Steganography Methods and some application (The hidden Secret data in Image)

Sadoon Hussein Abdullah

Biology Depart , Science Collage , University of Mosul , Mosul , Iraq

(Received 1 / 6 / 2008 , Accepted 12 / 4 / 2009)

Abstract

This paper will take an in-depth look at Steganography technique, and brief history of Steganography and take one of this technique (least significant techniques (LSB)) to implement encoding secret data in images JPG kind by using MATLAP version 7. The paper will close by looking at how we can use Steganography in an open-systems environment such as the Internet, as well as some of the tools and resources available to help us accomplish this.

Introduction

What is Steganography ?

The word steganography is of Greek origin and means "covered, or hidden_writing" Steganography is defined by Markus Kahn [2] as follows, "Steganography is the art and science of communicating in a way which hides the existence of the communication. In contrast to Cryptography, where the enemy is allowed to detect, intercept and modify messages without being able to violate certain security premises guaranteed by a cryptosystem, 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 message present".

Steganographic technologies are a very important part of the future of Internet security and privacy on open systems such as the Internet.

A Detailed Look at Steganography

In this section we will discuss Steganography at length. We will start by looking at the different types of Steganography generally used in practice today along with some of the other principles that are used in Steganography. We will then look at some of the Steganographic techniques in use today.

To start, In hiding process the following concept may be observed [3] :

Cover-object: refers to the object used as the carrier to embed messages into. Many different objects have been employed to embed messages for example images, audio, and video as well as file structures.

Stego-object: refers to the object, which is carrying a hidden message. So given a cover object, and a message the goal of the steganographer is to produce a stego object which would carry the message. The studying of communications security includes not just encryption but also traffic security, A steganographic message (the *plaintext*) is often first encrypted by some traditional means, producing a *cipher text*. Then, a *cover text* is modified in some way to contain the cipher text, resulting in *stagiest*. For example, the letter size, spacing, *typeface* or other characteristics of a covertext can be manipulated to carry the hidden message; only the recipient (who must know the technique used) can recover the message and then decrypt it.[4].

Lets look at what a theoretically perfect secret communication (Steganography) would consist of. To illustrate this concept, we will use three fictitious characters named Alice, Bop and Wendy. As shown in figure -1- .

Alice wishing to send a secret message *m* to Bob. In order to do so, she “embeds” *m* into a *cover-object c*, to obtain the *stego-object s*. The stego-object *s* is then sent through the public channel. Wendy who examines all messages in the channel, should not be able to distinguish in any sense between cover-objects (objects not containing any secret message) and stego-objects (objects containing a secret message). In this context, *steganalysis* refers to the body of techniques that aid Wendy in distinguishing between cover-objects and stego-objects. It should be noted that Wendy has to make this distinction without any knowledge of the secret key which Alice and Bob may be sharing and sometimes even without any knowledge of the specific algorithm that they might be using for embedding the secret message.[9].

As Fabien A.P. Petitcolas [7] points out, "in a 'perfect' system, a normal cover should not be distinguishable from a stego-object, neither by a human nor by a computer looking for statistical patterns." In practice, however, this is not always the case. In order to embed secret data into a cover message, the cover must contain a sufficient amount of redundant data or noise. This is because the embedding process.

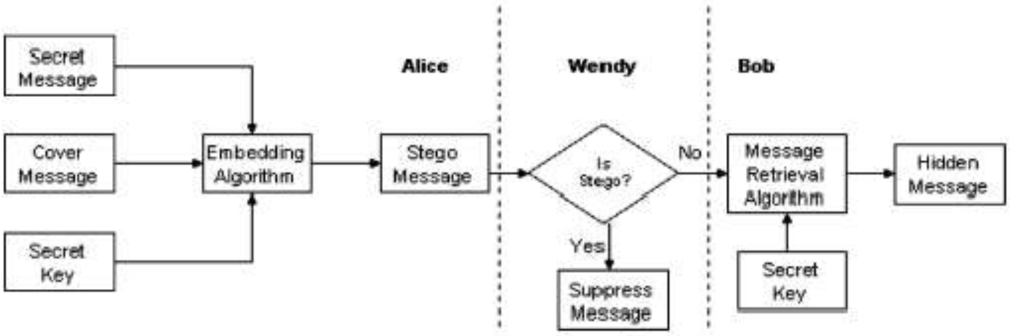


Figure 1. Framework for Secret Key Passive Warden Steganography. Alice embeds secret message in cover image (*left*). Wendy the warden checks if Alice's image is a stego-image (*center*). If she cannot determine it to be so, she passes it on to Bob who retrieves the hidden message based on secret key (*right*) he shares with Alice.

