Audio Steganography to Protect the Confidential Information: A Survey

Elham Zinaly
University College of Nabi Akram
No. 1283 Rah Ahan Street
Tabriz, Iran

Avaz Naghipour
University College of Nabi Akram
No. 1283 Rah Ahan Street
Tabriz, Iran

ABSTRACT
Steganography is a science to hide information, it hides a message to another object, and it increases the security of data transmission and archiving it. In the process of steganography, the hidden object in which data is hidden the carrier object and the new object, is called the steganography object. The multiple carriers, such as text, audio, video, image and so can be mentioned for steganography; however, audio has been significantly considered due to the multiplicity of uses in various fields such as the internet. For steganography process, several methods have been developed; including work in the temporary and transformation, each of has its own advantages and disadvantages, and special function. In this paper we mainly review and evaluate different types of audio steganography techniques, advantages and disadvantages.

Keywords
Audio steganography, Cryptography, Image steganography, Information steganography, LSB method, Security.

1. INTRODUCTION
In the conditions that human life has been changed to the current transformed form, human life style was transformed, and anomalies of urban life transformed with this transformation. Making the security of information and communication to exchange messages and information with classification of confidential, particularly in military, security and informational communications has been possessed great importance, and the development of information and communication technology in recent decades, and its influence and comprehensive in most civil actions such as economic, social, cultural, scientific and information has opened much larger space in relation to the issue of security of information and communication. The cryptography aims to protect confidentiality and message integrity that make it by cryptography. In this paper, given its aim that is the review and evaluation of different methods of audio steganography to protect the confidentiality of messages; first, methods and perspectives provided in steganography field were described. Then, by using them, the study of audio steganography that has more advantages compared to other methods was conducted. Some of these benefits include a good balance between security and quality, resisting the attacks of watermarking and no need to additional computational cost.

The outline of the paper is as follows: In Section 2 we give introduction of steganography. Section 3 discusses components of steganography. Section 4 presents features of information hiding techniques. Overview of audio steganography methods and methods provided for audio steganography are studied in Sections 5 and 6 respectively. Section 7 discussions the paper. The paper ends with a brief conclusion.

2. INTRODUCTION OF AUDIO STEGANOGRAPHY
The main purpose of steganography is to hide the reality in the communicating. So that sender of a secret message hides it into a digital content (image, audio, and video) and only the recipient can extract this message. Because the hidden message is completely carried out with remaining the quality for host, all those who have access to communication channels understand to deliver the digital content. Fig. 1 is presented for the best understanding about the subject of information hiding [1], and then each of the branches of security systems has been studied separately. Steganography compared to encryption has more security, since it is impossible that an unauthorized person can access to the content of messages. But steganography denies the existence of the message. The purpose of encryption is to protect the confidentiality and integrity of the message that it produces with coding it. While steganography pursues the same objectives by encrypting the message, it causes to select and hide the message bits among bits of host by utilizing a type of password. Also, in some cases, the message before embedding into the host using encryption algorithms is encrypted and then is hidden. In fact, three very strong protective layers will be made to access the message by steganography: first, the relationship is inappreciable that is the main objective in steganography, and therefore it will not be so simple to pass the first hurdle. If it is doubted that there are not any information into a host, the second step will be to find the hiding algorithm. This means that the location and hiding sequence of information should be determined. But at this stage, because using a key named secret key to embed the messages, it is necessary to detect this key, and so passing this stage will be difficult. If the previous two stages are passed successfully, then encrypted text was obtained, and therefore now encryption issues are considered at this stage. There are different ways to hide the information, such as text, audio, image and video, which the least significant bit (LSB) method are commonly used and JPEG files are the best format on the network and the internet that have high quality and good compression. During the use of audio, a new method is a discrete wavelet transform (DWT), which it has the benefits, including independence from sound when data will recovery, high capacity hiding, complete recovery and high quality audio of sound hidden. Various methods have been proposed for audio steganography that each of these methods tries to offer the best quality and high security. Table 1 shows the comparison of three methods of watermarking, steganography and cryptography. In this paper, a comprehensive study will be implemented about steganography and especially audio ones, and also different methods for steganography will be evaluated.

