STEGANOGRAPHY UNDER VARIOUS MEDIA

Er. KUMUD GUPTA, ASSISTANT PROFESSOR,
DEPARTMENT OF COMPUTER SCIENCE AND APPLICTION
D.A.V COLLEGE FOR GIRLS, YNR.

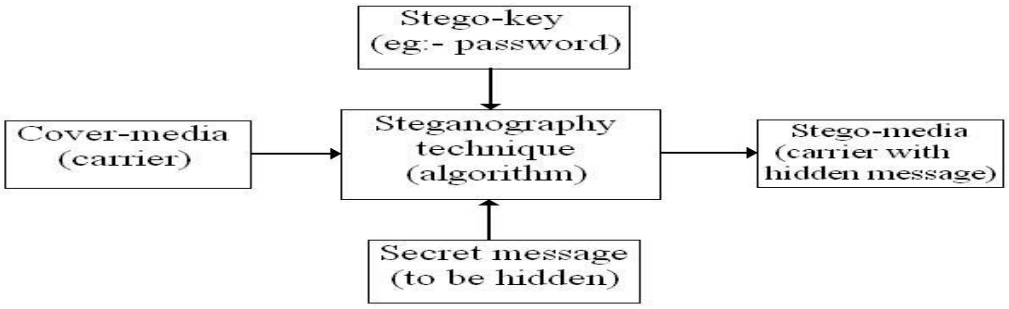
ABSTRACT

Steganography comes from the Greek and literally mean, "Covered or secret writing". Although related to cryptography, they are not the same. Steganography's intent is to hide the existence of the message, while cryptography scrambles a message so that it cannot be understood.

Steganography is one of various data hiding techniques, which aims at transmitting a message on a channel where some other kind of information is already being transmitted. The goal of Steganography is to hide messages inside other "harmless" messages in a way that does not allow any "enemy" to even detect that there is a second secret message present. The only missing information for the "enemy" is the short easily exchangeable random number sequence, the secret key, without the secret key, the "enemy" should not have the slightest chance of even becoming suspicious that on an observed communication channel, hidden communication might take place.

In the field of steganography, some terminology has developed,

The adjectives ***Cover, Embedded*** And **stego** were defined at the Information Hiding Workshop held in Cambridge, England. The term ``cover'' is used to describe the original, innocent message, data, audio, still, video and so on. When referring to audio signal Steganography, the cover signal is sometimes called the ``host'' signal.



The following formula provides a very generic description of the pieces of the steganographic process:

cover_medium + hidden_data + stego_key = stego_me.

Keywords: Steganography, Cryptography, Random numbers logic etc

Introduction

Steganography is the art and science of hiding information in ways that prevent the detection of hidden messages. Steganography, derived from Greek, literally means "Covered writing." It includes a vast array of secret communications methods that conceal the message's very existence. These methods include invisible inks, microdots, character arrangement, digital signatures, covert channels, and spread spectrum communications. Steganography and cryptography are cousins in the spy craft family. Cryptography scrambles a message, so that it cannot be understood. Steganography hides the message, so it cannot be seen. A message in cipher text, for instance, might arouse suspicion on the part of the recipient while an "invisible" message created with steganographic methods will not. Modern steganography's goal is to keep hidden message's mere presence undetectable, but steganographic systems because of their invasive nature, leave behind detectable traces in the cover medium.

Image steganography has been widely studied by researchers. There are a variety of methods used in which information can be hidden in images. Some of them are described here given By :-

Lee et al. [14], Chan et al. [15], Chang et al. [16], and Hsu et al. [17].

In image steganography almost all data hiding techniques try to alter insignificant information in the cover image. For instance, a simple scheme proposed by ***Lee et al. [14],*** is to place the embedding data at the least significant bit (LSB) of each pixel in the cover image.

Other familiar data hiding techniques use the transformation domain of digital media to hide information discussed by ***Chang et al. [16] and Hsu et al. [17].*** Functions such as the discrete cosine transform (DCT) and the discrete wavelet transform (DWT) are widely applied ***by Chang et al. [16], and Hsu et al. [17].*** These methods hide the messages in the significant areas of the cover image, which makes them robust against compression, cropping and other image processing attacks. Signal Processing: ***An International Journal, Volume 1 : Issue (1).***

Steganography under Various Media:

In the following three sections we will try to show how steganography can and is being used through the media of text, images, and audio. Often, although it is not necessary, the hidden messages will be encrypted. This meets a Requirement posed by the ``Kerckhoff principle'' in cryptography. This principle states that the Security of the system has to be based on the assumption that the enemy has full knowledge of the design and implementation details of the steganographic system.

