This article will appear in the Proceedings of the 1998 IEEE Information Technology Conference, Syracuse, New York, USA, September 1st - 3rd, 1998.

Steganalysis: The Investigation of Hidden Information

Neil F. Johnson and Sushil Jajodia

Center for Secure Information Systems, George Mason University, MS:4A4, Fairfax, Virginia 22030-4444

Abstract

The goal of steganography is to avoid drawing suspicion to the transmission of a hidden message. If suspicion is raised, then this goal is defeated. Discovering and rendering useless such covert messages is a new art form known as steganalysis. In this paper, we provide an overview of some characteristics in information hiding methods that direct the steganalyst to the existence of a hidden message and identify where to look for hidden information.

I. INTRODUCTION

Steganography literally means, "covered writing" and encompasses methods of transmitting secret messages through innocuous cover carriers in such a manner that the existence of the embedded messages is undetectable. Carriers of such messages may resemble innocent images, audio, video, text, or any other digitally represented code or transmission. The hidden message may be plaintext, ciphertext, or anything that can be represented as a bit stream.

Commercial use of steganographic techniques has evolved into digital watermarking. Watermarking does not necessarily conceal the knowledge of the hidden information other than from the human senses. A broad overview of data embedding and watermarking methods is available in [18]. Additional readings, software, and resources used in researching steganography and digital watermarking is available at http://isse.gmu.edu/~njohnson/Steganography.

II. STEGANOGRAPHIC METHODS

The Internet provides an increasingly broad band of communication as a means to distribute information to the masses. Such information includes text, images, and audio to convey ideas for mass communication. Such provide excellent carriers for hidden information and many different techniques have been introduced [1, 3, 12]. Other carriers for hidden information include storage devices [2] and TCP/IP packets [9].

An early approach to hiding information is in text. Invisible inks prove to be a popular medium. Computers bring more capability to information hiding. The layout of a document may also reveal information. Documents may be marked identified by modulations in the positions of lines and words [4]. Adding spaces and "invisible" characters to text provides a method to pass hidden information. An interesting way to see this is to add spaces and extra line breaks in an HTML file. Web browsers ignore these "extra" spaces and lines, but revealing the source of the web page displays the extra characters. For an additional text-based hiding techniques and an algorithm for mimicking the statistical distribution of text to pass information see [20].

Many different methods of hiding information in images exist. These methods range from Least Significant Bit (LSB) or noise insertions, manipulation of image and compression algorithms, and modification of image properties such as luminance. An introduction to steganography and its application to digital images is available from [12].

Other, more robust, methods of hiding information in images include application of the transform domain that take advantage of algorithms and coefficients from processing the image or its components to hide information. These methods hide messages in significant areas of the cover image which makes them more robust to attacks such as compression, cropping, and some image processing than the LSB approach. Many transform domain variations exist; one type is to use the discrete cosine transformation (DCT) as a vehicle to embed information in images. Transformations can be applied over the entire image [5] to blocks through out the image [17, 19], or other variations. Many of these transformation techniques require use of the original, unmarked image to extract the watermark. In [10] a number of papers propose techniques that do not require using the original to extract the watermark [16]. A method that proposes a combination of these techniques from LSB insertion to spread spectrum disbursement of data is described in [14]. A survey of transform domain techniques can be fount in [11].

The LSBs and transforms can also be applied to hide information in audio and video with virtually no impact to the human sensory system. In audio, small echoes can be added or subtle signals can be masked by sounds of higher amplitude [7, 8]. Unused space in file headers of image and audio can be used to hold “extra” information.

Taking advantage of unused or reserved space to hold covert information provides a means of hiding information without perceptually degrading the carrier. The way operating systems store files typically results in unused space that appears to be allocated to a file. For example, under Windows 95 operating system, drives formatted as FAT16 (MS-DOS compatible) without compression use cluster sizes of around 32 kilobytes (K). What this means is that the minimum space allocated to a file is 32K. If a file is 1K in size, then an additional 31K is "wasted" due to the way storage space is allocated. This "extra" space can be used to hide information without showing up in the directory.