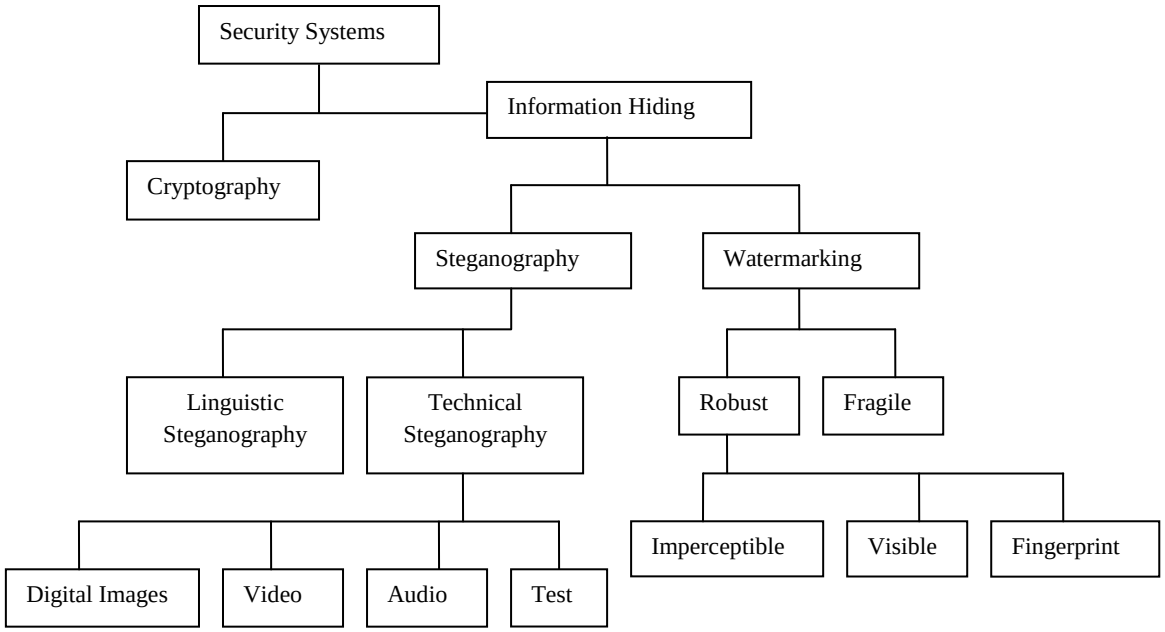

Fig 1: Overview of information hiding [1]

Table1. Comparison of three methods of watermarking, steganography and cryptography [2]

Index	**Watermarking**	**Steganography**	**Cryptography**
Carrier object	Images/Video	Each media file	Files based on image
Confidential information	Watermark	Each type file	Text
Secret key	Arbitrary	Arbitrary	Obligate
Visual field	Possible or Impossible	Never	Yes
Objective	To protect copyright	Secret communication	Protecting
Security	High	Superior	High
Capacity	Low	High	High

3. THE COMPONENTS OF STEGANOGRAPHY

The steganography algorithm consists of two phases, the embedding intelligence information in the receiver, and extracting them in the transmitter. Fig. 2 shows how to embed the information, image, video and audio are used. The secure data is became in code with a key and stored on multimedia intermediates. The combination of the secure information coded and multimedia intermediate which is the covering is called steganography combination. In the extraction phase, the secret information encrypted is extracted from the covering and then the decoded methods help to recover the information.

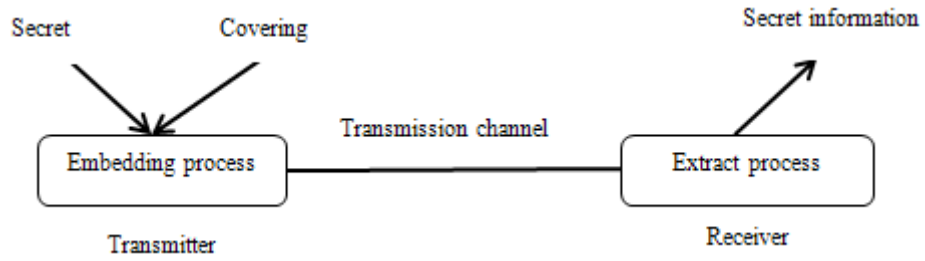

Fig 2: The components of steganography [3]

4. FEATURES OF INFORMATION HIDING TECHNIQUES

Steganography technique covers the information within image, audio and text. The advantages and disadvantages of different features in each of techniques exist. The relative importance of each components is based on their applications [3].

