Telecommun Syst (2015) 59:291–304
DOI 10.1007/s11235-014-9936-x

Genetic algorithm for fragile audio watermarking

Mazdak Zamani · Azizah Bt Abdul Manaf

Published online: 11 December 2014
© Springer Science+Business Media New York 2014

Abstract A novel genetic-concept-based algorithm is proposed for fragile audio watermarking to reduce the distortion of least significant bits substitution and consequently improve the peak signal-to-noise ratio (PSNR) and increase the payload of results. The result of testing shows that in comparison with ordinary substitution techniques and other presented techniques, the payload is considerably increased and PSNR (as an indicator of imperceptibility) is noticeably raised.

Keywords Artificial intelligence · Digital data hiding · Multimedia security · Steganography · Fragile audio watermarking

1 Introduction

Watermarking and steganography techniques embed information in a media in a transparent manner. Steganography is a form of security through obscurity; the science and art of hiding the existence of a message between sender and intended recipient [1]. Steganography is the study of methods for hiding the existence of secondary information in the presence of primary information in a way which neither affects the size nor results in perceptual distortion [2–5]. Steganography is a technique for covert information, but watermarking may not hide the existence of the message from third persons. Watermarking usually requires robustness to withstand attacks intended to remove or destroy the hidden message from the watermarked media, as well as preserving the carrier signal quality. This makes watermarking appropriate for those applications where the knowledge of a hidden message leads to a potential danger of manipulation. In fragile watermarking, a digital watermark is fails to be detectable after the slightest modification. Fragile watermarks are commonly used for tamper detection (integrity proof) [6–9].

A fundamental tradeoff exists between three key variables which restrict watermarking designers: robustness, payload and imperceptibility. However in some applications, computational time in determining the efficiency of a steganography technique is a crucial factor. The payload of an information hiding technique indicates the amount of data that an information hiding technique can successfully embed without introducing perceptual distortion in the changed media. Payload is usually measured in bits per sample (bps) or bits per second (BPS) [10–14].

Imperceptibility is the perceptual similarity between the audio signal before embedding (host) and after embedding (stego). In audio watermarking, imperceptibility is evaluated as an audible distortion caused by signal modifications. There are different methods to measure imperceptibility (listening test, PSNR, and so on). In order to meet the fidelity constraints of the embedded information, the perceptual distortion introduced by embedding should not be above the estimated threshold based on the human auditory system [15,16].

The ability of watermark to withstand intentional and unintentional attacks is measured as robustness. Unintentional attacks are generally common data manipulations, whereas intentional attacks include media degradations such as resizing and filtering attacks. Robustness is required in some applications such that a set of signal processing modifications is predefined, while in some other applications robustness is not desirable; these latter techniques are so called fragile audio watermarking techniques [17,18].

M. Zamani (✉) · A. B. A. Manaf
Advanced Informatics School (Postgraduate School),
Universiti Teknologi Malaysia (International Campus),
Jalan Semarak, 54100 Kuala Lumpur, Malaysia
e-mail: zamani.mazdak@gmail.com

A. B. A. Manaf
e-mail: azizaham.kl@utm.my

www.Matlabi.ir

The contribution of this research will be improving the imperceptibility more than other techniques and increasing the payload almost to double size which has never been met in other works.

2 Rationale

Initially watermarking techniques were developed for images. But later on researchers became interested in developing techniques for audio. Recently, several algorithms for audio watermarking have been presented. Audio watermarking is more secure due to the small number of audio steganalysis methods, since most watermarking techniques are for still images [19]. To compare the most robust techniques and the highest payload techniques, substitution techniques and spread spectrum techniques are introduced.

2.1 Substitution technique

Basic substitution systems try to encode secret information by substituting insignificant parts of the cover by secret message bits; the receiver can extract the information if he has knowledge of the positions where secret information has been embedded. Since only minor modifications are made in the embedding process, the sender assumes that they will not be noticed by a passive attacker. These approaches are common in watermarking and are relatively easy to apply in both image and audio. A surprising amount of information (e.g. the payload of techniques based on LSB, which is the most popular substitution technique, are more than 40,000 BPS) can be hidden into carriers imperceptibly by this technique [20–22].

Ji et al. [23] enhanced PSNR value for LSB watermarking with a secluded statistic peculiarity. They found the best mapping function between host and secret file at global scope and as a result were able to minimize the degradation of the stego. Their work improved the PSNR by 1.63 dB compared to the simple LSB. Given that the payload of their proposed method is quite high, the amount of improvement is acceptable. The problem with their proposed method is its imperceptibility: at about 45 dB, it is not high enough for some applications.

Wang et al. [24] developed a genetic algorithm to hide message data in the K-rightmost LSBs of the host. However the computation time needed to find the optimal result could be huge when K is large. Also an improved hiding technique to obtain a high-quality embedding is developed based on the concept of perceptual modeling. The PSNR obtained by the GA approach is very high, which makes the quality of the embedding result acceptable. Their work improved the PSNR by 0.55 dB on average compared to the simple LSB. Given that their proposed method does not provide high payload, the amount of improvement is relatively low.

Liu et al. [25] proposed a variable depth substitution technique for data hiding which improves the PSNR by 2.33 dB on average compared to the simple LSB. While the amount of improvement is acceptable, the payload is low.

Wu et al. [21] applied the LSB substitution and genetic algorithm (GA) to improve the quality of the watermark. They developed two different optimal substitution strategies: one is the global optimal substitution strategy and the other is the local optimal substitution strategy.

Their work improves the PSNR by 0.56 dB on average compared to the simple LSB. While the payload of their proposed method is quite high, the amount of improvement is low. Another problem with their method is its imperceptibility: at about 33 dB, it is not high enough for some applications.

2.2 Spread spectrum technique

Kalker et al. [26] defined spread spectrum techniques as a means of transmission in which the signal occupies a bandwidth in excess of the minimum necessary to send the information; the band spread is accomplished by means of a code which is independent of the data, and synchronized reception with the code at the receiver is used for despreading and for subsequent data recovery.

