American Journal of Engineering and Applied Sciences

Original Research Paper

Comparison in Cover Media under Stegnography: Digital Media by Hide and Seek Approach

Shruti

Department of Computer Science, Guru Nank Dev University, Gurdaspur, India

Article history
Received: 16-07-2015
Revised: 21-03-2016
Accepted: 22-03-2016

Email: shruti.jairath@gmail.com

Abstract: In this world every person have some secrets which they want to hide from eavesdropper: So hiding is done in such a way that nobody will able to retrieve the secretes and this will be done by stegnography. As protection is a major demand, hence stegnography become today's security technique. Embedding secret message in cover digital media so hacker will not detect it. In this study I m going to discuss about embedding technique as well as its certain properties like security, robustness and capacity. Moreover through this research article comparison of cover media is also being discussed.

Keywords: Classical Stegnography System, DCT, DCT Coefficient, Hide and Seek Approach, Kerckhoff's Principle, Steganalysis

Introduction

Stegnography is a method of hiding secrete communication; so that stegnographic system embed data n unrecognizable cover media which is not visible by eavesdropper. In earlier days people use invisible inks to convey secret message but in computers it is implemented by stegnography. Basically hiding message concept originates from the cover image's duplicate bits (those that can be modified without destroying that medium's integrity) (Anderson and Petitcolas, 1998). In embedding process, data is hide in those duplicate bits in cover image. So that presence of secrete message will not be detected by eavesdropper. Even hidden message may not revealed by cover images but there are some statistical changes in cover image so these are detected by eavesdropper which is known as statistical steganalysis. This paper will discuss about stegnograpgy and the various methods to detect it by statistical steganalysis. We can also hide information by watermarking or else provide an overview of detection algorithms (Petitcolas *et al*., 1999; Fridrich and Goljan, 2002). Here, I present recent research and the various detection algorithm to detect distortion in cover image.

The Basics of Embedding

Three different criteria in information-hiding systems combined with each other: Capacity, security and robustness (Chen and Wornell, 2001). Capacity means that how much data is to be embedded so that it will not detect by eavesdropper and robustness means that amount of modification is done by system. Information hiding generally done by watermarking also. A watermark system has one main motive is to provide robustness, so that no body van come to know about its secret information unless until after degrade the image. Stegnography can provide high embedding capacity and security but sometimes large amount of hidden data may cause distortion in cover image. A classical steganographic system's security relies on the encoding system's secrecy. An example of this type of system is a Roman general who shaved a slave's head and tattooed a message on it. After the growth of hair then slave will go to convey the secret message (Johnson and Jajodia, 1998a). In this technique there is a problem that we can shave anyone's head so here is a lack of security. Modern steganography attempts to be detectable only if secret information is known-namely, a secret key (Petitcolas *et al*., 1999). This is similar to Kerckhoffs' Principle in cryptography, which based upon cryptography key based techniques (Kerckhoffs, 1883). For secret message to remain undetected so we must secure the cover image area which is not used in embedding so that difference will be calculated.

Information theory allows us to be even more specific on what it means for a system to be perfectly secure. Here I m going to present model which shows that how we can protect hidden message from eavesdropper (Cachin, 2002).

