## Chapter XVIII

# Steganography and Steganalysis

Merrill Warkentin, Mississippi State University, USA

Mark B. Schmidt, St. Cloud State University, USA

Ernst Bekkering, Northeastern State University, USA

## Abstract

*In the digital environment, steganography has increasingly received attention over the last decade. Steganography, which literally means "covered writing," includes any process that conceals data or information within other data or conceals the fact that a message is being sent. Though the focus on use of steganography for criminal and terrorist purposes detracts from the potential use for legitimate purposes, the focus in this chapter is on its role as a security threat. The history of stenography as a tool for covert purposes is addressed. Recent technical innovations in computerized steganography are presented, and selected widely available steganography tools are presented. Finally, a brief discussion of the role of steganalysis is presented.*

## Introduction

In the digital environment, steganography has received increasing attention over the last decade. The steganography process conceals the fact that a message is being sent, thereby preventing an observer from knowing that anything unusual is taking place. Neil F. Johnson of the Center for Secure Information Systems at George Mason University defines steganography as "the art of concealing the existence of information within

seemingly innocuous carriers" (Johnson, 2003, p. 2). Much of this attention has focused on the use of steganography for illegitimate purposes by terrorists and criminals, culminating in news stories about Al Qaeda's use of the technique in its communications. The extent of actual use by terrorists remains to be seen and, so far, has never been (publicly) proven. Yet, it has been suggested by government officials in the US and elsewhere that Al Qaeda and other organizations are hiding maps and photographs of terrorist targets and are also posting instructions for terrorist activities on sports chat rooms, pornographic bulletin boards, and other Web sites.

The preoccupation with stenography as a tool for covert purposes can be explained by reviewing its history. Though the term itself is based on the Greek word for "covered writing," the term was first used in the 14th century by the German mathematician Johannes Trithemius (1606) as the title for his book *Steganographia*. On the surface, the book presents a system of angel magic, but it actually describes a highly sophisticated system of cryptography. The actual hiding of information is much older. In ancient Greece, messages might be tattooed on slaves' shaved heads and then their hair would be allowed to grow back before they were sent out as messengers. A more benign form of information hiding was inscribing messages on the wooden base of wax tablets, rather than on the surface of the wax itself (Jupitermedia Corporation, 2003). More recent forms of hiding messages were used in World War II when spies and resistance fighters used milk, fruit juice, or even urine to write invisible coded messages. Heating the source document would reveal the writing, which had turned invisible to the naked eye after the unusual form of ink had dried up. Thus, the history of steganography has long been associated with an air of secrecy, far removed from peaceful and productive purposes.

## Steganography Today

More technical forms of steganography have been in existence for several years. International workshops on information hiding and steganography have been held regularly since 1996 (Moulin & O'Sullivan, 2003). However, the majority of the development and use of computerized steganography has occurred since 2000 (Cole, 2003). Modern technology and connectivity have put steganographic capabilities within the reach of the average person with a computer and an Internet connection (Bartlett, 2002). Steganography does not necessarily encrypt a message, as is the case with cryptography. Instead, the goal is to conceal the fact that a message even exists in the first place (Anderson & Petitcolas, 1998). In today's fast-paced, high-tech society, people who want to send hidden messages have very efficient methods of getting a message to its destination with the use of computerized tools that encode a message in a graphic, sound, or other type of file.

## New Practices for an Ancient Technique

With the onset of the digital age, many new and innovative mechanisms became available for information hiding. Steganographic techniques and software focused on hiding