Steganography in Text

The illegal distribution of documents through modern electronic means, such as electronic mail, means such as this allow infringers to make identical copies of documents without paying royalties or revenues to the original author. To counteract this possible wide-scale piracy, a method of marking printable documents with a unique codeword that is Indiscernible to readers, but can be used to identify the intended recipient of a document just by Examination of a recovered document.The techniques they propose are intended to be used in conjunction with standard security measures. For example, documents should still be encrypted prior to transmission across a network. Primarily, their techniques are intended for use after a document has been decrypted, once it is readable to all. Tex. Three features are described in the following subsections:

Line-Shift Coding:

In this method, text lines are vertically shifted to encode the document uniquely. Encoding and decoding can generally be applied either to the format file of a document, or the bitmap of a page image.

Word-Shift Coding:

In word-shift coding, codewords are coded into a document by shifting the horizontal locations of words within text lines, while maintaining a natural spacing appearance. This encoding can also be applied to either the format file or the page image bitmap. The method, of course, is only applicable to documents with variable spacing between adjacent words, such as in documents that have been text-justified. As a result of this variable spacing, it is necessary to have the original image, or to at least know the spacing between words in the unencoded document.The following is a simple example of how word-shifting might work.

Feature Coding:

A third method of coding data into text is known as feature coding. This is applied either to the bitmap image of a document, or to a format file. In feature coding, certain text features are altered, or not altered, depending on the codeword. For example, one could encode bits into text by extending or shortening the upward, vertical endlines of letters such as b, d, h, etc.

Alternative Methods:

Alternative, interesting, major three text-coding methods of encoding data are:

- Open space methods, similar to the ones given
- Syntactic methods that utilize punctuation and contractions
- Semantic methods that encode using manipulation of the words themselves

The syntactic and semantic methods are particularly interesting. In syntactic methods, multiple methods of punctuation are harnessed to encode data. For example, the two phrases below are both considered correct, although the first line has an extra comma:

bread, butter, and milk

bread, butter and milk

Alternation between these two forms of listing can be used to represent binary data.

Steganography in Images

In this section we deal with data encoding in still digital images. In essence, image steganography is about exploiting the limited powers of the human visual system (HVS).

Some Guidelines to Image Steganography:

Before proceeding further, some explanation of image files is necessary. To a computer, an image is an array of numbers that represent light intensities at various points, or pixels. These pixels make up the image's raster data. An image size of 640 by 480 pixels, utilizing 256 colors (8 bits per pixel) is fairly common. Such an image would contain around 300 kilobits of data. Digital images are typically stored in either 24-bit or 8-bit per pixel files. 24-bit images are sometimes known as true color images. Obviously, a 24-bit image provides more space for hiding information; however, 24-bit images are generally large and not that common. A 24-bit image 1024 pixels wide by 768 pixels high would have a size in excess of 2 megabytes. As such large files would attract attention were they to be transmitted across a network or the Internet, image compression is desirable. However, compression brings with it other problems, as will explain shortly. Alternatively, 8-bit color images can be used to hide information. In 8-bit color images, (such as GIF files), each pixel is represented as a single byte. Each pixel merely points to a color index table, or palette, with 256 possible colors. The pixel's value, then, is between 0 and 255.

Image Compression:

Image compression offers a solution to large image files. Two kinds of image compression are lossless and lossy compression. Both methods save storage space but have differing effects on any uncompressed hidden data in the image. Lossy compression, as typified by JPEG (Joint Photographic Experts Group) format files, offers high compression, but may not maintain the original image's integrity. This can impact negatively on any hidden data in the image. This is due to the lossy compression algorithm, which may ``lose'' unnecessary image data, providing a close approximation to high-quality digital images, but not an exact duplicate. Hence, the term ``lossy'' compression

Image Encoding Techniques:

Information can be hidden many different ways in images. Straight message insertion can be done, which will simply encode every bit of information in the image.

