MJoSHT
Malaysian Journal of Science, Health & Technology
[mjosht.usim.edu.my]

USIM

Article

An Adaptive Multi Amplitude Thresholds Embedding Algorithm for Audio Steganography

Ahmed A. Alsabhany[1,a], Farida Hazwani Mohd Ridzuan[1.b], A. H Azni Haslizan Ab Halim[1,c]

[1, 2, 3]*Faculty of Science and Technology, Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800 Nilai, Malaysia*
E-mail: [a]ahmad88sabhany@gmail.com, [b]farida@usim.edu.my, [c]ahazni@usim.edu.my

***Abstract*— Secure data transfer is an ever-growing challenge. Steganography refers to the art and science of secret communication. Audio steganography provides a smart solution for secure data transferring. The idea is based on hiding a secret message inside an audio signal. The challenge is to produce a stego file that can embed more data without decreasing the quality of the audio file, which may compromise the existence of the embedded message. As the technology advances, old methods become vulnerable and easy to compromise. Several methods include an encryption algorithm to add a second line of defense against a probable eavesdropper. This paper proposes a novel algorithm that aims to improve the perceptual transparency, robustness, and capacity of the embedded message. The main idea of the algorithm is to avoid embedding in the silent or near to silent intervals of the audio samples, and to scramble the embedded message as much as possible. The algorithm uses AES-256 to increase the level of robustness. The proposed algorithm is uniquely adapted to the message size and expected to perform better than the existing algorithms.**

***Keywords*— audio steganography; LSB; amplitude threshold; embedding algorithm.**

I. INTRODUCTION

A secure data transfer refers to protecting data from unauthorized use, access, modification, disclosure, disruption, or destruction. One of the methods of data security is to hide the data itself. Hiding the data will conceal the fact that there exists a communication between two parties. Several articles proposed methods developed based on two main components which are cryptography and steganography [1]–[8]. Cryptography is defined as the science of secret writing with the goal of hiding the meaning of a message [9]. On the other hand, steganography is defined as the art and science of writing hidden messages in such a way that no one can realize there is a hidden message in data except for the sender and the intended recipient [10]. If the added noise become noticeable by the attacker, the steganography technique fails [6]. Few types of files can be used as cover, such as image, audio, and video [4], [11], [12].

Audio steganography is basically concealing a secret data inside an audio file. Audio steganography is the most challenging data hiding technique [13] because audio steganography targets the Human Auditory System (HAS), whereas image and video steganography targets the Human Visual System (HVS), and HAS is more sensitive to noise than HVS [6]. An audio steganography method is evaluated based on three main parameters: (1) capacity, (2) perceptual transparency, (3) and robustness [14]. Capacity refers to the amount of secret data that be embedded inside a cover file, perceptual transparency refers to the degree of secrecy that the embedded message retains, (i.e. the difference between the audio file before embedding the secret message and after should be negligible), and robustness refers to the ability of the embedded message to resist the attacks [1], [6], [15].

In the recent years, audio steganography has been an important topic [16] and many researchers were attracted to the idea, thus several techniques have been proposed in order to hide data inside an audio cover file. Some proposed methods [10], [17], [18] are based on the fact that embedding in silent intervals of the audio may cause some noticeable noise that may compromise the existing of the secret message. Similarly, embedding in the high amplitude audio samples may conceal the embedded message better [10]. In this paper, a novel approach is proposed for embedding any data format in the highest amplitude audio samples and avoid embedding in silent or near to silent samples. Embedding the data in the highest amplitude samples improves the imperceptibility features. In addition to improving imperceptibility, the proposed method includes AES encryption to reinforce the robustness of the method. The structure of the paper is as follows: the related works are presented in section 2, the proposed algorithm is presented in section 3, the expected results are presented in section 4, and finally the conclusion and future work are presented in section 5.

