Analysis of Image Steganography Techniques in Secure Online Voting

Lauretha Rura, Biju Issac, Manas Kumar Haldar
School of Engineering, Computing and Science
Swinburne University of Technology (Sarawak Campus)
Kuching, Malaysia
{lrura, bissac, mhaldar}@swinburne.edu.my

***Abstract*— As the demand of more secure yet efficient image steganography tools has increased, different steganography approaches have been proposed. In this paper some of those approaches are examined in order to identify the appropriate image steganography technique to be implemented in an electronic voting system. Different set of criteria has been set into place to analyze and evaluate the strengths and weaknesses of the presented steganography approaches.**

Keywords - image steganography, electronic voting, Least Significant Bit, palette-based, F5, spread spectrum

I. INTRODUCTION

Steganography is science of hiding information in communications, where apart from the sender and receiver, others would not know the existence of hidden information [1]. In 2007, Hong and Hong stated steganography pays less attention to its robustness against intentional attacks since it is focusing more on the data insertion capability [2]. However, as information technology evolved and more threats arose, it is necessary to develop more secure steganography algorithms.

Steganography is better than cryptography in its ability to offer less suspicious way of hiding a secret. Therefore, steganography is proposed to be used to secure the data communication in the election procedure, as its purpose is to maintain a secret communication between two parties. This scheme provides secret communication accessible by encoding a secret message to various types of cover data such as text, images, audio, video file format. Each cover data has multiple methods to hide the secret message.

Unlike cryptography, the output data of steganography (stego-object) would still look the same as its input data. As a result, it would be difficult to identify and interpret the hidden secret in the stego-object. For electronic voting system implementation, both image and text steganography are appropriate candidates. They have a higher degree of redundancy, which allow larger size data to be encoded into the cover file. Other than that, they would less likely raise any suspicion because they are the most common transmitted data between the voter and the server.

Image steganography also offers better encoding technique as it can securely hide the secret message by securely transferring a hidden secret in a digital image file. Fig. 1 illustrates the encoding and decoding secret message mechanism of image steganography.

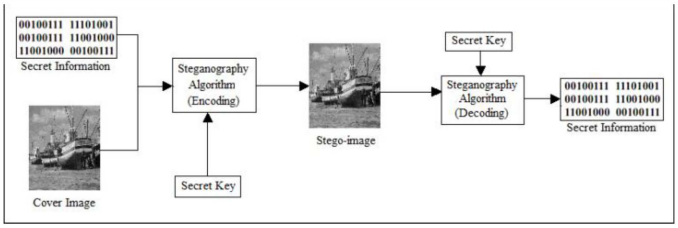

Figure 1. Image Steganography Mechanism

This paper is organized as follows. Section 2 discusses on image steganography, section 3 discusses different models, section 4 outlines related work, section 5 discusses results observed or achieved and section 6 is the conclusion.

II. IMAGE STEGANOGRAPHY

Image is a collection of pixels displayed row by row horizontally in a grid. Each pixel consists of a color and it is often represented as bits. Most monochrome and grayscale images have 8-bit depth and RGB images are represented as 24-bit image. In implementing image steganography for an electronic voting system, optimum cover image size is essential. Cover image with a greater bit depth would tend to become too large to be sent over the Internet. It could decrease the system's performance and increase the possibility of secret message being detected. Therefore, image compression holds a very important role in determining steganography techniques to be used. There are two types of image compression – lossy compression and lossless compression. In lossy compression, the original data cannot be exactly recovered. Therefore, the possibility of encoded message being lost is higher. Lossless compression on the other hand, preserves any information of the original image and embeds the secret data by using mathematical formulas. As a result, the stego-object of a PNG and BMP cover image file would have a bigger size compare to the JPEG cover image file. Image steganography can be separated into three types based on its cover image file format, such as image (spatial) domain, transform (frequency) domain and cover image in quantization format [3, 4]. In general, transform domain is more robust compared to image domain technique and cover image in quantization format. It eliminates the possibility of message being destroyed during the compression process when the excess image data is removed (lossy compression) and has lower computational cost compared to images in quantization format. Some of techniques of image steganography are described in the following section in more detail.