Another method of hiding information in file systems is to create a hidden partition. These partitions are not seen if the system is started normally. However, in many cases, running a disk configuration utility (such as DOS's FDISK) exposes the hidden partition. These concepts have been expanded and a novel proposal of a steganographic file system [2]. If the user knows the file name and password, then accesses is granted to the file; otherwise, no evidence of the file exists in the system.

Protocols in the OSI network model have vulnerabilities that can be used to hide information [9]. TCP/IP packets used to transport information across the Internet have unused space in the packet headers. The TCP packet header has six unused (reserved) bits and the IP packet header has two reserved bits. Thousands of packets are transmitted with each communication channel, which provides an excellent covert communication channel if unchecked. The ease in use and abundant availability of steganography tools has law enforcement concerned in trafficking of illicit material via web page images, audio, and other files being transmitted through the Internet. Methods of message detection and understanding the thresholds of current technology are necessary to uncover such activities.

III. STEGANALYSIS

Hiding information within electronic media requires alterations of the media properties that may introduce some form of degradation or unusual characteristics. These characteristics may act as signatures that broadcast the existence of the embedded message, thus defeating the purpose of steganography.

Attacks and analysis on hidden information may take several forms: detecting, extracting, and disabling or destroying hidden information. An attacker may also embed counter-information over the existing hidden information. Due to space limitations we will look at two methods: detecting messages or their transmission and disabling embedded information. These approaches (attacks) vary depending upon the methods used to embed the information in to the cover media.

Our goal is not to advocate the removal or disabling of valid hidden information such as copyrights, but to point out approaches that are vulnerable and may be exploited to investigate illicit hidden information.

Some amount of distortion and degradation may occur to carriers of hidden messages even though such distortions cannot be detected easily by the human perceptible system. This distortion may be anomalous to the "normal" carrier that when discovered may point to the existence of hidden information. Steganography tools vary in their approaches for hiding information. Without knowing which tool is used and which, if any, stegokey is used; detecting the hidden information may become quite complex. However, some of the steganographic approaches have characteristics that act as signatures for the method or tool used. For more information on terminology and concepts with respect to steganalysis see [13].

IV. DETECTING HIDDEN INFORMATION

Unusual patterns stand out and expose the possibility of hidden information. In text, small shifts in word and line spacing may be somewhat difficult to detect by the casual observer [4]. However, appended spaces and "invisible" characters can be easily revealed by opening the file with a common word processor. The text may look "normal" if typed out on the screen, but if the file is opened in a word processor, the spaces, tabs, and other characters distort the text's presentation.

Images too may display distortions from hidden information. Selecting the proper combination of steganography tools and carriers is key to successful information hiding. Some images may become grossly degraded with even small amounts of embedded information. This “visible noise” will give away the existence of hidden information. The same is true with audio. Echoes and shadow signals reduce the chance of audible noise, but they can be detected with little processing.

Only after evaluating many original images and stego-images as to color composition, luminance, and pixel relationships do anomalies point to characteristics that are not "normal" in other images. Patterns become visible when evaluating many images used for applying steganography. Such patterns are unusual sorting of color palettes, relationships between colors in color indexes, exaggerated "noise"

An approach used to identify such patterns is to compare the original cover-images with the stego-images and note visible differences (known-cover attack [13]). Minute changes are readily noticeable when comparing the cover and stego-images. In making these comparisons with numerous images, patterns begin to emerge as possible signatures of steganography software. Some of these signatures may be exploited automatically to identify the existence of hidden messages and even the tools used in embedding the messages. With this knowledge-base, 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. However, in some cases recurring, predictable patterns are not readily apparent even if distortion between the cover and stego-images is noticeable. See [13] for examples of unique signatures of steganography tools as applied to images.

A number of disk analysis utilities are available that can report and filter on hidden information in unused clusters or partitions of storage devices. A steganographic file system may also be vulnerable to detection through analysis of the systems partition information.

Filters can also be applied to capture TCP/IP packets that contain hidden or invalid information in the packet headers. Internet firewalls are becoming more sophisticated and allow for much customization. Just as filters can be set to determine if packets originate from within the firewall's domain and the validity of the SYN and ACK bits, so to can the filters be configured to catch packets that have information in supposed unused or reserved space.

