Volume 4, Issue 2, February 2014 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Digital Image Steganography: A Survey

Hardikkumar V. Desai
Research Scholar,
Singhania University,
Pacheri Bari, Jhunjhunu-333515,
Rajasthan, India.

Apurva A. Desai
Professor & Head
Department of Computer Science,
Veer Narmad South Gujarat University,
Surat, India.

ABSTRACT - Today Information and communication technology is growing rapidly because Internet is creating a borderless world. Therefore, digital communication has become essential part so security is a prime concern to communicate secretly. As a result, the security of information against unauthorized access has become a prime objective. The paper focuses on survey of various steganography techniques

Keywords: Steganography, Hide & Seek, JSteg, OutGuess0.1, OutGuess0.2, F3, F4, F5

I. INTRODUCTION

The term steganography refers to the art of covert communications. Steganography can take data confidentiality to a whole new level, since it embeds message and pictures within another object by tweaking its properties, making the existence of the messages practically undetectable. One of the oldest examples of Steganography dates back to around 440 BC in Greek History. Herodotus, a Greek historian from the 5th century BC, revealed .Some examples of its use in his work entitled "The Histories of Herodotus". One elaborate example suggests that Histaeus, ruler of Miletus, tattooed a secret message on the shaven head of one of his most trusted slaves. After the hair had grown back, the slave was sent to Aristagorus where his hair was shaved and the message that commanded a revolt against the Persians was revealed [1]. Astonishingly the method was still used by some German spies at the be-ginning of the 20th century [2]. Herodotus also tells how Demeratus, a Greek at the Persian court, warned Sparta of an imminent invasion by Xerxes: he removed the wax from a writing tablet, wrote his message on the wood underneath and then covered the message with wax. The tablet looked exactly like a blank one (it almost fooled the recipient as well as the customs men).

Gaspar schotti [3], in his book of Steganography, entitled schola steganographica, published in 1665. schotti drew extensively upon the work of Johannes Trithemius (1462-1562), a German Monk and early researcher in steganography and cryptography, steganographic research continued to develop in the fifteenth and sixteenth centuries .Bishop John Wilkins – later the master of Trinitity college, Cambridge – devised a number of steganographic processes that ranged from coding messages in sheet music and string knots to invisible inks. Auguste kerckhoff's Cryptographic Militaire appeared in 1883 and was followed by Charles Briquet's Les Filigraines (1907) a historical dictionary of watermarks (Cochran 2000, 11-12; Sellars 1999, 4-5).

Figure (1), is an example of Digital Steganography, Image (a) is to be embed within image (b) and create stegogramme image (c), which we called cover or carrier image is a same as original image (b). So that it is not easy to detect that image is hidden and receiver at the other end can transmit original image (message).

(a) (b) (c)

[*fig.1 :(a)Image to be Hide (b): Cover/Carrier Image (c): Cover /Carrier Image (with Stegogramme*)]

Three different aspects in information hiding systems challenge with each other: Capacity, Security, and Robustness [4]. Capacity refers to the amount of information that can be hidden in the cover medium. Security to an eavesdropper inability

to detect hidden information and robustness to the amount of modification the stego medium can withstand before an adversary can destroy hidden information. Information theory allows us to be even more specific on what it means for a system to be perfectly secure.

II. STEGANOGRAPHIC TECHNIQUES

There are many techniques for steganography are available and are widely used among the researcher some of the popular techniques are described below.

A. HIDE AND SEEK

Steganography is applicable to all data objects that contain redundancy. People often transmit digital pictures over email and other Internet communication, and JPEG is one of the most common formats for images. Moreover, steganographic systems for the JPEG format seem more interesting because the systems operate in a transform space and are not affected by visual attacks [5]. 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. Neil F. Johnson and Sushil Jajodia [6] showed that steganographic systems for palette-based images leave easily detected distortions.

In the encoding process of this technique can be implemented by taking the first pixel of the image and obtaining its LSB value. This is typically achieved by calculating the modulus 2 of the pixel value. This will return a 0 if the number is even and a 1 if the number is odd, which effectively tells us the LSB value. Then compare this value with the message bit to be embedded. If they are already the same, then nothing to do, but if they are different then replace it. This process continues until the values are encoded. While, in decoding process, Encoder replaced the LSBs of the pixel values in sequence; the order is known to be used to retrieve the data. Therefore just calculate the modulus 2 of all the pixel values in the stegogrammes, and able to reconstruct original data.

