Interested in learning more about security?

SANS Institute
InfoSec Reading Room

This paper is from the SANS Institute Reading Room site. Reposting is not permitted without express written permission.

Steganalysis: Detecting hidden information with computer forensic analysis

With the wide use and abundance of steganography tools on the Internet, law enforcement authorities have concerns in the trafficking of illicit material through web page images, audio, and other files. Methods of detecting hidden information and understanding the overall structure of this technology is crucial in uncovering these activities. Digital image steganography is growing in use and application. In areas where cryptography and strong encryption are being outlawed [1], people are using steganography to avoid the...

Copyright SANS Institute
Author Retains Full Rights

AD

SANS/GIAC Practical Assignment
for GSEC Certification
written by Pierre Richer
Version 1.4b

Steganalysis: Detecting hidden information with computer forensic analysis

Table of Contents

1) Introduction…… 2

2) What is steganography?…… 2

   a) -Why is steganography used?…… 3

   b) -Tools used to hide information…….. 4

3) What is steganalysis?……..4

   a) -Types of attacks used by the steganalyst……4

   b) -Steganography signatures…… 5

   c) -Visual detection……6

4) Detecting hidden information with various tools……..6

5) Conclusion……9

6) References……...10

© SANS Institute 2003, Author retains full rights

© SANS Institute 2003, As part of the Information Security Reading Room. Author retains full rights.

Introduction

With the wide use and abundance of steganography tools on the Internet, law enforcement authorities have concerns in the trafficking of illicit material through web page images, audio, and other files. Methods of detecting hidden information and understanding the overall structure of this technology is crucial in uncovering these activities.

Digital image steganography is growing in use and application. In areas where cryptography and strong encryption are being outlawed [1], people are using steganography to avoid these policies and to send these messages secretly.

In this paper I shall give a brief definition of steganography and steganalysis in general to provide a good understanding of these two terms, but more importantly, I shall talk about how to detect the existence of hidden information such as innocent looking carriers of digital media like text, JPEG images, and MP3 audio files with the help of various tools.

What is steganography?

The word steganography comes from the Greek name "steganos" (hidden or secret) and "graphy" (writing or drawing) and literally means hidden writing. Steganography uses techniques to communicate information in a way that is hidden.

Steganography hides the existence of a message by transmitting information through various carriers. Its goal is to prevent the detection of a secret message.

The most common use of steganography is hiding information from one file within the information of another file. For example, cover carriers, such as images, audio, video, text, or code represented digitally, hold the hidden information. The hidden information may be plaintext, ciphertext, images, or information hidden into a bit stream. The cover carrier and the hidden information create a stego-carrier. A stegokey, such as a password, is additional information to further conceal a message. An investigator who does not possess the name of the file and the password cannot know about the file's existence.

For example, the result of information hidden within a cover image is a stego-image, and the result of information hidden within a video is a stego-video and so forth. The process may be defined as follows:

cover medium + hidden information + stegokey = stego-medium [2]

© SANS Institute 2003, As part of the Information Security Reading Room. Author retains full rights.

Most people would probably detect no loss in the quality of the image. Therefore, an image posted on the Internet could contain a secret message and avoid suspicion. An article in *USA Today* [3] claimed that terrorist groups are using steganography to communicate without being detected. According to experts, the article lacked technical information to prove these claims. But, of course, there are many other ways that steganography is being used by people with harmless motives. For example, some photo agencies will use steganography to create digital "watermarks" of their pictures to protect their trademark.

Steganography is different from cryptography. Cryptography enciphers or garbles files to hide the information. A decryption key or password is needed to retrieve the information. A drawback to cryptography is that there are many ways to retrieve this encrypted information once it has been discovered. The most obvious example is by knowing about its existence, investigators can apply the many softwares available to decrypt the hidden information. Another obvious way is to obtain the password or decryption key from the owner.

Why is steganography used?

There are many reasons why steganography is used, and it is often used in significant fields. It can be used to communicate with complete freedom even under conditions that are censured or monitored. It can also be used to protect private communications where the use of the cryptography is normally not allowed or would raise suspicion.

There are also at least two techniques that are part of steganography.

Watermarking:

- Protects copyright owners of digital documents by hiding a signature in the information in a way that even a modified part of the document conserves the signature.