information and messages in audiovisual files such as graphics files, sound files, and video files. Insignificant and unused parts of these files were replaced with the digital data for the hidden information. The information itself could be protected even further by use of cryptography, where the information was converted into a form incomprehensible without knowledge of the specific cryptographic technique and key. This highlights an important difference between steganography and cryptography. The ultimate goal of cryptography is hiding and protecting the content of information, whereas steganography hides the presence of information itself. Another difference is the mode of transmission. Cryptographic messages can be transported by themselves. In steganography, to hide information, the secret content has to be hidden in a cover message. Whereas physical covers were used to hide information in the past, both the cover and hidden content can now be in digital form. Audiovisual files are ideal covers for several reasons. First, these types of files tend to be quite large in comparison to other file types, providing more opportunity for hiding information successfully. By keeping the ratio of hidden digital data to cover data low, the probability of discovery decreases. Furthermore, audiovisual files require special software to detect the presence of hidden information. ASCII-based data can be detected with simple string comparisons, whereas detection of hidden data in pixels and waves depends on detection of statistical anomalies. In other words, an unencrypted message could be detected in a text-based message by a relatively unsophisticated search engine or spybot, requiring less processing power than for detection in the audiovisual covers.

From a technical perspective, data can easily be hidden in an image file. One such technique to hide data is called *least significant bit* (LSB) insertion (Kessler, 2001). The LSB approach allows the last bit in a byte to be altered. While one might think that this would significantly alter the colors in an image file, it does not. In fact, the change is indiscernible to the human eye. As an example, consider three adjacent pixels (nine bytes) with the following RGB encoding:

10010101 00001101 11001001
10010110 00001111 11001010
10011111 00010000 11001011

An LSB algorithm could be used to hide the following nine bits 101101101. The last, or least significant, bit in each byte may be adjusted. The underlined bits indicate a change (notice that it is not necessary to change some of the bits).

10010101 00001100 11001001
10010111 00001110 11001011
10011111 00010000 11001011

The above example demonstrates that to hide nine bits of information, the algorithm only needed to change four of the nine least significant bits in these nine bytes. Because changing the last bit causes an extremely small change in the color of a pixel (speck of

color), the change in the graphic is imperceptible to the human eye. For a more detailed description of this process, please visit http://www.garykessler.net/library/steganography.html.

A special case of steganography can be the use of watermarking, which can have beneficial applications as well. Some watermarks are prominently displayed in an attempt to discourage or prevent unauthorized copying or use, while other watermarks are hidden intentionally. When the message, identification, or graphic is an attribute of the file itself and hidden from regular users, the technique can be termed digital watermarking. In pure steganography, the content of the information added to the file has its own significance. Furthermore, steganographic content is intended to be easily separated from the cover file with the proper software tools by the intended audience. Hidden watermarks can be considered to fall within the bounds of steganography. In our discussion, we will use this expanded definition use of the term steganography, and include the use of hidden watermarks. Good watermarks should be impossible to be separated from the file in which they have been inserted. Finally, steganographic message content can lose its significance when the information becomes outdated or stale. Ideally speaking, the protection afforded by hidden watermarking would last indefinitely.

## Relationship to Cryptography

Cryptography is concerned with creating electronic artifacts (typically data files) that are encoded and cannot be interpreted by an intercepting party. As such, an encrypted message often appears to be random or unintelligible. However, the goal of sending steganographically hidden messages is to appear normal — the artifact might appear to be "normal." However, steganography is often used in conjunction with cryptography to create an especially tough challenge to those responsible for enterprise security. The two technologies are completely independent, and the use of one in no way dictates the use of the other. But we ask the reader to remember that wherever steganography is used, the security of the content will be enhanced by the use of cryptography.

## Steganography Tools

Steganography tools are readily available and their simplicity has made it relatively easy for terrorists and other criminals to hide data in files (Schmidt, Bekkering, & Warkentin, 2004). There are several steganography tools that are publicly available, many of which are available over the Web at no cost. The interested reader is directed to http://www.jjtc.com/Steganography/toolmatrix.htm. This site, maintained by Neil F. Johnson, currently lists and describes more than 140 examples of steganographic software publicly available.