Steganography used in electronic communication include steganographic coding inside of a transport layer, such as an file, or a protocol, such as UDP. Usually, files meant for internet means are put into media types that are lossless, such as FLAC, WAV, and BMP and other file types.[4]

In practice, there are basically three types of steganographic protocols used.[2] Pure Steganography, Secret Key Steganography and Public Key Steganography.

(1)- Pure_Steganography is defined as a steganographic system that does not require the exchange of acipher such as a stego-key. This method of Steganography is the least secure means by which to communicate secretly because the sender and receiver can rely only upon the presumption that no other parties are aware of this secret message. Using open systems such as the Internet, we know this is not the case at all.

(2)- Secret Key Steganography is defined as a steganographic system that requires the exchange of a secret key (stego-key) prior to communication. Secret Key Steganography takes a cover message and embeds the secret message inside of it by using a_secret key (stego-key). Only the parties who know the secret key. can reverse the process and read the secret message. Unlike Pure Steganography where a perceived invisible communication channel is present, Secret Key Steganography exchanges a stego-key, which makes it more susceptible to interception. The benefit to Secret Key Steganography is even if it is intercepted, only parties who know the secret key can extract the secret message.

(3)- Public Key Steganography takes the concepts from Public Key Cryptography as explained below. Public Key Steganography is defined as a steganographic public key during the encoding process and only the private key, which has a direct mathematical relationship with the public key, can decipher the secret message. Public Key Steganography provides a more robust way of implementing a steganographic system because it can utilize a much more robust and researched technology in Public Key Cryptography. It also has multiple levels of security in that unwanted parties must first suspect the use of steganography and then they would have to find a way to crack the algorithm used by the public key system before they could intercept the secret message.

Methods of Encoding the Messages by Using Steganographic Techniques :

1. Encoding Secret Messages in Text

Encoding secret messages in text can be a very challenging task. This is because text files have a very small amount of redundant data to replace with a secret message. Another drawback is the ease of which text based Steganography can be altered by an unwanted parties by just changing the text itself or reformatting the text to some other form (from .TXT to .PDF,etc.). There are numerous methods by which to accomplish text based Steganography. I will introduce a few of the more popular encoding methods below.

A- Line-shift encoding involves actually shifting each line of text vertically up or down by as little as 3 centimeters. Depending on whether the line was up or from the stationary line would equate to a value that would or could be
down encoded into a secret message.

B- Word-shift encoding works in much the same way that line-shift encoding works, only we use the horizontal spaces between words to equate a value for the hidden message. This method of encoding is less visible than line-shift encoding but requires that the text format support variable spacing.

C- Feature specific encoding involves encoding secret messages into formatted text by changing certain text attributes such as vertical/horizontal length of letters such as b,d, T, etc.This is by far the hardest text encoding method to intercept as each type of formatted text has a large amount of features that can be used for encoding the secret message.

All three of these text based encoding methods require either the original file or the knowledge of the original files formatting to be able to decode the secret message.

2-Encoding Secret Messages in Audio

Encoding secret messages in audio is the most challenging technique to use when dealing with Steganography. We will now look at three of the more popular encoding methods for hiding data inside of audio. They are low-bit encoding, phase-coding and spread spectrum.

(A)- Low-bit encoding embeds secret data into the least significant bit (LSB) of the audio file. The channel capacity is 1KB per second per kilohertz (44 kbps for a44 KHz sampled sequence). This method is easy to incorporate but is very susceptible to data loss due to channel noise and resampling.

(B)- Phase coding substitutes the phase of an initial audio segment with a reference phase that represents the hidden data. This can be thought of, as sort of an encryption for the audio signal by using what is known as Discrete Fourier Transform (DFT), which is nothing more than a transformation algorithm for the audio signal.

(C)- Spread spectrum encodes the audio over almost the entire frequency spectrum. It then transmits the audio over different frequencies which will vary depending on what spread spectrum method is used.

3- Encoding Secret Messages in Images

Coding secret messages in digital images is by far the most widely used of all methods in the digital world of today. This is because it can take advantage of the limited power of the human visual system (HVS). Almost any plaintext, cipher text, image and any other media that can be encoded into a bit stream can be hidden in a digital image. With the continued growth of strong graphics power in computers and the research being put into image based Steganography, this field will continue to grow at a very rapid pace.