II. RELATED WORKS

Ahmed et al. [10] proposed a novel method to increase the capacity and to enhance the robustness of the hidden data. The main idea behind this method is to avoid embedding in the silent periods of the host audio signal. After selecting a threshold value, the algorithm calculates the amplitude value of each sample. The sample will only be used for embedding

only if that amplitude value is above the threshold. This method also increases the depth of embedding to the 8th bit, which improves the robustness. The main shortcoming of this method is that the message retrieval can be difficult [15]. Moreover, since the embedding happens only in non-silent segments of the host audio, a large number of samples (silent samples) will be excluded from embedding which will decrease the embedding capacity [15]. This method uses only one threshold for embedding which makes it prone causing to either additional noise or capacity loss. A high threshold can lead to an extreme capacity loss, whereas the low threshold can lead to lower quality of the audio (additional noise).

Srivastava and Rafiq [17] proposed a novel method based on two main ideas. Firstly, the samples are divided into two main categories based on the samples' amplitude. The samples that have a value below the designated threshold will be considered as silent samples, and therefore excluded from the embedding. Only the samples above that threshold are used for embedding. Secondly, the actual modification is done only in the two LSBs. The method compares up to three bits of the secret message with three MSBs of the sample. If a mismatch is found between the first bit of the message and the first MSB of the current sample, then 00 is inserted in the LSB. If a mismatch is found at the second MSB then 01 will be inserted in the LSB. This method in that case implicitly stores the value of the second bit of the message, since the second bit in the audio sample is the inverted value of the bit in the message. The third case occurs when a mismatch is found in the third bit. The LSB value will be modified to 10, obtaining two MSB of the sample and an inverted third bit. This method depends only on the security of the embedding method and lacks of an encryption mechanism. Hence, if the embedding technique is compromised by an attacker, the confidentiality of the message will be violated.

Wakiyama et al. [18] proposed two methods that are based on the amplitude value of the sample. The first method selects two thresholds to avoid embedding in the silent samples. If the amplitude is lower than the lowest threshold, no data will be embedded. If the amplitude is between the thresholds, one bit is used for embedding. If the amplitude is above the highest threshold, two bits are used for embedding. The second method uses the average amplitude of the samples before and after the selected sample. If the amplitude of that sample is higher than the average amplitude, two bits are used for embedding. No encryption algorithm was implemented to enhance the security.

Vimal and Alex [1] proposed a method based on three techniques: (1) Huffman encoding, (2) RSA encryption, and (3) dual randomness embedding technique. Huffman encoding is less vulnerable to steganalysis than standard encoding. Moreover, Huffman encoding provides lossless data compression which increases the capacity of the method. The encoded message is then encrypted using RSA. In the dual randomness technique, the sample number and the LSB positions used to embed the message bits are determined. The sample number is determined by the value of the 3 MSBs in the current sample plus the current sample number. The position of the selected bits is determined by the first 2 MSBs. The position of the bit is selected among 3 LSBs only, and the maximum number of bits embedded in the sample is one. This method obtains an excellent degree of robustness and security. However, since the algorithm embeds one bit at most per sample, and not all the samples will be used for embedding, this may decrease the capacity substantially, even with the consideration of the compression rate of the Huffman encoding. Moreover, randomness may lead to embedding in silent intervals of the audio, which may produce some noticeable noise.

Therefore, based on the analysis of related works, there are several weaknesses of existing studies which are: (1) the potential risk of using only one threshold, (2) the lack of encryption technique, and (3) the randomness or embedding in silent intervals of the audio. The proposed algorithm is created to improve these weaknesses.

III. The Proposed Algorithm

In order to achieve imperceptibility, the idea of the proposed algorithm is to select samples that obtained high audio (loud) signals so that the added noise is camouflaged by the original audio signal. This algorithm comprises three main steps: (1) Huffman encoding, (2) AES encryption, and (3) the novel embedding algorithm. Since all data formats are stored in bytes; this algorithm is capable of embedding all kinds of data into an audio signal. The algorithm uses uncompressed audio files (.wav) audio files with CD-quality (44100 samples/second; each is encoded 16 bits). Figure 1 shows the phases of which are Huffman Encoding, AES encryption, and the novel embedding method.