- Prevents discovery by marking in a hidden and unique way every copy of a confidential document.

Cover channel:

- Allows people to communicate secretely by establishing a secret communication protocol.

- Allows non-authorized communication through authorized communication of a firewall.

© SANS Institute 2003, As part of the Information Security Reading Room. Author retains full rights.

Tools used to hide information

There are two possible groups of steganographic tools: the image domain and the transform domain.

Image domain tools include bit-wise methods that apply least significant bit (LSB) insertion and noise manipulation. The tools used in this group are StegoDos, S-Tools, Mandelsteg, EzStego, Hide and Seek (versions 4.1 through 1.0 for Windows 95), Hide4PGP, Jpeg-Jsteg, White Noise Storm, and Steganos. The image formats used in these steganography methods cannot be lost and the information can be rearranged or recovered.

The transform domain tools include those groups that manage algorithms and image transforms such as Discrete Cosine Transformation (DCT).

The DCT is a technique used to compress JPEG, MJPEG and MPEG in which pixel values are converted to frequency values for further processing. This process makes it difficult for visual analysis attacks against the JPEG images. [11]

These two methods hide information in more areas of the cover and may manipulate image properties such as luminance or the color palette. These methods will allow more hidden information (about 30 percent the size of the carrier) in a carrier file. JPEG images are used on the Internet because of their compression quality, which does not degrade the image.

What is steganalysis?

Steganalysis is the discovery of the existence of hidden information; therefore, like cryptography and cryptanalysis, the goal of steganalysis is to discover hidden information and to break the security of its carriers [4].

Types of attacks used by the steganalyst

Stego-only attack: Only the stego-object is available for analysis. For example, only the stego-carrier and hidden information are available.

Known cover attack: The original cover-object is compared with the stego-object and pattern differences are detected. For example, the original image and the image containing the hidden information are available and can be compared.

© SANS Institute 2003, As part of the Information Security Reading Room. Author retains full rights.

Known message attack: A known message attack is the analysis of known patterns that correspond to hidden information, which may help against attacks in the future. Even with the message, this may be very difficult and may be considered the same as a stego-only attack.

Chosen stego attack: The steganography tool (algorithm) and stego-object are known. For example, the software and the stego-carrier and hidden information are known.

Chosen message attack: The steganalyst generates a stego-object from some steganography tool or algorithm of a chosen message. The goal in this attack is to determine corresponding patterns in the stego-object that may point to the use of specific steganography tools or algorithms.

Known stego attack: The stegonography tool (algorithm) is known and both the original and stego-object are available.

Steganography signatures

Unusual patterns in the stego-image are obvious and create suspicion. For example, unused areas on a disk can be used to hide information. A number of disk analysis utilities such as EnCase [5] and ILook Investigator © [6] are available, which can report on and filter hidden information in unused clusters or partitions in storage devices.

Filters can also be applied to capture TCP/IP packets that contain hidden or invalid information in the packet headers. TCP/IP packets have unused space in the packet headers. The TCP packet header has six reserved or unused bits, and the IP packet header has two reserved bits [10]. Information can also be hidden in the unused bits found in the Type of Service (TOS) Field and Flags of IP headers. Other methods to hide information under TCP/IP are exploiting the optional fields in IP headers, Timestamp, and Time to Live (TTL). These techniques can also be applied to other protocols such as Novell NetWare [13]. Thousands of packets are transmitted with each communication channel, which provide an excellent way to communicate secretly. This technique of hiding information is unsafe because TCP/IP headers might get overwritten in the routing process, and reserved bits could be overwritten, thus rendering the hidden information useless.

The technology of firewalls is also greatly improving. For example, you can set filters to determine if packets are coming from within the firewall's domain. Also, with the validity of the SYN and ACK bits, the filters can be configured to catch packets that have information in presumed unused or reserved space, just like you can set certain firewalls to exclude such packets with spoofed addresses.

© SANS Institute 2003, As part of the Information Security Reading Room. Author retains full rights.

Visual detection

By looking at repetitive patterns, you can detect hidden information in stego images. These repetitive patterns might reveal the identification or signature of a steganography tool or hidden information. Even small distortions can reveal the existence of hidden information.

