International Journal of Engineering & Technology, 8 (1.7) (2019) 288-292

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Review on Steganography Methods in Multi-Media Domain

Roshidi Din[1*], Massudi Mahmuddin[2], Alaa Jabbar Qasim[3]

School of Computing, College of Arts and Sciences,
Universiti Utara Malaysia,
06010 Sintok, Kedah, Malaysia
**Corresponding author, E-mail: roshidi@uum.edu.my*

Abstract

Steganography is one area in information security that is able to conceal the secret message in any media to avoid the intruders. In this paper, the review of steganography is done in certain media such as image, text, audio, and video. It analyses some of the techniques that applied steganography to discover the development of the techniques to cover a secret message. It is expected that this paper is able to describe the implementation of steganography by previous researchers on their efforts.

Keywords: *Information hiding, Image steganographic, Cover file, Embedding Data*

1. Introduction

Basically, the historical meaning of steganography is originated from the Greek words *steganos* (secret) and *graphy* (written), in which the former is understood linguistically while the latter is practically an art. However, the combination of both words means that it is a way for secret information to be hidden and protected from parties that might seek for material, destructive or political interests. Apparently, it is the way in which data can be disguised in a fake container that appears to be visible and exposed to unauthorized parties. Besides, it also carries the meaning of the art and science of concealing secret data in any medium that enables the embedded data to be hidden of its existence. [1-5].

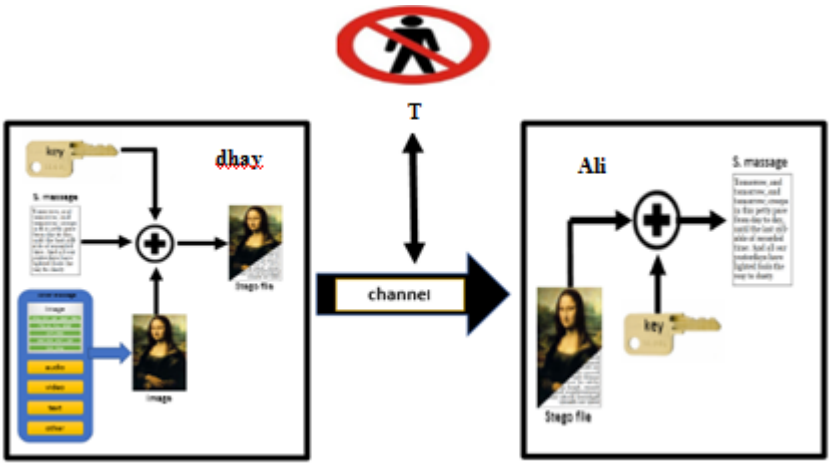

Fig. 1: General Principle of Image Steganographic System

2. Steganography

Steganography is the process of hiding confidential information in any kind of public data and can be integrated into encryption for confidential data. It becomes more difficult to discover the authorized bug as the information can be hidden in various digital media. To begin with, the confidential information to be hidden must be transferred to a binary system. Digital media is also transferred to the binary system for easy handling and integration. In general, digital steganography is a multidisciplinary field because it synthesises digital signal and data compression methods, signal coding theory, information theory, digital communication theory, cryptography, digital signal processing, and the theory of human visual perception. All these are utilised to conciliate information security.[6-9]. In the security system there are three classifications in steganography in encryption and watermark, which will be explained in Fig. 1 below.

There is one general principle of steganographic system described by the following example which contains the sender (Dhay) and the receiver (Ali). The sender intends to send a message to the receiver, and he must choose a cover file (c) in order to hide the secret message (m). In this atmosphere, to secure the recipient does not have any discovery of the unauthorised persons to know the information sent. Hence, to increase the safety of the sender (Dhay), stego key (k) is used.
Dhay gets the stego file (k) which must be indistinct from the cover file (c) neither by a human nor by a computer system. The stego file (s) represents the original (cover) file (c) along with the secret message (m) which is embedded inside this cover file.

2.1. Steganographic Methods in Image Medium

One of the most common and commonly used types of information hiding is through hiding information inside an image. This method is applied by replacing the less affected bits with the image and replacing them with the letter bits to be included in the image. It is most common because the data representation is uncomplicated (by simply flipping the low order bit, the image itself remains completely unaltered visually), [10, 11]. Steganographic methods that alter image files for information hiding is shown in Table 1.