The most common approaches to information hiding in images are:

- Least significant bit (LSB) insertion
- Masking and filtering techniques
- Algorithms and transformations

Least Significant bit insertion:

The least significant bit insertion method is probably the most well known image steganography technique. It is a common, simple approach to embedding information in a graphical image file. A simple conversion from a GIF or BMP format to a lossy compression format Such as JPEG can destroy the hidden information in the image. When applying LSB techniques to each byte of a 24-bit image, three bits can be encoded into each pixel. For example, the letter A can be hidden in three pixels. Assume the original three pixels are represented by the three 24-bit words below:

(00100111 11101001 11001000)
(00100111 11001000 11101001)
(11001000 00100111 11101001)

The binary value for the letter A is (101101101). Inserting the binary value of A into the three pixels, starting from the top left byte, would result in:

(00100111 1110100***0*** 1100100***1***)
(00100111 11001000 11101001)
(1100100***1*** 0010011***0*** 11101001)

The emphasized bits are the only bits that actually changed. The main advantage of LSB Insertion is that data can be hidden in the least and second to least bits and still the human eye would be unable to notice it.

Masking and Filtering:

Masking and filtering techniques hide information by marking an image in a manner similar to paper watermarks. Because watermarking techniques are more integrated into the image, they may be applied without fear of image destruction from lossy compression.

Steganography in Audio

Because of the range of the human auditory system (HAS), data hiding in audio signals is especially challenging. The HAS perceives over a range of power greater than one billion to one and range of frequencies greater than one thousand to one. Also, the auditory system is very sensitive to additive random noise. Any disturbances in a sound file can be detected as low as one part in ten million (80dB below ambient level).

Methods of Audio Data Hiding:

Low-bit encoding:

Similarly to how data was stored in the least-significant bit of images, binary data can be stored in the least-significant bit of audio files. Ideally the channel capacity is 1kb per second per kilohertz; so for example, the channel capacity would be 44kbps in a 44kHz sampled sequence. Unfortunately, this introduces audible noise. Of course, the primary disadvantage of this method is its poor immunity to manipulation. Factors such as channel noise and resampling can easily destroy the hidden signal. A particularly robust implementation of such a method results in a slight amplitude modification of each sample in a way that does not produce any perceptual difference. Their implementation offers high robustness to MPEG compression plus other forms of signal manipulation, such as filtering, resampling and requantization.

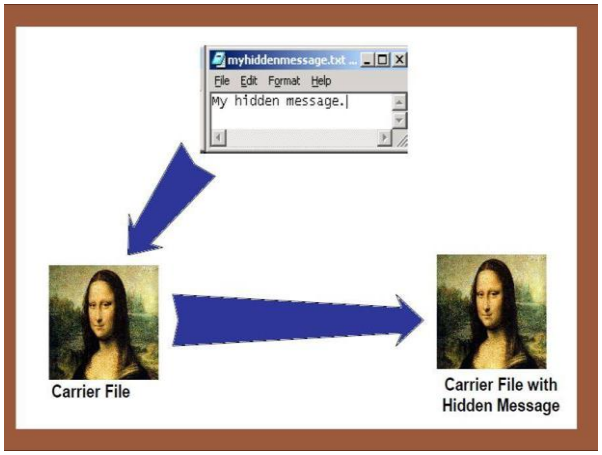
Phase coding:

The procedure for phase coding is as follows:

- The original sound sequence is broken into a series of N short segments.
- A discrete Fourier transform (DFT) is applied to each segment, to break create a matrix of the phase and magnitude.
- The phase difference between each adjacent segment is calculated.
- For segment S0, the first segment, an artificial absolute phase p0 is created.
- For all other segments, new phase frames are created.
- The new phase and original magnitude are combined to get a new segment, Sn.

- Finally, the new segments are concatenated to create the encoded output.

Echo data hiding: Echo data hiding embeds data into a host signal by introducing an echo. The data are hidden by varying three parameters of the echo: initial amplitude, decay rate, and offset, or delay. As the offset between the original and the echo decreases, the two signals blend. At a certain point, the human ear cannot distinguish between the two signals, and the echo is merely heard as added resonance. This point depends on factors such as the quality of the original recording, the type of sound, and the listener.



By using two different delay times, both below the human ear's perceptual level, we can encode a binary one or zero.

IMAGE 1:

IMAGE SHOWING HIDDEN MESSAGE BEHIND A PICTURE.

