of the deceptive attempts they are exposed to are of such a low degree of complexity that the attackers in those cases, mostly persons with substance abuse problems,

hardly spent time developing any kind of relationship but went directly to the attack. In some cases, involved parents may try to deceive a social worker in a manner in accordance to this model, for example, in custody conflicts. They also mentioned that this model could be the basis for education on how children can become victims to deceptions online.

Research Documentation

As an important part of the research process, papers were written and published in peer-reviewed conferences, journals or as book chapters. How each paper relates to the individual research areas is described in the research process above. The papers included in this thesis have all undergone minor changes and updates, mostly in order to conform to the thesis template.

The following six papers and one book chapter are included in the thesis. The papers are presented chronologically in the order that they were written.

Paper 1 **Social Engineering Audits Using Anonymous Surveys – Conning the Users in Order to Know if They Can Be Conned**

Marcus Nohlberg
In *Proceedings of the 4th Security Conference*, Las Vegas, USA, March 2005. ISBN 0-9729562-5-5.

Paper 2 **User-centered security applied to the development of a management information system.**

Marcus Nohlberg and Johannes Bäckström
In *Information Management and Computer Security* vol. 15, issue 5. ISBN: 978-1-84663-696-7

Book Chapter 1: Why Humans are the Weakest Link

Marcus Nohlberg
In Gupta, M. and Sharman, R. *Social and Human Elements in Information Security: Emerging Trends and Countermeasures*, IGI Global, Hershey, PA, USA. ISBN: 978-1-60566-036-3.

Paper 3 **The cycle of deception - a model of social engineering attacks, defenses and victims**

Marcus Nohlberg and Stewart Kowalski
In *Proceedings of the Second International Symposium on Human Aspects of Information Security and Assurance (HAI-*

SA 2008), Plymouth, UK, July 2008. ISBN: 978-1-84102-189-8.

Paper 4 **Non-Invasive Social Engineering Penetration Testing in a Medical Environment.**

Marcus Nohlberg, Stewart Kowalski and Kerstin Karlsson
In *Proceedings of the 7th Security Conference*, Las Vegas, USA, June 2008. ISBN: 978-1-935160-01-4.

Paper 5 **Measuring Readiness for Automated Social Engineering**

Marcus Nohlberg, Stewart Kowalski and Markus Huber
In *Proceedings of the 7th Security Conference*, Las Vegas, USA, June 2008. ISBN: 978-1-935160-01-4.

Paper 6 **Phishing with Gifts as Bait: Measurement and Analysis of Phishing Attacks within a University Environment**

Martin Boldt and Marcus Nohlberg
Submitted to the *International Journal of Information Security*.

Social Engineering and Phishing

This chapter contain detailed descriptions of “social engineering”, as well as Phishing, which is an important type of social engineering attack. The section on social psychology in Book Chapter 1 in Part II, is a recommended read, as it explains many of the reasons why we are susceptible to social engineering.

Social engineering is a technique in which an unauthorized person manages to pose as an insider or an authority to successfully obtain access to information or resources (Kajava & Siponen, 1997). A hacker can use social engineering to access other valuable data to benefit the hacker in further attacks (Hasle, et al. 2005). Mitnick provided our favorite definition in an interview by Tanneeru (2005):

> “Social engineering is using manipulation, influence and deception to get a person, a trusted insider within an organization, to comply with a request, and the request is usually to release information or to perform some sort of action item that benefits that attacker. It could be something as simple as talking over the telephone to something as complex as getting a target to visit a Web site, which exploits a technical flaw and allows the hacker to take over the computer.”

A social engineering attack focuses primarily on people’s vulnerability, and is based almost entirely on using “the principle of easiest penetration” (Pfleeger & Pfleeger, 2003). The greatest threat is that no matter how secure the system is in itself, it is never more secure than its users (Granger, 2001; Mitnick & Simon, 2002 etc.). Social engineering can be used instead of, or in combination with, threats and bribes. The classic social engineer aims at not leaving any traces, and generally leaving as little of an impression as possible, and thus threats and bribes are not favorite weapons of choice (Mitnick & Simon, 2002). However, foreign intelligence officers, for example, can still use them (Syrén & Malmström 2001).

Social engineering is used because it is often much easier to simply ask someone, a mark (the person being targeted by the perpetrator), for information, than to prepare and conduct a complicated software or hardware attack (Granger, 2001; Mitnick & Simon, 2002).

Models Describing Social Engineering

In the literature, most of the attacks are described using the typical attack cycle as presented in Figure 14.

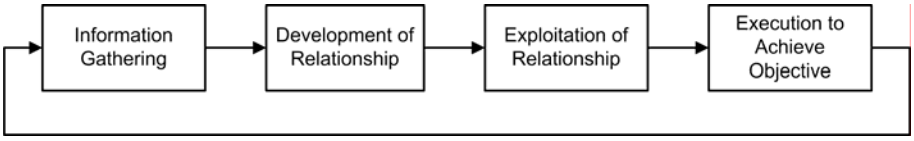


Figure 14. The Social Engineering Attack Cycle (Mitnick & Simon, 2002).

The description of this cycle is derived from Gartner (2002a). The first step is to gather information, for example, from public sources, such as phone books, web pages, and so on, or from other, previous social engineering attacks. This information will be used to develop a relationship with the target.

The second step is to develop a relationship by trying to create rapport and using the natural tendency of humans to be somewhat trusting and helpful.

The third step is to exploit the relationship by getting the target to reveal information, such as credit card numbers, passwords, secret information, and so on. This information can be the ultimate goal of the attack, or the starting point of the next stage.

The fourth step is the execution in which the attacker tries to achieve the end goal, or iterates into further cycles. It is possible that attacks consist of several cycles.

A Conceptual Model of the Social Engineering Attack

Another perspective is to look at a conceptual model of the typical attack as in Figure 15.

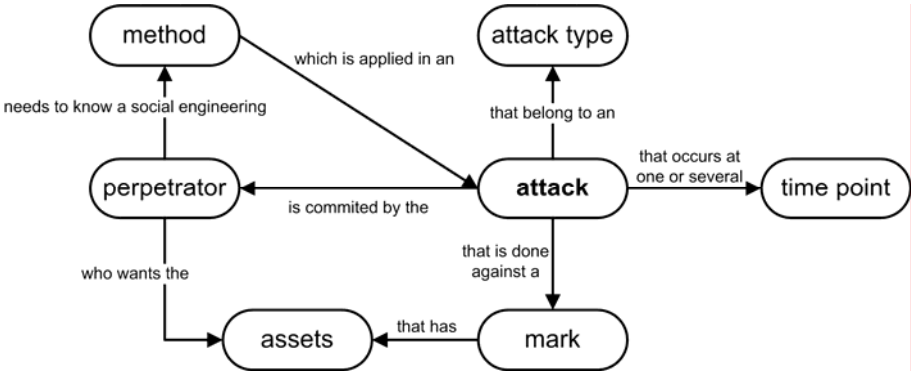


Figure 15. A Conceptual model of the social engineering attack.

The attack is performed by a perpetrator using a certain method and a specific type of attack against a mark. The aim is to acquire something of value possessed by the mark. Every attack has one or more time points, based on whether or not the attack can be carried out again, or if the deception is a long con with several points of contact. The different kinds of attacks are described in more detail on page 53.

The Cycle of Deception

One of the contributions of this thesis is an improved set of models, describing social engineering, named "the Cycle of Deception". Although this model and its origins are described in greater detail in Paper 3, Part II, a description of the model follows. Its aim is to improve the "social engineering attack cycle", as described above, making it a more useful model, both for professionals and those interested in learning more about social engineering. It consists of three different cycles; the attack, the victim, and the defender cycles, which are then merged into a model describing the complete cycle of deception.

The Attack Cycle

The attack cycle concerns the behavior of the attacker, and the actions he or she will take in an attack. In Figure 16, the stages of the attack cycle are described.

Figure 16. The Attack Cycle starts with Goal & Plan.

An attack must have a purpose, a goal, and a plan how to reach it. This is where traditional criminological knowledge becomes relevant. The four classic traits that an attacker must possess are method, motive, opportunity, and means (Pfleeger & Pfleeger, 2003). In order to be able to carry out an attack, the perpetrator must know what kinds of attacks and which methods are possible. Some methods are obvious and require no great cunning or planning, while others require certain skills or knowledge. There are three basic

ways to acquire this knowledge. A perpetrator might have prior knowledge about a method, it could be searched for specifically to use in an attack, or it might be found by chance. A perpetrator could discover an attack method that works well, on the first try, or he/she could chance on a book or text describing attacks without any prior intention of using such information. It is notable here that Sunderland's Differential Association Theory (DeMelo, 2007) states that once a potential perpetrator learns the methods required, he or she can easily pick up the required motive from just about anyone. Therefore, by learning the methods required it is probable that the perpetrator will also find the motives needed. The criminal culture, as discussed by Ferrell (1995), can be seen as the major factor determining crime. In fact, one of the flaws of traditional criminological reasoning is that the contemporary culture is sometimes neglected in the consideration of criminological analysis. The criminal subculture spans more than simply proximity, something that is ubiquitous in a connected world. It also concerns motives, drives, rationalizations and attitudes, as well as certain appearances, group specific language and self presentation, and style (Ferrell, 1995). *Map & Bond:* The stage in which the attacker tries to obtain information needed for the attack. This can be done by using traditional social engineering techniques, such as dumpster diving or desktop hacking, or by searching the web for data and studying other open sources of information. However, this information can also be obtained when the attacker befriends the victim or someone with usable knowledge, and uses manipulative techniques to get this person to divulge the information needed, or to "prepare" the victim for the next step. In order to create a deceptive relationship, the attacker uses influence techniques, for example, authority, scarcity, liking and similarity, reciprocation, commitment and consistency, social proof, and involvement (Cialdini, 1993). The influence techniques then exploit certain social psychological weaknesses, as suggested in the taxonomy put forth by Jordan and Goudey (2005). In other words, the victim is manipulated into trusting the attacker. *Execute:* During the execute-step, the attacker does something clearly illegal, or not allowed, for example, asks the target to submit his or her login information, or sends the nefarious e-mails. *Recruit & Cloak:* The term, cloak, refers to the actions performed to conceal the illegal activities used in the execution of the attack. Such actions can be to continue with the "friendship" to normalize the illegal activities, some kind of move that makes the victim seem untrustworthy, or more advanced techniques to conceal the crime. In some cases, the victim can be recruited to either work for the attacker or act as the perpetrator's ambassador/reference. *Evolve/Regress:* This is when the attacker learns from the process and creates an internal justification for his/her actions. At this stage, there are basically two choices for the attacker. If the process has been successful thus far, the attack evolves, moving into another phase, or, if the results have been unsuccessful thus far, the attack regresses, which means to

either stop the attack or return it to a more basic level in order to try again for success.

The Defense Cycle

The defense cycle describes the general options available to the defender, who could, in some cases, be the same person as the victim, or security professionals, or similar, in an organization. This section is based on the work of Kowalski (2002), which has provided the terms and definitions and identified the flow.

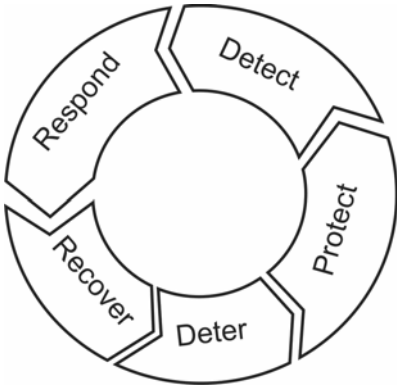


Figure 17. The Defense Cycle (adapted from Kowalski, 2002), which starts with Deter.

The description of defenses, given by Kowalski (2002), has been adapted to a circle to match the model in Figure 17. Several examples of implementations, which can, of course, consist of many other measures, are given below. The description is based on what the defender must do to successfully provide defenses. For example, having a good, public defense policy, or a reputation of reporting illegal incidents to the police, can *deter* an attacker. In addition, keeping the availability of sensitive data to a minimum, educating employees about the risks and methods of attackers who try to bond with them, as well as providing a strong policy on how to act, are measures that *protect* the organization. Furthermore, running a surveillance of the network communication can reveal when sensitive data are being sent or accessed, and having well-educated employees who know when they are being asked illicit questions, helps to *detect* an attack. Furthermore, making it easy to report social engineering incidents and not attaching any social or professional stigma to such an act, as well as making the employees aware of how they can be manipulated by an attacker, enables a defender to *respond* to an ongoing attack. Also, knowing the value of your data, reporting attacks and having a well-designed policy, means that a victim can *recover* from the attack and learn from it. Hopefully the attacker can be found and prevented from evolving and attacking you, or others, in the future.

The Victim Cycle

The victim cycle is focused on the behavior of the targeted person, the individual victim of the attack. A common mistake when analyzing crime is that too much focus is on the attacker, and not enough is on the victim. In fact, many crimes could be more readily prevented by focusing more on the victim than the attacker. The flow is described in Figure 18.

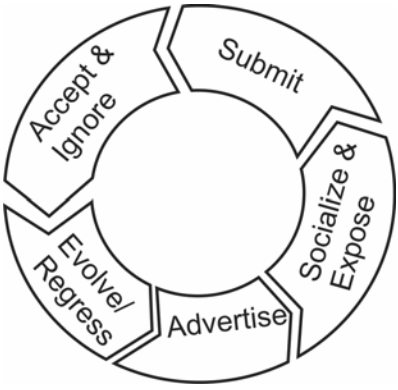


Figure 18. The Victim Cycle, which starts with Advertise.

By having something of value and making it known, either knowingly or unknowingly, the victim *advertises* its suitability as a target. Furthermore, by *socializing* with the criminal, the victim sets him/herself up for deception, and *exposing* valuables makes them accessible to the attacker. When the actual crime is being executed, the victim *submits* to it, for instance, by revealing the secret information. After the crime has been executed, the victim can choose to *accept* it, for example, through believing that it was not so "serious", or simply by *ignoring* it, either knowingly, or by actually being unaware of it. The victim can learn from the crime and *evolve* into someone who will be harder to victimize in the future. However, it is also possible that the victim can *regress*, becoming someone who accepts the role of victim and thus easier prey in the future.

The Cycle of a Social Engineering Attack

When the three different cycles are merged and a target in the center is added, a more holistic view of the prerequisites of a social engineering attack appears, as illustrated in Figure 19. One of our theories is that in order to achieve a "successful" social engineering attack, all the steps in all the cycles have to fall into place. The attacker must succeed with the first three steps for the attack to be successful, and with the fourth and fifth to be able to continue attacking in the future. This is based on the reasoning that if the attacker is unable to provide a plan and a method for the attack, it will most

likely fail. Furthermore, if the attacker cannot learn about the potential victim, or perform the attack, it will fail. In addition, if the attacker is unable to conceal the attack, he/she will most likely be caught, and, if the attacker, through internal rationalization, judges that the attack was not a "good" experience, he/she will most likely not continue.

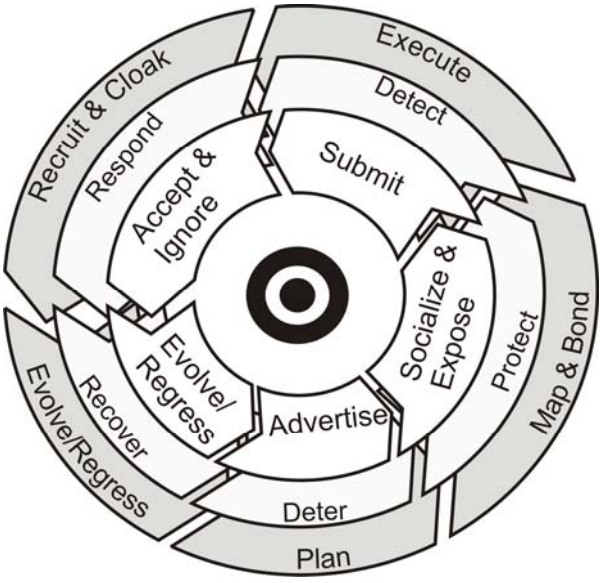


Figure 19. The Cycle of Deception, which starts with Advertise/Deter/Plan.

The same reasoning applies to the defender. If any one of the steps in the defense cycle is adequate enough to stop the attacker, then the attack will obviously fail or lead to the capture of the attacker. In contrast, if no single part of the cycle can stop the attacker, then the attack will not fail due to the activities of the defender. With regard to the victim cycle, we assume that the victim must submit in each of the sections of the model for the attack to succeed.

Potential Targets

Mitnick and Simon (2002) provide a list of typical targets (in social engineering often referred to as "marks") for social engineering attacks. They are:

- People who are unaware of the value of information, such as administrative assistants, receptionists, security guards, etc.
- People with special privileges, such as technical support, system administrators, etc.
- Manufacturer/vendor: Organizations that manufacture hardware, software, etc., which could be of interest for hackers.

- Specific departments. This could be accounting, human resources or other departments with potentially valuable information.

In general, typical marks are those who lack a certain insight into security, who work with helping others, have high access rights or specific knowledge, or who have access to something valuable, either information or economic value. This basically means that almost everyone with access to any part of the system is a potential target (Harl, 1997). The marks can also be organizations, but it is notable that in the actual attack, the target is always a human as social engineering is an exchange between humans.

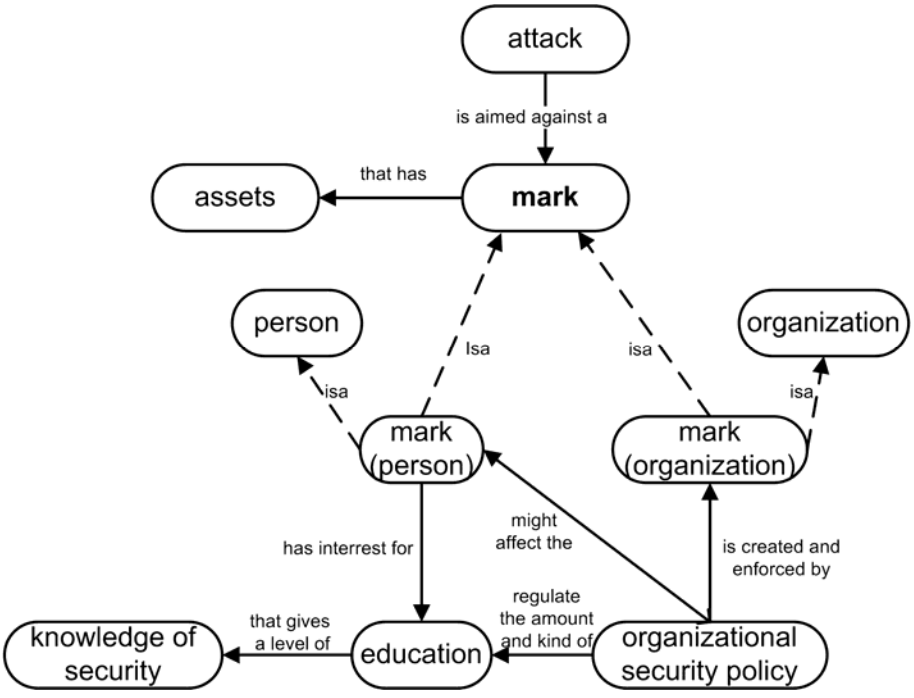


Figure 20. Conceptual model of the mark.

In our conceptual model of the mark, as seen in Figure 20, the mark has some kinds of knowledge that is used to counter the attack, gained either from personal interest, required education or previous experience. This knowledge can be sorted into two separate categories:

Every person has a certain security awareness. Within this awareness, we do not place specific security knowledge, but rather a level of suspicion, gullibility, caution, and an inherent paranoia we find, to some degree, in most people. The factors affecting awareness are discussed in greater detail by other researchers.

An understanding of the current security policy: Not every person is affected by a security policy. Some organizations do not have one, some individuals

are unaware of any policy, and others are simply not affected by any kind of policy. However, those individuals actually controlled by policies have some knowledge about them, and also follow them to an extent. While people cannot be expected to rigidly follow security polices, simply understanding them possibly provides a certain level of protection if the policy is well written. One of our preferred methods of describing organizational security work, and thus by extension the policy, and indeed the effects of security in general, from both an organizational, and an individual perspective, is the SBC-model used by Kowalski (1994). The areas covered in the SBC-model are those that should also be included in a security policy, and the aspects that should regulate the education of employees. Any such education should cover both general security and security awareness training.

To build further on security awareness, a person has a certain kind of knowledge of attacks based on human weaknesses. While such knowledge alone does not provide sufficient protection, it can be assumed that knowing about attacks and how they are performed leads to better protection than ignorance. The other kind of knowledge is centered on security information with regard to technical attacks, and includes traditional computer security areas such as firewalls, anti-virus programs, and so on.

Social Engineering Attacks

There is a vast selection of social engineering attacks and some of the classic examples are presented below. In addition, the whole picture is described using a conceptual model of a social engineering attack, as illustrated in Figure 21. In general, an attack can either be a short con or a long con. While a short con is an attack with a single contact between perpetrator and mark, a long con is an attack with several contacts between mark and perpetrator.

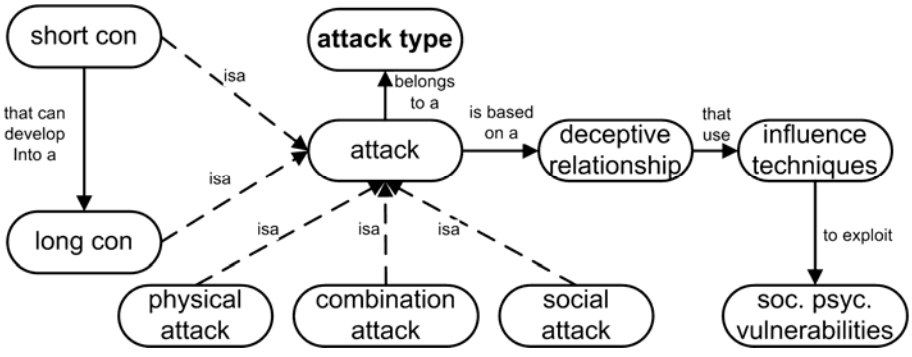


Figure 21. Conceptual model of the attacks that can be used in social engineering.

Each attack can belong to one of three different categories:

The physical attack is primarily carried out in the real world. Its intent is often essentially to gather information that cannot be found in other sources

such as online searches. This kind of attack is sometimes done in order to facilitate another, more advanced, attack. The favored methods of physical attacks are:

- *Dumpster Diving*: In which the perpetrator goes through the mark's garbage to find information (Granger (2002) and Gupta (2002)).
- *Theft*: The perpetrator can steal physical information, or computers containing the information. This is often not considered information theft, but computer theft instead.
- *Extortion*: This is carried out either by threats or actual violence against the victim or his/her loved ones.
- *Desktop Hacking* (Gupta, 2002): By looking at the office environment, valuable information can often be found. For example, passwords, and so on, might be written on post-it notes and cunningly hidden under the keyboard perhaps. Often the attacker can also see the password by "shoulder surfing", which is watching while the mark types the password.

The Social Attack primarily utilizes social techniques, but can also employ technical means. The main characteristic, however, is that the social attack uses deceptive relationships of some kind to be successful. In order to create a deceptive relationship, a perpetrator uses influence techniques, described in greater detail in Part II, Book Chapter.

The classic example of a social attack is to simply call the mark on the phone, explain that there is a problem with the network and ask the mark for some assistance in a series of complicated, technical corrective measures. When the mark finds that the process is too complicated, the perpetrator offers to help if the mark simply shares his or her login information. This is an example of a direct attack (Mitnick & Simon, 2002).

A social attack can also be based on a pretext relationship developed on deceptive terms between perpetrator and mark. Such a relationship can be based on, for example, romance, business or friendship. Furthermore, quite a lot of research is available on how to easily develop a strong relationship that can be exploited, which is described further in Part II, Book Chapter 1.

A more indirect approach involves the perpetrator getting the mark to ask the perpetrator for help. This is known as a reverse social engineering attack and consists of three phases (Granger, 2002 and Mitnick & Simon, 2002). The first is to sabotage the system and actually create some kind of disturbance. This can be of the simplest kind, for instance, unplugging a network cable, or more advanced, such as creating network interference through technical attacks. The next phase involves advertising, in which the perpetrator lets the mark know that the perpetrator can solve that kind of problem. This can be done by sending out e-mails, handing out business cards, using posters or

similar. Lastly, the solve-phase is when the perpetrator solves the problem, but not without attending to his/her own interests. For instance, the problem might be "solved" if the mark submits login information, or installs a specific piece of software, and so on.

Social attacks can also be carried out using technical means, such as e-mail or instant messaging. In such an attack, the perpetrator would, for example, aim to get the mark to click on a certain link, or install a piece of malicious software (Gulati, 2003).

A combination of technical and social attack methods is used, for example, when creating a Road Apple. This is when an attacker leaves a USB memory stick, or a CD with a tempting text (such as “Salaries 2008” or “my nude pictures”), outside a building, to entice a mark’s curiosity into using the item in his/her computer (Stasiukonis, 2006).

We consider that Phishing is a social rather than a technical attack as the most important part of the eventual success of a Phishing attack is the social aspect of it, the message and the context.

Protection against Social Engineering

The literature seems to agree on one thing; there is no “silver bullet” protection against social engineering. Education is the most commonly recommended means of protection, particularly if combined with a decent security policy (Hancock 1996; Mitnick & Simon, 2002; Gupta, 2002; Granger, 2002 etc.). Mitnick and Simon (2002) also provide several guidelines about what should be taught to the users regarding social engineering. These include what kinds of attacks can occur, how to detect them, and where to report them. There is also a lesson on not to trust everyone.

Hiner (2002) presents some clear guidelines with regard to education in protection against social engineering attacks, as illustrated in Figure 22:

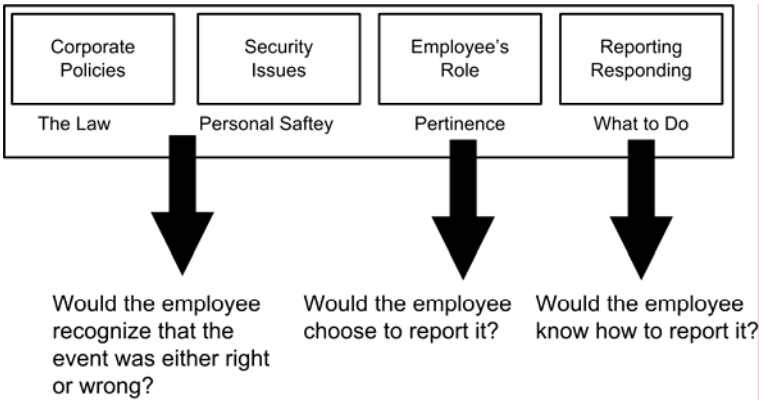


Figure 22. Education leads to defense (Hiner, 2002).

One should begin by thinking about how the employees in the organization would act “if an unfamiliar person who looked out of place sat down in a cubicle and started working on a computer.” (Hiner, 2002). Then one should consider these three questions:

- Would any of your employees become suspicious about this event?
- Would any employee choose to report it?
- Would any employee know how, and who it should be reported to?

If the answer to any question is no, then further education is needed using the organization’s security policy as a foundation. Figure 22 illustrates the different areas of the policy that should cover each of the questions above.

Other examples of important aspects that should be considered when building a defense against social engineering are (Hiner, 2002):

- Conduct background checks when hiring employees.
- Screen temporary and ancillary workers.
- Establish a clear reporting process for security problems.
- Open the lines of communication between physical security and the IT department.
- Monitor employee behavior patterns for abnormal activities and access violations.
- Lock out terminated employees immediately.
- Create a positive work environment, which will reduce the number of disgruntled employees.
- Publish a formally written, company security policy stating that the IT department will never ask for a user's password.
- Require ID badges for employees and mandate that an employee with a badge always accompanies visitors.

In Gartner (2002b), there is a collection of suggested protective approaches:

- Have clear, consistent, comprehensive and enforceable security policies.
- The single strongest defense against social engineering attacks is educated employees.
- Establish procedures that eliminate *any* exchange of passwords.
- Avoid using passwords or authentication questions that an attacker can easily discern with a little research.
- Security plans must be coordinated with physical/organizational security.

Gragg (2002) has a different approach to protection, and proposes “A multi-layered defense against social engineering”. Gragg argues the need for Social Engineering Land Mines, SELM, used together with a defense in several layers:

Foundational Level: Security Policy Addressing Social Engineering

The foundation of any security is a thorough security policy (Gragg, 2002). Such a clear policy strengthens the users’ resistance to social engineering, and, if the policy is strict enough, leaves users without any other option than to deny the Social Engineers’ requests. Another interesting point made by Gragg (2002) is that a strict security policy increases the users’ resistance to persuasion because they feel supported by the guidelines.

Parameter Level: Security Awareness Training for all Users

Gragg (2002) recommends training for all employees, using the security policy as a basis. The specific issues for protection against social engineering are:

- Know what has value
- Friends are not always friends
- Passwords are personal
- Uniforms are cheap

Fortress Level: Resistance Training for Key Personnel

Key personnel (those who work with helping others, especially external parts) should have more resistance training than other users. The two main points for key personnel are that they must be able to realize when someone is trying to manipulate them and that they are vulnerable to such manipulation (Gragg, 2002).

Persistence Level: Ongoing Reminders

Results from education and training do not last forever. Gragg (2002) recommends constant and creative reminders of the risks.

Gotcha Level: Social Engineering Land Mines (SELM)

A SELM is setup in the system to detect and stop social engineering attacks. These SELMs can be implemented to be used in several ways. Some examples by Gragg (2002) include:

- *The Justified Know It All.* This is an employee who has been given the task of questioning every person he or she does not know why they are on the floor. Besides questioning everyone, this employee should also have a decent knowledge of the security risks.
- *Call Backs by Policy.* Whenever any questionable request is made by phone, personnel should call back and check that the number they are calling belongs to someone with suitable authorization. If, for any reason, a call back is not possible, they should make a security log and be authorized to decline the demand.
- *Please Hold by Policy.* As a social engineer tends to use pressure, surprise or overloading to persuade the target. Therefore, users should be instructed to put any doubtful person on hold for a while, to give the user time to think, and perhaps to discuss the request with a manager or colleague.
- *Key Questions.* Gragg (2002) argues the need for a *three-question rule* (three questions that only the real employees could know, such as the name of certain pets) to use as a means of identification. These should be easy to remember, and available in a database for the personnel. Another means of control is the *bogus question* that implies false knowledge, which the real employee can correct, but the social engineer cannot.

Offensive Level: Incident Response

There must be a well-defined protocol to use as soon as a social engineering attack is discovered. This should be part of an incident response unit, which immediately informs the employees that an attack is in progress and what to expect. The unit also starts investigative work to identify the social engineer, and the target (Gragg, 2002).

While Gragg's (2002) "Multi-layered defense against Social Engineering" in theory provides quite extensive protection against social engineering, it is also a rather extensive commitment for the organization, especially considering the fact that many organizations do not provide any information at all about the risks of social engineering. In fact, the approach may be too expensive and complicated to be feasible.

Phishing

Before explaining Phishing further, it is important to examine the difference between Phishing and social engineering. In our opinion, the difference lies within the scope of the attacks, and the delivery. A social engineering attack targets a single, often specifically selected person (or organization), whereas a traditional Phishing attack employs techniques used by spam in order to

target thousands, or even millions, of users. Thereby, one could say that Phishing is social engineering using data-mining. The difference is, however, not always clear. In addition, more sophisticated and precise Phishing attacks, in general, and Automated Social Engineering, in particular, both serve to obscure the differences. In fact, one can argue that social engineering is an important part of most Phishing attacks, as they often, to some extent, focus on deceiving humans (Ollmann, 2004). A simple way to distinguish between them is that Phishing is simply social engineering in combination with the techniques from spam.

In this thesis Phishing is regarded as Jakobsson (2005, p. 3) regards it: "Phishing can be described as the marriage of technology and social engineering". We consider Phishing to be an attack mainly against the human element, and therefore a subset of social engineering.

What Phishing is

One of the organizations working against Phishing, the Anti-Phishing Working Group, defines Phishing as:

> "Phishing is a criminal mechanism employing both *social engineering* and *technical subterfuge* to steal consumers' personal identity data and financial account credentials. Social-engineering schemes use spoofed e-mails purporting to be from legitimate businesses and agencies to lead consumers to counterfeit websites designed to trick recipients into divulging financial data such as usernames and passwords." (Anti-Phishing Working Group, 2008).

Phishing typically uses less personal means than a telephone for message delivery, for example, e-mail or instant messages. Phishing is basically deceiving people into believing that someone of authority, and with legitimate reasons, needs their personal information, or that they must install a piece of software. The two primary goals are (Post- och telestyrelsen, 2007a):

- Acquire personal information.
- Get the user to install programs.

Phishing should not be confused with Pharming, which is a technique for misdirecting users to fraudulent sites or proxy servers, typically through DNS hijacking or poisoning (Anti-Phishing Working Group, 2008). As this is primarily a technical attack, it is not covered further in this thesis. Spear Phishing

Spear Phishing

Spear Phishing is a relatively new technique that does not use the wide attack patterns of Phishing, but instead sends highly targeted e-mails. The trick is to make the sender seem like someone the mark actually knows, or is associated with. A good method of gathering this data is data mining. While the goal of Phishing is to steal information from an individual, the goal of Spear Phishing is not only to steal an individual's data, but may also be to gain access to a specific organization's computer system (Microsoft, 2006).

This specific targeting makes Spear Phishing much more dangerous than ordinary Phishing, and professional attackers probably prefer to use it in order to get financial gains, trade secrets or even military information (O'Brien, 2005).

Spear Phishing could be regarded the "perfect" mix of social engineering and Phishing, and it also seems a lot more efficient and dangerous, than ordinary Phishing (O'Brien, 2005). It uses a higher degree of authority and the technique of the attackers pretending to be someone the mark is associated with. In order to strengthen spear Phishing even further, another context element can be added; that the victim anticipates receiving the message. This is described by Jakobsson (2005). The idea is to send a message that is not only from a person who can be expected to send such a message, but also in a context and time in which the recipient would anticipate receiving such a message. For example, sending a false e-mail from eBay directly after the user has placed a realistically winning bid on an auction site. This can be a devastatingly efficient attack. Similar attacks have been tried using data from social networks to create attacks emulating they are from people the victims know, something referred to as Social Phishing (Jagatic, et al. 2007).

Spy-Phishing

A "Spy-Phishing" attack consists of the attacker sending an e-mail, or a link, where the mark can download or execute a piece of software, which the installs itself on the marks computer, monitoring traffic until the mark visits a specific web-site. When the mark visits this site the software becomes active, and sends the login info etc. to the attacker. It is thus a combination of Spyware and Phishing that Trend Micro (2006) believes will be very common in the future.

Examples of Phishing Attacks

While it has always been a goal of computer criminals to acquire data and access to other resources, the name Phishing does not have a long history. The word is derived from the analogy of fishing for information by using e-mails as lures, and combined with the "classic" hacks "phreaking", using a

child's toy to get free access to telephone systems (Trend Micro, 2005). The term was first mentioned online in 1996, and in the media in 1997 (Ollmann, 2004). In the early days, the primary goal of the attacks was America Online accounts, which were then used to trade for other services, such as pirated software (Ollmann, 2004).

Most of the communication channels used over the Internet can be used for Phishing attacks, but the most common examples of Phishing attacks are those conducted by e-mail.

Dear eBay User,
During our regular update and verification of the accounts,
we couldn't verify your current information.
Either your information has changed or it is incomplete.
If the account information is not updated to current information
within 5 days then, your access to bid or buy on eBay will be suspended.
go to the link below,
and re-enter your account information.

Click here to update your account.
*****Please Do Not Reply To This E-Mail As You Will Not Receive A Response*****
Thank you
Accounts Management

Copyright©1995-2005 eBay Inc.

Figure 23. Phishing example – e-mail text (Anti-Phishing Working Group, 2005).

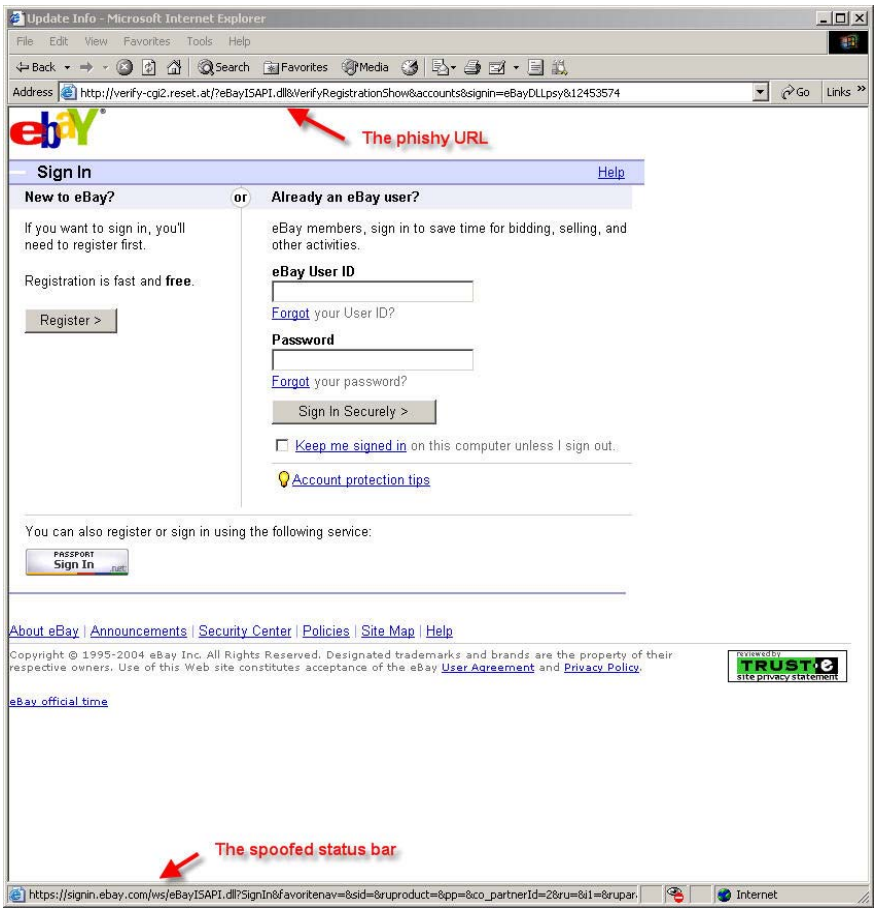


Figure 24. Phishing example – fake web page (Anti-Phishing Working Group, 2005).

An example of a Phishing e-mail sent to thousands of eBay customers (as well as other Internet users) can be seen in Figure 23. The attackers' goal is to get the receiver to click on the attached link, which will lead to an official looking but fake webpage (as seen in Figure 24). The mark can try to login using his/her login information, and thus, unknowingly, submit the login information to the attackers.

This example uses several of the techniques often employed within Phishing e-mails, a full list is (Ollmann, 2004, p.6):

- Emails that appear and sound official
- Copies of legitimate corporate emails with minor URL changes
- HTML based emails used to obscure target URL information
- Standard virus/worm attachments to emails

- A plethora of anti-spam detection inclusions
- Crafting of “personalized” or unique email messages
- Fake postings to popular message boards and mailing lists
- Use of fake “Mail From:” addresses and open mail relays to disguise the source of the email

Web-based Delivery

By using malicious web-site content, an attacker can perform a Phishing attack against an unknowing mark. This can be carried out either on a web-site run by the attacker, or by embedding code on a third-party site (Ollmann, 2004). The techniques for this described by Ollmann (2004, p. 7) are:

- The inclusion of HTML disguised links within popular web-sites, message boards.
- The use of third-party supplied, or fake, banner advertising graphics to lure customers to the Phishers’ web-site.
- The use of web-bugs (hidden items within the page – such as a zero-sized graphic) to track a potential customer in preparation for a Phishing attack.
- The use of pop-up or frameless windows to disguise the true source of the Phishers’ message.
- Embedding malicious content within the viewable web page that exploits a known vulnerability within the customer’s web browser software and installs software of the Phishers’ choice (e.g. key-loggers, screen-grabbers, back-doors and other Trojan horse programs).
- Abuse of trust relationships within the customer’s web-browser configuration to make use of site-authorized scriptable components or data storage areas.

Other examples of attacks using web pages include using fake banner advertising and obscuring the mark’s destination after clicking on the banners (Ollmann, 2004).

