[UNIVERSITY OF COLORADO AT COLORADO SPRINGS,COLORADO SPRINGS, COLORADO]

An Overview of Digital Steganography

CS5910 Fall 2010

Sheila Jeruto Magut

12/8/2010

Contents

1. Abstract

The purpose of this paper is to provide an overview to the ancient field of steganography. The first part of the paper delves into the history of steganography. The rest of the paper outlines the various forms of Steganography and the techniques employed today. We then move into a discussion on detection and practical applications for steganography.

2. Introduction

Although the focus of research and study has mainly been on improving ways of making secret information illegible to the unintended recipient (essentially Cryptography), more focus should be out on actually concealing this secret information so that the unintended recipient does not even know such information exists. This is the essence of Steganography. The difference between the two areas of security is that while cryptography seeks to encrypt information, steganography seeks to encode information. The idea behind Steganography is stuffing secret information inside of a carrier medium in order to safely transmit the information with minimal chance that it will be found.

3. What is Steganography?

Steganography can be defined as the art and science of writing hidden messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the

message [1].It is a form of security through obscurity. Simply put, steganography is simply a way of concealing the existence of secret communication.

Steganography is not to be confused with cryptography. Cryptography involves the scrambling of communication to make illegible to unintended recipient. An encrypted message is very conspicuous. One may not know the intended meaning of the message, but it is obvious that it exists. Steganography makes an attempt to hide the fact that the secret communication even exists, thereby not drawing attention to it. It replaces bits of unused data into the file- (i.e. graphics, sound, text, audio, or video) with some other bits that have been obtained secretly or unauthorized manner.

Below is a diagram that summarizes the Steganographic process[10]:

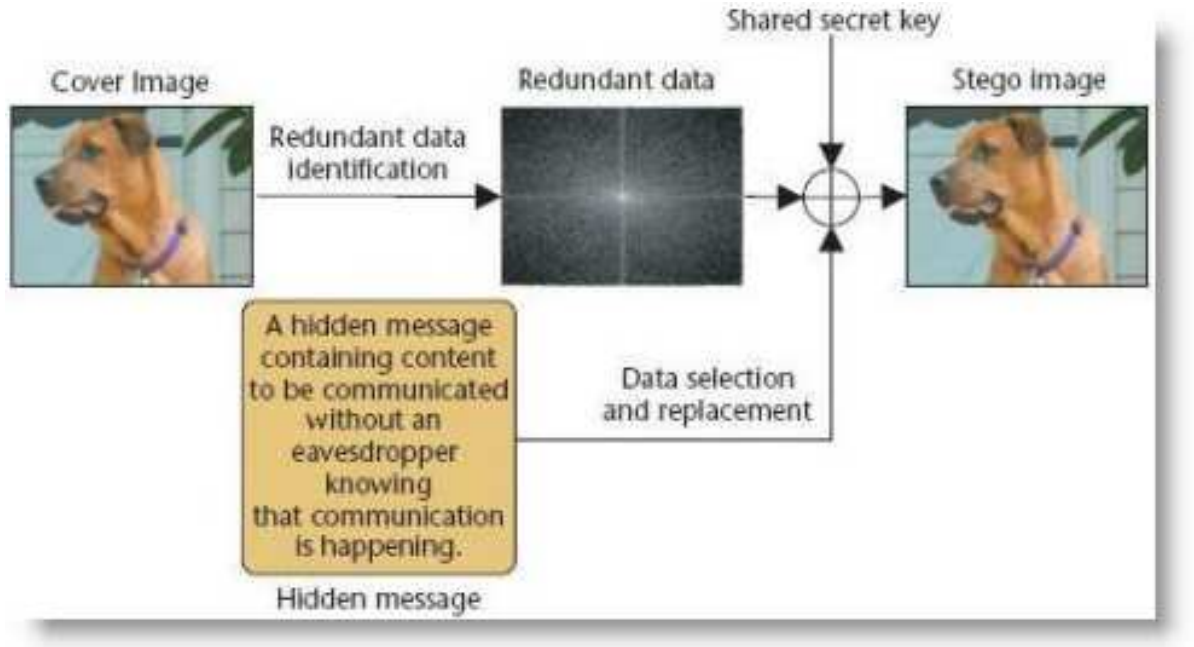

Figure 1 The Steganographic process

4. History of Steganography

Steganography comes from the Greek words **Steganós** (Covered) and **Graptos** (Writing).Steganography been used in various forms for the past 2500.The practice can be traced back to the Golden Age in Greece where it is believed that the Miletus ruler, Histaeus used steganography to send a secret message to his friend Aristagorus to arrange a revolt against the Persian. Histaeus shaved the head of his most trusted slave, tattoed the message on his scalp [2].Once the hair had a grown back, the slave was sent to Aristagorus to deliver the hidden message.

