DETECTING SOCIAL ENGINEERING

by Michael D Hoeschele & Marcus K Rogers

Center for Education and Research in
Information Assurance and Security,
Purdue University, West Lafayette, IN 47907-2086

Running head: DETECTING SOCIAL ENGINEERING

Detecting Social Engineering

Michael Hoeschele & Marc Rogers

CERIAS

Purdue University

{mhoesche, mkr}@cerias.purdue.edu

Recitation Building, 656 Oval Drive

West Lafayette, IN 47907

December 07, 2004

Abstract

This paper proposes a theoretical solution to the problem of Social Engineering (SE) attacks perpetrated over the phone lines. As a byproduct real time attack signatures are generated, which can be used in a cyber forensic analysis of such attacks. Current methods of SE attack detection and prevention rely on policy and personnel training, which fails because the root of the problem, people, are still involved. The proposed solution relies on computer systems to analyze phone conversations in real time and determine if the caller is deceiving the receiver. This Social Engineering Defense Architecture (SEDA) is completely theoretical as the technologies employed are only in the proof of concept phase, but they are all proven to be tractable problems.

Keywords: social engineering, defense architecture, forensics

Introduction

Social Engineering (SE) presents an interesting problem to cyber forensics that has not been researched enough; this is evident by the lack of formal published research on the topic. The research that does focus specifically on SE, as it is used to commit computer related crimes, primarily define SE and the techniques that can be used to breach security. No research has been preformed on forensic analysis of these attacks. The only type of solution or ideas presented to detect or trace attacks is the education of personnel on SE techniques. An example of this is found in the Information Security Management Handbook. The chapter entitled Social engineering: The forgotten risk by Rogers and Berti (2002) discusses how to prevent SE attacks. Most of their suggestions are security policies and training of personnel to be aware of what an attack may look like, both of which are still dependant upon the human element. Dolan (2004) makes similar suggestions for detecting SE attacks. He claims, "Successful social engineering attacks rely on the employees of an organization. To contain such an attack, employees must be well trained and familiar with common SE techniques" (p. 6).

In 1997 the Client Server Computing magazine published the article Liar, Liar by Julie Bort. In this article Bort states, "Of the 384 respondents who confessed to being attacked over the last year, social engineering was to blame in 15 percent of the cases--the second largest cause" (p. 40). While this is somewhat outdated, it clearly shows that SE is neither a new or small problem. This was the only statistic available that directly measured the prevalence of SE attacks. Other more current publications like the FBI/CSI annual Computer Crime and Security reports, from 2002, 2003, and 2004, showed some vague statistics on telecom fraud that could include SE. More details on these reports are provided in the Financial Findings section (Gordon, et. al., 2004; Power, 2002; Richerdson, 2003).

Bort states in her article, “There is no hardware or software that can defend your information systems against a human being telling a convincing lie (Bort, p. 40).” This was true in 1997 but Raskin et. al. (2004) are currently researching the problem of lie detection in natural language and are making progress. They have proven it to be a tractable problem and created a working model.

Statement of Problem

Social Engineering (SE) attacks are a current threat to organizations large and small, for profit and non-profit alike. However, the focus of most cyber forensics is on attacks perpetrated through computer systems alone. This results in a very limited idea of how to perform a forensic analysis of a SE attack. There are no documented signatures for such attacks or systems to log activity associated with SE. This means that even if it can be determined that a SE attack occurred, it is very unlikely that the crime can be traced back to the perpetrator in such a way that they can be prosecuted in a court of law. These attack signatures are somewhat elusive because of the nature of SE and the avenues of attack used, most commonly the telephone system. There are other avenues that can be used such as face-to-face conversations, but they are much more rare as the risk incurred to the attacker greatly increases when entering the target premises. These are the factors that lead this paper to focus primarily on the telephone system to derive an architecture that generates attack signatures. The attack signatures will be generated in real time allowing the logging architecture to also serve as a defense architecture much in the same way a Network Intrusion Detection System works.

