ISSN:2320-0790

An International Journal of Advanced Computer Technology

A Comparative Study of Various image Steganographic Techniques Used for Information Hiding

SAMIKSHA K. JICHKAR, MAHIP M. BARTERE

G.H.Raisoni College Of Engineering & Management, Amravati, SGBAU University, Amravati, Maharashtra 444602, India

Abstract: The staggering growth in communication technology and usage of public domain channels (i.e. Internet) has greatly facilitated transfer data. However, such open communication channels have greater vulnerability to security threats causing unauthorized information access. Traditionally, encryption is used to realize the communication security. Today HTML pages and Spam Emails are also being used for steganography. With the help of steganography we can hide the information which is paramount and confidential for us. Steganography is the art of hiding the fact that communication is taking place, by hiding information in other information. Many different carrier file formats can be used. For hiding secret information in images, there exist a large variety of steganographic techniques some are more complex than other and all of them have respective strong and weak points. Steganography is a technology where modern data compression, information theory, spread spectrum, any cryptography technologies are brought together to satisfy the need for privacy on the internet. This paper intends to give through understanding and evolution of different existing image steganographic techniques of data hiding in spatial, transform and compression domains. It also discusses that which image format is most appropriate or best for our algorithm and how can we perform compression on that.

***Keywords*:** Steganography, image steganography, spatial domain, Transform domain.

1. Introduction

Steganography is a technique to hide data inside a cover medium in such a way that the existence of any communication itself is undetectable as opposed to cryptography where the existence of secret communication is known to everyone but is indecipherable. In this modern era, internet offers great convenience in transmitting large amounts of data in different parts of the world. Steganography is the science that communicates secret data in an appropriate multimedia carrier, e.g., image, audio, and video files. Steganography can be used to exchange secret information in an undetectable way over a public communication channel, whereas watermarking can be used for copyright protection and tracking legitimate use of a particular software or media file.

Image files are the most common cover medium used for steganography. One of the most common and naive methods of embedding message bits is LSB replacement in spatial domain where the bits are encoded in the cover image by replacing the least significant bits of pixels. JPEG is the most popular image format used over the Internet and by image acquisition devices, and therefore we use JPEG as our choice for steganography. There are many real life applications of steganography. Apparently, during the 1980's, Margaret Thatcher became so irritated at press leaks of cabinet documents that she had the word processors programmed to encode their identity in the word spacing, so that disloyal ministers could be traced. Similar

techniques are now undergoing trials in an electronic publishing project, with a view to hiding copyright messages and serial numbers in documents.

Essentially, the information-hiding process in a steganographic system starts by identifying a cover medium's redundant bits (those that can be modified without destroying that medium's integrity).The embedding process creates a *stego medium* by replacing these redundant bits with data from the hidden message.

In the steganography system scenario, before the hiding process, the sender must select the appropriate message carrier (i.e. image, audio, video, text) and select the effective secret message as well as the robust password (which supposed to be known by the receiver). The effective and appropriate Steganography algorithm must be selected that able to encode the message in more secure technique. Then the sender may send the Stego file by email or chatting, or by other modern techniques. The Stego file is the carried message with the secret information. After receiving the message by the receiver, he can decode it using the extracting algorithm and the same password used by the sender. The Steganography system scenario is shown in the Figure 1.

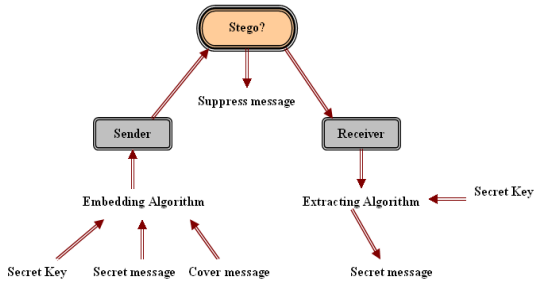

Fig.1 Steganography System

IMAGE

1. Image steganography:

As stated earlier, images are the most popular cover objects used for steganography. In the domain of digital images many different image file formats exist, most of them for specific applications. For these different image file formats, different steganographic algorithms exist. Image compression techniques are extensively used in steganography. Among the two types of image compressions, lossy compression and loss less compression; lossless compression formats offer more promises. Lossy compression compression may not maintain the original image's integrity. Lossless compression maintains the original image data exactly, hence it is preferred. Example of Lossy compression format is JPEG format files. Examples of Lossless compression formats are GIF and BMP formats.