Other accounts of the practice of Steganography in Greek history involved inscribing a secret message on a wooden tablet and covering it with wax. The inscribed tablet would then be easily transported to the intended recipient without arousing any suspicion of the presence of the secret message. This method is said to have been used by the Spartan king Demaratus in 440 BC, to send a secret message warning of impending attacks on Greece.

More recent accounts state that the American military forces practiced Steganography in World War II to conceal secret information from the Japanese. They used the Navaho language, an Athabaskan language spoken in the southwest United States spoken by the Navajo. This was a safer and faster mode of communication than radio communication. Safer because the Japanese had

never heard this language before and therefore could not interpret, and faster because no form of encryption was required nor used.

Null Ciphers were also very popular in the 19th century as a form of Steganography. This form of passing hidden information is still very much in use today especially by prison gangs. Null Ciphers are a classic way to hide a message in another without the use of a complicated algorithm. For example[5] :

PRESIDENT'S EMBARGO RULING SHOULD HAVE IMMEDIATE NOTICE. GRAVE SITUATION AFFECTING INTERNATIONAL LAW. STATEMENT FORESHADOWS RUIN OF MANY NEUTRALS. YELLOW JOURNALS UNIFYING NATIONAL EXCITEMENT IMMENSELY.

Above is a classic example of the use of Null Ciphers. The German Embassy in Washington, DC, sent these messages in telegrams to their headquarters in Berlin during World War I[5]. Reading the ***first*** character of ***every*** word in the first will give the following hidden text:

PERSHING SAILS FROM N.Y. JUNE 1

5. Steganography Techniques

There are many techniques employed in Steganography. These include:

I. Physical Steganography :

This involves using a physical medium to hide messages with. Historical example date back to the Greco-Roman era outline in the History section above, involving the use of wax tables and shaven heads.

More recent examples of physical steganography include the use of microdots. During and after World War II, espionage agents used photographically produced microdots to send information back and forth. Microdots were typically minute, approximately less than the size of the period produced by a typewriter. WWII microdots needed to be embedded in the paper and covered with an adhesive (such as collodion). This was reflective and thus detectable by viewing against glancing light. Alternative techniques included inserting microdots into slits cut into the edge of post cards[1].

II. Printed Steganography

Digital steganography output may be in the form of printed documents. A message, the plaintext, may be first encrypted by traditional means, producing a ciphertext. Then, an innocuous covertext is modified in some way to as to contain the ciphertext, resulting in the stegotext. For example, the letter size, spacing, typeface, or other characteristics of a covertext can be manipulated to carry the hidden message. Only a recipient who knows the technique used can recover the message and then decrypt it.

Another example is that of a picture you stare at for a long time trying to see the image and your eyes finally adjust to it and you see a different shape appear out of the image.

III. Network Steganography

Network steganography is a recently introduced art that described all information concealing methods that may be used to hide data in the normal data transmission of users. Unlike the typical steganographic methods which utilize digital media (images, audio and video files) as a

carrier for hidden data, network steganography utilizes communication protocols' control elements and their basic intrinsic functionality. As a result, such methods are harder to detect and eliminate[1]

Typical network steganography methods involve modification of the properties of a single network protocol. Such modification can be applied to the PDU (Protocol Data Unit),to the time relations between the exchanged PDUs, or both (hybrid methods)

There has been an increase in the use of Voice over Internet Protocol (VoIP) communication as a means of practicing Steganography, essentially hiding secret message in voice data. Take this scenario as an example:

A very famous actor (VFA) has a brief conversation with a well-known director (WKD) over Skype, an application that lets them make free voice calls over the Internet. They discuss the medical problems of VFA's cat in great detail. When the conversation is over, WKD's computer has a sleazy new addition—in a folder on his desktop, there is a picture of a nude teenager, along with her mobile number and the date and time at which WKD will meet her at VFA's pool party for a photo session[9].