Although the power of the signal to be transmitted can be large, the signal-to-noise ratio in every frequency band will be small. Even if parts of the signal could be removed in several frequency bands, enough information should be present in the other bands to recover the signal. Thus, a spread spectrum makes it difficult to detect and/or remove a signal. This situation is very similar to a watermarking system which tries to spread a secret message over a cover in order to make it impossible to perceive. Since spread signals tend to be difficult to remove, embedding methods based on a spread spectrum should provide a considerable level of robustness [26–33].

Although the substitution technique is not as robust as other watermarking techniques like the spread spectrum technique, the payload of substitution techniques is incomparably higher. In general, the payload of substitution techniques is more than 40,000 BPS, while more robust techniques like the spread spectrum have a negligible payload that is only about 4 BPS [34–39].

3 Problem statement

A trade-off is required between the payload and robustness, at the same time keeping the quality of the watermarking algorithm at an acceptable level. It is not achievable to get a high payload and high robust technique at the same time in watermarking. Hence, if it is desired to have a robust watermarking algorithm, its payload will be low; and vice versa,

www.Matlabi.ir

a watermarking algorithm with high payload embedding is usually very fragile [19].

Many applications exist which do not require high robustness in watermarking. Apart from robustness that is largely not desirable for substitution techniques (e.g. those algorithms are labeled fragile audio watermarking algorithms) [11,19], the only remaining measure to achieve high payload is imperceptibility. Although substitution techniques are comparatively well-known for achieving high payload, they have yet to satisfy the imperceptibility condition. The distortion caused by substitution degrades the quality.

Therefore, to take advantage of a potential high payload, imperceptibility should be retained. In other words, while improving imperceptibility, a higher payload should be achieved.

The general question this research aims to answer is:

RQ: How can the quality of audio watermarking techniques be improved?

This question contains the two following sub-questions:

Sub-RQ 1: How can the PSNR be maintained at a high level when a higher payload is desired?

Sub-RQ 2: How to maintain a high payload when a higher PSNR is desired?

An efficient algorithm could reduce the distortion caused by the substitution of sample bits. This reduction of distortion directly improves the PSNR, and indirectly increases the payload. If the PSNR is improved, the difference between the prior and improved PSNR could be loaded by embedding more data, which would indirectly increase the payload.

4 Proposed step

In the substitution techniques, some insignificant bits of the host file, called LSBs, are replaced with the same number of message file bits. This substitution will probably change the decimal value of the sample. The difference between host and stego for each sample in decimal values is called the error of that sample. Consequently, this will have an effect on the PSNR.

A modification step is proposed in addition to the standard steps of ordinary substitution techniques that alters the sample bits in order to decrease the amount of error in the sample, thus the PSNR is improved in the sample and as a result in the final PSNR.

For example, assume that each sample has 16 bits and the payload is 3 bits per sample (bps). Therefore, three LSBs are used to embed the message bits and other thirteen bits are used in the proposed modification step. All possible modifications in the proposed step are as follows:

- Following algorithm can modify the sample if
  - Host bits are 010 and message bits are 111
  - Host bits are 001 and message bits are 110 or 111
  - Host bits are 000 and message bits are 101 or 110 or 111

From (fourth bit in the byte) TO (the end of the byte OR a bit whose value is 1)

DO reset current bit to 1;

IF (the loop above is finished because of the second condition)

DO reset current bit to 0.

- Following algorithm can modify the sample if
  - Host bits are 101 and message bits are 000
  - Host bits are 110 and message bits are 000 or 001
  - Host bits are 111 and message bits are 000 or 001 or 010

From (fourth bit in the byte) TO (the end of the byte OR a bit whose value is 0)

DO reset current bit to 0;

IF (the loop above is finished because of the second condition)

DO reset current bit to 1.

- There is no chance to modify a sample if
  - Host bits are 011 OR host bits are 100

For example, if 3 LSBs of a sample are substituted with 3 bits of a message stream, the proposed modification step decreases the difference between the decimal value of the host and stego as follow:

Original host chromosome:	$10000000_2 = 128_{10}$	before embedding
Embedding message bits:	111	3 bps payload is supposed
Initial stego chromosome:	$10000111_2 = 135_{10}$	current difference: 135−128=7 (not modified yet)
1st iteration of mutation:	$10001111_2 = 143_{10}$	current difference: 143−128= 15 (modified but worse result)
2nd iteration of mutation:	$10011111_2 = 159_{10}$	current difference: 159−128= 31 (modified but worse result)
3rd iteration of mutation:	$10111111_2 = 191_{10}$	current difference: 191−128= 63 (modified but worse result)
4th iteration of mutation:	$11111111_2 = 255_{10}$	current difference: 255−128=127 (modified but worse result)
5th iteration of mutation:	$01111111_2 = 127_{10}$	current difference: 128−127= 1(modified with the best result)

www.Matlabi.ir

5 Evaluation of the proposed step

Since there are many possibilities that are not used in the modification step, only modified possibilities are listed in Table 1. The first column of the table shows the decimal value corresponding to the host bits which are going to be replaced with message bits. The second column of the table shows the decimal value corresponding to the message bits which are going to be substituted. The third column of the table shows the value of the next bit before modification. The fourth column of the table shows the decimal value corresponding to the first 4 host bits (before modification). The fifth column of the table shows the decimal value corresponding to the first 4 stego bits (after modification). The sixth column of the table shows the weight of the next bit after modification. The seventh column of the table shows the decimal value corresponding to the first 4 stego bits (after modification). The eighth column of the table shows the amount of error before modification. That is the difference between the 4-bits value of the host before modification and the 4-bits value of the stego before modification. The ninth column of the table shows the amount of error after modification. That is the difference between the 4-bits value of the host before modification and the 4-bits value of the stego after modification. The amount of improvement is listed in the tenth column.