1.1 Image Definition:

To a computer, an image is a collection of numbers that constitute different light intensities in different areas of the image. This numeric representation forms a grid and the individual points are referred to as pixels. Most images on the Internet consists of a rectangular map of the image's pixels (represented as bits) where each pixel is located and its color. These pixels are displayed horizontally row by row.

1.2 Image Domain:

Image steganography techniques can be divided into two groups: those in the Image Domain and those in the Transform Domain. Image – also known as spatial – domain techniques embed messages in the intensity of the pixels directly, while for transform – also known as frequency – domain, images are first transformed and then the message is embedded in the image. Image domain techniques encompass bit-wise methods that apply bit insertion and noise manipulation and are sometimes characterized as "simple systems". The image formats that are most suitable for image domain steganography are lossless and the techniques are typically dependent on the image format.

1.2.1 Least Significant Bit:

In this technique, the bits of the secret data are embedded in the least significant bit of certain bytes of the cover image. Secrecy can be implemented by taking secret key into consideration that will specify the bytes that need to be manipulated. It usually considers BMP files as they use lossless data compression. Least significant bit (LSB) insertion is a common, simple approach to embedding information in a cover image. Least significant bit (LSB) insertion is a common, simple approach to embedding information in a cover image. The least significant bit (in other words, the 8th bit) of some or all of the bytes inside an image is changed to a bit of the secret message. When using a 24-bit image, a bit of each of the red, green and blue color components can be used, since they are each represented by a byte. In other words, one can store 3 bits in each pixel. An 800 × 600 pixel image, can thus store a total amount of 1,440,000 bits or 180,000 bytes of embedded data. For example a grid for 3 pixels of a 24-bit image can be as follows:

(00101101 00011100 11011100)
(10100110 11000100 00001100)
(11010010 10101101 01100011)

When the number 200, which binary representation is 11001000, is embedded into the least significant bits of this part of the image, the resulting grid is as follows:

(00101101 00011101 11011100)
(10100110 11000101 00001100)
(11010010 10101100 01100011)

Although the number was embedded into the first 8 bytes of the grid, only the 3 underlined bits needed to be changed according to the embedded message. On average, only half of the bits in an image will need to be modified to hide a secret message using the maximum cover size. Since there are 256 possible intensities of each primary color, changing the LSB of a pixel results in small changes in the intensity of the colors. These changes cannot be perceived by the human eye - thus the message is successfully hidden. With a well-chosen image, one can even hide the message in the least as well as second to least significant bit and still not see the difference.

In the above example, consecutive bytes of the image data – from the first byte to the end of the message – are used to embed the information. This approach is very easy to detect. A slightly more secure system is for the sender and receiver to share a secret key that specifies only certain pixels to be changed. Should an adversary suspect that LSB steganography has been used, he has no way of knowing which pixels to target without the secret key.

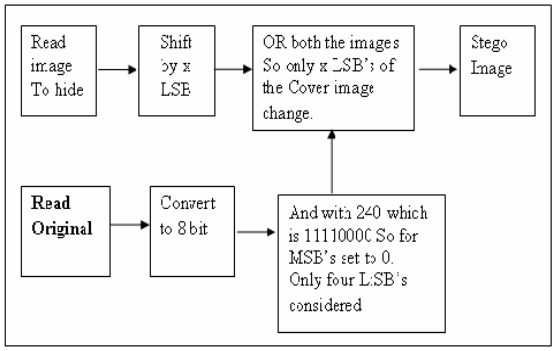

Fig.2 Block Diagram for implemented Logic of LSB Embedding

1.2.2 The LSB Algorithm

1) Select *cover-object* (GIF/BMP) *c* as an input.
2) Encode the *c* in binary.
3) The Secret Message, *m.*
4) Encode the *m* in binary.
5) Choose one pixel of the *c* randomly.
6) Use a pixel selection to hide information in the *c*.
7) Save the new image (*Stego-object*) s.