III. Discussion Of Different Models

A. Lease Significant Bit (LSB)

LSB is implemented with a characteristic similar to text steganography. In text steganography, a message can be hidden in every n^{th} character of a passage. Likewise in LSB the secret information is embedded in the least significant bit (8^{th} bit) or in all of the pixels of an image. In 24-bits PNG image there are three different color components, red, green and blue. When a steganography method is applied to an RGB image, each of these three components is used because each of them is represented by bytes. For example, in 200 x 200 pixel image, a total number of 120000 bits or 15000 bytes of secret data can be embedded. The following example shows how a letter S can be encoded in eight bytes of three pixels in a 24-bit image.

S: 01010011
Pixels: (11001000 11101001 00100111)
(11001000 11001000 11101001)
(00100111 11101001 11001000)
Result: (11001000 11101001 0010011**0**)
(1100100**1** 11001000 1110100**0**)
(00100111 11101001 11001000)

In this example, there are three least significant bits that were modified in order to hide the secret data. These changes in the bits cannot be detected by the human eye because our eyes are not sensitive enough to notice the differences in the LSB of two images.

LSB technique does not remove any information of the original image; instead it inserts the secret data in the original image's least significant bits of each pixel. LSB is suggested to be used for embedding secret data in a large size image. Due to this characteristic, the chances of secret data protection in the cover image being detected would be higher, as it is easier to see the difference of both original image and its cover image even if the data embedded is not visible to human eyes. To tackle this problem, hidden secret could be encoded and decoded with a pair of keys that can be distributed to both sender and receiver as another layer of security.

B. Discrete Cosine Transformation (DCT)

JPEG file format are widely known to use lossy image compression. This type of compression actually removes image data in a way that is not readily perceptible to the eye. However, this would result in altered image data in some part of the image itself that may result in some of the hidden message being destroyed. Some properties of this compression method have been evaluated to generate a steganographic approach for JPEG image file format. One of these evaluations is a study to make the cover image with embedded message invisible to the human eye by applying Huffman encoding to compress the data scheme. The other study carries lossy compression in the DCT phase and quantization phase. This is done as there are rounding errors that occur in the coefficient data that are not noticeable to the human eye [5]. This property could be used to hide the messages. F5 algorithm is based on this as shown in Fig 2.

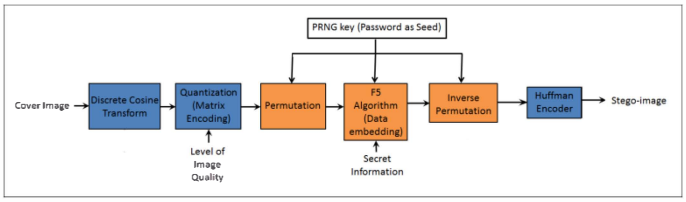

Figure 2. Message Encoding Process of F5 Steganography Algorithm. [6]

C. Palette-based Image

A palette-based image consists of a color palette and a set of color indexes. The color palette contains list of color pixels in the image whereas the color indexes are pointers to those color palette list that specify the RGB colors in the image [7]. In palette-based image steganography method, secret data can be embedded in either the bits of palette or the image data. This is because the color indices in each color palette are not the values of the color in the image. This technique permutes the colors in an image palette in a specific order together with the secret data. Using this method, Human Visual System (HVS) attack can be avoided. The selection of cover image is crucial as its color palette modification in the image could result in completely different color pixels if the selected cover image has very dissimilar neighboring palette pixels. This problem could be avoided by sorting the color palette, adding visually similar colors to the color palette with dissimilar neighboring colors, generating sub-palettes based on the similarity of colors in palette-based image, or by using greyscale cover image because it has bigger range of similar color pixels [8] as shown in Fig. 3.

Figure 3. A part of color indices in a palette-based image (mark in square) that shows each pixels color similarity with its neighboring pixels.