Table 1: Review of Image steganographic implementation

Method	Summary	Advantages and Disadvantages
Spatial Methods (substitution techniques) [13]	In the substitution technique of steganography, The secret information parts are replaced by the handy Y-Bytes substitution in the cover file without causing any radical change in the cover file In addition to that Technique LSB substitution technique is relatively quick and easy to use, It is the most common and most commonly used technology for hiding digital information, [10-14]	• Advantages • Image quality is not changed for any algorithm that uses spatial method.[14] • The large capacity of data can be stored. [14] • Disadvantages • amount of additive noise maybe creeps in the image[15] • Editing the image leads the image to lose its secret data. Less robust.[14] • Secret data can be modified by the intruder during the communication.[14]
Transform method [13]	Unlike spatial domain techniques (e.g. LSB technique) transform (frequency) domain techniques You hide confidential information in part of the cover file for you are considered frequency domain techniques More powerful than attacks compared to the first type spatial domain techniques Here, most of robust steganographic systems known today on frequency domain techniques. [13, 16-18]There are many transforms Used in this technique to set the signal in the frequency band Discrete cosine transforms (DCT), discrete 22 wavelet transform (DWT), and discrete Fourier transform (DFT) are methods used as mediums to embed secret data in digital images. [19-27].	• Advantages • To hide data in most significant areas of the cover-image, it makes them more robust from attack than LSB.[14] • It can be applied changes for the whole image [11] • secret and embedded data will be spread across the entire image and will not be concentrated on one certain area or region [11] • Disadvantages • These method types are computationally complex.[14] • where you add a slight noise to frequency domain components it changes the whole image rather than changing only this part of the image.
Spread Spectrum method [13]	define spread spectrum communication as" the process of spreading the bandwidth of a narrowband signal across a wide band of frequencies". In spread spectrum steganography, the frequency domain of the cover file is a communication channel and the secret message as a signal that is transmitted through it. Since the secret message is spread through a wide frequency band, this technique is relatively robust against stego file modification or message removal [28-32]	• Advantages • Robustness against narrow band interference • Relatively high security • Coexistence of several signals the receiver can separate each user based on code • No need of frequency planning as all user uses same BW • Wide band signals less prone to interference, less prone to fading • Disadvantages • Increased complexity • Needs synchronization between T & R Large BW
Statistical method	These techniques embed only one bit of secret data in a cover file. Therefore, it is known as "1-bit" steganography scheme. If "1" is hidden in a cover file, some statistical characteristics (e.g. entropy and probability distribution) of this cover file must be changed significantly to clearly indicate the existence of a message. However, if the hidden bit is "0", the cover file is left unmodified. Therefore, this technique entirely depends on the ability of the receiver to differentiate between changed and intact cover files[1, 4, 33-35]	
Distortion method	Most technologies are designed to hide information by making it unclear, which means that the recipient does not need the original cover file to find confidential information Only in this type Distortion Techniques The recipient needs a copy of the cover file in order to compare it in order to extract the secret information hidden in the file stego	

2.2. Steganographic Methods in Audio Medium

This study discusses digital audio in the computer and modern technologies in order to know that the sound, as a voice signal, can be converted in multiple ways into a format that interprets the computer in binary representation (1 and 0).

Table 2: Review of Audio medium steganographic

Method	Summary
LSB Coding [37, 40, 43, 44]	LSB is considered as the easiest technique in implemented in information hiding of digital audio. LSB Coding can be done by simply replacing the LSB of each sampling point by hidden data
Phase Coding[[37, 40, 43, 44]	Phase Coding works by substituting the phase of an initial audio segment with a reference phase, this phase represents the hidden data. Instead of breaking a signal down into individual samples, the parity coding method breaks a signal down into separate regions of samples and encodes each bit from the secret message in a sample region's parity bit. If the parity bit of a selected region does not match the secret bit to be encoded, the process flips the LSB of one of the samples in the region Parity Coding
Spread Spectrum [37, 40, 43, 44]	Spread spectrum (SS) is technique designed to encode any stream of information via spreading the encoded data across as much of the frequency spectrum as possible. Even though, there is interference on some frequencies, SS allows the signal reception.

