MP3 Steganography: Review

Mohammed Salem Atoum[1], Subariah Ibrahim[2], Ghazali Sulong[3] and Ali M-Ahmad[4]
[1,2,3,4] Faculty of Computer Science and Information Systems ,
Universiti Teknologi Malaysia
Skudai, Johor bahru, Malaysia 83100

Abstract
Steganography has existed as the science and art for hiding information in a way that the secret message cannot be deciphered by others, except the sender and receiver. All digital files such as audio, image and text files can be utilized for hiding secret information. Audio file can provide a good hiding medium because of its high data transmission rate and high degree of redundancy. Many formats such as MP3 have been utilized in audio information hiding, however to date there are limited review on the use of MP3 file as an audio file information hiding format. In this paper, we present a thorough analysis on the techniques used in audio files technologies, with more emphasis on MP3 steganography technique. This is to make available comprehensive information on the strengths and weaknesses of the MP3 file steganography techniques and detail comparison information for the research community on information protection.
Keywords: *Steganography, MP3, Information Hiding, Secret message, LSB.*

1. Introduction

The increasing Internet usage stems from the growing availability of the global communication technology that has led to electronically induced information gathering and distribution. However, the challenge it presents in terms of information security is enormous. Every Internet user interest lies in having a secure transaction, communication and information across the transmission link, but in reality, much communication are infiltrated, jabbed and altered. Information confidentiality was enacted by the CIA as one of the key principles of a secure communication and if abused attracts penalty. However, many communications still fall short of achieving a secured information transmission across the global network (the Internet). The need to secure information within the global network is of paramount importance so that user information is preserved until it reaches its destination undisclosed.

A lot of sensitive information goes through the Internet on frequent basis. This information could be military codes, government dealings, and personal data, the route, sender/ receiver, the content of such information requires that they are protected against hacking and infiltration. Therefore, providing a secure framework that conceals information content and sender/receiver identity should be an urgent matter of interest. There are two known approaches to information confidentiality; they are cryptography and steganography [1]. Cryptography has long existed as the method for securing data; it works with set of rules that transforms information into unrecognizable format. The rules are used to serve for authentication purposes, because only the one who knows the rules can decipher the encrypted information [2]. The advent of steganography provides more security features since the information is disguised in the sense that the information does not give away its content and identity of sender and receiver within the communication link. Cryptography and steganography techniques both make use of data encryption approach but Cryptography encrypts plainly its secret message thereby making the content and the user's details vulnerable to exploitation. Steganography technique protects both information content and identity of a person's transmitting the information, whereas only information is concealed with cryptography [3].

Steganography operates by embedding a secret message which might be a copyright mark, or a covert communication, or a serial number in a cover such as a video film, an audio recording, or computer code in such a way that it cannot be accessed by unauthorised person during data exchange. A cover containing a secret data is known as a Stego-object [3]. After data exchange; it is advisable for both parties (sender and receiver) to destroy the cover in order to avoid accidental reuse. The basic model of a steganographic system is shown in the Figure 1 [4]. The model contains two inputs and two processes, the inputs are a cover medium and secret message both can be any image, audio, video and so on. Two processes contain embedding and extracting processes.

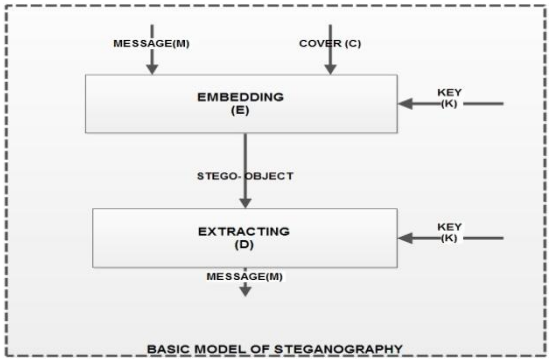


Fig. 1 Basic Model of Steganography

The embedding process is used to hide secret messages inside a cover; the embedding process is protected by a key if using secret key steganography and public key Steganography types, or without a key if using pure Steganography type. When using key, only those who possess the secret key can access the hidden secret message, while the extracting process is applied to a possibly modified carrier and returns the hidden secret message. Until recently, steganography utilized image files for embedding information across the Internet network. However recently, its use has been extended to audio steganography. The usage of audio signal as an embedding platform for information hiding is due to the fact that it has sophisticated features that allow information hiding, though difficult its robustness counts. In audio steganography, various signal processing techniques can be utilized to hide information in an audio file in such a way that it cannot be visually interpreted [5]. This approach has brought about the growing research interest in the use of digital audio signal for embedding information. The sensitiveness of audio files to delay presents more challenges to the design objective of steganography. There are three fundamental properties to the design of steganography. They are: 1) imperceptibility, 2) robustness and 3) capacity. However, there are other properties such as computational time that must be considered when dealing with different types of applications (information) such as broadcast monitoring applications in a global network. In most cases it requires real time processing and thereby cannot tolerate any form of delay [6].