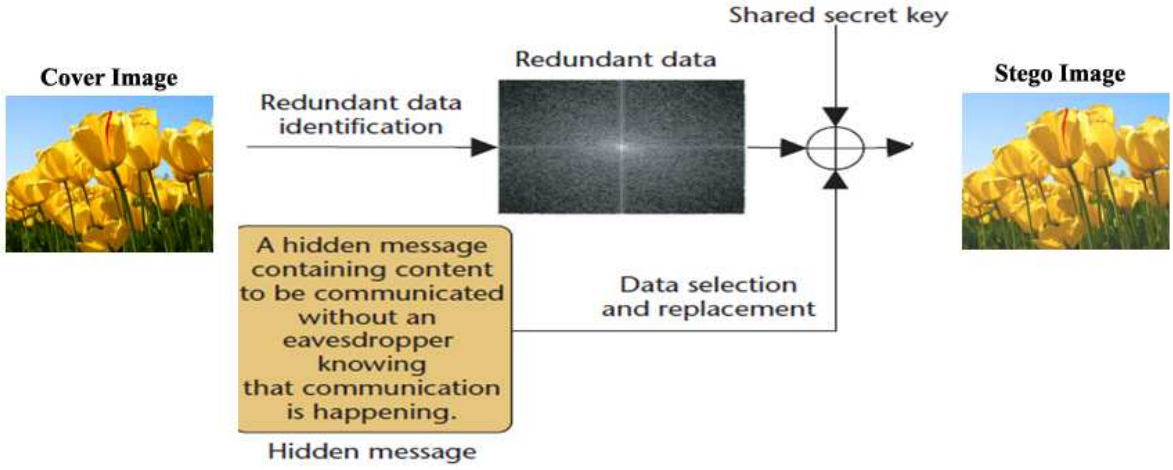

Fig. 1. Modern steganographic communication. The encoding step of a steganographic system identifies redundant bits and then replaces a subset of them with data from a secret message

In the above illustrated model adversary know about the complete encoding system but unaware about the secret key. His or her task is to devise a model for the probability distribution of all possible cover media and all possible stego-media. Only then eavesdropper will come to know about the cover image which contain secret content and which is not. If system come with any decision rule then that system is perfectly secure. Stegnographic system can determines that how a secret message is encoded. Suppose Alice needs to send secret message then he used to send message along with his secret key. So by key we can embed message in redundant bits. So when bob receives that message which is send by Alice then he used shared key so that he will come to know about the message. Figure 1 shows an overview of the encoding step; as mentioned earlier, statistical analysis can reveal the presence of hidden content (Westfeld and Pfitzmann, 1999; Johnson and Jajodia, 1998b; Farid, 2002; Lyu and Farid, 2002; Provos, 2001).

Hide and Seek Approach

Though setgnography can be implemented in any digital object either image, audio or video but here we deal with only jpeg format images. People generally use internet to transfer the images to one end to another end so jpeg is most popularly used format. Moreover system used jpeg images and they are generally not affected by visual attacks (Westfeld and Pfitzmann, 1999) (Visual attacks mean that you can see steganographic messages on the low bit planes of an image because they overwrite visual structures; this usually happens in .BMP images). In this study I presented pallet based system which is less affected by visual attacks (Johnson and Jajodia, 1998b). Let's look towards some algorithm that can covert image in detectable way. I compare the different systems and contrast in between their relative effectiveness.

Discrete Cosine Transform

Color images uses discrete cosine transformation so that we can convert 8×8 pixel blocks of the image into 64 DCT coefficients. The DCT coordinates $F(u, v)$ of an 8×8 block of image pixels $f(x, y)$ are given by, where $C(x) = 1$ when x equal 0 and $C(x) = 1$. Following is the equation which actually shows the quantization of the process:

$$F(u,v)=\frac{1}{4}C(u)C(v)\left[\sum_{x=0}^{7}\sum_{y=0}^{7}f(x,y)*\cos\frac{(2x+1)u\pi}{16}\cos\frac{(2x+1)v\pi}{16}\right],$$

$$F^{Q}(u,v)=\left\lfloor\frac{F(u,v)}{Q(u,v)}\right\rfloor$$

where, $Q(u,v)$ is a 64-element quantization table. We use LSB in order to embed secrete message. The modification of a single DCT coefficient affects all 64 image pixels. In some image format like .giff the image visual structure exist. Steganographic systems those updates least-significant bits of these image formats are generally prone to visual attacks (Westfeld and Pfitzmann, 1999). This is false for JPEG type format. All the updation in frequency domain instead of updating in spatial domain, so that's why there is no visual attacks in jpeg format images. As in JSteg algorithm it generally used to replace LSB of Discreet Cosine Transformation with the message which is to be kept hidden and it does not require and secret data.