The data were secreted among the bits of a digital Voice over Internet Protocol conversation. In this new era of steganography, the mule that coconspirators are using is not the carrier itself but the communication protocols that govern the carrier's path through the Internet. The major

advantage with this is that the longer the communicators talk, the longer the secret message (or more detailed the secret image) they can send. Most strikingly, the concealment occurs within data whose inherent ephemerality makes the hidden payload nearly impossible to detect, let alone stop.

IV. Digital Steganography

Modern steganography entered the world in 1985 with the advent of the personal .The advents of new technology make this a seemingly limitless art form. Digital Steganography utilizes digital media (images, audio and video files) as a cover for hidden data. The basic idea involves hiding or stuffing bits or bytes into various forms of media, and then transmitting that media across a network. There are now over 800 known digital steganography tools available at little or no cost[1].

There are numerous ways to conceal information in digital steganography. The most common are LSB and Injection[8].

a. Least Significant Bit (LSB)

The Least Significant Bit (LSB) is a technique used in digital Steganography. The least significant bit term comes from the numeric significance of the bits in a byte. The high-order or most significant bit is the one with the highest arithmetic value (i.e., 2^7=128), whereas the low-order or least significant bit is the one with the lowest arithmetic value (i.e., 2^0=1).

The *least significant bits* are the vulnerable areas of the file and are the playing ground for Steganography. These bits/bytes can be substituted with the information to be hidden without significantly altering the file [8].

A simple example of LSB substitution[5]; to conceal the character 'G' across the eight bytes below of a carrier file (the least significant bits are underlined):

1001010<u>1</u> 0000110<u>1</u> 1100100<u>1</u> 1001011<u>0</u>

0000111<u>1</u> 1100101<u>1</u> 1001111<u>1</u> 0001000<u>0</u>

A 'G' is represented in the American Standard Code for Information Interchange (ASCII) as the binary string 01000111. These eight bits can be written to the LSB of each of the eight carrier bytes as follows:

1001010<u>*0*</u> 0000110<u>1</u> 1100100<u>*0*</u> 1001011<u>0</u>

0000111<u>*0*</u> 1100101<u>1</u> 1001111<u>1</u> 0001000<u>*1*</u>

b. Injection

As the name suggests, this method involves simply injecting the information to be hidden in the carrier file (a file which contains hidden information) [8]

6. Digital Steganography Methods

I. Images

Steganography in images involves hiding certain secret information within pictures. An image file can be said to be simply a binary file containing a binary representation of the color of each pixel (picture element) comprising the image.

Images typically use either 8-bit or 24-bit color schemes. Steganography is best done using BMP (Bitmap) images as they use the 8-bit color scheme. Bitmap images basically use arrays of RGB (i.e., the three primary colors red, green and blue) values to comprise the total image. Each color is denoted by a byte (8-bits).

Steganography takes advantage of this binary color representation. Information hiding is done using LSB, the last bit or two of each RGB value is substituted with bits from the secret message. For instance, if you have a secret message stored in plain text, you could read that message in two bits at a time, and simply force the next RGB values of the next pixel to match those two bits [1]. Although this inevitably changes the final color of the pixel, the overall effect on the image is likely unnoticeable to the average human eye [3].

In this manner, the entire secret message can be encoded into the picture, with minimal effect on the image. Similarly, the intended recipient can decode the message by extracting the last bit(s) of each RGB value and concatenating them together to derive the secret message.

Although Steganography can be performed on other image files types such as Jpegs, it is more difficult than with Bitmap image file types. This is due to the fact that Jpegs use “lossy” compression to reduce the disk size of the overall file, which means that the expanded image is very nearly the same as the original but not an exact duplicate. Jpegs do not store RGB values for each pixel in the image. In the case of jpegs, the easiest way to do this might be to change

an entire pixel's value to match that of the secret method. As long as the method of choosing the pixels to manipulate is understood by both sender and receiver, the method will work just as well [4].

There are many good examples everywhere of how Steganography is employed in images. Figure 2 below is a map of the Burlington, Vermont, airport This image has been embedded into various images including the pictures below (Figure 3 and Figure 4) [5].

Figure 2.This map is hidden the images below

Figure 3.A GIF file containing the airport map.

Figure 4 A JPEG carrier file containing the airport map.

II. Audio

Steganography in audio files is pretty much the as in image files, i.e. stuffing bits/bytes. However, here the goal is to avoid audible distortion.