B. JSteg

The JSteg algorithm was developed by Derek Upham and is essentially as same as a copy of the Hide & Seek algorithm, because it employs sequential least significant bit embedding. In fact, the JSteg algorithm only differs from the Hide & Seek algorithm because it embeds the message data within the LSBs of the DCT coefficients, rather than its pixel values [7].

In encoding process, JSteg algorithm shows an alternative method for calculating the LSB value of the coefficient by using mod 2. The result is replaced with the value .No key is used for this algorithm. So long as the decoder knows that the embedding took place in the DCT domain, it will be capable of extracting the message successfully. The main difficulty of not using a key is when we try to determine LSBs when extracting the message. Without a key, it is impossible to know the length of the message to extract, so the loop is typically run for the entire duration of the image to ensure that the entire message is extracted. Though, in the decoding process, the decoding process functions by converting the stegogrammes to the DCT domain. It then avoids the same coefficient values that the encoding algorithm avoids, and retrieves the hidden message from the LSBs of all the other coefficients sequentially.

C. OutGuess 0.1

In much the same way that embedding the message data sequentially using the Hide & Seek method was not considered very secure, neither was the fact that the JSteg algorithm embedded in the same fashion. The first version of Outguess, designed by Neils Provos [8], improved the JSteg algorithm by scattering the embedding locations over the entire image according to a PRNG on image. This is very similar to the way that the randomized embedding approach improved the Hide & Seek algorithm. While decoding process, for OutGuess 0.1 is might expected. Firstly, the stegogramme is converted to the DCT domain, before being shuffled in the encoding process. Then retrieve the message data by extracting the LSBs from all the coefficients whose values are neither a DC value, nor a 0 or a 1.

D. OutGuess 0.2

OutGuess 0.1 algorithm is considered much more secure than the Hide & Seek and JSteg algorithms because it not only embedded the information in a more discrete area of the image (DCT coefficients), but it also scattered the locations of embedding by using a PRNG to shuffle the ordering of the coefficients. However, after the release of the algorithm, steganalysts were able to find a serious error in the technique that left statistical artifacts in the stegogrammes. As a result, Neils Provos [9] created a revised version of the OutGuess 0.1 algorithm, called OutGuess 0.2. It was ensure that the statistical properties of the cover image were maintained after embedding, such that stegogrammes looks statistically similar to a clean image. This would make it harder for steganalysts to calculate the likelihood that their suspect image is a stegogramme.

The algorithm is exactly the same for OutGuess 0.2 as it was for OutGuess 0.1. The difference lies in what happens after the information has been embedded. In OutGuess 0.2, corrections are made to the coefficients such that they appear similar to that of a clean image in terms of frequencies of the values. This is known as statistics aware steganography.

E. F3

As an alternative to the OutGuess 0.2 algorithm, AndreasWestfeld designed an algorithm Called F3 [10]. It was considered even more secure. The reason for this is that it did not instantiate the same embedding process as the JSteg and OutGuess algorithms. Instead of avoiding embedding in DCT coefficients equal to 1, the F3 algorithm permitted

embedding in these regions, at the same time as it would still avoid embedding in zeros and the DCT coefficients. The algorithm still embedded the message data sequentially .Another change with this algorithm was that it did not embed directly in the least significant bits of the DCT coefficients, but instead took the absolute value of the coefficients first, before comparing them to the message bits. If both the absolute value of the coefficient, and the message bit were the same, then no changes are made. If they are different, then the absolute value of the DCT coefficient is reduced by 1. An implication of this however, is that zero values are often created which the decoding algorithm will not be programmed to extract data from. The F5 algorithm worked around this by re-embedding mi when the result is that a zero DCT coefficient is created.

In encoding process F3 Technique the key part is that absolute value of the current coefficient is taken such that it can then be compared with the message bit. If they are both the same then we can move on to the next coefficient and attempt to embed the next value. However, decoding process for F3 is slightly less complicated that the encoding process because it does not concern the shrinkage issue. If the coefficient value is equal to zero, it does not attempt to retrieve.