Before diving into coding techniques for digital images, a brief explanation of digital image architecture and digital image compression techniques should be explained.

As Duncan Sellars [8] explains "To a computer, an image is an array of numbers that represent light intensities at various points, or pixels. These pixels make up the images raster data." When dealing with digital images for use with Steganography, 8-bit and 24-bit per pixel image files are typical. Both have advantages and disadvantages, as we will explain below.

8-bit images are a great format to use because of their relatively small size. The drawback is that only 256 possible colors can be used which can be a potential problem during encoding. Usually a gray scale color palette is used when dealing with 8-bit images such as (.GIF) because its gradual change in color will be harder to detect after the image has been encoded with the secret message. 24-bit images offer much more flexibility when used for Steganography. The large numbers of colors (over 16 million) that can be used go well beyond the human visual system (HVS), which makes it very hard to detect once a secret message, has been encoded. The other benefit is that a much larger amount of hidden data can be encoded into a 24-bit digital image as opposed to an 8-bit digital image. The one major drawback to 24-bit digital images is their large size (usually in MB) makes them more suspect than the much smaller 8-bit digital images (usually in KB) when sent over an open system such as the Internet.

Digital image compression:

Digital image compression is a good solution to large digital images such as the 24-bit images mentioned earlier. There are two types of compression used in digital images, lossy and lossless. Lossy compression such as (.JPEG) greatly reduces the size of a digital image by removing excess image data and calculating a close approximation of the original image. Lossy compression such as wavelet compression is usually used with 24-bit digital images to reduce its size, but it does carryone major drawback. Lossy compression techniques increase the possibility that the uncompressed secret message will lose parts of its contents because of the fact that lossy compression removes what it sees as excess image data. Lossless compression techniques, as the name suggests, keeps the original digital image in tact without the chance of loss. It is for this reason that it is the compression tech nique of choice for steganographic uses. Examples of lossless compression techniques are (.GIF and .BMP). The only drawback to lossless image compression is that it doesn't do a very good job at compressing the size of the image data.[2]

As follows image steganography algorithm classified into four different parts:

A. LSB

A simple way of steganography is based on modifying the least significant bit layer of images, known as the *LSB technique*. In the LSB technique, the least significant bits of the pixels is replaced by the message which bits are permuted before embedding. In some cases (Fridrich et al. [9]) LSB of pixels visited in random or in certain areas of image and sometimes increment or decrement the pixel value.

A simplified example with a 24-bit image 1 pixel:
(00100111 11101001 11001000)
Insert 101:
(00100111 11101000 11001001)
red green blue

B. DCT

DCT (discrete cosines transform) is used in JPEG compression. Embedding in DCT domain is simply done by altering the DCT coefficients, for example by changing the least significant bit of each coefficient [3].One of the limitation in DCT domain happened when 64 coefficients are equal to zero. Values will have an effect on the compression rate. So the number of bit one could embed in DCT domain is less that the number of bits one could embed by the LSB method. Also embedding capacity becomes dependent on the image type used in the case of DCT embedding. . There are different methods for altering the DCT coefficients that reviews of jsteg please refer to [3].

C. Frequency Domain

Another transform domain for embedding is frequency domain. They first decor relate the image by scrambling the pixels randomly, which in effect whitens the frequency domain of the image and increases the number of transform coefficients in the frequency domain thus increasing the embedding capacity [3]. Note the result is a salt and pepper image.