V. DISABLING STEGANOGRAPHY

Detecting the existence of hidden information defeats the steganography's goal of imperceptibility. Methods exist that produce results which are far more difficult to detect without the original image for comparison. At times the existence of hidden information may be known so detecting it is not always necessary. Disabling and rendering it useless seems to be the next best alternative [13]. With each method of hiding information there is a trade off between the size of the payload (amount of hidden information) that can be embedded and the survivability or robustness of that information to manipulation.

The distortions in text noted by appended spaces and "invisible" characters can be easily revealed by opening the file with a word processor. Extra spaces and characters can be quickly stripped from text documents.

The disabling or removal of hidden information in images comes down to image processing techniques. For LSB methods of inserting data, simply using a lossy compression technique, such as JPEG, is enough to render the embedded message useless. Images compressed with such a method are still pleasing to the human eye but no longer contain the hidden information. An explanation of JPEG compression and the relation to steganography can be found in [12].

Tools exist to test the robustness of information hiding techniques in images. These tools automate image-processing techniques such as warping, cropping, rotating, and blurring. Examples and evaluation of these tools are found in [13] and [15]. Such tools and techniques should be used by those considering making the investment of watermarking to provide a sense of security of copyright and licensing just as password cracking tools are used by system administrators to test the strength of user and system passwords. If the password fails, the administrator should notify the password owner that the password is not secure.

Hidden information may also be overwritten. If information is added to some media such that the added information cannot be detected, then there exists some amount of additional information that may be added or removed within the same threshold which will overwrite or remove the embedded covert information. A variation of this approach is explored in [6] to the aspect of counterfeiting watermarks.

Audio and video are vulnerable to the same methods of disabling as with images. Manipulation of the signals will alter embedded signals in the noise level (LSB) which may be enough to overwrite or destroy the embedded message. Filters can be used in an attempt to cancel out echoes or subtle signals but becomes this may not be as successful as expected. A possible "brute force" combination of attacks on echo hiding in audio can be found in [15].

Caution must be used in hiding information in unused space in files or file systems. File headers and reserved spaces are common places to look for "out of place" information. In file systems, unless the steganographic areas are in some way protected (as in a partition), the operating system may freely overwrite the hidden data since the clusters are thought to be free. This is a particular annoyance of operating systems that do a lot of caching and creating of temporary files. Utilities are also available which "clean" or wipe unused storage areas. In wiping, clusters are overwritten several times to ensure any data has been removed. Even in this extreme case, utilities exist that may recover portions of the overwritten information.

As with unused or reserved space in file headers, TCP/IP packet headers can also be reviewed easily. Just as firewall filters are set to test the validity of the source and destination IP addresses, the SYN and ACK bits, so to can the filters be configured to catch packets that have information in supposed unused or reserved space. If IP addresses are altered or spoofed to pass covert information, a reverse lookup in a domain name service (DNS) can verify the address. If the IP address is false, the packet can be terminated. Using this technique to hide information is risky as TCP/IP headers may get overwritten in the routing process. Reserved bits can be overwritten and passed along without impacting the routing of the packet.

VI. Comments and Conclusion

This paper provided an overview of steganalysis and introduced some characteristics of steganographic software that point signs of information hiding. This work is but a fraction of the steganalysis approach. To date general detection techniques as applied to steganography have not been devised and methods beyond visual analysis are being explored. Too many images exist to be reviewed manually for hidden messages so development of a tool to automate the process will be beneficial to analysts.

The ease in use and abundant availability of steganography tools has law enforcement concerned in trafficking of illicit material via web page images, audio, and other transmissions over the Internet. Methods of message detection and understanding the thresholds of current technology are under investigation.

Success in steganographic secrecy results from selecting the proper mechanisms. However, a stego-medium which seems innocent enough may, upon further investigation, actually broadcast the existence of embedded information.

Development in the area of covert communications and steganography will continue. Research in building more robust methods that can survive image manipulation and attacks continues to grow. The more information is placed in the public's reach on the Internet, the more owners of such information need to protect themselves from theft and false representation. Systems to recover seemingly destroyed information and steganalysis techniques will be useful to law enforcement authorities in computer forensics and digital traffic analysis.