Social Engineering

Aaron Dolan defines SE as "using relationships with people to attain a goal (p. 2)." For the purposes of this paper, the term SE will be discussed specifically as the malicious intent of cyber attackers attempting to illegally compromise an organization's assets (Dolan, 2004). It should be noted that the types of organizations under SE attacks are not limited to faceless multinational corporations. Educational institutes, banks, and even the corner video store are at risk.

Methods

Social engineers, akin to computer hackers, take advantage of weaknesses in the system in place to keep enemies out, and assist insiders. As Dolan (2004) stated:

> "Social engineers use tactics to leverage trust, helpfulness, easily attainable information, knowledge of internal processes, authority, technology and any combination there of. They often use several small attacks to put them in the position to reach their final goal. Social engineering is all about taking advantage of others to gather information and infiltrate an attack. The information gained in a phone book may lead to a phone call. The information gained in the phone call may lead to another phone call. A social engineer builds on each tidbit of information he or she gains to eventually stage a final, deadly attack. A successful social engineering attempt could result in great financial loss for the target company. A motivated attacker will be willing to gain information in any way possible" (p. 3).

One very important reason that SE is so easily achieved is that people in general have the desire to help others, and gain satisfaction from it (Dolan, 2004). One of the most fundamental skills of a social engineer is the ability to establish trust. This can be accomplished by masquerading as someone an employee should trust. Typically most of the information necessary to steal someone's identity for a SE attack is publicly available. Reverse phone look-up directories are available on the web free of charge, (e.g. http://www.reversephonedirectory.com/). Once a phone number or address is obtained the other follows effortlessly. Many organization's web pages hold vast quantities of information such as organizational charts, which provides social engineers with a target from which to steal an identity or claim association. Another common technique is for a social engineer to ask to be transferred to an employee from one of the main telephone operators. The receiver of the transfer call does not have a phone number that is posted publicly, so anyone that calls them is already considered an insider to the receiver of the call as they were able to find the number, assumedly from the company phonebook. This can prove to be a strong enough credential to allow more internal numbers to leak out, furthering the social engineer's cache of information.

A few more examples of SE attacks that Wendy Arthurs stated in her paper are as follows (Arthurs, 2001):

> "IT Support – Somebody, claiming to be from the company's IT support group, phones a user and explains that he is fault finding on the network. He has limited the fault to within the users department but he needs a user ID and password from that department to finish tracing the problem. Unless the user has been properly educated in security practices, he will very likely give the "trouble-shooter" his information.

> Manager – The social engineer, using a perceived position of authority, phones the help desk demanding to know why he can't log on with his password. He then intimidates the help desk into giving him a new password by telling them he only has a limited time to retrieve some information for a report to the company vice president. He may also threaten to report the help desk employee to his supervisor.
>
> Trusted Third Party – The social engineer phones the help desk, claiming to be Susan Sly, the vice-president's executive assistant, stating the vice-president has authorized her to collect the information. If the help desk employee balks, she threatens job loss or a report to the employee's supervisor" (p. 2).

It can be seen from these examples that a large portion of the attacks are committed over the phone and rely on the fact that the receiver of the call has to take the caller's word that they are who they say they are. There is typically no form of authentication other than verbal quizzes regarding information only an employee should know.

Motivation

The motivation for SE attacks varies between a sense of curiosity and challenge seeking, to directed attacks with the intent to compromise an organization's assets. The Hackers Manifesto gives us a brief insight into why hackers desire to break into secure areas. The claim is that of curiosity and a quest for knowledge. While harm may not be the intent, it is clear that great harm can be caused by such activities. Some social engineers also seek a challenge, something to test their skills on (Mentor, 1986).