F. F4

The main drawback with F3 was the reality that it effectively embedded more zeros than ones as a result of the shrinkage mechanism. This meant that when the statistical properties of the stegogrammes are examined through its some object of embedding became visible. This is much the same as what happened in the JSteg implementation except a slightly different pattern is derived. In addition to this, steganalysts also found that more odd coefficients existed in F3 stegogrammes than even coefficients. This now meant that there were two weaknesses that could be examined when viewing the histogram of a suspect image. F4 was developed to remove these properties such that the histogram would appear similar to that of a clean image.

In F4 encoding, The F4 algorithm eliminates the two weaknesses of F3 in one stroke by mapping negative coefficients to the steganographic value, where even-negative coefficients = steganographic 1, odd-negative coefficients = 0, even-positive coefficients = 0. Put more simply, this means that now, if to embed a 0 in a DCT coefficient equal to -3, the result will remain -3, where as it would have been modified to -2 using F3. This means that the bit-flips now occur with roughly the same probability, so the histogram for the stegogramme will not appear unstructured in terms of its frequency distribution. The technique is to simplified approach to the encoding process of the F4 algorithm. While, in decoding process, stegogramme is converted to obtain the quantized DCT coefficients. It ensures that to skip the DC values and any value equal to zero. However, all other values are used to retrieve the message data in accordance with the encoding algorithm.

G. F5

The F5 algorithm [11] is predominantly the same as the F4 algorithm, at least in terms of its strategy for encoding the message data. However, the F5 algorithm was designed in an attempt to improve on the F4 algorithm by minimizing the disturbance caused when embedding the message data. This was achieved by introducing matrix encoding, and the algorithm was the first known stego-system to make use of this technique.Hamming coding is then used to embed potentially more than one bit per value by making no more than 1 change to the coefficients. This is denoted to the number of bits that are to be embedded [12].The F5 algorithm is therefore much more secure than the F3 and F4 algorithms.

III. LITERATURE REVIEW

B. P Fitzmann [13], analyze that today most of communication occurs electronically .There have been advancements utilizing digital multimedia signals as vehicles for steganographic communication. These signals, which are typically audio, video or still imagery, cover signals. Schemes where the original cover signal is needed to reveal the hidden information are known as cover escrow.

K. Tanaka et.al [14], suggested in his method that dot patterns of the ordered dither pixels are controlled by the information bits to be concealed. This system accommodates 2 KB of hidden information for a bi-level 256 x 256 image, yielding a payload of data or information hiding ratio of one information bit to four cover image bits. An information hiding ratio of 1:6 is obtained for tri-level image of the same size, the method has high payload but is restricted to dithered images and is not resistant to errors in stego image.

N. Provos and P. Honeyman [15], that least significant bit (LSB) insertion, masking and filtering techniques and transformations are the most common approaches. Each of these can be applied to various images, with varying degrees of success. Each of them suffers to varying degrees from operations performed on images, such as cropping or resolution decrementing or decrease in the color depth.

Bas et.al [16],Li,C.,Wang[17],Liao et.al[18] , suggested a great variety of steganographic methods on fractal compression principles, but greatest robustness is ensured by means of the methods suggested by[17,18] their approach are directly manipulate the code of compressed image. Building in secrecy increase the given approaches (by means of efficient stegodetector development) will provide high level of protection. Diversity of interchangeable image fragments allows coding of hidden data. For this purpose in every case of comparison with the rank block all domain candidates should be described: one should determine the set of candidates and compare with each element of the set corresponding numerical index. In this case it is sufficient to note, that the losses of secret information are due to the errors of non-correspondence of indices while building-in and restoration. On the other hand the increase of staganographic efficiency can also be achieved by increase of the number of indices.

B. Wohlberg and G. de Jager [19], in his review of fractal image coding literature analyze that fractal coding is suitable for watermarking and steganography. The fundamental principle of fractal coding consists of representing an image by a contractive transform of which the fixed point is close to that image block.

Davern and Scott in [20], proposed the use fractal image compression technique to identify the domain blocks and range block of image fractal image compression technique is used to find the self similar structure in the image and they choose a block from one of the two domain sets depending on whether the data bit they are embedding is a one or a zero.

Po-Yueh Chen, Wei-En Wu [21], suggested that by hiding more information in the edge portions, image quality is improved while maintaining the same embedding capacity. This merit results from the fact that human eyes can rarely percept trivial differences in the edge regions. The embedding capacity is adjustable according to various demands of individual users. In addition to the improvement on image quality, the proposed approach provides respectable security as well.