The rest of this paper is organized as follows: A detailed introduction of audio steganography and MP3 file structure, the existing methods for MP3 steganography, then discussion and conclusion.

2. Audio steganography

The techniques of Steganography were originally developed and used for images. Researchers in the field then started studying on how the techniques can be used on audio media. Hence, the introduction and development of the known algorithms for audio steganography was founded. As the known steganography techniques are mostly used for images there are not many methods for audio steganalysis. Thus, audio Steganography provides considerably better security [5].

In audio Steganography, many types of file can be used as a cover of steganography such as Waveform Audio File Format (WAVE, or more commonly known as WAV due to its filename extension) or MPEG-1 or MPEG-2 Audio Layer III (MP3). Similarly, secret messages that are embedded can be of secured types such as text or speech. MP3 is the most popular compression format for digital audio. In steganography, which uses MP3 as a cover, secret message can be embedded during compression and after compression [12-13].

This section explains and discusses MP3 file structure, MP3 encoding and MP3 frames header

2.1 MP3 file structure

The content of MP3 files depends on the type of encoding used. The common structures of MP3 files consist of three components, they are Tag, Padding bytes and Frames, and these are shown in Figure 2.

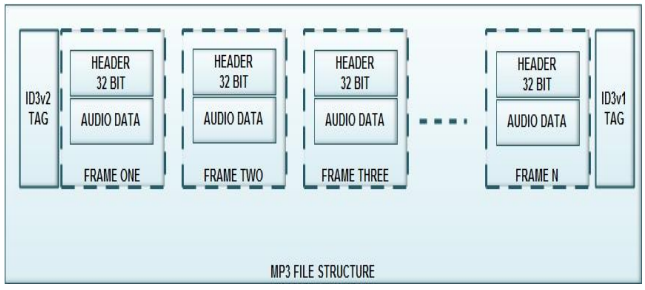


Fig. 2 MP3 File Structure

Tags are of two types, the ID3v1 and the ID3v2, and the later usually utilize the end section of the file to post-pended information. The length of ID3v1 is 128 bytes separated to seven fields which composed of the artist name, album, song title and genre. Its drawbacks are that it has a static size and also lack flexibility of implementation. Also, not all MP3 files accept the ID3v1 tagging system. However, its second ID3v2 presents a more adaptive standard since it is flexible and has a tagging system that pre-pended information before it is sent [7]. The ID3v2 tags consist of its own frames which are capable of storing various bits of data, for instance, the standard of character strings such as the artist name, song title or more advanced information concerning the way the file was programmed are all the data that is embedded in the signal. The advantages of the ID3v2 tags are that

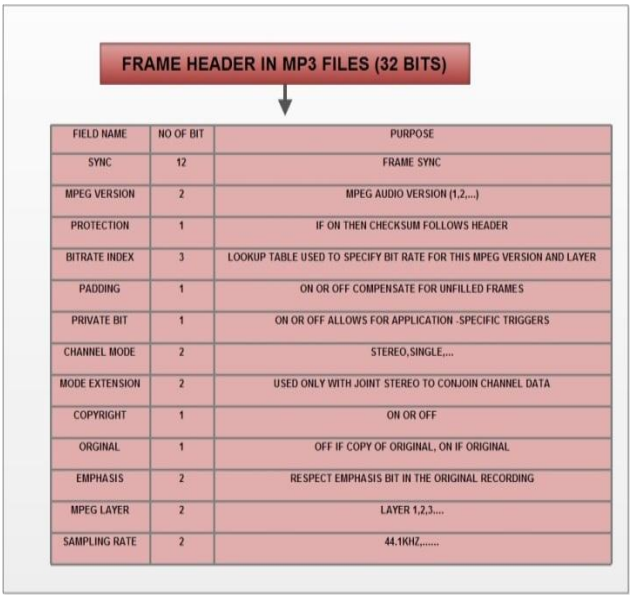
useful hints to the decoder are provided prior to transmission and that there is no size limit to information capacity provided in its pre-pending system [8].
The Padding byte provides additional data embedding; the data provided are added to the frame. Its working principle is that on the event of the encoding, additional data are evenly filled to the frame; this byte can be found in Constant Bit Rate (CBR) so as to ensure that frames are of the same size [9].

2.2 MP3 Encoding

MP3 encoding refers to quality enhancer of both compressed sound and the size of compressed sound file or compression ratio. The three encoding bit rates used by different encoders include the CBR, Variable Bit Rate (VBR) and Average Bit Rate (ABR).
CBR refers to a standard encoding mechanism used by basic encoders. In this encoding mechanism, each frame used the same bit rate in the audio data. The bit rate is fixed in the whole of MP3 file, as the same number of bits is used for each part of the MP3 file. However, the quality of MP3 is variable. These techniques can be used to predict the size of encoded file and can be calculated by multiplying the bit rate chosen to encode with the length of a song [10].
VBR is a technique that can keep the quality of audio files during the encoding process. In this technology the quality of the sound can be specified but the size of the sound file remains unpredictable [10].
ABR is a mode that uses higher bit rate for the part of music by choosing the encoder adds bits. The final result showed that the quality is higher than CBR. Moreover, the average file size remains predictable [10].