The acoustic signal is generated by sampling the samples of analog and continuous signals at a specified rate within a specific waveform PCM (Pulse Code Modulation), which is the most widely used and most popular for storing digital sounds or transforming live sounds into digital formats [10, 11].
Meanwhile, audio uses digital audio formats such as MPEG, AVI MIDI,WAVE.[12, 13]. Hence, to secretly embed data into digital audio file, there are several methods that have been introduced earlier. The lists of the methods [13-18] are summarized in Table 2.

2.3. Steganographic Methods in Video Medium

This type of data carrier is only recently known. The segregation of video into audio and images or frames has resulted in the data hiding method efficiency. In comparison with other types, the use of video information is more convenient (combination of pictures) to be used as a carrier for information hiding through video steganography uses formats such as AVI, MPEG, Mp4, H.264 and others [10, 19-22] which is shown in Table 3.

Table 3: Review of Video medium steganographic

METHOD	SUMMARY
TRANSFORM DOMAIN EMBEDDING [1, 4, 33, 35, 47, 49]	This is a more complex way of hiding information in an image. Various algorithms and transformations are applied on the image to hide information in it. DCT (Direct Cosine Transformation) is one such method, which is used in JPEG compression algorithm to transform successive 8x8 pixel blocks of the image, into 64 DCT coefficients each (). DCT helps separate the image into parts (or spectral sub-bands) of differing importance (with respect to the image's visual quality), embedding in DCT domain is simply done by changing DCT coefficients, for example by changing the least significant bit of each coefficient. One of the constraints of embedding in DCT domain is that many of the 64 coefficients are equal to zero and changing too many zeros to non-zero values will have an effect on the compression rate. That is why the number of bits one could embed in DCT domain, is less that the number of bits one could embed by the LSB method
LEAST SIGNIFICANT BIT INSERTION [3, 37]	This is a very popular method because of its simplicity, in this method; the LSB bit of 1 byte in the image is used to store the secret data. The resulting changes are too small to be recognized by the human eye The extraction of the data from the video file illustrated by read the video frames, then the desired frame can be identified by it sequence number, after identifying the frame number the extraction function can be started.
CONSIDERING VIDEO AS SEPARATE IMAGES [3, 37]	In this method, each video frame is considered as a separate image, in which information is hidden. The main advantage of this method is the possibility of using the algorithms used in image steganography and watermarking for video, but it requires a large amount of computation.
REAL-TIME VIDEO STEGANOGRAPHY [3, 37]	This technique involves hiding the information on the output image of the instrument (such as image displayed by an electronic advertising billboard). If the pixel colors of the blocks are similar, it changes the color characteristics of a number of these pixels to a certain extent, so data information is hidden in the image. In the following section the embedding of data into video-based steganography. The embedding of data within video file start by selecting the desired video, after selecting the video, the system should read all the video frames and assign frame number to each frame and then the desired frames can be selected for further processing.

2.4. Steganographic Methods in Text Medium

One of the least types that contains noise is the text is used to hide the data. However, there are disadvantages of the techniques that will be discussed. Hiding in the text is usually of a small capacity in the data, and this will be discussed in this part as well. In a primitive Word processor where spaces have fixed size, a bit can be hidden at the end of each sentence by appending one or two spaces to the sentence, where one space indicates a hidden 0 and two spaces indicate a hidden 1. Since a sentence ends with a period, every period in the text, even those in a context such as “Mr. Smith,” hides one data bit and must be followed by one or two spaces. Appending one or two spaces to the end of each line is also a simple data hiding method. Such spaces do not show up when the text is printed but can be easily identified by the Word processor.
There are various methods of text steganography. The first method is Selective Hiding (SELH) that hides the characters on the first or any specific location of the words to combine that characters and help in extracting the text. Nevertheless, this technique requires a huge amount of plain text. The second method is HTML Web Pages (HWP) that hides the text by using the attributes in HTML. The character is then used to retrieve the original text. The third method, Hiding Using Whitespaces (HUW), has a smaller number of whitespaces between words which can determine the whitespace is 0. However, if there are more numbers of whitespaces between words, this may be determined by the last method, Semantic Hiding (SEMH), that uses synonyms to hide the message [41]. A potential problem may arise when the text is processed by programmers that remove extra blank spaces. Text steganography techniques can be classified into three basic categories.[23, 24]