Instant Messaging and IRC

As many new clients for Instant Messaging, IM, and IRC, allow for dynamic content, they are likely to be used in much the same way as e-mail is today (Ollmann, 2004). The trend is that IM will be attacked more frequently in the future (Symantec, 2006).

Defense against Phishing

One of the most important tools for strengthening the defenses against Phishing is education (Ollmann, 2004).

There is a lot of focus on informing the users on the proper behavior to avoid being tricked by Phishing techniques. In general, the advice can be summarized into five separate points to remember (derived from Microsoft (2005a), Post- och telestyrelsen (2006b), FraudWatch International (2006), Ollmann (2004) etc.):

- Never reveal sensitive information in an e-mail or Instant Message.
- Be wary of clicking on links in messages.
- Check whether the webpage is genuine or not, and that the information you submit is protected.
- Keep an eye on your account balance.
- Keep your computer updated and use a firewall and anti-virus software.

Even if these tips seem simple, it is apparent that security education has been less successful with regard to providing protection against Phishing. We see examples of problems with traditional education methods (Kumaraguru, et al. 2007) that recommend simulated attacks carried out in combination with targeted education afterwards, aimed at those who need it. This might be complicated by suggestions that if users are trained to look for certain indicators of secure communications, they may more easily fall for attacks spoofing such indicators (Downs, Holbrook, & Cranor, 2006). There are indications that better general knowledge about how the Internet and web works can be efficient in providing resistance to Phishing attacks (Downs, Holbrook, & Cranor, 2007). Other novel examples of Phishing education include the use of comics, as described by Srikwan and Jakobsson (2008). There are also more technical approaches for protection against Phishing attacks. The most obvious approach is using anti-Phishing software, which somehow informs the user whether he/she is at risk or not. Some of the newer web browsers have this built in, and there are software programs that can be downloaded for protection against Phishing.

At an organizational level, as well as a systems administrative level, there are a lot of strategic decisions that can be made to improve protection against Phishing, as discussed by Ollmann (2004). In general, they deal with building an infrastructure that does not lend itself to being vulnerable to Phishing attacks, for example, by employing encryption, digitally signed e-mails, using strong, token-based authentication, monitoring the system, as well as strict host and linking conventions.

As these are mostly technical solutions, they are not covered further in this thesis.

Impact of Phishing and New Threats

Stolen data can have many uses. Credit card information can be used to purchase goods and services, ATM card information can be used to duplicate ATM cards for the withdrawal of cash. Account information can be used to steal information or for acting as another user online (Trend Micro, 2006).

Actually making reliable estimates on the extent of the success of Phishing is complicated, because many of the victims do not know they have been fleeced (Hansell, 2004). Calculating the real costs is also complicated, as it is not well known how successful the attacks are. Reasonably reliable sources mention costs in the area of $1.2 Billion a year in the US alone, and that 57 million Americans had received these fraudulent e-mails in 2003 (Gartner, 2004).

Historic findings indicate a continued increase in Phishing activity. Symantec (2006) reports a rise from 5.70 million daily Phishing attempts in the first half of 2005 to 7.92 million daily attempts in the last half of 2005. The Anti-Phishing Working Group (2008) reports an increase from 8829 unique submitted Phishing reports in December 2004, to 15244 reports in December 2005, and in January 2008 there were 29284 reports.

Symantec (2006) also expects a future increase in Phishing attacks, as well as an increase in Instant Messaging Phishing attacks. Trend Micro (2006) warns for the future increase in ever more sophisticated and targeted Phishing attacks. The general trend seems to be that for computer criminals the motivation for the attacks is no longer fun, or bragging rights, but instead, economic criminals are carrying out attacks for financial gain (Trend Micro, 2006). There are also suggestions that organized crime is behind Phishing attacks (Hansell, 2004, Jackson Higgins, 2008).

Why Social Engineering and Phishing Works

There are a number of psychological issues that can be used to create the perfect pretext for an attack. These are described in more detail in Part II, Book Chapter 1. In that chapter, sufficient examples are given on manipulative techniques in general, and also on how those techniques can be used by attackers in a social engineering context.

Research Results

This section describes and elaborates on the results gathered during the research process. In it we combine what has been gained from the individual papers into a greater picture. The results are ordered into the three pillars as illustrated above, but the big picture is also described by synthesized results.

Understanding

Early on in the research process it was not certain that the human element of security in general and social engineering in particular was such a major problem. In the first paper, (Åhlfelt & Nohlberg, 2005) we conducted a case study on systems administrators in health care and it was apparent that the human element was a major problem. After studying the area further, it was apparent that humans share a lot of common weaknesses. These weaknesses have been rather well known in other areas of research, mostly in those dealing with social psychology and in marketing. Our results in this area are the concrete description of how these common weaknesses can be, and are, used in a social engineering context. This is described in depth in the Book Chapter in Part II. We have also used several of these techniques when planning, and indeed when realizing attacks in other research projects, for example, the highly successful attack described in Paper 6, Part II. For the design of that attack we used the findings from the Book Chapter, Part II, in order to create a successful attack.

In order to more fully describe how a social engineering attack works, the cycle of deception was created, as described in Paper 3, Part II and shown in Figure 25. It describes the actions, or inactions, of the victim, the defender and the attacker in a social engineering attack. The cycle of deception can be, and has been, used as a teaching tool to help explain the somewhat complex iterative nature of the social engineering attack, and to give a more thorough view of those involved in the deception.

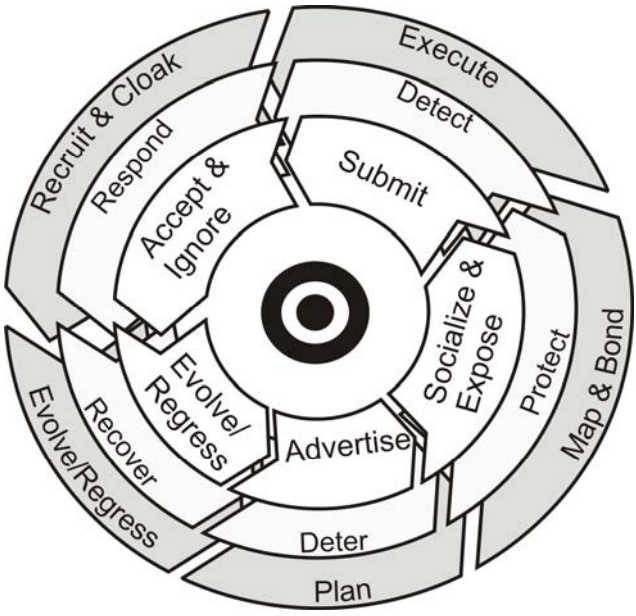


Figure 25. The Cycle of Deception.

Measuring

The first problem associated with measuring susceptibility to social engineering attacks is defining what constitutes a good way of assessing social engineering. We have tried several approaches, some traditional and others more novel, each of which has its own advantages and disadvantages. In order to highlight their relative strengths, we have created comparison matrices that are illustrated in Tables 2 and 3.

While there will probably never be a silver bullet approach to measuring social engineering that fits all scenarios, these matrices serve as a starting point when choosing a method of measuring that suits the needs of the organization best.

The matrices have been kept as simple as possible, and the methods are ranked from 1 – 5, of which 5 is the most suitable approach for addressing a particular problem area. The rankings have been chosen from our own experiences as users of the methods and what we have learned from reading about others using similar methods.

The first matrix, Table 2, concerns the results gained from an organizational perspective, using the classic values of Behavior, Attitude, and Knowledge, used by many security researchers, for instance, Kruger and Kearney (2006). The results are ranked on the basis of the information obtained about the organization as a whole. The focus on the whole organization is the reason why “Phishing Users”, for example, is given a higher grade than “Actual

S.E.", which is simply because it covers far more users than it could feasibly cover with "Actual S.E." This can be seen in Table 2.

Table 2. Organizational perspectives on measuring.

		Areas measured		
Measure using	**Read more**	**Behavior**	**Attitude**	**Knowledge**
Phishing Users	Paper 6	**5**	1	1
Deceptive Survey-ing	Paper 1	4	**5**	**5**
Asking users	Paper 4	3	4	4
Asking managers	Paper 5	1	3	3
Actual S.E.	Jones (2003) etc.	2	2	2

It may be difficult to learn about the attitude of the users when deceiving them, but learning about it is easier when asking them questions. As the "Deceptive Surveying" covers more employees, it is given a higher score than "Asking Users", even if the interviews used in "Asking Users" result in more qualitative knowledge. The problem associated with both "Phishing Users" and "Actual S.E." is that they do not provide much information about the attitude; instead they give excellent information about the behavior of the users. The same is true with regard to the knowledge area, where the same constraints apply as with the attitude area.

The second matrix in Table 3 deals with what we consider to be other important areas when selecting a testing method. "User Coverage" concerns the number of users typically included in a test. "Individual Coverage" deals with how well each tested individual's knowledge is covered. "Expense" concerns the actual cost of conducting a test based on the method chosen. This is certainly a field in which there can, and will, be huge variances depending on the complexity of the measurement, number of people involved, and on how the count is done. For example, should the cost of lost productivity, while a survey is being answered, be included? Should it also include the loss of trust among employees? In this latter example, the focus should be on the credible cost associated with employing some external part to perform the study, based on the man-hours needed to devise and carry it out. This is simply due to the fact that other costs are too individual to calculate in a meaningful way.

The "Ethical Impact" aspect is considered to be the potential trust/distrust problems that can arise from an organization's own employees after the study. For instance, a highly deceptive study may infuriate the employees, making them distrustful of the organization in the future, while interviews may create a more positive and trusting environment. The "Learning Impact"

aspect concerns how much positive learning the study can generate in an organization. More learning is generated if more users are involved and affected by the study. However, other factors also influence the process, such as the wake-up call that a simulated attack can evoke in the employees. This can happen whether they are attacked themselves, or a trusted colleague is deceived and they are later informed about it.

Table 3. Other important aspects of measuring.

	Aspects of Measuring				
Measure using	**User Coverage**	**Individual Coverage**	**Expense**	**Ethical Impact**	**Learning Impact**
Phishing Users	4	1	4	2	4
Deceptive Surveying	**5**	2	**5**	3	1
Asking users	3	4	1	**5**	3
Asking managers	2	3	3	4	2
Actual S.E.	1	**5**	2	1	**5**

These studies indicate that trying to find the “best” method of measuring susceptibility to social engineering is not meaningful. However, the tables above can be of some guidance. The best method is the one that provides the information needed according to the constraints of the organization. For example, if the budget is a major constraint, a deceptive survey can be devised for quite a low cost. If an actual test on as many users as possible is important, then Phishing is a good choice. However, if only a small group can be tested, or if security is highly crucial and the budget is of no major concern, actual social engineering testing is a good choice.

We believe there are good reasons for those active in penetration testing to try other methods than the small-scale actual social engineering tests that are quite common today. Trying alternative methods is easy and cheap, and the learning process among the users can be activated in novel ways.

Protecting

For many, the single most interesting part of this thesis is the section on protection, as there is always great interest in a silver bullet solution to any problem. At the beginning of this research process, the goal and hope was to find an efficient and simple solution to the problem of social engineering. As the research has progressed, it has become apparent that such a solution is unlikely. The human mind is a fragile and fickle thing. There are some weaknesses, that have been exploited for hundreds of years, which still work well today. While there are certain efficient and simple solutions to specific

kinds of social engineering, such as letting the suspected perpetrator of a telephone deception know that the call is being recorded, the risk associated with these specific preventive measures is that they are limited to a certain kind of attack. This is one of the fundamental problems with most of the protection advice other researchers have given. A summary of protective tips and tricks can be found on page 55 and page 64. An extensive method for regulating an organization in order to protect it against social engineering attacks is the "Multi-layered defense against social engineering" by Gragg (2002). It would probably provide increased protection for those organizations that have the commitment to introduce and enforce it. However, all of these tricks and tips share a common problem. If the preventive measures are generally known, the attackers can circumvent them. An example of this is teaching users to always enter the wrong user-ID and password at the first try if they suspect the site is a Phishing site. This would work very well for a short time, until the attackers became aware of it and started to ask everyone to enter their information twice. It would also provide the victims with a false sense of security; that by following a set of heuristics the attackers will not be able to trick them. This false sense of security may be one explanation for the high rate of susceptibility among the security experts presented in Paper 6, Part II. Thus, teaching the easy tricks may be dangerous. Consequently, we are not able to present a simple solution to all the problems of social engineering, but we can submit some generally good ways of improving protection against social engineering, as well as a set of future possibilities.

The first step towards protection against social engineering attacks is knowing what constitutes them, and we argue that the thesis provides a good starting point for this knowledge. Accordingly, it is probably more efficient to train users than simply to educate them. In order to train them, users must be exposed to attacks and make choices, good and bad, on their own, rather than passively take part in educational programs. One way of achieving this is constantly conducting various penetration tests, such as those discussed above. These will keep the users alert, unless they done to an excessive degree, in which case the users will grow weary of the tests.

One of the most important steps in protecting against social engineering attacks is involving the managers in general and upper level management in particular. This can be facilitated by talking to managers about security in the ways discussed in Nohlberg and Bäckström (2007). It is important to involve them, and a good way of doing this is to adapt a language suitable for that particular audience. An even better way is to use a management information system that provides up-to-date information on all aspects concerning information security adapted to the demands of the management users. Our trials with the prototype of such a system, presented in Paper 2, Part II, indicated great interest in that system. It is notable that many of the

components needed for a management information system are available as open resources, which means that development costs can be kept low.

Involving the managers and trying to test the users are the foundation of protection. Since the attackers are flexible and can adapt easily, it is dangerous to be set in fixed patterns. In addition, there is need for a more holistic approach, which can be found in using the cycle of deception as a starting point for preparing the defenses and understanding the readiness of the organization. This was tested in a study, presented in Paper 5, Part II, with good results. By understanding the consequences of the cycle of deception, non-standard defenses can be devised by each organization. In attempting to address the first three steps of the cycle, attacks can be avoided, and by working on the last two, it is possible to bring the attackers to justice and strengthen the organization in the future. It is also important to remember that security is, at best, regarded as a nuisance among the employees, and, at worst, they loathe it. It is a good idea to find ways to motivate the stakeholders, be they users or management. Examples of good motivation could be financial rewards, public acknowledgement, or, perhaps even, privileges. Users and managers will probably not suddenly be motivated to adhere to security guidelines unless they are given incentives.

Since there are almost always financial constraints to all security work, the cycle of deception can also be used to rank and prioritize security work, and as a tool with which to follow up improvements and measurements.

To summarize the concrete advice on protecting, the following is a set of crucial areas that should be covered:

- Learn about what social engineering is and what general risks you may face.
- Remember that attacks will constantly change and adapt, so there is a risk in training users on methods for spotting deception that are too specific.
- Involve managers and get their support and cooperation. Unless managers care about security and behave securely themselves, it is hard to get the ordinary users involved.
- Train the users and let them experience attack-like scenarios
- Consider, in a holistic way, the organization’s readiness to face social engineering threats.
- Acquire an overview of the organizational readiness in order to be able to address weak areas.

Concluding Discussion

This chapter contains the discussion part of the thesis, together with some ideas that might be good starting points for further research. Early in the research process, a mind-map was created containing the possible areas associated with the human element of security, as illustrated in Figure 1. That model has been updated with the areas covered in this thesis, as can be seen in Figure 26. The areas we covered are presented in bold text and rounded rectangles.

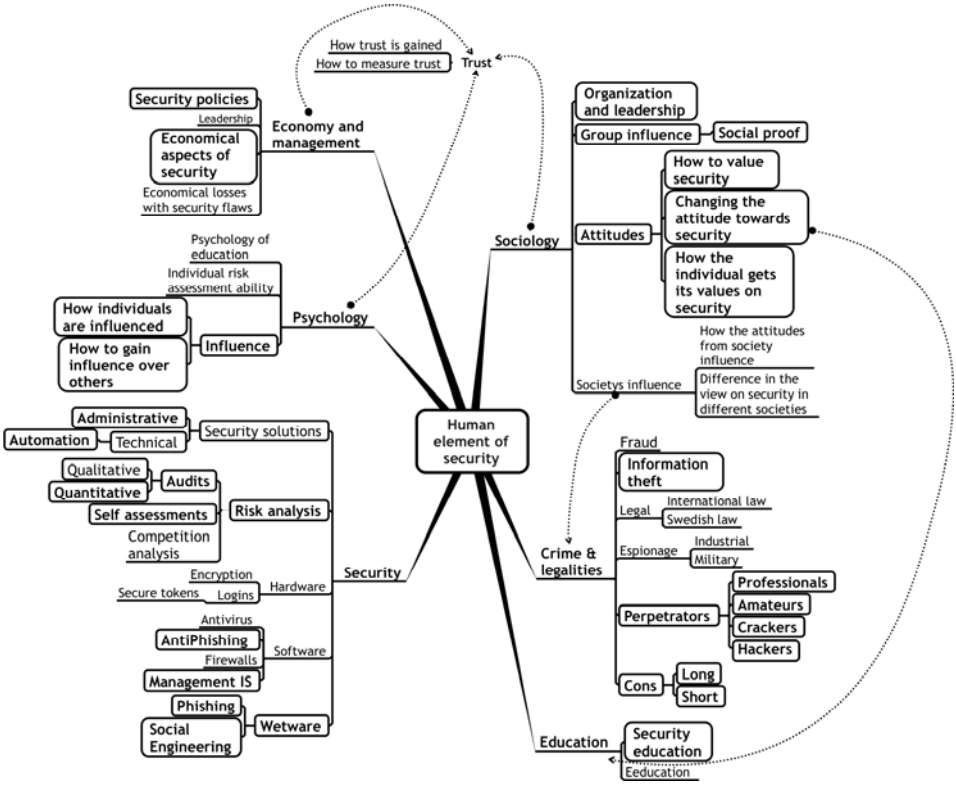


Figure 26. The human element of security mind-map revisited.

Many areas were not examined in depth in this thesis, but covering all aspects was never the aim. Instead, the intention was to use a broad approach to the area, which, we argue, has been achieved in this research. Although

other research areas have been visited, the foundation has been in information security. One of the obvious reasons for not covering all the areas was time constraints, but also that we had insufficient background information to include all aspects in an adequate way. For instance, legal aspects of social engineering are complex and require a deep understanding of legal research. We have attempted to confine ourselves to the areas in which our previous knowledge would be most beneficial.

In revisiting the section of security models, the SBC-model is used to describe which areas have been examined in this thesis, as Figure 27 illustrates.

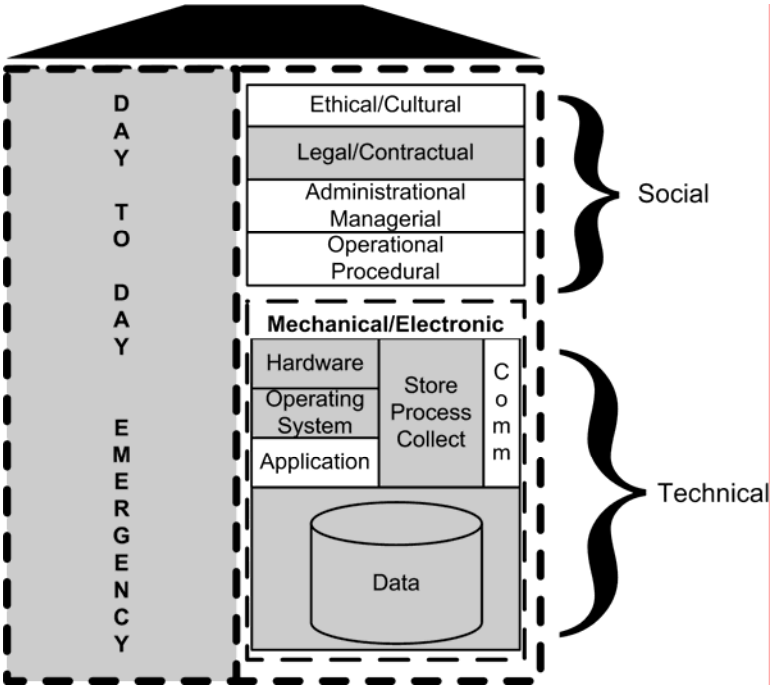


Figure 27. The SBC-model (Kowalski, 1994). The white areas have been studied in this thesis.

Most of our work in the thesis has been based in the social area, except for the Legal/Contractual aspect, which has been outside the scope of this thesis. Some technical areas, mainly “Communications”, have also been influenced because of the impact of Phishing attacks and the suggested future threat of Automated Social Engineering. The Applications area has been examined to some extent in our design of a management information system for information security.

Method

The three different areas studied in this thesis cover a broad and relevant section of the field referred to as social engineering. The study of these areas also necessitated the use of a selection of methods, ranging from qualitative to quantitative, as well as from surveys and actual penetration tests to interviews and observations. The decision to use these methods was made consciously because the studied area had not been previously examined to any great extent. It was therefore appropriate to try to study it from as many aspects as were relevant for this thesis. We do argue that using several methods gave good results, probably better than we would have obtained using only a single method. We have worked hard to ensure an acceptable level of ethics in our research, even though the research area itself deals with attackers not concerned with ethics. This has led to some constraints in the research, mostly on the kinds of studies that have been possible, but it has been manageable and not influenced the results in any major way. In a world without research ethics, it is possible that the knowledge levels in areas related to social engineering would be further advanced, but it would be a sad and dangerous world.

Results

The results gained from this research are, in some cases, not surprising, but it is good to have them confirmed. Other results indicate aspects that were quite unexpected. Early on, our hope was to be able to present a set of easy, rule of thumb guidelines that could prevent social engineering attacks. It seems that this was not only a naïve hope, but also something unfeasible. Instead, our results are spread over three different areas, as discussed above.

The first step to improve protection against social engineering is to find a structure for the work. The cycle of deception is an excellent starting point, as most parties can quickly understand it. The second step is to involve both managers and users. Managers can be involved by being informed in a way that suits their interests and needs, preferably by using a management information system. While education is often heralded as the best solution to the social engineering threat, one of the included papers indicates that even highly trained users can fall victim to social engineering, which in that case involved the Phishing technique. If users with several years security training and a personal interest in security can fall for an attack, it is unfeasible to assume that conventional education, which often spans a couple of hours or a day at most, would work. We argue that users should instead be trained by using simulated real attacks, such as Phishing, but they also need to be heard. It is wasteful, and perhaps even risky, not to listen to users. A good way of listening to them is having anonymous incident reporting. An even better way is also carrying out regular interviews with a selection of users,

where they are both heard and informed. This was done in a study presented in Paper 4, Part II, and the results indicate that such a soft approach has its benefits.

It is important to focus on the underlying mechanics of the potential attacks, and address those using the suggestions above. It would be easy for us to provide simple solutions for particular attacks, such as trying false logins and passwords on potential Phishing sites, since being able to login with a fake ID means that the site is a fake. While this is an excellent suggestion for a very specific attack, the perpetrators would certainly adapt their strategy by informing all users that the first login attempt was wrong and they should try again. Suddenly the highly successful defensive technique becomes a dangerous aid for the attacker. Furthermore, it would not be feasible to continue the arms race of "trying faulty logins", as it would be unlikely that users would attempt more than one false login. In the very nature of security work lies the problem of being a step behind the attackers. Consequently, in this research process, we have attempted to strengthen the knowledge base regarding social engineering that we rely on as security professionals and researchers. More specific examples of defenses can, and hopefully will, be devised from the framework suggested in this thesis.

Contributions

The results in this thesis are an improvement of the general knowledge of social engineering, in both the academic world and among practitioners. We have achieved this by merging knowledge from other disciplines, and by attempting novel approaches to both protection and social engineering reviewing.

The concrete contributions from this research are:

- A description of the socio-psychological factors that make humans susceptible to attacks, see Book Chapter in Part II. This chapter explains how humans can be, and are, manipulated in information security attacks. This chapter is especially useful for those with limited previous knowledge in either social psychology or information security.
- A model describing the cyclical patterns of attacks, defenses, as well as the victims' actions and inactions. This model can have several concrete uses. For example, it has been used as a teaching aid for educating in social engineering. It has also been used to obtain an overall understanding of defenses and preventive measures for conducting readiness studies. Furthermore, the model can be used to create automated social engineering AI-bots. There is also a possibility that it could be used to model and describe other crimes committed with intent, even if this is outside the scope of the thesis. The model is further described in Paper 3, Part II.

- A set of conceptual models describing certain key concepts of social engineering: the perpetrator, the mark, a general view of the attack, as well as the different types of social engineering attack. These are described in Chapter 2, Part I.
- Recommendations of the different approaches for conducting social engineering penetration testing. These were derived from several methods that were attempted in the real world during this thesis work. The methods used are presented in Papers 1, 4, 5 and 6 in Part II. This is further merged into a comparison on page 68, where recommendations for penetration tests based on different prerequisites can be found. These matrices should be useful for professionals working with audits and penetration tests, as well as for other academics in the field.
- Recommendations for protective measures, including aspects for protecting end users, as well as those concerning management information systems for information security, and details on how to inform managers about security issues. Papers 1, 3, 4, 5 and 6 in Part II concern user level protection. The papers concerning management information systems on security information, and the art of involving managers in information security are Paper 2 in Part II and Nohlberg and Bäckström (2007). These recommendations are useful for organizations that want to improve their protection, and for other academics studying penetration testing. The paper on the design of a management information system for information security contains several relevant and useful design heuristics that can be used for designing information systems in general, and those related to security in particular. If a software solution can be used, to some extent, to protect against social engineering, it would be ironic if it did not consider the humans using it. In order to achieve good usability for our interface, we developed a user-centered security concept that mainly focuses on usability, which is, in fact, an often overlooked but key security concept. If, for instance, the technical products that are employed in an organization are hard to use, new ways of attacking the users trying to understand the products will be created. It was therefore very relevant to present design heuristics that can help to design not only a security management information system, but perhaps more secure software in general.
- The concept of automated social engineering, ASE, has been coined and described within the context of this thesis work. Automated social engineering is one of the possible epic threats to global Internet use in general, and Web 2.0 in particular. It combines the massive numbers of attacks used, for instance, in common spam and the sophistication and targeting of social engineering. As current readiness for this appears to be low even in security conscious organizations, as described in Paper 5, Part II, it is

quite possible that ASE may be astonishingly successful and cause major problems if we do not create better awareness and readiness for this threat.

Scientific Quality

We use the criteria for qualitative research, suggested by Lincoln and Guba (1985), in the discussions of the scientific quality in this section. It is notable that some of the criteria are interwoven, so in certain cases the same arguments may apply.

Credibility

Credibility refers to the trustworthiness of the research, the information sources used, and that sources or references for the findings and conclusions drawn from the material used can be found (Lincoln & Guba, 1985). Our means of addressing this have been to allow the interviewees to read the interview transcriptions and make corrections and changes to any misinterpretations, thereby ensuring that we understood them correctly. The surveys have been studied with reasonable statistical tools. Where there are problems with the studies that may have influenced the results, this has been clearly indicated and discussed in the respective papers. Other related work has been identified, and where material from other sources has been used, it has been clearly referenced. The research process has also been described in detail, making it possible to see how the research has been conducted. Furthermore, the contributions of the papers and the book chapter have been published and subjected to a peer-review process. This means that the research conducted, its conclusions and findings have been exposed to critical reviews. The contributions have also been discussed with colleagues, both within academia and those working as practitioners in the security field. A further means of validating credibility is the actual, real world use of some of the findings, most notably the management system interface.

Conformability

Conformability concerns the extent to which the results cohere with and are supported by the data, that is, the objectivity of the report (Lincoln & Guba, 1985). We address this by documenting how all data have been collected and which techniques were used. In addition, we have documented our reasoning behind the conclusions drawn from the data gathered. As our studies examined different areas and used different approaches, we describe the complete picture and the conclusions that can be drawn when all the results of this thesis are combined. The received submissions to conferences, journals and books entail that our results have been examined and accepted.

Dependability

It is important that the findings are consistent and could be repeated in a similar setting, with similar subjects, or that the same conclusions could be drawn by other researchers using the same data (Lincoln & Guba, 1985). In order to provide dependability, we have documented our findings so that the data can be studied further. We have also described the methodological considerations, which questions were asked and how the studies have been conducted. This makes it possible to conduct the same studies in similar settings, and it is our belief the results would be comparable. By publishing our results, other researchers have had the possibility to examine our conclusions based on the data gathered, and in that manner the conclusions have been discussed. We thus claim to have achieved dependability.

Transferability

Transferability concerns the degree of applicability in other contexts of our research results (Lincoln & Guba, 1985). The research focus has been on human weaknesses in a security setting. These weaknesses are fundamentally the same no matter which organization a person is connected to; be they educational, health care, high-tech, medical or other. This is due to the fact that human weaknesses, as discussed in Book Chapter 1, Part II, are common traits among all humans, to some extent, even if certain cultural (both societal and organizational) aspects might influence which of the weaknesses is most easily exploited in a particular person. This creates a great deal of transferability among organizations. We have also used literature from other areas than information security, in order to strengthen our research and hopefully provide results that can be useful in other domains. We argue that our results possibly could also be transferred into other research areas, although they primarily relate to information security. The deception cycle could, for instance, perhaps be helpful in criminology, where it may be used, in a general way, as a basis for describing other kinds of frauds and crime with intent. We have tried to make it easier for others to decide whether or not this research suits their goals by describing our findings and data in a thorough manner that is even accessible to the reader without a background in information security.

Relevance

Relevance deals with having solved problems, current references, clear premises and a description of the thesis context (Lincoln & Guba, 1985). This is perhaps the most important aspect. Is this research relevant? We do believe that we advance the research in the field by employing a broader approach to security and contributing to a sometimes overlooked field of research within it. We also provide models that explain phenomena in the research field. The

contributions are described on page 76, but in general, the positive reception both in academia and among professionals indicates that this research is relevant and extends the body of knowledge.

Future Work

In this work several different perspectives on social engineering have been tried and studied. This has been beneficial, since academics have not studied the field to a great extent previously. The three different pillars of this research; Understanding, Measuring and Protecting, all served their purposes in providing the broad understanding that was the goal of this thesis. However, each area could alone be studied further by an aspiring researcher. There are some concrete examples of areas of study that we consider especially relevant.

- The cycle of deception is a tool that can be useful in many other areas of crime, and could be studied in relation to different crimes. For example, the deceptions carried out by prisoners to get compliance from their wardens, the manipulative techniques parents use against social workers in custody conflicts, and many other issues not related to information security. The model could also be used to create specific, social engineering training programs, and serve as a tool for creating social engineering artificial intelligence bots, as described below.
- The psychological and social-psychological aspects of security are still research areas in their infancy, and they have a wide selection of research opportunities with a great deal of potential. For instance, there are indications that certain people are more susceptible to deception than others. This is indicated in many studies on deception which reveal that about 30 % of the subjects are deceived. If we could identify those in an easy way and tailor specific training for them, the overall security would benefit tremendously. It is possible these people could be identified by traditional psychometric methods, such as personality types, and so on.
- The penetration testing methods described in the measuring section of this work could be studied further and applied in more, and different, organizations. They could also be used as benchmarks against other methodologies, both manual and more automated, to enable more precise recommendations. The methods could also be complemented with other means of penetration testing. There ought to be many other ways of trying to measure susceptibility.
- The cultural aspects, such as the values of the society in which the attacker and the victim live, tend to be somewhat overlooked in security. This may be due to the proverbial “hot potato” that cultural studies can become if not carried out correctly. In an increasingly globalized society it would

be improper for security specialists not to consider the area. Attackers would be more than happy to exploit it – something they are possibly doing already. There are several studies that may be useful as a starting point for this, most notably the "The World Value Survey", in which a large number of nations were surveyed on a large selection of areas; religion and values, sociodemographics, natural identity, religion and morale, politics and society, family, work, environment and perceptions of life. Each of the eight areas examined in this survey influences the social engineering attack, most importantly, how the attack should be performed to be successful. A would be perpetrator planning to attack a specific country would be well advised to spend some time reading through the data collected in "The World Value Survey" to prepare which kind of attack should be selected and how it should be deployed. Culture can also include organizational culture. While this is an area we know more about, it is still well worth studying further.

- AI-bots are rather simple pieces of software that are easy to program and develop, and, at the same time, can emulate basic human interaction (Walentowicz & Mozuraite Araby, 2008). They mainly work by pattern recognition, and are available both as free software and as more advanced commercial software. We have seen great potential in the use of AI-bots in research projects related to this. For instance, they have been successfully used as a training tool in teaching employees about security in general (Walentowicz & Mozuraite Araby, 2008), but we have also seen their potential in other areas. We believe they can be used to train people in experiencing social engineering attacks and how an attacker would work, while being educated at the same time. This would have the benefit of providing a much richer experience for the user, while at the same time keeping down costs and avoiding ethical dilemmas. The same bots could also be used for social engineering penetration testing, for example, when using social networks or online chats, a technique attackers will most likely be using in the future. If the bots were sufficiently developed, they could be used as a communication firewall, scanning messages and chats to detect and warn when deceitful communication is at hand. This, however, is a goal that lies sometime in the future, but the other suggested uses, mentioned above, can reasonably be developed and used within a short time. That carrying out successful social engineering without using much technology is simple does not mean that doing even more successful social engineering while using technology is not possible. We need to prepare the defenses and our organizations. This should be achieved partly by adapting to the threats of today, but also by realizing that the threats of tomorrow are going to be far more sophisticated. If we do not prepare for them today, we may be swept away by the oncoming torrent of attacks.

References

Adams A. & Sasse M. (1999). Users are not the Enemy: Why users compromise computer security mechanisms and how to take remedial measures, *Communications of the ACM*, 42(12), pp. 40-46.

Anti-Phishing Working Group (2005). eBay- 'UpdateYour Account.'. [Online]. Anti-Phishing Working Group. Available from: http://www.antiphishing.org/phishing_archive/05-03-05_Ebay/05-03-05_Ebay.html [Accessed 20 Nov 2008].

Anti-Phishing Working Group (2008). Phishing Activity Trends Report Q1/2008 [Online]. Anti-Phishing Working Group. Available from: http://www.antiphishing.org/reports/apwg_report_Q1_2008.pdf [Accessed 20 Nov 2008].

Bakhshi T., Papadaki M. & Furnell S. (2008). A Practical Assessment of Social Engineering Vulnerabilities. *Proceedings of the Second International Symposium on Human Aspects of Information Security & Assurance (HAISA 2008)*, Plymouth, UK, pp. 12-23.

Bank, D. (2005). 'Spear Phishing' Tests Educate People About Online Scams. [Online]. The Wall Street Journal. Available from: http://online.wsj.com/public/article/SB112424042313615131-z_8jLB2WkfcVtgdAWf6LRh733sg_20060817.html?mod=blogs [Accessed 20 Nov 2008].

Barret, N. (2003). Penetration testing and social engineering: hacking the weakest link. *Information Security Technical Report.* 8(4), pp. 56–64.

Berne, E. (1996). *Games people play: The psychology of human relationships.* New York, USA: Ballantine Books.

Biros, D. (2004). "Scenario Based Training for Deception Detection." *Proceedings of the 1st Annual Conference on Information Security Curriculum Development*, Kennesaw, GA, USA, pp. 32-36.

Björck, F. (2005). *Discovering Information Security Management.* PhD Thesis, Department of Computer and Systems Sciences, University of Stockholm and Royal Institute of Technology, Stockholm, Sweden.

Björck, F. & L. Yngström (2001). IFIP World Computer Congress / SEC 2000 Revisited. *In WISE 2, Proceedings of the IFIP TC11 WG 11.8 Second World Conference on Information Security Education 2001*, Perth, Australia, pp. 209-223.

Brandon, M. (2003-12-10). IT-säkerhet till varje pris för svenska storföretag [Online]. IDG. Available from: http://www.idg.se/2.1085/1.51395 [Accessed 20 Nov 2008] (in Swedish).

Brostoff S., Sasse A. & Weirich D. (2002). Transforming the ”weakest link”: A Human-computer Interaction Approach to Usable and Effective Security, *BT Technology Journal* 19(3), pp. 122-131.

Cao, J., Lin, M., Deokar, A., Burgoon, J. K., Crews, J. M., & Adkins, M. (2004). Computer-based Training for Deception Detection: What Users Want. *Proceedings of the second NSF/NIJ Symposium on Intelligence and Security Informatics (ISI 2004)*, Tucson, AZ, pp 163-175.

Cialdini, R. (1993). *Influence: the psychology of persuasion.* New York, USA: Quill.

Conti G., Ahamad M. & Stasko J. (2005). Attacking Information Visualization System Usability Overloading and Deceiving the Human. *Symposium On Usable Privacy and Security (SOUPS)*. Available from: http://cups.cs.cmu.edu/soups/2005/2005proceedings/p89-conti.pdf [Accessed 20 Nov 2008].

Dalrymple, M. (2005). Auditors Find IRS Workers Prone to Hackers. [Online]. AP. Available from: http://www.infosecnews.org/hypermail/0503/9684.html [Accessed 20 Nov 2008].

DeMelo, D. (2007). Sutherland's Differential Association. [Online]. Available from: http://web.archive.org/web/20070306161622/http://home.comcast.net/~ddemelo/crime/differ.html [Accessed 21 Nov 2008].

Dodge, R., & Ferguson, A. (2006). Using Phishing for User Email Security Awareness. In Fischer-Hübner, S., (Ed.), Rannenberg, K. (Ed.), Yngström, L. (Ed.), Lindskog, S. (Ed.), In *Proceedings of the IFIP TC-11 21st International Information Security Conference (SEC 2006)*, pp. 454-458. New York, NY: Springer Science + Business Media Inc.

Dourish, P. & Redmiles, D. (2002). An approach to Usable Security Based on Event Monitoring and Visualization, In *NSPW '02: Proceedings of the 2002 workshop on New security paradigms*, ACM Press, pp. 75-81.

Downs, J., Holbrook, M. & Cranor, L. (2006). Decision Strategies and Susceptibility to Phishing. In *Proceedings of the Second Symposium on Usable Privacy and Security*. SOUPS '06, vol. 149. New York, NY, USA: ACM Press, pp. 79-90.

Downs, J., Holbrook, M. & Cranor, L. (2007). Behavioral Response to Phishing Risk. In *Proceedings of the 2nd Annual eCrime Researchers Summit*, October 4-5, 2007, Pittsburgh, USA, pp. 37-44.

Gartner (2002a). There Are No Secrets: Social Engineering and Privacy" (TU-14-5662). [Online]. Gartner. Available from: http://www.gartner.com/gc/webletter/security/issue1/index.html [Accessed 21 Nov 2008].

Gartner (2002b). Protect Against Social Engineering Attacks, TG-14-7359. [Online]. Gartner. Available from: http://www.gartner.com/gc/webletter/security/issue1/article2.html [Accessed 20 Nov 2008].

Gartner (2004). Gartner Study Finds Significant Increase in E-Mail Phishing Attacks. [Online]. Gartner. Available from: http://www.gartner.com/5_about/press_releases/asset_71087_11.jsp [Accessed 20 Nov 2008].

Gragg, D. (2002). A Multi-Level Defense Against Social Engineering [Online]. SANS Institute. Available from: http://www.sans.org/rr/papers/index.php?id=920 [Accessed 20 Nov 2008].

Granger, S. (2001). *Social Engineering Fundamentals* [Online]. Security Focus. Available from: http://www.securityfocus.com/infocus/1527 [Accessed 20 Nov 2008].

Grazioli, S. (2004). Where did they go wrong? An analysis of the failure of knowledgeable Internet consumers to detect deception over the Internet. *Group Decision and Negotiation*, 13(2), pp. 149-172.

Gulati, R. (2003). The Threat of Social Engineering and Your Defense Against It [Online]. SANS Institute. Available from: http://www.securitytechnet.com/resource/security/hacking/1232.pdf [Accessed 20 Nov 2008].

Gupta, A. (2002). The Art of Social Engineering [Online]. InformIT. Available from: http://www.informit.com/articles/article.aspx?p=28802 [Accessed 20 Nov 2008].

Ferrell, J. (1995). Culture, Crime, and Cultural Criminology. *Journal of Criminal Justice and Popular Culture*, 3(2), pp. 25-42.

Flechais, I. & Sasse, A. (2005). *Usable Security*, in Lorrie, F.C. & Simson, G., (Eds) *Security and Usability*, Sebastopol, USA: O'Reilly Media.