D.Wavelet Transform

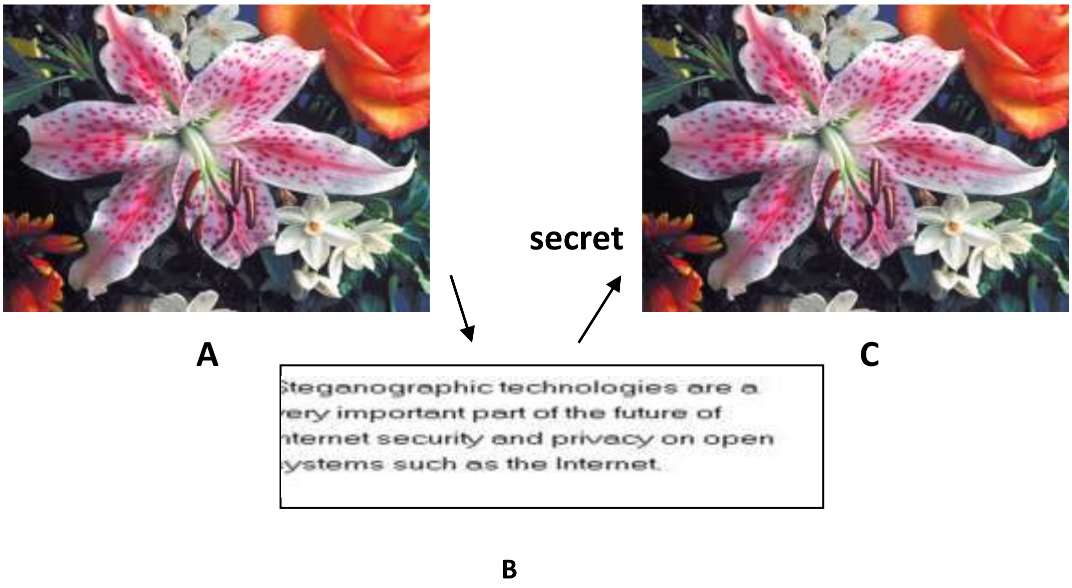
The wavelet transform is a transformation to basis functions that are localized in frequency. The wavelet compression methods are better at representing transients, such as an image of stars on a night sky. This means that "elements of some data signal that are transient can be represented by a smaller amount of information than would be the case if some other transform, such as the more widespread discrete cosine transform, had been used" [3]. Wavelet compressions are good for transient signal characteristics but not for smooth, periodic signals.

The steps are to take the DCT or wavelet transform of the cover image and find the coefficients below a specific threshold. Replace these bits with bits to be hidden (for example, use LSB insertion) and then take the inverse transform and store it as a regular image.

Application Software to implement embed secret data in image by using LSB technique

After explain the steganography and its method we will take an one example program hidden data in image kind JPG by using MATLAB 7,the result of program in figure (2).

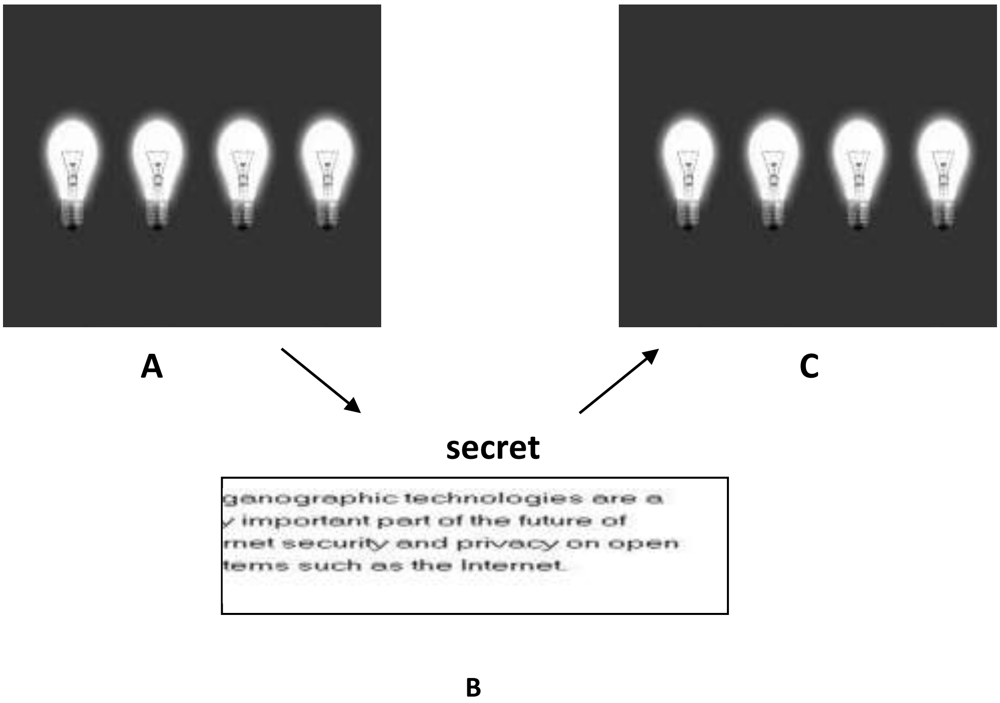
```
>>  secret = imread ('pratihar.JPG');
>>  coverImage = imread ('waterfall.jpg');
>>  numSignificantBits = 1;
>>  coverLSBZero = uint8(bitand(double(coverImage), repmat(255 – 2^numSignificantBits+1,size(coverImage))));
>>  stegoImage = uint8(double(coverLSBZero) + double(secret) / 2^(8 numSignificantBits));
>>  figure, image (cover Image),title ('Image that is used as  CoverImage ')
>>  figure, image (secret),title ('Image that will be hidden ')
>> figure, image (stegoImage),title ('Stego Image, i.e)
>>imwrite(stegoImage,'stego.jpg','JPEG')
```