The more troublesome type of attack is that of a social engineer with a purpose. The motivation could come from a recently fired employee seeking vengeance or a seasoned social engineer locating information for money. This type of social engineer has a better chance of succeeding and causing considerable damage in the process as they have much greater motivation to succeed and potentially more resources at their disposal (Dolan, 2004).

Targets

The types of targets for any SE attack range from product designs to personal employee information. However, it is important to note that these are only the end targets of the attack, and many small pieces of information must be obtained before the final target can be reached. The types of targets used to reach a final goal can include company policy or protocol, company phone books, organization trees, or server names. Most of this type of data is either easily obtainable with a call or is publicly available on the company website. The real challenge is protecting such seemingly trivial information it in such a way as not to interfere too greatly with day-to-day workflow (Rogers & Berti, 2002).

Current solution

The current solution to SE attacks is employee training to resist such attacks coupled with a thorough security policy. This solution is fundamentally flawed as it relies on humans to patch the security holes. The flaw is that human trust is the vulnerability exploited, which is deeply imbedded in US culture and is difficult to overcome (Rogers & Berti, 2002). A good social engineer will convince you that he or she deserves the requested data, and you are hurting the company by not complying.

Training may stop most menial or juvenile attackers, but a seasoned social engineer will never be thought of as a risk when employees talk to them, so the training will never be triggered. This can be seen with the examples of SE attacks in the Social Engineering section above (Rogers & Berti, 2002).

Policy

Employing security policy is a good idea in the sense that after someone has broken the policy it is easy to point to the policy and show how they violated it. However, it is not realistic to use policy alone to prevent break-ins. Security policy that classifies data into different levels of sensitivity is the most beneficial to prevent SE attacks. This provides a greater level of security for data that warrants it by forcing a higher level of credentials to gain access. The end result is more work for the social engineer, which will deter most inexperienced and casual social engineers. However, an experienced, or more importantly, persistent social engineer will keep working until they have all the necessary credentials to obtain approval for access. Allen (2003) states the following as policy based counter measures to SE attacks in his paper:

> "Security policy - A sound security policy will ensure a clear direction on what is expected of staff within an organization. For example, support teams should only offer assistance for a defined range of activities.
>
> Limit Data leakage - Reducing the amount of specific data available will ensure that the attack is not an effortless exercise. For example websites, public databases, Internet registries, and other publicly accessible data sources should only list generic information, such as main organization phone number and job

titles instead of employee name(s), for example 'site administrator' instead of 'Joe Bloggs'" (p. 5).

Training

Currently the most effective deterrent to SE is training employees to resist such attacks. Training ranges from a multi day seminar once a year to every day reminders in the form of posters and mailings. The idea is if everyone in the company is aware of how a social engineer executes an attack and gains trust, they will be able to detect an attack in progress and stop it. Employees will also be told not to release certain types of data over the phone, like password or employee ID number. One of the problems with this is a good social engineer never shows any signs of an attacker, they very quickly assimilate into the company and at worst appear as a employee trying to go the extra mile or in need of help. It is not realistic to expect employees to be the primary defense against SE attacks, but it is logical to make them aware of it. At worst it will help in the acceptance of any systems put in place to defend against SE attacks (Rogers & Berti, 2002).

Financial Findings

The closest thing to reports of the prevalence and trends of SE attacks was in the CSI/FBI Computer Crime And Security Survey (Gordon et. al., 2004; Power, 2002; Richardson, 2003). The survey has an attack grouping titled Telecom Fraud. However, based on the data it seems that this is more likely this is in reference to a company's phone switch being hacked. Regardless, the report shows that percentage of respondents detecting telecom fraud decreased from 17% in 1999 to 11% in 2000, then to 10% in 2001, and to 9% in 2002. The percentage did rise to 10% in 2003, but the increase seems negligible. The finical loss due to telecom fraud is listed at $773,000 in 1999, $4,028,000 in 2000, $9,041,000 in 2001, $346,000 in 2002, and $701,500 in 2003. While it is difficult to determine if SE attacks are encapsulated within Telecom Fraud the data points to the problem becoming less important. The reason it does not seem plausible that this data reflects SE attacks is that no current method exists to determine if a social engineer has attacked an organization. No logs are generated when a call is placed, so there is no way to check after an attack has occurred. It is the job of a social engineer to complete an objective without sounding any alarms so they can return later for more information if necessary. This is different than a computer attacker whom might leave a back door for later use, in that case there is something to find or detect, like unsolicited outbound traffic. A social engineer will keep all the data he or she needs to masquerade as an employee again later, this cannot easily be detected.