There are a number of techniques that can be employed in concealing information inside audio files

i. Low Bit Encoding :This method is similar to the LSB method used in images. The main problem with this technique is that it is usually to the human ear.

ii. Spread Spectrum: this involves appending random noise to the signal, the information is hidden inside a carrier medium and transmitted across the frequency spectrum

iii. Echo Data Hiding: this does exactly what the name suggests; using echoes in sound files. This is done by simply adding extra sound to an echo inside an audio file. This method is preferred because it can actually improve the sound of the audio inside an audio file.

7. Practical Uses

Watermarking

Steganography can be used for digital watermarking, where a message (being simply an identifier) is hidden in an image so that its source can be tracked or verified (for example, Coded Anti-Piracy), or even just to identify an image.

Branding/Copyright

Some printer companies such as HP and Xerox use Steganography in their printers for copyrighting and branding purposes. In this case, tiny dots containing printer serial numbers, date and time stamps are printed on each page.

Alleged use by terrorist

It is alleged that Al Qaeda terrorists used Steganography in planning and executing the 9/11 attacks on America. These claims were later thought to be false as no evidence to support them were found.

Alleged use by intelligence services

It is also alleged that Russian foreign intelligence service use Steganography to communicate with spies abroad

8. Comparison with Cryptography

Steganography is not be confused with Cryptography. The goal of Cryptography is to prevent an unintended recipient from determining the meaning of the secret message. The practice involves scrambling the message thereby rendering it illegible to the unintended recipient. The goal Steganography is to prevent the unintended recipient from eve suspecting that a secret message exists. The difference between steganography and cryptography is that in cryptography, one can tell that a message has been encrypted, but he cannot decode the message without knowing the proper key. In steganography, the message itself may not be difficult to decode, but most people would not detect the presence of the message.

The advantage of steganography, over cryptography alone, is that messages do not attract attention to themselves. Plainly visible encrypted messages, no matter how unbreakable, will

arouse suspicion. Therefore, whereas cryptography protects the contents of a message, steganography can be said to protect both messages and communicating parties.

It is prudent to recognize that steganography is a field that need not, nor should not, stand alone .Employing Cryptography with Steganography yields a powerful form of security, as it becomes very difficult to detect. Encrypting a secret message first and the concealing it in a given medium, yields the ultimate in private communication.

9. Limitations of Steganography

Steganography faces similar constraints as cryptography. For instance, if Alice wants to send an image with a hidden message to Bob, she must first privately agree with Bob on a method of steganography. Under the encryption model, Bob can be fairly sure when he's got some ciphertext. However, in the steganography model, it will be difficult for Bob to know when an image is just an image[7].

Now,let us assume that Alice borrows Bob`s neglects to inform him to pay special attention to every 73rd byte in the images she sends him. Because Bob is oblivious to Alice's steganographic efforts, the large number of pictures he receives from her will only decrease the chance that Bob will let Alice borrow his digital camera again[7].

The size of the medium restricts the amount of data that can be effectively hidden in the medium. The fewer constraints that exist on the integrity of the medium, the more potential it has for hiding data. For example, this paragraph is constrained by the rules of the English language and a specific topic of discussion. It would be difficult for me to hide a secret message in this paragraph due to the limited number of ways one can reasonably alter this text under those constraints[7].

10. Detection (Steganalysis)

The field of study into detecting messages hidden using steganography is known as **Steganalysis** [6]**.** It is a relatively young research discipline, with little research before the 90s.Steganalysis can be thought of as the official countermeasure to Steganography [7].It is a very treacherous and tedious process and unless the encoding format used to hide data is well-defined and common, it can be almost impossible to detect the presence of hidden data in any medium.

Unlike with cryptography where the medium is known to contain the secret message and therefore the focus is on extracting the secret message, with steganography, the medium is unknown to contain the given message. In fact, the medium could contain no secret information at all. The major task is to determine whether or not a secret message.

There are many methods that can be used to detect Steganography such as:

- Viewing the file and comparing it to another copy of the file found on the Internet

- Listening to the file. This is similar to the above method used for trying to manually detect Steganography in picture files.

The basic technique employed in Steganalysis is statistical analysis of an image or audio file. This can help to determine whether the statistical properties of the files deviate from the expected norm.

Steganalysis techniques can be classified in a similar way as cryptanalysis methods, largely based on how much prior information is known [5].