Now, we seek to formulate mathematically how the PSNR can be improved by proposed modification step:

$$T_{PSNR} = \sum_{\text{each } \theta} \omega(\theta) * \rho(\theta)$$

where

T_{PSNR} : total improvement in PSNR
θ : error rate
$\omega(\theta)$: improvement in PSNR by θ improvement in error
$\rho(\theta)$: quantity of β

Then

$$\omega(\theta) = \text{PSNR}(\beta) - \text{PSNR}(\alpha) \quad \theta = \beta - \alpha$$
$$\rho(\theta) = N(\theta)/A$$

where

$N(\theta)$: quantity of possibilities whose error rate is improved by θ
A : number of all possibilities

Then

$\text{PSNR}(\beta) = 10 * \log(\text{MAX}^2 / \text{MSE}_\beta)$
MAX : maximum value of sample
MSE_β : Mean-Squared-Error of β

Table 1 Summarized tables for the possibilities of 3 bit per sample

3 Bits decimal value of host	3 Bits decimal value of message	Next bit decimal value before modification	4 Bits decimal value of host before modification	4 Bits decimal value of stego before modification	Next bit value after modification	4 Bits decimal value of stego after modification	Error before modification	Error after modification	Amount of improvement	Amount of improvement in PSNR
0	5	8	8	13	0	5	5	3	2	4.44
0	6	8	8	14	0	6	6	2	4	9.54
0	7	8	8	15	0	7	7	1	6	16.90
1	6	8	9	14	0	6	5	3	2	4.44
1	7	8	9	15	0	7	6	2	4	9.54
2	7	8	10	15	0	7	5	3	2	4.44
5	0	0	5	0	8	8	5	3	2	4.44
6	0	0	6	0	8	8	6	2	4	9.54
6	1	0	6	1	8	9	5	3	2	4.44
7	0	0	7	0	8	8	7	1	6	16.90
7	1	0	7	1	8	9	6	2	4	9.54
7	2	0	7	2	8	10	5	3	2	4.44

www.Matlabi.ir

For example, the total improvement in the PSNR for 3 bps is calculated as follows:

$$T_{PSNR} = \sum_{\text{each } \theta} \omega(\theta) * \rho(\theta) \quad \theta = 2, 4, 6$$

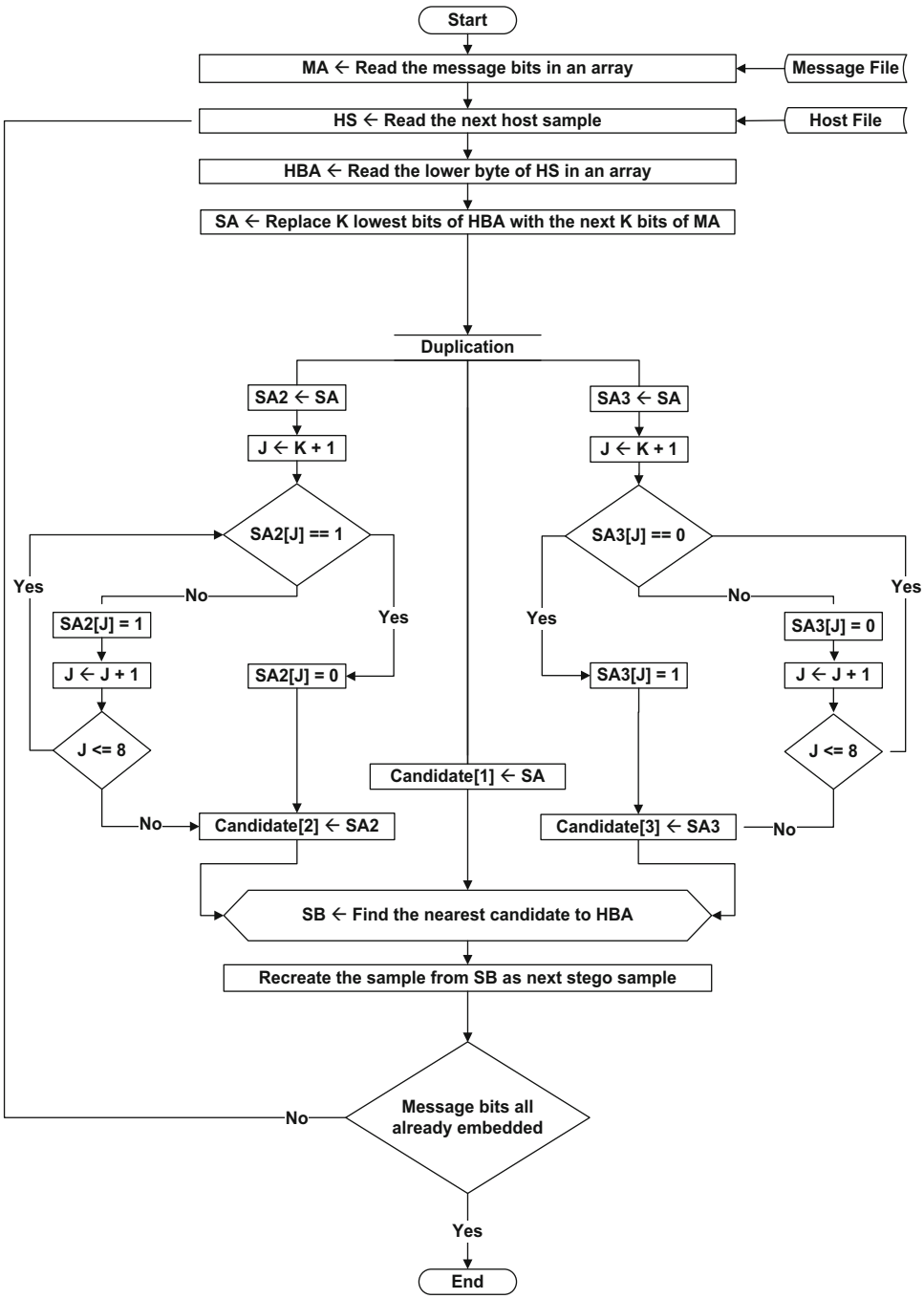
There are 64 possibilities in all, as referred to in the summarized information in Table 1,

$$\begin{aligned} T_{PSNR} &= \omega(2) * \rho(2) + \omega(4) * \rho(4) + \omega(6) * \rho(6) \\ &= 4.43 * (6/64) + 9.54 * (4/64) + 16.90 * (2/64) \\ &= 4.436 * 0.09 + 9.54 * 0.59 + 16.90 * 0.52 \\ &= 1.54 \end{aligned}$$

6 Proposed method

As Fig. 1 shows, message bits are firstly read in an array. Host bits are read sample by sample for embedding. Since the method is based on LSB, and each sample consists of two bytes in the format used in WAV files, the algorithm separates the lower byte for manipulation and stores it in an

Fig. 1 Research flow chart



www.Matlabi.ir

array. Depending on the preferred bit rate for the method, K lowest bits of the current read host byte are substituted with K bits of message.