Success of Current Solution

Unfortunately it is impossible with the available data to determine what effect the current methods of attack prevention and detection have had on the success of SE attacks. It could be reasoned that the general lack of data shows the inadequacy of current methods of detection. If there were no Network Intrusion Detection Systems how would one measure the amount of attacks? This inability produce data logically points to the current solution's overwhelming inadequacy.

Case Study

There are no case studies of SE attacks subverting the current solution. It is very unlikely that this is because it has never happened, but rather that those that are beaten don’t report it. Rogers and Berti state as a possibility that SE “attacks the intelligence of the victim and, as such, there is a reluctance to admit that it has occurred” (2002, p. 52).

Proposed Solution

To mitigate the risk of SE attacks perpetrated over the telephone consider the following. A Social Engineering Defense Architecture (SEDA) that both detects attacks as they occur over the phone and generates logs so attacks can be traced back to the attacker in a forensic analysis of the attack. The system focuses on the telephone medium as Dolan (2004) and Gragg (2003) showed that most attacks are carried out over the phone. This method will work because it detects attacks based on intent and deception instead of attack target. Detecting a SE attack based on target is difficult because social engineers typically first pursue targets with seemingly very little importance as explained in the Social Engineering section. However, this trivial information is then used to obtain more sensitive and well-guarded information. This system will prevent both the early and later stages of a SE attack by detecting lying and deception. Any company employee that is being deceived by a caller should be made aware of it regardless of whether the caller is a social engineer or not.

Intrusion Detection

The primary purpose of the SEDA is to make call recipients within the company aware of callers that are attempting to deceive them or obtain information they don’t have permission to. The muscle of the system is a text-independent voice signature authentication system. Markowitz defines text-independent verification as “[accepting] any spoken input, making it possible to design unobtrusive, even invisible, verification applications that examine the ongoing speech of an individual” (Markowitz, 2000, p. 69).

She then states, "The ability of text-independent technology to operate unobtrusively and in the background makes it attractive for customer related applications, because customers need not pause for a security check before moving on to their primary business objective" (Markowitz, 2000, p. 69). The result is a system of authentication that hinders workflow as little as possible. The voice signatures will be linked to a database of personal information possibly including name, corporate association, job title, and all phone numbers called from. The type of information gathered would depend on the needs of the organization employing the SEDA. The main advantage of this tool is that social engineers often use identity theft to build trust as can be seen in the Social Engineering section above. This system would prevent a social engineer from claiming to be an employee, even if they have all the right information pass as one in conversation. It would also make life much more difficult for a social engineer that keeps changing their name. The first time the attacker calls the name they use will be associated with their voice signature, so if they want to call under another name they would have to modify their voice. While this is a problem, most attackers will be deterred by the system, and those that do modify their voice will still have to deal with the attack detection systems described below. This and further problems are discussed in greater detail below in the Caveats and Problems section.