An easy-to-use but effective steganography tool is SpamMimic. SpamMimic can be used by anyone with access to the Web without even downloading any software. To disguise a message, one can visit http://spammimic.com/ and type a message. The website will then create a message that looks like spam, but actually contains the covert message. The primary advantage for the communicating parties is that spam is typically ignored by





*Table 1. Examples of steganography tools*

| Tool | File types | Cost | Web address |
|---|---|---|---|
| Camouflage | Several | Free | http://www.downseek.com/download/5746.asp |
| Invisible Secrets v4.0 | JPEG, PNG, BMP, HTML, and WAV | $39.95 | http://www.stegoarchive.com/ |
| SecurEngine 2.0 | BMP, JPG, and TXT | Free | http://www.freewareseek.com |
| Camera/Shy | GIF, and Web pages | Free | http://sourceforge.net/projects/camerashy/ |
| Stegdetect (XSteg) | Detects the presence of steganography in JPEG | Free | http://packages.debian.org/cgi-bin/download.pl |
| MP3Stego | MP3 | Free | http://www.petitcolas.net/fabien/software/index.html |

many authorities and their systems such as Echelon and Carnivore (Clark, 2001). The recipient of the "spam" can then visit spammimic.com to decode the message. Since the message is unique and no digital signatures are generated, signature-based spam filters will not block these messages beyond the level of the normal false-positive messages blocked by filters. On inspection, the encoded SpamMimic messages also appear very innocuous and lack "red flag" words such as Viagra, porn, prescriptions, cheap, and so forth. This makes classification of encoded messages based on keywords equally unlikely.

Another user-friendly steganography tool is called S-Tools, which is also publicly available at no cost. BMP, GIF, and WAV files can be used as "host" files for steganographically embedded (hidden) messages. The *graphical user interface* (GUI) of S-Tools makes it intuitive to hide files simply by dragging them over the host image window. For an added level of protection, S-Tools can also encrypt its hidden file prior to creating the new image.

# Steganalysis: The Process of Detecting Hidden Messages

How can an intercepted artifact (data file, message, video stream, etc.) be evaluated to determine if an embedded hidden message is present? Detection of steganographic

content is the counterpart of hiding information through steganography. Just as cryptographers are involved in both sides of the coin — developing more secure codes and cracking adversaries' codes — and just as virus authors and antivirus software vendors are engaged in a continuous struggle for dominance, so is the field of steganography characterized by two faces. Steganography is used by those hiding messages and also by those detecting hidden messages. Research in steganography focuses on developing new steganographic techniques, but it is also focused on detection and deciphering of stenographic content.

This process is termed steganalysis and is often explained in the terms of the "Prisoner's Dilemma." In this analogy, two prisoners, Alfred and Bob, attempt to send each other secret messages. Each secret message is hidden in an innocuous message carried and inspected by the warden. Different scenarios lead to different results. One scenario is the assumption that there has been the possibility of data exchange between the prisoners before imprisonment. If Alfred and Bob can communicate and exchange a key before imprisonment, the key can be used to hide the existence of the message, and separating the hidden message from the cover is only possible if the key is known to an interceptor. Of course, the presence of hidden information can still be suspected or detected, even if the actual message cannot be obtained.

An alternative scenario relates to the role of the warden. The warden may have a passive role where he only checks the cover message for hidden content, but does not actively change the cover message. The warden will deliver the message as long as he does not notice anything unusual about the cover message and/or discovers the presence of hidden content. In this case, the safety of the hidden message relies solely on its ability to remain hidden from detection. On the other hand, if the warden actively alters the cover message, even if only slightly, then he may potentially destroy the hidden message. In this case, the safety of the hidden message relies not only on its ability to remain hidden, but also in resistance to changes in the cover. For example, writing messages in invisible ink might escape the attention of the warden, but would not survive photocopying.

## The Computer Security Student gets the Bad Guy

Brokerage houses, including the fictitious Bull Run Investments, are privy to a great deal of sensitive financial information before it is publicly reported. It is illegal for employees to use "insider information" to achieve financial gain for themselves. Further, it is illegal for brokerage employees to supply such information to family, friends, or others before that information is released to the public. Any such illegal release of information would constitute a major security threat to Bull Run Investments, and could jeopardize its future.

Shady Profit was a brokerage employee with a questionable past. The U.S. *Securities and Exchange Commission* (SEC) has been on the trail of Mr. Profit for quite some time. It seemed that nearly every time there was a corporate merger where Bull Run Investments was involved, Mr. Profit's girlfriend, Imma Rich, would purchase large quantities of the acquired company prior to the announcement, thereby profiting on the sale after the price increased.