Table 4: Review of Text medium steganographic

METHOD	SUMMARY
SYNTACTIC METHOD [38]	Moreland discussed about text steganography by using punctuation signs such as full stop or period (.), comma (,), semi colon (;), quotes („") etc. in the text of encoding a secret message. The use of punctuation sign is quite common in a normal English text and hence, it becomes difficult for the intruder to recognize the presence of the secret message in the text document. [16]
LINE SHIFTING METHOD [11, 12, 16, 45, 52]	In this method, the line of text is vertically shifted to some degree and information is hidden by creating a unique shape of a text.
WORD SHIFTING [39]	In this method, the information is hidden by shifting the words horizontally or by changing the distance between the words.
ABBREVIATION [39]	In this method very, little information is hidden in the text ; example : only a few bytes of data can be hidden.
SYNONYM METHOD [12, 16]	In this method, certain words along with their synonyms are used to hide the secret message in the text.
WORD SPELLING [3]	This method is used for hiding data in English text. In this method, word spelling in US English are spelt differently from UK English.

3. Conclusion

The above table with Steganography techniques of data in this study is able to be observed in the first type of Image Steganography, and its multiple technologies of high imperceptibility. This study can apply them in substitution domain method, transform domain, and spread spectrum either in statistical method or the imperceptibility. Currently, this study has led to utilising Robustness to hide data in images. However, the most powerful and safest way is to use the transform domain.
Meanwhile, the intermediate way is spread spectrum. The rests are substitution domain method, statistical method, and distortion method. They are considered the weakest and least secured, and through this discussion the first type of Steganography is steganography in image. Hence, it can be concluded that the spread spectrum method is the strongest as it contains (High Imperceptibility, Medium Robustness, and High Payload Capacity). The weakest technique, however, is Distortion method because it contains Low Imperceptibility, Low Robustness, and Low Payload Capacity.

Table 5: Summary of multi steganographic medium

TYPE OF STEGANOGRAPHY	METHOD	ADVANTAGE	DISADVANTAGE
IMAGE STEGANOGRAPHY	Substitution domain	High Imperceptibility and High Payload Capacity	Low Robustness
	Transform domain	High Imperceptibility and High Robustness	Low Payload Capacity
	Spread spectrum	High Imperceptibility, Medium Robustness and High Payload Capacity	
	Statistical method	Medium Imperceptibility	Low Robustness and Low Payload Capacity
	Distortion method		Low Imperceptibility, Low Robustness And Low Payload Capacity
AUDIO STEGANOGRAPHY	LSB Coding	High Payload Capacity and Low computational complexity	Low Robustness
	Phase Coding	Basic technique	Low Payload Capacity and easily by the attack
	Parity Coding	Medium Robustness and more of a choice in encoding the secret bit	Low Robustness
	Spread Spectrum	Difficult to detect by unauthorized users and provide a high level of Robustness It is one of the requirements of good data concealment	delay when applied
	Echo Hiding	easy to implement.	prone to inevitable mistakes, such as the echo from the host signal
VIDEO STEGANOGRAPHY	Transform domain embedding	It is a difficult way to hide data	Used on a method DCT Which affect the rate of data that can be hidden compared to the method LSB
	Least significant bit insertion	very popular method because of its simplicity	You need a lot of time to figure out the video frame in which the data was hidden in order to extract it
	Considering video as separate images	can using the algorithms used in image steganography and watermarking for video,	requires a large amount of computation
	Real-time video steganography	provide more security	delay when applied
TEXT STEGANOGRAPHY	Syntactic method	The amount of information is very few compared to other methods	Hidden data can easily be found for ordinary people
	Line shifting method	text OCR (character recognition) This method is appropriate and good	When OCR applies hidden data lost
	Word shifting	Specify spaces between words to fill in data	In the case of someone who knows the design of the algorithm it will be easy to detect it
	Abbreviation	it's a kind of any abbreviation present	limited only for small data
	Synonym method	use different terms of words that hide data properly.	It takes a lot of time to change one word to another
	Word spelling	good method for data hiding not only for electronic document but also for printing text.	less secure than new synonyms text method

Acknowledgement

The author would like to thank the Ministry of Higher Education Malaysia in funding the grant under the Fundamental Research Grant Scheme (FRGS), S/O code 13576, and Awang Had Salleh Graduate Grant (S/O Code: 15817) with School of Research and Innovation Management Centre, Universiti Utara Malaysia, Kedah for the administration of this study.