In the ordinary method of LSB substitution, embedding is already done, but the proposed method would still modify some other bits in order to reduce the amount of error caused by substitution.

The proposed method uses the concept of the genetic algorithm for error reduction. It was preferred not to use genetic algorithm in the usual way because that is generally time-consuming. In contrast, the proposed genetic-concept-based method creates new children intelligently and avoids creating those children which are not potentially going to be selected. This is a result of the study having been done manually in order to find some good possibilities (as children) of mutation. The genetic parts of the proposed technique are done as follow:

6.1 Offspring

In the duplication step corresponding to the offspring in the genetic algorithm, apart from the original stego which is already considered as a child or candidate, two more candidates are going to be born with two different formulas for mutation.

6.2 Mutation

Both formulas work similarly, except that the first formula mutates the given chromosome (stego sequence) in order to decrease the decimal value of the stego byte (or chromosome), whereas the second formula mutates the given chromosome (stego sequence) in order to increase the decimal value of the stego byte.

6.3 Fitness function

Fitness function is called to calculate the nearest decimal value of candidates to the original value of host. The input of fitness function consists of three children which are the original stego sample (which is the original sample with substituted bits by message bits), first mutated child, and second mutated child. The function finds the child which has the nearest decimal value to the original sample. When that is found, that will be selected to construct the final version of the stego sample. This procedure lasts until the message bits all are embedded.

7 Experimental results: improvement in imperceptibility

The proposed algorithm, Genetic Substitution Based Audio Watermarking (GSBAW), was implemented and applied. The improvement in imperceptibility by GSBAW is presented in an experimental way. The payload is kept the same and the PSNR is increased. For the same host, different sizes of messages were selected, embedded by different bits per sample rate algorithms spread across almost the entire host, then their PSNRs were calculated.

As is shown in Tables 2 and 3, the same message is embedded into the same host by simple LSB and the proposed technique (GSBAW). The obtained PSNR with the proposed technique is about 3 dB better than the simple LSB for four bits per sample rate. The degree of improvement for the same situation except for the 2 bits per sample rate is about 2.2 dB, and for the 3 bits per sample rate is about 2.8 dB. As Fig. 2 shows, the PSNR is almost the same for all different tested host files. Also, as Fig. 3 shows, there is a clear improvement between the same bits per sample rate algorithms, with GSBAW.

Table 4 shows the average amount of improvement in PSNR (dB) in other techniques compared with our proposed technique (GSBAW).

7.1 Visual representation of improvement in imperceptibility

Furthermore, as shown in Table 5, two samples are watermarked by proposed technique (GSBAW) and visual repre-

Table 2 Same payload with increased PSNR for 1, 2, and 3 bps

WAV Size (KB)	Message size (KB) 1 bps	PSNR 1 bps by simple LSB	Message size (KB) 2 bps	PSNR 2 bps by GSBAW	PSNR 2 bps by simple LSB	Message size (KB) 3 bps	PSNR 3 bps by GSBAW	PSNR 3 bps by simple LSB
817.68	50.06	72.03	99.99	67.22	64.91	150.20	61.50	58.81
824.39	50.06	72.26	99.99	67.42	65.15	150.20	61.68	59.03
34,623.04	2, 154.10	72.39	4, 285.16	67.60	65.37	6, 171.03	62.05	59.26
122,233.54	7, 628.00	71.89	15, 161.59	67.10	64.89	22, 343.79	61.43	58.33
Average PSNR		72.14		67.33	65.08		61.67	58.86
Difference				2.25			2.81	

www.Matlabi.ir

Table 3 Same payload with increased PSNR for 4, 6, and 8 bps

WAV size (KB)	Message size (KB) 4bps	PSNR 4 bps by GSBAW	PSNR 4 bps by simple LSB	Message size (KB) 6 bps	PSNR 6 bps by simple LSB	Message size (KB) 8 bps	PSNR 8 bps by simple LSB
817.68	203.96	55.34	52.52	303.34	40.40	407.21	28.15
824.39	203.96	55.47	52.72	303.34	40.58	407.21	28.27
34,623.04	8, 516.86	55.63	51.62	12, 129.78	40.67	16, 339.53	28.07
122,233.54	29, 501.50	55.44	52.72	35, 155.72	41.64	57, 024.04	27.12
Average PSNR		55.47	52.40		40.82		27.90
Difference		3.07					