FraudWatch International (2006). Tips to Protect Yourself from Phishing Scams. [Online]. FraudWatch International. Available from: http://www.fraudwatchinternational.com/phishing-fraud/phishing-protection/ [Accessed 20 Nov 2008].

Harl (1997). People Hacking: The Psychology of Social Engineering. [Online]. Available from: http://packetstormsecurity.nl/docs/social-engineering/aaatalk.html [Accessed 20 Nov 2008].

Hancock, B. (1996). Can You Social Engineer Your Way into Your Network? *Network Security* 1996(4), pp. 14-15.

Hansell, S. (2004). Organized crime may be behind Phishing. [Online]. New York Times. Available from: http://www.sfgate.com/cgi-bin/article.cgi?f=/chronicle/archive/2004/03/29/BUG8F5S1011.DTL [Accessed 20 Nov 2008].

Hasle, H., Kristiansen, Y., Kintel, K. & Snekkenes, E. (2005). Measuring Resistance to Social Engineering. In *Information Security Practice and Experience: First International Conference, ISPEC 2005*, Singapore, April 11-14 (2005), vol. 3439 of *Lecture Notes in Computer Science*, Springer, pp. 132-143.

Hermansson, M. & Ravne, R. (2005). Fighting Social Engineering. [Online]. University of Stockholm. Available from: http://www.dsv.su.se/research/seclab/pages/pdf-files/2005-x-281.pdf [Accessed 20 Nov 2008].

Heylighen, F. & Joslyn, C. (2000). The law of requisite variety. [Online]. Principia Cybernetica web. Available from: http://pespmc1.vub.ac.be/REQVAR.html [Accessed Nov 1 2008].

Hiner, J. (2002). Lock IT Down: Change your company's culture to combat social engineering attacks. [Online]. Available from: http://techrepublic.com.com/5100-1035_11-1047991.html# [Accessed 20 Nov 2008].

ISACA (2004). Is auditing procedure security assessment-penetration testing and vulnerability analysis. [Online]. ISACA. Available from: http://www.isaca.org/ContentManagement/ContentDisplay.cfm?ContentID=18750 [Accessed Nov 20 2008].

ISO/IEC (1999). Information Technology Security Techniques: Evaluation Criteria for IT Security, Parts 1 – 3 (No. 15408-1:1999). Geneva, ISO/IEC.

Jackson Higgins, K. (2008). Cybercrime, Cosa Nostra-Style. [Online]. Dark Reading. Available from: http://www.darkreading.com/document.asp?doc_id=159015 [Accessed 20 Nov 2008].

Jagatic, T., Johnson, N., Jakobsson, M., & Menczer, F. (2007). Social phishing. *Communications of the ACM*, 5 (10), pp. 94-100.

Jakobsson, M. (2005). Modeling and Preventing Phishing Attacks. [Online]. School of Informatics & Dept. of Computer Science, Indiana University. Available from: http://www.informatics.indiana.edu/markus/papers/phishing_jakobsson.pdf [Accessed Nov 20 2008].

Jakobsson, M., Finn, P. & Johnson, N. (2008). Why and how to perform fraud experiments. *IEEE Security & Privacy Magazine*. 2008 March-April; 6(2), pp. 66-68.

Jones, C. (2003). Social Engineering: Understanding and Auditing [Online]. SANS Institute. Available from: http://www.sans.org/rr/whitepapers/engineering/1332.php [Accessed Nov 20 2008].

Jordan, J. & Goudey, H. (2005). The signs, signifiers and semiotics of the successful semantic attack. [Online]. 14th Annual EICAR Conference. Available from: http://papers.weburb.dk/archive/00000135/01/EICAR_Paper_-_myles's_edits.pdf [Accessed 20 Nov 2008].

Kajava, J. & Siponen, M. (1997). Social Engineering - IT Security Threat of Informatics [Online]. IRIS 20. Available from: http://web.archive.org/web/20040422210025/http://iris.informatik.gu.se/conference/iris20/9.htm [Accessed 20 Nov 2008].

Kelly, M. (2007). Chocolate the key to uncovering PC passwords [Online]. The Register. Available from: http://www.theregister.co.uk/2007/04/17/chocolate_password_survey/ [Accessed 20 Nov 2008].

Krebs, B. (2005). Paris Hilton Hack Started With Old-Fashioned Con [Online]. Washington Post. Available from: http://www.washingtonpost.com/wp-dyn/content/article/2005/05/19/AR2005051900711_pf.html [Accessed 20 Nov 2008].

Kruger, H. & Kearney, W., (2006). A prototype for assessing information security awareness, *Computers & Security* 25(4), pp. 289–296.

Kowalski, S. (1994). *IT Insecurity: A Multi-disciplinary Inquiry*. PhD Thesis, Department of Computer and Systems Sciences, University of Stockholm and Royal Institute of Technology, Stockholm, Sweden.

Kowalski, S. (2002). Value Based Risk Assessment: The Key to a Successful Security Target for the Telecommunication Industry, *3rd International Common Criteria Conference* (ICCC) Ottawa, 2002.

Kumaraguru, P., Rhee, Y., Sheng, S., Hasan, S., Acquisti, A., Cranor, L.F. & Hong, J. (2007): Getting users to pay attention to anti-phishing education: evaluation of retention and transfer. In *Proceedings of the 2007 Anti-Phishing Working Groups eCrime Researchers Summit 2007*, pp. 70-81.

Lee, A. & Harley, D. (2002). Back to the Future – Fresh Approaches to Malware Management. *EICAR Conference Proceedings 2002*, pp. 76-109.

Levine, R. (2003). *The power of persuasion.* Hoboken, USA: John Wiley & Sons Inc.

Lincoln, Y. & Guba, E. (1985). *Nauralistic inquiring.* Beverly Hills, USA: Sage Publications.

Marett, K., Biros, D., Knode, M. (2004). Self-efficacy, Training Effectiveness, and Deception Detection: A Longitudinal Study of Lie Detection Training, *Lecture Notes in Computer Science*, Volume 3073, Jan 2004, pp. 187–200.

May, T. (2001). *Social Research: Issues, Methods and Process.* Buckingham, UK: Open University Press.

Microsoft (2006). Spear phishing: Highly targeted phishing scams [Online]. Microsoft. Available from: http://www.microsoft.com/protect/yourself/phishing/spear.mspx [Accessed 20 Nov 2008].

Mitnick, K. & Simon, W. (2002). The Art of deception: Controlling the Human Element of Security. Indianapolis, USA: Wiley Publishing, Inc.

Nohlberg, M. & Bäckström, J. (2007). Talking Security to Managers: How to Do it. In *Proceedings of the 6th International Conference on Perspectives in Business Information Research 2007*. Tampere, Finland.

Näckros, K. (2005). *Visualising Security through Computer Games: Investigating Game-Based Instruction in ICT Security: an Experimental Approach.*. PhD Thesis, Department of Computer and Systems Sciences, University of Stockholm and Royal Institute of Technology, Stockholm, Sweden.

O'Brien, T. (2005). Gone Spear-Phishin'. [Online]. The New York Times. Available from: http://www.nytimes.com/2005/12/04/business/yourmoney/04spear.html?ex=1291352400&en=2f313fc4b55b47bf&ei=5088&partner=rssnyt&emc=rss [Accessed 20 Nov 2008].

Ollmann, G. (2004). The Phishing Guide [Online]. Next Generation Security Software Ltd. Available from: http://www.ngssoftware.com/papers/NISR-WP-Phishing.pdf [Accessed 20 Nov 2008].

Orgill, G., Romney, G., Bailey, M., Orgill, P. (2004) The Urgency for Effective User Privacy-education to Counter Social Engineering Attacks on Secure Computer Systems, In *Proceedings of SIGITE Conference'2004*, pp. 177-181.

Parker, D. (1998). *Fighting computer crime: A new framework for protecting information.* New York, USA: John Wiley & Sons, Inc.

Pfleeger, C. & Pfleeger, S. (2003). *Security in Computing* (3rd ed). Upper Saddle River, USA: Prentice Hall.

Post- och telestyrelsen (2007a). Vilka är riskerna? [Online]. Post- och telestyrelsen. Available from: http://www.pts.se/pts/Templates/Page.aspx?id=27294&epslanguage=SV [Accessed 20 Nov 2008] (in Swedish).

Post- och telestyrelsen (2007b). Hur skyddar jag mig? [Online]. Post- och telestyrelsen. Available from: http://www.pts.se/sv/Internet/Internetsakerhet/For-hemmet/E-posta/Hur-skyddar-jag-mig/ [Accessed 20 Nov 2008] (in Swedish).

Rogers, M. (2000). A New Hacker Taxonomy [Online]. University of Manitoba. Available from http://homes.cerias.purdue.edu/~mkr/hacker.doc [Accessed 20 Nov 2008].

SIS (2003). SIS Handbok 550. Terminologi för informationssäkerhet. Stockholm, Swede: SIS Förlag AB (in Swedish).

Stasiukonis, S. (2006). Social Engineering, the USB Way [Online]. DarkReading. Available from: http://www.darkreading.com/security/perimeter/showArticle.jhtml?articleID=208803634 [Accessed 20 Nov 2008].

Srikwan, S. & Jakobsson, M. (2008). Using Cartoons to Teach Internet Security. *Cryptologia* 32(2), pp. 137-154.

Symantec (2006). Symantec Internet Security Threat Report Tracks Notable Rise in Cybercrime Activity [Online]. Symantec. Available from: http://www.symantec.com/about/news/release/article.jsp?prid=20060307_01 [Accessed 20 Nov 2008].

Syrén, H. & Malmström, K. (2001). Handbok för Försvarsmaktens Säkerhetstjänst, Informationsteknik Hotbeskrivning (H SÄK IT Hot). Stockholm, Sweden: Försvarsmakten (in Swedish).

Tanneeru, M. (2005). A convicted hacker debunks some myths. [Online]. CNN. Available from: http://edition.cnn.com/2005/TECH/internet/10/07/kevin.mitnick.cnna/index.html [Accessed 20 Nov 2008].

PWC (2007). The global state of information security 2007 [Online]. PriceWaterhouseCoopers. Available from: http://www.pwc.com/nz/security/pwc_GISS2007.pdf [Accessed 20 Nov 2008]

Thomson, M.E., von Solms, R. (1998) Information security awareness: educating your users effectively. In *Information Management & Computer Security* 6(4), pp. 167-173.

Trend Micro (2005). Hook, Line and Sinker [Online]. Trend Micro. Available from: http://www.trendmicro.com/NR/rdonlyres/8329E15A-B0B5-4392-AF55-C2E2B9A1601E/17124/PhishingPaper_FINAL.pdf [Accessed 20 Nov 2008].

Trend Micro (2006). The Trend of Threats Today: 2005 Annual Roundup and 2006 Forecast. [Online]. Trend Micro. Available from: http://us.trendmicro.com/imperia/md/content/us/pdf/threats/securitylibrary/trendannualroundup.pdf [Accessed 20 Nov 2008].

Tygar, J & Whitten, A. (1998). Usability of Security: A Case Study. [Online] School of Computer Science, Carnegie Mellon University. Available from: http://reports-archive.adm.cs.cmu.edu/anon/1998/abstracts/98-155.html [Accessed 20 Nov 2008].

Tzu, Sun (1910). The Art of War. Translated from the Chinese By Lionel Giles. [Online]. Available from: http://www.chinapage.com/sunzi-e.html [Accessed 20 Nov 2008].

Vroom C. & von Solms R. (2004). Towards information security behavioural compliance. Computers and Security 2004, 23(3), pp. 191-198.

Walentowicz, S. & Mozuraite Araby, R. (2008). Using Chatbots within Information Security Education. Masters Thesis, Department of Computer and Systems Sciences, Stockholm University/Royal Institute of Technology of Stockholm.

Wilson, T. (2007). Eight Faces of a Hacker. [Online]. Darkreading. Available from: http://www.darkreading.com/document.asp?doc_id=120800 [Accessed 20 Nov 2008].

Yngström, L. (1996). *A Systemic- Holistic Approach to Academic Programmes in IT Security*, PhD Thesis, Department of Computer and Systems Sciences, University of Stockholm and Royal Institute of Technology, Stockholm, Sweden.

Åhlfeldt, R-M. & Nohlberg, M. (2005). System and Network Security in a Heterogeneous Healthcare Domain: A Case Study. In Proceedings of the 4th Security Conference, Las Vegas, USA, 30 – 31 March 2005.

Part II: Publications

Part II includes six papers and one book chapter each representing different parts and stages of the research. They are presented in chronological order. The Book Chapter and Paper 3 present the result from the aspect of Understanding. Papers 1, 4 – 6 are mostly focused on Measuring. Paper 2 is mostly focused on Protecting, even if this area is also covered, to some extent, in the Measuring section.

Paper 1 Measuring	**Social Engineering Audits Using Anonymous Surveys – Conning the Users in Order to Know if They Can Be Conned** Marcus Nohlberg
Paper 2 Protecting	**User-centered security applied to the development of a management information system.** Marcus Nohlberg and Johannes Bäckström
Book Chapter 1 Understanding	**Why Humans are the Weakest Link** Marcus Nohlberg
Paper 3 Understanding	**The cycle of deception - a model of social engineering attacks, defenses and victims.** Marcus Nohlberg and Stewart Kowalski
Paper 4 Measuring	**Non-Invasive Social Engineering Penetration Testing in a Medical Environment.** Marcus Nohlberg, Stewart Kowalski and Kerstin Karlsson
Paper 5 Measuring	**Measuring Readiness for Automated Social Engineering** Marcus Nohlberg, Stewart Kowalski and Markus Huber
Paper 6 Measuring	**Phishing with Gifts as Bait: Measurement and Analysis of Phishing Attacks within a University Environment** Martin Boldt and Marcus Nohlberg

Social Engineering Audits Using Anonymous Surveys – Conning the Users in Order to Know if They Can Be Conned[1]

Marcus Nohlberg

Abstract

It is important to know the security readiness of any organization in order to strengthen it. One often neglected aspect of security is the human element, which is often attacked by "social engineering" techniques. This paper studies to what extent users are aware and susceptible to common social engineering attacks, and if a quantitative approach to penetration testing of social engineering can be used. By employing a quantitative study under the false pretense of studying "micro efficiency", an organization with above average skilled users was surveyed on three classic social engineering cons. The results indicate that the approach could be useful as a part of, or as a stand alone auditing technique. The human element is not only vulnerable, but vulnerable to the extent that it shadows most other security measures. The author argues for the necessity of education in order to counter the serious threat of social engineering, since, in many cases, it complies with the principle of adequate protection.

Key words: Social engineering, cons, fraud, security awareness, information security, vulnerability testing, penetration testing, auditing.

---

[1] A version of this paper was published in *Proceedings of the 4th Security Conference*, Las Vegas, USA, May 2005, ISBN 0-9729562-5-5.

Introduction

This is a paper about the human element of security, and more specifically about social engineering. This term was made famous by the hacker icon Kevin Mitnick, who managed to gain access to several high security government systems, not by using high tech password crackers or obscure bugs in the systems, but by using a con man's approach to obtaining information. By piecing this information together he managed to get the access he wanted. His most frequently used tool was the telephone and a well-delivered ruse.

The unique issue with social engineering is that social engineering attacks are quite different from the majority of the technical attacks in that they have a clear aim. The vast majority of the attacks and threats to security are "script kiddies", viruses, Trojans and other broad attacks, and thus done without a clear aim (Mitnick & Simon, 2002). Social engineering attacks, however, always have a clear purpose. It can be to acquire specific information, or even a login. Those doing it can be hackers, such as Mitnick was, doing it just for the curiosity, or high tech information brokers, doing it to steal information. It can even be foreign intelligence, doing it to prepare for war (Syrén & Malmström, 2001). It is, however, usually done with intent (Mitnick & Simon, 2002).

It is obvious that social engineering, to some extent, poses risks to organizations. Traditionally, when examining risks in organizations, one approach in information security is to do a penetration test. The dilemmas with penetration testing and social engineering are discussed by Barrett (2003), where the conclusion is that it is preferable to use an audit style that has results and objectives that are clear and can be accepted by both subjects and company. They should also not lead to discipline or dismissal for the individuals. A more traditional approach to social engineering auditing is argued by Jones (2003).

This paper tries to find an alternative approach to investigating just how aware the users are of a couple of classic social engineering cons, by using a con for the investigation.

Social Engineering

Social engineering is the technique in which an unauthorized person manages to pose as an insider or an authority to successfully get access to information or resources (Kajava & Siponen, 1997). A hacker can use social engineering to access other valuable data to benefit the hacker in further attacks (Rusch, 1999). Perhaps the best definition is by Granger (2001):

> "…social engineering is generally a hacker's clever manipulation of the natural human tendency to trust. The hacker's goal is to obtain information that will allow him/her to gain unauthorized access to a valued system and the information that resides on that system.

A social engineering attack focuses primarily on the people's vulnerability, and is based almost entirely on using the "principle of easiest penetration" (Pfleeger, 2003). The greatest threat is that no matter how secure the system is in itself, it is never more secure than its users (Granger, 2001; Mitnick & Simon, 2002). Social engineering can be used instead of, or in combination with, threats and bribes. The classic social engineer aims towards not leaving any traces, and generally leaving as little of an impression as possible, and thus threats and bribes are not favorite weapons of choice (Mitnick & Simon 2002). They can still be used, for instance by foreign intelligence officers (Syrén & Malmström, 2001).

Social engineering is used because it is often much easier to simply ask someone for information, than to prepare and conduct a complicated software or hardware attack (Granger, 2001; Mitnick & Simon, 2002).

Protection against Social Engineering

Literature seems to agree on one thing; there is no "silver bullet" protection against social engineering. Education is the most commonly recommended means of protection, particularly if combined with a decent security policy (Hancock 1998; Mitnick & Simon, 2002; Gupta, 2002; Granger, 2002). Mitnick & Simon (2002) also provide a couple of guidelines about what should be taught to the users regarding social engineering. These include the kind of attacks which can occur, how to detect them and where to report. There is also a lesson on not to trust everyone.

Aim

The aim of this paper is to answer the following questions:

Do users have a spontaneous awareness of common social engineering cons? That is: do they fall for cons?

Can a quantitative survey be used as a basis for a social engineering audit of an organization?

Research Approach

This quantitative study was done by means of a questionnaire, sent to a large number of subjects within the tested organization. Because of the sensitive information being examined, the questions could not be answered under the context of a security related questionnaire. This is because the respondents

would most likely act quite differently if they knew that the questionnaire was aimed at exposing their possible lack of concern for security.

Thus, a smoke-screen investigation was established. In order to make the respondents think and act as they probably would normally do, helpfully and with a normal approach to people contacting them, they were told that the research was about their efficiency in hypothetical everyday situations (albeit somewhat unusual) at work. As this investigation was said to be about work efficiency, it probably could have encouraged the respondents to answer in a slightly less suspicious manner than normal, as they would most likely want to have been perceived as efficient. This would, however, probably only have mimicked the powerful influence of a skilled social engineer.

The choice of using a quantitative study in this way was made because the difference among the subjects is quite hard to compare if the basis is interviews. One major characteristic among social engineering assaults is that most users in an organization, no matter what their role, have some kind of information that is useful for the social engineer. Therefore, the value of a broad base is high. The author has been unable to find any record of a similar research approach to a social engineering vulnerability study, thus this can be regarded as an explorative study.

The questionnaire contained six multiple choice questions. Out of these six, three questions were classical social engineering cons, and three were quite ordinary business situations. Standard questions, such as gender, age group, and type of work were also asked in order to facilitate data comparison later on.

In order to facilitate a high number of subjects, the study was conducted over the Internet. There is a probability that the choice of a web based survey influenced the result somewhat, as users might have felt more at ease filling out documents than answering questions over the web. Since the organization has used web based surveys in the past, and the subjects were quite accustomed to using the Web, it should, however, not have been a major influence. It is notable that the questions were asked in Swedish, and it is hard to translate all the nuances of the Swedish language into English.

The Test Subject

The nature of the research is such that details about the subject cannot be divulged, but a short description is given below.

The subject is an influential IT consultancy firm in Sweden. While the inherent need for security in a consultancy firm is rather high, it cannot be compared to that of, for instance, a company in the defense sector. However, as some consultants work for customers with various high security projects,

there is a general need for security, and a particular need for high security in some cases.

The Selection of Answers

A common selection of answers was designed. They were to be the same for every question, in order to provide the subjects with the sense of a standard survey feeling.

In order not to have a middle choice, as recommended by Patel and Davidson (1994), six different answers were chosen:

- I instantly help.
- I ask a number of complementary questions before I help.
- I ask complementary questions and ask if I can come back to the caller.
- I ask the person to contact a colleague.
- I ask the person to contact my director/superior.
- I deny helping the person completely

Each of these was chosen to reflect a probable reaction to common questions in everyday work. Since one of the prerequisites for this survey was that the participants were unaware of the real subject for the questions, an introduction was fabricated and written about the subject to be researched, namely "micro efficiency".

Distribution of the Survey

To reach the subjects, the author sent an e-mail to all the employees at the company, giving a short introduction and background to the survey. The recipients were urged to visit a webpage, where they were greeted with the following text (which has been slightly altered to assure the privacy of the organization):

> A study about micro efficiency.
>
> I am working with my final thesis for my MBA. This thesis will survey your company. I have chosen to work with organizational theory. One of the exciting new fields within organizational theory is the study of "micro efficiency", the measuring of how efficient companies are with small tasks during an ordinary day at work. The study is based on standardized questions, which I have adjusted slightly to adapt to your line of work, and standardized answers. The standardization of the questions might make them sound slightly strange, but it is important for the study that you try hard to answer as you would have done in the situation, in order for me to get a correct image of [company names] micro efficiency for my thesis. Some of the questions might not be applicable to your current position at work, but do try to envision how you would have acted any-

way. The survey takes approximately five minutes to answer, and it is important that as many as possible answer.

Your answers are completely anonymous, and cannot be tracked back to you.

Thank you for your help!

The Questions

In the beginning of the survey a couple of diagnostic questions were asked. They were age group, gender and type of work (administrative, management, consultant – technical, consultant – managerial, consultant – analytical, other).

The questions not related to social engineering are not included in this paper, but were mixed in with the other questions, as numbers 2, 5 and 6 in order.

After completing the survey, users were shown a text thanking them for their participation.

Results

Only basic statistical analysis was used on the data, most notably the Chi-Square test (c^2 (p)).

The formula for c^2 is:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Formula 1. Chi-Square test.

Where *O* is the observed frequency and *E* is the expected frequency.

In order to facilitate easier comparisons to age, a separation was done between “old” and “young” in some cases. “Old” was considered to be employees over 40 (alternatives 4, 5 and 6) and “Young” was under 40 (alternative 1, 2 and 3).

Survey Response Statistics

The e-mail about the survey was sent to approximately 300 subjects. Out of these, 130 looked at the survey, and 110 answered it completely.

This resulted in a rather low answer frequency, 37 %, but there were reasons for that. One is that the time for the survey was short, approximately one

week. Another is that the organization had recently been exposed to several surveys, making the employees hesitant to partake in a further study. A third reason for the low answer rate is probably that this survey was not being conducted by the organization, which made it easier to pass over. While a higher percentage of respondents would have been preferable, the nature of the research topic is such that the very existence of susceptible subjects is interesting in itself, even if the answer frequency of the organization as a whole is not that high.

The Questions Not Related to Social Engineering

There were three questions not related to the actual subject, but included in order to improve the "smokescreen". The answers from these were not analyzed, but a short conclusion can be that the recipients generally were quite helpful.

Diagnostic Questions

In the survey, a number of diagnostic questions were asked: age, gender and type of work.

Table 1. Age.

Age	**Count**	**Percent**
15 - 19	1	1 %
20 - 29	16	15 %
30 - 39	43	39 %
40 - 49	31	28 %
50 - 59	19	17 %
60 -	0	0 %
	110	**100 %**

The age distribution reveals nothing unexpected, as can be seen in Table 1. The organization's employees are mostly aged between 30 – 50 years. c^2 (p) is at 0.41 indicating that the age can be regarded as random among the genders.

In a fashion typical of many other IT related organizations, there are more males (76 %) than females (24 %).

Table 2. Type of work.

Type of work	Count	Percent
Administrative	6	5 %
Management	14	13 %
Consultant – technical	36	33 %
Consultant – managerial	22	20 %
Consultant – analytical	30	27 %
Other	2	2 %
	110	**100 %**

It is apparent that age influences what kind of work an employee has, c^2 (p) = 0.0007, which makes it highly unlikely that the difference is due to randomness alone. Table 2 presents the numbers.

The type of work is also influenced by gender, as the c^2 (p) for that comparison is 0.02.

The Questions Related to Social Engineering

The responses that can be said to indicate real security risks are answer 1,"I instantly help" and answer 2, "I ask a number of complementary questions before I help". Answer 2 is included since the author argues it is unlikely that a person with a basic desire to help could not be persuaded by a skilled social engineer, even if questions would be asked.

Alternative 3 is not considered a risk, since most social engineers would not give out information so that they could be contacted later, instead, they would choose a new target. Alternatives 4 and 5 could be considered security risks, as they are not hindering the attack, rather directing it to a new target. In this work they are considered as successful attempts at stopping the attack, based on the presumption that the attacker would be directed to a colleague or supervisor more skilled in deflecting a social engineering attack.

Alternative 6 is truly the most effective measure against the attack, but also the worst approach if there is a valid reason for the request. It is considered a valid defense in this study.

Question 1 - "Acquiring Information"

Question number 1 is based on a con where the objective is to acquire more information; about the company, and/or the employee. The subject has no information regarding whether the e-mail address really belongs to a valid recipient. While there might, in fact, be a valid external accountancy firm, their accounts can have been hacked, or more probably the attacker could be using a domain purchased to be similar to that of the real accountancy firm.

Who would consider a con such as this, and what would be the purpose? One of the most obvious is a headhunting firm, using this con to get information about interesting candidates. Not only would they get to know the percentage of billable hours, but also information about the customers and the projects. This would be quite valuable information when headhunting for the best, or most profitable, employees.

The question:

> You sit at your desk working. It is an ordinary Thursday in late September, and you do not really have all that much to do. You get an e-mail saying that a new person at [the company's] external accounting firm has started to work on the time reports, and she needs your preliminary time report mailed to her e-mail at the accounting firm as soon as possible, preferably today, as the end of the quarter is coming up and the owners have requested financial information as soon as possible.

Table 3. Question 1.

	Count	Percent	Answer
1	59	54 %	I instantly help.
2	21	19 %	I ask a number of complementary questions before I help.
3	15	14 %	I ask complementary questions and ask if I can come back to the caller.
4	2	2 %	I ask the person to contact a colleague.
5	12	11 %	I ask the person to contact my director/superior.
6	1	1 %	I deny helping the person completely
	110	**100 %**	

The data from Question 1 is presented in Table 3. The c^2 (p) analysis reveals that there is a difference between how males and females answer this question. There is a tendency, to some extent, for females to be more prone to “help out”, and to thus be vulnerable, (the c^2 (p) value is 0.06. Age is of no major influence, with a c^2 (p) value of 0.17, nor is type of work an influence, with a c^2 (p) value of 0.65. The data reveal that 73 % of the subjects would most probably send their time reports to the social engineer, and 54 % would instantly send them.

Question 2 - “Technical”/”Online Social Engineering”

Question 2 deals with a type of attack which could be described as a “Technical”/”Online social engineering” attack. The object is to get the mark to download a piece of software, which is most probably a Trojan, and then to log in again, allowing the trojan to transfer the login information, or other sensitive data, to the attacker in some manner.

This same kind of attack has many variations, from faked user-name pop-ups, to far more advanced software manipulation. It can be quite hard to notice, and thus the requirement to only install trusted software, from trusted sources, is obvious. An attack like this could be used against most organizations, but would only be successful in the organizations where the employees enjoy a certain degree of freedom with their own computers.

The question:

> You have been working with a major project for a customer for a long time. The project has been quite stressful in periods, and calmer in others. It has been a very rewarding project, and you think it is a little sad that it is soon to end. The project has involved a number of offices in several countries, as well as many people. The project also uses special software, especially developed for this project. You do not know that much about how it works, as you have mostly been involved at an overall level.
>
> At the moment you are sitting trying to finish the last of this week's documentation in order to leave for the weekend. You get a call from a "Fredrik", from technical support, which is stationed in Oslo. He explains that you seem to be running an old version of the software, something that has given them severe problems as it hinders the encryption from working correctly, because it uses another system for the public keys. Fredrik directs you to a link where a more recent version can be found and asks you to install it as soon as possible and to log in again to stop the problems.

Table 4. Question 2.

	Count	Percent	Answer
1	48	44 %	I instantly help.
2	36	33 %	I ask a number of complementary questions before I help.
3	22	20 %	I ask complementary questions and ask if I can come back to the caller.
4	1	1 %	I ask the person to contact a colleague.
5	2	2 %	I ask the person to contact my director/superior.
6	1	1 %	I deny helping the person completely
	110	**100 %**	

Table 4 shows the results from question 2. What kind of work the subject does seems not to influence the answers, c^2 (p) is at 0.76. Nor does age make any difference, c^2 (p) is 0.38 and gender is also without influence, c^2 (p) at 0.37.

It is highly probable that this attack would also be successful. The results show that 77 % of the subjects say they would download the patch, install it, and login again. Almost half of the users would do it without asking any questions.

Question 3 – “Simply Asking for Login Information”

This is a slightly more complicated version of the most classic social engineering con of all, simply asking for login information. There is an added element of complexity, as well as several constraints that could influence how respondents feel at the time. However, it is also a situation that is a clear violation of the most basic computer security principle; do not share your login information.

The question:

> It is snowy outside, and you are enjoying your Christmas vacation that is longer than you have had in some time. This is going to be a much needed vacation. You have been extra careful ensuring that your tasks at work have been divided among colleagues, so that everyone knows who is responsible for what while you are on vacation. You have also informed your most important customer, and he has, with the usual protests, accepted your vacation.
>
> You have just poured a glass of mulled wine when the phone rings. You curse your forgetfulness about leaving the phone on before answering. A very apologetic woman presents herself as a newly employed technician of your customer. She tells you that she is very sorry for calling you, but because she is going to upgrade the customer’s backup system she needs to know where the files you have worked on can be found. She also wants to know which computer they are located in and what folder, as well as what user-name you have, in order to guarantee that they will be backed up in the new system. You give her this information, and she thanks you and wishes you a Merry Christmas.
>
> An hour later she calls again, almost ready to cry. Something has gone wrong with your backup, she suspects, as it has been locked in the system, and she is the only one left over Christmas. It also seems that only your files are causing trouble. She asks if you would consider coming into the customer’s office, checking this, something you refuse. She then wonders if you can give her your login information to enable her to quickly check that the files have not been destroyed so that you can both celebrate Christmas in peace.

Table 5. Question 3.

	Count	Percent	Answer
1	14	13 %	I instantly help.
2	25	23 %	I ask a number of complementary questions before I help.
3	23	21 %	I ask complementary questions and ask if I can come back to the caller.
4	17	15 %	I ask the person to contact a colleague.
5	12	11 %	I ask the person to contact my director/superior.
6	19	17 %	I deny helping the person completely
	110	**100 %**	

A statistical analysis of the result, as presented in Table 5, shows that type of work influences the answers, c^2 (p), since the comparison of the results with type of work gives a value of 0.05. However, the variation is very low when reviewing the numbers and comparing the groups only shows a difference of a couple of percentage points. The minor difference in how the groups respond is interesting however. A reasonable assumption could have been that technical personnel should have been far less likely to divulge their passwords than, for instance, administrative personnel. This, however, is not the case. The difference is only 3 %. Still, there were not enough answers to fully investigate those differences. Age has no considerable impact, c^2 (p) is 0.21, and gender does not have an influence either with a c^2 (p) of 0.27.

More than one third of the subjects would have given out their login data, even if 23 % would have wanted to ask complementary questions. On the other hand, 17 % would have refused outright (interestingly, management consultants are by far the most resistant to this, 41 % refused, much more than other categories).

Analysis and Discussion

Approaching the subject of social engineering was done with some concerns. It is never easy to study humans, and to study their weaknesses is even harder.

Validity of the Results

There are a number of concerns with the results that are worth mentioning. The answer rate was low, a mere 37 %. While such a low response rate could have been a major concern in another type of study, it is not as critical in this study as it can be considered a broad vulnerability test. Therefore the very existence of users who are vulnerable is interesting, but a higher answer rate would, of course, have been preferable. However, if only the numbers, and

not the percentages, are studied, the results are still relevant in the author's opinion.

Perhaps the major concern with the study is the false pretext under which the survey was conducted. The motivation for doing a study based on efficiency, resulting in a comparable number, may have spurred the competitive instinct, or the will to help someone writing a thesis, by supplying some good answers. This could have made some subjects prone to answer more helpfully than they possibly would do in a real life situation. This is something that, in the author's opinion, is compensated by the fact that a skilled social engineer is far more persuasive than a survey. While a survey and a human social engineer cannot be directly compared, the results from the survey indicate to what extent users have an instinct allowing them to react when exposed to foul play. Perhaps the results also indicate the value users put on potentially valuable information.

Another possible influence is that some of the questions were somewhat complicated. Still, the very first question, which is probably the one that users pay attention to the most, was short and easy to read. It gave positive results with a trend similar to the other questions.

The Extent of Vulnerability to Social Engineering

The first research question asked was: Do users have a spontaneous awareness of common social engineering cons? In other words: do they fall for cons?

It would seem that while some are aware of the cons, or have at least a degree of suspicion, the number of unaware and susceptible users is by far large enough to warrant using social engineering as an attack. Very few users answered in such a way in which it could be suspected that they reacted with an awareness of something being wrong, except in the case of the question that clearly asked for the password. One reason for the success of social engineering is that many humans live with the assumption that nothing bad will happen to them as individuals. They might see that there could be a risk involved, but not something that affects them.

The results from the three different questions are discussed below:

Question 1 - "Acquiring Information" was quite successful. The reason for this is most likely due to the influence of two factors: time reports are not in any way considered to be sensitive information, and the subjects are, as many consultants, not attentive with regard to time reports and their handling.

Question 2 - "Technical"/"Online Social Engineering" uses an effective combination to achieve results. The arguments used by "Fredrik", the social

engineer, are based on a common tactic; confusing information, the need for haste, and a favor being done for both the perpetrator and the target. The same arguments and method of contact are, however, typically used by most real systems administrators calling users with valid requests. One other issue that could have influenced the answers is whether the subjects assumed they knew "Fredrik" or not, but to what extent this was an influence is hard to say. Still, one would have to know someone quite well to trust a voice over the phone recommending the actions that were suggested in Question 2.

Question 3 – "Simply Asking for Login Information" was also, to the author's surprise, successful. A part of the reason is the inability among humans to assume they are under threat even if the situation clearly is a threat. The attack consists of several of the classic techniques for influencing humans, as presented by Rusch (Rusch, 1999). Nevertheless, people should be alarmed when asked, more or less outright, for login data. The author's conclusion about the results of question 3 is that the security policies are no match for a comfortable evening at home. Comfort is judged as more important than security.

Using a Quantitative Survey for Auditing

The second research question was: Can a quantitative survey be used as a basis for a social engineering audit of an organization?

In creating the survey, the author's informal hypothesis was that some instances would possibly be susceptible to social engineering cons, but that the vast majority of users would see through them. The results indicate that users are more vulnerable than the author had expected them to be. In fact, the three different social engineering attacks that were suggested in the study would have been quite efficient. The study has identified what can only be described as critical security flaws. This implies that there is a lot of potential use for this methodology when doing security audits. One reason for using it is that a very large number of persons can be, at least to some extent, audited, thus giving a far greater quantitative basis for the human element in a security audit. Another reason for using it is that many of the ethical dilemmas about doing actual social engineering auditing on real persons can be avoided. Still, using a quantitative study does not give the whole picture, but rather a result which could be used for further, more thorough studies later on, if needed.

The Big Picture and Concluding Remarks

There are a couple of conclusions to be drawn from this study:

Social engineering works. And it seems to work well. It would even seem that it works so well that there really is not that much use for technical at-

tacks, unless used in cooperation with social engineering. This is something argued by Mitnick and Simon (2002). The second conclusion is that if these attacks are effective among technology savvy users, the success rate among less skilled users could be even higher.

While this was a study done in only one organization, the author believes that these results could be indicative for many organizations. This leads to some noteworthy complications. It would seem that much of the investment in protective hardware and software is, to some extent, wasted. If almost 15 % of the employees in an organization would easily give out their login information over the phone, it seems that the need for advanced protective software and hardware should be a second priority to education about social engineering, in the author's opinion. This is not to say that there is no need for software and hardware protection, only that the countermeasures to social engineering, primarily education, are inexpensive and provide protection against a, potentially, very serious threat. The risk of social engineering attacks should not be underestimated, nor forgotten, simply because protection against it cannot be bought as a product, but rather requires an investment in education and dedication from the organization.

In this work social engineering has been studied from a theoretical and, to some extent, practical approach. However, social engineering is a form of art, where a good con artist plays his mark like an instrument. Just like the salesmen at the fair, the subjects are drawn in by the magic done by the con artist. This is very hard to replicate using ordinary tests. It is also very hard to protect against, but informing that there is a risk, as well as giving some advice about how to act, might go a long way.

The result from this study implies some interesting theories. One theory is that the resources spent on security are not spent where they would have the greatest effect, as the users seem to be a greater security flaw than the author, and obviously many security experts, previously thought. A conspiracy minded individual could easily suspect that the reason there is so little information on, and protection against, social engineering is that it is relatively cheap to protect against, to some extent. In addition, there is more money to be made selling technical, or software, solutions that aim to improve the technical elements of security, than to simply educate the users.

Social engineering is an interesting field of information security, and it is interesting to note how unexplored it is. For example, Pfleeger (2002), which is probably one of the most influential books recently published on security in computing, only makes passing notes on social engineering as a threat.

Perhaps the extent of the threat must be known before action is taken, and perhaps this paper can be a part of that process.

References

Barret, N. (2003) Penetration testing and social engineering: hacking the weakest link. *Information Security Technical Report.* **8** (4), 56 – 64.

Granger, S. (2001) *Social Engineering Fundamentals* [Online]. Security Focus. Available from: http://www.securityfocus.com/printable/infocus/1527 [Accessed 18 Sep 2003]

Gupta, A. (2002) *The Art of Social Engineering* [Online]. InformIT. Available from: http://www.informit.com/isapi/product_id~%7B29543BE0-E535-4ADC-8FA5-E2C16295C8A1%7D/content/index.asp [Accessed Oct 10 2003].

Hancock, B. (1998) *Can You Social Engineer Your Way into Your Network?* [Online]. Computer Fraud & Security. Available from: *http://www.isoc.org/isoc/conferences/inet/99/proceedings/3g/3g_2.htm#s1* [Accessed 10 Oct 2003].

Jones, C. (2003) *The Social Engineering: Understanding and Auditing* [Online]. SANS Institute. Available from: http://www.sans.org/rr/whitepapers/engineering/1332.php [Accessed Nov 10 2004].

Kajava, J. & Siponen, M. (1997) *Social Engineering - IT Security Threat of Informatics* [Online]. Available from: http://iris.informatik.gu.se/conference/iris20/9.htm [Accessed Oct 10 2003].

Mitnick, K. & Simon, W. (2002) *The Art of deception: Controlling the Human Element of Security.* Indianapolis: Wiley Publishing, Inc.

Patel, R. & Davidson, B. (1994) *Forskningsmetodikens grunder.* Lund: Studentlitteratur. (In Swedish).

Pfleeger, C. (2003) *Security in Computing* (3rd ed). Upper Saddle River: Prentice Hall.

Rusch, J. (1999) *The "Social Engineering" of Internet Fraud* [Online]. Paper Presented at 1999 Internet Society Annual Conference: http://www.isoc.org/isoc/conferences/inet/99/proceedings/3g/3g_2.htm [Accessed 10 Sep 2003].

Syrén, H. & Malmström, K. (2001) *Handbok för Försvarsmaktens Säkerhetstjänst, Informationsteknik Hotbeskrivning (H SÄK IT Hot).* Stockholm: Försvarsmakten. (In Swedish).

User-Centered Security Applied to the Development of a Management Information System[1]

Marcus Nohlberg and Johannes Bäckström

Abstract

Purpose - To use user-centered security development of a prototype graphical interface for a management information system dealing with information security with upper-level management as the intended users.