4.1 Security

The information security subject can be very important for security and private organizations in a country or city. And it for each section can be used based on the importance of information steganography. Steganography method may suffer from various active or inactive attacks. If there are confidential data, which is more than a random guessing, can discover existing some safe hidden breaker and steganography systems. On the other hand, can be said that steganography is a technique to hide the unsafe information.

4.2 Capacity of the Information Hiding

The capacity of the information hiding refers to the amount of information that can be hidden in cover medium without deteriorating the integrity of the cover image [1]. It in terms of bits per pixel is represented.

4.3 Clarity of Perception

The hiding a message in the cover needs to make some noise and parasite within a cover image. It is very important that the hiding the object occurs without the reducing the perceptual quality of the covering object. After hiding, confidential information in the image should not be changed and it must not see with a clear view that information are hidden. In fact, the hidden image should be very similar to the original and cannot differentiate them, and must maintain the integrity of the original image. In a secret communication application, if an attacker observe some distortion that causes suspicion of the presence of data in the image, technical support shows that hidden breaker has failed, even if the attacker is able to extract messages [4]. For functional applications where its perceptual clarity of the data hidden is not critical, allowing more deviation of hidden object in image could increase strength.

4.4 Robustness

One of the main objectives of robustness is to acquire steganography. It refers to the degree of difficulty required to determine whether the image contains hidden information or not. Robustness acts very good to protect copyright. Because the hackers will make an effort to destroy and filter out any text in embedded images [5].

Table 2. Comparison of techniques in spatial and frequency domains

Feature	Spatial domain	Frequency domain
Amplitude of function	Based on pixel	Transforming image to frequency domain
Capacity	High	Depends to the image structure
Stability	Low stability against pre-processing	High stable
Invisibility	Depends to the hidden data and the tissue of image	It is not visible at majority of image structures
Cost	With low computational cost	With high computational cost

5. AUDIO STEGANOGRAPHY ETHODS

Steganography to hide the secure data must be strong enough and does not cause fundamental changes in the carrier. The hiding information, in addition to its application in steganography to protect property rights, is a means to make a hidden relationship. Depending on the type of function, researchers the different methods have suggested in the spatial domain and convert to audio steganography [6, 7, 8]. In the most of steganography methods, LSB methods are used to hide data. Steganography is performed on the audio that like other types based on the file structure could be done in the spatial or frequency domains. Even the nature-inspired algorithms such as genetic algorithm are used for the audio steganography [9], that a new approach to audio steganography is presented. Table 2 shows the comparison of methods in spatial and frequency domains. The various methods of audio steganography are:

1. LSB coding
2. Balance coding
3. Phase coding
4. Spread spectrum
5. Echo hiding

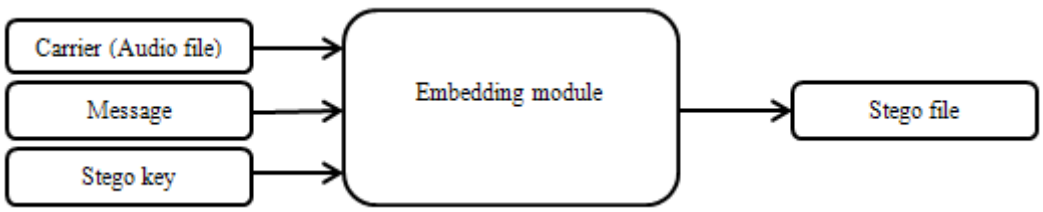

Fig 3: Basic model of audio steganography [10]

The basic model of audio steganography consists of Carrier (Audio file), Message and Password is shown in Fig. 3.