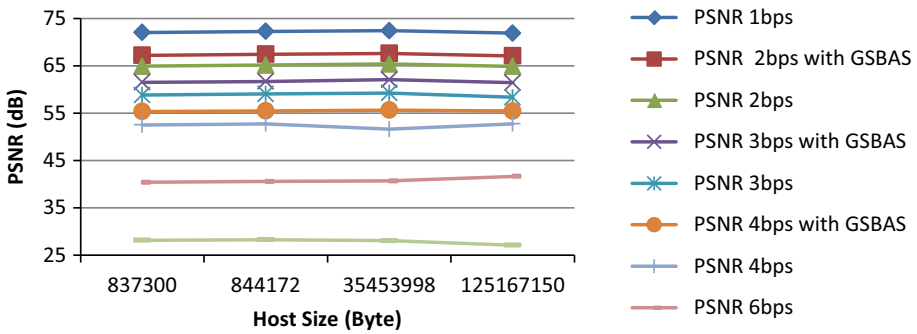


Fig. 2 Almost the same PSNR for all tested host files

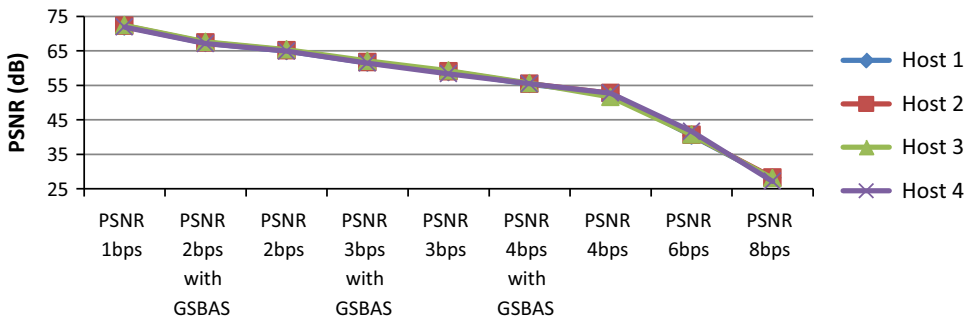


Fig. 3 Improvements in PSNR by GSBAW

Table 4 Comparison of average amount of improvement in PSNR

	Average amount of improvement in PSNR (dB)
Ji's method	1.63
Wang's method	0.55
Wu's method	0.56
Liu's method	2.33
GSBAW	3.07

Table 5 Comparison of average amount of improvement in PSNR

	WAV size (bytes)	Message size (bytes)	PSNR with proposed method
Sample 1	817.68	203.96	55.34
Sample 2	414.69	13.72	73.05

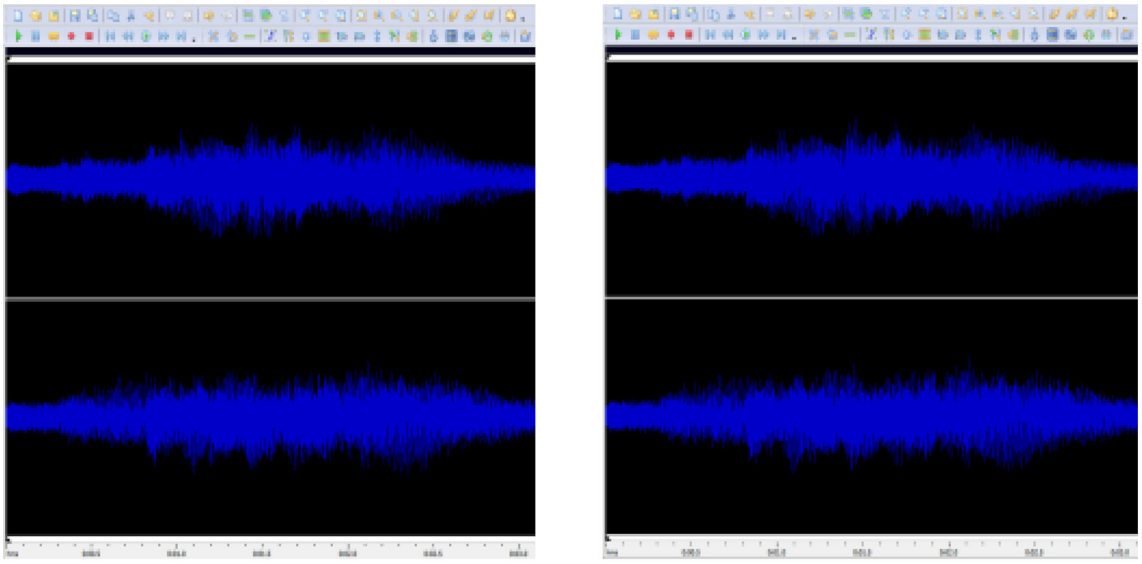
sentations of audio files are shown in Figs. 4, 5, 6, 7, 8 and 9.

8 Experimental results: improvement in payload

The improvement in payload by GSBAW is presented in an experimental way. PSNR is maintained when the payload is increased. To show how the proposed technique can improve the payload, the embedding algorithm was run on the same host to maintain the PSNR. The result is shown in Table 6.

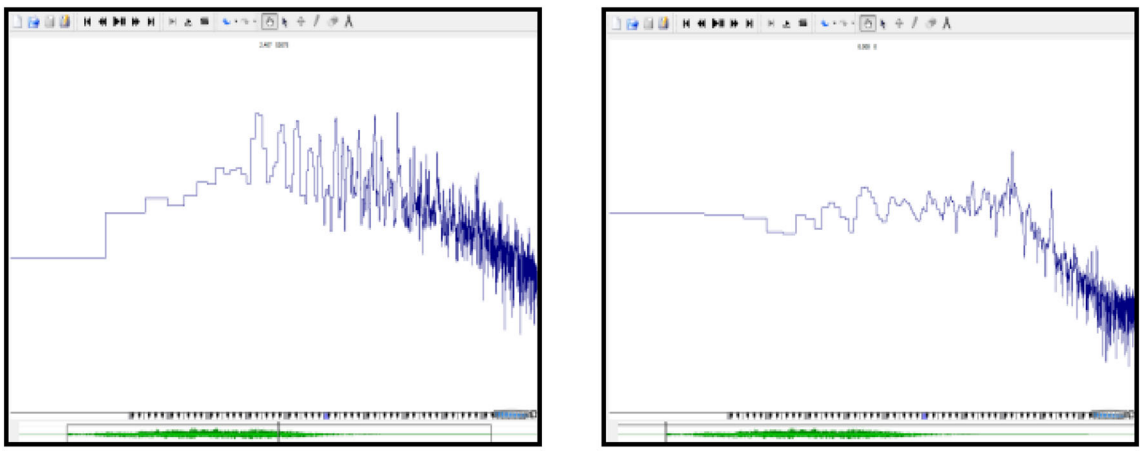
Referring to Table 6, imperceptibility is maintained and payload is increased using the proposed technique (GSBAW) to embed messages into the same audio host (in each row). If a certain level of imperceptibility is desired (for example in Table 5, 55 dB is supposed for PSNR), the payload can be increased from 18 to 25 % of the audio host size by the proposed technique. In other words, for instance in the