2.3 MP3 Frame Headers

MP3 frame header consists of bits 0 and 1, it can either start with 0 or 1. In most cases a frame header is always set to one (1) and if in that state it is referred to as block synchronization, see Figure 3 for illustration.
The sync is a series of bits that represents a header. These bits make up the frame which composed of 12-bits for the ones. The frame does not necessarily need to have unique headers in order to establish longer data block. However, some conditions must have to be followed in order to recognize a long byte data block such as 4 byte data block as a header. A more detailed discussion on the conditions of deciding on how to determine a header was investigated by [9]. For instance, the 4 byte block starts through the Sync [11] and does not violate any of the conditions stated by [9]. Although, the frame size cannot be easily determined, some approaches that utilize the beginning and the end of the frame are implemented. This is achievable only if the headers of the frame are identical in terms of structure and content.



FIELD NAME	NO OF BIT	PURPOSE
SYNC	12	FRAME SYNC
MPEG VERSION	2	MPEG AUDIO VERSION (1,2,...)
PROTECTION	1	IF ON THEN CHECKSUM FOLLOWS HEADER
BITRATE INDEX	3	LOOKUP TABLE USED TO SPECIFY BIT RATE FOR THIS MPEG VERSION AND LAYER
PADDING	1	ON OR OFF COMPENSATE FOR UNFILLED FRAMES
PRIVATE BIT	1	ON OR OFF ALLOWS FOR APPLICATION .SPECIFIC TRIGGERS
CHANNEL MODE	2	STEREO,SINGLE,...
MODE EXTENSION	2	USED ONLY WITH JOINT STEREO TO CONJOIN CHANNEL DATA
COPYRIGHT	1	ON OR OFF
ORGINAL	1	OFF IF COPY OF ORIGINAL, ON IF ORIGINAL
EMPHASIS	2	RESPECT EMPHASIS BIT IN THE ORIGINAL RECORDING
MPEG LAYER	2	LAYER 1,2,3....
SAMPLING RATE	2	44.1KHZ,.....

Fig. 3 MP3 Frames Header

The volume of a specified frame in bytes can be obtained using the following equation [9];
Frame Size = 144* Bitrate / (Sample Rate + Padding) (1)
Bit rate is measured in bits per second, Sample Rate refers to sample rate of the original data and Padding is the additional data that is added to the frame during the encoding process so as to evenly fill the frame [9].

3. MP3 Steganography Methods

The methods of embedding information in MP3 audio file can be divided into two ways, the information are either embedded while the information is compressed or after compression of information is done [12-13].

3.1 Embedding during Compression

Embedding the information while compression of the information is executed can be achieved through four different ways. They are the least significant bit (LSB), stage coding, echo hiding, and Spread Spectrum (SS) [9].

3.1.1 Least Significant Bit (LSB)

The Least significant bit technique is about the first and simplest technique that was used to embed secret messages to audio files or other form of material medium of securing information. The LSB technique works by converting audio file and secret messages into stream of bits, the bits from secret messages are then subsequently embedded into the audio file (which can be of any type). This is done by changing the LSB bit of audio by one or more bits to tune

with the secret message after which transmission is enabled to the receiver [14]. This procedure is done in order that the secret message is received at its destination undisclosed. At the receiving end, the receiver extracts the secret message using sequence of sample indices used in the embedding process. A maximum of 1 Kbps per I KHZ is the required capacity for audio files and it is one of the limitations of the LSB. The work of [15] proposed that before the secret messages are inserted in audio signals, that the higher bit indices should be used as the alternative to the traditional LSB. This is because this approach does not transform the volume of the file securely transmitted and it is as well suitable for any type of audio file format, unlike the conventional LSB. However, some other parameters can be set for the quality of sound which depends on the size of the audio to be transmitted and the length of the secret message. The work presented by [16] recorded a significant contribution in LSB capacity. In the same year [17] proposed and adopted an algorithm for increasing capacity of LSB technique. Their algorithm presented a method whereby the secret message bits are inserted into multiple and variable LSBs in this approach using 16 bits per sample. The approach recorded up to 7 bits LSBs compared to the conventional 4 bits LSBs and achieved a capacity of 5.563 bits per sample. They made further progress by adopting an adaptive approach that utilized only the MSB of the cover samples in a way that if it is 0, 6 bits LSB is used otherwise 7 bits LSB (when in 1 state) is used. This adaptive approach recorded success of up to 6.574 bits per sample.