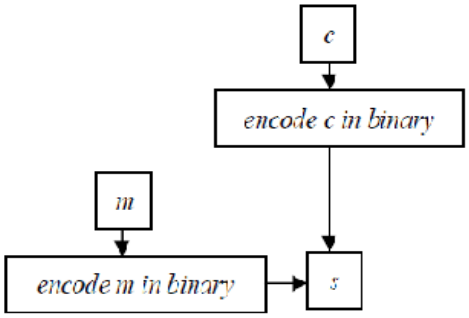

Fig.3 The LSB Algorithm

A. LSB in BMP

The BMP file format also called bitmap or DIB file format (for *device-independent bitmap*), is an image file format used to store bitmap digital images. Since BMP is not widely used the suspicion might arise, if it is transmitted with an LSB stego. When image are used as the carrier in Steganography they are generally manipulated by changing one or more of the bits of the byte or bytes that make up the pixels of an image. The message can be stored in the LSB of one color of the RGB value or in the parity bit of the entire RGB value. A BMP is capable of hiding quite a large message. LSB in BMP is most suitable for applications, where the focus is on the amount of information to be transmitted and not on the secrecy of that information. If more number of bits is altered, it may result in a larger possibility that the altered bits can be seen with the human eye. But with the LSB the main objective of Steganography is to pass a message to a receiver without an intruder even knowing that a message is being passed is being achieved.

A.1 General BMP format properties

1) A bitmap format that can be uncompressed, or compressed with RLE
2) BMP files are.
   - In 1-bit black and white.
   - 8-bit greyscale.
   - 16-, 24- or 32-bit RGB color.
   - Or 4- or 8-bit indexed color.
3) BMP files don't support CMYK color.
4) Transparency is supported for individual pixels as in GIF files.
5) Alpha channels are supported in new versions of BMP.

B. LSB in GIF

Graphics interchange format also known as GIF is one of the machine independent compressed formats for storing images. Since GIF images only have a bit depth of 8, amount of information that can be hidden is less than with BMP. Embedding information in GIF images using LSB results in almost the same results as those of using LSB with BMP. LSB in GIF is a very efficient algorithm to use when embedding a reasonable amount of data in a grayscale image. GIF images are indexed images where the colors used in the image are stored in a palette. It is sometimes referred to as a color lookup table. Each pixel is represented as a single byte and the pixel data is an index to the color palette. The colors of the palette are typically ordered from the most used color to the least used colors to reduce lookup time. Some extra care is to be taken if the GIF images are to be used for Steganography. This is because of the problem with the palette approach. If the LSB of a GIF

image is changed using the palette approach, it may result in a completely different color. This is because the index to the color palette is changed. The change in the resulting image is noticeable if the adjacent palette entries are not similar. But the change is not noticeable if the adjacent palette entries are similar. Most applications that use LSB methods on GIF images have low security because it is possible to detect even moderate change in the image. Solutions to these problems could be

1. Sort the palette so that the color difference between consecutive colors is minimized
2. Add new colors, which are visually similar to the existing colors in the palette.
3. Use Gray scale images. In a 8 bit Gray scale GIF image, there are 256 shades of gray. This results in gradual changes in the colors and it is hard to detect.

B.1 General GIF format properties

1) Can be compressed to a small size**.**
2) ·Are commonly used for images presented on the web.
3) GIF files allow only 8-bit indexed color.
4) GIF files use lossless LZW compression.
5) GIF files support transparency.
6) Animated GIF files can be created by sequences of single images.
7) GIF files can be saved in an interlaced format that allows progressive download of web images (low-resolution version of an image first then gradually comes into focus the rest of the data is downloaded**.**

C. **JPEG**

The term actually stands for "Joint Photographic Experts Group," because that is the name of the committee that developed the format. But you don't have to remember that because even computer nerds will think you're weird if you mention what JPEG stands for. Instead, remember that a JPEG is a compressed image file format. JPEG images are not limited to a certain amount of color, like GIF images are. Therefore, the JPEG format is best for compressing photographic images. So if you see a large, colorful image on the Web, it is most likely a JPEG file. While JPEG images can contain colorful, high-resolution image data, it is a lossy format, which means some quality is lost when the image is compressed. If the image is compressed too much, the graphics become noticeably "blocky" and some of the detail is lost. Like GIFs, JPEGs are cross platform, meaning the same file will look the same on both a Mac and PC.

C.1 General JPEG format properties

1) Are commonly used for photo.
2) Can be compressed to a smaller size.
3) JPEG files allow only 8 - 24-bit indexed color.
4) JPEG files use lossy compression.

2. Example:

A. *The original image (before hiding)*

Figure 4. An image before embedding.

B. *The image after hiding*

Figure 5. The image after embedding.

Table 1: Comparison of GIF, BMP & JPEG Images.