With the spread of the internet and the emergence of new technologies, and the possibility of unauthorized access to information of individuals and organizations has been causing concerns. Therefore, there is a need for secure communication. To create a secret and secure communication can be used steganography. There are two fundamental concepts at steganography, including the host signal and the watermark signal. The host signal is a signal that the data is embedded in it; this signal should not contain any additional data. Created signal after embedding data is called watermark signal that contains the hidden messages [11]. These two signals should not be audibly and quality differentiable and have not any perceptible difference. Important features of these systems include the robustness, clarity (transparent), watermark bit rate (capacity), security, complexity and how to detect them. In fact, a perceptual similarity between the host and watermark signal is called clarity. Robustness is the ability of the algorithm to detect the watermark signal after changes in the common processes of signal. Watermark bit rate is the number of bits watermarked per unit of time that can be embedded in a host signal, but enemy fails to understand that is hidden in the signal. Complexity refers to the amount of computations and speed required in embedding and detection algorithms, and the number of improvisers and detectors used in the system. Some algorithms need the host signal to detect data that is called non-smart detection, and the detection without need the main signal is called the blind or intelligent [12]. Here we will discuss the methods of steganography in detail, and we will compare the fundamental features of audio steganography with basic algorithms.

5.1 Concealment methods in the interim

5.1.1 Coding method of least significant bit

This method is a more basic one to hide information. Watermark data bits are embedded in the least significant bits of carrier voice samples. Its embedding capacity is high, but by changing bits of samples that are used to embed the watermark bits, an additive Gaussian white noise with low robustness is created. On the other hand, since the human auditory system is very sensitive to additive white Gaussian noise, and this causes to limit the number of least significant bits that can be changed without changing the sound quality. It can be begun to embed watermark data from the first sample, but in this case the watermark data is not secure. To increase the security, a pseudo-random sequence can be used to select the sample that we want to embed the watermark data in them. Many ways are provided to improve robustness and reduce noise from adding watermark. In [13] genetic algorithms are used to choose the best layer for embedding watermark bits per sample, watermark data bits are embedded in random layers and above the layer of the LSB, which the embedding in the top layer causes to increase the robustness against additive noise. In [14], a fixed layer was used for embedding the watermark bits and other bits of the sample are altered so that they have the smallest difference with the original sample. To reduce the noise generated in the silent part, the method of reducing error was used and error of the watermark sample embedding is distributed on 4 next samples. It can be embedded the watermark in six layers with this algorithm and can be created an embedding error of as much as four layers embedded in the base LSB. In [11], significant bits were applied to select the embedding layer and the desired number of samples.

5.1.2 Echo data hiding

One method of steganography is the embedding bits of the message in an audio signal as adding the artificial reverberation. This technique is known as a method of Echo hiding steganography. In hiding information with echo, information was embedded in the host signal by voice echo. Information will be hidden by changes in these three features; the initial range, lowering speed and delay. When the distance between the original signal and the echo signal decreases, both of them mixed together. At a certain point, the human ear cannot distinguish the difference between the two signals, which in this case, echo is heard as a resonant at the original signal. It depends on factors such as the quality of the original sound recording and quality of the listener. The adjustable parameters of Echo data hiding are shown in the Fig. 4.