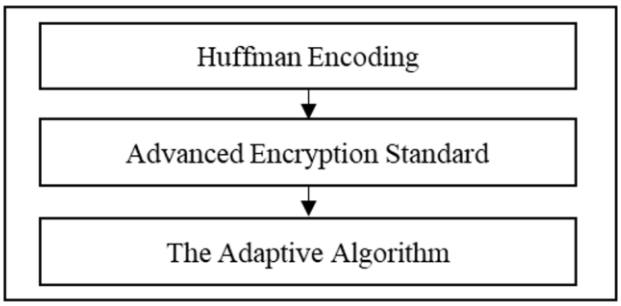

Fig. 1 The phases of the algorithm

A. Huffman Encoding

In this phase, the data are grouped as a stream of bytes. Then the ASCII representation of each byte is retrieved. After that, the Huffman encoding algorithm is performed on the ASCII characters. Huffman encoding provides lossless compression, which improves the capacity of the overall method. Before forwarding the result to the next phase, the algorithm regroups the results in a stream of bytes.

B. Advanced Encryption Standard (AES)

In this phase, AES is performed on the message based on a 256-bit key, which is agreed on between the sender and recipient. AES provides a high level of security and speed compared to other encryption algorithms [9], [19]. The resulted cipher is then converted back into an array of bits T_i. Finally, the total number of bits (the array length) is calculated and added to 50. A number of 50 bits size is large enough to represent the number of embedded bits. Then the array size is incremented by 50, the array is shifted by 50,

and then the length is inserted without being encrypted in the first 50 bits of the array. This process is called the length insertion.

C. The Adaptive Multi Amplitude Thresholds Embedding Algorithm

The algorithm embeds six bits maximum into an audio sample. The position on which the embedding occurs is denoted by *q*. In this phase, seven thresholds are considered for embedding. The algorithm first starts by embedding in the samples of high Amplitudes. At each level *k*, the lower threshold is calculated by setting the *kth* position to 1 and all other bits to 0. Figure 2 shows an example of calculating the lower threshold in three cases, when k equals 15, 11, and 9. The lower threshold *LT* equals (2^k), where the upper threshold *UT* equals (2^{k+1}). The maximum lower threshold is the 15 and at each step, the counter is decreased by 1, till k = 9 where the threshold equal 512. Samples with amplitude values below 512 are considered silent or near to silent samples, therefore the algorithm avoids embedding at those samples.

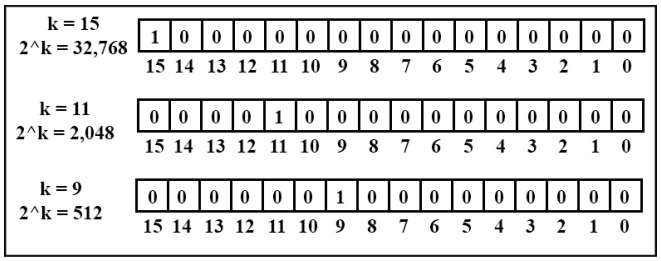

Fig. 2 Thresholds calculation

The decimal representation of a sample j is considered as the amplitude value AVj. At each level of the seven levels, the algorithm compares AVj with LT and UT, and only if AVj is more than LT and less than UT the algorithm substitute one LSB at a time, with one bit of the encrypted message. The algorithm is adaptable to message size because it can provide embedding in deeper layers. Moreover, the behaviour of the algorithm provides a scrambling mechanism, which increases the security of the embedded message. The pseudo code for the algorithm is presented in Figure 3.

```
q represents the layer of embedding
T represents the array of the message bits
i represents the index of the array T
n represents the length of T
k represents the level of amplitude
m represents the total number of audio samples
j represents the audio sample number
AVj represents the decimal representation of j
LT represents the lower threshold, where LT = 2^k
UT represents the upper threshold, where UT=2^(k+1)

for (q=0; q<7; q++){
        for (K=15; K>8; K--){
                LT = 2 ^ k;
                UT = 2 ^ (k+1)
                for(i=0; i<n; i++){
                        for(j = 0; j<m; j++){
                                If (AVj ≥ LT && AVj < UT){
                                J[q] = T[i]}
                        }
                }
        }
Report q value and prompt the user to continue or not
If (no){ exit;}
}
```