A trade-off between robustness, capacity and imperceptibility that audio Steganography technique requires was investigated by [18]. Their work reported that the trade-off between noise acceptance and capacity is dependent on higher bit indices which inherently results in imperceptibility of the embedded secret message. However, a previous report by [19] proposed that lifting scheme produces an ideal rebuilding filter banks such as the Int2lnt, and thus decreases the fault rate in wavelet domain steganography. A resulting outcome of their work was fewer than 100 Kbps and the capacity of up to 200 Kbps was achieved. Another effort to address robustness of LSB is the work of [20-21]; they utilized genetic algorithm (GA) and RSA respectively, for the substitution technique in order to minimize computational time of conventional LSBs. The LSB technique through GA was utilized to embed the encrypted secret message into multiple, vague and deeper layers of audio signals to achieve higher capacity and better robustness in transmission of secret messages. Although LSB is simple, but the inefficiency of the LSB coding increases the signal to noise ratio of the sound file. Moreover through the process of using either LSB (one layer) or LSB (multi-layers) the embedded encrypted information are most likely to get lost [21]. However the efficacy of genetic algorithm was disputed by [20]. In their work RSA algorithm was utilized in place of the GA algorithm because the GA has recorded increase in noise accumulation.

Overall, robustness of the LSB can be achieved through the implantation of a redundancy technique alongside the encoding of the encrypted secret message. This redundancy technique is a promising approach to the reduction of transmission rate for LSBs.

3.1.2 Phase Coding

The phase coding technique was introduced to address the problem of LSB as their noises accumulate during information hiding. In Phase coding, the noise level is not obvious to the human hearing. This technique presents an alternative approach to entering technical bits when secret messages are embedded to the audio signal. It works to encode secret messages in stage shifts for the stage spectrum of digital signals, so as to achieve indiscernible noise to signal ratio. In Figure 5 the signal encoding processes of the phase coding are illustrated [14].

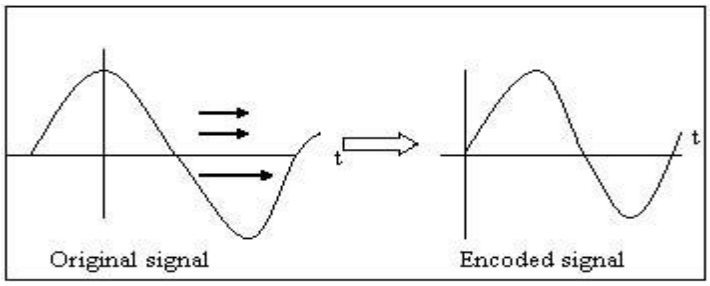


Fig. 4 Phase Coding [14]

The following steps are the coding stages implemented by the phase coding technique during encrypted secret message embedding [14]:

- Separate the original sound signal into smaller segments of lengths equal to the size of the secret message that is to be encoded.
- Apply Discrete Fourier Transform (DFT) to each segment in order to create a matrix of stages and Fourier transform magnitudes.
- Calculate Stage difference among adjacent segments.
- Embed an undisclosed secret message in the first segment of audio file by using the following equation

$$phase\_new = \begin{cases} \pi/2 & if \ \ message \ \ bit = 0 \\ -\pi/2 & if \ \ message \ \ bit = 1 \end{cases} \quad (2)$$

- Compute new stage matrix by using new stage of the initial segment and original stage dissimilarities.

- By using new segment matrix in addition to new magnitude matrix, reconstruct the sound signal with reference to the opposite DFT and then simultaneously, re-concatenate the sound fragments.

To extract the undisclosed secret message from audio file, the receiver uses the distance between the end to end of the piece and the DFT to get the stages, in order to extract information. Since the inception of phase coding till this day, much research effort is directed towards a robust steganography technique. In this regard, the phase coding has recorded contribution in its robustness to noise but we will mention very few progresses made so far with phase coding steganography technique. The report of [22] introduced stage shifting scheme within audio signals. The approach proposed on trimming down the correlation complexities that exists as a result of PN indication per each sub-band and the undisclosed secret message. The problem of their scheme lies in achieving quality signal at the receiving end. However, if signal processing is done before hand in such a way that the original audio signal is processed before embedding secret messages a more effective scheme will be achieved. This approach will yield more robustness than approaches that did not take into consideration the pre-processing of signals before they are used to embed the information to be transmitted.

[23] Developed a novel technique for stage coding that inserts data bits using changed stage modification on simultaneous basis for more capacity and robustness of stage coding. Their approach was of two stages, the first stage being the selection of the frequency that was varied through each frame of indicator. Though very promising but it provides an enabling attack space, since the information can be sniffed to give away initial spot of bit flow even after a flow of frames have been transmitted. The second stage of their approach records the difference between the stage values of end-to-end frequencies, which was used to insert undisclosed data in its place of complete stage value. The advantage of their work was to reduce noise level in the processed signal.