Design/methodology/approach - The intended users were studied in order to understand their needs. An iterative design process was used where the designs were first made on paper, then as a prototype interface and later as a final interface design. All was tested by subjects within the target user group.

Findings - The interface was perceived as successful by the test subjects and the sponsoring organization, Siguru. The major conclusion of the study is that managers use knowledge of information security mainly for financial and strategic matters which focus more on risk issues then security issues. To facilitate the need of managers the study presents three heuristics for the design of management information security system interfaces.

Research limitations/implications - This interface was tested on a limited set of users and further tests could be done, especially of users with other cultural/professional backgrounds.

Practical implications - A useful set of heuristics that can be used in development of management information systems as well as other practical tips for similar projects.

[1] A version of this paper was published in *Information Management and Computer Security" volume 15, issue 5, 2007,* ISSN: 0968-5227.

Originality/Value - This paper gives an example of a successful user-centered security development process. The lessons learned could be beneficial in software development in general and security products in particular.

Key words: User-Centered Security, Information Security, Management Information System, Usability

Paper type Research paper

Introduction

While security has been a concern almost since the beginning of the history of computers, it is during the last couple of years that the problem has been communicated to a broader audience than merely systems administrators and technicians. Security has been a major operational issue for a long time, and the costs have continued to rise, as have the number of incidents. In 1998, 32% of British companies suffered some kind of information security incident, and in 2004 that number had risen to 74 % of all companies and 94 % of the major companies (Department of Trade and Industry, 2004).

New laws and regulations, such as Sarbanes-Oxley, make managers more responsible for security. The widespread media coverage of viruses, DOS-attacks, computer crime, and so on, also adds to the attention paid to security. Business partners and stakeholders demand good information security if they are going to conduct business with a company (Rasmussen, 2002). This all makes security a business problem.

Many managers are neither technicians nor particularly knowledgeable in information security. Most of their knowledge comes distilled from a specialist who informs them about the current situation, as discussed by the authors in a previous paper (Nohlberg & Bäckström, 2007). Managers only receive second-hand information since they themselves are unable to get any kind of impartial data from the organization using the systems they have today. From the point of view of managers this makes security vastly different from, for instance, economical data that can be found from several sources in a company. The few security solutions that gives some general security information aimed towards users rather than specialists are often regarded as too complicated or too difficult to use, as is understood from the authors' professional experience and as discussed more in general by Furnell, et al. (2006). Hence, there is a need for an application that provides managers with an overview and understanding of information security that is aimed towards their specific needs and interests. This approach also fits well with the strive to avoid "The 10 deadly sins of information security", as argued by von Solms and von Solms (2004), where security is argued to be a corporate governance responsibility, as well as a business issue.

The purpose of this study was to construct a usable interface for information security-monitoring software with upper-level managers as target users. In order to construct a usable interface, it was important to get to know the users, their situation, their view on information security, and what kind of information they need.

This study was supported by the company, Siguru, a small start-up company developing information security software, and is a part of the development process for the forthcoming product.

User-centered security

Almost as long as we have had computers and computer networks, there has been ongoing work with the development of programs to make them more secure. The focus of this work has been to generate powerful tools to protect our systems. The same attention has not been paid to making the users understand the programs and making the same people understand the importance of a secure behavior when using computers and computer networks (Whitten & Tygar, 1998; Flechais & Sasse in Cranor & Garfinkel 2005; Dourish & Redmiles, 2002).

Since the mid-nineties there has been a growing interest among researchers in the information security-area who have called for a more user-centered approach to information security. There is an increasing number of articles on the subject (E.g. Simon & Zurko (1996); Holmström (1999); DePaula, et al. (2005)).

The term “user-centered security” was coined by Simons and Zurko (1996) at the proceedings of the ACM-conference in 1996, and it can be seen as a key component of the movement for user centered development of information security applications. The term refers to “Security models, mechanisms, systems and software that have usability as a primary motivation or goal” (Simon & Zurko, 1996, page 27).

Design principles for development of information security applications

The same basic principles of usability that apply to other applications apply to information security applications. A great foundation for all usability design is the design principles developed by Donald Norman (Norman, 2002). They describe a number of heuristics that are likely to enhance the usability of a product, that is:

- Feedback, giving the user some sort of information of what his or her action has lead to will make the user more aware of the status of the system.

- Visibility, it is easier to label a control that only has one function. The label of a control can be used by the user to remember the control's function. If a control has many functions there is an immediate risk that the labeling will be ambiguous.
- Constraints, if the designer is able to constrain the number of actions that a user can carry out at a specific moment, the designer is also able to minimize the number of errors that the user can carry out at the same moment.

Other principles specifically considered in this project are described below.

According to Berson (in Carnor & Garfinkel, 2005), as little text as possible should be used to explain facts to the user. At the same time, it is important for the user to understand what is being communicated by the design, though the amount of text should be kept at a minimum. Every word, button and pixel should have a pedagogic ulterior motive (Berson in Carnor & Garfinkel, 2005).

The complexity of the interface should be kept as low as possible. It is also important *not* to design an application for all potential tasks that a user might be willing to undertake. The design should rather focus on the tasks most likely to be undertaken by the user (Berson, 2005).

An interface which gives the users too much information, information at the wrong time or in an unsuitable way, will be perceived as confusing by the user. If the amount of information is too small, there is a risk that the user will not discover potential security threats (Long & Moskowitz, 2005; Berson, 2005).

Try to teach the user simple tricks. For instance, in the web browser Firefox, the address bar turns yellow when the user enters a site that uses SSL (a protocol for transmitting data safely over the Internet). The user knows that the current page is a secure site when he/she sees this, without having to interact with numerous dialogue boxes (Berson, 2005). Using simple tricks like this is mainly positive, although the designer must continuously consider whether the user really needs this information about the program. A clue about when to use these tricks is when there is some change in the state of the program that the user needs to know about (Berson, in Cranor & Garfinkel, 2005).

Research methodology

The first task was to learn what the potential users of the product would actually want to know about security. This knowledge was obtained through a number of interviews and by scenario testing, described in more detail in a previous paper (Nohlberg & Bäckström, 2007). In general, the results indicated a specific interest in knowing about security from a financial and stra-

tegic perspective, grouped in sections of security information, rather than an interest in detailed data. In fact, security was perceived by the managers mostly as financial risks.

An interview was conducted with representatives from the sponsoring company, Siguru, in order to obtain a broader understanding of the product the interface was supposed to be used with, its limits and possibilities. This formed the first guidelines on how the interface was supposed to be designed.

When the information from the interviews had been collected, a “lo-fi” (low fidelity) prototype was constructed on paper. The design of the prototype was created on the basis of the information from the interviews and information regarding the design of interfaces that was found in the literature survey. The “lo-fi” prototype was hand drawn on paper, in order to quickly generate a sketch of the interface while at the same time communicating with the subjects that this was an early prototype hoping it would make them more willing to provide improvement suggestions.

User tests on the “lo-fi” prototype were made on potential target users. The first subject was in charge of the information security in a major governmental organization in Sweden. The second subject is the chairman of an information security company. The two subjects did not take part in the interviews. The user tests were recorded with a video camera. The tests consisted of six tasks and ten questions. The tasks were conducted first (except for the first question “What are your impressions of the first page”) and then followed up by questions.

The user tests were analyzed through a task log. The purpose of the task log was to find out where the test persons experienced difficulties with the prototype, why they experienced these difficulties, and what could be done to eliminate these difficulties (Hackos & Redish, 1998). Through this procedure, it was possible to find concrete improvements for the interface as well as investigate the test person’s mental model of how the system should work.

The “lo-fi” prototype was updated to a “hi-fi”(high fidelity) prototype using the feedback from the interviews, the results, and taking into consideration the inferences from the “lo-fi” tests. The “hi-fi” prototype was interactive and built using Macromedia Flash. The prototype thus emulated a working interface. In addition, the tests were done on a computer, in contrast to the tests of the “lo-fi” prototype, which were done on paper.

The “hi-fi” tests were carried out on three potential end product users. The users all had management positions within their respective company. The first two test persons are managers of a science park, while the third person is the MD of a mobile application company. The persons in the user tests for

the “hi-fi” prototype had not taken part in the interviews nor the user tests on the “lo-fi” prototype. The user tests were recorded with a video camera.

The user tests consisted of five tasks and thirteen structured questions connected to the tasks that had been performed. These questions were then complemented with follow up questions, depending on the answers of the interview subjects. The user tests were analyzed through a task log.

The “hi-fi” prototype tests resulted in an updated design of the interface and a number of requirements for how the program should behave and how it should be implemented in the organization, as well as a set of design heuristics.

Results

In this section, the three stages of development are presented, as well as represented by figures. These figures show only a small part of all the screens in the product, and are, of course, in higher resolution and color in the actual software. The purpose of the data displayed in the figures is to visualize the software UI, the data does not represent any actual organization. The features and functions behind the data are proprietary to Siguru, and thus not discussed further here. Where the decisions have been made with a basis in CHI literature, (then the source is referenced, in other cases it is decisions made during the study with conclusions from the prototypes as a background.

The “lo-fi” prototype, as seen in Figure 1, was constructed with support from theory and information gained through interviews. The major design decisions supported by theory and the interviews were:

- Every piece of information is (in most cases) just one or two clicks away. This is made possible through the flap system and supports "recognition rather than recall" (Nielsen, 1994).
- Three flaps were added, policy, education and inventory, in order to match the typical managers mental model of information security, which was found during the interviews and the scenario attached to the interviews (Nohlberg & Bäckström, 2007).
- Information that was not regarded to be highly important for the novice users was hidden under the triangles to support Nielsen’s principle of minimalist design (Nielsen, 1994). This was also done to minimize the amount of text and pictures, which would lead to reduced complexity and reduced clutter. Through this, the user primarily gets an overview of the current situation rather than details, which was one thing the interviewees said they wanted.
- The language was adjusted to suit managers rather than technicians.

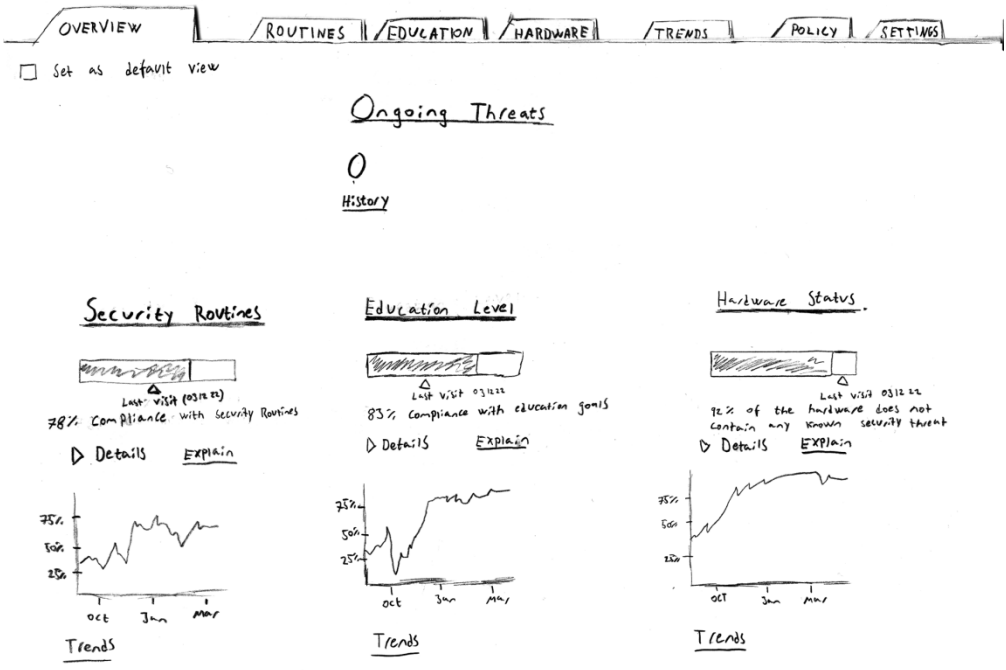


Figure 1. The "lo-fi" version of the overview page.

The "lo-fi" prototype, as seen in Figure 1, was developed on paper and tested. The major points of the feedback from the subjects were:

- The subjects preferred having as little information as possible at the beginning, but also stressed that it was important to be able to access additional information in an easy way.
- The subjects want to be able to review why an incident happened and what can be learnt/improved from that. Thus, the history should include when and why did the incident happen, how much harm was done, what did it cost, and other specific conditions at the time of the incident.
- The subjects want to be able to see different threats that occur at a certain time/period so that they can make strategic decisions based on this information. Therefore, the history page should include an option that lets the user compare the threats during specific periods.
- It is important for managers to see the consequences of their investments in information security. Therefore a flap for money and resources was added to the "hi-fi" prototype.
- The subjects want to be able to see how severe a single threat/attack has been to the organization, in order to be able to make new decisions based

on this information. Therefore the threats part of the inventory should include “consequences” of threats.

Figure 2. The “hi-fi” version of the overview page.

The “hi-fi” prototype was further developed based on the results of the “lo-fi” tests and the interviews mentioned above; a screenshot can be seen in Figure 2. The major points of feedback gained from the subjects were:

- A contextual help for each function that the users can access at any time will enhance the interaction with the system and provide guidance to the users when they need it.
- If the colors of the bars change depending on the status of the bar (e.g., if the value of a bar is critical, it should be red), then the users will more easily interpret the value of the bar. Therefore, if a value of a bar is acceptable, it should be in one color, if it is not acceptable, it should be in another, and if it is close to not being acceptable it should be in a third color.
- When a user expands information that concerns a single subject, it should be made clear whether the information that he is expanding is connected to the information above or if it is new information. Therefore, information that gives an overview of something should be made clearer and separate from information regarding one individual person
- According to the subjects, it is important to be able to use the software to follow up the goals of the company, expenses, and so on, and thereby facilitate their ability to make strategic decisions. Therefore, the trends

graph should have an indication of how the different values relate to the reference values of the company. All the subjects stressed that it was very important that the program should support strategic and financial decisions, since that is a very important aspect of managers' responsibility.

- All the subjects were satisfied with the amount of information that the interface presented.
- All the subjects stressed that it was important to involve their subordinates in the system. By doing this they were likely to be more motivated to act in a secure way and at the same time they would not feel as monitored as they would if they were not involved.

All the subjects stressed that the information that was important to managers was the kind that gives an overview of the current situation rather than details about information security, and that the information should be presented using a vocabulary that managers can understand.

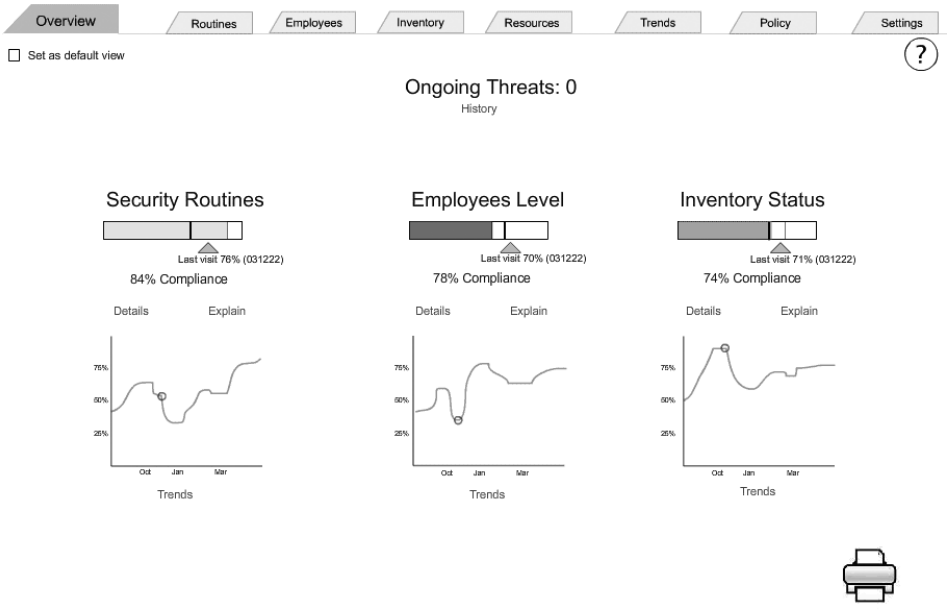


Figure 3. The final version of the overview page.

This input was used when developing a final version of the interface; a screenshot of the final version can be seen in Figure 3 above.

Conclusions and discussion

Managers are interested in security, but often find it hard to grasp the knowledge needed to fully understand and make decisions about security. Therefore, the security information given to managers should be adapted to the

specific needs of the target user, rather than the technical possibilities or the requirements of technical personnel. This project revealed several key characteristics of what managers want to know about security:

1. Managers are more interested in the overall status of the information security of the company than the details. This does not mean managers are uninterested in details, but that they want details only when they are needed. Managers also tend to group together areas of information security.
2. Managers consider information security on a more strategic and financial level than security specialists, who tend to focus more on risks.
3. Managers do not only see security from the security perspective, but also consider the possibility of making other gains, such as increased efficiency, minimized downtime, and so on.

In order to cater to the specific needs of managers, three design heuristics for user centered security design aimed at managers were developed during development of the interface. They were based on the works of Norman (2002) and further developed in this context:

1. Provide overview information very early in the program. The typical manager does not have the time or the knowledge to make this overview by himself/herself.
2. Do not overwhelm the user. Normally a manager is not interested in the details of the information security and/or does not have time to read this sort of information. If the manager wants the information, the manager is likely to find it.
3. Provide information in a way that is familiar to the manager. Use wording that the user understands. Provide contextual help for expressions that must be presented in a technical way.

This project aimed towards developing an interface to display security related information to managers. The user centered process for creating the interface has been successful. The concept and the interface have been appreciated by both the subjects and the company, and will now be used as the basis for developing the actual “Siguru”-product.

From a managerial perspective, it is important to know where the educational and economic resources should be spent to secure proper information security in the entire organization. This kind of security information system might be able to prevent people from acting in an insecure way, since it will help managers make the right investments, be they in technology, resources, or education.

In the future, the Siguru-software will also consist of an education module, to be used by each and every employee. It is believed this will help the employees to educate themselves in security, as well as improve general awareness of security. With a management information system like the one proposed in this study as a foundation, security education can be transformed from a mere sidetrack to a critical process, the same way that information security might be transformed by helping the decision makers understand, and be active in the process. Information security might finally be integrated in the normal decision processes of managers. This is the first step in achieving a really secure organization – when those making the decisions both care about security and understand it, and a good way to stop committing the 10 deadly sins of information security management (von Solms & von Solms, 2004).

Acknowledgments

The authors would like to thank the subjects for giving freely and willingly of their time. The authors would also like to thank colleagues and friends for support during the process, in particular Benkt Wangler, Stewart Kowalski, Thomas Ekström from The Logic Planet AB and Alexander Backlund.

References

Berson, J. (2005), Zone Alarm: Creating Usable Security Products for Consumers, in Lorrie, F.C. and Simson G., *Security and Usability*, O' Reilly Media, Sebastopol, CA.

Department of Trade and Industry (2004) "United Kingdom's Department of Trade and Industry's Information Security Breaches Survey 2004", www.pwc.com/uk/eng/ins-sol/publ/pwc_DTI-InfoSecutiry-Survey2004-Exec.pdf (Accessed 10 May 2006).

DePaula, R., Ding, X., Dourish, P., Nies, K., Pillet, B., Redmiles, D., Ren, J., Rode, J. & Silva, F.R. (2005), "Two Experiences Designing for Effective Security", cups.cs.cmu.edu/soups/2005/2005proceedings/p25-depaula.pdf (Accessed 15 November 2006).

Dourish, P. & Redmiles, D. (2002), *An approach to Usable Security Based on Event Monitoring and Visualization*, New Security Paradigms Workshop 02.

Furnell, S.M., Jusoh, A., Katsabas, D. and Dowland, P. (2006), *Considering the Usability of End-User Security Software*, Proceedings of 21st IFIP International Information Security Conference (IFIP SEC 2006), Karlstad, Sweden, 22–24 May 2006, pp. 307–316.

Flechais, I. & Sasse, A.M. (2005), Usable Security, in Lorrie, F.C. and Simson, G., *Security and Usability*, O' Reilly Media, Sebastopol, CA.

Hackos, J.T. & Redish, J.C. (1998), *User and Task Analysis for Interface Design*, John Wiley & Sons, Inc., New York, NY.

Holmström, U. (1999), "User-centered design of security software", www.hft.org/HFT99/paper99/Design/5_99.pdf (Accessed 10 October 2006).

Long, C.A. & Moskowitz, C. (2005), Simple Desktop Security with Chameleon, in Lorrie, F.C. and Simson, G., *Security and Usability,* O' Reilly Media, Sebastopol, CA.

Nielsen, J. (1994), Heuristic evaluation, in Nielsen, J., and Mack, R.L., *Usability Inspection Methods,* John Wiley & Sons, New York, NY.

Nohlberg, M., Bäckström, J. (2007), *Talking Security to Management: How to Do it.* In Proceedings of the 6th International Conference on Perspectives in Business Information Research 2007. Tampere, Finland.

Norman, D. (2002), *The Design of Everyday Things,* Basic Books, New York, NY.

Rasmussen M. (2002), "IT-Trends 2003: Information Security Standards, Regulations and Legislation", images.telos.com/files/external/Giga_IT_Trends_2003.pdf, (Accessed 5 August 2006).

Simon, R.T & Zurko, M.E. (1996), *User-centered security,* Proceedings of the UCLA conference on New security paradigms workshops September 17 – 20, portal.acm.org/citation.cfm?id=304859 (Accessed 5 August 2006).

Von Solms, B., Von Solms, R. (2004), The 10 deadly sins of information security management, *Computers & Security* 23 (5), pp. 371 – 376.

Tygar, J.D & Whitten, A. (1998), "Usability of Security: A Case Study", reports-archive.adm.cs.cmu.edu/anon/1998/abstracts/98-155.html (Accessed 2 August 2006).

Why Humans are the Weakest Link[1]

Marcus Nohlberg

Abstract

This chapter introduces the concept of social psychology, and the forms of deception to which humans are prone to fall. It presents a background of the area as well as a thorough description of the most common and important techniques of influence. The chapter also provides more practical examples of potential attacks, the kinds of influence techniques being used, a set of recommendations on how to defend against deception, and a discussion about future trends. The author hopes that an understanding of why and how the deceptive techniques work will give the reader new insights into information security in general and deception in particular. These insights can be used to improve training, to discover influence earlier, or even to gain new powers of influence.

Key words: Information Security, Hacker, Sociology of Computing, Social Engineering, Deception, Fraud, Influence, E-Commerce Risks, Electronic Commerce Security, Internet Security, Security Threats

[1] A Version of this book chapter was published in Gupta, M. and Sharman, R. *Social and Human Elements in Information Security: Emerging Trends and Countermeasures*, IGI Global, Hershey, PA, USA. ISBN: 978-1-60566-036-3.

Introduction

A computer crime starts, and ends, with a human, no matter which method of attack is chosen. Many successful computer crimes could have been prevented if the people involved had been more vigilant, more security conscious, or aware of their own weaknesses. This chapter deals with human weakness. It can be perceived as a "how-to-manual" for the aspiring attacker, but it could just as well be perceived as a "know-yourself" guide that both individuals and professionals can use to improve their personal and organizational defenses. It might also give a little more understanding for the victims. When researching successful attacks from the comfortable position of the outside observer, most of us are prone to throw the first stone against what can be seen as gullible humans. The fact is that almost everyone is susceptible to the techniques and weaknesses described in this chapter, simply because the attacks play on human emotion rather than logic.

Background

We humans are complicated beings, with some interesting shortcuts in our behavior. In recent years, there have been multiple studies on deception in general and influence in particular. These studies have been done, amongst others, in the field of economics and most notably in social psychology. In order to stay as close to the human element as possible, this chapter focuses on the social psychological aspects that can be practically used by the attacker. Although there are ample theories and much work is being done with a more theoretical application, this chapter focuses on the techniques that the perpetrators might use. Cialdini (2001) has written one of the most influential books in this area, and this chapter follows his use of the six basic rules of influence, together with some other added aspects of influence. In order to facilitate a better understanding of the concepts, examples are given, both from the literature and from real life. When applicable, the terms are tied together with information security as far as possible. Not all the information here is from research, some is also added from online sources, and includes guides about what to explore and attack which are written for the aspiring social engineer. While this information has not been judged against academic standards, it is still relevant, because it is the information attackers will try to use for their attacks, and therefore important to know.

Deception is a powerful tool for any attacker, but also for any parent, teacher, salesman, or most of us in our everyday lives. We buy and sell goods, we court romantic interests and we try to raise our kids in a good way without them loathing us too much when we try to get them to do their chores. In all of these examples, and many more, deception is the key element. Deception can be defined as:

> "Everything done to manipulate the behavior of the other side, without their knowledge of the friendly intent, for the purpose of achieving and exploiting an advantage is deception. The "what" of deception is the manipulation of behavior. The "why" is to exploit the advantage achieved." (Feer, 2004).

There are two different kinds of deception. One is dissimulation, which concerns the hiding of the truth (Bowyer, 2003). The truth can be hidden in three ways. One of these is to mask the information, for instance, by hiding nefarious features in a piece of software. Another way is to repackage the information, for instance, by hiding a Trojan horse in legitimate software. Finally, information can be dissimulated by dazzle, to shock or surprise, for instance, by sending nude pictures in an e-mail. The other kind of deception, simulation, deals with exhibiting false information. Simulation can be done by mimicking, which is spoofing or imitating reality, for instance as in a Phishing attack. It can also be done by inventing, which is the creation of a new reality, for example, false messages from Microsoft that a certain bug must be patched as soon as possible. The final method of simulation is decoying, which is a diversion created to divert from the real object, such as the false warning of a different attack than the one to which you are being exposed at the time.

Humans and Deception

Most of us, and indeed probably you, the reader, consider ourselves exceptionally resistant to manipulation. We are better than the average at detecting lies, and can spot a con a mile away. When asked about our friends' susceptibility to deception, however, we find them to be far more gullible (Levine, 2003). Obviously we are misjudging our own capacities, as influence in general is highly effective, which is proven by the huge profits it generates for advertisers, corporations, and religious groups, among others, that use these techniques.

The reason people misjudge their own abilities to spot deception is because of the "lie detector bias" where individuals almost always overestimate their ability to detect lies (Marett, et al. 2004). This is further complicated by the truth bias, which is the widespread assumption that most people are telling the truth (Martin, 2004). Humans also tend to think that bad things, such as death, accidents, crime, natural disasters, and so on, generally only happen to others (Levine, 2003). To further highlight our vulnerabilities, another interesting weakness in the human psyche is the "fixed-action patterns". They are most easily studied in animals, where certain specific conditions, "trigger features", trigger a predetermined response. For instance, a certain breed of bird will instantly start to care for any egg-like object, even if it is obviously not an egg but perhaps a painted volleyball instead (Levine, 2003). While "fixed-action patterns" might seem impractical, for animals in particular they

save time, energy, and mental capacity. Even for humans, certain "fixed-action patterns" are beneficial, for instance, giving thanks when receiving a gift, complying with what police officers say, and so on. In normal circumstances, the "fixed-action patterns" are usually correct and beneficial to us. They start to become a major problem when someone induces us to use them as a weapon against ourselves.

So basically, we humans believe that we are good at spotting lies, that people seldom lie to us and that bad things mostly happen to others. We are also mostly blissfully unaware that we have certain "fixed-action patterns" that will make us react almost without thinking to certain requests. This is the stage for the great game of influence.

The Basics of Influence

The easiest, and often the most efficient way of influencing others is simply to be kind. A bit of kindness goes a long way, since most average users really want to be helpful (Granger, 2002). A slightly more advanced method is to add the illusion of a reason behind the request. The illusion of a reason can be just as effective as a good reason. When people were asked to do something, it was found that simply using the word "because" in the question was just as effective as using it together with an actual motivation (Cialdini, 2003). In order to develop deeper skills of influence, any single one, or combination of, the following techniques can be used.

Authority

People are likely to respond obediently to authority. We are generally brought up to respect authority and ever since childhood it has been beneficial to obey the authorities; both in school, at home, in church, in the army, and in the workplace. Listening to authority is seldom detrimental to anyone. In extreme circumstances this can push people to do dreadful things, as was shown by the famous Stanley-Milgram experiment (Obedience to Authority Study). In the study, subjects thought that they were administering electric shocks to another (fake) subject in order to punish them for errors. The real study was to test their willingness to administer painful, or even potentially lethal, doses of electricity while being told to do so by an authoritative test supervisor. The study showed that a disturbingly high percentage (65 %) were willing to continue the experiment even though they, to the best of their knowledge, were administering extremely painful and potentially lethal doses of electricity to another subject who was screaming in pain and complaining about intense chest pains (Blass, 2002).

However, authority is not only someone telling us what to do. Other aspects than verbal orders also influence who we think is a person of authority. One

example of this is the uniform. The wearing of a uniform is a cheap and simple way for someone to be perceived as a person of great authority (Mitnick & Simon, 2002). Uniforms can be of the obvious kind, police uniform, doctor's coat, soldier's uniform, but perhaps the most effective kinds of uniforms are those we would not normally perceive as uniforms. Examples of these kinds of uniforms include the clothing worn by technicians, maintenance personnel and cleaners. These groups of people tend to often have full access to most areas, frequently at times when few or no other employees are present. They are also often employees of another organization engaged as subcontractors. Since such people have full access, and are rarely questioned, it makes them a risk element. Reasonably normal clothing is also a kind of uniform, especially the style and message of the outfit. For instance, a nice, tailored suit sends a different message than a "nerdy" Linux t-shirt, but both are efficient as uniforms in a particular context. Another kind of uniform is the title of a person. For example, an impressive title, such as professor, doctor, lord, sir, and so on, can influence the amount of authority we perceive that someone has (Cialdini, 2003). Achieving legitimate titles often takes years of hard work, but acquiring a fake title takes only seconds. Even fake diplomas can be bought cheaply, making it even more difficult to judge the value of a mentioned title.

Other examples of items that make us perceive someone as having authority are purely material artifacts, such as those associated with wealth, for example, fancy clothing, jewelry, and expensive cars, as well as certain other human traits such as length and tone of voice. Humans are easily influenced by these things, and having the right clothes can make a big difference, something which is well known by con men (Cialdini, 2003).

Except for the sad fact that the imagery in hip-hop videos actually works well to influence our perceptions of the artists as important, the practical consequences of this human weakness for uniforms and fancy attributes are that an attacker could benefit from using either a specific uniform to make desktop hacking easier, or, for instance, using specific titles to make a social engineering attack over the telephone more efficient. This was made chillingly obvious in a study where nurses were called on the telephone by a person introducing himself as a doctor responsible for one of the patients. The presumed doctor then proceeded to tell the nurse to administer a dangerously high dosage of medicine to the patient. Without requesting further identification most nurses, 95 %, complied, and were stopped by the researchers on their way to the medicine cabinet (Levine, 2003).

Scarcity

When people are told that something they want is in short supply, they tend to want it even more. The information that others might be competing for the

same thing triggers the sense of competition. This can be observed in advertisements everyday, where terms as “limited supply” are frequently used. Time is always a stress-inducing factor, it is efficient to make the market believe that time is in limited supply, thus leaving less time for reflection (Cialdini, 2003). Our reactions to scarcity also mean that the things which are hard to possess are valued higher and perceived as better than those that are easy to possess. This has interesting consequences for how people value information that is banned or made secret. When information is banned, humans have a greater desire to acquire it. In addition, they also have a more favorable attitude towards it than before it was banned. Furthermore, humans have a greater interest in what has become scarce, than what has always been scarce (Cialdini, 2003). That people value banned information more is a noteworthy piece of information for organizations that begin to employ stricter secrecy policies, or that have a rigorous security classification. It also explains some of the basics for the so called “hacker culture”. Information wants to be free, because if it is secret, it must be interesting. It also means that information might actually be more secure if it is not classified as secret at all. The very classification of secret makes people want it, because then it is limited, and if it is limited, it is good. This principle can also explain why people lust to get into exclusive night clubs, even if queuing to get in can take all night, why people work so hard to be accepted into more or less exclusive social clubs, why the value of art increases when the artist is dead, and why almost everything nowadays is sold in “limited editions”, products ranging from sodas to cars. This is because if supplies really are limited, we do not want to miss the chance of buying the product. Scarcity works because we learn, historically, that the good things really are in short supply. And if they are in short supply, we lose the freedom of choice, something we, as humans, resent (Cialdini, 2003).

Scarcity can be used by attackers providing a “limited service offer” or pressing the time factor: “Sure, I could help, but I’m leaving soon, so we’ll have to fix it quickly”. Another consequence of scarcity is that making information harder to get could actually make more users interested in it, in fact, making it less secret.

Liking and Similarity

People favor others that are like themselves. If we share similarities, then we are prone to react favorably to a person similar to ourselves only because of the similarity. Another particularly influencing factor here is the physical attractiveness of a person. A person who is very attractive can be perceived purely as an attractive person, when attractiveness is the dominant characteristic of the person. This is referred to as the “halo effect” and it makes attractiveness a very influential factor (Cialdini, 2003). In fact, an attractive

physical appearance can make us believe that the person is smarter, kinder, stronger, and of a higher moral character. However, we are also oblivious of our mostly automated preference towards attractive people (Levine, 2003). If you are blessed with an attractive physical appearance, you will find that influencing people is easier.

There can be several different kinds of similarity, for instance, dressing in a similar way as others and sharing individuals' background and interests. Thus, in the choice of how to dress when attempting to deceive, it is basically a choice between using authority, or dressing like the victims and using similarity (Cialdini, 2003). The importance of liking is also emphasized in Neuro-Linguistic Programming, NLP, where much focus is on developing rapport between people. In NLP, rapport means being "in sync" with the person you are talking to. The common techniques are the matching of body language, breathing (frequency) and maintaining eye contact (O'Connor & McDermott, 1996). Creating rapport increases liking, and is a powerful weapon of influence.

Other ways to increase liking is to have frequent contact with the target, as familiarity increases liking, a tactic which is also used in examples by Mitnick and Simon (2002). What is interesting here is that familiarity works without victims realizing that it occurs. Thus we tend to like people frequently featured in the media, or those we often see at work, for no other reason than that we see them often. An effective method to achieve liking quickly among strangers is to share a common "enemy". This is something most army recruits have experienced when sharing the dislike for certain officers is a sure way to get conversation started. If the attacker manages to leverage himself and the victim into a situation in which they cooperate in order to gain mutual benefits, such as helping each other, liking will also increase. As our senses are linked together with our overall experience of a situation, it is also, interestingly enough, effective to meet while eating. The positive experience of the food will strengthen liking. Most importantly, it is important to avoid meeting under adverse conditions, as the negativity of the circumstances will affect the liking between the persons involved, as does being the bearer of bad news. We are also easily affected by compliments, even if we realize that the compliments are given with an ulterior motive (Cialdini, 2003).

This knowledge could be used by an attacker to befriend the target, in order to build liking and rapport between them, for instance, by sharing an enemy (perhaps the boss), or by sharing a remarkable number of interests, and having a similar background. Why are most car salesmen so similar to their customers, with children roughly the same age?

Reciprocation

The rule of reciprocation is hard wired in us, and might indeed be the very reason we are humans; our ancestors learned to share which led to civilization. The rule is quite simple: if someone does a favor for us, we feel that we must repay that favor, even if we did not ask for it. It is more or less an automatic reaction, frequently used, and abused, for instance, by car salesmen. They may tell a customer that they are doing them a favor by lowering the price, or by including rust proofing, or even by selling them the car without any commission. This induces the customer to feel obliged to repay the salesman, and what better way to do so than to buy a car?

Reciprocation is a very powerful technique that in many cases can be directly responsible for successful influence (Cialdini, 2003). One of the classic examples is the flowers that are handed out by Hare-Krishnas. The flower is free, they say, but it is customary to give a small donation in return. Even if the receiver of the flower does not want it, or even likes the Hare-Krishnas, the person will feel obliged to return the favor, and to give a donation. In fact, this technique is so powerful that it is one of the major reasons for the success of the Hare Krishnas (Cialdini, 2003). The same technique is behind the free taste samples one often encounters in super-markets. Not only are the customers able to taste the product, the sample also has the aura of a gift around it, making it hard for people to resist buying the product after receiving a sample as a gift from the pleasant sales lady.

What should be noted here especially is that people's sense of reciprocation will stand even if the gift is very small, and the request for return is far greater than what would be reasonable (Cialdini, 2003). A variation of the reciprocation rule is referred to as the "rejection-then-retreat" technique. It consists of making an initial, extreme, offer that is sure to be rejected, and then retreating to a lower, more sensible, request that was the goal of the initial offer. An example would be asking someone to buy a $50 painting to support the arts, and upon rejection offering them a $5 set of postcards. Not only does the "rejection-then-retreat" technique increase the possibility of the request being accepted, it is also more probable that the target will carry out the request, and succumb to such requests in the future (Cialdini, 2003).

An attacker could use this by stating that he has helped the victim in a small matter without prior request, or by giving the victim privileged information that he did not ask for.

Commitment and Consistency

No one wants to be known as a failure. If a person has promised to do something, he will try his best to do it, so as to not be regarded by his peers as untrustworthy. Therefore, people try hard to act in ways consistent with pre-

vious behavior and with the choices they have made. Similarly, people find they are more willing to abide by their decisions when they have been made public in some way, when a stand has to be taken. This is why a gambler is far more certain of the odds after placing a bid than before, and also why so many charities collect signatures on lists (Cialdini, 2003).

In order for a commitment to be at its most effective, it should be active, public, and demand a certain degree of effort. In addition, if a person is to accept responsibility for it afterwards, it should also be made without strong outside pressures (Cialdini, 2003). This has the interesting spin-off effect that it is actually harder to convince someone to cooperate for a longer period of time using a large bribe, or a really violent threat, than it is using a smaller bribe and a more feasible threat. This is something that was well known during the cold war, when most recruited traitors were not actually paid a great deal of money. It was more effective for the foreign power to get classified information for relatively little money, as the traitors could then justify their treason not just for monetary gain but also with ideological support. This would make them feel more personally responsible and induce them to have a greater commitment to the relationship, which made them easier to exploit as resources for a long time.

A person wanting to use this knowledge to influence someone could try to get the target to express public support for the concept, while making the expression of the support not too easy. If a bribe were offered, it should be relatively small, and any threat made should be of a reasonable kind, not too spectacular, but threatening enough to “tip the edge”. If the intimidation is too threatening, the mark will not feel obliged to follow through as soon as the immediate risk is removed.

Social Proof

When people have to decide on the proper behavior in a situation they are uncertain about, they observe the actions of people in their vicinity, especially those similar to themselves. Usually it is correct to do the same thing as the people around you. This is the phenomenon known as “social proof”. Social proof can cause people to do things not in their own self-interest, such as purchasing products because of their popularity, or sharing passwords with coworkers because “everyone else in the department is doing it”. What is even worse, it can lead to a phenomenon referred to as pluralistic ignorance (Cialdini, 2003). Pluralistic ignorance is when people try to see how everyone else is acting, leading to a situation where no one acts at all. This is exemplified, most horrifyingly, in cases where crimes are committed in front of many witnesses but no one acts to help the victim, or when someone becomes ill in the middle of the street and no one stops to check if the person is all right. On the other hand, when someone actually stops to check if the

person is all right, several others might help out almost immediately, as I myself discovered while helping an elderly lady who had fallen off her bike. After the first couple of volunteers had arrived, the crowd started to snowball, and soon people had to be told to leave in order not to create a traffic hazard.

This could have a major impact on the security of any organization, because people will adapt to the general attitude towards security in the organization, rather than to what is written in a policy. Even if management wants to have a high degree of security, the employees can nullify any attempts, unwittingly, by social proof. Examples of this are organizations where the sharing of passwords, while expressly forbidden in the policy, still is a sign of trust among employees. Not sharing would stigmatize a person as untrusting, paranoid, and not a part of the group, since sharing is regarded as a matter of trust (Brostoff, et al. 2002).

An attacker could use this to enforce the techniques of persuasion by telling the target that everyone else is doing what the target is being asked to do, such as revealing login information. If there is proof, or if the target believes this to be true, it would be very hard to resist the demand.