Input: Hidden message, cover image

Output: stego image
while data is to be escaped
do
get next of DCT coefficient from cover image
if
DCT ≠ 0 and DCT ≠ 1
then
get next of LSB from message
replace LSB of DCT with LSB of message
end if
insert DCT into stego-image
end while

Sequential

Derek Upham's JSteg was the first steganographic system for JPEG images. In this LSB will be replace by message's data (see above algorithm) (Kelley, 2001). The algorithm does not require a shared secret; as a result, anyone who knows the steganographic system can retrieve the message hidden by JSteg. Andreas Westfeld and Andreas Pfitzmann noticed that

Changes in LSB will be detected by steganalysis (Westfeld and Pfitzmann, 1999). They observed about the information of large amount of data which is embedded will change in the color frequencies in histogram. It will change the LSB in image. The colors are addressed by their indices i in the color table; we refer to their respective frequencies before and after embedding as ni and ni*. Given uniformly distributed message bits, if n2i > n2i+1, then pixels with color 2i are changed more frequently to color 2i + 1 than pixels with color 2i + 1 are changed to color 2i. As a result, the following relation is likely to hold:

$$|n2i - n2i+1| \geq |n2i^* - n2i+1^*|$$

Embedding will reduce the difference between consecutive pixels. The same is true in the JPEG data format. Instead of measuring color frequencies, we observe differences in the DCT coefficients' frequency. Figure 2, displays the histogram before and after a hidden message is embedded in a JPEG image. I saw a reduction in the frequency difference between coefficient -1 and its adjacent DCT coefficient -2. We can see a similar reduction in frequency difference between coordinates 2 and 3.

Probability of secret message message will indicate by evaluating the value of p. It will start at the beginning of image and evaluating the incremental growth as in Fig. 3, shows the message created by JSteg. Higher the probability indicates the presence of secret message in cover media and dropping in above graph will shows no secret message.

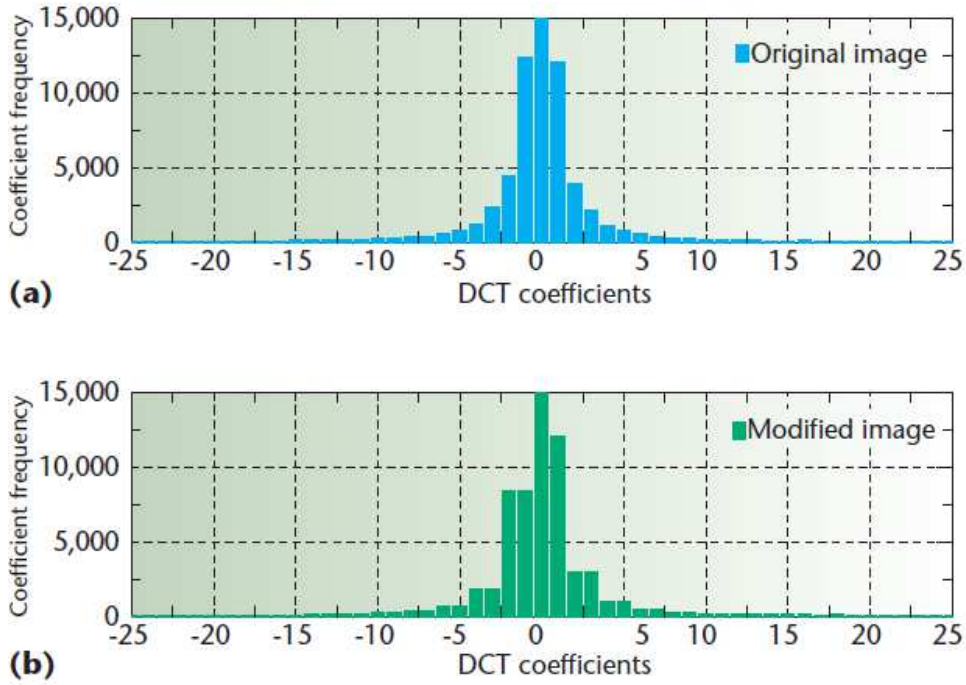

Fig. 2. Showing frequency histograms. Sequential changes which are presented in (a) original and (b) modified image's Least-Sequential Bit (LSB) of Discrete Cosine Transformation (DCT) coefficients tends to equate the frequency of adjacent DCT coefficients in the histograms