Having observed different literature approaches to achieve capacity and robustness of the phase coding, researchers deduces that the selection of stage as an alternative to the use of the amplitude of the signal increases the noise resistance of the Stage signal. It is also important to note that if the degree of difference between the end to end frequencies of stage encoding is calculated beforehand, the performance of the steganography technique can be tuned to achieve better performance. More also with increased signal length a success in the transmission of the signal can be achieved without loss of information.

3.1.3 Echo Hiding

Echo hiding involves the hiding of information in the audio file by creating an echo with a separate signal. It is similar to the spread spectrum in that there is an increase in data transfer rate and robustness compared to other methods. The echo hiding approach does not induce noise to the signal especially at the embedding process that most approaches sustain noise [14]. The embedding process for echo hiding utilizes three parameters that include amplitude, decay rate, and offset (delay time) from the original signal to predict the transmission process. A threshold level is set for all the parameters using human hearing information to fine tune the system. The challenge of this approach lies in the difficulty to discern the echo from the main signal. The binary offset used in echo hiding are of two types: 0 and 1 bit. The varied signal can either take the 1 bit that represents the offset value or the 0 bit that represents the binary offset [14]. Figure 5 shows the parameters and their thresholds.

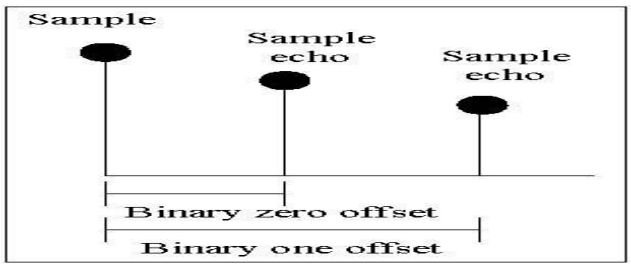


Fig. 5 Binary offset [14]

The original signal is broken down into blocks with the use of the parameters and its thresholds as the encoding progresses. The challenge in echo hiding lies with the fact that there is only one original signal amidst all the block of signals that are simultaneously encoded until the last signal is completed. At the receiver end the information are deciphered through the signal cluster in sequence of steps used to encode them till the original signal is obtained. A robust approach to deciphering the information was investigated by [24]. The approach utilizes the function of automatic correlation of the spectrum of the signal to decode. The work of [25] proposed ways to revise and evaluate a number of alternates for echo hiding. Using T-codes the secret message rescue rate of echo hiding are developed. The T-code is a subset of all possible Huffman code sets, a set of self-coordinating codes. Though promising but the limitations of this approach are that a weak structure of security is created and its file decrypting process requires prior information of the coding progression in order to achieve decrypting.

3.1.4 Spread Spectrum

In the Spread Spectrum (SS) approach to information hiding, the information is embedded in the audio file by intersecting the spectrum of the undisclosed secret

message with the spectrum of the audio signal. This makes the SS similar to LSB coding since the undisclosed secret message bits are spread randomly within the entire audio file [26]. However, the undisclosed secret message bits spread in the bits of the audio file differs from that of the LSB coding. This is for the fact that the undisclosed secret message spectrum is inserted in frequency spectrum of the audio file by using the code that is self-sufficient of the actual signal with the information. This approach presents challenges to the bandwidth capacity of the transmission link [27].

Audio steganography has recorded two versions of Spread Spectrum techniques, they are: the direct-sequence spread spectrum (DSSS) and frequency-hopping spread spectrum (FHSS). The DSSS works in a way that the secret message is distributed through an unchanging process determined by the chip rate. Subsequently, the signals are adjusted using a pseudo-random generated signal. In FHSS approach, the recurrence spectrum of the audio files is altered, this is due to the fast bounds that exist among the frequencies [26]. Since efficiency is the limiting factor for spread spectrum techniques, the spread spectrum has attracted a lot of research contribution with respect to its efficiency. Notable modification registered for SS is the quantized spectrum values of audio layer III that embeds undisclosed information into audios. The efficiency of the SS used in literature are derived from the technology of spectrum shift, were the distributions of human speech are located in unheard places of music spectrum. Another means of achieving SS efficiency are derived through the spread of the spectrum when harmonics of human voice and music are sorted from each other. In this way, conversion of the audio communication system is achieved based on SS data that is designed to conceal the secret message. The DSSS technique has also been enhanced through aspects of the frequency aimed at improving the carrier frequency of the binary stage shift keying (BPSK) signal [27].

Prior to the work of [27-28] attempted to address the efficiency of the SS through their proposed approach that embeds undisclosed secret message in audio signal using SS in a sub-band domain approach. Their aims were to improve on the transparency and robustness of the SS technique. For the transparency feature, modification is made to psycho-acoustic representations of the signal during audio compression to manage the loudness of announced deformation when embedding is carried out. The robustness feature in terms of sturdiness was achieved through determining the attribute collection and organization difficulty of the hiding process in order to maximize the survivability of the data.