D. Spread Spectrum

Similar to patchwork (a pattern encoding technique), spread spectrum method hides data by spreading it throughout the cover image. Both methods are suitable for encoding a small amount of secretive information. Spread spectrum communication can be defined as the process of the spreading the bandwidth of a narrowband signal across a wide band of frequencies [5]. This can be accomplished by adjusting the narrowband waveform with a wideband waveform such as white noise. After spreading, the energy of the narrowband signal in any one frequency band is low and therefore difficult to detect. The message is stored in noise of the original image and combined with it to generate the stego image. As an image is not likely to have noise, the embedded message in the stego-image will not be noticeable by the human eye or by computer analysis without access to the original image [9]. The

encoding process of spread spectrum is shown in Fig. 4 and the process is shown below [10]:

1. Create encoded message by adding redundancy via error-correcting code.
2. Add padding to make the encoded message the same size as the image.
3. Interleave the encoded message.
4. Generate a pseudorandom noise sequence, *s*.
5. Use encoded message *m*, to modulate the sequence, generating noise, *n*.
6. Combine the noise with the original image, *i*

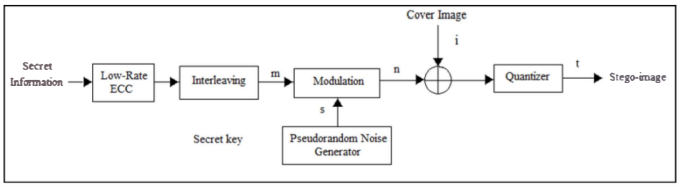

Figure 4. Encoding Process of Spread Spectrum [10].

Spread spectrum communication relates to time and frequency. Similarly, in image processing spread spectrum relates to space and data rates. The steps of spread spectrum decoding process are executed as below [10] and illustrated in Fig. 5:

1. Filter the stego-image *t*, to get an approximation of the original image, *i'*.
2. Subtract the approximation of the original image from the stego-image to get an estimate of the noise *n'*, added by the embedder.
3. Generate the same pseudorandom noise sequence, *s'*.
4. Demodulate by comparing the extracted noise with the regenerated noise.
5. Deinterleave the estimate of the encoded message *m'*, and remove the padding. Use error-correcting decoder to repair the message as needed.

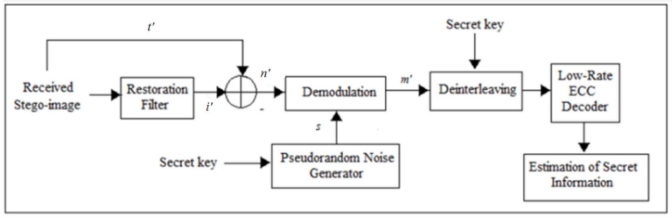

Figure 5. Decoding Process of Spread Spectrum [10].

IV. Related Work

Hayati et al., [11] examined different steganography and steganalysis tools specifically for digital forensic investigation. For this they are making use of more than 40 tools in total that varies from open source code, freeware, shareware and also commercial tools. In image steganography tools analysis, they concluded that the most popular cover image file format in both open source code and freeware or shareware is BMP whereas PNG is the least popular cover image file format. Besides that, Mathkour et al., [12] compared different steganography techniques and also tools in more detail. The investigation is based on seven image steganography technique, such as LSB, Pseurandom permutation, Patchwork, Palette-based using the palette order, Palette-based using the image data, modulating the relative size of two DTC coefficients and manipulating the LSB's of the DCT coefficient in contrast to some criteria. The parameters looked at are visibility level of hidden information, detectability, robustness of embedded data, capacity of secret information, domain type of steganography techniques and cover image file format dependency. The result of their comparison is illustrated below in Table 1 [12]. In our paper however, we included another method; Spread Spectrum.

TABLE I. Image Steganography Techniques Comparison [9]

Criteria / Technique	Domain Type	File Format Dependency	Delectability	Level of Visibility	Robustness	Embedded Message Capacity
LSB	Spatial	Yes	High	Visible	Low	High
Pseurandom Permutation	Spatial	Yes	High	Visible	Med	High
Patchwork	Spatial	Yes	Med	Invisible	Low	High
Palette-based using palette order	Spatial	Yes	High	Visible	Low	Low
Palette-based using image data	Spatial	Yes	High	Visible	Low	Low
Relative size of DCT coefficients modulation	Transform	Yes	Low	Invisible	Med	Low
DCT coefficients' LSB manipulation	Transform	Yes	Med	Invisible	Med	Med