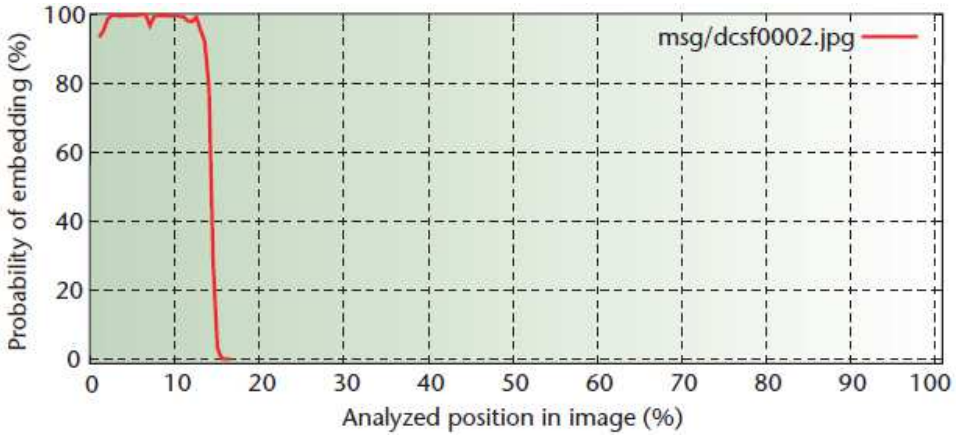

Fig. 3. The high probability in image showing large content of stegnography. With JSteg, it can be possible to calculate the message length

Steganography-Detection on the Internet

Now the question arises that how can we use steganalytic method in real world application for example if somebody claims that there are various stegnographic content posted on internet (Kelley, 2001; McCullagh, 2001; Kelley, 2002)? So to find out such type of claims we develop steganalytic framework (Provos and Honeyman, 2002) that gets off jpeg images those contain stegnography content.

Steganographic Systems in Use

To test the framework, I searched for three stegnographic system that can hide information in jpeg images and these are: JSteg (and JSteg-Shell), JPHide and OutGuess. All are using Least significant bit method for data hiding and these can be detected by statistical analysis. JSteg-Shell is a Windows user interface to JSteg first founded by John Korejwa. It used to encrypt and compress before JSteg embeds the data. JSteg-Shell uses the RC4 stream cipher for encryption (but the RC4 key space is restricted to 40 bits). JPHide is a steganographic system Allan Latham first developed which uses Blowfish as a PRNG (Schneier, 1993; Latham, 1999). Version 0.5 (there's also a version 0.3) support extra compression in hiding messages, so it uses different headers to store embedding information. Before the content is embedded, the message will be blowfish encrypted by the sender.

Detection Framework

Stegdetect is an utility that can hide messages by using jpeg images with JSteg, JPHide and OutGuess 0.13b. Stegdetect's output lists the steganographic systems it finds in each image or writes "negative" if there will be no any hidden content found. We collect Stegdetect's detection sensitivity against a set of 500 non-stego images (of different sizes) and stego images (from different steganographic systems). On a 1,200-MHz Pentium III processor, Stegdetect uses web crowler on a 10 MBit/s network. Stegdetect negative rate will depend upon the capacity of embedded data. Smaller will be the message, then it is harder to detect. Stegdetect can easily capture those images which contain stegnographic content be using JSteg. For JPHide, detection depends also on the capacity of embedded data and the compression quality of the JPEG images. Furthermore, JPHide 0.5 minimize the hidden message size by applying compression. Figure 4, shows the results of detecting JPHide and JSteg. In JSteg W will not able to detect message smaller than 50 bytes. False negative rate is about 100%. Larger the message size upto 150 bytes, false rate will be lesser than 10%. But in JPhide both are independent and the false-negativerate is at least 20%. For Outguess false rate is about 60%, a high false-negative rate, higher will be positive rate.

Finding Images

To improvise ability to detect stegnographic images, so for this we need to find those images which contain hidden messages. So for this I select images from eBay auctions (due to various news reports) (Kelley, 2001; McCullagh, 2001) and discussion group from Usenet for analysis (IABIL, 2001). To get images from eBay auctions, a Web crawler that can easily detect jpeg images is the obvious choice. When research started there were no web crawler available so many researchers develop Crawl, a simple, efficient Web crawler that makes a local copy of any JPEG images it stores on a Web page. It perform Depth first and has two key features:

Images and Web pages can be matched against regular expressions; then a match can be include or exclude from the search.