Deduced from the discussions on SS are that with the SS techniques the transparency of audio quality, good survivability and better embedding capacity can be achieved. More also, the SS technique can offer better data transmission rate and it has the ability to maintain high level of robustness in comparison to the LSB coding and phase coding techniques.

3.2 Embedding After Compression

The process of embedding information after compression has not been extensively researched because of the difficulty that has to do with embedding in a compressed signal. The compressions of signal before it uses a material medium for hiding information results in poor spread of secret messages in steganography process and likewise result in poorly transmitted signal, that is, the signal sound quality [9]. However, the other aspect of the embedding that a deal with embedding during compression is does not affect sound quality. This means that the signal to noise ratio of the technique of embedding the secret message during compression is low. Some example of the media files for embedding after compression is unused audio data and used audio data [9, 12, 29-30].

3.2.1 Embedding in Unused audio data

The frame of the unused audio data comprises of unused header bit, padding byte stuffing, before all frames and between frames.

3.2.1.1 Embedding in Unused Header Bit

The MP3 frame headers are made up of fields such as the private bit, original bit, copyright bit, and emphasis bit but its usage are mostly omitted in some MP3 players. These fields are the important aspect of the frame that aids the interpretation of information concealed within the audio signal. They can be properly used to embed undisclosed massage by replacing the bit stream of undisclosed massage through the bits in the field. However, if in the process of replacing the bit stream with the bit in the field fails, the actual content of the secret message received within the frame will be lost and that will make the signal recovery more challenging [9]. The work by [32] highlighted on the possibility that audio steganography can achieve good capacity and robustness through the use of 4 bits in each header frame of the audio signal to embed secret messages.

3.2.1.2 Embedding in Padding Byte Stuffing

Padding byte stuffing was recently established as one of the techniques for steganography. Its approach is relatively straightforward in terms of implementation. It represents fine regular storage capability and has the ability to program 1 byte of information for each frame as long as padding bytes are accessible. The MP3 file is a given example of the material medium that can well utilize the padding byte stuffing method because it can allow for

hundreds of frames in one secret message, especially when the filling bytes cannot take any more audio information [32].

3.2.1.3 Embedding in Before All Frames

Before all frames (BAF) was developed by [29]. Their approach embeds text file to MP3 file. The text file is encrypted by using RSA algorithm to increase the security of undisclosed secret message. The first frame will be filled with encrypted information. This process is repeated sequentially until the frame headers are filled. The capacity of about 15 KB is utilized when encryption algorithm is used otherwise it takes about 30 KB for the MP3 file. Even though there are chances of the secret message being sniffed, for this approach, its advantages are enormous, for instance, the method of padding and the unused bit even after the frames must have been filled, provides more encoding capability.

3.2.1.4 Embedding in Between Frames

[30] Developed steganography technique that embedded between frames (BF). It also embeds text file to MP3 file like the BAF and encrypts information in bits format by using RSA algorithm in order to increase the protection of concealed secret messages. The BF differs from the BAF in the way the text files are inserted into the frames. It does not start with the first frame it sees but selects the frame of its choice. On the other hand, the capacity of the BF in comparison to the BAF utilizes the capacity of about 40 MB with encryption algorithm but requires 80 MB on original format. BF likewise provides good capacity for embedding text file in more capacity but it is still prone to attack.

We draw inferences based on the literatures accessed that the method of embedding information after compression is a challenging task since the embedding process is done after compression and the text file are located in the unused bit location and not in the audio data. This technique provides a platform that is prone to attack because the content of the secret message sent can be easily deciphered by a third party sniffing through the communication link. It also provides only limited capacity for secret message hiding. However, if the LSB technique is used to insert speech in MP3 file with the use of 2, 3 and 4 bit exchange in audio data (8-bit for sample), the problem of capacity can be resolve. In addressing the problem of security, the use of key as the lock for concealed secret message is a foreseeable approach that can achieve maximum security for concealed secret messages.

3.2.2 Embedding in Used audio data

In [33-34] proposed the use of M4M and M16MA for inserting undisclosed secret message in audio data. These algorithms were developed based on M16M in Image types. The M4M is a mathematical function that maps 2 bit of the undisclosed secret message in the required slot in a precise manner using a pseudo random number for inserting secret message bits in a random process. The algorithm worked in a self-determining manner, that is, the nature of the data to be concealed is determined afore hand in order that the best approach of insertion was utilized. As a result the concealed audio signal was created through lowest amount degradation. The M16MA was also developed for deciding on the embedding location. It used some statistical function to map every 4 bit of the undisclosed secret message in the pre-determined locations. It also utilizes a pseudo randomly generated number to embed secret message bits to its location on random basis. As a result of its self-determining concept the least possible degradation of concealed audio signal was created.

4. Discussion

The capacity, robustness and imperceptibility requirement for Steganography are the important features that characterize the strength and weaknesses of the MP3 techniques for achieving information hiding. We further summarize the strength and weaknesses of the compression techniques that has been used so far for MP3 Steganography in a tabular format that is self-explanatory.