He et al., [13] on the other hand, compared different image steganography tools with transform domain image steganography technique such as – F5, Jteg, Outguess and Outguess+. They emphasize on the impact of security measurement for different cover images quality factor and embedded data capacity. From the experiments they concluded three characteristics of its security measurement [13]:

1. If a JPEG image undergoes double compression with a different quality factor from its own, statistical distortion may be introduced even if no secret message is embedded.
2. If a smaller compression quality factor than the cover image's quality factor is used in steganography, security measurement takes bigger values. With secret message increasing, measurements vary in a complex way.
3. If a bigger quality factor than the cover image's quality factor is used in steganography, the result shows that the impact on the measurement is small.

V. Discussion and Results

Comparison of implementations of different steganography model on various standard colored images can be based on four different steganography methods described in the previous sections against three visual requirements model proposed by Johnson et al. called the magic triangle. It consists of imperceptibility of hidden information, robustness

against attacks and capacity of embedded message in the stego-image with less damage on cover image. The result presented in this section may vary if other cover image files are used. In this experiment we are using the standard 512 x 512 with 8-bit depth image for GIF image file format and 24 bit depth images for other file formats.

The experiments were done for three types of image steganography schemes based on the provided cover images' file format – transform domain, image domain, and cover image in quantization format along with four different techniques – such as F5 (JPEG), Spread Spectrum (JPEG), LSB (PNG) and Palette-based Image Algorithm using Palette Order (GIF). A set of 8-bit (GIF) and 24-bit image depths images are encoded with the range of text file starting from 100 bytes to 2000 bytes (2 kB) as the input secret message. Each image is represented in two types of digital images, RGB and greyscale with 8 bits length for each pixel.

A. Imperceptibility

Level of visibility in a stego-image is subject to the size of the hidden information, the file format of cover image and also its color tones [12]. Different sets of values on each element could determine the quality of the stego-image itself. The three aspects of image steganography technique evaluation are connected very closely to each other. Due to steganography's own character to secretly hide information in the carrier file without damaging the file itself, most of the approaches of image steganography are able to tackle the visual attacks (Human Visual System). Since imperceptibility depends strongly on the image characteristics, in this paper we will be using two sets of cover images (colored and greyscale images) and analyze the total number of their stego-images as a comparison of imperceptibility. Fig. 6 shows the result of color increment for different stego-images.

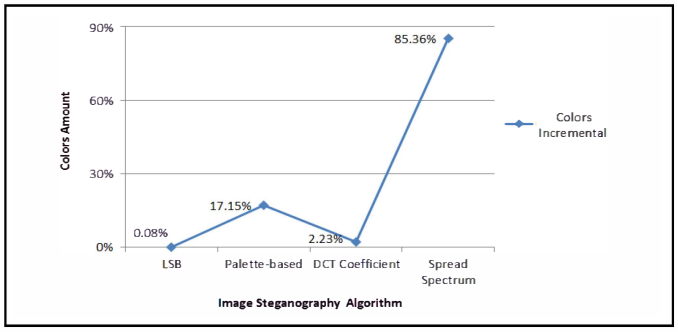

Figure 6. The color incremental for a few sets of stego-images from different image steganography methods.

B. Robustness (against security attacks)

There are two types of attacks in image steganography – visual and statistical attacks. To evaluate the quality of an image steganography approach, its robustness against these attacks must be considered. The stego-image must stay undamaged after the embedding process of hidden data. However, most of them do not support the exclusion of statistical attacks. These statistical attacks vary from Chi Square Test, RS Analysis, Binary Similarity Measures Test (BSM), DCT Domain Steganalysis, etc. All of them are considered to test the robustness of each image steganography methods proposed. Attacks and analysis on hidden information may take several forms, such as detecting, extracting, and disabling or destroying hidden information [14]. Refer to table II. In general, based on the available information, those attacks include analyzing the location of the embedded message, calculating the length of the encoded secret message and also estimating the input parameter of the algorithm including the secret generated key, or even the encoded message itself [15]. In this paper, the comparison is done based on its vulnerability level against steganalysis tools such as VSL (Virtual Steganographic Laboratory).