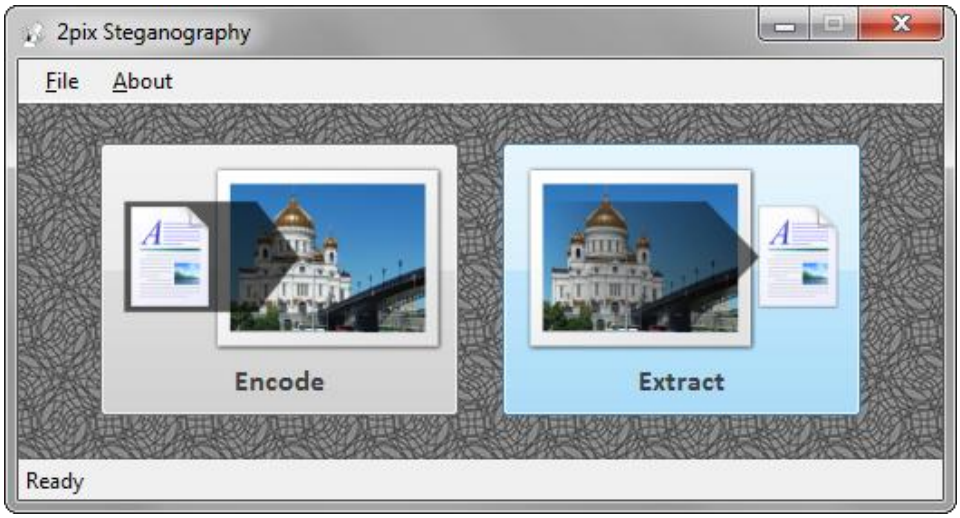


IMAGE 2.

SHOWING THE ENCODED IMAGE AND DECODED IMAGE.

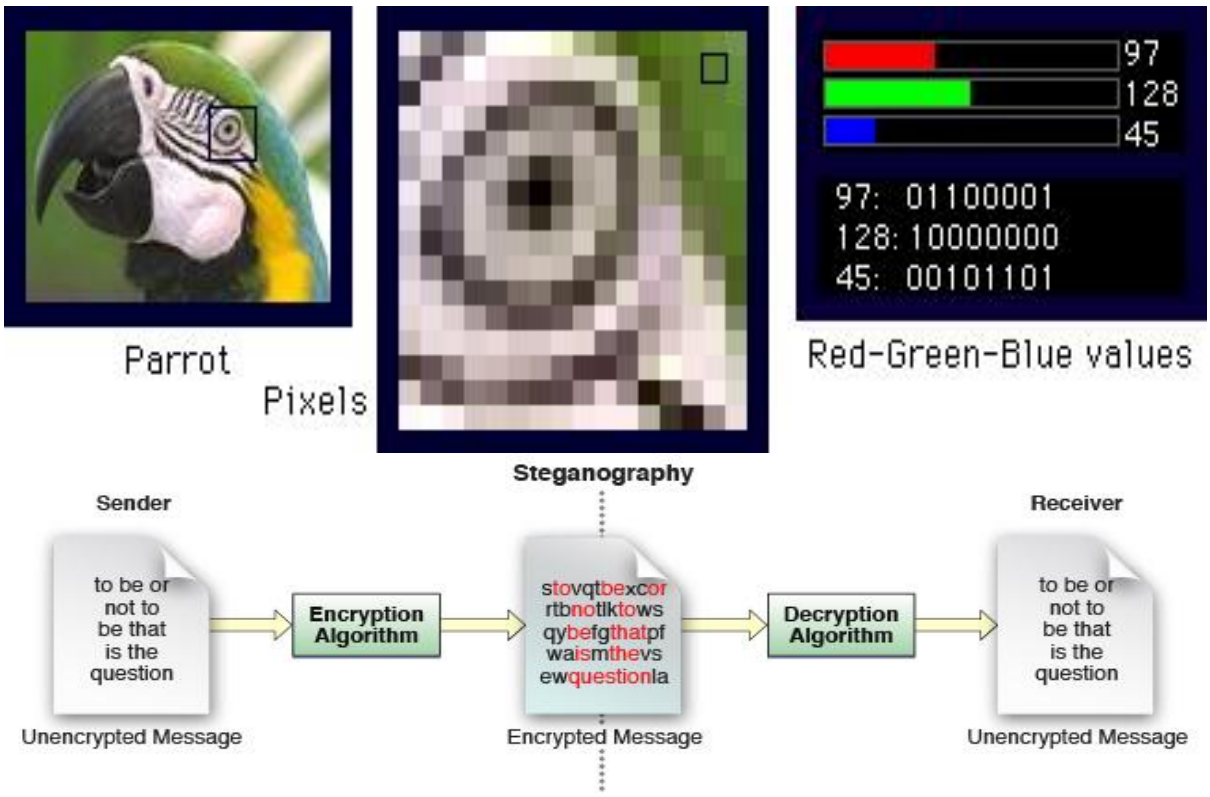


IMAGE 3 AND IMAGE 4:CODED IMAGE

Steganalysis

Whereas the goal of steganography is the avoidance of suspicion to hidden messages in other data, steganalysis aims to discover and render useless such covert messages.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. Here two methods are looked into: 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.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.