Other Weaknesses

When someone asked to perform something has very little interest in it, then that person generally has a low degree of involvement. Because such people are detached from the task they are being asked to perform, they may be especially easily influenced by logical reasons for the task, urgency, or authority. Examples of people with low involvement can be security guards, cleaners, or receptionists (Harl, 1997). These groups of people do not care as much about the quality of the arguments, but more about the quantity; the more the better (Harl, 1997). In contrast, people with a high degree of involvement, for example, systems administrators, are persuaded more by the quality of the arguments than the quantity (Harl, 1997).

Another powerful factor to elicit the desired compliance is to use strong affect (Gragg, 2002). If the victim is feeling a heightened sense of anger, surprise, or anticipation, he will be less likely to think through the arguments presented to him. This can be done either by aggravating the mark or simply by surprising him with a demand that is completely unanticipated. Similar to surprise is overloading (Gragg, 2002). When someone has to deal with a great deal of information and does not have enough time to think about it, the ability to think critically about the situation is lowered. An example of this would be to present and require a lot of technical information from a person with very little technical knowledge. The basis for all of the more advanced deception tricks is to use deceptive relationships (Gragg, 2002). A

very powerful psychological trigger is establishing a relationship with someone, solely for exploitation. This can be done effectively by sharing information and a common enemy, as previously discussed under liking. The attacker does this by using techniques, for a long time, to create rapport, actually building up a (false) relation with the target, befriending her, and then slowly starting to use the relationship for nefarious gain. This technique was especially popular with foreign intelligence services, as it also leads victims to rationalize their actions internally, thus being more committed to the case.

How to Act When Influencing Others – a Practical Example

The above techniques and examples might sound convincing, but in order to illustrate how they can be used, an example follows. It is based on the premises that a perpetrator either wants to obtain information, in the form of login information, from the mark (victim), or to get the mark to perform some action at the perpetrator's request. This is a classical social engineering attack. The techniques should work best when trying to influence someone from a Western culture. Many of the same techniques can be used against persons from other cultures too, but they might, due to cultural differences, be ineffective or even insulting (Levine, 2003).

The perpetrator begins by either creating a person of authority, or by exploiting existing relationships. If the perpetrator knows the mark, or someone who knows the mark, he can use this, or make it up, but a actual reference is far more useful. If that is not possible, the perpetrator may create a person of authority, such as a doctor, researcher, or some other successful person, as suggested above. In this case, the attacker chooses to be a systems administrator:

The attacker describes himself as a senior systems administrator (Authority) from a high profile consultancy firm hired to investigate critical network problems of the organization (Scarcity). He phones the mark, introduces himself, notes the accent of the target, and asks where the target is from. Whatever city the mark answers, the attacker's wife is from the same town (Similarity). He then asks if the mark would consider spending a couple of minutes helping him fix the network (Commitment). The attacker then starts to describe the problem with the network, by using technical jargon and ample statistics (Authority). He explains that the mark's computer must be taken offline for a couple of days, maybe a week, to fix the problem. This is if the mark cannot help them with some technical services, the way many of his colleagues have today (Social proof), notably by typing in an increasingly complicated series of commands in the DOS-prompt (Overloading). The perpetrator then offers to do the mark a favor by fixing the problem, as he is to leave for a week's vacation in a couple of minutes (Scarcity). The mark must do a small favor for the perpetrator (Reciprocation), which is not to tell

anyone of this, because it could lose the attacker his job due to the mark's really strict boss (Liking, by finding common enemy). The best way of fixing the situation is for the mark to bring his computer to the fictitious office of the attacker, just an hour away by car, to bring his personal ID papers, and a signed letter of recommendation from a co-worker, as well as a written history of what the mark has done with his computer over the last year, as is the policy in the consultancy firms. Otherwise, perhaps, if it can be kept just between them, the mark could just give the attacker his login information (Contrast).

This is a simple, and classic, example of a social engineering attack. As demonstrated above, it is, however, deceptively simple, since it uses most of the manipulative techniques available, although it does not delve too deeply into any one of them. Against the right kind of mark, using the right kind of setting, this attack is highly efficient.

How the Attacker Can Be Persuasive

Levine (2003) believes that there are three key elements to being persuasive as a person. They are authority, honesty, and likeability. The other techniques described above can be used to strengthen influence, but on an interpersonal level only these three are crucial. There are some easy ways to strengthen the way the mark perceives the offerings of the attacker in these elements.

If actually meeting the mark face to face, it is always important to maintain eye contact. This will make the attacker seem far more honest and authoritative. While maintaining eye contact, it is also useful for the attacker to act as if he is engrossed in what the mark is talking about. It is, however, not good if the attacker actually is engrossed, as this will limit his perception. When preparing for an attack, the perpetrator will carefully consider the clothes he will wear. They are a kind of uniform, signaling authority, and will be carefully selected to reflect the particular kind of authority the attacker aims for. Classic examples of this are doctor's coats, police uniforms, but it should not be forgotten that normal clothing is also a kind of uniform. For instance, a nice tailored suit sends a different message than a "nerdy" Linux t-shirt, although either is efficient as a uniform in a particular context.

Speaking with confidence and using ample technical jargon will make the attacker seem more knowledgeable, and therefore more authoritative, especially if the mark knows little about the area the perpetrator is talking of. The same is valid with regard to statistics; therefore, the attacker can use them to further his argument, as people tend to believe more in arguments supported by statistics, even if the statistics are false or irrelevant. The attacker should also always show both sides of the argument, as this will make him seem

less pushy and more honest. The attacker will try to find similarities with the mark, such as the same hobbies, kids the same age, have relatives in the mark's hometown, and so on. He will also mimic the behavior and speech patterns of the mark somewhat. This builds rapport, which leads to liking.

Be wary of new acquaintances displaying several of the above characteristics.

Defending Against Deception

In this section, you, the reader, is given a concrete set of tips on how to avoid being influenced. While simply reading about the techniques and vulnerabilities presented in this chapter will make you more resistant to manipulation, simply theorizing around the concepts is of limited use to organizations and those responsible for security. Levine (2003) suggests two basic approaches to enhance resistance. The first is "the sting". People are put in situations where they are influenced into acting against their own preferences, and when the people comply, they are informed of the influence tactic and what has just happened. This has the benefit of pushing the subjects out of their comfort zone, making their vulnerability more obvious to them. What is critical here is that the subjects should be made to acknowledge their own personal susceptibility (Levine, 2003).

The second method is a little less intrusive than "the sting", and more manageable in a business context. The goal here is to expose the subjects to weaker forms of persuasions, which then act much as an inoculation does in an immune system; they prepare it for the real threat. The most important issue for consideration here is obtaining support both from management and in the information security policy for such efficient counter measures as "the sting" and inoculations. When support has been acquired, a small roll-out, especially of inoculations is preferred, and in high risk scenarios, stings can also be enacted. While it can sound cruel and unethical, it is also one of the easiest ways of practicing some kind of resistance to these attacks. Deception against one's own employees has been used, with some success, at both West Point Military Academy (Dodge & Ferguson, 2006), and the New York State (Bank, 2005). In the West Point case, students were sent an e-mail from a person claiming to be a Colonel, ordering them to click on an attached link to verify their grades. This approach received 80 % compliance among the students, who were later informed of the risks of their acts. In the case of the New York state, 15 % of the employees tried to enter their passwords into a special online "password checker" after receiving an e-mail from the "Office of Cyber Security and Critical Infrastructure Coordination", urging them to do so. A follow-up of this a couple of months later, using a similar approach, received a lower compliance rate (8 %).

In order for you to be better prepared for attacks using any of the specific vulnerabilities discussed above, a short guide of defenses is given below.

The best method of defense against authority is to remove the element of surprise from it. Be suspicious of authoritative power, and remember the influential power of authority. There are two questions that might help with this: is the person really someone of authority? If yes, then how truthful do you think that person is (Cialdini, 2003)?

While the scarcity principle is easy to learn about, it is hard to counter. One method is trying to learn how to recognize the feeling when the competitive cogs in our brains start to whirl, but this might not be enough. Learning to think about the scarce object from a more utilitarian standpoint can also help. Do we want the object because it is rare, or do we believe that the object will be better because it is rare? Then we should remember that rare things are rarely better (Cialdini, 2003).

Due to the vast spectra of possibilities that influence liking, it is hard to develop a broad spectrum of defenses. Instead, Cialdini (2003) recommends a simple approach. Allow yourself to be swept away by the liking of others, but when it comes to decisions, consider how long the person who is asking you to make a decision has been in contact with you, and whether or not you like him to a reasonable extent based on this time. If you adore someone who is trying to get you to give him some information after only knowing him for a couple of minutes, there is probably foul play in the works.

Reciprocity is a very effective influence technique, and very hard to defend against. A overly strict rule against accepting any kind of gifts will make you seem socially awkward. A more efficient method is to redefine the gifts given into their real meaning. A sales person giving you a gift is really exposing you to marketing, and thus you do not need to return the favor. A stranger on the telephone offering to help you with something you did not request or know that you needed is most likely up to no good.

To protect against consistency, you also have to reach inside yourself. Cialdini (2003) mentions two kinds of methods for spotting when someone is exploiting your consistency. The first is identifying when we get the feeling that we are being pushed into performing actions we know we do not want to perform. The second method is considering whether or not we would make the same commitment, if we could travel back in time.

Social proof is something very often useful to you. In fact, in most new situations you would be well advised to follow the behaviors of others. There are, however, certain situations when social proof can lead to you being tricked into performing harmful acts. In order to avoid this, there are two strategies. Be aware of what are obviously faked situations, such as those found in advertisements showing groups of people praising a product, or

when someone claims that the people around you are doing things you doubt are being done. It should also be remembered that the actions of others are not to be taken as the sole reason for your actions (Cialdini, 2003).

Future Trends

This is an area that has been studied extensively in other fields of science than information security. There is ample material in areas ranging from literature to social sciences, marketing and economics. While there is a broad range of researchers working in the field, the area of information security remains relatively unexplored. Although problems with software, networks and other technical artifacts will no doubt be of significant importance in the foreseeable future, there is a growing trend with regard to more human aspects of information security. It is notable that an industry icon, such as Bruce Schneier, has started to become interested in the field, and there are emerging academic conferences such as the HAISA (Human Aspects of Information Security & Assurance) conference.

One of the challenges for those of us working in the field of the human element of security is how we can argue both that our work is important and whether or not it actually improves security. This has always been easier for technical products, as they often can argue efficiency based on statistics. When the buyer has to choose from a product promising a 99.99 % protection against “millions” of computer viruses, or an education program that might prevent one case of social engineering, the choice is often simple for the purchaser. It is thus, sadly, our task as researchers, professionals and students to help point out that the single attack might very well be the most damaging attack imaginable, far more than a random virus attack.

The increased attention gained in this field will probably bring greater awareness, among both professionals and ordinary users. When users become more resilient towards the easy ruses, such as those described in this chapter, the attackers will have to either become more advanced themselves, which is quite hard due to the increased complexity of the skills needed, or find other ways to attack. If perpetrators have to develop their skills of influence to a level high enough to affect even people who are well aware and trained against such techniques, they might as well leave the field of crime and seek more lucrative employment as influence professionals, such as salesmen or politicians.

Conclusion

This chapter has shown how easy it is to use influence to get people to do things that they may not want to do. The goal has been to give concrete examples of well-established techniques and methods, together with practical

uses. Hopefully the reader now has a greater insight into both the manipulation techniques used by computer criminals and the techniques used by the everyday deception professionals.

One of the major points of this chapter is just how easy the techniques are to learn and to implement. In fact, just by reading through this chapter, you, the reader, probably now have most of the tools needed to influence people around you to a far greater extent than before. This is, of course, knowledge that should be used with some caution. While in most cases it is rather easy to influence people, the counter reaction from those who have just understood that they have been manipulated is generally rather severe. Good relationships are not built on deception.

Nevertheless, there is a lot of merit in using deception, albeit on a small scale, against one's users and subordinates in an organization. As long as deceiving one's own employees and co-workers is practiced with afterthought and a clear goal, with ample feedback and information given, it might serve well as an educational and training tool. Do remember that these are the techniques the bad guys are using. If we do not prepare against the methods used, we will easily fall victims to them.

References

Bank, D. (2005). 'Spear Phishing' Tests Educate People About Online Scams. *The Wall Street Journal*. Retrieved March 2, 2006, from: http://online.wsj.com/public/article/SB112424042313615131-z_8jLB2WkfcVtgdAWf6LRh733sg_20060817.html?mod=blogs

Blass, T. (2002). The man who shocked the world. *Psychology Today*. Retrieved March 9, 2006, from: http://www.psychologytoday.com/articles/pto-20020301-000037.html

Bowyer, B. (2003). Toward a Theory of deception, *International journal of. intelligence and counterintelligence*, 16, 244-279.

Brostoff S., Sasse A. & Weirich D. (2002). Transforming the "weakest link": A Human-computer Interaction Approach to Usable and Effective Security, *BT Technology Journal* 19(3), 122-131.

Cialdini, R. (2001). *Influence: Science and Practice*. Needham Heights, MA: Allyn & Bacon.

Dodge, R., & Ferguson, A. (2006). Using Phishing for User Email Security Awareness. In Fischer-Hübner, S., (Ed.), Rannenberg, K. (Ed.), Yngström, L. (Ed.), Lindskog, S. (Ed.), *Proceedings of the IFIP TC-11 21st International Information Security Conference (SEC 2006)* (pp. 454-458). New York, NY: Springer Science + Business Media Inc.

Feer, F. (2004). Thinking About Deception. Retrieved March 11, 2006, from: http://www.d-n-i.net/fcs/feer_thinking_about_deception.htm

Gragg, D. (2002). A Multi-Level Defense Against Social Engineering. *SANS Institute*. Retrieved September 17, 2003, from: http://www.sans.org/rr/papers/index.php?id=920

Granger, S. (2001). Social Engineering Fundamentals. *Security Focus.* Retrieved September 18, 2003, from: http://www.securityfocus.com/printable/infocus/1527

Harl (1997). The Psychology of Social Engineering. Retrieved March 12, 2006, from: http://searchlores.org/aaatalk.htm

Levine, R. (2003). *The Power of Persuasion*. Hoboken, NJ: John Wiley & Sons Inc.

Marett, K., Biros, D., Knode, M. (2004). Self-efficacy, Training Effectiveness, and Deception Detection: A Longitudinal Study of Lie Detection Training, Lecture Notes in Computer Science, Volume 3073, Jan 2004, pp. 187 – 200.

Martin, B. (2004). Telling lies for a better world? *Social Anarchism*, 35, 27-39.

Mitnick, K. & Simon, W. (2002). *The Art of deception: Controlling the Human Element of Security.* Indianapolis: Wiley Publishing, Inc.

O'Connor, J. & McDermott, I. (1996). *Principles of NLP*. London, UK: Thorsons.

The Cycle of Deception – a Model of Social Engineering Attacks, Defenses and Victims[1]

Marcus Nohlberg and Stewart Kowalski

Abstract

In this paper we propose a model for describing deceptive crimes in general and social engineering in particular. Our research approach was naïve inductivist and the methods used were literature study and interviews with the lead investigator in a grooming case, as we see many similarities between the techniques used in grooming and those used in social engineering. From this we create cycles describing the attacker, defender, and victim, and merge them into a model describing the cycle of deception. This model is then extended into a possible deception sphere. The resulting models can be used for educating about social engineering, creating automated social engineering attacks, facilitating better incident reporting, and for understanding the impact and economical aspects of defenses.

Key words: Social engineering, fraud, deception, security models, computer crime

[1] A version of this paper was published in *Proceedings of the Second International Symposium on Human Aspects of Information Security and Assurance (HAISA 2008)*, Plymouth, UK, July 2008, ISBN: 978-1-84102-189-8.

Background

Social engineering is a term for techniques used to con, or trick, victims into giving an attacker sensitive information, or that make them perform actions at the behest of the attacker. A good definition is given by Kevin Mitnick in an interview by Tanneeru (2005):

> "Social engineering is using manipulation, influence and deception to get a person, a trusted insider within an organization, to comply with a request, and the request is usually to release information or to perform some sort of action item that benefits that attacker. It could be something as simple as talking over the telephone to something as complex as getting a target to visit a Web site, which exploits a technical flaw and allows the hacker to take over the computer."

Since many users do not believe that anyone would ever trick or con them, because they are not "rich and famous", and that hackers "cannot do much damage anyway" (Brostoff, et al. 2002) these attack techniques are often quite successful. This is further complicated by the fact that most users do not understand how security works, and therefore construct their own, often incorrect, models (Adams & Sasse, 1999). There are a lot of studies on the "gullibility" of users, both academic and non-academic. One example is a study conducted by Treasury Department inspectors, where one third of the Internal Revenue Service (IRS) employees gave away their logins and passwords to auditors who called pretending to be computer technicians (Dalrymple, 2005). This and several other studies demonstrate a high degree of susceptibility to social engineering attacks.

Social engineering as a term has been used for some time in the security sector. We have found references dating back to 1995 (Winkler & Dealt, 1995), and there was, of course, the boom of notoriety of social engineering in connection with the case of Kevin Mitnick (Mitnick & Simon, 2002). His warrants, and later imprisonment, led to a great deal of publicity for the technique. One of the more influential contributions from Mitnick in an academic setting is the social engineering attack cycle, SEAC, as seen in Figure 1 below (Mitnick & Simon, 2002). This descriptive cycle is frequently used both by security professionals and academics when describing social engineering attacks.

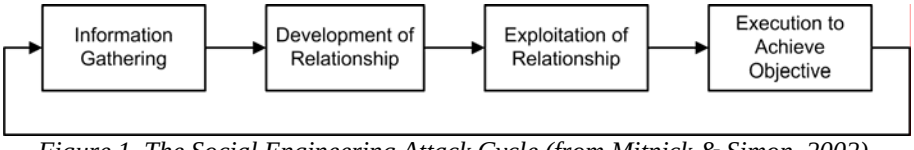


Figure 1. The Social Engineering Attack Cycle (from Mitnick & Simon, 2002).

In this paper we propose a new conceptual model of the social engineering attack cycle, which includes descriptions of both defenders and victims. This

new cycle addresses the predominant problems with the existing model. We find that the existing model is overly simplistic, while, at the same time, it is obscure. It is quite common that those using the model put too great an emphasis on describing a step-by-step approach while giving little support for the iterative reality of most attacks. SEAC does not provide any suggestions for proper protection, making the model of limited use. Our aim is to provide a new model of the social engineering attack cycle that can be used as both a teaching aid and a framework for developing a holistic protection strategy.

The paper is structured with an introductory section on the research area and problem followed by a discussion on the method used. The model is then presented and the paper concludes with a brief discussion concerning the strengths, weaknesses and possible use of the model.

Method

The general approach in this study was of the naïve inductivist (Kowalski, 1994) as illustrated in Figure 2.

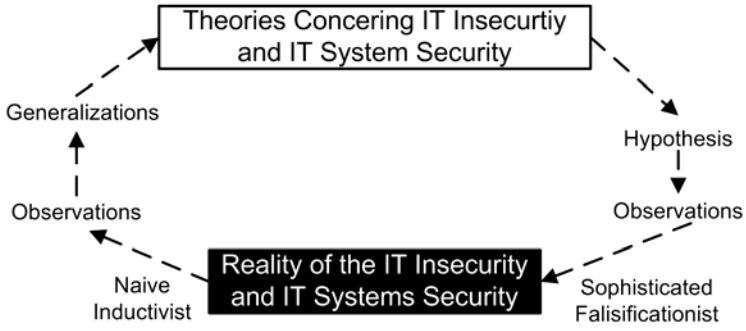


Figure 2. Naïve inductivist and sophisticated falsificationist approaches to studying IT systems Security and IT insecurity (Kowalski, 1994, pp. 4).

We started out observing the problem for the need of an improved model of social engineering attacks and the assumption that this could be created using the body of knowledge existing, but also by studying actual crimes. The study was made on criminals using grooming, and a new model was created using generalizations from the data gathered. The study was based on literature and interviews. The literature survey has been ongoing for several years, and has covered a large part of the written material on the subject. This has given an understanding of the concept and what it comprises, but also of the current misconceptions and areas where improvements are needed. In order to understand more about how attackers actually use social engineering, we set out to find public, well documented, cases of social engineering. This proved to be difficult. The attacks were either poorly documented or purely anecdotal. This is a common problem with security research; it is difficult to

obtain access to well documented data about successful attacks, as most organizations tend to keep those secret, if they even know about the attack. In order to address this problem we looked outside of traditional social engineering attacks for other crimes using similar patterns and techniques. We found that the attack patterns used in grooming matched social engineering. Grooming is the term used when an adult, most often a man, tries to convince a child or a minor to commit sexual acts (O'Connell 2003). The advantage of studying grooming is that once grooming is reported to the police, a thorough, and often public, investigation is conducted. In our study we chose to focus on the most infamous Swedish groomer, who during a period of several years developed an ever increasingly efficient method of grooming using a combination of advanced manipulative techniques and technology (The Local, 2005). In order to study this case we analyzed the public records of the trial, but also conducted telephone interviews with the lead investigator of the case. The first was an open interview aiming to obtain background information of the case. The second interview was semi-structured, as described by May (2001), and developed with the SBC-model (Kowalski, 1994) as a foundation for the questions asked. The SBC-model is a model of information security that covers both technical and social aspects of security. The questions were designed to gain specific knowledge while remaining open enough to enable us to ask complimentary questions and have a dialogue with the subject. The subject was informed about the usual ethical issues before the interview and, in order to ensure correctness, given the opportunity to read the transcripts. A couple of minor misunderstandings were subsequently corrected.

From the data gathered by the interviews, the steps in the attacks were identified and described. By analyzing these steps and using knowledge gained from the literature study, a model for the attack cycle was developed, followed by the victim cycle. They were complemented by a defense cycle adapted from Kowalski (2002). These three cycles were then merged into a complete model, which was later discussed with security professionals and especially well received by students when used as a teaching aid.

Results

When studying the grooming attacks, it was apparent that the attacks had eight steps. The steps have been somewhat simplified in the description below, but the general sense is maintained.

1. Create a pretext, in this case a fictional female model.
2. Contact the victims, or set up web pages so the victims could contact the attacker believing he was the female model.
3. Obtain interest, gather information and get compliance from the victims.

4. Move the victim to an unfamiliar location physically, use information gathered earlier.
5. Carry out the crime, in this case sexual abuse or rape.
6. Contact the victims again afterwards, get the victims to agree that what happened was OK.
7. Try to recruit the victim to find new victims.
8. Relive the crime, from stored data/pictures, and deny that any crime was committed if confronted, for instance, by parents or the victims.

Some of the steps in the description above are specific for grooming, but five generalizable steps in the attack cycle are apparent and used for the attack cycle.

The attack cycle

The attack cycle concerns the behavior of the attacker, and the actions he or she will take in an attack. The stages of the attack cycle are described in Figure 3.

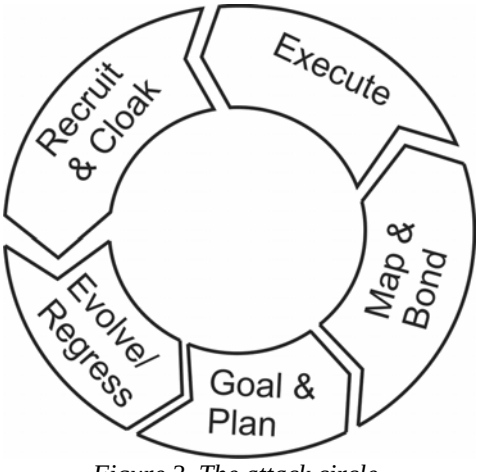


Figure 3. The attack circle.

Goal & Plan: The attacker must have a purpose with the attack, a goal, and a plan how to reach it. This is where traditional criminological knowledge comes into play. The four classic traits that the attacker must possess are method, motive, opportunity, and means (Pfleeger & Pfleeger, 2003). In order to be able to carry out an attack, the perpetrator must know what kinds of attacks are possible, the methods. Some methods are obvious and require no great cunning or planning, while others require certain skills or knowledge. There are three basic ways to acquire this knowledge. The method might be known beforehand, it could be searched for specifically to use for

attacks, or it might be found by chance. The perpetrator could discover an attack method that works well on the first try, or he might find a book or text describing attacks without any prior intention of performing an attack. It is notable here that Sunderland's Differential Association Theory (DeMelo, 2007) states that once a potential perpetrator learns the methods required, he or she can easily pick up the required motive from just about anyone. Therefore, by learning the methods required it is probable that the perpetrator will also find the motives needed. The criminal culture, as discussed by Ferrell (1995), can be seen as the major factor determining crime. In fact, one of the flaws of traditional criminological reasoning is that the contemporary culture is sometimes neglected in the consideration of criminological analysis. The criminal subculture spans more than simply proximity, something that is ubiquitous in a connected world, it also concerns motives, drives, rationalizations and attitudes, as well as certain appearances, group specific language and self presentation, and style (Ferrell, 1995). *Map & Bond:* The stage in which the attacker tries to obtain information needed for the attack. This can be done by using traditional social engineering techniques, such as dumpster diving or desktop hacking, or by searching the web for data and studying other open sources of information. It can also, however, be when the attacker befriends the victim or someone with usable knowledge, and uses manipulative techniques to get him or her to divulge the information needed, or to "prepare" the victim for the next step. In order to create a deceptive relationship, the attacker uses influence techniques, for instance, authority, scarcity, liking and similarity, reciprocation, commitment and consistency, social proof, and involvement (Cialdini, 1993). The influence techniques then exploit certain social psychological weaknesses, as suggested by the taxonomy put forth by Jordan and Goudey (2005). In other words, the victim is manipulated into trusting the attacker. *Execute:* The execute-step is where the attacker does something clearly illegal or not allowed, for instance, asking the target to submit his or her login information, or the sending of the nefarious e-mails. *Recruit & Cloak:* Cloak is the term for the actions performed after the execution, actions that hide the illegal activities. Such an action can be continuing with the "friendship" to normalize the illegal activities, moves to make the victim seem untrustworthy, or more advanced techniques to hide the crime. In some cases, the victim can be recruited to either work for the attacker or act as the perpetrator's ambassador/reference. *Evolve/Regress:* This is where the attacker learns from the process and creates an internal justification for what has happened. At this stage, there are basically two choices for the attacker. Either the attack evolves, moving into another phase, if the process has been successful thus far, or, if the results have been unsuccessful, the attack regresses, which is either to stop the attack or move it to a more basic level in order to try for success again.

The defense cycle

The defense cycle describes the general options available to the defender, who could, in some cases, be the same person as the victim, or security professionals in an organization or similar. This section is based on the work of Kowalski (2002), from which the terms and definitions have been taken and the flow identified.

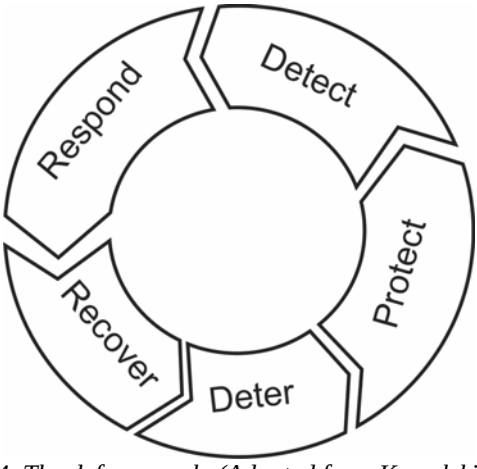


Figure 4. The defense cycle (Adapted from Kowalski, 2002).

The description of defenses given by Kowalski (2002) has been adapted to a circle to match this model, as shown in Figure 4 above. Several examples of implementations, which can, of course, consist of many other measures, are given below. The description is based on what the defender must do to be successful in providing defenses. Having a good, public policy, or a reputation of reporting incidents to the police, can *deter* an attacker. In addition, making little sensitive data available, educating employees about the risks and methods of attackers who try to bond with them, as well as providing a strong policy on how to act, are measures that *protect* the organization. Running a surveillance of the network communication, can reveal when sensitive data are being sent or accessed, and having well-educated employees who know when they are being asked illicit questions, helps to *detect* an attack. Furthermore, making it easy, and without any social or professional stigma to report social engineering incidents, and making the employees aware of how they can be manipulated into acting on behalf of an attacker, enable a defender to *respond* to an ongoing attack. Also, knowing the value of your data, reporting attacks and having a well-designed policy, means that a victim can *recover* from the attack and learn from it. Hopefully the attacker can be found and prevented from evolving and attacking you, or others, in the future.

The victim cycle

The victim cycle is focused on the behavior of the individual victim, the person, in the attack. A common mistake when analyzing crime is to focus too much on the attacker, and to forget the victim. In fact, many crimes can be more readily prevented by focusing on the victim rather than the attacker. The flow is described in Figure 5.

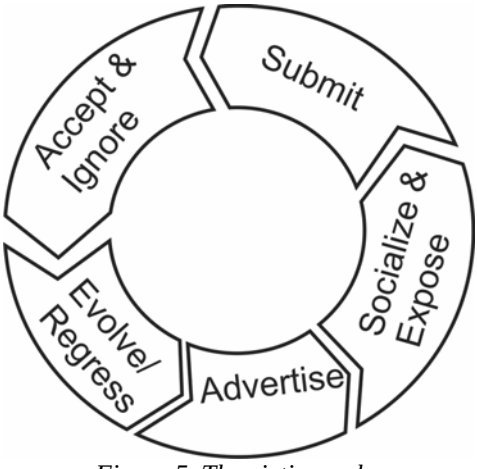


Figure 5. The victim cycle.

By having something of value and making it known, either knowingly or unknowingly, the victim *advertises* its suitability as a target. Furthermore, by *socializing* with the criminal, the victim sets itself up for deception, and *exposing* valuables makes them accessible to the attacker. When the actual crime is being executed, the victim *submits* to it, for instance, by giving out the secret information. After the crime has been executed, the victim can choose to *accept* it, for example, through believing that it was not so "serious", or simply by *ignoring* it, either knowingly, or by actually being unaware of it. The victim can learn from the crime and *evolve* into someone who is harder to victimize in the future. However, it is also possible that the victim can *regress*, turning into someone who accepts the role of victim and becomes easier prey in the future.

Adding the element of control

The attack circle could be perceived as the attackers' efforts to reach the target in the center, and the way to reach the center is through increasing control. Once the attacker has gained enough control through the process, and the risk has been reduced to an acceptable level, according to the attacker, the attack can be performed. The level of acceptable risk is individual for each attacker. This is illustrated in Figure 6.

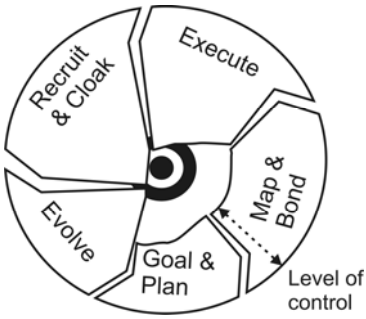


Figure 6. The attack cycle with added control.

This aspect of control is also true for the victim and the defender, where they both strive to have a higher level of control than that of the attacker.

The cycle of a social engineering attack

When merging the three different cycles and adding a target in the center, a more holistic view of the prerequisites of a social engineering attack appears, as illustrated in Figure 7. One of our theories is that in order to have a "successful" social engineering attack, all the steps in all the cycles have to fall into place. The attacker needs to succeed with the first three steps of the attack in order to be successful, and with the fourth and fifth to be able to continue attacking in the future. This is based on the reasoning that if the attacker is unable to provide a plan and a method for the attack, he will most likely fail. If he cannot learn about the potential victim, or perform the attack, he will fail. In addition, if the attacker is unable to conceal the attack, he will most likely be caught, and, if the attacker, through internal rationalization, judges that the attack was not a "good" experience, he will most likely not continue.

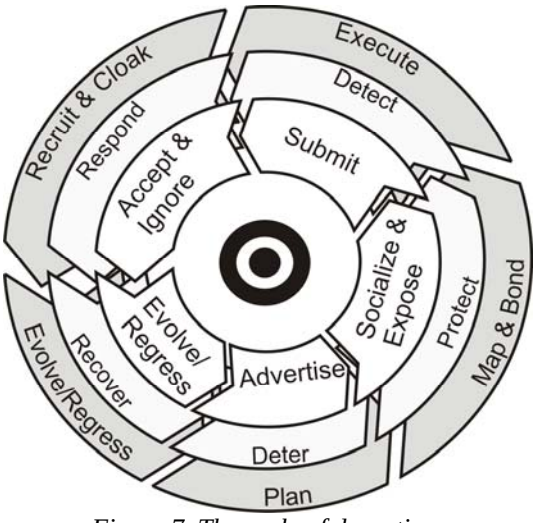


Figure 7. The cycle of deception.

The same is true of the defender. If any one of the steps in the defense cycle is good enough to stop the attacker, then the attack will obviously fail or lead to the attacker being caught. In contrast, if no single part of the cycle can stop the attacker, then the attack will not fail due to the activities of the defender. Looking at the victim cycle, we assume that the victim must submit in each of the sections in the model for the attacker to succeed. There are perhaps exceptions to this assumption, but based on sound reasoning it should be mostly true. This gives a possibility to consider the economic impact of this model. Choosing to invest in sections that are among the first three will stop the crime from happening. Investing in the fourth and fifth might stop the crime from happening in the future. When considering purchasing defenses or educating the users, considering where to place that particular investment in the cycle of deception is relevant. This knowledge enables us to see if the investment will have the intended consequences.

The spherical view of deception

The earlier descriptions presented here are simplified and omit the fact that most attacks span several cycles and include several smaller attacks that are parts of the larger attacks. Therefore, there is a need for a third dimension in the description. This added information is useful, for instance, when trying to create a piece of software using this model, but also to illustrate a more complete image of the attack.

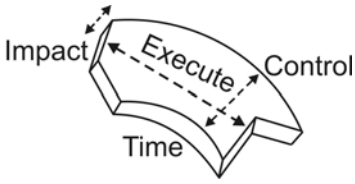


Figure 8. A 3D view of the Execution Phase.

The additions in Figure 8 above are the element of control, the time, and the impact. The impact refers to how noticeable the attack is for the victim and controlling organization. The goal for the attacker is to keep the impact as low as possible, while maintaining a high level of control in a short time. The separate cycles in the whole attack each belong to one of the general attack cycles. For instance, there might be several smaller attacks performed in order to facilitate the first step (Goal & Plan) of a larger attack. Using this imagery and extending it into a more holistic description, we obtain the spherical view of attacks, presented in Figure 9.

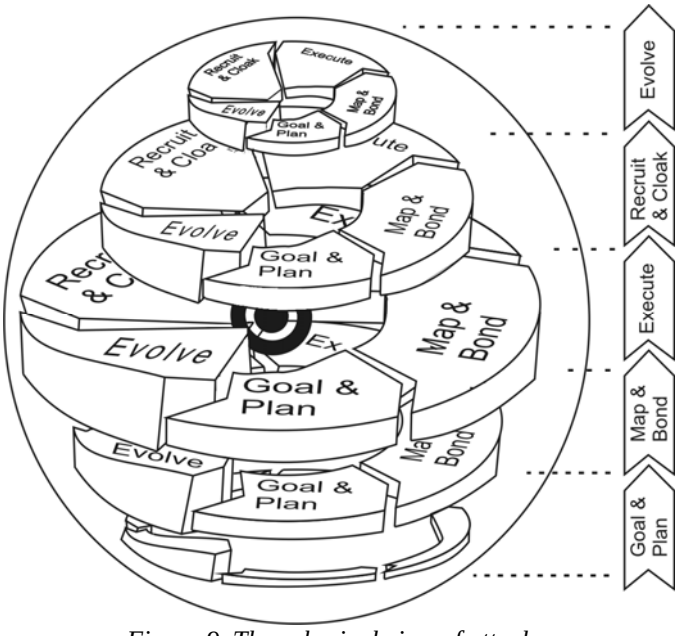


Figure 9. The spherical view of attacks.

This model adds a third dimension to the description. The target is reached in the centre of the sphere, in the execution phase.

Discussion

This proposed model has several uses. It is can be used for education in social engineering. As the model covers activities by all involved – victims, attackers, and the protection organization – it provides a holistic and unders-

tandable view. It also facilitates a deeper understanding of the criminal process, and perhaps, most importantly, it facilitates understanding how to develop a protection strategy. This model is a good tool for prioritizing and understanding the implications of spending in a certain area while neglecting other areas.

From an academic point of view, this model presents an excellent starting point for security researchers trying to position themselves. When looking at this model from a computer science perspective, it can be used to facilitate the implementation of an automated social engineering AI-bot. By using this model, the attack flow of the AI-bot, as well as the target points, that is, where the AI-bot has gained enough information to move on, can be implemented. This allows for the creation of sophisticated AI-bots which in turn can be used to train users to avoid falling for real attacks, as well as be used for social engineering penetration testing. Security professionals can use the model when studying and visualizing the readiness of an organization. The model can also be used to investigate and understand attacks made against the organization, and to make it easier for the potential victims to report their experiences, as they can be guided in the process by the steps involved in this model. This can provide improved incidence reports, something that is quite important in fighting social engineering.

One of the potential problems with this model is that it is developed from a study of grooming, which is not the same crime as social engineering. After comparing what is known about social engineering and our case of grooming, it is, however, apparent that these two attacks share many of the same methods, manipulative techniques, and the same flow of the attack, even if the end goals and motivations are quite different. In fact, our model might be valid and have the same merits for other crimes and nefarious acts purposefully carried out by rational perpetrators. The model has a slight focus on examples given in traditional social engineering, with an attacker having interaction with a victim, but it should also be valid for more technical attacks, such as phishing. The major difference is that those attacks often have a shorter timeframe, so the parts of the cycle might be moved through quicker. Whether an attacker calls or uses a spear-phishing attack makes no major difference for the cycle, however. The attacks that are not covered by this model are those that are based almost purely on random successes from exposure to large numbers of users. For instance, the most basic form of phishing, or malware, where the attack is a generic and non-specific message used against a group of non-tailored victims, obviously has very little “Map & Bound”, apart from the message sent out to deliver the attack. The model could be argued to cover these attacks too, but we feel that the greatest use of the model comes when applying it to crimes with a more specific intent than random frauds.

In the future this proposed model could be studied further, and analyzed in more types of attacks and crimes than social engineering to validate a broader use for it. There is a possibility that this model can help us understand and prevent more crimes than social engineering.

References

Adams A. & Sasse M. (1999), Users are not the Enemy: Why users compromise computer security mechanisms and how to take remedial measures, Commun. ACM 42.

Brostoff S., Sasse A. & Weirich D. (2002), Transforming the ”weakest link”: A Human-computer Interaction Approach to Usable and Effective Security, BT Technology Journal 19(3), 122-131.

Cialdini, R. (1993), Influence: the psychology of persuasion. New York, Quill, ISBN: 0688128165.

Dalrymple, M. (2005), Auditors Find IRS Workers Prone to Hackers. [Online]. AP. Available from: http://www.infosecnews.org/hypermail/0503/9684.html, (Accessed 6 Mar 2006)

DeMelo, D. (2007), Sutherland's Differential Association. [Online]. Available from: http://home.comcast.net/~ddemelo/crime/differ.html, (Accessed 29 Jan 2007)

Ferrell, J. (1995), Culture, Crime, and Cultural Criminology. Journal of Criminal Justice and Popular Culture, 3(2) (1995) 25-42.

Jordan, J. & Goudey, H. (2005), The signs, signifiers and semiotics of the successful semantic attack. Presented at the 14th Annual EICAR Conference, St.Juliens/Valletta, Malta, 2005.

Kowalski, S. (1994), IT Insecurity: A Multi-disciplinary Inquiry. Diss. University of Stockholm. Report series No. 94-040, Stockholm.

Kowalski, S. (2002), Value Based Risk Assessment: The Key to a Successful Security Target for the Telecommunication Industry, 3rd International Common Criteria Conference (ICCC) Ottawa, 2002.

May, T. (2001), Social Research: Issues, Methods and Process. Buckingham: Open University Press, ISBN: 0335206123.

Mitnick, K. & Simon, W. (2002), The Art of deception: Controlling the Human Element of Security. Indianapolis: Wiley Publishing, Inc., ISBN: 076454280X.

O’Connell, R. (2003), A typology of child cybersexploitation and online grooming practices. [Online]. University of Central Lancashire: Preston. Available at: http://www.uclan.ac.uk/host/cru/docs/cru010.pdf, (Accessed 11 Dec 2007)

Pfleeger, C. & Pfleeger, S. H. (2003), Security in Computing (3rd ed). Upper Saddle River: Prentice Hall, ISBN: 0130355488.