TABLE II. ROBUSTNESS OF DIFFERENT IMAGE STEGANOGRAPHY METHODS AGAINST STATISTICAL AND VISUAL ATTACKS

	Visual Attack	**Statistical Attack**	**Steganalysis**
LSB	Low	Yes	RS Analysis
Palette-based	Low	Yes	BSM Test
DCT Coefficient	Medium	Yes	BSM Test
Spread Spectrum	High	Yes	BSM Test

C. Capacity

Another important aspect of stego-image is capacity. It is the maximum message size that can be embedded subject to certain constraints [16]. A good image steganography method must ensure the amount of data embedded in the image does not alter the quality of the stego-image. For example, LSB technique which is the most well-known method because of its higher data encoding capacity has rather low computational complexity in its algorithm. This would make it more vulnerable to statistical attacks compared to the other techniques. Besides that, this aspect of evaluation criteria is highly dependent on the quality of the cover image. For stego-image generated with palette-based method, its cover image has a very limited numbers of colors (256 colors) in which data could be inserted. Based on our experiment on palette based image steganography tool SteganoGifPaletteOrder, we derived the maximum size of embedded data as 200 bytes. The comparison we have done in this paper is based on its capability to produce the best stego-image for different sets of secret data from the stego-images size. Fig. 7 shows data size comparison for different image steganography methods.

As observed from multiple execution of the program based on different bits of image depths, a stego-image size pattern is identified. The method from transform domain – F5 algorithm has smaller size of stego-image compared to other algorithms in image domain techniques, which is a combination of domain and quantization approach. The reason is that image domain techniques preserve the information of images. Its pixels will only be altered by inserting bits of data on the least significant bits of the image, unlike transform domain algorithms which allow its own image resizing but yet still store the secret data well. Other than that, for the algorithm in quantization format, limited input data size is allowed.

Therefore palette-based method cannot satisfy the requirements of image steganography in online voting system.

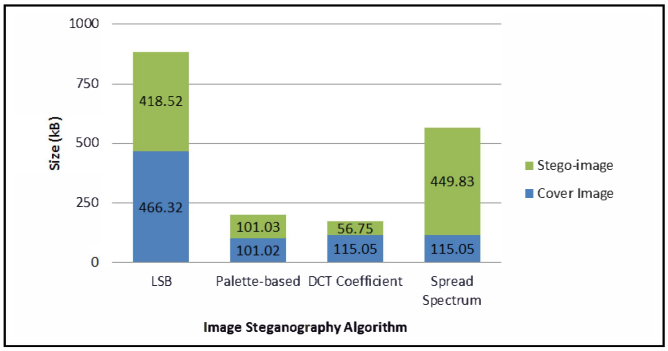

Figure 7. Comparison of initial and stego-image size.

Based on the observations, we have come to a conclusion that by implementing F5 algorithm, the stego-image's size will eventually decrease if the size of embedded data is increased. Refer to fig. 8. This could eliminate the problem of large input data. Besides that, the modification of a single DCT coefficient affects all 64 image pixels which could solve the issue of visual attack [17]. Therefore, based on those previously described aspects, transform domain techniques with 24-bits image depths are the most appropriate algorithm to be implemented in electronic voting system.

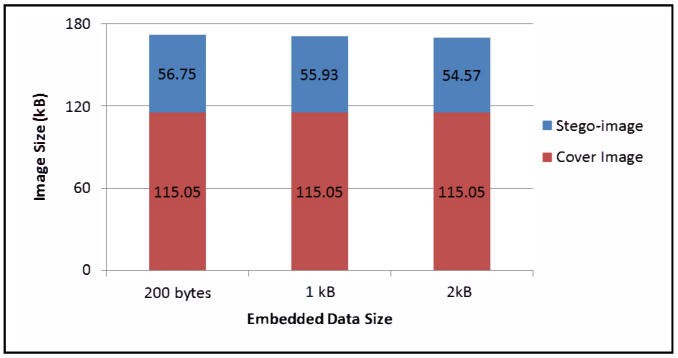

Figure 8. Comparison of initial and stego-image size with different secret data size on F5 (DCT Coefficient) Algorithm.