The next important part of the SEDA is a voice-to-text engine. This converts the entire voice conversation into text in real time. There are examples of some systems, that while not accurate or reliable enough, were able to convert voice into text (Karat et. al., 1999; Lia & Vergo, 1997). While this is not directly related to preventing SE attacks it allows other powerful attack detection algorithms to analyze the conversation. The important note about this system is that it will need to be very fast and accurate. If the text cannot be sent for analysis quickly enough, the attacker could obtain the information necessary before the recipient of the call can be notified that an attack is in progress. Also, the voice-to-text system must be robust enough to deal with less than perfect telephone connections. It should be noted that to assist forensics analysis all conversations originating outside of the company phone switch are recorded and the text of the conversation is linked with the caller and receiver voice signatures. Because of the need to record conversations for security purposes each caller would have to opt-in by selecting a certain number when they first call in. This removes the expectation of privacy associated with telephone conversations, and prevents the organization using the SEDA from breaking the Wire Tap Act (US Department of Justice, 1986). If a caller does not want to opt-in to a secure call they will be transferred to an operator highly trained in resisting SE attacks, who will explain the purpose of the opt-in.

The “brains” behind the SEDA are the textual conversation analysis tools. Based on the current computational demands of these tools they will probably have to be implemented on different servers to analyze the conversation in parallel. The first content analysis tool is Raskin et al., Natural Language Processing program to determine if a person is lying. While this program is currently not ready to be implemented into the SEDA, his research and progress shows that the problem of parsing a conversation and determining if someone is lying is tractable; it is a matter of time before it becomes a usable application (Raskin et. al., 2004).

Another content analysis tool that would be much more simple and narrow in its purpose could search the conversation for certain strings used in common attacks. This is synonymous to virus definitions. If the caller says, “Please read me your username and password,” it is clear that either the caller has bad intent, or an employee is violating security policy, neither is acceptable. These rules would have to be customized depending on desired security and other company policies. Figure 1 is a decision tree that maps out how the SEDA is structured. It should be noted that no action, other than notifying the receiver of the call, is taken when an attack is detected. Further research is needed to determine the best course of action.

Attack Signature Generation

As with computer attackers, some attacks get through no matter what is done to prevent them. In the case of a very skilled social engineer breaking into an organization, the SEDA provides call logs to perform forensics analysis of the attack. As noted above, every conversation originating outside the company phone switch is recorded in text format with the voice signatures linked for caller identification. Text format will be used to provide both smaller amounts of storage needed and to allow forensic investigations to scan for clues without having to convert voice back into text. These logs of conversations will allow a forensic analyst to trace a cyber crime back from the final target to the attacker. The call logs could also be used in conjunction with other forensics methods to track an attacker.

Caveats and Problems

One of the major problems with the proposed solution is its inability to handle voice modulation. If an attacker were to pipe his or her voice through a voice modulation device during every call in an attack the ability to link the calls together in a forensic analysis would be greatly decreased. However, voice modulation would have no effect

on the SEDA's ability to detect deception based on conversation content. A resourceful social engineer would not be able to bypass all of the SEDA's levels of protection. Problems could also arise if the system were unable to handle poor telephone connections, for example, from an older cell phone. If this could not be handled it would serve as another form of voice modulation. A related problem is replay attacks; recording someone's voice and playing it into the phone to authorize, then continue the conversation. Text-independent voice signatures can be generated over the entire call to prevent this, and "many commercial speaker-verification systems look for telltale auditory signals, distortions, exact matches, and other indications that a recording has been used" (Markowitz, 2000, p. 68).
some voice signature systems already detect such attacks (Markowitz, 2000, p. 69).

Another problem lies in how to introduce such a product to the business world. Since there is no current method to meter or gauge the damage of SE attacks, it is impossible to show why something like this is necessary or a finically sound investment. One solution is to conduct a study to generate such data based on what security personnel think is the amount of damage cause by SE attacks.

Future Research

There are several areas within the proposed solution that are candidates for further research. First is the handling of internal calls. The proposed solution simply treats all calls equally regardless of the location of the caller. For certain businesses this could be a great problem because of the quantity of internal calls. If there were a way to streamline internal calls then the amount of processing power necessary could decrease greatly.