Tanneeru, M. (2005) A convicted hacker debunks some myths. [Online]. CNN. Available at: http://edition.cnn.com/2005/TECH/internet/10/07/kevin.mitnick.cnna/index.html (Accessed 28 Mar 2008)

The Local (2005), Man sexually abused "at least sixty girls". [Online]. Available at: http://www.thelocal.se/2756/20051228, (Accessed 1 Dec 2007)

Winkler, I. & Dealt, B. (1995), Information Security Technology? …Don’t rely on it A case Study in Social Engineering. [Online]. Proceedings of the Fifth USENIX UNIX Security Symposium. Available at: http://www.usenix.org/publications/library/proceedings/security95/full_papers/winkler.ps, (Accessed 12 Apr 2004)

Non-Invasive Social Engineering Penetration Testing in a Medical Environment [1]

Marcus Nohlberg, Stewart Kowalski and Kerstin Karlsson

Abstract

This paper proposes a soft approach for social engineering penetration testing. By using the SBC model as a foundation, questions related to the social element of security were asked in semi-structured interviews to a group of subjects. The answers were analyzed and presented in an uncomplicated graph. The purpose was to study the feasibility of letting the users participate, instead of exploiting their weaknesses. It was found that the approach of interviewing the subjects rendered interesting, and relevant, results, making it an approach that should be studied further due to its apparent gains: less ethically troublesome penetration testing, increased awareness, improved coverage and novel information as added bonuses.

Key words: Social Engineering, SBC model, Penetration Tests

[1] A version of this paper was published in *Proceedings of the 7th Security Conference*, Las Vegas, USA, June 2008. Ed. G. Dhillon. Washington DC: Information Institute Publishing, USA, ISBN: 978-1-935160-01-4.

Introduction

The paper begins with an introduction to social engineering, penetration testing and the SBC model. It then presents our method and the results, concluding with a discussion of this approach.

Social engineering is the manipulation of humans in order to get them to, more or less willingly, give out information or access to an asset (Mitnick & Simon, 2002). It can be in a simple form, such as simply asking for the information, or in more complex forms with the use of complicated ruses. It might also be to use technical means to improve the efficiency, or the scope of the attack. Phishing is an example of this, where manipulative techniques from social engineering are used in combination with techniques used in spam to form an efficient attack vector affecting thousands, if not millions, of potential marks at once (Jakobsson, 2005).

For a long time, a cornerstone in information security has been the penetration test. From this, new insights can be gained into the weaknesses, and strengths, of the information security. The focus has historically been on testing the network, firewall, and other technical aspects. The dilemmas with penetration testing and social engineering are discussed by Barrett (2003), where the conclusion was that it is preferable to use an audit style which has results and objectives that are clear and can be accepted by both subjects and companies. The testing should also not lead to discipline or dismissal for the individuals. More concrete examples of this are not given by Barrett (2003), however, making it hard to properly judge his suggestions as being other than reasonable.

Another academic approach to social engineering audit was used by Hasle, et al. (2005), who employed an approach to social engineering penetration that tried to test a larger population. They performed two tests; the first was a survey where the users were asked to submit their login information in order to authenticate if they were to win a prize. The second test was an e-mail which, when sent out, triggered a login box. The study by Hasle, et al. (2005) is interesting, but we argue that they were testing resistance to Phishing attacks, rather than what we consider pure social engineering. For instance, in none of their tests was human interaction used. Another approach is the one used in Nohlberg (2005) where the subjects were asked to answer a set of questions under a false pretext (in this case “micro efficiency”), in order to get them to answer somewhat truthfully about their actions in certain security related situations. A more traditional approach to social engineering auditing is argued by Jones (2003), where the auditor is advised to actually perform social engineering attacks on the users. A similar approach is used by Orgill, et al. (2004), where they actually have a person trying to manipulate his way to information from the employees of the tested organization. They do this test in two stages. The first is to let the person wander

around submitting employees to a written questionnaire about security, logins, and so on, and the second stage is to try to gain physical access to the perimeters. Both approaches are disturbingly efficient. The results show that 81 % of the subjects asked revealed their login name. Furthermore, 59 % also disclosed their passwords. Very few employees asked for identification or questioned the auditor. The auditor also managed to get unrestricted, physical access to the building.

Dalrymple (2005) describes the highly successful internal audit on social engineering done by the IRS, where they called a select number of users under some pretext, requesting their passwords, which 35 % of the employees gave out.

The classic approach, as used by Orgill, et al. (2004), definitely has it uses, but the flaws are that it is costly (since it takes a lot of time to perform), and it might be perceived as more ethically questionable among the employees. One can also question the educational aspect. Will tricking a subset of all users make those who were not audited identify with the colleagues who were actually conned and learn from their mistakes, or would they stick to the "lie detection" bias (Marett, et al. 2004), believing that they themselves would not fall for "tricks like that"?

For the audit professionals, Information Systems Audit and Control Association, ISACA (2004) gives a list of areas that should be tested when doing a social engineering audit. They suggest that the four areas to test are:

- Test of Controls – a general overview of the organization, can give a basic knowledge usable in further tests.
- Telephone Access – to use a set of well known attacks to test the resistance of the organization against attacks over the telephone.
- Garbage Viewing – to see if there is any sensitive information being thrown away (dumpster diving).

Desktop Review – Check the user's workplace. Merge the data from the social engineering audits with other audits.

The guidelines given by ISACA (2004) present a basis for testing that could be perceived as ethical, at least by the organization. However, in our opinion, the attacks suggested and the general set-up seem to provide little useful data. In addition, the approach leaves out any user input and feedback from the subjects, thus limiting the knowledge gained to that from a small set of attacks, and the learning experience of the subjects is absent. The approach also only tests how the users would act in a specific setting and does not give any general information about values, knowledge, attitudes, and so on.

One of the novel approaches that inspired us was Stanton, et al. (2005), where they did a survey asking employees about their views on security, and

then contrasted the results with an actual audit done by experts. The results from Stanton, et al. (2005) indicate that one of the key success factors lay in the clear communication that the management takes security seriously and that the employees believe they are accountable for their own actions. While this approach covers more areas of information security than the social, and was done in a survey, the idea was inspiring.

The SBC-model

In order to create questions on the social element of security, we needed a model describing security that covered more than technical issues. In our opinion, the frequently used SIS-model (SIS, 2003) seems to be more a model of the hierarchical set-up of an information security organization than a good description of the term, information security. It is especially lacking in the sections covering social elements of security, as they are not included. In fact, the efficiency of social engineering attacks against an organization that has modeled its security closely on the SIS-model would probably be great, due to the attacks falling “between the cracks”. A preferred model of security is the SBC (Security by Consensus) model proposed by Kowalski (1994), which gives a more useful description of security, as seen in 0. In the SBC model, a greater emphasis is put on a holistic approach, thus including the social aspects. In the SBC model, the owner or user of a system is perceived to create opportunities of becoming a victim by not protecting the systems they use or own. It is notable here that the perpetrators are not included in the model due to the fact that it is almost impossible to collect enough data on the perpetrators to enable a crime prevention program for IT crime (Kowalski, 1994).

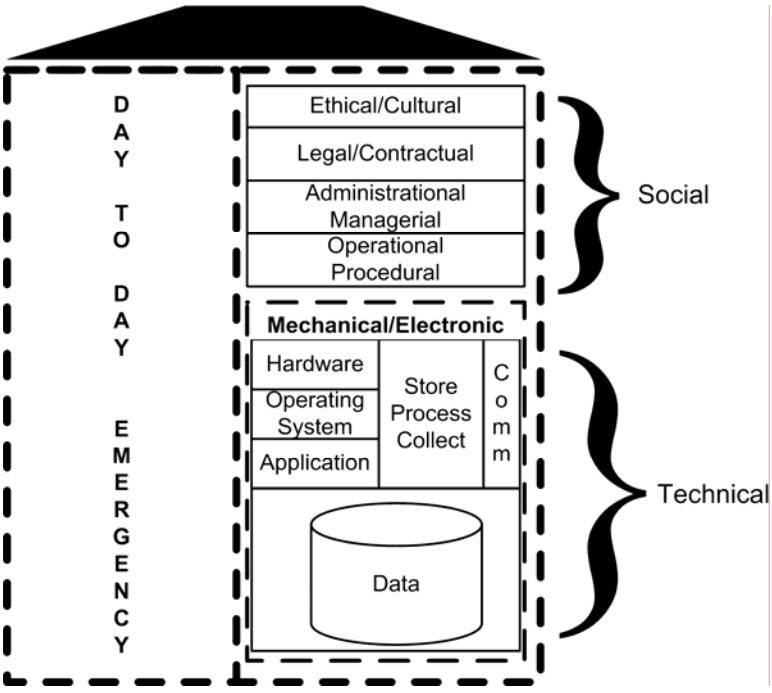


Figure 1. SBC Model, from Kowalski (1994, p. 19).

The SBC model can be used to analyze security at every level, from individual to national. This flexibility combined with the inclusion of the social elements meant that the SBC-model was the best one for this study.

Another advantage of the SBC model is that there is a well-defined SBC checklist that can be used to analyze a computer crime, or to report a crime, or just to understand an event better. This checklist was expanded on and improved by Tarimo (2006), who created a checklist for general information security work, with the SBC model as a basis. That checklist is, however, mostly intended for use by people actually working with the information security, and not a tool to test subjects such as in the context of this study.

Another problem is that most users do not understand how security works, and therefore construct their own, often incorrect, models (Adams & Sasse, 1999). The "old" way of managing information security has led to two specific problems (Adams & Sasse, 1999 p. 45):

> (a) users' lack of security awareness, and
>
> (b) security departments' lack of knowledge about users, producing security mechanisms and systems that are not usable. These two factors lower users' motivation to produce secure work practices. This in turn reinforces security departments' belief that users are "inherently insecure"

and leads to the introduction of stricter mechanisms, which require more effort from users.

The use of the SBC model to visualize the whole area of information security is to facilitate the creation of a mental model amongst the subjects, and indeed the readers of the results of the study. According to Sasse (1997), users tend to create mental models of the functions and behavior of the system that are relevant to the user. The SBC model uses common sense terminology and therefore makes it easy for the user to be directed in the creation of a mental model of security that is as intended by the researcher. With a correct mental image of security, awareness can be improved and less strict mechanisms might be needed.

The areas we focus on in this study are those in the social group (taken from Kowalski, 1994):

- Ethical/Cultural: Educational measures to deal with ethical or cultural problems.
- Legal/Contractual: Legal contingency structure, knowledge of contractual and legal requirements.
- Administrative/Managerial: Measures activities focused on control, the formulation of policy, and regulations. Risk and vulnerability analysis.
- Operational/Procedural: Concrete activities for security.

The study

In this study, our aim was to see if a soft approach to penetration testing could be used to complement, or even replace, traditional penetration testing when it comes to the human element of security. We wanted to test a pragmatic approach where we involve the users and get their thoughts and feedback on security related areas, in order to avoid the ethical problems of more aggressive penetration testing.

The research question was: Can interviews of employees on their awareness of the social areas of information security as proposed by the SBC model give useful information like traditional penetration tests do?

Method

In the beginning of this study we struggled with the choice of method. In a longer study, observation, a case study, or even action research could have been interesting, but we could only get a limited access to the nurses, making a qualitative method suitable. The beauty of the qualitative method is that you get an insight into people’s world and views. While a stricter approach, such as grounded theory, was tempting, we realized that this was a preliminary study, with the goal of developing a pragmatic set of questions and guidelines that could be used by professionals, and not just researchers.

As we wanted to get input from the users, while still allowing them to think freely, to some extent, we chose to use a semi-structured interview as described by May (2001). With a semi-structured interview, the questions are prepared in advance, but the researcher can ask complementary questions and have a dialogue with the subject. In order to facilitate elaboration, certain possible follow up questions were prepared beforehand. Since we suspected that the subjects would be unwilling to consider that they behave insecurely, we also asked about what their colleagues would do. This also had the benefit of covering more subjects.

The questions were developed with the SBC model as a foundation. A preliminary test was done on the interview questions, after which slight changes were made on the order in which the questions were asked and how the questions were formulated. By listening to the recording, the researcher became more aware about how the wording and how the formulation of the questions could influence the subject. While the process of developing questions using the SBC model as a background can be used for most studied organizations, the questions themselves should be adapted to the organization and the people studied. As the subjects studied mostly work with electronic journals, the questions were developed with that in mind. The goal was to ask them questions that they understood, about areas that were relevant to them. It was no use asking them complex questions that they could not answer, it would only lead to weakened rapport. The validity of the study is ensured through the description of the process, that several persons were interviewed, by the opportunity for the subjects to provide feedback from reading the transcriptions, as well as the fact that the results are quite reasonable.

The interviews

The interviews all started with an explanation of the study, ethical aspects, and so on. During the interviews, other questions than the pre-developed ones were asked, which was expected in advance. The researcher asking the questions is not only trained in security but also a trained nurse, which made the interviews easier. Each interview lasted between 35 and 55 minutes. They were taped and later transcribed and then sent to the subjects to ensure that there were no major misunderstandings or misquotes. The interviews were subsequently analyzed using qualitative methods. The transcriptions were sent using e-mail due to practical reasons, but it is notable that poor e-mail security might endanger the anonymity of the subjects. In the answers, different themes and categories were apparent, and, in some cases, the subjects replied in such a way that the answers could easily be compared; in those cases a comparative analysis was made.

Ethical aspects

This study was done in connection with a larger research project, Melior, where the aim is to study different aspects of electronic patient journals in Swedish healthcare. The involved organizations approved their employees' participation in this study. The department heads were contacted and given a description of the study, which they approved. They then facilitated contact with a suitable nurse who could participate in the study.

The nurses were contacted, given a description of the project and what the interview and method would consist of, and asked if they would like to participate.

At the start of the interview, the subjects were again informed about the aim and method of the study, both orally and in a written document. They were also informed that the interviews would be taped and the tapes stored, but that they would remain anonymous both in the study and on the tapes. The subjects were also informed about how the material would be published and also that they could discontinue the study at any time, without needing to give a reason. Both the subject and the researcher then signed the document. The subjects were also offered a chance to see the transcriptions from their own interviews to ensure that there were no misunderstandings or misquotes.

Results

This section contains three sets of results. The first part is the procedure, the method, for doing this kind of penetration test. The second is the questions we developed, and the third set of results is those we got from using the questions in the health care organization, which are included as an example of the kind of real world data that can be gathered from using this approach.

The procedure for the penetration testing

The procedure is described as a numbered list, as there is some importance in doing the steps in the right order. The steps are taken from our own process, but improved based on our experience from this study.

1. Get the permission you need from those involved.
2. Randomly select subjects if possible, or find out who you can use from the organization.
3. Study the organization's domain; learn about typical flaws, if any.
4. Using the SBC model as a basis, develop questions and follow-up questions covering the relevant areas.

5. Interview the subjects using semi-structured interviews and remember ethical demands.
6. Analyze the interviews based on answer frequency, content, and context.
7. Using the SBC model, try to visualize strong and weak areas of the information security based on this test. Be sure to include relevant thoughts and fears from the interviews in written materials too.

The process is deceivingly simple, but it worked well for us.

The questions

Ten questions were asked, with between two and ten prepared follow-up questions to each. The questions were asked in Swedish, but have been translated into English for this paper. They were not always asked exactly as phrased below, and there might be slight differences in meaning between the languages which were lost in translation.

While our aim was to study the Social section, it was apparent that some areas in the Technical section were also of interest. In 0 the areas we focused on are in white, while the other areas are shaded. In the questions below, the intended focus areas from SBC are in brackets.

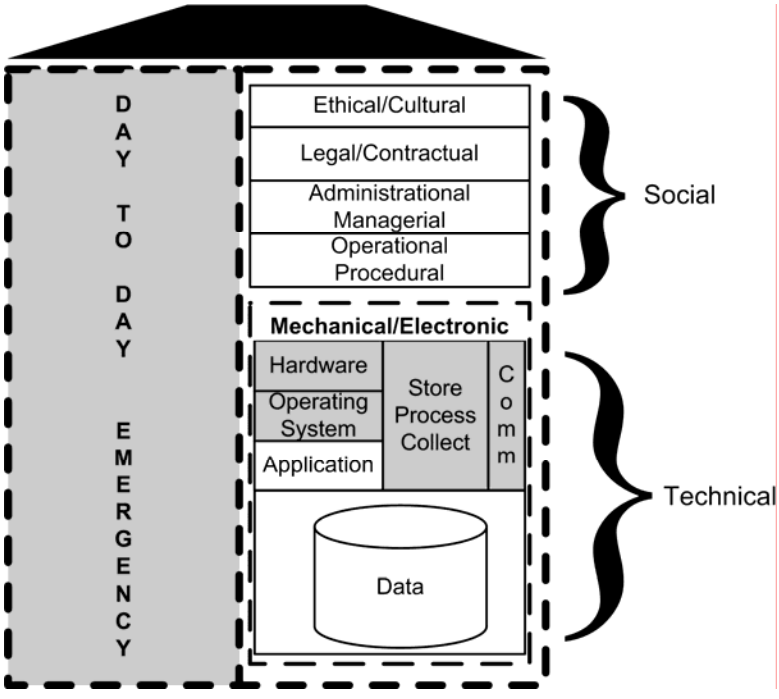


Figure 2. The areas of the SBC model that were studied.

1. Introduction

First I would like you to tell a bit about yourself, your professional background, and your computer experience.

- Do you have any other education?
- How long have you been working with Melior [The electronic journal system]?

2. Electronic journals and availability (Technical/Cultural/Application).

Could you tell a bit about how you feel about using electronic journals?
Follow-up questions:

- How is the access to data when using the journal?
- Are the systems ever down?
- Do you find log-ins time consuming?
- From which mode do you log in?

3. Passwords (Cultural, Operational).

Could you tell a bit about what is important to you when choosing a password.
Follow-up questions:

- What do you do if you forget passwords?
- If you write down passwords, where do you store them?
- Do you reuse the same passwords inside or outside the hospital?

4. Log-ins, log-outs and log-files (Cultural/Legal/Contractual).

Could you tell a bit more about log-ins and especially log-outs?
Follow-up questions:

- Do people always log-out? If not, how common is it?
- Why do you believe that someone does not log out?
- What kind of consequences do you think not logging out might have?
- Do colleagues know or use each other's logins in the different systems? What do you feel about that?
- Do you use shared log-ins anywhere?
- If so, where are the passwords stored?
- How do you feel about everything you do being stored in log-files?
- When using Internet at the hospital, do you consider that those pages you visit are also logged, and how do you feel about it?
- Is there anything you would like to add about passwords and log-ins?

5. Short scenarios (Operational, Cultural).

You will now be given a couple of short scenarios to comment about, and you are welcome to think out loud about them, and ask if there is anything more you would like to know.

Follow-up questions:

- A systems administrator phones you, asking for your user-name and password. How would your colleagues react, and how would you react?
- A guy from the IT department comes to fix some error on the computers. He is wearing a sweater with the logo of the hospital and the IT department. How do you think that you and your co-workers would react, if he sat down by a computer connected to the network and started to work?
- (Today there are a lot of web-sites for special interests, where many people have a lot of friends that they can communicate with. If they access such a site when they are at the hospital, and a friend sends a link to something like a cool motorbike or cute dog using that site, would people be less prone to clicking that link if they were at work than they would be at home? How would your colleagues think, and how would you think?

6. What do you believe is the probability of something like that happening in your workplace? (Operational, Cultural)

Follow-up questions:

- Have you ever heard about, or experienced a similar incident?
- Strange phone calls?
- Someone coming to your work, asking "the wrong questions"?
- Strange e-mails?

7. Tell me about the threats that you believe exist to the patient's data, and the computer system as a whole, now that more and more information is available from the same source. (Ethics, Operational, Legal)

Follow-up questions:

- Why do we protect patient data, and what would happen if we did not?
- What would happen if access to Melior and the patient data would get into the wrong hands?
- Who would be interested?
- What would the information be used for?
- How do you think someone would go about getting access to the information systems, that is, rights to change the data?
- Who would want to do that?
- What would it be used for?
- What could happen?

8. Tell me more about how, where, and by whom you have been trained, or educated, about information security? (Legal, Contractual)

- The security policy?
- Can you read about this somewhere?
- Is it a continuous education?

9. Is there anything you would like to add on this subject, perhaps something has come to your mind during the interview?

10. Last question: Could you give an estimate on how many times a day you write your password? (Operational)

Results from using the method

Seven interviews were conducted with nurses that came from six different departments. Six of the nurses were females, and one was male. Their ages were between 25 and 50 years, and they had been nurses from 4 to 26 years. The distribution of the subjects seems to match the age/gender distribution in the nursing community. Four nurses worked with computer related tasks in their departments, but none had any specific computer education. They were self-taught.

Interviewing the subjects gave us ample examples of potential security breaches. A brief summary of the kind of data we received is described below, in order to illustrate some of the flaws found by this study. They are presented below in a manner modeled after the SBC model.

Ethic/Cultural: The awareness of ethical demands was high, and the subjects expressed a serious concern for secrecy concerning patient data. There is a tradition of keeping patient data safe. There is also, however, a tradition of unsafe behavior, for instance, poor password management, an inability to lock unsupervised computers, and so on. The stated awareness is good, but the actions are not. The focus on functionality meant that passwords were selected so that they were easy to remember and quick to type.

Legal/Contractual: The law clearly states it is illegal to study patient information that one is not entitled to read. This made the nurses aware of the importance of keeping passwords safe, and they had an intention of locking computers in order to not be legally responsible for the actions of others using their account. While there was information on security, the documents were not known by most of the subjects, and their actual influence on the day-to-day security work is probably very slight.

Administrative/Managerial: The subjects rarely acted with the written policies as support; they only had a vague idea of what the policy consisted of and where it could be found. According to the subjects' understanding, the policies are more concerned with regulating the relationship between employer and employee when it comes to not conducting personal business during working hours. As personal use of the Internet is forbidden, it is also not regulated, for example, there are no rules regulating receiving files from friends. The users did not know of any regulations about how often to change passwords, how to create a good password, and so on.

Operational/Procedural: The subjects had a specific vision of the threats to their information. The threats were from staff, patients, and relatives. Other attackers seemed highly unlikely. The consequences of an attack were primarily loss of personal information, but the concern for the integrity of data, as well as the availability of data was quite low. While there was an appreciation of the influences one's own actions might have on the security, the awareness of what one could actually do to improve security was very low. Technical attacks were generally not understood at all, and there was little awareness of attacks aimed against the social aspect.

Technical: The login procedure to the network was slow and tedious, and the systems lacked any automatic screen locks, as discussed above.

General Trends: The awareness of the importance of password quality is apparent. The actual quality of passwords, however, is not good. The argument among the subjects is that this is because of complicated routines. The knowledge of legal and ethical demands does seem to at least improve awareness of password importance, but not enough to change behavior.

It is notable that the users are an apparent risk. They do not see any links between their own behavior and potential external attacks. They know little about the security documents that are available, and the documents they have read give them only limited information. This is because the documents have not been created to give guidelines for the usage of the information systems. They tend to focus on protecting the patient journals rather than protecting the systems themselves.

This study shows that two suggested social engineering attacks would be quite successful: Asking for login information over the phone and pretending to be an IT technician to gain access. The relative inexperience of the subjects on using the Internet meant that there were no conclusive answers on whether or not they would click on links. The major reason for their stiff resistance to this was probably that the scenario was too far away from their situation for them to grasp it.

Using the data acquired in the interviews, we made a judgment on the readiness of the areas surveyed based on frequency of answers, and the impact of the vulnerabilities. It can be seen in Figure 3 below.

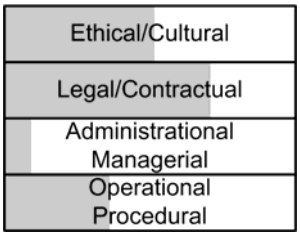


Figure 3. State of the social area.

The dark areas in Figure 3 represent coverage, i.e. the darker an area, the more protected it is.

Conclusion

Our aim in this study was to see if it was possible to get useful and relevant information by asking the users, instead of performing a penetration test on them. We argue that our result indicates that there is a surprising amount of information available from simply asking the users, instead of trying to trick them. The major advantages of this approach is that there are no major ethical concerns and that the users are asked to reason and think about security, hopefully creating a learning process simply by involving them in a non-judgmental manner. The very audit gives at least a selection of users increased awareness of possible attacks and the status of the information security situation. Hopefully the users will be asked by their colleagues or they will ask them about security after the audit. It also manages to capture certain risks that are quite hard to ask about in a more strictly formalized environment. The users were not afraid to report believing that both they themselves and their colleagues could potentially fall for certain attacks. It is interesting to contrast this to some of the arguments put forth by Nohlberg (2005) about the need to deceive subjects in order to get an accurate view of security. This might depend on the subjects. Subjects who claim to have a greater understanding of technology and security might have more "pride", while people who do not have the same “pride” might be more honest, or even too honest. Their lack of knowledge might make them believe that they and their coworkers are far more gullible than they actually are, and that the attacks are super-efficient. This is why the auditor needs to have a grasp of information security, to be able to gently and without bias bring potentially too far-fetched and biased ideas and thoughts in line with reality. We did not experience any problems with this.

The weaknesses of the approach is that it is quite resource intensive. With slightly more formalized interview guidelines and simplified reporting, a lot of time could be saved. Another problem is that in most cases only a small selection of employees can be surveyed, but that is a problem shared with all but the automated penetration tests.

One of the future improvements is to develop a metric on how to judge the vulnerabilities in every category. It would be quite interesting to compare the results from this very “soft” approach to penetration testing to those from other methods. By doing this it would be possible to judge more fairly the merits of each method, but in such a study it is important to assess more than the results. Increased awareness among the subjects and their colleagues, attitude change, ethical concerns, and time spent are also important factors that could be examined in a future study. We believe there is a problem in

using just one method, because all methods probably have their own merits and specific contexts in which they are useful.

In general, this approach could be useful for most organizations. The SBC model allows the results to be presented in such a way that they are easy to understand by non-specialists, and the survey of the organization can also be complemented by technical assessment to give a complete overview of the status. The approach suggested in this paper can either be used as a pre-study before a traditional penetration test if the organization sees the need for the traditional approach, or as a stand alone tool to gain an understanding of the status of the social section of information security.

The scary fact is that a skilled social engineer is successful in far too many cases today. Our primary task now as security specialists is not simply to see that there are vulnerabilities, but instead to understand more about the situation of our users, and improve that. Only by understanding their situation can we limit the success of social engineering in the future. We will not build better security by pointing fingers at our employees, but by holding hands with them.

References

Adams A. & Sasse M. A. (1999) Users are not the Enemy: Why users compromise computer security mechanisms and how to take remedial measures, Communications of the ACM, 42, 12, 40 – 46.

Barret, N. (2003) Penetration testing and social engineering: hacking the weakest link. Information Security Technical Report. 8, 4, 56 – 64.

Dalrymple, M. (2005) Auditors Find IRS Workers Prone to Hackers. [Online]. AP. Available from: http://www.infosecnews.org/hypermail/0503/9684.html [Accessed 30 Mar 2008].

Hasle, H., Kristiansen, Y., Kintel, K. & Snekkenes, E. (2005) Measuring Resistance to Social Engineering. In Proceedings of the First International Conference on Information Security Practice and Experience, 2005, Singapore, April 11-14, volume 3439 of Lecture Notes in Computer Science, pp. 132-143. Springer, 2005.

ISACA (2004) IS Auditing Procedure Security Assessment-Penetration Testing and Vulnerability Analysis. [Online]. ISACA. Available from: http://www.isaca.org/ContentManagement/ContentDisplay.cfm?ContentID=18750 [Accessed May 1 2008].

Jakobsson, M. (2005) Modeling and Preventing Phishing Attacks. Phishing Panel of Financial Cryptography 2005.

Jones, C. (2003) The Social Engineering: Understanding and Auditing [Online]. SANS Institute. Available from: http://www.sans.org/rr/whitepapers/engineering/1332.php [Accessed May 1 2008].

Kowalski, S. (1994) IT Insecurity: A Multi-disciplinary Inquiry. Diss. University of Stockholm. Report series No. 94-040, Stockholm, Sweden.

Marett, K., Biros, D. & Knode, M. (2004) A Longitudinal Study of Lie Detection Training, Lecture Notes in Computer Science, Volume 3073, pp. 187–200. Springer, 2004.

May, T. (2001) Social Research: Issues, Methods and Process. Buckingham: Open University Press.

Mitnick, K. & Simon, W. (2002) The Art of deception: Controlling the Human Element of Security. Indianapolis: Wiley Publishing, Inc.

Nohlberg, M. (2005) Social Engineering Audits Using Anonymous Surveys – Conning the Users in Order to Know if They Can Be Conned. In *Proceedings of the 4th Security Conference*, Las Vegas, USA, Mar 30 – 31 2005.

Orgill, G., Romney, G., Bailey, M. & Orgill, P. (2004) The Urgency for Effective User Privacy-education to Counter Social Engineering Attacks on Secure Computer Systems, In *Proceedings of SIGITE Conference'2004*. pp.177-181.

Sasse, A. (1997) Eliciting and describing users' models of computer systems. PhD Thesis. University of Birmingham, Faculty of Science. Birmingham, UK:

SIS 2003. SIS Handbok 550. Terminologi för informationssäkerhet. SIS Förlag AB. Stockholm (in Swedish).

Stanton, J., Yamodo-Fagnot, I., Stam, K. (2005) The Madness of Crowds: Employees Beliefs about Information Security in Relation to Security Outcomes. In Proceedings of the 4th Security Conference, Las Vegas, USA, 30 – 31 March 2005.

Tarimo, C. (2006) ICT Security Readiness Checklist for Developing Countries : A Social-Technical Approach. Diss. University of Stockholm. Report series No.06-017, Stockholm.

Measuring Readiness against Automated Social Engineering[1]

Marcus Nohlberg, Stewart Kowalski and Markus Huber

Abstract

This paper presents the result of a case study of the readiness of four large Swedish multinational corporations to deal with automated social engineering attacks. A preliminary study to review how the security policy of a large corporation deals with social engineering attacks was performed. The results from this study were combined with a conceptual model of social engineering when constructing a new interview protocol and a grading scale. This interview protocol was designed to measure the readiness of an organization to deal with social engineering attacks in general, and in this case with automated social engineering in particular. Four interviews were conducted with senior security managers and senior employees. Results indicate that no organization was over 60% on the readiness scale and thus all are considered at risk of attack.

Key words: Automated social engineering, social engineering, readiness, security readiness measurements, web 2.0 security, cycle of deception, online social networks.

---

[1] A version of this paper was published in *Proceedings of the 7th Security Conference*, Las Vegas, USA, June 2008. Ed. G. Dhillon. Washington DC: Information Institute Publishing, USA, ISBN: 978-1-935160-01-4.

Introduction

Over the past years we have shifted more and more responsibility for information security decisions to end-users. Many of these decisions, at least to the information security professional, seem trivial, such as whether you should trust the old Nigerian gentleman who wants you to help him move funds from his native country, for a large fee? Did you really win a million dollars in a lottery even though you did not buy a ticket? Should you click on the attached file in order to see the nude star? Yet, we spend a great deal of time, and resources, educating our users to ignore these trivial attack forms on the Internet. Most users who care about security can learn to recognize these attacks and to become rather good at spotting them. But, the world is changing. One of the more difficult attacks to protect end-users from and educate them about is the social engineering attack. The success rate of a social engineering attack is hard to measure, but some research indicates its success rate is high (Nohlberg, 2005). The reason social engineering is not used more often is that it is expensive, due to the fact that it is labor intensive and can take a great deal of time. It takes time to gather information, to build a relationship, to exploit it, and to maintain it. From the attackers' economical standpoint, it is often more economical to send out 10,000 e-mails and hope to catch one unsuspecting end user than to spend time developing a good relationship with their next victim. But what happens when the attackers can combine the economical benefits of automated attacks with the high success rate of social engineering? We get a gap between what we can control and what the attackers can do. The introduction of automated social engineering might change the face of the relationship based web, web 2.0, as we know it; it might indeed even change the use of the Internet. How do you protect yourself, not from strangers, but from your friends? How can you tell if someone is an AI chat-bot or an old friend if they know the same things?

In this case study, we attempted to see what the readiness of organizations is to automated social engineering attacks. In order to measure the readiness of organizations to deal with automated social engineering attacks, we designed a conceptual model of a social engineering attack and tested the model by doing in-depth interviews with the CSOs and high ranking security specialists in four Swedish multinational companies.

The paper is divided into four sections. Following the introduction, a review of the background of social engineering, including a presentation of the conceptual model of social engineering is provided in section two. While section three presents the case study, including methods and results, section four lists some conclusions and suggestions for further study.

Background

In this section we present background information on the subjects covered in this paper.

Social Engineering

Social engineering is a term used for techniques to con, or trick, users into revealing data, or login information, to individuals who are not authorized to receive it. Since many users do not believe that anyone would ever trick or con them, because they are not "rich and famous", and that hackers cannot do much damage anyway (Brostoff, et al, 2002), these attack techniques are often highly successful. The success rate of social engineering is also due to the fact that most users do not understand how security works, and therefore construct their own, often incorrect, models (Adams & Sasse, 1999). In a study performed by Treasury Department inspectors, one third of the Internal Revenue Service (IRS) employees submitted their login and password to auditors who called pretending to be computer technicians (Dalrymple, 2005). There have been other studies on the "gullibility" of users, and to what extent they submit information when being attacked by perpetrators using Phishing and social engineering attacks. The results have generally indicated that users are quite susceptible to these kinds of attacks.

Social engineering as a term has been used for quite a while in the security sector. We have found references dating back to 1995 (Winkler & Dealt, 1995) – and there was, of course, the boom of notoriety of social engineering in connection to the case of Kevin Mitnick (Mitnick & Simon, 2002). His imprisonments lead to a great deal of publicity for the technique, and his subsequent books have also served to give further attention to the technique.

Web 2.0

Finding a widely accepted definition of Web 2.0 is difficult since the term is meant to refer to different aspects of successful online services in the post dot-com era. One aspect among others is the new and rediscovered technologies that are utilized in web services today: Ajax, SOAP, WSDL, CSS, RSS, and XML. These technologies are important and an extensive topic on their own. Furthermore, there is a great deal of literature in the area already, including publications that deal with the security aspects of Web 2.0 in particular (Lawton, 2007). Web 2.0 is in many cases discussed as a business concept, which is not surprising since the term itself was created at a business conference (O'Reilly, 2005). Another view on Web 2.0 is the discussion on user-generated data and information, which is seen as the major success factor of Web 2.0 applications. Since the term itself was created during a business conference (O'Reilly, 2005) and because Web 2.0 applications

have become the driver for the online industry, Web 2.0 is in many cases discussed as a business concept.

Online Social Networks or Social Networking Sites (SNSs) are a fast growing subgroup of Web 2.0 applications. Users of SNSs can maintain an online profile, foster social relationships, and exchange media files with other users. The number of users and signups are critical for the commercial success of SSNs services. Service providers therefore design their services in order to increase the number of new signups and target broader user scopes.

Automated Social Engineering

When looking at the latest trends in online use, we see that social networking sites, for instance Facebook, are being used by ever increasing numbers of people (ComScore, 2007). We see the same with other online services, such as blogs, which by an evolved usability in comparison to manually updated web pages attract a much wider selection of users.

More and more companies are trying to embrace these social networks, and are in fact encouraging their employees to use them. As the usage has increased, so has the awareness of security vulnerabilities, but the scientific focus lies mainly at privacy concerns. Rosenblum (2007) gives an introduction to the various risks related to the disclosure of personal information in Online Social Networks. Gross, et al. (2005) analyzed the online behavior of 4,000 Carnegie Mellon University students and concludes that the students have not been aware of the ways their personal information could be exploited. These two papers discuss the security issues based on privacy risks; a broader approach is the ENISA position paper (Hogben, 2007) that focuses on more threat categories.

What seems to be missing in the current research is the big picture. Traditional social engineering is more efficient if the attacker has more background information. The traditional ways of getting background information have been to use online searches, or to do dumpster diving or other similar techniques. With the prevalent use of SNS there is, however, ample information, more or less freely, and conveniently, available on the network sites. This makes the data gathering part of social engineering far easier. So easy, in fact, that it is quite obvious that the attacks can and will be automated, which we have seen signs of when studying online sexual predators (Nohlberg & Kowalski, 2008). Special software, so called Internet bots, is already being used for commercial purposes. These bots are currently used to spread spam messages within Online Social Networks. A recent study based on anonymized headers of 362 million messages exchanged by 4.2 million users of Facebook, stated that 43% of all messages analyzed were Spam (Golder,

et al. 2007). Existing SSN bots can be easily adapted to perform automated social engineering attacks, ASE.

When we face the threat of automated social engineering, a number of interesting things happen. One of the major constraints on using social engineering has always been that it is costly, as it requires a lot of human interaction, while other attack vectors, such as Phishing, require very little human time per victim. From a purely economical perspective, this made attackers decide not to use social engineering and instead rely on cheaper attacks. When social engineering is automated, however, we reach the point where attacks that are very successful (Nohlberg, 2005) are also cheap. Imagine not being able to know if the person you have a relationship with online, who knows all about your background, and with whom you share friends, interests, and hobbies – perhaps you even went to the same school – is a malicious automated social engineering bot or a real, human friend. And an attacker can, in fact, have an almost unlimited number of bots, learning the social networks, and building groups of "friends" online. Much like the current botnets, which are threatening due to their size and network communicating possibilities, the automated social engineering bots will be threatening due to the fact that they will know so many people and so much about them. Their threat will lie in abusing the social networks, and not the Internet traffic networks.

The Cycle of Deception

In order to describe the complexity of a social engineering attack, we use the cycle of deception (Nohlberg & Kowalski, 2008). This model was developed to present different perspectives on what goes on in each part of an attack, from the perspective of the involved parts, which are the attacker, the victim, and the protector. The model is presented in Figure 1.

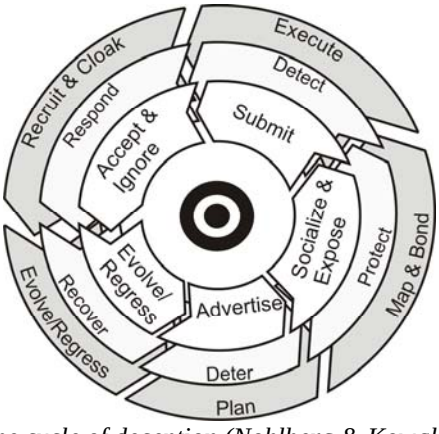


Figure 1. The cycle of deception (Nohlberg & Kowalski, 2008).

In each of the steps in the model, there are a set of requisites that must be met if the attack is to be successful, or if the attacker will be able to do it again. If any one of these is not met, the attack will fail. Therefore, one can use this model, for instance, to build defenses or to map and describe an attack.

The most central part of the model represents the victim. By having something of value and making it known, either knowingly or unknowingly, the victim *advertises* its suitability as a target. By *socializing* with the criminal, the victim sets itself up for deception, and by *exposing* valuables they make them accessible to the attacker. When the actual crime is executed, the victim *submits* to the crime, for instance, by giving out the secret information. After the crime has been executed, the victim can chose to *accept* the crime, for instance, by believing that it was not so "serious", or by simply *ignoring* it, either knowingly or by actually not knowing about the crime. By learning from the crime, the victim can *evolve* into someone who is harder to victimize in the future, but it is also possible that the victim can *regress*, turning into someone who accepts the role of victim and is an easier prey in the future.

The second part of the model represents the defenses between the victim and the attacker. By having a good, public policy, or a reputation of reporting incidents to the police, you, the defender, can *deter* an attacker. In addition, by making little sensitive data available, and educating the employees about the risks and methods of attackers who bond with them, as well as providing a strong policy on how to act, you *protect* the organization. Furthermore, running a surveillance of the network communication, can reveal when sensitive data are sent, or when sensitive data are accessed, and by having well educated employees that know when they are asked illicit questions, you *detect* an attack. By making it easy and attaching no social or professional stigma to reporting social engineering incidents, and by making the employees aware of how they can be manipulated into acting on behalf of attackers, you are able to *respond* to an ongoing attack. By knowing the value of your data, having attacks reported, and having a well-designed policy, you can *recover* from the attack and learn from it. Hopefully you are able to find the attacker to prevent him from evolving and attacking you, or others, in the future.