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. 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 the 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.

Steganography and Steganalysis Tools

Steganography Tools:

1. MP3Stego

MP3Stego will hide information in MP3 files during the compression process. The data is first compressed, encrypted and then hidden in the MP3 bit stream.

METHOD/IMAGE 1:

- http://www.petitcolas.net/fabien/software/MP3Stego_1_1_17.zip

2. JPHide and JPSeek

JPHIDE and JPSEEK are programs which allow you to hide a file in a jpeg visual image. There are lots of versions of similar programs available on the internet but JPHIDE and JPSEEK are rather special.

METHOD/IMAGE 2

- http://www.snapfiles.com/php/download.php?id=101911

3. GIFShuffle

The program **gifshuffle** is used to conceal messages in GIF images by shuffling the colour map, which leaves the image visibly unchanged. **gifshuffle** works with all GIF images, including those with transparency and animation, and in addition provides compression and encryption of the concealed message.

METHOD/IMAGE 3:

- http://www.darkside.com.au/gifshuffle/

WbStego

WbStego is a tool that hides any type of file in bitmap images, text files, HTML f
Adobe PDF files.

METHOD/IMAGE 4

http://www.wbailer.com/wbstego

StegoVideo

MSU StegoVideo allows to hide any file in a video sequence. When the program was created, different popular codec's were analyzed and an algorithm was chosen which provides small data loss after video compression. You can use MSU StegoVideo as VirtualDub filter or as standalone .exe program, independent from VirtualDub.

http://compression.ru/video/stego_video/index_en.html

Steganalysis Tools

Steganography Analyzer Artifact Scanner (StegAlyzerAS)

StegAlyzerAS gives you the capability to scan the entire file system, or individual directories, on suspect media for the presence of Steganography application artifacts. And, unlike other popular forensic tools, you can 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.

Steganography Analyzer Signature Scanner (StegAlyzerSS)

StegAlyzerSS gives you the capability to scan every file on the suspect media for the presence of hexadecimal byte patterns, or signatures, of particular Steganography applications in the files.

Digital Invisible Ink Toolkit

This project provides a simple Java-based steganography tool that can hide a message inside a 24-bit color image so that knowing how it was embedded, or performing statistical analysis, does not make it any easier to find the concealed information.

Conclusion

This paper provides 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.

Bibliography:

- Wikipedia contributors.Steganography [Internet].Wikipedia, The free encyclopedia; 2010 Feb 16.
- URL: http://zone-h.org
- N.provos & P.Honeyman, “Hide & Seek: An Introduction to Steganography,” IEEE Security and privacy.
- D.Artz , “Digital Steganography : Hiding Data within Data,” IEEE Security and Privacy
- “ The Science of secrecy ;Steganography”
    - URL: www.channel4.com/plus/secrecy/page1b.html
- *Andersen R.J. and Petitcolas* F.A.P. *"On the limits of steganography".* IEEE Journal of Selected Areas in Communications, Special Issue on Copyright and Privacy Protection 16 No.4, 474-481, 1998
- *Westfeld A. and Pfitzmann A.* "Attacks on Steganographic Systems". Lecture Notes in Computer Science, vol. 1768, Springer-Verlag, Berlin, pp. 61-75, 2000.
- Steganography Software”. Lecture Notes in Computer Science, vol.1525, Springer Verlag, Berlin, pp. 273-289, 1998

- An International Journal, Volume 1 : Issue (1)

- Chan Y. K. and Chang C. C. “Concealing a Secret Image Using the Breadth First Traversal Linear Quad tree Structure”. IEEE Proceedings of Third International on Cooperative Database Systems for Advance Applications pp. 194-199,2001 Lee, Y. K. and Chen L. H. “High Capacity Image Steganographic Model”. IEEE
- Lee, Y. K. and Chen L. H. “High Capacity Image Steganographic Model”. IEEE Proceedings Vision, Image and Signal Processing, pp. 288-294, 2015.

- Kelton W. D. and Law A. “Simulation Modeling and Analysis”. Mcgraw-Hill Science, USA,2006

- Koval Oleksiy, Voloshynovskiy Svyatoslav, Holotyak Taras and Pun Thierry. “Information-theoretic Analysis of Steganalysis in Real Images”. International Multimedia Conference Proceeding of the 8th workshop on Multimedia and security, Geneva, Switzerland,Pages: 11 – 16, 2014