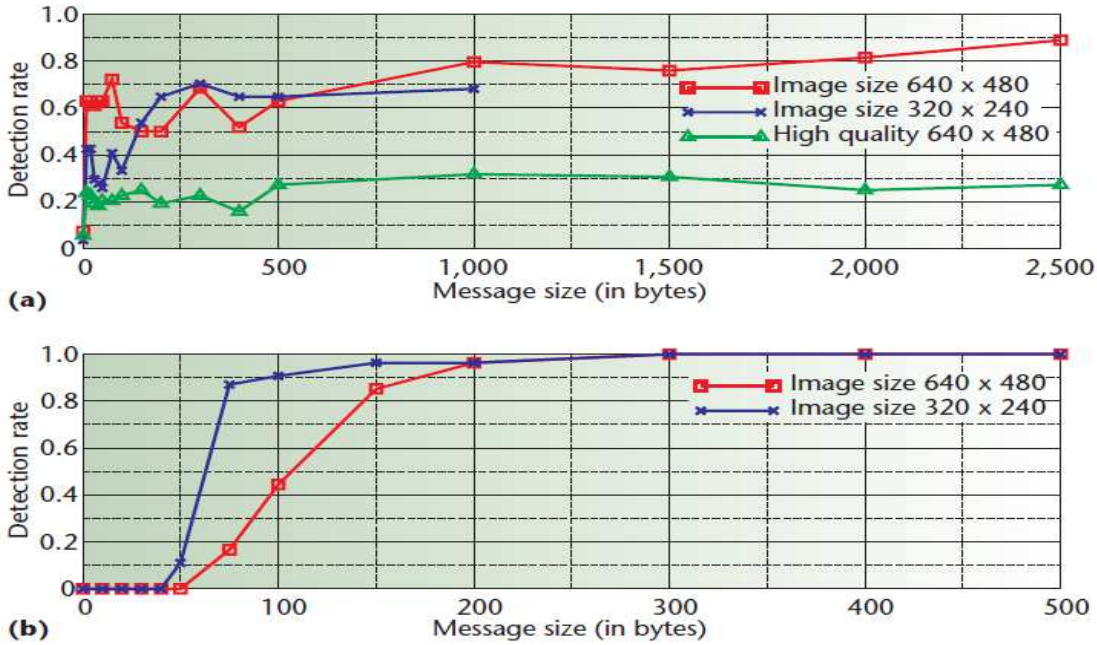

Fig. 4. Using Stegdetect over the Internet. (a) JPHide and (b) JSteg generate different results for different test images and hidden messages

Table 1. Percentage of (false) positives for analyised images

Test	Ebray	Usenet
JSteg	0.003	0.007
JPHide	1.000	2.000
OutGuess	0.100	0.140

Minimum and maximum can also be a barriers which can exclude smaller size images to contain hidden messages. I was downloaded around two hundred images linked to eBay auctions. To automate detection, Crawl uses stdout to report successfully retrieved images to Stegdetect. After processing the two hundred images with Stegdetect, I come to know about 1% of all images may contain hidden content. JPHide was detected most often contain hidden messages (Table 1).

I augmented my study by analyzing an another one hundred images from a Usenet archive. Most of them are false positive. Axelsson (1999) applied the base-rate fallacy to intrusion detection systems and presented that high percentage of false positive can impact on system efficiency. The situation is quite similar for Stegdetect. I evaluated true positive rate, in which images contain hidden content in real-as follows, where P(S) is the probability of steganographic content in images and P(S) is its complement. P(D|S) is the probability that we'll detect an image that has steganographic content and P(D|S) is the false-positive rate-which does not contain stegnographic content. Conversely, P(D|S) = 1- P(D|S) is the false-negative rate. To maximize the true-positive rate, I must maximize the numerator or decrease the denominator. Increasing detection system is not possible by increasing false-positive rate and vice-versa. We assume that P(S)- the probability that an image contains steganographic content- is very much less as compared to P(S), the probability that an image contains no hidden content. As a result, the false-positive rate P(D|S) is the dominating term in the equation; reducing hence the best way to increase the true-positive rate. The assumptions for false positive rate also dominates the computational cost to check hidden content. For the detection system to be practical, keeping the false-positive rate must be low.