Another area that needs further research is other methods of detecting SE attacks. There are currently products in the market that tout lie detection based on voice analysis such as Nemesysco's Layered Voice Analysis (2004). This could be integrated to provide a further level of attack detection. Another attack detection method that could be added is to listen for account names adjacent to suspicious comments, then flag the account.

As mentioned earlier, research is needed to determine how to handle attacks when they are detected. While this may vary depending on the security goals of the organization, certain attack response guidelines need to be formulated to aid in the creation of countermeasures.

The last area of future research is SE forensics. With the new logs being generated by the SEDA, forensics tools need to be created to parse log files and find clues. In addition, policy on how to conduct a SE cyber forensic investigation using the logs needs to be created. Essentially all the research that has been done and is currently in progress on cyber forensics needs to be done in SE forensics.

Conclusion

The main strength of the SEDA is it will take the human element out of determining a person's identity over the phone. Callers will be identified as employees or outsiders; this alone is crucial to preventing SE attacks. The ability to detect deception also means a social engineer will not be able to appeal to someone's emotions or try to bully him or her into an action. Another great strength is the log files that are generated that allow for forensic analysis of SE attacks. This opens up an entire new field in cyber forensics. The weakness is that voice modulation makes it possible for one person to call many times under different names and not be tracked. While this is a clear problem that needs to be addressed with voice modulation detection, it does not undermine the entire SEDA. Unfortunately, the proposed solution cannot be implemented currently, so it cannot be determined how affective it will be.

Despite the limitations, the SEDA addresses two major problems that have gone unanswered for several years, how to detect and how to perform a forensic analysis of a SE attack.

References

Allen, M. (2/18/2003) *The Use of 'Social Engineering' as a means of Violating Computer Systems*. Retrieved October 10, 2004, from http://www.sans.org/rr/catindex.php?cat_id=51

Arthurs, W. (8/2/2001) *A Proactive Defense to Social Engineering*. Retrieved October 10, 2004, from http://www.sans.org/rr/catindex.php?cat_id=51

Bort, J. *Liar, Liar*. Client Server Computing, Vol. 4 Issue 5, 1997

Dolan, A. (4/8/2004) *Social Engineering*. Retrieved October 10, 2004, from http://www.sans.org/rr/catindex.php?cat_id=51

Gragg, D. (3/13/2003) *A Multi-Level Defense Against Social Engineering*. Retrieved October 10, 2004, from http://www.sans.org/rr/catindex.php?cat_id=51

Gordon, L., Loeb, M. P., Lucyshyn, W., & Richardson, R. (2004) *CSI/FBI Computer Crime and Security Survey*. Retrieved September 10, 2004, from http://www.gocsi.com/

Karat, C-M., & Halverson, C., & Horn, D., & Karat, J. (1999). *Patterns of Entry and Correction in Large Vocabulary Continuous Speech Recognition Systems*. Retrieved October 10, 2004, from http://portal.acm.org/citation.cfm?id=303160&coll=ACM&dl=ACM&CFID=29958187&CFTOKEN=58331754

Lai, J., & Vergo, J. (1997). MedSpeak: *Report Creation with Continuous Speech Recognition.* Retrieved October 10, 2004, from http://portal.acm.org/citation.cfm?id=258829&coll=ACM&dl=ACM&CFID=29958187&CFTOKEN=58331754

Markowitz, J. A., (September 2000). *Voice Biometrics*. Communications of the ACM, Vol. 43, No. 9. Retrieved October 10, 2004, from http://portal.acm.org/citation.cfm?id=348995&coll=ACM&dl=ACM&CFID=29958187&CFTOKEN=58331754

Mentor, The. (1986) The Hackers Manifesto. Retrieved October 10, 2004 from http://www.geocities.com/SiliconValley/Heights/1926/mentor.html

Nemesysco (2004). *The LVA (Layer Voice Analysis) Technology*. Retrieved December 7, 2004 from http://www.nemesysco.com/technology-lvavoiceanalysis.html

Power, R. (2002) *CSI/FBI Computer Crime and Security Survey*. Retrieved September 10, 2004, from http://www.gocsi.com/

Raskin, V. H., Christian F. & Triezenberg, Katrina E. (2004). Semantic forensics: An application of ontological semantics to information assurance. 42nd Annual Meeting of the Association for Computational Linguistics, Barcelona, Spain, Association for Computational Linguistics.