An interesting pattern emerged where Ms. Rich would purchase the stock the morning that Mr. Profit got word of the merger. The SEC knew that Mr. Profit was not telling Ms. Rich the news in person because she worked across town and she was not allowed in the Bull Run Investments office. Furthermore, the SEC had a tap on Mr. Profit's phone and no insider information was given over the phone, so it started investigating Mr. Profit's e-mail. The messages sent to Ms. Rich appeared to be innocuous. There were standard joke e-mails, several personal messages planning lunch, and a few links to Web sites. Everything seemed legitimate — the jokes were just jokes, the personal messages didn't contain any merger information, and the links were to photographs of restaurants on Mr. Profit's Web site. It was not until Kari Catchum, the fresh college graduate at the SEC who had taken a class in computer security, investigated the photographs that the SEC was able to build a case against Mr. Profit.

As it turns out, as soon as Mr. Profit received word of a merger, he edited his Web site with Microsoft FrontPage. He modified an image file of a certain restaurant by adding a ciphertext message to it using S-Tools (a program used for steganography). He would then send an e-mail to Ms. Rich asking her to join him for lunch at a restaurant and include a link to the photograph of the restaurant on his Web site. Ms. Rich would then follow the link, find the photograph, use S-Tools to obtain the ciphertext hidden in the picture, and finally decode the ciphertext to plaintext using their agreed-upon algorithm. Armed with the decoded company name, Ms. Rich would then purchase the stock before the public had access to the news. It was a brilliant scheme until Ms. Catchum investigated the photographs using steganalysis methods.

# Conclusion

Stenography can be used in business environments, well beyond the prevailing image of steganography as a tool for spies, criminals, and terrorists. Though steganography offers potential for legitimate and illegitimate purposes, the focus of *chief security officers* (CSOs) and other managers of IT security is naturally on the illicit uses who may threaten organizational resources. If disloyal employees might pass trade secrets to competitors via standard e-mail protocols, without any apparent wrongdoing, the enterprise's security is threatened. If international terrorists might embed a secret message within a legitimate video broadcast or Web site display, national security is endangered. Constant vigilance must be exercised in order to ensure that the enterprise perimeter is not violated through steganographic methods. The cat-and-mouse game is likely to continue, so highly-secure environments (military, competitive industries, etc.) are advised to ensure that their staffs includes individuals with training and responsibility for such analysis.

# References

Anderson, R. J., & Petitcolas, F. A. P. (1998). On the limits of steganography. *IEEE Journal on Selected Areas in Communications, 16*(4), 474-481.

Bartlett, J. (2002, March 17). The ease of steganography and camouflage. *SANS Information Security Reading Room*. Retrieved October 29, 2003, from http://www.sans.org/rr/paper.php?id=762

Clark, E. (2001). A reason to love spam. *Network Magazine, 16*(20), 1.

Cole, E. (2003). *Hiding in plain sight: Steganography and the art of covert communication*. Indianapolis, IN: Wiley Publishing, Inc.

Johnson, N. F. (2003). *Steganography*. Retrieved December 15, 2004, from http://www.jjtc.com/stegdoc/steg1995.html

Jupitermedia Corporation. (2003). *Steganography*. Retrieved Aug. 31, 2003, from http://www.webopedia.com/TERM/S/steganography.html

Kessler, G. C. (2001). Steganography: Hiding data within data. Retrieved July 21, 2005, from http://www.garykessler.net/library/steganography.html

Moulin, P., & O'Sullivan, J. A. (2003). Information-theoretic analysis of information hiding. *IEEE Transactions on Information Theory, 49*(3), 563-593.

Schmidt, M. B., Bekkering, E., & Warkentin, M. (2004, April 14-16). On the illicit use of steganography and its detection, In *Proceedings of the 2004 ISOneWorld International Conference*, Las Vegas, NV (pp. 1-10).

Trithemius, J. (1606). *Steganographia.* Retrieved November 17, 2005, from http://www.esotericarchives.com/tritheim/stegano.htm