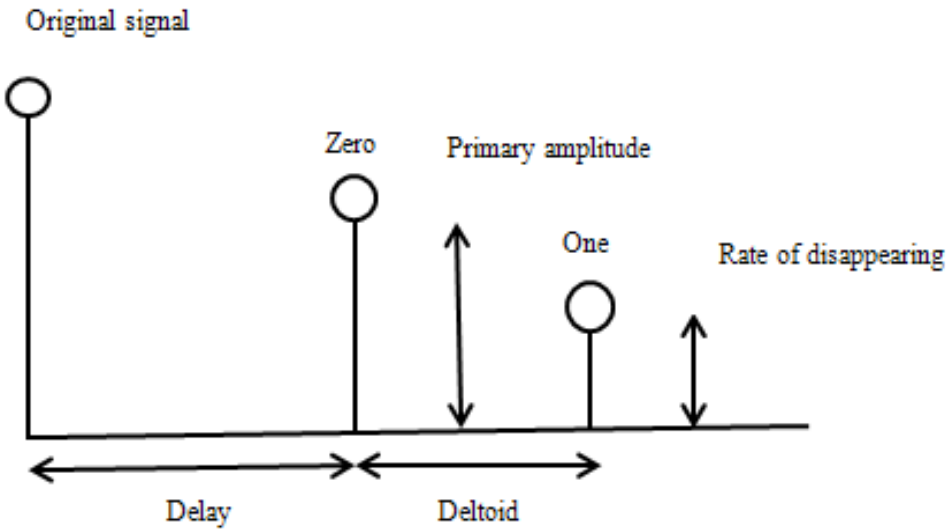

Fig 4: Adjustable parameters of Echo data hiding

As the delay (offset) between the original signal and the echo decreases, the two signals are merged, and the human ear cannot distinguish between the two signals at a certain point. Therefore echo is understood as an added resonance. It is difficult to determine the exact point and it the quality recording of the original audio, the type of sound to make echo and the listener is depended. We calculate this delay bout 1 millisecond for many voices and listeners. Coder uses two delay times; one to display binary digits 1 (offset) and another for showing binary digits 0 (offset + delta), so the process of embedding has the two modes. Both of them are located in subliminal latency where the human ear cannot detect and differentiate them [14].

5.1.3 Balance coding

Balance coding is a steganography technique for strong voice. Instead of breaking a signal to unique samples, this method breaks a signal to the discrete samples, and the embedding every bit of secret messages of one bit creates the balance. If the balance bit of a selected area has not been matched with together, confidential bits are encoded. Reverse process of the LSB is one of examples in the region. Therefore, transmitter has a more of one selection to encode secret bits.

5.2 Hiding in the transformation range

5.2.1 Phase coding

The human ear is not sensitive to absolute shifting the phase but for relative shifting is very sensitive. Phase coding uses this feature to hide the data. First, the original signal $\boldsymbol{C_0}$ is broken down into $\boldsymbol{M}$ blocks that lengths (N) of each block as much as twice the number of bits of data. Then, the matrix of phases $\boldsymbol{F_{0j}[\omega_k]}$ and the matrix of amplitudes $\boldsymbol{|A_{0j}[\omega_k]|, 0 \leq k \leq \frac{N}{2} - 1}$ for all the blocks are calculated. The matrix with the differences in phase between the $\boldsymbol{M}$ neighbor blocks is computed:

$$\boldsymbol{\Delta F_{oj+1}[\omega_k] = F_{0j+1}[\omega_k] - F_{0j}[\omega_k], \forall j, k} \quad (\mathbf{1})$$

The data is encoded in phase spectrum of the first block:

$$F_{w0}[\omega_k] = (-1)^{m[k]+1\frac{\pi}{2}}, \quad (2)$$

$$\text{for} \quad m[k] \in \{0,1\}, \quad 0 \leq k \leq \tfrac{N}{2} - 1$$

In order to ensure the inaudibility of the phase changes between the individual blocks, the phase differences in each of the blocks have to be adjusted:

$$F_{wj+1}[\omega_k] = F_{wj}[\omega_k] - \Delta F_{oj+1}[\omega_k], \forall j, k \quad (3)$$

Finally, an inverse Fourier transform is carried out on the modified phase spectrum F_w and the original domain $|A_0|$ to restore the audio signal. In order to detecting is required that the data length be known at the receiver. After synchronization on the first block, the first block Fourier transform is calculated and bits are read from phase information of first block. Another way to manipulate the phase is the phase modulation. In phase modulation, the original signal is divided into blocks and data is embedded in each block by phase modulation. Fig. 5 shows the signals before and after Phase coding procedure.