Guang-Yu Kang et.al[22], presented a universal steganalysis scheme based on fractal compression and Affinity propagation clustering. In their experiments, they evaluate the feasibility of taking the code of fractal image compression as features to test whether a new image is with hidden messages or not the drawback of their scheme is the relatively high time complexity, because the fractal image compression is a timeconsuming technique.

Mohammed Abbas Fadhil [23], in his system he hides a message into a stego-image by using a mapping table and some other tables to map different values of pixels in the image to the alphabetic letters of the message. One of the strong points in the system is that the system does not make any changes/distortion in the stego-image, where most steganography systems suffer from this point. Also the system tries, in its operations, to mix the properties of some cryptographic systems to provide additional security to the hidden message. The experiments on the suggested system proved that it is an easy and efficient Steganography system with a good security for the message. Therefore, this system can be considered as a Steganograpgy-Cryptograpgy system and it can be used effectively in these two fields.

Provos and Honeyman [24], Provos carried out an extensive analysis of JPEG images downloaded from e-Bay. Using his steganalytic software, he identified several thousands of "suspicious" images embedded with J-steg and JP Hide & Seek. A dictionary attack was then applied in an attempt to recover the hidden message. Although this experiment did not reveal presence of any secret messages, it does not prove that criminals are not using steganography for communication.

Simons [25], the study of this subject in the scientific literature may be traced to Simmons, who in 1983 formulated it as the "Prisoners Problem". In this scenario, Alice & Bob are in Jail, and wish to hatch an escape plan; all their communications pass through the Warden Willie; and if Willie detects any encrypted messages, he will frustrate their plan by throwing them in to Solitary confinement. So they must find some way of hiding their cipher text in an innocuous looking cover text, as in the related field of cryptography, we assume that the mechanism in use is known to the Warden and so the security must depend solely on a secret key that Alice and Bob have somehow managed to share.

Harmsen and Pearlman[26],showed that LSBM work as a low pass filter on the histogram of the image, which means that the histogram of the stego image contains fewer high frequency components compared with the histogram of its cover. Based on this property, the authors introduced a detector using the center of mass (com) of the Histogram Characteristic Function (HCF).

Swanson, Zhu and Tewfik [27], utilizes an approach of perceptual masking to exploit characteristics of the human visual system (HVS) for data hiding.

IV. Conclusion

Steganography is very popular technique for data hiding as compare to other data hiding techniques. As we saw different techniques its attempts to improve the existent one and it's hard to recognize and break the code. Every technique has its own pros and cons. we analyze that Hide & Seek and JSteg algorithm follows the sequential least significant bit for embedding. While, OutGuess 0.1 & OutGuess 0.2 are better as compare to Hide & Seek and JSteg because it scatters the locations for embedding by using a PRNG to shuffle the ordering. Moreover , F3 is advanced compared to all others technique as it embedded in DCT coefficients equal to 1 and avoids embedding in 0.The algorithm still embedding the message data sequentially. Further, F4 is enhanced but it results in shrinkage mechanism. But, in F5 disturbance is minimized by using the matrix encoding scheme. Whereas davern scott, barnsley use fractal image compression technique in which self similarity of structure in image is found through fractal compression for data hiding. As there are a number of other applications capturing interest in the subject of information hiding. Military, intelligence agencies as well as criminals require uninterrupted communications. Law enforcement and intelligence agencies are interested in understanding these technologies and their weaknesses, so as to detect and monitor hidden messages.

References

[1] Herodotus, the Histories, chap. 5 book entitled Terpsichore, & chap .7 book entitled Polymnia. London, England: J. M. Dent & Sons, Ltd, 1992, ISBN 0-460-87170-6, translated by George Rawlinson, introduction by Hugh Bowden.

[2] B. Newman, Secrets of German Espionage. London: Robert Hale Ltd, 1940.

[3] P. Gasparis Schotti e Societate Jesu Schola steganographica Working Paper in International Studies Centre for International Studies Dublin City University Working Paper 6 of 2008.

[4] B. Chen and G.W. Wornell, "Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding," IEEE Trans. Information Theory, vol. 47, no. 4, 2001, pp. 1423–1443.