The third part of the model is the attacker. An attacker must have a *goal* for an attack, and a *plan* how to reach the goal. In order to get background information, the attacker *maps* the organization and *bonds* with potential victims. When the attacker asks for secret information or gets the victims to do something that they should not do, it is an *execution* of the attack. After the execution, the attacker tries to *cloak* the attack so that it cannot be easily discovered, and perhaps to *recruit* the victim to help in future attacks. The

last step is for the attacker to *evolve*, if the attack has been successful and an internal justification can be created. If not, the attacker can *regress* into either performing a more basic attack or to stop the attack.

Case Study

Method

A literature study was conducted at the beginning of the research to obtain an overview of existing research on social engineering. Attention was particularly given to formalized models of social engineering, which might serve as a base for automated social engineering. While social engineering was found to remain rather unexplored in the field of information security, other areas of science ranging from social science to economics provided ample material.

Preliminary case study

In order to get background information about the problem area of ASE, a pre-case study was carried out at an international high tech company by analyzing their security policies and interviewing experts. The respondents were carefully selected to ensure that views and opinions diversify: a technologically-open-minded security generalist on the one hand and an experienced senior security manager, on the other. The questionnaire consisted of ten, both open and closed, questions. The interviews each had a timeframe of 30 minutes and were held in private, in the company's conference rooms. To ensure that the interviewees could speak freely and without constraints, the ethics involved in the interviews was explained to them beforehand.

The case study showed that the evaluated company did not have any policies or directives that address the current threat of automated social engineering intuitively. The interviews of the experts led to the conclusion that the company is currently unaware of the Web 2.0 phenomenon, or how Web 2.0 affects Information Technology as a whole. Although one interviewee stated that the usage of SSNs is noticeably increasing among employees, the security expert had little knowledge about possible social engineering attacks. We came to the conclusion that in order to obtain a good understanding, new and more concrete questions were needed. Moreover, the lack of knowledge in the area of Web 2.0 made it clear that new questions have to be based on well-established and more fundamental principles, which are approachable by all professionals in the field of information security.

The Interviews

The construction of the questionnaire was based on the assumption that automated social engineering will follow the flow of the Cycle of deception either by design or necessity. Hence the questions were constructed to cover all 15 areas of the cycle of deception. Due to time constraints in the interviews only one question was asked in each case. The intention of the interviews was to get a deeper understanding of the area based on the tailored subset of possible questions. The questionnaire was written in Swedish, as the respondents were native speakers. Personal contacts made it possible to choose three senior security managers and one security management consultant out of a pool of security professionals. Furthermore, the organizations were deliberately selected to cover various industry sectors:

Organization 1: International high-tech telecom company

Organization 2: International energy provider

Organization 3: International IT company with a focus on web 2.0

Organization 4: International health care company

Each interview had a time frame of 30 - 45 minutes and started with a briefing on the ethics involved: anonymity, the possibility to discontinue the interview without giving reasons, and that the data gathered would not be misused in any way. Two interviews were performed face to face while the other two were conducted via the telephone. In addition to a short briefing on the interview's ethics, the respondents were asked if they agreed to the interviews being recorded with a voice recorder. Following each interview the answers were transcribed and analyzed. The assessment of responses was based on the well-established Behavior/Attitude/Knowledge triad and inspired, amongst others, by the security awareness testing done by H.A. Kruger, et al. (2006). In order to obtain comparable results, each answer was then rated with the matrix in table 1 below.

Table 1. The Ranking Matrix.

	Behavior	Attitude	Knowledge
None	0	0	0
Informal	+1	+1	+1
Formal	+1	+1	+1

Hence, the best possible score for an organization is six points in each 15 areas of the cycle of deception, which is equivalent to 90 points in total. We hypothesized that this arrangement enables an overall assessment of the current readiness of each organization as well as a granular assessment of specific weaknesses.

After the interviews were transcribed and analyzed, the data were used to create polar graphs, where each category from each organization was drawn. This was then analyzed, together with calculation of averages, and so on.

Results

This study has yielded two different sets of results. The first is the results of the question regarding the readiness of the studied organizations to defend themselves against automated social engineering attacks, and the second set of results is from the method used to study the readiness.

Results on readiness

After rating the organizations in accordance with the scoring matrix, there were some emerging patterns. It seems that the organizations, in general, have a higher level of readiness in the traditional defenses than on protecting and covering the victims, and the least on preventing attackers. This is not a surprising observation, but it is notable.

Looking at the attacker cycle, we see that three of the organizations do not have an updated image of who might attack them, nor do they try to prevent mapping from attackers. The exception is one of the organizations. Instead of using rules and regulations, this company makes it harder for an outsider to map the organization by providing all possible web 2.0 services to the employees within the company, thus eliminating the need for them to use external services. The organizations do have training, but the social engineering training is almost always only theoretical and without practical exercises. Social engineering is also mostly a small part of the security training, and no one has any training on automated social engineering, as could be expected. When it comes to the risks associated with using web 2.0 sites, the organizations generally assume that the old rules “should” cover it. However, as they have a basic understanding of the risks associated with it, they instigate improvements and say they are, in several cases, just a short time away from establishing new policies and rules about this. However, as they do not have any policy today, it is an unregulated area in practice. The organizations understand the use of real examples from their companies when teaching social engineering, but these are rarely used for more than anecdotal stories. They agree that using actual examples would make it easier for their employees to understand the actual risk or impact of social engineering attacks.

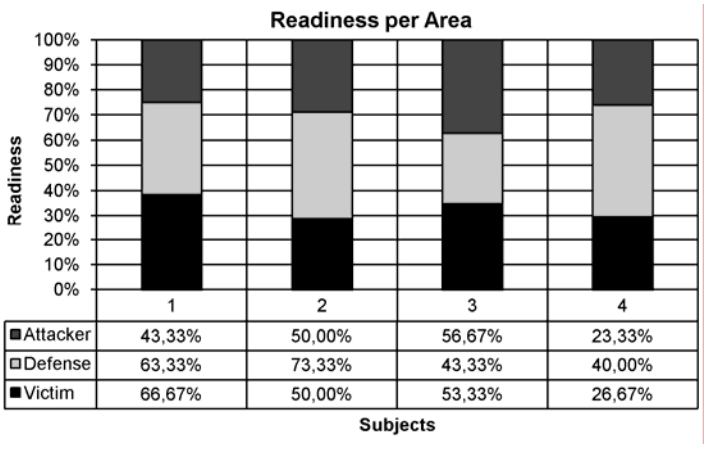
The defense cycle is one with more well-rounded defenses, but three of the organizations do not have education based on the different types of jobs their employees might have, indicating that the education might be too broad to be applicable for their users. After all, there are vast differences in the risks and

potential attacks facing different job positions. Most organizations have an incident reporting system, but it is not possible to be anonymous in all of them. The organizations agreed on the importance of anonymous reporting as traditional reporting might make a potential victim who feels embarrassed reluctant to report an incident. One measure that might make certain attacks considerably harder to carry out is to use the separation of duties method, something that the organizations feel is important, but which has only been used in some limited cases. Even though the organizations think the method is good, they do not know specifically how often it is used. However, it seems that it is mostly used in financial systems, rather than in security ones. The organizations all classify information and they all share the same problem with it: their employees are not using the classification correctly, but far too indiscriminately. They tend to class almost all documents as sensitive or secret, rendering the information classification somewhat useless. If everything is ranked as secret, no one will treat the really secret documents differently than the regular ones. The organizations are all aware of this and are at least preparing to educate and train users in this matter, but seem less than hopeful that it will be successful.

When looking at the victim cycle, we see that, generally speaking, there are no good policies on what, where, and how work related activities and work positions can be discussed online. While some work is being done to regulate this, the most prepared organization, number three, has solved it by having massive internal systems for discussion and participation. They still, however, lack any regulation of the use of *external* systems, rendering them also quite vulnerable. The organizations all consider that their employees mostly adhere to rules and regulations on security, but there does not appear to be any formal verification of this; it is mostly a "feeling" the organizations have. There is quite a difference when it comes to how frequently the employees are educated and trained. The most common solution seems to be to have at least one training session a year, or at least once for every employee. However, only two organizations have any form of continuous education and constant security updates for their employees. The other two organizations are highly vulnerable with the possibility of changes in policy, and rules, and the risks of threats taking more than a year for the employees to be informed about. They are thus slow to react to new and rapidly evolving security threats. It is good to see that the organizations all consider themselves to be well prepared to handle their staff after a possible attack, making it harder for employees to be recruited for long term complicity.

When comparing the cycles, as seen in Figure 2, we see that their averages are similar, but that there are more activities on defenses (55 %) than on victims (49 %), and least of all on the attackers (43 %). This is to be expected, as the activities on defenses are similar to many common security approaches today. Still, it is notable that both the attacker and the victim side

can be addressed with many inexpensive activities, and can therefore be improved using small resources.



	1	2	3	4
Attacker	43,33%	50,00%	56,67%	23,33%
Defense	63,33%	73,33%	43,33%	40,00%
Victim	66,67%	50,00%	53,33%	26,67%

Figure 2. Readiness per Area.

Results on the method used

The questions were selected to be representative of the parts of the deception cycle. They were also targeted against what could be considered reasonable steps in preparing against automated social engineering attacks. As there are no established standards on how to protect against this, we formulated a selection of questions based on traditional social engineering and what we know about automated social engineering. In some cases we asked questions that seemed similar, in order to triangulate some areas. Due to time-constraints, we could only ask one question in each of the categories. This could be considered insufficient if more specific conclusions are to be drawn, but sufficient enough for an early, and general, study such as this one.

The questions asked and their categories (translated from Swedish).

Here we present the questions we asked.

Attacker

Goal and Plan: Do you have an understanding of who might want to attack your organization, and why?

Map & Bond: Do you use any techniques/routines to make it harder for an attacker to map your organization and your employees?

Execution: Do you have any training, both theoretical and practical, to teach your employees to handle social engineering and deceptive techniques?

Recruit & Cloak: Do you have rules forbidding employees to accept invitations to and to use social network, and communication sites being run outside of your organization?

Evolve/Regress: If you have any education on social engineering, do you use real examples, both successful and unsuccessful, from your own organization?

Defense

Deter: Do you have a legal text on the confidential material that your users are not allowed to read informing them that possessing this confidential material is illegal?

Protect: Do you have separation of duties, meaning that it requires more than one person for sensitive tasks in the business system and for security work?

Detect: Is your social engineering education adapted to the students' individual needs, for instance, with examples that identify the specific risks of their specific positions?

Respond: Do you have routines to handle suspicions and the reporting of incidents from the employees where they can be anonymous if they choose to be?

Recover: Do you know which of your information is sensitive and which is not? Do you use information classification?

Victim

Advertise: Do you have rules regulating what your employees can expose on the Internet regarding their work, assignments, responsibilities, and so on?

Socializing & expose: Do you have rules and routines for what, with whom, how, and when your employees can discuss work online?

Submit: Is your business culture such that the employees comply with security regulations and routines in practice?

Accept & Ignore: Do you have ongoing education and awareness training of social engineering, where the employees are frequently and regularly reminded about new and old threats and risks?

Evolve: Do you have routines that offer support and specific training for someone who has been a victim of an attack?

After analyzing the answers to the questions, graphs were drawn, as can be seen in Figure 3.

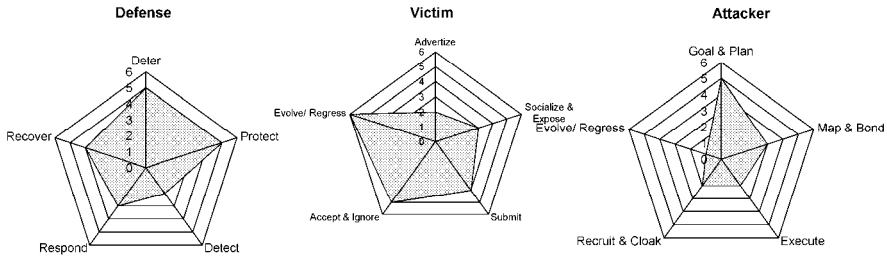


Figure 3. Sample Polar Graphs (Organization 1).

From an overall perspective, we see that this method of asking questions and the way of ranking provide usable data. The organizations that have a self-proclaimed high level of security consciousness also score higher with this ranking method; while the more "relaxed" organizations score lower, as illustrated in Figure 4.

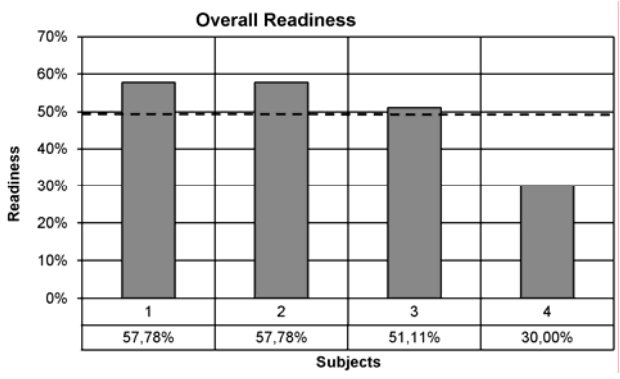


Figure 4. Overall Readiness.

All in all, the organizations scored mostly above 50 %, but nonetheless below 60 %. This indicates that even the most security conscious organizations still have some way to go before they have a sufficient level of readiness, and the organization which is least ready has quite a long way to go. The underlying model does, however, indicate that providing a higher level of protection, for instance, by spending resources on the victim side, can be quite inexpensive.

Conclusions

When we conducted the preliminary study, we interviewed a diverse group of security specialists from a major international telecom company, and ana-

lyzed the organization's security policy. We did this with the goal of finding out whether they were prepared for automated social engineering attacks in the future. What we learned was that they really knew little about this threat, and the answers we received were of little use in trying to judge their state of readiness. So we tried another approach. We would ask them questions about their routines and policies, things they knew, that we judged were important for readiness. The security specialists do not need to know about the threat, if they already have reasonable active measures of protection. What we learned was that while there is a certain readiness, it varies greatly, and most of the studied organizations have some weaknesses that might make them especially vulnerable to attacks. However, no single organization scored high enough to be considered without risk.

We found that there was a surprisingly high balance between the three cycles, even though the victim perspective tended to be the one least focused on. In general, it is positive to see that the organizations all have a degree of readiness, but it is also worrying that they are all, to some extent, vulnerable to future attacks. Awareness of the threat is, in general, quite low, which is hardly surprising since it has not been given a great deal of-attention in research or in the popular press. Still, there are instances when it has started to draw attention to itself, such as the new Russian service of letting AI chatbots chat with women to make it easier for men to find partners online (Carr, 2007).

The model and the questions presented here can be used as a baseline for preventive measures in order to improve protection to such a level that it is possible to avoid being an easy mark for both automated social engineering and the traditional form of attack. By having good routines and education in all the areas, the likelihood of damaging attacks can be minimized. By visualizing the results in these graphs, we get an image that users can quickly and easily understand and grasp. It has proved to be appreciated in informal tests. This approach also has several other strong points. It is cheap, in man-hours, giving quite visual results, as well as being easy to communicate to the interested parties. Perhaps, most of all, the very model makes it easy to extend it into a more complete method of measuring readiness. It could, for example, be quite easy to extend with more questions, covering a wider base and thus giving a more complete view of the readiness, even if it would mean that the interviews would take a lot longer to conduct. It appears that the method is valid.

The model of deception has been developed for social engineering, but it is quite possible that it could also be valid for other attacks, and this whole approach, might, in addition, be thus adaptable to other areas of security work.

Our future efforts will involve work on using AI chat-bots to educate users on social engineering. We will then use the cycle of deception and the form of questions we developed in this study as a basis for the training. The same training and education that uses AI chat-bots can also be used to analyze answers given in an online chat between the AI chat-bots and a security representative for the organization, thus automating one part of the process for doing readiness testing like this.

In general, information security professionals have often been more reactive than proactive when dealing with security issues. It is the hope of the authors that with more research and attention to the threat of automated social engineering we can stay one step ahead of the attackers rather than one step behind.

References

Adams A. & Sasse M. A. (1999) Users are not the Enemy: Why users compromise computer security mechanisms and how to take remedial measures, Communications of the ACM, 42, 12, 40 – 46.

Brostoff S., Sasse M.A. & Weirich D. (2002). Transforming the ”weakest link”: A Human-computer Interaction Approach to Usable and Effective Security, BT Technology Journal 19, 3, 122-131.

Carr, N. (2007). The oldest profession combined with the newest technology. [Online]. The Guardian. Available from: http://www.guardian.co.uk/technology/2007/dec/13/internet.crime, (Accessed 1 May 2008)

ComScore (2007). Social Networking Goes Global, July 2007. [Online]. ComScore. Available from: http://www.comscore.com/press/release.asp?press=1555, (Accessed 1 May 2008)

Dalrymple, M. (2005) Auditors Find IRS Workers Prone to Hackers. [Online]. AP. Available from: http://www.infosecnews.org/hypermail/0503/9684.html [Accessed 10 Nov 2008].

Golder, S. A., Wilkinson, D. & Huberman, B. A. (2007). Rhythms of Social Interaction: Messaging within a Massive Online Network, 3rd International Conference on Communities and Technologies (CT2007). East Lansing, MI. June 28-30, 2007.

Gross, R., Acquisti, A. &Heinz, H. J. (2005). Information revelation and privacy in online social networks. In Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society (Alexandria, VA, USA, November 07 - 07, 2005). WPES '05. ACM, New York, NY, 71-80.

Hogben, G. (2007). ENISA Position Paper No.1 Security Issues and Recommendations for Online Social Networks, ENISA October 2007

Kruger, H.A. & Kearney, W.D., (2006). A prototype for assessing information security awareness, Computers & Security 25, 4, 289 – 296.

Lawton, G. (2004). Web 2.0 Creates Security Challenges, Computer Volume 40, Issue 10, Oct. 2007, 13 - 16.

Mitnick, K. & Simon, W. (2002) The Art of deception: Controlling the Human Element of Security. Indianapolis: Wiley Publishing, Inc., ISBN: 076454280X.

Nohlberg, M. (2005). Social Engineering Audits Using Anonymous Surveys – Conning the Users in Order to Know if They Can Be Conned. In Proceedings of the 4th Security Conference, Las Vegas, USA, 30 – 31 March 2005.

Nohlberg, M. & Kowalski, S. (2008). The cycle of deception - a model of social engineering attacks, defenses and victims, will be published in Proceedings of the International Symposium on Human Aspects of Information Security & Assurance (HAISA 2008).

O'Reilly, T. (2005). What Is Web 2.0 -- Design Patterns and Business Models for the Next Generation of Software. [Online]. Available from: http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html, (Accessed 1 May 2008)

Rosenblum, D. (2007). What Anyone Can Know: The Privacy Risks of Social Networking Sites, IEEE Security and Privacy, 5, 3, 40-49.

Winkler, I. & Dealt, B. (1995). Information Security Technology? Don't rely on it A case Study in Social Engineering. [Online]. Proceedings of the Fifth USENIX UNIX Security Symposium. Available at: http://www.usenix.org/publications/library/proceedings/security95/full_papers/winkler.ps, (Accessed 1 May 2008)

Phishing with Gifts as Bait: Measurement and Analysis of Phishing Attacks within a University Environment [1]

Martin Boldt and Marcus Nohlberg

Abstract

The increasing use of computers in today's society is, in addition to positive aspects, also associated with negative effects. One example being the escalated threat from computer-oriented fraud and criminal activity, such as phishing. These attacks often target the human in front of the computer instead of the actual computer itself by using various social engineering techniques. In this paper we describe a study based on a highly advanced phishing attack that was launched on 354 selected Bachelor degree students. The selection of these students was based on three groups; students with no technical education; students with firm technical, but no security education; and students with both firm technical and security education.

The results from the study show that 29 % in the first group were vulnerable to the attack, while this figure was 57 % and 61 % for the second and third group respectively. These results indicate that more technical and security education means greater vulnerability for social engineering and advanced phishing attacks. Our conclusion is that the security students were overly confident in their knowledge and their ability to spot attacks. This is especially alarming since the most common countermeasure against social engineering threats is education. In the conclusion of this paper we give suggestions on how security education could be improved to address this problem.

Key words: Phishing, social engineering, information security.

[1] A version of this paper has been submitted to the *International Journal of Information Security*.

Introduction

As computers become increasingly more commonplace in today's society we also see a parallel increase in computer based fraud and crime [16]. Just as normal computer users benefit from the positive aspects of computers in the form of conveniently available e-commerce and e-banking services, so do antagonistic actors involved in illegal activities. These villains have traditionally exploited computer users' poor security concern regarding the security of their computers and network. By exploiting, for instance, users' operating systems to distribute viruses, spyware programs, or spam they could create revenues. When the protection techniques against these technical threats mature, and increasingly more people use personal firewalls and anti-virus software, the attackers need pioneering ways to attack their targets. Based on the path of least resistance [15] the attackers therefore start to exploit the gullible human in front of the computer instead of the actual computer itself. By using *social engineering* tricks the attackers manage to extract confidential information directly from the people that are authorized to access it, usually by telling some plausible lie [1]. The main problem with social engineering from the attackers point of view is that social interaction takes a lot of time in each attack. In order to make attacks efficient, they use a combination of social engineering and spam, referred to as *phishing*. In phishing the attackers send, for instance, large amounts of e-mails that aim at tricking computer users into revealing sensitive information, e.g. their bank credentials. Social engineering is a completely different type of attack compared to more technical attacks, such as computer worms or network intrusions, that is, firewalls and anti-virus software will not help against social engineering threats. Therefore, the most common protection technique against social engineering attacks has been user education. By informing computer users about the threats and hazards that are associated with computer usage on the Internet it is thought that the users will become better at identifying such scams [12].

The aim of this work is to investigate whether or not computer security education really is a good protection technique against phishing attacks. Therefore we exposed unwitting persons who have a firm understanding of the computer security field to an advanced phishing attack. However, the attack was of course set up in a controlled fashion. Previous phishing studies focus on a broader cross-section of users, often with little or no security education at all. Such studies often conclude with recommendations pointing towards user education as a key countermeasure against various types of social engineering attacks. That is also why we wanted to investigate whether or not the use of security education really helps computer users protect themselves against phishing attacks.

In the following section we describe central concepts such as social engineering and phishing in more detail. This is followed by a description of the method used. We then present our results, as well as the analysis of these results together with the discussion.

Background

Social engineering is a technique in which an unauthorized person manages to pose as an insider or an authority to successfully get access to information or resources [9]. A hacker can use social engineering to access other valuable data to benefit the hacker in further attacks [7]. Perhaps Mitnick gave the most useful definition in an interview by Tanneeru [18]:

> "Social engineering is using manipulation, influence and deception to get a person, a trusted insider within an organization, to comply with a request, and the request is usually to release information or to perform some sort of action item that benefits that attacker. It could be something as simple as talking over the telephone to something as complex as getting a target to visit a Web site, which exploits a technical flaw and allows the hacker to take over the computer."

A social engineering attack focuses primarily on the people's vulnerability, and is based almost entirely on using "the principle of easiest penetration" [15]. The greatest threat is that no matter how secure the system is in itself, it could never be more secure than its users [6] [12]. In addition to this, social engineering can be used instead of, or in combination with, threats and bribes. The classic social engineer aims towards not leaving any traces, and generally leaving as little of an impression as possible, and thus threats and bribes are not favorite weapons of choice [12]. Social engineering is used because it is often much easier to simply ask the target, for information, than to prepare and conduct a software or hardware attack [6] [12].

Phishing is an example of this. Manipulative techniques from social engineering are used in combination with techniques used in spam, to form an efficient attack vector affecting vast numbers of computer users at once [8].

The difference between social engineering and phishing lies within the scope of the attacks, and the delivery. A social engineering attack is targeted towards a single, often specifically selected person (or organization). While the attackers spend notable resources on attacking that individual, a phishing attack traditionally employs techniques similar to those used by spam in order to target thousands, or even millions, of users. Thus, very little resources are spent on each individual, making the phishing attack use the economy of scales. The traditional example of phishing has been an e-mail supposedly sent from your bank, asking you to log in to a new site and update your information. Upon doing so, you unknowingly submit your login

information to the attackers. While this technique has proven remarkably efficient before, the attackers are turning to more sophisticated attacks in order to increase the gain from their attacks. One example of this is *spear phishing*, which does not use the wide attack patterns of phishing, but instead the sending of highly targeted e-mails created from the use of data mining. The trick is to make the sender seem like someone the mark actually knows, or is associated with [11], ideally at a time when the mark is expecting to receive the message [8]. This specific targeting makes spear phishing much more dangerous to computer users than ordinary phishing, and professional attackers are probably more prone to use it in order to get financial gains, trade secrets or even military information [14]. Spear phishing could be seen as the "perfect" mix of social engineering and phishing, and it seems that it is also a lot more efficient, and dangerous, than ordinary phishing [14].

Related Work

In order to obtain a better sense of the actual level of security in an organization, it is common to use penetration testing. Historically, most penetration testing has focused on technical flaws, such as computer networks, operating systems, and other software vulnerabilities, but testing the human element of security is more difficult. Penetration testing can be used to proactively identify and address security vulnerabilities, thus often making it an integral part of information security. According to Barrett, it is preferable to use an audit style, which has results and objectives that are clear and can be accepted by both subjects and company [3]. Furthermore, these results should not lead to discipline or dismissal measures against the individuals.

An academic approach to a phishing audit was done by Hasle, et al. in which they performed two tests against a subset of users [7]. The first was a survey where the users were asked to submit their login information in order to authenticate if they were to win a prize. The second test was an e-mail sent out which triggered a login box. They also have an excellent discussion on ethics. In their tests approximately 30 % of the targets submitted their passwords, However, they mention little about the characteristics of their subjects, and we assume that they were ordinary users rather than specialists. It is also notable that no subjects informed the administration about the e-mails they received. A further interesting and somewhat similar study dealing with were done by Jagatic, et al. [19]. In this study they used data from social networks when trying to phish students from their university. They also validated the accounts, and found that when using data gathered from the social network in order to pose as someone the victim knew in the attack, 72 % of the users submitted valid logins. When doing a more traditional attack not

using the data from the social network (but claiming that the attacker was an unknown person from the same university), 16 % fell for the attack.

Both West Point Military Academy and the New York State has used penetration testing in order to train the users [2] [5]. Their approach has been to teach and train their employees by purposely testing them with phishing attacks. In the case of West Point, students were sent an e-mail from a person claiming to be a Colonel, ordering them to click on an attached link to verify their grades. This approach got 80 % compliance from the students, who were later informed of the risks of their behavior. In the case of the New York state, educational materials were sent to the employees, but later on 15 % of the employees tried to enter their passwords into a special online "password checker" after receiving an e-mail from the "Office of Cyber Security and Critical Infrastructure Coordination", urging them to do so. A follow-up, several months later, using a similar approach, got a lower compliance rate (8 %), indicating that the users had learned from the first experience. A less scientific, but still relevant study, was the update to the famous "Chocolate for passwords"-study in which both office workers and IT-professionals were offered candy if they gave away their passwords, an offer that about two thirds of the subjects accepted[10]. The obvious problem with this study is of course that they did not validate the passwords, making it quite possible that the subjects were social engineering the researchers out of chocolate by making up passwords. When we created our study we struggled with ethical constraints and how we should set-up the study. Now there is a very well written paper by Jakobsson, et al. [20] that explains the pitfalls on constructing phishing studies, and gives good advice that is useful for any researcher aiming to do studies in phishing. The difficulty of teaching security in general, and phishing in particular, is discussed in [21]. They use a novel approach by teaching security through the use of cartoons, and mentions the fact that there often is a great discrepancy between what users know, and what they practice. This might be more apparent when the subjects are a group that assumes that they know quite a lot, such as our security student subjects.

In the following section, we describe the phishing investigation that was carried out at Blekinge Institute of Technology, a Swedish university, during the winter of 2007.

Method

In this study we included subjects from the following three bachelor programs; *information security*, *computer game programming*, and *nursing*. All three bachelor programs include first, second and third year students and all of these students were included in the study. The students that were accepted from the security and game development programs needed similar grades to

be accepted. These figures were 11.25 and 10.78 respectively, where 0 is lowest and 20 the highest possible score. The grades needed for acceptance in the nursing program, on the other hand, were higher at 14.87. In total the study included 354 subjects divided between the three educational programs and classes, as shown in Table 1.

Table 1. Number of students in each class of the three Bachelor programs included in the study.

Educational program	1st year students	2nd year students	3rd year students
Information security	15	16	20
Game development	20	20	13
Nursing	82	75	93

Obviously we included the security students in the study to be able to measure whether their security education increases their resistance against phishing attacks. All security students are trained in understanding the nature of security threats and how techniques used to address them work. When it comes to Phishing, all security students had carried out a mandatory assignment in a previous course. The goal with this assignment was to teach the students how to distinguish legitimate e-mails from Phishing attempts, which taught them what clues to look for when trying to detect Phishing attacks.

The game development students were included since they have a firm technical understanding, but lack any explicit security training. We included the nursing students since they lack both the technical and security education. To be able to measure the impact that a security education has on users when exposed to a phishing attack, we actually performed such an attack. However, the phishing attack was executed under anonymous and strictly controlled forms. After the attack, all subjects were informed about the investigation and they were also encouraged to provide feedback through an anonymous survey.

Constructing the Phishing Attack

We wanted to create a highly sophisticated phishing attack relying on spear phishing techniques. The actual scam was based on an e-mail sent out by an imaginary person, Håkan Svensson, promising the subjects an early Christmas gift from the university if they registered themselves on a web page using their university accounts. The phishing scam was written in reasonably good Swedish, which up till now has been quite rare. Even though it made use of several manipulation techniques, it still included some weaknesses that should have been perceived as somewhat strange by attentive subjects. First of all, we spoofed the sender's e-mail address so that it would be impossible to send an answer to it without receiving an error message. Second-

ly, we wrote the e-mail in the fictitious name of Håkan Svensson, when no such person was employed at the university. While it is unfeasible to assume that every student should know every person employed at the university, even at a small one such as BTH, it is possible to search for employees on the university website. Third and last, the login page accepted not only valid student accounts, but also invalid ones, that is, the subjects could log in using *any* username and password. These university accounts are used for accessing the student e-mail server and other university systems, for example, when registering or discontinuing courses. However, the accounts are also used when accessing Windows and Unix computers at the university, both physically and remotely over the Internet, which of course is very interesting for attackers. It is also quite possible that the students use the same username and password at other sites too, as argued by Kelly [10], which make them even more interesting for an attacker.

Our phishing study is based on a scenario where an attacker wants access the university computers for questionable purposes. To reach this goal, the attacker chooses to identify one of the teachers' computers that includes some sort of security vulnerability, for example, a software vulnerability. By exploiting this vulnerability, the attacker obtains access to the computer where he/she installs a small web server and database, allowing the computer to act as a fake login server. To set up the login page the attacker simply makes an identical copy of the page that the students face when logging in to check their student e-mail. At this stage the attacker controls a fake login page that is identical to the university login page, and which is also placed on the university network domain, further imposing authenticity. Next, the attacker sends out an e-mail message through publicly available e-mail lists, one such list exists for each class at the university, convincing the students to login on the fake login page with their usernames and passwords. As soon as the students login, the attacker collects their login credentials. In the final step, the attacker cleans up all traces on the exploited machine and then logs in on the hundreds of computer to which he now has access.

When we began constructing this phishing study, we made sure that the person who designed the attack did not have any knowledge at all about the technical infrastructure at the university. This allowed us to ascertain that an attacker without any connections to the university could actually setup such an attack, rendering an authentic real world scenario.

The Phishing E-mail

Sweden has been relatively spared from any large scale phishing attacks and the few we have received have been of poor quality, Many Swedish people therefore tend to think that phishing only happens in English, and only for a couple of specific targets, such as banks. We therefore wanted to see how efficient a spear-phishing attack written in Swedish would be against some-

thing other than a bank account. Another reason for choosing Swedish was to match the communication pattern of the tested subjects, since they rarely receive e-mails from the school in English.

As shown in Figure 1, we designed the e-mail message to closely mimic a well thought out and planned spear-phishing e-mail properties. The idea was to use some classic manipulative techniques inspired by Cialdini and Nohlberg, such as authority, by claiming the message was from a senior researcher, scarcity, by claiming the gifts were a limited time offer and by having limited amounts of the most sought after goods [4] [13]. The products were selected to range from cheap and common, to somewhat more exclusive. On the webpage, the number available of the most expensive product rapidly decreased as more and more students logged in to emulate scarcity. The products were also selected to be reasonable as promotional gifts, and not too expensive, as this could be suspicious. The time available for the students to log-in and claim their products was short; they had to do it as quickly as possible so that they would not miss the chance. The e-mail was sent out on a Friday afternoon, with the assumption that the students would have a harder time asking a teacher or contacting support, which thus mimicked a real attacker's logic.

Hi
My name is Håkan Svensson, and I am a new employee at Blekinge Institute of Technology and I will work with the marketing of some specially selected educational programs within [security | game development | nursing]. As a part of this and due to this year's good result, we would like to offer all students in these selected programs a commercial gift from the university as an early Christmas present. You can choose between a t-shirt, baseball cap, laptop bag or a fleece hat. Some of the products are in limited supply and will be served on a "first come, first serve" basis.
In order to simplify the handling of the orders this will be done digitally. You log in using the URL below and select which product that you want, size and color. You have to login so that we can know that you are a student on one of selected educational programs.
https://apshsv.tek.bth.se

In order for us to guarantee delivery of your product before the holidays you must log in and register on the 14th of November at the latest. The products will be delivered to [your city] at around December 15th. More information about the delivery will be sent to you in a later e-mail.

Best regards,
Håkan Svensson, Associate Professor

Figure 1. A translation of the e-mail message written in Swedish that was sent out to the subjects.

Implementation

The actual implementation of the fake login page was made in PHP, a script language commonly used on the Internet. We also used a MySQL database to store statistics regarding the attack. The fake login page was placed on a well-protected computer running the OpenBSD operating system and Apache 2 web server. This computer was connected to one of the teacher networks at the university, according to the real world scenario. Another

advantage with this arrangement was that the network traffic was kept within the university network perimeter as much as possible, in order to protect the transmitted data as far as possible. In the real world scenario the attacker would simply store victims' usernames and passwords as they log in to the fake page. However, we did not collect the subjects' login credentials. Instead, we validated them against the university LDAP system, which allowed us to distinguish valid university accounts from invalid ones. For each validated account we also extracted what educational program the subject was registered on and which class he/she belonged to. This allowed us to keep track of exactly how many valid logins were made, and which class each of these subjects belonged to, that is, we could see exactly how many students from each class that logged in using their university account. In addition to this we also kept track of all invalid logins, but of course without any coupling to educational programs and classes. It is important to stress that the login credentials were never stored and that all validation was done in an automatic manner, that is, no one would ever see these credentials during the study.

All network communication was encrypted using SSL to protect the confidentiality of the login credentials, for example, between the web server and the LDAP server. As can be seen in Figure 1, the URL to the fake login page also uses SSL encryption since it begins with https. For this to work we had to rely on a self-signed certificate instead of a stronger certificate that is signed by a trusted third party, such as VeriSign. The use of self-signed certificates is usually a bad idea due to the fact that you cannot trust the content in them to be correct. It is similar to issuing your own ID card instead of applying for one from the government. In fact, the only advantage with self-signed certificates includes the ease of creation and that they are free of charge. However, since you cannot trust the content in the certificate to be correct, all web browsers show a warning message every time a non-trusted certificate is used, that is, one not issued by a trusted third party. A rule of thumb is that one should be careful when visiting websites using self-signed certificates since it is not possible to rule out that it is a fake version of the site, possibly included in a scam. Most servers at Blekinge Institute of Technology, as at most other Swedish universities, unfortunately use self-signed certificates, for instance, in student e-mail servers. Consequently, the users are exposed to these warning messages every time they log in to these sites, for example, every time they check their e-mails. Due to the frequent display of these warnings most students learn to simply ignore them, similar to the 'boy who cried wolf' scenario. This, however, is good news for an attacker, since he/she can use this user ignorance against the students when including self-signed certificates in a scam. It should be pointed out that Blekinge Institute of Technology has now finally begun to implement certificates issued by a well-respected third party.

Execution

The phishing e-mail was sent out to game and security students on a Friday afternoon in mid-November 2007, with a deadline specified to Tuesday the following week. Once the deadline had passed, we closed down the fake login page and informed the students about the study. After analyzing the results from the game program and security students, we wanted to include a third group of subjects with little or no technical experience at all. We therefore subjected three classes of the nursing program to the same attack. The same e-mail was also sent out on a Friday, two weeks after the initial attack on the security and game students. We chose to include the nursing students since they lack the technical based computer science education of the other subjects. Furthermore, the nursing students also lack the information security education of the security students. In addition, the nursing students are located at a different campus in another city, than the security and game students. This is important since we wanted to find a group that has no, or at least very limited, contact with the first group, which was the fact in this case. Since there was no prior public announcement of the study and it received no media attention beforehand, we can assume that the nursing students did not know about the study at all.

When the deadline of the second study had passed, we informed the nursing students that the e-mail message was part of a well controlled and scientific phishing study. This information was sent out using a teacher's official university e-mail address and the students were encouraged to reply if they had any questions. At this time we also announced that we would arrange an open seminar about phishing in which the results from the study would be presented. Furthermore, we allowed all subjects involved in the study to participate in an anonymous web-based survey.

Follow-up Survey

In order to know more about why the students did or did not fall for the attack, we asked them to fill out an online survey. It was anonymous and consisted of a number of questions on the students' background information, but also questions about what they thought of the study, if they felt they had learned from it, if it had been offensive to them to be exposed to this kind of test, what they had felt about the e-mail, the site, and so on. The web survey used LimeSurvey, an open source survey software, for its design. A pilot study of the survey was tested on five different persons, aged between 15 and 66 who had no connection to the previous phishing study, in order to check that the questions were understandable. Based on their feedback, minor changes were made before an e-mail was sent to the subjects asking them to partake in the survey. After about a week, a reminder was sent out to

the students. The data gathered were analyzed in LimeSurvey and Microsoft Excel.

Ethical Concerns and Informing the Subjects

This kind of study is quite sensitive and associated with ethical concerns. Consequently, we spent a lot of time discussing the ethics of performing such a study and how to design it. Firstly, we obtained the required authorization from those responsible for the educational programs and the heads of the departments at the university. We used no specific inside knowledge or data when designing or deploying the attack. All background data used was gathered using public sources. While the students' real log-in credentials were validated, they were not saved but always encrypted, and they never left the school's own computer systems, that is, we could at no time see any of this information. Anonymity, and security, were guaranteed.

We also made sure the students were informed, as quickly as possible after the study was finished, that they had been part of a controlled phishing scam. This information was sent out through e-mail, and the students were asked to reply if they had any questions or comments. However, nobody chose to reply. After the students were informed, we arranged an open seminar on social engineering and phishing in general, and this study in particular. During this seminar we tried to find out how the subjects had experienced the phishing scam and how they felt about the study.

In addition to the public seminar, all the subjects also had the possibility of voicing any concerns/critique in the anonymous survey that was published after the study. With regard to the ethics of the study, the survey revealed that the majority of students found the study had been completely non-intrusive, or only slightly disturbing, as Figure 2 illustrates. However, there is a limited group of students that found the study violated their privacy, but there were no written remarks from this group explaining why, even though the survey revealed a large number of positive comments.

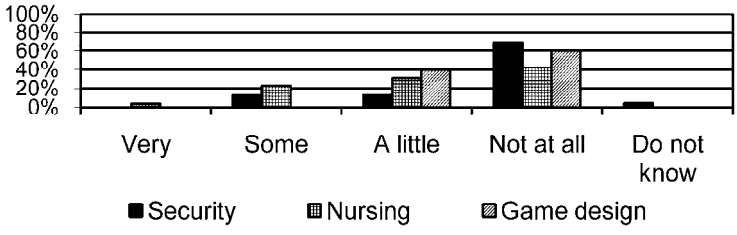


Figure 2. Do you find this kind of study violating?

The survey also revealed, see Figure 3, that the vast majority of subjects felt it was good that the university had conducted this study. Should the universi-

ty execute more studies, the vast majority of students would prefer fewer but bigger tests instead of smaller ones occurring more frequently.