Richardson, R. (2003) *CSI/FBI Computer crime and security survey*. Retrieved September 10, 2004, from http://www.gocsi.com/

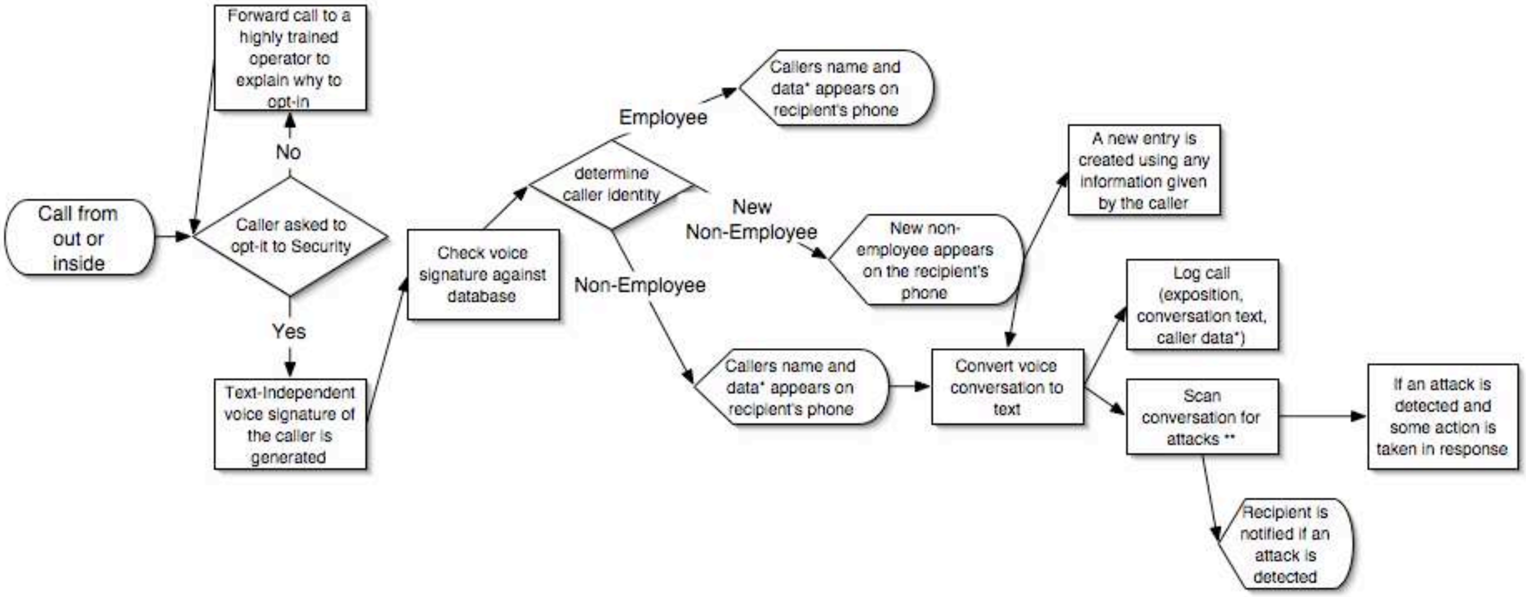
Rogers, M. & Berti, J. (2002) Social engineering: The forgotten risk, In H. F. Tipton & M. Krause (Eds.), *Information security management handbook,* Vol. 3, 4th Edition (pp. 51-63). New York: CRC Press LLC