References

[1] M. Kumar, "Steganography and steganalysis of joint picture expert group (JPEG) images," University of Florida, 2011.
[2] B. Li, J. He, J. Huang, and Y. Q. Shi, "A survey on image steganography and steganalysis," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, pp. 142-172, 2011.
[3] J. H. Bhavsar and I. Khan, "Techniques of Steganography and Steganalysis," *Available at SSRN 2029407*, 2012.
[4] Y. J. Chanu, K. M. Singh, and T. Tuithung, "Image steganography and steganalysis: A survey," *International Journal of Computer Applications*, vol. 52, 2012.
[5] K. Bailey and K. Curran, "An evaluation of image based steganography methods," *Multimedia Tools and Applications*, vol. 30, pp. 55-88, 2006.
[6] B. G. Banik and S. K. Bandyopadhyay, "Review on Steganography in Digital Media."
[7] R. Samagh and S. Rani, "Data Hiding using Image Steganography."
[8] R. J. Anderson and F. A. Petitcolas, "On the limits of steganography," *IEEE Journal on selected areas in communications*, vol. 16, pp. 474-481, 1998.
[9] N. Johnson and S. Jajodia, "Steganalysis of images created using current steganography software," in *Information Hiding*, 1998, pp. 273-289.
[10] B. G. Banik and S. K. Bandyopadhyay, "Review on Steganography in Digital Media," 2015.
[11] R. Samagh and S. Rani, "Data Hiding using Image Steganography," 2015.
[12] N. Cvejic, *Algorithms for audio watermarking and steganography*: Oulun yliopisto, 2004.
[13] P. Jayaram, H. Ranganatha, and H. Anupama, "Information hiding using audio steganography–a survey," *The International Journal of Multimedia & Its Applications (IJMA) Vol*, vol. 3, pp. 86-96, 2011.
[14] N. Cvejic and T. Seppänen, "Increasing Robustness of LSB Audio Steganography by Reduced Distortion LSB Coding," *J. UCS*, vol. 11, pp. 56-65, 2005.
[15] N. Cvejic and T. Seppanen, "Increasing robustness of LSB audio steganography using a novel embedding method," in *Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004. International Conference on*, 2004, pp. 533-537.
[16] C. Parthasarathy and S. Srivatsa, "Increased robustness of LSB audio steganography by reduced distortion LSB coding," *Journal of Theoretical and Applied Information Technology*, vol. 7, pp. 080-086, 2005.
[17] M. Nosrati, R. Karimi, and M. Hariri, "Audio steganography: a survey on recent approaches," *world applied programming*, vol. 2, pp. 202-205, 2012.
[18] S. Moon and R. Kawitkar, "Data security using data hiding," in *Conference on computational intelligence and multimedia applications, 2007. International conference on*, 2007, pp. 247-251.
[19] F. A. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding-a survey," *Proceedings of the IEEE*, vol. 87, pp. 1062-1078, 1999.
[20] V. M. Wajgade and D. S. Kumar, "Enhancing data security using video steganography," *International Journal of Emerging Technology and Advanced Engineering*, vol. 3, pp. 549-552, 2013.
[21] S. Bhattacharyya, "A survey of steganography and steganalysis technique in image, text, audio and video as cover carrier," *Journal of global research in computer science*, vol. 2, 2011.
[22] A. Al-Frajat, H. Jalab, Z. Kasirun, A. Zaidan, and B. Zaidan, "Hiding data in video file: An overview," *Journal of Applied Sciences(Faisalabad)*, vol. 10, pp. 1644-1649, 2010.
[23] O. G. Roshidi Din, Alaa Jabbar Qasim, "Analytical Review on Graphical Formats Used in Image Steganographic Compression," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. Vol 12, No 2, p. pp. 441~446, November 2018.
[24] A. J. QASSIM and Y. SUDHAKAR, "Information Security with Image through Reversible Room by using Advanced Encryption Standard and Least Significant Bit Algorithm," 2015.