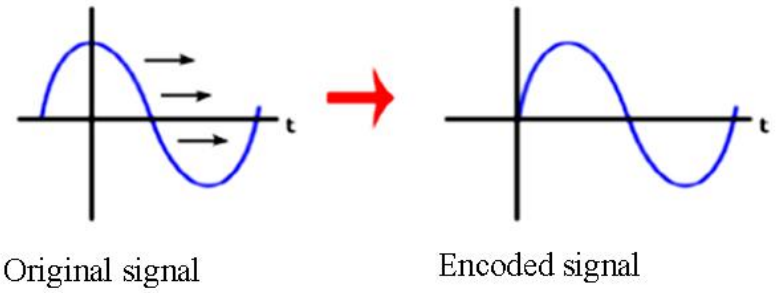

Fig 5: The signals before and after Phase coding procedure

5.2.2 Spread spectrum

In this way, a pseudo-random noise code that is independent of the data, to be used as a modulation wave to spread the signal energy in a much larger bandwidth to the bandwidth of information. On the receiving side with the help of the pseudo-noise code, signal becomes narrow. Sequence of spread spectrum can be added to voice samples of hosts in the time domain, the Fourier fast transform (FFT) coefficients and or wavelet domain. If the embedding is carried out in the conversation domain, watermark should be located at coefficients that are fixed against common attacks such as compression of range, re-sampling, low pass filtering and other common techniques of signal processing. In this method, data spreads on the many number of coefficients and distortion with the help of the masking effects of the human auditory system to be kept below the minimum perceptible distortion. Fig. 6 shows a general model for a spread spectrum.

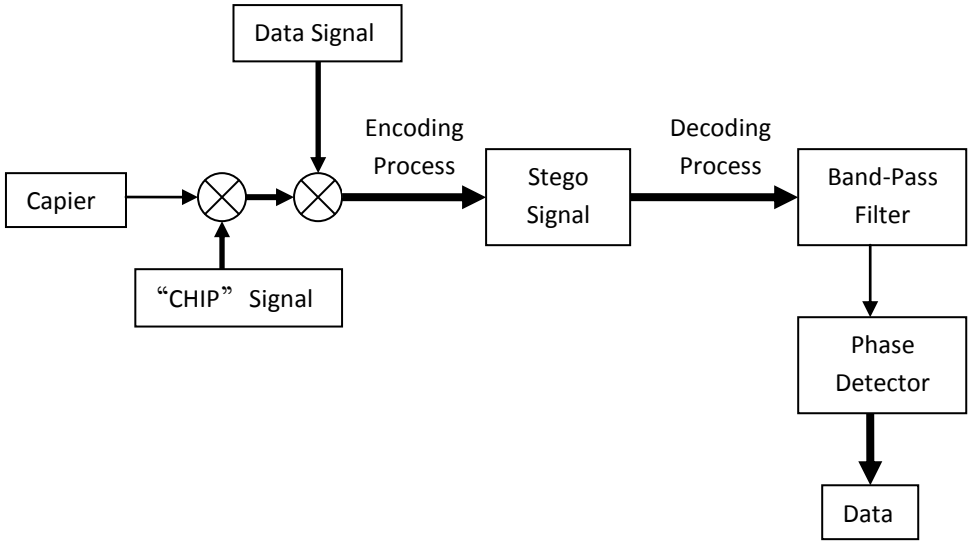

Fig 6: General model of spread spectrum [15]

There are two types of technology of spread spectrum that are called direct sequence spread spectrum (DSSS) and frequency hopping spread spectrum (FHSS). DSSS is a technology that applies a pseudo-random noise code which is independent of the signal that will be transmitted, and it is used to multiply or spread narrow-band signal into a signal with a wider bandwidth. On the receiving side, the correlation between the pseudo noise (PN) sequence used in the transmitter and broad-band signal to calculate the original narrow-band signal is applied. Other technology is spread spectrum FHSS, which the carrier frequency jumps from one channel to another at different times, and so it can prevent the entry of noise in some frequency. FHSS system can usually carry more data than to DSSS system, because the signal is narrow-band. Furthermore, the attack on the FHSS system compared to DSSS system is more difficult, because in FHSS system, the competitor must find both the time and frequency slots to synchronize.

5.2.3 The domain of discrete wavelet transform

Audio steganography in domain of discrete wavelet transform (DWT) by [16] is presented. Information with signal was embedded in the LSBs of wavelet coefficients of audio signals. In order to improve the security of embedded data, [17, 18] applied a hearing threshold during information steganography in integer wavelet coefficients, while in the [19] to avoid the hiding in the silent parts of the audio signal is expressed. Furthermore, an algorithm was carried out to increase the capacity and security of a method based on discrete wavelet packet transform with adaptive embedding in the least significant bits. Therefore, the algorithm determines the number of bits that can safely hide in each sample.