VII. References

[1] R. Anderson, (ed.), *Information hiding: first international workshop*, Cambridge, UK. *Lecture Notes in Computer Science*, vol. 1174, Berlin Heidelberg New York: Springer-Verlag, 1996.

[2] R. Anderson, R. Needham, A. Shamir, "The Steganographic File System", *Proc. Information Hiding Workshop*, Portland, Oregon, USA, April 1998. To be published.

[3] W. Bender, D. Gruhl, N. Morimoto, A. Lu, "Techniques for Data Hiding", *IBM Systems Journal* Vol. 35, No. 3&4. MIT Media Lab, pp. 313-336, 1996.

[4] J. Brassil, L. O'Gorman, N.F. Maxemchuk, S.H. Low, "Document Marking and Identification using Both Line and Word Shifting", *Infocom 1995*, Boston, April 1995, pp. 853-860.

[5] I. Cox, J. Kilian, T. Shamoon, T. Leighton, "A Secure, Robust Watermark for Multimedia", In: [1] pp 185-206, 1996.

[6] S. Craver, N. Memon, B. Yeo, N.M. Yeung, "Resolving Rightful Ownerships with Invisible Watermarking Techniques", Research Report RC 20755 (91985), Computer Science/Mathematics, IBM Research Division, 1997.

[7] E. Franz, A. Jerichow, S. Moller, A. Pfitzmann, I. Stierand, "Computer Based Steganography: How it works and why therefore any restrictions on cryptography are nonsense, at best", In: [1] pp. 7-21, 1996.

[8] D. Gruhl, W. Bender, and A. Lu. "Echo Hiding". In [1], pp. 295-315, 1996.

[9] T.G. Handel, M.T. Stanford, III. "Hiding Data in the OSI Network Model", In: [1] pp. 23-38, 1996.

[10] *Proc.IEEE International Conference on Image Processing* (ICIP'97), Santa Barbara, California, October 26-29, IEEE Press, 1997.

[11] N.F. Johnson, Z. Duric, S. Jajodia, "The Role of Digital Watermarking in Electronic Commerce", submitted for publication to *ACM* issue on electronic commerce. To be published.

[12] N.F. Johnson, S. Jajodia, "Exploring Steganography: Seeing the Unseen", *IEEE Computer*, February 1998, vol. 31, no. 2, pp.26-34

[13] N.F. Johnson, S. Jajodia, "Steganalysis of Images Created Using Current Steganography Software", *Proc. Information Hiding Workshop*, Portland, Oregon, USA, April 1998. To be published.

[14] L.M. Marvel, C.G. Boncelet, Jr., C.T. Retter, "Reliable Blind Information Hiding for Images", *Proc. Information Hiding Workshop*, Portland, Oregon, USA, April 1998. To be published.

[15] F. Petitcolas, R. Anderson, M. Kuhn, "Attacks on Copyright Marking Systems", *Proc. Information Hiding Workshop*, Portland, Oregon, USA, April 1998. To be published.

[16] A. Piva, M. Barni, F. Bartolini, V. Cappellini, "DCT-Based Watermark Recovering Without Resorting to the Uncorrupted Original Image", in [10], 1997 pp. 520-523.

[17] G.B. Rhoads, "Method and Apparatus Responsive to a Code Signal Conveyed Through a Graphic Image", U.S. Patent 5,710,834, January 20, 1998. Held by Digimarc Corporation, http://www.digimarc.com

[18] M.D. Swanson, M. Kobayashi, A.H. Tewfik, "Multimedia Data-Embedding and Watermarking Technologies", *Proc.of the IEEE*, vol. 86, no. 6, June, 1998, pp. 1064-1087.

[19] M.D. Swanson, B. Zhu, A.H. Tewfik, "Transparent robust image watermarking", *Proc. IEEE International Conference on Image Processing* (ICIP96), Piscataway, NJ. IEEE Press, vol. III, 1996, pp. 211-214

[20] P. Wayner, *Disappearing Cryptography*, Chestnut Hill, MA: AP Professional, 1996.

Neil F. Johnson may be reached at njohnson@jmu.edu