You can analyze these patterns by comparing the original cover images with the stego images and try to see differences. This is called a known-cover attack. By comparing numerous images, patterns become possible signatures to a steganography tool. A few of these signatures might identify the existence of hidden information and the tools used to embed the messages. With this information, if the cover images are not available for comparison, the derived known signatures are enough to imply the existence of a message and identify the tool used to embed the message.

Detecting hidden information with various tools

1) Guidance Software, Inc.
http://www.guidancesoftware.com

"Award winning and validated by the courts, EnCase allows law enforcement and IT professionals to conduct a powerful, yet completely non-invasive computer forensic investigation. EnCase features a intuitive GUI that enables examiners to easily manage large volumes of computer evidence and view all relevant files, including "deleted" files, file slack and unallocated space. The solution effectively automates core investigative procedures, replacing archaic, time-consuming and cost-prohibitive processes and tools."

Guidance Software, Inc.

In EnCase investigators must identify and match the MD5 hash value of each suspected file. They must import or build a library of hash sets (in this particular case, a steganography software) with the library feature in EnCase [5]. The hash sets will identify stego file matches.

Also, investigators must be careful when they create hash sets to discover steganography, to prevent false positives. For example, investigators must use safe hash sets to filter harmless files from their investigation. System files that have not been modified since installation are included in a safe hash set.

© SANS Institute 2003, As part of the Information Security Reading Room. Author retains full rights.

2) ILook Investigator
http://www.ilook-forensics.org/

"ILook Investigator© is a forensic analysis tool used by thousands of law enforcement labs and investigators around the world for the investigation of forensic images created by many different imaging utilities."

ILook Investigator

3) Comparing the MD5 hash values of two files with the program, md5sum.exe. This program is found easily on the Internet.

Investigators use MD5, an algorithm, to generate a 128-bit fingerprint of an file, irregardless of its size. Because there are 10 exponent, 38 possible hash values, it is unlikely that files would have the same hash value. Furthermore, right now, manufacturing a file that generates a particular hash value is "computationally infeasible." Therefore, at the moment, files are identified reliably through their MD5 hash value.

As shown in Fig. 1, I have compared the MD5 hash value of an original photo (cloudysunset.bmp) with the same photo containing the hidden information using the steganography tool, S-Tools. You will notice the different MD5 hash values of the two photos. Obviously, this tells the investigator that the image contains information embedded within the file and will require further analysis from the investigator.

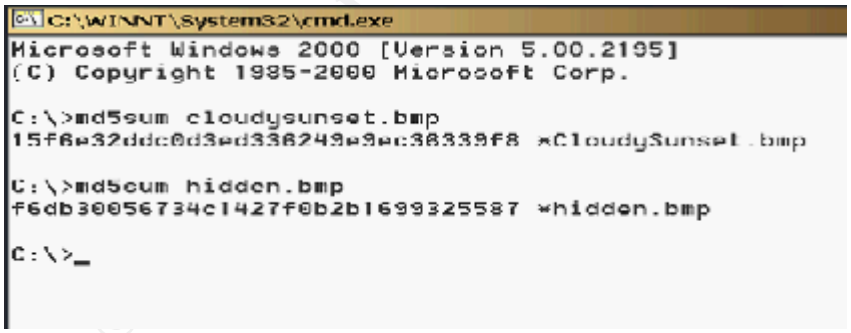

Fig. 1

© SANS Institute 2003, As part of the Information Security Reading Room. Author retains full rights.

4) Detecting hidden information with Stegdetect and Xsteg (freeware) [7]

Stegdetect was written by Niels Provos in 2001, the author of the steganography program called Outguess. Stegdetect is reliable in detecting JPEG images that have content embedded with JSteg, JPHide and OutGuess[9].

Stegdetect also contains a utility using brute-force attack that launches dictionary attacks against JSteg and JPHide. This utility is called Stegbreak. Xsteg is the GUI (Graphical User Interface) to Stegdetect.