Table 1: Comparative between MP3 Steganography Methods

Methods	Techniques name	Summary	strength	weakness
Embedding During Compression	Least significant bit	•Is the oldest and simplest techniques in audio steganography its embed secret message in audio after convert the cover and message to bit stream and modified the least bit of audio bit stream with bit in secret message	•simplest way •large capacity	• Low robustness
	Phase coding	•Phase coding works by substituting the phase of an initial audio segment with a reference phase, this phase represents the hidden data. The phase of subsequent segments is adjusted in order to preserve the relative phase between segments.	• Robust against signal processing manipulation and data retrieval •needs the original signal	•low data transmission • complex
	Echo hiding	•in echo technique embedding data into host a host audio signal by introducing an echo; the hidden data can be adjusted by two parameters: amplitude and offset , the two parameters represent the magnitude and time delay for the embedded echo, respectively the embedding process uses two echoes with different offset, one to represent the binary datum" one" and the other to represent the binary datum "ZERO"	• Resilient to lossy data compression algorithms	•Low robustness • Low security and capacity
	Spread spectrum	•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	•High level robustness	•Can introduce noise •Vulnerable to time scale modification
Embedding After Compression	Unused bit	•In this technique the can embedded secret message in unused bit in headers of frame. In the frame header can founded two or three bit unused can embedding secret message in it.	•Simple	•Low robustness and security
	Padding byte stuffing	•In the MP3 frames can found some of byte stuffing uses this for make the all frame in MP3 is the same size. In embedding process can search to find this byte and replacement this byte with a byte in secret message.	•Efficient •Simple	•Not all MP3 found just in ABR or VBR. •Low robustness and security
	Between frames	•In this algorithm can embedding the secret message after the end of frame previously and before the start of next frame.	•High capacity. •Simple	•Low robustness and security
	Before all frames	•Before all frames technique can be embedding secret message before all frames start.	•Low capacity.	•Low robustness and security

5. Conclusions

This paper presented a review of existing techniques that has found usage in MP3 steganography. We gave a detailed presentation of the two approaches commonly used in MP3 steganography, which are embedding after compression and embedding during compression. The various techniques that have been proposed for achieving capacity, robustness and imperceptibility for the two approaches were discussed alongside with their strength and weaknesses. We drew conclusion on each technique based on the lapses that were observed for the proposed methods and suggested on a better approach that could likely offer better results. However, overall, success can be achieved for embedding after compression if the encoding and decoding processes are not executed during the embedding and extraction process. Finally, we propose that to achieve security of concealed information as it travels through the communication link the techniques that embed information after compression are the best.

Acknowledgment

This work was supported by Universiti Teknologi Malaysia(UTM), Johor, Malaysia under the VOT:Q.J13000.7128.00J29.

References

[1] Lentij J., "Steganographic Methods", Department Of Control Engineering And Information Technology, Budapest University. Periodica Poltechnica Ser. El. Eng. Vol.44, No. 3–4, P. 249–258 (2000), Url: Http://Www.Citesseer.Ist.Psu.Edu/514698.Html.

[2] Katzenbeisser S., Peticotas F., "Information Hiding Techniques For Steganography And Digital Watermarking", Artech House Inc.2000.

[3] Petitcolas F.A, Anderson R.J., Kuhn M.G., "Information Hiding – A Survey", Ieee, Special Issue On Protection Of Multimedia Content: 1062-1078, July, 1999.

[4] Cacciaguerra S., Ferretti S., "Data Hiding: Steganography And Copyright Marking", Department Of Computer Science, University Of Bologna, Italy, Url: Http://Www.Cs.Unibo.It/~Scacciag/Home-File/Teach/Datahiding.Pdf.

[5] Nedeljko C. (2004). Algorithms For Audio Watermarking And Steganography. Acta Universitatis Ouluensis. Series C., 2004..

[6] Andres G. (2002). Measuring And Evaluating Digital Watermarks In Audiofiles. Washington Dc. 2002.

[7] Supurovic P., "Mpeg Audio Compression Basics", Url: Http://Www.Chested. Chalmers.Se\~Kf96svgu\, 1998.

[8] M. Nilsson, "Id3 Tag Version 2.4.0 - Main Structure", November 2000, Available In Internet Www.Id3.Org/Id3v2.4.0-Structure.Txt .

[9] L. Maciak And M. Ponniah And R. Sharma, "Mp3 Steganography", 2008

[10] Strnad Peter, Gingold Peter, "Lyrics3 Tag V2.00", Jun 1998, Http://Www.Id3.Org/Lyrics3200.Html

[11] M. Nilsson, "The Private Life Of Mp3 Frames", Available In Internet Www.Id3.Org/Mp3frame.Htm

[12] Chan, P. (2011). Secret Sharing in Audio Steganography. Industrial Research.