Fig. 3 The pseudo code of the embedding algorithm

The extraction operation is the exact reverse, instead of writing the audio sample the algorithm reads and reconstructs the array of bits. However, the main difference is that after extracting the first 50 bits which are the length of *T*, the algorithm stops and modify the value of *n* to the remaining bits by subtracting 50 from the decimal value represented in the first 50 bits. Figure 4 shows an example of embedding at the LSB (q = 0) and when the amplitude level is 11 (k = 11).

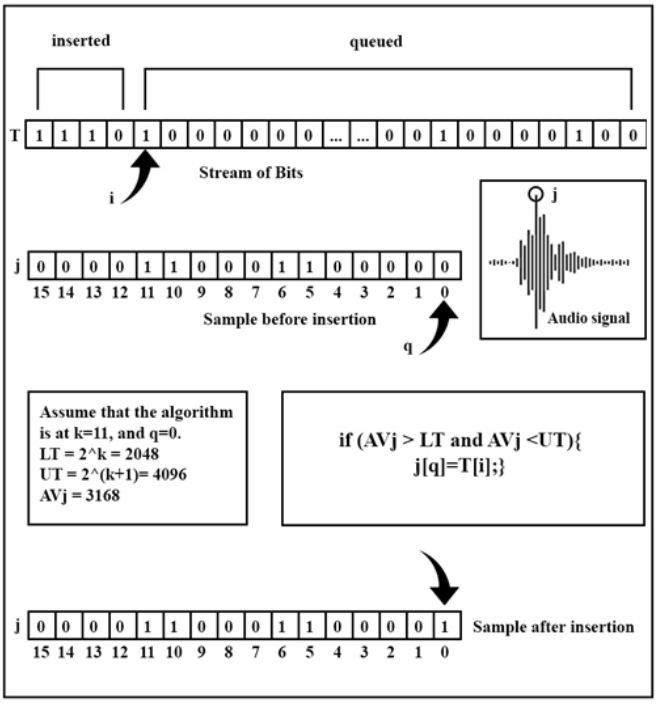

Fig. 4 An example of embedding

IV. Expected results and Discussion

The objective of the proposed algorithm is to improve the perceptual transparency, the robustness, and the capacity.

The embedding is carried out in the "noisy" samples to avoid detection by listening. The advantage of the samples selection mechanism is that it provides a high level of scrambling, which makes the data harder to be extracted by an eavesdropper. However, the mechanism will put more restrictions on samples to be selected. Consequently, the capacity could drop because not all available samples will be selected for embedding. Nonetheless, the algorithm compensates for the resulted loss of capacity by increasing the layer of embedding. In addition, the algorithm uses a lossless compression technique to increase the capacity. The algorithm is expected to improve perceptual transparency, robustness, and capacity.

V. CONCLUSIONS

The proposed algorithm provides a high level of security based on three main features: (1) avoid embedding in the silent or near to silent samples, (2) the samples selecting criteria provide a unique scrambling mechanism, and (3) AES encryption reinforces the message with a robust last line of defense. As this algorithm aims to improve the security feature, it is expected that this will cause a small capacity decrement. Therefore, in order to solve the capacity issue, the algorithm adapts to large message sizes by increasing the number of LSBs used for embedding based on the size of the message. The increment of the number of LSBs will allow the algorithm to embed more bits per sample (up to six). In addition, using Huffman encoding (lossless compression) may also compensate for capacity decrement.

For future works, an experimental evaluation will be carried out for the proposed algorithm in order to evaluate its performance. Moreover, a comparison study will be carried out against the most related methods to capture the main variation in the algorithm's performance.

REFERENCES

[1] J. Vimal and A. M. Alex, "Audio steganography using dual randomness LSB method," in 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies, ICCICCT 2014, 2014, pp. 941–944.