Conclusion

As Stegnography is one of the security technique, which is used to hide secrets in plain sight. In this study I represents stegnography by various types of cover images like .bmp, .giff and .jpeg and out of them .jpeg is concluded to be the best one because as far as security concerns .jpeg is best cover media because when we try to embed message in .jpeg cover image then quality of image will not suffer at all i.e., image will not be distorted. In this study I work on steganalysis also by using various methods like JSteg, JP Hide and Outguess and I take two properties i.e., negative or positive. This will be ranked if and only if msg. is not detected or is message is detected. Out of various methods of steganalysis and JP Hide is prove to be the best method. But sender needs to be careful regarding the length of message because as far I judge if message is large enough then it is easy to detect by any method so that's

why I prefer that message which is to be hidden it must be of shorter length so that it is not easily detected. At last ultimately I concluded from my research that Positive rate is as low as possible just to make secure communication between sender and reciver.

Funding Information

The authors have no support or funding to report.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Anderson, R.J. and F.A.P. Petitcolas, 1998. On the limits of steganography. IEEE J. Selected Areas Comm., 16: 474-481. DOI: 10.1109/49.668971

Axelsson, S., 1999. The base-rate fallacy and its implications for the difficulty of intrusion detection. Proceedings of the 6th ACM Conference on Computer and Communications Security, Nov. 01-04, ACM, Singapore, pp: 1-7. DOI: 10.1145/319709.319710

Cachin, C., 2002. An information-theoretic model for steganography. Cryptology ePrint Archive, Report 2000/028

Chen, B. and G.W. Wornell, 2001. Quantization index modulation: A class of provably good methods for digital watermarking and information embedding. IEEE Trans. Inform. Theory, 47: 1423-1443. DOI: 10.1109/18.923725

Farid, H., 2002. Detecting hidden messages using higher-order statistical models. Proceedings of the International Conference on Image Processing, Sept. 22-25, IEEE Xplore Press, pp: II-905-II-908. DOI: 10.1109/ICIP.2002.1040098

Fridrich, J. and M. Goljan, 2002. Practical steganalysis-state of the art. Proceedings of the SPIE Photonics Imaging, Security and Watermarking of Multimedia Contents, (WMC' 02), SPIE Press, pp: 1-13.

IABIL, 2001. The internet archive: Building an 'Internet Library'.

Johnson, N.F. and S. Jajodia, 1998a. Exploring steganography: Seeing the unseen. Computer, 31: 26-34. DOI: 10.1109/MC.1998.4655281

Johnson, N.F. and S. Jajodia, 1998b. Steganalysis of images created using current steganographic software. Proceedings of the 2nd International Workshop on Information Hiding, Apr. 14-17, Springer, USA, pp: 273-289. DOI: 10.1007/3-540-49380-8_19

Kelley, J., 2001. Terror groups hide behind web encryption. USA Today.

Kelley, J., 2002. Militants wire web with links to jihad. USA Today.

Kerckhoffs, A., 1883. La cryptographie militaire (military cryptography). J. Sci. Militaires.

Latham, A., 1999. Steganography: JPHIDE and JPSEEK.

Lyu, S. and H. Farid, 2002. Detecting hidden messages using higher-order statistics and support vector machines. Proceedings of the 5th International Workshop on Information Hiding, Oct. 7-9, Springer, The Netherlands, pp: 340-354. DOI: 10.1007/3-540-36415-3_22

McCullagh, D., 2001. Secret messages come in Wavs. Wired News.

Petitcolas, F.A.P., R.J. Anderson and M.G. Kuhn, 1999. Information hiding-a survey. Proc. IEEE, 87: 1062-1078. DOI: 10.1109/5.771065

Provos, N. and P. Honeyman, 2002. Detecting steganographic content on the internet. Proceedings of the Network and Distributed System Security Symp., (SSS' 02), Internet Soc.

Provos, N., 2001. Defending against statistical steganalysis. Proceedings of the 10th Conference on USENIX Security Symposium, (SS' 01), USENIX Association, pp: 323-335.

Schneier, B., 1993. Description of a new variable-length key, 64-bit block cipher (Blowfish). Proceedings of the Fast Software Encryption, Cambridge Security Workshop, Dec. 9-11, Springer, U.K., pp: 191-204. DOI: 10.1007/3-540-58108-1_24

Westfeld, A. and A. Pfitzmann, 1999. Attacks on steganographic systems. Proceedings of the 3rd International Workshop on Information Hiding, Sept. 29-Oct. 1, Springer, Germany, pp: 61-76. DOI: 10.1007/10719724_5