www.Matlabi.ir



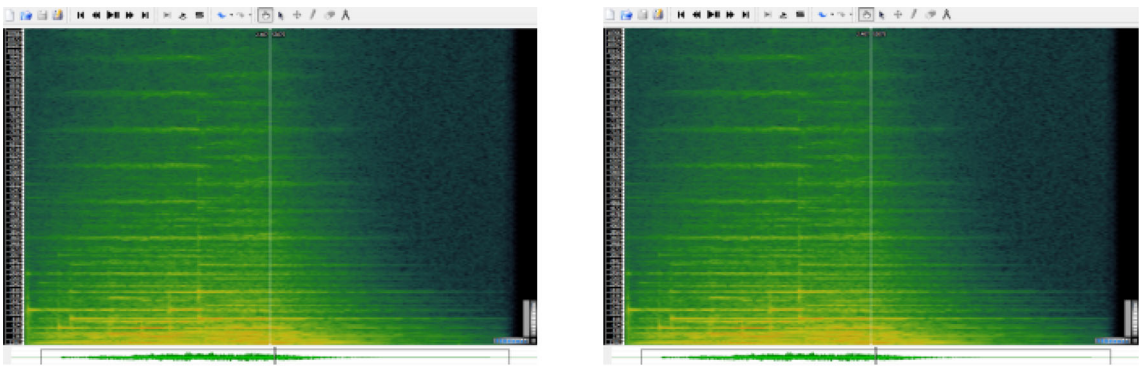
(a) Sample No. 1 before embedding **(b)** Sample No. 1 after embedding

Fig. 4 Waveform visual representation in order for Sample No. 1 to compare the audio file before and after embedding



(a) Sample No. 1 before embedding **(b)** Sample No. 1 after embedding

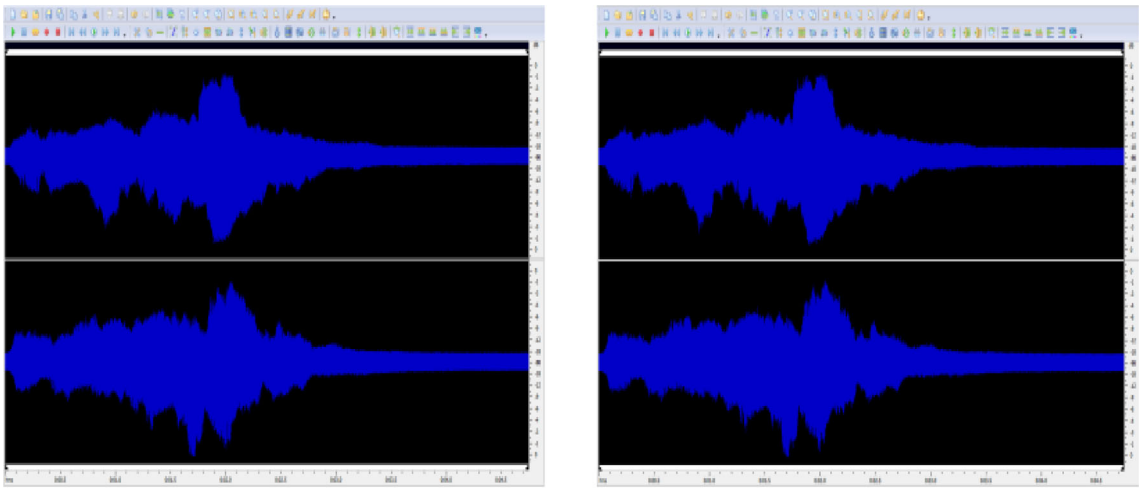
Fig. 5 Spectrum visual representation in order for Sample No. 1 to compare the audio file before and after embedding



(a) Sample No. 1 before embedding **(b)** Sample No. 1 after embedding

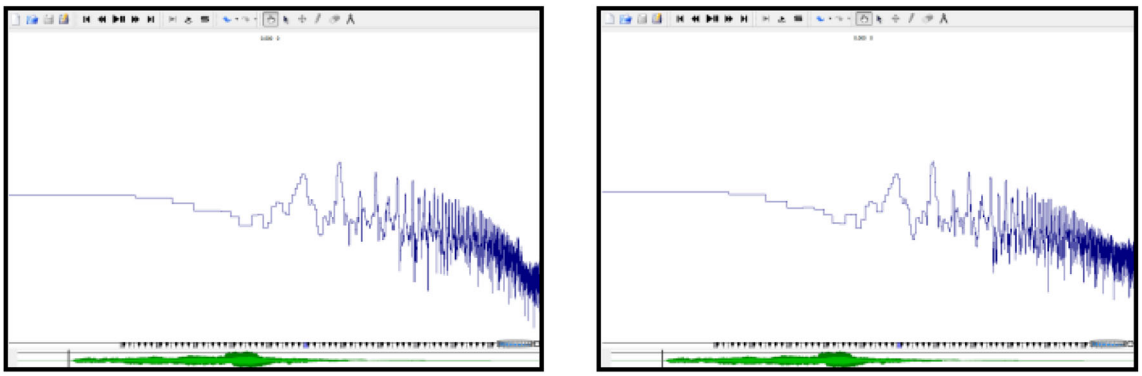
Fig. 6 Spectrogram visual representation in order for Sample No. 1 to compare the audio file before and after embedding