Figure 2 shows some sample output from Stegdetect. It also shows an estimate in the detection process. The output from Stegdetect lists the steganographic systems found in each image or negative results if no steganographic content is detected. Stegdetect reveals the level of certainty of the detection with one to three stars.

hd/test1.jpg : outguess (***)
hd/test2.jpg : outguess(old)(***)
hd/test3.jpg : jsteg(***)
hd/test4.jpg : negative
hd/test5.jpg : jphide(*)

Fig. 2

5) Detecting hidden information with file compression [8]: I have tested this method by compressing an original bitmap file and the same file used as a carrier file (S-Tools) by using a simple compression software such as WinZip (URL: http://www.winzip.com/ddchomea.htm) This method of detection would be categorized as a known-cover attack.

By comparing the properties of the two files, I noticed that the original file compressed better than the carrier file.

6) Detecting hidden information with Stego Watch from WetStone Technologies Inc. (commercial product).

"Stego Watch is available as both an outsourced monitoring/scanning service as well as a stand-alone software package. The service option is convenient for customers wishing to leave the administration of the program in the hands of expert WetStone staff. The software package is for those customers who choose to make the steganography function internal to their system. Either selection comes with the technical support and expertise of the Stego Watch team."

WetStone Technologies, Inc.

© SANS Institute 2003, As part of the Information Security Reading Room. Author retains full rights.

Reference taken from URL:
http://www.wetstonetech.com/stegowatchdatasheet.pdf

The program, Stego Watch, from WetStone Technologies examines images. It tries to determine through a mathematical model if steganography is used or if the image has been tampered. For example, organizations can scan their networks routinely for suspicious activities. [12]

7) Spam mimic transforms a text message into e-mail spam. If you encounter such a suspicious e-mail on someone's computer, you can copy the spam and paste the text in decode mode on the spam mimic Web site http://www.spammimic.com/ and voilà the secret message is revealed. There is, however, a new option added to this program. Someone can encode a message with a password. The Web site indicates that it doesn't use strong password encryption; therefore, you can try any password cracking tool found on the Internet to see if it will reveal the hidden information.

8) Foundstone offers its Forensic Toolkit, which is used to examine the files on a disk drive for unauthorized activity. It lists files by their last access time, searches for access times between certain time frames, scans the disk for hidden files and data streams. Furthermore, it dumps files and security attributes, reports on audited files, and discovers altered ACL's. It also sees if a server reveals too much info via NULL sessions.
http://www.foundstone.com/knowledge/proddesc/forensic-toolkit.html

Conclusion

Steganography certainly has some beneficial advantages. It is an effective tool for protecting personal information, and organizations are spending a lot of energy and time in analyzing steganography techniques to protect their integrity. However, steganography can also be detrimental. It is hindering law enforcement authorities in gathering evidence to stop illegal activities, because these techniques of hiding information are becoming more sophisticated.

Although steganography is becoming more advanced, it is still a science that is not well-known. But it may become very popular in the near future. Its use on the Internet is certainly promising. That is why law enforcement authorities must continually stay abreast of this technology, because there will always be some new program to hinder their efforts.

© SANS Institute 2003, As part of the Information Security Reading Room. Author retains full rights.

References :

[1] Andy Oram
The McCain-Kerrey "Secure Public Networks Act"
URL: http://www.cpsr.org/cpsr/nii/cyber-rights/web/mccain-kerrey.html

[2] Neil F. Johnson and Sushil Jajodia
"Steganalysis of Images Created Using Current Steganography Software"
URL: http://www.jjtc.com/ihws98/jjgmu.html

[3] Jack Kelley
"Terror groups hide behind Web encryption" USA Today, Feburary 2001.
URL: http://www.landfield.com/isn/mail-archive/2001/Feb/0038.html

[4] Deborah Radcliff "Computer World"
URL:
http://www.computerworld.com/securitytopics/security/story/0,10801,71726,00.html

[5] Guidance Software, Inc.
URL: http://www.encase.com/products/software/encaseforensic.shtm

[6] IRS Criminal Investigation Electronic Crimes Program ILook Investigator © Elliot Spencer
URL: http://www.ilook-forensics.org/

[7] Outguess; Steganography Detection with Stegdetect.
URL: http://www.outguess.org/detection.php

[8] Fabian Hansmann
"Fighting steganography detection" 04 January 1997
URL: http://www.woodmann.com/fravia/fabian2.htm