[2] O. C. Abikoye, K. S. Adewole, and A. J. Oladipupo, "Efficient Data Hiding System using Cryptography and Steganography," *Int. J. Appl. Inf. Syst.*, vol. 4, no. 11, pp. 6–11, 2012.

[3] B. Bhardwaj, "Analysis & Evaluation of Digital Audio Steganography," *Int. J. Innov. Res. Sci. Technol.*, vol. 2, no. 11, pp. 613–617, 2016.

[4] M. Jyotheeswari, "A Novel Steganographic System for Data Hiding in Video / Audio," *Int. J. Comput. Appl.*, vol. 82, no. 11, pp. 31–36, 2013.

[5] M. Asad, J. Gilani, and A. Khalid, "Three layered model for audio steganography," *Proc. - 2012 Int. Conf. Emerg. Technol. ICET 2012*, pp. 270–275, 2012.

[6] M. Asad, J. Gilani, and A. Khalid, "An enhanced least significant bit modification technique for audio steganography," *Proc. - Int. Conf. Comput. Networks Inf. Technol.*, pp. 143–147, 2011.

[7] A. M. Meligy, M. M. Nasef, and F. T. Eid, "A Hybrid Technique for Enhancing the Efficiency of Audio Steganography," *Int. J. Image, Graph. Signal Process.*, vol. 8, no. 1, pp. 36–42, 2016.

[8] N. Sinha, A. Bhowmick, and B. Kishore, "Encrypted Information Hiding using Audio Steganography and Audio Cryptography," *Int. J. Comput. Appl.*, vol. 112, no. 5, pp. 49–53, 2015.

[9] C. Paar and J. Pelzl, *Understanding Cryptography: A Textbook for Students and Practitioners*. Springer–Verlag, 2010.

[10] M. A. Ahmed, M. L. M. Kiah, B. B. Zaidan, and A. A. Zaidan, "A Novel Embedding Method to Increase Capacity and Robustness of Low-bit Encoding Audio Steganography Technique Using Noise Gate Software Logic Algorithm," *J. Appl. Sci.*, vol. 10, no. 1, pp. 59–64, 2010.

[11] R. Tanwar and M. Bisla, "Audio steganography," *ICROIT 2014 - Proc. 2014 Int. Conf. Reliab. Optim. Inf. Technol.*, pp. 322–325, 2014.

[12] V. L. Reddy, A. Subramanyam, and P. C. Reddy., "A Novel Approach for Hiding Encrypted Data in Image , Audio and Video using Steganography," *Int. J. Comput. Appl.*, vol. 69, no. 15, pp. 37–44, 2013.

[13] V. Sharma, "LSB Modification Based Audio Steganography Using Trusted Third Party Key Indexing Method," in *2015 Third International Conference On Image Information Processing (ICIIP)*, 2015, pp. 403–406.

[14] F. Djebbar, B. Ayad, K. A. Meraim, and H. Hamam, "Comparative study of digital audio steganography techniques," *EURASIP J. Audio, Speech, Music Process.*, vol. 2012, no. 1, pp. 1–16, 2012.

[15] F. Djebbar, B. Ayad, H. Hamam, and K. Abed-Meraim, "A view on latest audio steganography techniques," *2011 Int. Conf. Innov. Inf. Technol. IIT 2011*, pp. 409–414, 2011.

[16] H. Li, Z. Qin, X. Zhang, and X. Wang, "Auditory cryptography security algorithm with audio shelters," *Procedia Eng.*, vol. 15, pp. 2695–2699, 2011.

[17] M. Srivastava and M. Q. Rafiq, "A Novel Approach to Secure Communication Using Audio Steganography," *Adv. Mater. Res.*, vol. 403–408, pp. 963–969, 2011.

[18] M. Wakiyama, Y. Hidaka, and K. Nozaki, "An audio steganography by a low-bit coding method with wave files," in *Proceedings - 2010 6th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIHMSP 2010*, 2010, pp. 530–533.

[19] W. Stallings, *Cryptography And Network Security Principles And Practice Fifth Edition*. 2011.