www.Matlabi.ir



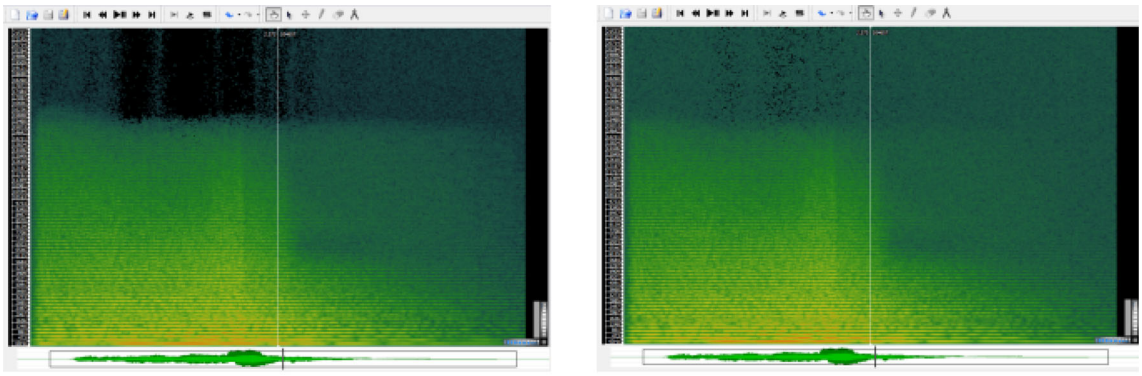
(a) Sample No. 2 before embedding **(b)** Sample No. 2 after embedding

Fig. 7 Waveform visual representation in order for Sample No. 2 to compare the audio file before and after embedding



(a) Sample No. 2 before embedding **(b)** Sample No. 2 after embedding

Fig. 8 Spectrum visual representation in order for Sample No. 2 to compare the audio file before and after embedding



(a) Sample No. 2 before embedding **(b)** Sample No. 2 after embedding

Fig. 9 Spectrogram visual representation in order for Sample No. 2 to compare the audio file before and after embedding

last row, when a PSNR with a minimum value of 54 dB is desired and the audio host file is a WAV file whose size is 119.37 MB, a simple LSB is capable of embedding a message file whose size is only 21.82 MB. But using GSBAW the size of the embeddable message file is increased to 28.81 MB.

In Table 7, when the proposed technique is used for a message file whose size is 29.6 MB, the obtained PSNR is

www.Matlabi.ir

Table 6 Increased payload with the same PSNR

WAV size in Mega byte (MB)	Message size (MB) 4 bps	Message/host size	PSNR 4 bps by simple LSB	Message size (MB) 4 bps	Message/host size	PSNR 4 bps by GSBAW
0.80	0.15	0.18	53.97	0.20	0.25	55.34
0.81	0.15	0.18	54.18	0.20	0.25	55.47
33.81	6.03	0.18	54.40	8.32	0.25	55.63
119.37	21.82	0.18	53.42	28.81	0.24	55.44

Table 7 Larger message in the same host with the same PSNR

Tool	Host size (MB)	Message size (MB)	PSNR (db)
Simple LSB	119.4	12.7	56.1
		13.2	55.8
		14.4	55.1
		14.7	54.7
		16.2	54.6
		20.2	53.5
		25.0	53.3
		29.6	52.6
GSBAW	119.4	29.6	55.3

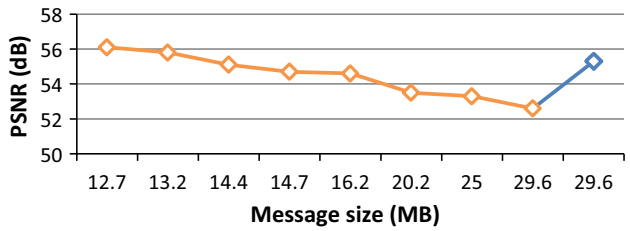


Fig. 10 The improvement in payload by GSBAW

55.3 dB. While using the simple LSB technique to get almost the same PSNR, the payload (message size) is almost half. The amount of payload obtained by the proposed technique is shown in Fig. 10.

Table 8 shows the amount of improvement in payload in other techniques [40–46] compared with our proposed technique (GSBAW).

Table 9 Average PSNR of GSBAW for 4 bps

WAV size (MB)	Message size (MB)	PSNR (db)
0.80	0.20	55.34
0.81	0.20	55.47
33.81	8.32	55.63
119.37	28.81	55.44
Average PSNR		55.47

As Table 9 shows, GSBAW is capable of embedding 4 bits per sample with an average PSNR of 55.47 dB.

9 Experimental results: listening test to determine the threshold of noise perception

To determine the threshold of noise perception, a listening test was performed. Twenty musicians and music fans who were familiar with the nature of the noise took part in the listening test. The result is shown in Table 10.

Obviously, the noise perception at 1 and 2 bps is negligible and can be considered as imperceptible. By contrast, the noise is quite perceptible at 6 and 8 bps. Noise perception at 6 and 8 bps rates is very high, and even with improvement will maintain perceptible.

On the other hand, noise perception at the 4 bps rate is neither so negligible (as at 1 and 2 bps) to be ignored, nor anywhere near as high as at 6 and 8 bps. Thus, 4 bps is selected as the threshold of noise perception in audio watermarking. The result is illustrated in Fig. 11.

Table 8 Comparison of improvement in payload

	Others techniques	Proposed technique
Harsh's Method [40]	PSNR: 62 dB payload: less than 1 bps	PSNR: 65 dB payload: 2 bps
Sos's Method [41]	PSNR: between 34 and 44 dB payload: 1/63	PSNR: 72 dB payload: 4/64
Foo's Method [42]	PSNR: 64.81 dB payload: 1 bps (in second layer)	PSNR: 65 dB payload: 2 bps (in first two layers)
Liu's Method [43]	PSNR: between 25 and 30 dB payload: 4 bps	PSNR: more than 52 dB payload: 4 bps
Cvejic's Method [44]	PSNR: 50 dB payload: 1 bps (less than 172 KBPS)	PSNR: more than 52 dB payload: 4 bps (more than 172 KBPS)
Nedeljko's Method [45]	PSNR: 55 dB payload: 1 bps (in fourth layer)	PSNR: 55.47 dB payload: 4 bps (in the first four layers)
Ahmad's Method [46]	PSNR: less than 40 dB payload: 6 bps	PSNR: more than 40 dB payload: 6 bps

www.Matlabi.ir

Table 10 The result of listening test

Sample No. (bps)	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9	No. 10	Average detection %
1	0	0	0	0	0	0	0	0	0	0	0
2	0	0	10	0	0	0	0	0	0	0	1
4	15	0	30	10	5	0	0	0	35	0	9.5
6	60	0	90	20	60	0	60	20	70	30	41
8	95	5	100	100	100	15	100	95	100	95	80.5