[13] Deng, K., Tian, Y., Yu, X., Niu, X., Yang, Y., & Technology, S. (2010). Steganalysis of the MP3 Steganographic Algorithm Based on Huffman Coding. Test, (1), 79-82.

[14] P.K. Singh, H. Singh, And K. Saroha, "A Survey On Steganography In Audio," Audio, 2009.

[15] Seppanen T., Cvejic N., "Increasing The Capacity Of Lsb-Based Audio Steganography", Ieee 0-7803-7713, 2002, Url:Www.Mediateam.Oulu.Fi/Puplications/Pdf/374.Pdf.

[16] M. Wakiyama, Y. Hidaka, And K. Nozaki, "An Audio Steganography By A Low-Bit Coding Method With Wave Files," 2010 Sixth International Conference On Intelligent Information Hiding And Multimedia Signal Processing, Oct. 2010, Pp. 530-533

[17] Kekre, H. B., Athawale, a, Rao, B. S., & Athawale, U. (2010). Increasing the Capacity of the Cover Audio Signal by Using Multiple LSBs for Information Hiding. 2010 3rd International Conference on Emerging Trends in Engineering and Technology, 196-201. IEEE. doi:10.1109/ICETET.2010.118

[18] Kaliappan Gopalan, Qidong Shi, "Audio Steganography Using Bit Modification – A Tradeoff on Perceptibility and Data Robustness for Large Payload Audio Embedding", Computer Communications and Networks (ICCCN), 2-5 Aug. 2010.

[19] S. Shirali-Shahreza, M. T. Manzuri-Shalmani, "Adaptive Wavelet Domain Audio Steganography with High Capacity and Low Error Rate", IEEE International Conference on Information and Emerging Technologies, 2007, 06-07 July 2007 pp 1-5

[20] Bhowal, K., Pal, a J., Tomar, G. S., & Sarkar, P. P. (2010). Audio Steganography Using GA. 2010 International Conference on Computational Intelligence and Communication Networks, 449-453. Ieee. doi:10.1109/CICN.2010.91

[21] Zamani, M., Manaf, A. A., & Ahmad, R. B. (2011). Knots of Substitution Techniques of Audio Steganography. Computer Engineering, 2, 370-374.

[22] H. Matsuka, "Spread Spectrum Audio Steganography using Sub-band Phase Shifting," IEEE Int. conf. Intelligent
Information Hiding and Multimedia Signal Processing (IIHMSP' 06), pp. 3-6, Dec. 2006 , Pasadena, CA, USA

[23] K. Shah, V.R. Lakshmi Gorty, and A. Phirke , "Audio Steganography Using Differential Phase Encoding", ICTSM 2011, CCIS 145,. Springer-Verlag Berlin Heidelberg 2011, pp. 146–151, 2011

[24] H. Ozer, "Steganalysis Of Audio Based On Audio Quality Metrics," Proceedings Of Spie", 2003, Pp. 55-66

[25] Sameer Mitra and Sathiamoorthy Manoharan, "Experiments with and Enhancements to EchoHiding ",Fourth international Conference on Systems and Networks Communications, DOI 10.1109/ICSNC.2009.76,pp 119-124, 2009 IEEE

[26] A. Delforouzi And M. Pooyan, "Adaptive Digital Audio Steganography Based On Integer Wavelet Transform," Circuits, Systems & Signal Processing, Vol. 27, Mar. 2008, Pp. 247-259.

[27] Z. Kexin, "Audio Steganalysis Of Spread Spectrum Hiding Based On Statistical Moment," Signal Processing, 2010, Pp. 381-384

[28] X. Li and H. H. Yu, "Transparent and robust audio data hiding in subband domain,"in Proceedings of the IEEE International Conference on Information Technology:Coding and Computing, 2000, pp. 74-79.

[29] Atoum, M. S., Rababah, O. A. A.-, & Al-attili, A. I. (2011). New Technique for Hiding Data in Audio File. Journal of Computer Science, 11(4), 173-177.

[30] Atoum, M. S., Suleiman, M., Rababaa, A., Ibrahim, S., & Ahmed, A. (2011). A Steganography Method Based on Hiding secrete data in MPEG / Audio Layer III. Journal of Computer Science, 11(5), 184-188.

[31] Mikhail Zaturenskiy, "Behind The Music: MP3 steganography", April 4, 2009, Url: http://www.cpd.iit.edu/netsecure09/MIKHAIL_ZATURENSKIY.pdf

[32] Koso A., Turi A., And Obimbo C., "Embedding Digital Signatures In Mp3s", From Proceedings 477 Internet And Multimedia Systems, And Applications, 2005

[33] Bhattacharyya, S., Kundu, A., Chakraborty, K., & Sanyal, G. (2011). Audio Steganography Using Mod 4 Method. Computing, 3(8), 30-38.

[34] Bhattacharyya, S, A Novel Audio Steganography Technique by M16MA. International Journal, 30(8), 26-34.2011.