[5] A. Westfeld and A. Pfitzmann, "Attacks on Stegano- graphic Systems," *Proc. Information Hiding 3rd Int'l Work-shop* , Springer Verlag, 1999, pp. 61–76.
[6] N.F. Johnson and S. Jajodia, "Steganalysis of Images Created Using Current Steganographic Software," *Proc. 2nd Int'l Workshop in Information Hiding,* Springer-Verlag, 1998, pp. 273–289.
[7] T. Zhang and X. Ping, "A Fast and Effective Steganalytic Technique Against JSteg-like Algorithms," *Proc. 8th ACMSymp. Applied Computing, ACM* Press, 2003.
[8] N. Provos. "Defending Against Statistical Steganalysis", Proceedings of the 10th USENIX Security Symposium, vol. 10, pp. 323-335, 2001.
[9] N. Provos. "Defending Against Statistical Steganalysis", Proceedings of the 10th USENIX Security Symposium, vol. 10, pp. 323-335, 2001.
[10] A. Westfeld. "F5 - A Steganographic Algorithm: High Capacity Despite Better Steganalysis", Lecture Notes in Computer Science, vol. 2137, pp. 289-302, 2001.
[11] A. Westfeld. "F5 - A Steganographic Algorithm: High Capacity Despite Better Steganalysis", Lecture Notes in Computer Science, vol. 2137, pp. 289-302, and 2001.
[12] M. Leivaditis. "Statistical Steganalysis", Master's thesis, Department of Computing, University of Surrey, 2007.
[13] B. P Fitzmann, "Trials of traced traitors." Information hiding, first international work shop, Lecture notes in computer science R. Anderson, Ed. Berlin, Germany: Springer Verlag 1996, vol. 1, pp= 49-64.
[14] K. Tanaka, Y. Nakamura and K. Matsui, "Embedding Secret Information in to a Dithered Multi Level Image," in Proc IEEE Military communications conf., Monterey, CA, 1990, pp-216 220.
[15] N. Provos and P. Honeyman. "Hide and seek: An Introduction to Steganography ", IEEE: security & Privacy, vol. 10, pp. 32-44, 2003.
[16] Bas, P., Chassery, J. M., Davoine, F. Using the Fractal Code to Watermark Images // Proc ICIP'98, 1998. – № 1. P. 469-473.
[17] Li, C., Wang, S. Digital Watermarking Using Fractal Image Coding // IEICE Trans. Fund., 2000. – № 6. – P.1286-1288.
[18] Liao, P., Chen, C., Chen, C., Pan, J. Interlacing Domain Partition for Fractal Watermarking // IIH-MSP'06, 2006. – P. 441-444.
[19] B. Wohlberg and G. de Jager, "A Review of Fractal Image Coding Literature." IEEE Transactions on Image Processing. Vol. 8, No. 12, pp.1716-1729 (1999)
[20] P. Davern and M. Scott. "Fractal Based Image Steganography," Information hiding: first international.
[21] Po-Yueh Chen*, Wei-En Wu "A Modified Side Match Scheme for Image Steganography", International Journal of Applied Science and Engineering no 7, 1: 53-60(2009).
[22] Guang-Yu Kang, Yu-Xin Su, Shi-Ze Guo, Rui-Xu Guo, and Zhe-Ming Lu ,"Steganalysis Using Fractal Block Codes and AP Clustering in Grayscale Images" JOURNAL OF ELECTRONIC SCIENCE AND TECHNOLOGY, VOL. 9, NO. 4, pp.312-314 DECEMBER (2011).
[23] Mohammed Abbas Fadhil," A Novel Steganography-Cryptography System", Proceedings of the World Congress on Engineering and Computer Science 2010 Vol I WCECS 2010, San Francisco, USA, October 20-22, 2010.
[24] N. Provos and Peter Honeyman "Detecting Steganographic Content on the Internet", CITI Technical Repot 01-11, August 2001, Submitted for publication.
[25] "The Prisoners Problem and the Subliminal Channel", GJ Simons, in Proceedings of CRYPTO 183, Plenum Press (1984) pp 51-67.
[26] J. Harsens and W. Pearlman, "Steganalysis of Additive-Noise Model Label Information Hiding", proc. SPIE Electronic Imaging, vol. 5020, pp-131-142, 2003.
[27] M. D. Swanson, B-zhu and A. H. Tewfik, "Robust Data Hiding for Images" in proc. IEEE Digital signal processing workshop, Loen, Norway, sept. 1996, pp-37-40.