	BMP	GIF	JPEG
File Type	Windows Bitmap	Graphical Interchange Format	Joint Photographic Expert Group
File Suffix	.BMP	.GIF	.JPEG,.JPG
File Size	Larger	Larger	Small
Resolution	Medium		High
Support Color	RGB	Grayscale	16 million color
Complexity	Very Simplistic	-	Quite Complex
Ideal For	Icons & Small images	-	Photo
Color Depth	1-32 bit color	8 bit color	8-24 bit color
Compression Algorithm	Lossless	Lossless	Lossy

Table 2: Comparison of LSB for GIF, BMP & JPEG Images

	BMP	GIF	JPEG
Efficient when amount of data reasonable	High	Medium	-
Percentage Distortion less resultant image	High	Medium	-
Steganalysis detection	Low	Low	-
Amount of embedded data	High	Low	-
Robustness against image manipulation	Low	Low	Medium
Invisibility	High	Medium	High
Robustness against statistical attacks	Low	Low	Medium
Independent of file format	Low	Low	Low
Payload capacity	High	Medium	Medium
Unsuspicious files	Low	Low	High

"High = 2, Medium = 1 and Low = 0"

Application:

There are many applications for digital steganography of images, including copyright protection, feature tagging, and secret communications. Image Steganography has many applications, especially in today's modern, high-tech World. Privacy and anonymity is a concern for most people on the internet. Image Steganography allows for two parties to communicate secretly and covertly. It allows for some morally-conscious people to safely whistle blow on internal actions; it allows for copyright protection on digital files using the message as a digital watermark. One of the other main uses for Image Steganography is for the transportation of high-level or top-secret documents between international governments. While Image Steganography has many legitimate uses, it can also be quite nefarious. It can be used by hackers to send viruses and Trojans to compromise machines, and also by terrorists and other organizations that rely on covert operations to communicate secretly and safely.

Conclusion:

Steganography is an effective way to obscure data and hide sensitive information. It allows an individual to hide data inside other data with hopes that the transfer medium will be so obscure that no one would ever think to examine the contents of the file. use of BMP image. To be able to hide a secret message inside a BMP file, one would require a very large cover image. BMP images of 800×600 pixels found to have less web applications. Moreover such uses are not accepted as valid. For this reason, LSB Steganography has also been developed for use with other image file formats. In the images of the kind BMP, data type that can be embedded is large and the image is not distorted because of the ability of this kind of images for carrying amount of data without notice. For the image of kind JPEG we find very medium data embedded, as if resistance to statistical attacks, robustness against image manipulation is low, and when we increase the amount of data the image becomes distorted and is subject to discovery.

References:

[1] Moerland, T., "Steganography and Steganalysis", Leiden Institute of Advanced Computing Science, www.liacs.nl/home/tmoerl/privtech.pdf.

[2] Artz, D., "Digital Steganography: Hiding Data within Data", IEEE Internet Computing Journal, June 2001.

[3] Eltyeb E.Abed Elgabar, Haysam A. Ali Alamin, "Comparison of LSB

Steganography in GIF and BMP Images ", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-4, September 2013.

[4] Artz, D., "Digital Steganography: Hiding Data within Data", IEEE Internet Computing Journal, June 2001.

[5] Atallah M. Al-Shatnawi, "A New Method in Image Steganography with Improved Image Quality", Applied Mathematical Sciences, Vol. 6, 2012, no. 79, 3907 - 391.

[6] Priya Thomas," Literature Survey On Modern Image Steganographic Techniques", International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 5, May - 2013 ISSN: 2278-0181.

[7] Y. Lee and L. Chen, "High capacity image steganographic model," IEE Proceedings-Vision, Image and Signal Processing, vol. 147, no. 3, pp. 288–294, 2000.

[8] W. Pennebaker and J. Mitchell, JPEG still image data compression standard. Kluwer Academic Publishers, 1993.

[9] Jiri Fridrich and Du Rui. Secure steganographic methods for palette images. In Inter'l Workshop on Information Hiding, pages 47–60, 1999.

[10] Provos, N., Honeyman, P, *Hide and seek: An introduction to steganography*, IEEE Security & Privacy Magazine 1 (2003) pp. 32-44

[11] E. Franz and A. Schneidewind., "Adaptive steganography based on dithering", in Proc. Of the 2004 workshop on Multimedia and Security, 2004. pp. 56-62.

[12] Yang *et al.*, "Adaptive data hiding in edge areas of images with spatial LSB domain systems", IEEE transactions on information forensics and security, vol. 3, no. 3, september 2008.

[13] I.-C. Lin et al, "Hiding data in spatial domain images with distortion tolerance", Computer Standards & Interfaces 31 (2009) 458–464.