VI. CONCLUSION

Cryptography focuses more in preserving the contents of a message as a secret. Steganography on the other hand, focuses on protecting the existence of a message to be a secret. The objective of both of these techniques is to protect the secret information from any party during data transmission process. For online electronic voting system, the combination of both schemes is appropriate for the security needs of the system. F5 creates smaller stego-image size, and it does not include a very complex mathematical calculation. Its image file format is widely transmitted over the Internet. We are planning to use the F5 technique in online voting system due to the results obtained from evaluating each image steganography methods based on the cover image format, such as – imperceptibility, robustness against attacks and capability. F5 also shows more stable characteristics compared to other methods.

REFERENCES

[1] N. Provos, P. Honeyman, "Hide and seek: an introduction to steganography," IEEE Security & Privacy, vol.1, no.3, pp. 32- 44, Oakland, California, USA, May-June 2003.

[2] H. J. Zhang, H. J. Tang, "A Novel Image Steganography Algorithm Against Statistical Analysis," Machine Learning and Cybernetics, 2007 International Conference , vol.7, pp.3884-3888, 19-22 Aug. 2007.

[3] Y. C. Li, P. Tsai, C. H. Lin, H. L. Yeh, C. T. Huang ,"Palette Partition Based Data Hiding for Color Images," Intelligent Information Hiding and Multimedia Signal Processing, 2009. IIH-MSP '09. Fifth International Conference, pp.620-623, 12-14 Sept. 2009.

[4] M. C Chen, S. Agaian, P. Chen, "Generalized Collage Steganography on Images", IEEE International Conference on Systems, Man and Cybernetics (SMC), IEEE, 2008.

[5] N. F. Johnson, S. Jajodia, "Exploring steganography: Seeing the unseen," Computer , vol.31, no.2, pp.26-34, Feb. 1998.

[6] D. Patel and T. Schulze, "F5 a steganographic algorithm," http://www-ec.njit.edu/~shi/courses/ECE643/Fall08%20%20Course%20project%20PPT%20files/F5%20Presentation.ppt, 2008.

[7] C. H. Tzeng, Z. F. Yang, W. H. Tsai, "Adaptive data hiding in palette images by color ordering and mapping with security protection," Communications, IEEE Transactions, vol.52, no.5, pp. 791-800, May 2004.

[8] T. Morkel, J. H. P. Eloff, M. S. Olivier, "An overview of image steganography," Proceedings of the ISSA 2005 New Knowledge Today Conference, pp.1-11, Johannesburg, South Africa, 2005.

[9] L. M. Marvel, C. G. Boncelet, C. T. Retter, "Spread Spectrum Image Steganography," IEEE Transactions on Image Processing, vol. 8, no. 8, August 1999.

[10] F. Brundick, L. Marvel, "Implementation of Spread Spectrum Image Steganography," Army Research Laboratory, ARL-TR-2433, March 2001.

[11] P. Hayati, V. Potdar, E. Chang, "A Survey of Steganographic and Steganalytic Tools for the Digital Forensic Investigator", In Workshop of Information Hiding and Digital Watermarking, 2007.

[12] H. Mathkour, B. Al-Sadoon, A. Touir, "A New Image Steganography Technique," Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference, pp.1-4, 12-14 Oct. 2008.

[13] J. He, S. Tang, T. Wu, , "On the Security of Steganographic Techniques," Image and Signal Processing, 2008. CISP '08. Congress, vol.5, pp.716-719, 27-30 May 2008.

[14] N. F. Johnson, S. Jajodia, "Steganalysis: the investigation of hidden information," Information Technology Conference, 1998. IEEE, pp.113-116, 1-3 Sept 1999.

[15] G. S. Prakash, "Measures for Classification and Detection in Steganalysis," M.Eng. Thesis, Faculty of Eng., Indian Inst. Of Science, Bangalore, India, 2006.

[16] R. K. M. Chandramouli, N. Memon, " Image steganography and steganalysis: Concepts and practice," In Digital Watermarking LNCS, vol. 2939, pp. 204-211. Springer Berlin/Heidelberg, 2004.

[17] N. Provos, P. Honeyman, "Detecting Steganographic Content on the Internet," Proceeding 2002 Network and Distributed System Security Symp., Internet Soc., 2002.