[9] Niels Provos, Peter Honeyman
"Detecting Steganographic Content on the Internet"
URL: http://www.citi.umich.edu/techreports/reports/citi-tr-01-11.pdf

[10] The SANS Security Essentials Lecture notes: IP Concepts I, IP Concepts II

[11] Uma maheswaran Sekarji
"DETECTION OF HIDDEN INFORMATION IN IMAGES USING NEURAL NETWORKS"; Department of Information Technology SASTRA Tanjore - 613 402, India
URL: http://www.know.comp.kyutech.ac.jp/STEG02/Papers/pdf-files/paper_uma.pdf

© SANS Institute 2003, As part of the Information Security Reading Room. Author retains full rights.

[12] Tom Kellen
"Hiding in Plain View: Could Steganography be a Terrorist Tool?"
November 19, 2001
URL: http://rr.sans.org/steg/plain_view.php

[13] Sue Adamkiewicz, Mike High, Hui Huang, Phillip Perry
"Steganography: The Unseen Threat to Information Security"
March 21, 2000
URL:
http://www.szgti.bmf.hu/~mtoth/download/Szteganografia/RandysStudents.pdf

Books:

Stefan Katzenbeisser, Fabien A. P. Petitcolas
INFORMATION HIDING techniques for steganography and digital watermarking,
Boston Massachusetts: Artech House, 2000.

Neil F. Johnson, Zoran Duric, Sushil Jajodia, *Information Hiding: Steganography and Watermarking – Attacks and Countermeasures,* Boston, Massachusetts: Kluwer Academic Publishers, 2001.

© SANS Institute 2003, Author retains full rights

© SANS Institute 2003, As part of the Information Security Reading Room. Author retains full rights.

Last Updated: June 19th, 2013

Upcoming SANS Training

Click Here for a full list of all Upcoming SANS Events by Location

Digital Forensics & Incident Response Summit 2013	**Austin, TXUS**	**Jul 09, 2013 - Jul 16, 2013**	**Live Event**
SANS London Summer 2013	**London, GB**	**Jul 09, 2013 - Jul 16, 2013**	**Live Event**
SANS Rocky Mountain 2013	**Denver, COUS**	**Jul 14, 2013 - Jul 20, 2013**	**Live Event**
SEC528 SANS Training Program for the CompTIA® New Advanced Security Practitioner (CASP) Certification	**Washington, DCUS**	**Jul 22, 2013 - Jul 26, 2013**	**Live Event**
SANS Mumbai 2013	**Mumbai, IN**	**Jul 22, 2013 - Jul 27, 2013**	**Live Event**
SANS San Francisco 2013	**San Francisco, CAUS**	**Jul 29, 2013 - Aug 03, 2013**	**Live Event**
SANS Boston 2013	**Boston, MAUS**	**Aug 05, 2013 - Aug 10, 2013**	**Live Event**
SANS SEC 560: Network Penetration Testing @ Bangalore 2013	**Bangalore, IN**	**Aug 05, 2013 - Aug 10, 2013**	**Live Event**
Critical Security Controls Summit	**Washington, DCUS**	**Aug 12, 2013 - Aug 18, 2013**	**Live Event**
Industrial Control Systems Security Training - DC	**Washington, DCUS**	**Aug 12, 2013 - Aug 16, 2013**	**Live Event**
SANS Virginia Beach 2013	**Virginia Beach, VAUS**	**Aug 19, 2013 - Aug 30, 2013**	**Live Event**
SANS Thailand 2013	**Bangkok, TH**	**Aug 19, 2013 - Aug 31, 2013**	**Live Event**
SANS Melbourne 2013	**Melbourne, AU**	**Sep 02, 2013 - Sep 07, 2013**	**Live Event**
SANS Capital City 2013	**Washington, DCUS**	**Sep 03, 2013 - Sep 08, 2013**	**Live Event**
Network Security 2013	**Las Vegas, NVUS**	**Sep 14, 2013 - Sep 23, 2013**	**Live Event**
SANS Canberra 2013	**OnlineAU**	**Jul 01, 2013 - Jul 13, 2013**	**Live Event**
SANS OnDemand	**Books & MP3s OnlyUS**	**Anytime**	**Self Paced**