Table 3. Comparison of audio steganography techniques

Methods	**Weakness**	**Firmness**	**Data hiding technique**
LSB	Easy to extract	Simple and easy to hiding information	LSB of each sample at audio is embedded by a bit of hidden information
Echo hiding	Low security of information and low capacity	Without the problem of additive noise	hiding information with introducing echo at cover signal
Balance coding	Easy to extract	More powerful than LSB	Change LSB of balance bit of samples
Discrete wavelet transform domain	Data recovery with losing	To make the high embedding capacity and clarity	Changing wavelet coefficient to hiding information
Spread spectrum	More bandwidth occupation	Best firmness and increasing the clarity	Spread information under all signal frequencies
Phase coding	Low capacity	Stable against signal processing options	To fluctuate the phase of cover signal

6. A REVIEW OF THE METHODS PROVIDED FOR AUDIO STEGANOGRAPHY

Because of the importance of information steganography in relation to the e-commerce as well as issues of safety and security in urban areas, the need for information and knowledge of steganography capabilities to deal with the vandalism that is possible through cryptography is felt. Depending on the application, researchers have suggested the different methods of audio steganography in the spatial and transformation domains [20, 21, 22]. One of these methods is observed in [23], the implementation of an audio steganography using Arduino cards is presented due to the use of minimum success of the LSB technique on a special communication channel. In [11], a new method to improve conventional LSB techniques for audio steganography against attacks of hidden breaker was presented, which expressed two steps to improve the usual technique of the LSB. The first was number of random bits of host messages to embed a secret message, and the second was the number of random sample containing the next confidential message bit. In [23], the combination of steganography and cryptography were used as an efficient tool with high security level. Also in [24], a method was proposed to increase the confidentiality of the secretive message of which can be noted Huffman coding, RSA cryptography, and dual random LSB. Table 3 gives the comparison of audio steganography techniques.

7. DISCUSSION

According to opinions expressed in this paper, information steganography can be used in many places, including in this century, among them can be referred to the security and intelligence agencies, private companies, government organizations (such as police, etc.) and or even personal uses. The object of steganography is the embedding information on a host of media, so that it causes the least suspicion about the existence of the message in the host mask. A successful algorithm is an algorithm that in addition to maintaining visual quality than its original sample has enough resistance against attacks to extract information. According to the studies that was done about audio steganography, in the most of the work; especially in recent years have used the LSB method. In the majority analyses of each of the studies, confidentiality and authentication and sustainability have not been considered together, of course, the target of each procedure is to increase the quality of steganography. We can make a good balance between quality and security in addition to the confidentiality and authentication and sustainability. Using meta-heuristic algorithms can help to increase storage capacity and security of steganography. Therefore, it can be used a method based on genetic algorithms to protect confidential information for audio steganography with high-capacity embedding. Because data by manipulating bits are embedded, therefore, this method acts in the spatial domain. In this way, information hiding capacity increases and also maintaining the hidden information in an image will increase. The objectives of the proposed method are: 1) decreasing the error difference between the cover and watermarking audio, 2) maintaining the low computational complexity and high performance in the proposal, 3) resistance against attacks with better quality and a good balance between safety and quality. For extra security, the logical operation XOR and for the sufficient balance between security and the quality, the optimization algorithms are used.

8. CONCLUSION

According this study based on steganography techniques in an audio cover with five different methods can be notice two viewpoints related to the steganography: 1) functionalist view and 2) communicational view. The first made closer it to the cryptography field using the technique domain of steganography and its nature is introduced as a security tool, and the second try to transfer concept to the special second person with this definition that the media is a message itself. This paper has a comparative state to some steganography methods provided. Considering what was said, an idea (project) for steganography as a whole, without implementation and other issues was propounded. Therefore, in the future works a new method of steganography with meta-heuristic algorithms or the combination of cryptography and steganography methods were offered based on first steps tried to guarantee the security of information, because, security of steganography procedures would be gradually reduced during time and advancement of computational techniques. For this reason, it needs to provide new methods in this field at any time.

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interests regarding the publication of this paper.

9. REFERENCES

[1] M. S. Subhedar, V. H. Mankar, Current status and key issues in image steganography: A survey, Computer Science Review, 13-14 (2014) 95-113.

[2] B. Chandel, S. Jain, Video Steganography: A Survey, IOSR JCE. 1 (2016) 11-17.