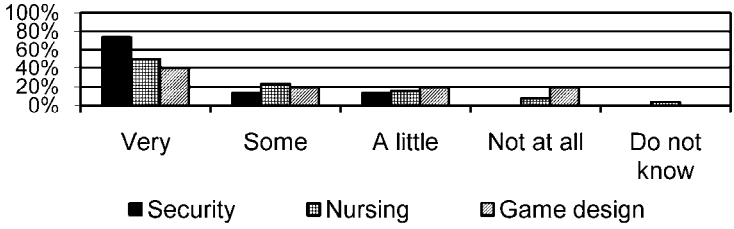


Figure 3. Do you think that it is good that the school does anonymous security studies like this one?

It is also important to consider that this type of study is probably the only possible way of examining how well equipped subjects are in protecting themselves against social engineering attacks. It is hard, if not impossible, to test for resistance against phishing attacks if, for instance, prior consent from the subjects must be acquired. Since the study was anonymous, and prior consent from those responsible had been obtained, and since the subjects had been provided with ample information and possibilities to voice their concerns afterwards, we conclude that we achieved an acceptable ethical level for this study.

Results

In this section we present the results from the phishing study followed by the results from the survey.

Experiment Results

The results of the phishing experiment show that 133 of the 354 students included in the study had logged in with their valid university accounts. This means that the total rate of submitted log-ins is 38 %. In addition to the subjects included in the investigation, three more students from other educational programs also logged in using their university accounts. Since we did not send any phishing e-mail to these students they probably received an “invitation” from friends in one of the three included Bachelor programs. The result, including these three additional students, totals 136 valid, “compromised”, university accounts. During the study there was a total of 24 false logins from the security and game developer students. Among the nursing students there were only 2 false logins.

Table 2 shows the statistics for the game development students, which reveal that an average of 57 % of the subjects logged in using valid accounts. There is a slightly higher representation of first year students than second and third

year students, but these differences are not statistically significant. All in all, 30 game development students fell for the attack.

Table 2. Number of game development students in each class that logged in using valid university accounts.

Educational program	No# of students with valid logins	No# of students in class	Per cent of students in class
1st year game students	12	20	60 %
2nd year game students	11	20	55 %
3rd year game students	7	13	54 %
Total	**30**	**53**	**57 %**

Table 3 shows the results for the security students. In total, 31 security students, or an average of 61 %, logged in with a valid university account. There is a slightly higher representation of first year security students than second and third year students, but these differences are too small to be statistically significant. Furthermore, it was not possible to find any decreasing trend among the students who fell for the scam that was related to which year they were studying in, for example, that more experienced students should be better at recognizing a phishing scam. In fact, the security students showed the highest average, 61 %, of the subjects in the three investigated Bachelor programs, which struck us as quite surprising at first. This, among other things, is considered in the discussion section below.

Table 3. Number of security students in each class that logged in using valid university accounts.

Educational program	No# of students with valid logins	No# of students in class	Per cent of students in class
1st year security students	10	15	67 %
2nd year security students	8	16	50 %
3rd year security students	13	20	65 %
Total	**31**	**51**	**61 %**

A total of 72, or 29 %, of the 250 nursing students included in the study, logged in using their university accounts, as shown in Table 4. The second year nursing students have a slightly higher representation than the first and third year students.

Table 4. Number of nursing students in each class that logged in using valid university accounts.

Educational program	No# of students with valid logins	No# of students in class	Per cent of students in class
1st year nursing students	25	93	27 %
2nd year nursing students	29	75	39 %
3rd year nursing students	18	82	22 %
Total	**72**	**250**	**29 %**

Survey Results

The survey had an answer rate at 17 %, since we received 58 complete answers from a population of 354, which we argue is acceptable for a web-based survey. It is also worth mentioning the inherent problem with asking the students to click on a link in an e-mail after having been exposed to a phishing study. It is possible that some felt it was another phishing attack and therefore unsafe.

The security students supplied 39 % of the answers , the nurses 44 % , and 17 % came from the game design students. When factoring in the total number of students, the security students had the highest answer rate. A majority of the answers came from third year students (49 %), followed by second year (37 %) and first year students (12 %). Most of the subjects were between 18 – 29 years old (87 %), and men were somewhat in the majority at 61 %. On average, the subjects had been using the Internet for 10 years, and no one had used it for less than five years.

When the subjects were asked about their security consciousness on the Internet, the nursing students had less confidence in their abilities than the security and game students, see Figure 4. It is clear that the security students and the game design students consider themselves to be quite security conscious, while the nursing students have a somewhat more humble attitude.

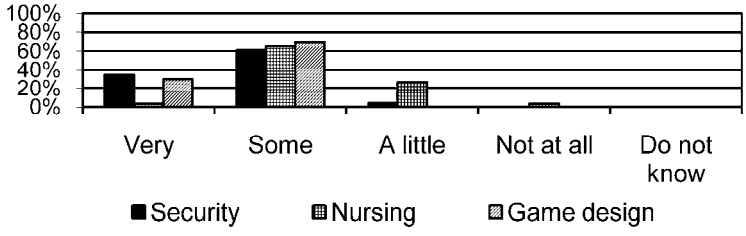


Figure 4. How security conscious do you consider yourself to be when using the Internet?

With regard to aptitude in recognizing frauds, the same trend is apparent in Figure 5, but in this question the security students rank themselves second to the perceived skills of the game design students.

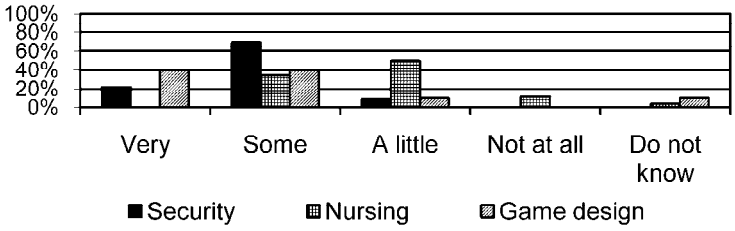


Figure 5. How good are you at spotting frauds?

The phishing e-mail was considered highly believable among all the students, but interestingly enough, the security students regarded it more credible than other students, while the nurses were the most skeptical, as Figure 6 illustrates.

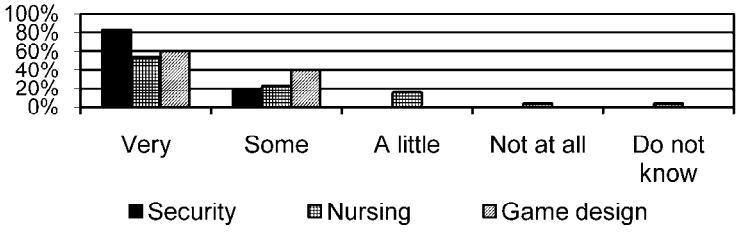


Figure 6. How credible did you think the e-mail was?

The website, which the students were directed to after clicking on the URL in the e-mail, was considered either very or to some degree credible by 61 % of the subjects. However, it is interesting that the security students found the website to be more credible than both the game and nursing students, as shown in Figure 7.

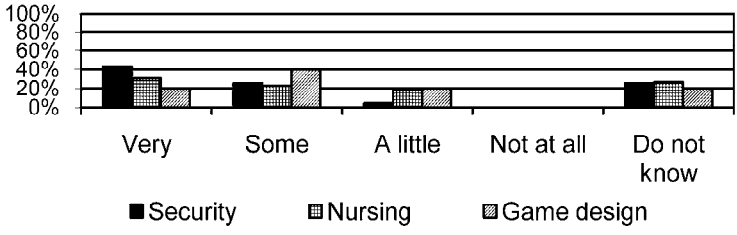


Figure 7. How credible did you think the website was?

The URL in the e-mail was perceived as far more trustworthy among the more technically knowledgeable students, than among the nurses, as Figure 8 reveals. This was probably due to the use of https and since a DNS name within the university network was used. Both of these details require a certain technical awareness to see and reflect upon.

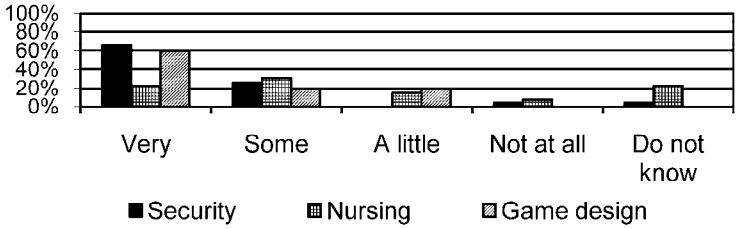


Figure 8. How credible did you think the URL in the e-mail was?

When asked if the products offered in the phishing e-mail were articles that the subjects actually had an interest in, it seems that the giveaways were somewhat attractive at best. The products offered might stereotypically be considered more attractive to male students than female ones, and therefore not as tempting to nursing students, who are predominantly female. Considering the survey answers, it does in fact seem that the nursing students found that the giveaways were more attractive than the security and game design students, as seen in Figure 9. However, it must be pointed out that only 10 % of the nursing students participated in the survey, that is, the opinion of the majority of nurses is unknown.

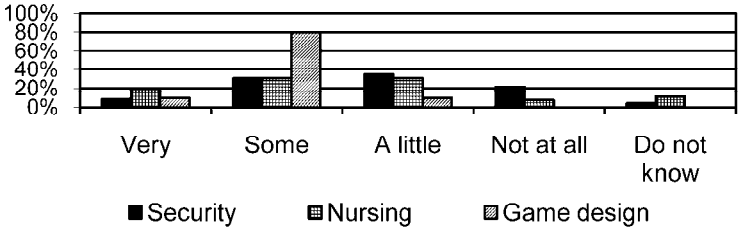


Figure 9. The attractiveness of the products offered.

Some subjects, especially the security students, believe that they have increased their security consciousness after being part of the study, as illustrated in Figure 10. This also applies to the nursing students, but to a lesser degree. However, the game students found the study to be of limited use. How a study is judged afterwards by its participants is still not a conclusive argument of its efficiency, but one could argue that the security students should have a better security understanding and can therefore better judge what they have learned.

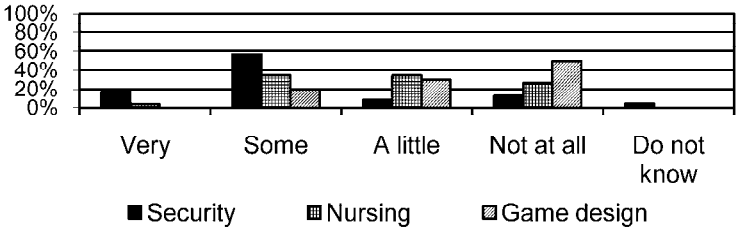


Figure 10. Do you consider yourself more security conscious after this study?

In order to avoid stigmatizing the subjects in the survey, we did opt to not ask them directly if they fell for the attack. Instead we asked them if "they submitted their actual login information to log-in to the website", which amounts to the same thing. The results are shown in Figure 11.

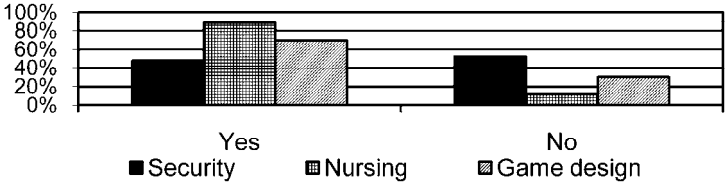


Figure 11. Did you enter your actual login information?

These results can be compared to those from the actual attack, shown in Figure 12. Clearly the nurses who fell for the attack were more prone to answer the survey. The same is true, but to a lesser extent, for both security and game design students. This might imply that some students, particularly nurses, did not read their e-mails (either the phishing attack, the survey e-mail, or both), but it might just as well indicate a greater interest in the study among those who "fell" for it.

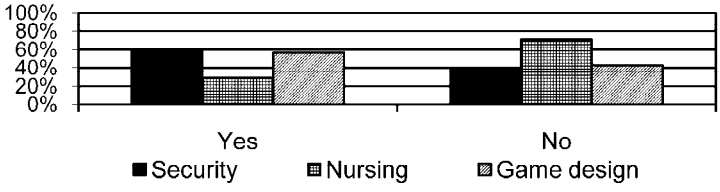


Figure 12. The actual distribution of those who fell for the attack judging from the passwords gathered.

Discussion

This study yielded interesting, and surprising results. Our first surprise came when we saw that the security students "fell" for the attack to a surprisingly high degree. Our response to this was to launch a new test on nursing students to see if the attack was "too good". As it eventuated, the attack was less successful on the nurses, indicating that those who should be the most

security conscious were, in fact, the least. It is very hard to ascertain exactly why the result turned out this way, since there are several factors that influenced the final conclusion of the study. We will, for instance, never know which of the subjects actually opened and read the e-mail. If we knew this, it would be possible to distinguish the students who read the e-mail but chose not to login, from those that did not even check their e-mail during the time-period of the study. However, we do know that all students rely on their e-mail to receive important information about their education, e.g. schedule changes. So, all students included in the survey check their e-mails on a regular basis.

During the seminar discussion with the students, and also in the survey answers, we found that the security students evaluated the attack based on their previous knowledge of technical aspects. For example, they noted the use of SSL, checked the validity of the SSL certificate, made reverse and WHOIS lookups on the domain name, and analyzed that the text in the e-mail was (reasonably) correct in language. Since none of these checks revealed anything unusual, they assumed that the e-mail was legitimate. By using these evaluation techniques they would identify traditional phishing attacks, but not this advanced form of spear phishing. Consequently, they used a correct, but insufficient, set of metrics on whether this was an attack or not. Since their judgment relied on only a limited set of tests, they were tricked without difficulty, as this attack was not easily spotted by those particular metrics. This is in accordance with the results from [23], that people trained to look for certain indicators of secure transactions are more likely to fall for attacks that spoof such indicators. People are good in handling risks they know, but poor in handling new ones. The subjects could have recognized the scam if they, for instance, verified the existence of the sender of the e-mail, tried to send an answer to the e-mail, or by trying to login using a fake username and password. If one, or more, of these techniques had been used it would have been obvious that the e-mail was strange and not to be trusted. On the other hand, recommendations to check indicators such as these are only temporary counter measures. If everyone checks the validity of a site using a fake login, the attackers would start to always tell the victims that the login will be incorrect the first time. It is a continuous arms race.

Based on the results from the survey, the nursing students assessed the attack in other ways than the security and game students did. The nursing students seemed to have a different, and in some cases flawed, mental image of what could be trusted or not. These flawed mental images are in line with the results from Sasse, where it was found that users tend to create their own, often incorrect, mental models of how security works [17]. One nursing student had complete faith in the sender of the e-mail "as he claimed to have been hired recently, he seemed like a trustworthy man since you have to assume that BTH does not employ just anyone". Another nursing student

claimed to trust the e-mail since all the e-mails sent to students must be authorized by a boss. There are other examples and they all show that many of the nursing students have a flawed mental image of how security works. It is interesting to note that despite these flawed mental models, the nurses, as a group, were the ones most resilient to the attack. Unfortunately, we cannot identify the reason for this.

At first we thought the difference between the student groups could be explained because the attractiveness of the gifts offered would be regarded differently among these groups. However, the survey indicates that this is not the case since all the student groups showed a similar interest in the products offered. Nevertheless, since only a minor portion of the students participated in the survey, the opinions of the majority of the subjects are unknown. An additional explanation could be that the nursing students do not check their e-mails as often as the security and game programming students, that is, they did not receive the phishing e-mail until the study was already finished. However, this is unlikely since all the students at the university, including the nursing students, rely on their e-mail to receive important information from their teachers regarding their education, for example, courses and exams.

It is also notable that, apparently, there were no warnings spread among the students. No one asked their teachers or other senior staff about the offer. There was no incident reporting to the university abuse list. Therefore, even if some of the students recognized the attack, they did not warn the university staff or their fellow students about it. In fact, some security students even tried to login using a fake name and password, which was accepted. However, in fear of losing the offered product, they then logged in a second time using their real username and password. We believe this is a good example of how the craving for a free gift can disable the security consciousness of humans. As they try to solve a primary task, in this case to receive a free gift, they do not consider the clues indicating that something is wrong.

The study showed that the subjects also recruited other students who were not in any of the educational programs that were part of the study. One could say that the phishing attack got a life of its own. The lure of free “stuff” was so great that non-target students were recruited by others who were already victims of the attack. We believe this unintended effect could be exploited in an attack designed specifically for that, for example, similar to network marketing or pyramid schemes. We believe that such an attack could result in an alarming scenario.

With regard to security education in general, the results from this study surprised us. The highly trained security students fared far worse than the nursing subjects, who we assumed would be much more likely to fall for these attacks. This is contrary to what one could expect considering, for instance

[22], the argument that a deep understanding of the web environment is associated with reducing the risk of falling for phishing attacks. It is probably undoubtedly true that the security and game design students are more knowledgeable about the Internet than the nursing students. The question then is why did we get these results? If we consider the traditional way of measuring awareness with behavior, attitude, and knowledge, we can assume that the security students have a very high degree of attitude and knowledge. However, it is possible that their high degree of attitude and knowledge makes them overly confident in their behavior, that is, since they know so much they believe they behave correctly. The nursing, and game development students on the other hand are aware of their lack of security knowledge, and are thus more restrictive in their behavior. Even though the subjects in this study might be young with little practical security experience, the security students do have enough interest, as well as several years of education, in this field to make it a career.

The results from this study indicate that we might need to rethink security education for ordinary users in general, and specialists in particular. Besides teaching computer users security in general, it is also important to teach them how to act in certain situations. Perhaps a lot of security work on empowering ordinary users with knowledge is wasted - what they might need to learn instead is how to act in given situations. A parallel could be drawn to the instructions given to industrial workers operating dangerous machinery. These workers do not need to know exactly how the machine works, only how to act so that they do not hurt themselves or others while using it. An approach that is more of a compromise might be better, however. An example of this is to combine the testing done in a phishing study such as this, with immediate education of the users afterwards, if they fall for the attack. This is referred to as embedded training, and recommended by Kumaraguru [24].

To summarize, the deployed study is one of the most advanced phishing attacks carried out in Sweden, especially considering the set-up and the formulation of the e-mail. There have been others with greater media attention, but they have been based on traditional Phishing attacks with a low degree of complexity. What is notable here is that it still did not take more than two days work to design and construct the attack, that is, the investment in time for a potential attacker is very low. Repeating the attack on several different organizations would only require a small overhead on the work. Considering the success rate and the amount of work required to launch the attack, the pure economics of spear-phishing is obvious. It yields extremely good results for the attacker, at a reasonable cost. We therefore think it is important to further analyze social engineering attacks; and as our knowledge about these attacks increases, so do our odds for developing measures that ordinary computer users can rely on as protection.

Lessons Learned

While this study provided interesting results, there are a number of issues that we believe should be noted as possible future improvements. One of our main concerns is that the three studies were not done in parallel. The first two attacks on the security students and the game design students were done in parallel, but the second study on the nurses was conducted a couple of weeks later. While we have no indications that there has been any leakage among the subjects, this is obviously not an optimal approach. There is also a possibility that the balance between risk and reward was skewed too far to "the attackers" advantage. Perhaps the students did not feel that their school logins was important enough to warrant any particular caution when faced with the possibility of getting "free stuff". We do believe that the students are reasonably cautious with regard to their accounts, at least on a par with the average student, even if the risks for personal monetary losses are slim to none. Trying to Phish student account data is, however, the one scenario that we could test; carrying out a test like this against, for instance, bank customers would be much more complicated. Offering the subjects more expensive, or inexpensive, gifts would probably raise suspicions or not raise any interest at all. We focused on finding reasonably valued gifts that a university would be likely to give their students. The selection of gift, the bait, in this scenario is certainly an important one and our selection might have been more attractive for the "geeks" than the nurses. The survey results did not indicate that this was the case, but a future study should probably try to find a selection of possible "gifts" that attract their subjects equally.

When designing this study, we made the decision not to use any techniques for tracing the reading of e-mails, such as the web bugs that Hasle, et al. [7] used. The reason for this is that it would have been easily detected among the security students, but the most obvious problem is that it would not have measured anything with reasonable accuracy. Many of the security students do not use HTML-supporting e-mail clients, rendering that approach void. In the same manner, it would have been pointless to look at the e-mail servers to see if they had read the e-mails. The only thing we would have been able to read there is if the students had downloaded the e-mail or not, but not if they had actually read them. Also, many students automatically forward all their e-mails to external accounts which are impossible for us to control. At least we do know that the university Spam filter did not stop the e-mail from reaching the subjects.

After considering that there was no feasible way for us to know just how many students had actually read the e-mail, we instead decided to let the study run for a longer time, so we could assume that the students would have had the chance to receive and read the e-mail. If, in a future study, one could

find a suitable way of knowing the extent to which the e-mails were read, this would of course be an improvement.

The three groups that we studied were different not only with regard to the educational program they studied, but also in age, gender, and so on, which might have influenced just how successful the attack was. In this study our goal was to compare these fairly diverse groups, but studying more homogenous groups, than our rather heterogeneous ones, could make it possible to draw stronger conclusions on those particular groups. This was not an option for us at the time since we did not have access to groups large enough to make that option a good choice.

The set-up of the attack was quite sophisticated, perhaps making it unnecessarily hard to spot. While there are ways that this attack could have been made "simpler", such as using an external server, or no certificates, we had to ensure the students data. A real attacker cares little about the privacy of the victims, but in a scientific study that must pass ethical scrutiny options are limited. The scenario is not, however, unfeasible, even if a real perpetrator could have made the attack even more difficult to spot. However, testing the resistance to a "perfect" Phishing attack would provide little data of interest. We tried to find a balance between ethical considerations and the creation of a feasible attack.

Conclusions

This study indicates that the security students were not more resistant than others to the highly advanced phishing attack. The security students might, in fact, be even less resistant than non-security educated users since they are overly confident in their knowledge based ability to spot frauds because of their education. The results further show that the phishing technique is an effective means of tricking people into divulging sensitive information, which on the other hand is nothing new. What is new, however, is that it appears that traditional security education does not equal secure behavior, and nor does firm technical knowledge, as the game programming students demonstrated.

It appears that all types of computer users risk being vulnerable to this kind of attack. Our results indicate that the nursing students were vulnerable to a lesser degree than their more technically oriented counterparts, their degree of vulnerability were similar to the degrees found in the study done by Hasle, et al. [7]. The security and game development students, on the other hand, were more vulnerable in comparison. We think that the most probable explanation for this is that these students are deceiving themselves into thinking that they are security conscious in their behavior, when that is not the case. Therefore, their security education is working against them, and the

group that is aware of lacking this knowledge is better off since they are more skeptical and thus act carefully.

This difference can be discussed at length, but the one clear conclusion we can draw is that in this case security education did not result in secure behavior. What we as researchers should do is to see if these results could be generalized into also being valid in a larger context. If so, we have to take this into account when designing our security education, for example, by making it less generic and instead focusing on certain secure workflows that could guide users to act correctly in intricate situations. Ordinary computer users do not need to understand exactly why they do something to act safely, but they do need to know how to act safely.

Future Work

In this work we found interesting results regarding the efficiency of security education for this specific form of social engineering attack. We believe that these indications need to be validated through additional studies. In fact, this whole phenomenon can be studied in settings other than schools, for example, both smaller and larger corporations and organizations. Obviously, it would also be interesting to do follow-up studies to see if the students actually behave more securely after being exposed to this study and attending the seminar. If a similar study to this one were to be executed, we recommend that measures be taken to separate the subjects that do not see the e-mail from those that read it and instantly reject it.

One interesting side effect we found in this study was the recruiting phenomena in which unaware subjects acted as ambassadors for the attacker by recruiting friends in the desire for free goods, somewhat similar to network marketing. It would be interesting to study what protection techniques could be used against a phishing attack that is designed to get people to recruit their friends to the site, thus bypassing many current security mechanisms.

References

1. Anderson, R. (2001). *Security Engineering.* Danvers MA, USA: John Wiley & Sons.
2. Bank, D. (2005). 'Spear Phishing' Tests Educate People About Online Scams. [Online]. The Wall Street Journal. Available from: http://online.wsj.com/public/article/SB112424042313615131-z_8jLB2WkfcVtgdAWf6LRh733sg_20060817.html?mod=blogs [Accessed 20 Nov 2008].
3. Barret, N. (2003). Penetration testing and social engineering: hacking the weakest link. *Information Security Technical Report.* 8(4), pp. 56–64.
4. Cialdini, R. (1993). *Influence: the psychology of persuasion.* New York, USA: Quill.

5. Dodge, R., & Ferguson, A. (2006). Using Phishing for User Email Security Awareness. In Fischer-Hübner, S., (Ed.), Rannenberg, K. (Ed.), Yngström, L. (Ed.), Lindskog, S. (Ed.), In *Proceedings of the IFIP TC-11 21st International Information Security Conference (SEC 2006)*, pp. 454-458. New York, NY: Springer Science + Business Media Inc.
6. Granger, S. (2001). *Social Engineering Fundamentals* [Online]. Security Focus. Available from: http://www.securityfocus.com/infocus/1527 [Accessed 20 Nov 2008].
7. Hasle, H., Kristiansen, Y., Kintel, K. & Snekkenes, E. (2005). Measuring Resistance to Social Engineering. In *Information Security Practice and Experience: First International Conference, ISPEC 2005*, Singapore, April 11-14 (2005), vol. 3439 of *Lecture Notes in Computer Science*, Springer, pp. 132-143.
8. Jakobsson, M. (2005). Modeling and Preventing Phishing Attacks. [Online]. School of Informatics & Dept. of Computer Science, Indiana University. Available from: http://www.informatics.indiana.edu/markus/papers/phishing_jakobsson.pdf [Accessed Nov 20 2008].
9. Kajava, J. & Siponen, M. (1997). Social Engineering - IT Security Threat of Informatics [Online]. IRIS 20. Available from: http://web.archive.org/web/20040422210025/http://iris.informatik.gu.se/conference/iris20/9.htm [Accessed 20 Nov 2008].
10. Kelly, M. (2007). Chocolate the key to uncovering PC passwords [Online]. The Register. Available from: http://www.theregister.co.uk/2007/04/17/chocolate_password_survey/ [Accessed 20 Nov 2008].
11. Microsoft (2006). Spear phishing: Highly targeted phishing scams [Online]. Microsoft. Available from: http://www.microsoft.com/protect/yourself/phishing/spear.mspx [Accessed 20 Nov 2008].
12. Mitnick, K. & Simon, W. (2002). The Art of deception: Controlling the Human Element of Security. Indianapolis, USA: Wiley Publishing, Inc.
13. Nohlberg, M. (2008). Why Humans are the Weakest Link. In Gupta, M. (Ed), Sharman, R. (Ed). Social and Human Elements in Information Security: Emerging Trends and Countermeasures, Idea Group, Inc.
14. O'Brien, T. (2005). Gone Spear-Phishin'. [Online]. The New York Times. Available from: http://www.nytimes.com/2005/12/04/business/yourmoney/04spear.html?ex=1291352400&en=2f313fc4b55b47bf&ei=5088&partner=rssnyt&emc=rss [Accessed 20 Nov 2008].
15. Pfleeger, C. & Pfleeger, S. (2003). *Security in Computing* (3rd ed). Upper Saddle River, USA: Prentice Hall.
16. Rosenberg, R. S. (2004). *The Social Impact of Computers* (3rd Ed), San Diego CA, USA: Elsevier Academic Press.
17. Sasse, M. (1997). Eliciting and Describing Users' Models of Computer Systems. Doctoral Thesis. University of Birmingham.
18. Tanneeru, M. (2005). A convicted hacker debunks some myths. [Online]. CNN. Available from: http://edition.cnn.com/2005/TECH/internet/10/07/kevin.mitnick.cnna/index.html [Accessed 20 Nov 2008].
19. Jagatic, T., Johnson, N., Jakobsson, M., & Menczer, F. (2007). Social phishing. *Communications of the ACM*, 5 (10), pp. 94-100.

20. Jakobsson, M., Finn, P. & Johnson, N. (2008). Why and how to perform fraud experiments. *IEEE Security & Privacy Magazine*. 2008 March-April; 6(2), pp. 66-68.
21. Srikwan, S. & Jakobsson, M. (2008). Using Cartoons to Teach Internet Security. *Cryptologia* 32(2), pp. 137-154.
22. Downs, J., Holbrook, M. & Cranor, L. (2007) Behavioral Response to Phishing Risk. In *Proceedings of the 2nd Annual eCrime Researchers Summit*, October 4-5, 2007, Pittsburgh, USA, pp. 37-44.
23. Downs, J., Holbrook, M. & Cranor, L. (2006). Decision Strategies and Susceptibility to Phishing. In *Proceedings of the Second Symposium on Usable Privacy and Security*. SOUPS '06, vol. 149. New York, NY, USA: ACM Press, pp. 79-90.
24. Kumaraguru, P., Rhee, Y., Sheng, S., Hasan, S., Acquisti, A., Cranor, L.F. & Hong, J. (2007): Getting users to pay attention to anti-phishing education: evaluation of retention and transfer. In *Proceedings of the 2007 Anti-Phishing Working Groups eCrime Researchers Summit 2007*, pp. 70-81.

DEPARTMENT OF COMPUTER AND SYSTEMS SCIENCES
Stockholm University/KTH
www.dsv.su.se
Ph.D. theses:

No 91-004 **Olsson, Jan**
An Architecture for Diagnostic Reasoning Based on Causal Models
No 93-008 **Orci, Terttu**
Temporal Reasoning and Data Bases
No 93-009 **Eriksson, Lars-Henrik**
Finitary Partial Definitions and General Logic
No 93-010 **Johannesson, Paul**
Schema Integration Schema Translation, and Interoperability in Federated Information Systems
No 93-018 **Wangler, Benkt**
Contributions to Functional Requirements Modelling
No 93-019 **Boman, Magnus**
A Logical Specification for Federated Information Systems
No 93-024 **Rayner, Manny**
Abductive Equivalential Translation and its Application to Natural-Language Database Interfacing
No 93-025 **Idestam-Almquist, Peter**
Generalization of Clauses
No 93-026 **Aronsson, Martin**
GCLA: The Design, Use, and Implementation of a Program Development
No 93-029 **Boström, Henrik**
Explanation-Based Transformation of Logic programs
No 94-001 **Samuelsson, Christer**
Fast Natural Language Parsing Using Explanation-Based Learning
No 94-003 **Ekenberg, Love**
Decision Support in Numerically Imprecise Domains
No 94-004 **Kowalski, Stewart**
IT Insecurity: A Multi-disciplinary Inquiry
No 94-007 **Asker, Lars**
Partial Explanations as a Basis for Learning
No 94-009 **Kjellin, Harald**
A Method for Acquiring and Refining Knowledge in Weak Theory Domains
No 94-011 **Britts, Stefan**
Object Database Design
No 94-014 **Kilander, Fredrik**
Incremental Conceptual Clustering in an On-Line Application
No 95-019 **Song, Wei**
Schema Integration: - Principles, Methods and Applications

No 95-050 **Johansson, Anna-Lena**
Logic Program Synthesis Using Schema Instantiation in an Interactive Environment
No 95-054 **Stensmo, Magnus**
Adaptive Automated Diagnosis
No 96-004 **Wærn, Annika**
Recognising Human Plans: Issues for Plan Recognition in Human - Computer Interaction
No 96-006 **Orsvärn, Klas**
Knowledge Modelling with Libraries of Task Decomposition Methods
No 96-008 **Dalianis, Hercules**
Concise Natural Language Generation from Formal Specifications
No 96-009 **Holm, Peter**
On the Design and Usage of Information Technology and the Structuring of Communication and Work
No 96-018 **Höök, Kristina**
A Glass Box Approach to Adaptive Hypermedia
No 96-021 **Yngström, Louise**
A Systemic-Holistic Approach to Academic Programmes in IT Security
No 97-005 **Wohed, Rolf**
A Language for Enterprise and Information System Modelling
No 97-008 **Gambäck, Björn**
Processing Swedish Sentences: A Unification-Based Grammar and Some Applications
No 97-010 **Kapidzic Cicovic, Nada**
Extended Certificate Management System: Design and Protocols
No 97-011 **Danielson, Mats**
Computational Decision Analysis
No 97-012 **Wijkman, Pierre**
Contributions to Evolutionary Computation
No 97-017 **Zhang, Ying**
Multi-Temporal Database Management with a Visual Query Interface
No 98-001 **Essler, Ulf**
Analyzing Groupware Adoption: A Framework and Three Case Studies in Lotus Notes Deployment
No 98-008 **Koistinen, Jari**
Contributions in Distributed Object Systems Engineering
No 99-009 **Hakkarainen, Sari**
Dynamic Aspects and Semantic Enrichment in Schema Comparison
No 99-015 **Magnusson, Christer**
Hedging Shareholder Value in an IT dependent Business society - the Framework BRITS
No 00-004 **Verhagen, Henricus**
Norm Autonomous Agents

No 00-006 **Wohed, Petia**
Schema Quality, Schema Enrichment, and Reuse in Information Systems Analysis
No 01-001 **Hökenhammar, Peter**
Integrerad Beställningsprocess vid Datasystemutveckling
No 01-008 **von Schéele, Fabian**
Controlling Time and Communication in Service Economy
No 01-015 **Kajko-Mattsson, Mira**
Corrective Maintenance Maturity Model: Problem Management
No 01-019 **Stirna, Janis**
The Influence of Intentional and Situational Factors on Enterprise Modelling Tool Acquisition in Organisations
No 01-020 **Persson, Anne**
Enterprise Modelling in Practice: Situational Factors and their Influence on Adopting a Participative Approach
No 02-003 **Sneiders, Eriks**
Automated Question Answering: Template-Based Approach
No 02-005 **Eineborg, Martin**
Inductive Logic Programming for Part-of-Speech Tagging
No 02-006 **Bider, Ilia**
State-Oriented Business Process Modelling: Principles, Theory and Practice
No 02-007 **Malmberg, Åke**
Notations Supporting Knowledge Acquisition from Multiple Sources
No 02-012 **Männikkö-Barbutiu, Sirkku**
SENIOR CYBORGS- About Appropriation of Personal Computers Among Some Swedish Elderly People
No 02-028 **Brash, Danny**
Reuse in Information Systems Development: A Qualitative Inquiry
No 03-001 **Svensson, Martin**
Designing, Defining and Evaluating Social Navigation
No 03-002 **Espinoza, Fredrik**
Individual Service Provisioning
No 03-004 **Eriksson-Granskog, Agneta**
General Metarules for Interactive Modular Construction of Natural Deduction Proofs
No 03-005 **De Zoysa, T. Nandika Kasun**
A Model of Security Architecture for Multi-Party Transactions
No 03-008 **Tholander, Jakob**
Constructing to Learn, Learning to Construct - Studies on Computational Tools for Learning
No 03-009 **Karlgren, Klas**
Mastering the Use of Gobbledygook - Studies on the Development of Expertise Through Exposure to Experienced Practitioners' Deliberation on Authentic Problems

No 03-014 **Kjellman, Arne**
Constructive Systems Science - The Only Remaining Alternative?
No 03-015 **Rydberg Fåhræus, Eva**
A Triple Helix of Learning Processes - How to cultivate learning, communication and collaboration among distance-education learners
No 03-016 **Zemke, Stefan**
Data Mining for Prediction - Financial Series Case
No 04-002 **Hulth, Anette**
Combining Machine Learning and Natural Language Processing for Automatic Keyword Extraction
No 04-011 **Jayaweera, Prasad M**.
A Unified Framework for e-Commerce Systems Development: *Business Process Patterns Perspective*
No 04-013 **Söderström, Eva**
B2B Standards Implementation: Issues and Solutions
No 04-014 **Backlund, Per**
Development Process Knowledge Transfer through Method Adaptation, Implementation, and Use
No 05-003 **Davies, Guy**
Mapping and Integration of Schema Representations of Component Specifications
No 05-004 **Jansson, Eva**
Working Together when Being Apart – An Analysis of Distributed Collaborative Work through ICT from an Organizational and Psychosocial Perspective
No 05-007 **Cöster, Rickard**
Algorithms and Representations for Personalised Information Access
No 05-009 **Ciobanu Morogan, Matei**
Security System for Ad-hoc Wireless Networks based on Generic Secure Objects
No 05-010 **Björck, Fredrik**
Discovering Information Security Management
No 05-012 **Brouwers, Lisa**
Microsimulation Models for Disaster Policy Making
No 05-014 **Näckros, Kjell**
Visualising Security through Computer Games
Investigating Game-Based Instruction in ICT Security: an Experimental approach
No 05-015 **Bylund, Markus**
A Design Rationale for Pervasive Computing
No 05-016 **Strand, Mattias**
External Data Incorporation into Data Warehouses
No 05-020 **Casmir, Respickius**
A Dynamic and Adaptive Information Security Awareness (DAISA) approach

No 05-021 **Svensson, Harald**
Developing Support for Agile and Plan-Driven Methods
No 05-022 **Rudström, Åsa**
Co-Construction of Hybrid Spaces
No 06-005 **Lindgren, Tony**
Methods of Solving Conflicts among Induced Rules
No 06-009 **Wrigstad, Tobias**
Owner-Based Alias Management
No 06-011 **Skoglund, Mats**
Curbing Dependencies in Software Evolution
No 06-012 **Zdravkovic, Jelena**
Process Integration for the Extended Enterprise
No 06-013 **Olsson Neve, Theresia**
Capturing and Analysing Emotions to Support Organisational Learning: The Affect Based Learning Matrix
No 06-016 **Chaula, Job Asheri**
A Socio-Technical Analysis of Information Systems Security Assurance A Case Study for Effective Assurance
No 06-017 **Tarimo, Charles N.**
ICT Security Readiness Checklist for Developing Countries:
A Social-Technical Approac**h**
No 06-020 **Kifle Gelan, Mengistu**
A Theoretical Model for Telemedicine
- Social and Value Outcomes in Sub-Saharan Africa
No 07-001 **Fernaeus, Ylva**
Let’s Make a Digital Patchwork
Designing for Children’s Creative Play with Programming Materials
No 07-003 **Bakari, Jabiri Kuwe**
A Holistic Approach for Managing ICT Security in Non-Commercial Organisations
A Case Study in a Developing Country
No 07-004 **Sundholm, Hillevi**
Spaces within Spaces: The Construction of a Collaborative Reality
No 07-005 **Hansson, Karin**
A Framework for Evaluation of Flood Management Strategies
No 07-007 **Aidemark, Jan**
Strategic Planning of Knowledge Management Systems
- A Problem Exploration Approach
No 07-009 **Jonsson, Martin**
Sensing and Making Sense
Designing Middleware for Context Aware Computing
No 07-013 **Kabilan, Vandana**
Ontology for Information Systems (O4IS) Design Methodology:
Conceptualizing, Designing and Representing Domain Ontologies

No 07-014 **Mattsson, Johan**
Pointing, Placing, Touching - Physical Manipulation and Coordination Techniques for Interactive Meeting Spaces
No 07-015 **Kessler, Anna-Maria**
A Systemic Approach Framework for Operational Risk- SAFOR
No 08-001 **Laaksolahti, Jarmo**
Plot, Spectacle and Experience: Contributions to the design and evaluation of Interactive Storytelling
No 08-002 **Van Nguyen Hong**
Mobile Agent Approach to Congestion Control in Heterogeneous Networks
No 08-003 **Rose-Mharie Åhlfeldt**
Information Security in Distributed Healthcare
- Exploring the Needs for Achieving Patient Safety and Patient Privacy
No 08-004 **Sara Ljungblad**
Beyond users:
Grounding technology in experience
No 08-005 **Eva Sjöqvist**
Electronic Mail and its Possible Negative Aspects in Organizational Contexts
No 08-006 **Thomas Sandholm**
Statistical Methods for Computational Markets
– Proportional Share Market Prediction and Admission Control
No 08-007 **Lena Aggestam**
IT-supported Knowledge Repositories:
Increasing their Usefulness by Supporting Knowledge Capture
No 08-008 **Jaana Nyfjord**
Towards Integrating Agile Development and Risk Management
No 08-009 **Åsa Smedberg**
Online Communities and Learning for Health
- The Use of Online Health Communities and Online Expertise for People with Established Bad Habits
No 08-010 **Martin Henkel**
Service-based Processes
- Design for Business and Technology
No 08-012 **Jan Odelstad**
Many-Sorted Implicative Conceptual Systems