- ***Steganography-only attack:*** The steganography medium is the only item available for analysis.
- ***Known-carrier attack***: The carrier and steganography media are both available for analysis.
- ***Known-message attack:*** The hidden message is known.
- ***Chosen-steganography attack:*** The steganography medium and algorithm are both known.
- ***Chosen-message attack:*** A known message and steganography algorithm are used to create steganography media for future analysis and comparison.
- ***Known-steganography attack***: The carrier and steganography medium, as well as the steganography algorithm, are known

I. Steganalysis Tools

There are various tools that can be used in Steganalysis. Below is a brief outline of a few of these tools.

Stegdetect

Stegdetect is a free program . It contains a database of images that are known NOT to have secret information embedded in them. When a questionable image is presented to the program, it compares the image to the store of images it knows are clean.

Steganography Analyzer Artifact Scanner (StegAlyzerAS)

StegAlyzerAS works by giving you the capability to scan an entire file system, or individual directories, on suspect media for the presence of Steganographically enhanced media. And, unlike other popular forensic tools, it has the capability to perform an automated or manual search of the Windows Registry to determine whether or not any Registry keys or values exist that can be associated with a particular Steganography application [8].

Steganography Analyzer Signature Scanner (StegAlyzerSS)

StegAlyzerSS is similar to SteAlyzerSS in that it has the capability to scan every file on a suspect media. However,it searches for the presence of hexadecimal byte patterns, or signatures, of particular Steganography applications in the files. If a known signature is detected, it may be possible to extract information hidden with the Steganography application associated with the signature[8].

11. Summary

Steganography is a large, interesting, and mostly unexplored subject of study. It is a field outside of the mainstream cryptography and system administration that are seen day after day. Where cryptography seeks encryption, steganography seeks obscurity. The two, however, are not necessarily mutually exclusive. When a secret message is encrypted before being hidden using a steganographic technique, the result is even more obscure, safe, and undetectable.

Although very little attention has been paid to the art, it is also quite. Steganography may, in fact, be all too real, as there have been several reports that the terrorist organization behind the September 11 attacks in America used steganography as one of their means of communication. It is therefore very necessary that more research and studies into the detection of Steganographic efforts , as security and privacy continue to increase in importance in the world today

References

[1] “Steganography”, November 2009. [Online]. Available: http://en.wikipedia.org/wiki/Steganography [Accessed: Dec. 07, 2010]

[2] Bryan Clair, “Steganography: How to Send a Secret Message”, October 2001. [Online]. Available: http://strangehorizons.com/2001/20011008/steganography.shtml [Accessed: Dec. 06, 2010]

[3] Gary C. Kessler, “Steganography: Hiding Data within Data”, September 2001. [Online]. Available: http://www.garykessler.net/library/steganography.html [Accessed: Dec. 07, 2010]

[4] Jeff Hinson,” An Introduction to Digital Steganography”, December 2009. [Online]. Available: cs.uccs.edu/~jhinson/cs591/files/jhinson_digital-steganography.pdf [Accessed: Dec. 07, 2010]

[5] Gary C. Kessler, “An Overview of Steganography for the Computer Forensics Examiner”, February 2004. [Online]. Available: http://www.garykessler.net/library/fsc_stego.html [Accessed: Dec. 07, 2010]

[6] “Steganalysis”, November 2009. [Online]. Available: http://en.wikipedia.org/wiki/Steganalysis [Accessed: Dec. 07, 2010]

[7] Donovan Artz, “Digital Steganography: Hiding Data within Data”, June 2001. [Online]. Available:http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=935180&userType=&tag=1 [Accessed: Dec. 07, 2010]

[8] Aelphaeis Mangarae [Zone-H.Org], “Steganography FAQ”, March 2006. [Online]. Available:http://infosecwriters.com/text_resources/pdf/Steganography_AMangarae.pdf [Accessed: Dec. 07, 2010]

[9] Józef Lubacz et al, “Vice Over IP: The VoIP Steganography Threat”, Feb 2010. [Online]. Available: http://spectrum.ieee.org/telecom/internet/vice-over-ip-the-voip-steganography-threat/0 [Accessed: Dec. 07, 2010]

[10] “Steganography: Art of Sending Secret Message. [Online]. Available: http://www.aboutonlinetips.com/what-is-steganographyand-how-to-embed-secret-messages-in-other-messages/ [Accessed: Dec. 08, 2010]