[3] N.F. Johnson, S. Jajodia, Exploring steganography: Seeing the unseen, Computer, 31 (1998) 26-34.

[4] R.B. Wolfgang, E.J. Delp III, Fragile watermarking using the VW2D watermark, In electronic imaging, International Society for Optics and Photonics, 1999, April, pp.204-213.

[5] M.D. Swanson, M. Kobayashi, A.H. Tewfik, Multimedia data-embedding and watermarking technologies, Proc. IEEE. 86 (1998) 1064-1087.

[6] D.C. Kar, C.J. Mulkey, A multi-threshold based audio steganography scheme, JISA.23 (2015) 54-67.

[7] M.B. Begum, Y. Venkataramani, LSB based audio steganography based on text compression, Procedia Eng. 30 (2012) 703-710.

[8] S. Banerjee, S. Roy, M.S. Chakraborty, S. Das, A variable higher bit approach to audio steganography, International Conference on Recent Trends in nformation Technology (ICRTIT), 2013, July, pp. 46-49.

[9] K. Bhowal, A.J. Pal, G.S. Tomar, P.P. Sarkar, Audio steganography using GA, International Conference on Computational Intelligence and Communication Networks, 2010, November, pp. 449-453, IEEE.

[10] P. Jayaram, H.R. Ranganatha, H.S. Anupama, Information hiding using audio steganography–a survey, IJMA. 3 (2011) 86-96.

[11] M. Asad, J. Gilani, A. Khalid, An enhanced least significant bit modification technique for audio steganography, International Conference on Computer

Networks and Information Technology (ICCNIT), 2011, July, pp. 143-147,IEEE.

[12] N. Cvejic, Algorithms for audio watermarking and steganography, Department of Electrical and Information Engineering, Information Processing Laboratory, University of Oulu, 2004.

[13] G. Nehru, P. Dhar, A detailed look of audio steganography techniques using LSB and genetic algorithm approach, IJCSI. 9 (2012) 402-406.

[14] N. Cvejic, T. Seppanen, Increasing robustness of LSB audio steganography using a novel embedding method, International Conference on Coding and Computing (ITCC), 2004, April, pp. 533-537, IEEE.

[15] W. Stallings, Data and computer communications, 8th Edition, Pearson Prentice Hall, 2007.

[16] G.P. TVS, S. Varadarajan, A novel hybrid audio steganography for imperceptible data hiding, International Conference on Communication and Signal Processing (ICCSP), 2015, April, pp. 0634-0638, IEEE.

[17] A. Delforouzi, M. Pooyan, Adaptive digital audio steganography based on integer wavelet transform, Circ. Syst. Signal Pr. 27 (2008) 247-259.

[18] S. Nehete, S.D. Sawarkar, M. Sohani, Digital audio steganography using DWT with with reduced embedding error and better extraction compared to DCT, Proceedings of the International Conference & Workshop on Emerging Trends in Technology, 2011, February, pp. 167-168, ACM.

[19] H.I. Shahadi, R. Jidin, High capacity and inaudibility audio steganography scheme, 7th International Conference on Information Assurance and Security (IAS), 2011, December, pp. 104-109, IEEE.

[20] Y. Kakde, P. Gonnade, P. Dahiwale, Audio-video steganography, International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2015, March, pp. 1-6, IEEE.

[21] M. Tayel, A. Gamal, H. Shawky, A proposed implementation method of an audio steganography technique, 18th International Conference on Advanced Communication Technology (ICACT), 2016, January, pp. 180-184, IEEE.

[22] V. Sharma, R. Thakur, LSB modification based audio steganography using trusted third party key indexing method, Third International Conference on Image Information Processing (ICIIP), 2015, December, pp. 403-406, IEEE.

[23] M. Tayel, A. Gamal, H. Shawky, A proposed implementation method of an audio steganography technique, 18th International Conference on Advanced Communication Technology (ICACT), 2016, January, pp. 180-184, IEEE.

[24] J. Vimal, A.M. Alex, Audio steganography using dual randomness LSB method, International Conference on Control Instrumentation Communication & Computational Technologies (ICCICCT), 2014, July, pp. 941-944, IEEE.