Steganographic technologies are a
ery important part of the future of
nternet security and privacy on open
ystems such as the Internet.

Figure 2. embed data in a jpg . (A) the unmodified original image ,(B). the secret image ,(C).the result with the secret data embedded in it.

can you use anther kind of image (i.e. gray) by same method above. Show figure (3).



ganographic technologies are a
y important part of the future of
rnet security and privacy on open
tems such as the Internet.

Figure 3. embed data in a gray. (A) the unmodified original image ,(B). the secret image ,(C).the result with the secret data embedded in it.

Applications for Steganography in an Open Systems Environment

In this section we will look at some of the possible applications for steganography and then close by pointing out some of the more popular steganographic tools available today.

The three most popular and researched uses for steganography in an open systems environment are covert channels, embedded data and digital watermarking.

A- Covert channels in TCP/IP involve masking identification information in the the Internet when absolute secrecy is needed for an entire communication process and not just one document as mentioned next.

B- Using containers (cover messages) to embed secret messages into is by far the most popular use of Steganography today. This method of Steganography is very useful when a party must send a top secret, private or highly sensitive document over an open systems environment such as the Internet.

C- By embedding the hidden data into the cover message and sending it, you than a harmless message other than the intended recipients.

Although not a pure steganographic technique, digital watermarking is very common in today's world and does use Steganographic techniques to embed information into documents. Digital watermarking is usually used for copy write reasons by companies or entities that wish to protect their property by either embedding their trademark into their property or by concealing serial numbers/license information in software, etc. Digital watermarking is very important in the detection and prosecution of software pirates/digital thieves.

References:

[1]. Davi Tassinari de Figuueiredo," Hide In Picture (HIP)", 2002,URL: http://hide-in-picture.sf.net/

[2]. SANS Institute 2002, by Bret Dunbar,"Adetailed look at Steganographic Techniques",2002.

[3]. H. Motameni, M. Norouzi, M. Jahandar, and A. Hatami, Labeling Method in Steganography, 2007

[4]. Wikipedia, Wikimedia Foundation, Inc., " Steganography ",2008,

[5] . Johnson, Neil F., "Steganography", 2000, URL: http://www.jjtc.com/stegdoc/index2.html

[6]. MEHDI KHARRAZI AND NASIR MEMON," Data Masking: A New Approach for Steganography?" 2005

[7]. Petitcolas, Fabien A.P., "Information Hiding: Techniques for Steganography and Digital Watermarking.", 2000.

[8]. Sellars, D., "An Introduction to Steganography".

[9] J. Fridrich and M. Goljan, "Digital image steganography using stochastic modulation", SPIE Symposium on Electronic Imaging, San Jose, CA,2003.

طرق علم الإخفاء وبعض تطبيقاته(إخفاء البيانات السرية داخل الصورة)

سعدون حسين عبدالله

قسم علوم الحياة ، كلية العلوم ، جامعة الموصل ، الموصل ، العراق

(تاريخ الاستلام: ١ / ٦ /٢٠٠٨ ، تاريخ القبول: ١٢ / ٤ / ٢٠٠٩)

الملخص

في هذا البحث سنلقي نظرة في تقنيات علم الإخفاء ومختصر التاريخي لهذا العلم،ثم نتطرق إلى تطبيق إحدى هذه التقنيات وهي تقنية البت الأخير لطمر البيانات السرية في الصور من نوع jpg باستخدام برنامج الماتلاب إصدار ٧ لبرمجة هذه التقنية ثم ننهي البحث بنظرة في إمكانية استخدام علم الإخفاء في بيئة الأنظمة المفتوحة مثل الانترنيت بالإضافة إلى بعض الأدوات والمصادر المتوفرة لمساعدتنا لانجازها.