US Department Of Justice. *18 U.S.C. 2511. Interception and disclosure of wire, oral, or electronic communications prohibited*. Retrieved October 26, 2004 from http://www.cybercrime.gov/usc2511.htm

Figure Caption

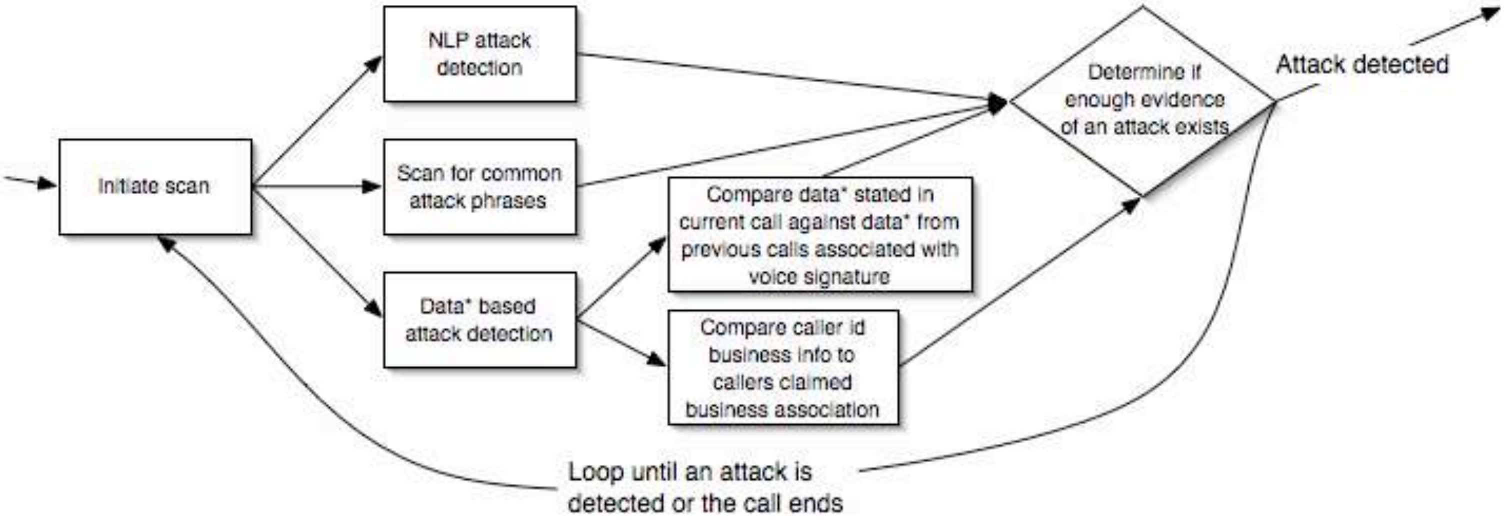
Figure 1. Social Engineering Defense Architecture decision tree

Figure 2. Expansion of attack detection processes


Call from out or inside
Caller asked to opt-it to Security
No
Forward call to a highly trained operator to explain why to opt-in
Yes
Text-Independent voice signature of the caller is generated
Check voice signature against database
determine caller identity
Employee
Callers name and data* appears on recipient's phone
New Non-Employee
New non-employee appears on the recipient's phone
A new entry is created using any information given by the caller
Non-Employee
Callers name and data* appears on recipient's phone
Convert voice conversation to text
Log call (exposition, conversation text, caller data*)
Scan conversation for attacks **
If an attack is detected and some action is taken in response
Recipient is notified if an attack is detected

*   This is all information associated with a caller

** See Figure 2


Initiate scan
NLP attack detection
Scan for common attack phrases
Data* based attack detection
Compare data* stated in current call against data* from previous calls associated with voice signature
Compare caller id business info to callers claimed business association
Determine if enough evidence of an attack exists
Attack detected
Loop until an attack is detected or the call ends

* This is all information associated with a caller