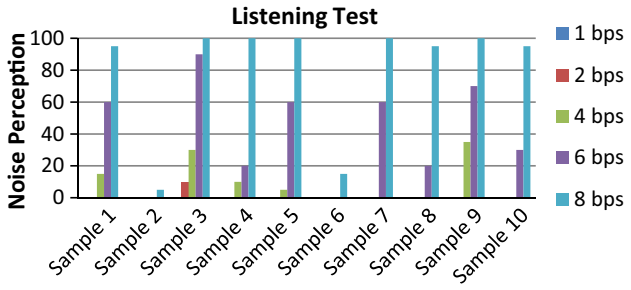


Fig. 11 The result of listening test

When the proposed technique (GSBAW) was implemented for 4 bps and the results were tested by listeners again, some samples were no longer tagged by listeners as noisy. This is illustrated in Table 11.

As illustrated in Fig. 12, there is an improvement in the result after using the genetic technique, and noise perception fell from 9.5 to 6 %.

10 Conclusions

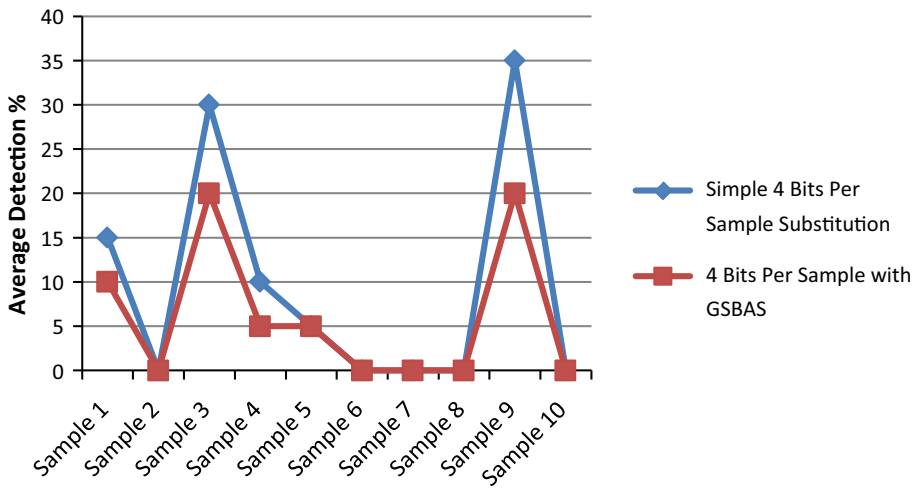
A novel algorithm was proposed as an additional modification step to substitute LSBs in audio watermarking in order to reduce the distortion of LSB substitution and consequently improve the PSNR of the resulting sound messages. This algorithm, which uses the concept of genetic algorithm, was implemented as a tool to embed various types of message files into a WAV file.

The improvement was demonstrated in both theoretical and practical ways. In the theoretical process, through a mathematical formulation, the number of possibilities whose bits were modified by the proposed algorithm was determined. Then the amount of improvement for all of the modified possibilities was calculated to get the average total expected improvement in PSNR. It was mathematically shown that the proposed method improves the PSNR of 2 bps payload techniques by 1.19 dB, 3 bps payload techniques by 1.54 dB, and 4 bps payload techniques by 1.63 dB.

Table 11 The comparison of listening test results

Sample No. (bps)	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9	No. 10	Average detection%
4 bps by simple LSB	15	0	30	10	5	0	0	0	35	0	9.5
4 bits Per sample by GSBAW	10	0	20	5	5	0	0	0	20	0	6

Fig. 12 The average detection for GSBAW and Simple LSB



www.Matlabi.ir

In the experiments to prove the improvement, a wide range of audio file samples were selected as host. In the first phase, the GSBAW was able perfectly to embed extra information into the same host while the PSNR remained the same. In the second phase, the GSBAW was able perfectly to embed the same size of information into the same host while the PSNR increased.

The imperceptibility of GSBAW was significantly better than other current substitution techniques of audio watermarking, and at the same time the payload was kept the same. It was shown that in practice the proposed method improved the PSNR of 2 bps payload techniques by 2.25 dB, 3 bps payload techniques by 2.80 dB, and 4 bps payload techniques by 3.07 dB.

Compared to other current substitution techniques of audio watermarking, the payload of GSBAW was much higher and at the same time the level of imperceptibility was maintained. Our research showed that the achieved 3.07 dB improvement in PSNR could double the payload of embedding.

Any watermarking technique is ineffective unless the noise caused by embedding the hidden message bits is imperceptible. On the other hand, it is also desirable to achieve the maximum possible payload in some applications. It is therefore important to find out the maximum achievable payload while maintaining imperceptibility.

To achieve this, a listening test was conducted to determine the threshold of noise perception in the Human Auditory System (HAS). It was observed that noise levels of 1 and 2 bps are absolutely imperceptible. On other hand, noise is quite perceptible at 6 and 8 bps. Hence, 4 bps was found to be the threshold for noise imperceptibility. In addition, this amount of improvement was able to decrease the amount of noise perception from 9.5 to 6 % in the listening test.

As a future work, proposed technique can be benchmarked with other recent techniques introduced in [47–60].

Acknowledgments This work is a part of a research that has been done and supported by Universiti Teknologi Malaysia, Malaysia.

References

1. Srivastava, M., & Rafiq, M. Q. (2011). A novel approach to secure communication using audio steganography. *Advanced Materials Research*, *403–408*(2012), 963–969.
2. Zamani, M., Taherdoost, H., Manaf, A. A., Ahmad, R. B., & Zeki, A. M. (2009). An artificial-intelligence-based approach for audio steganography. *MASAUM Journal of Open Probelms in Science and Engineering (MJOPSE)*, *1*(1), 64–68.
3. Zamani, M., Taherdoost, H., Manaf, A.A., Ahmad, R., & Zeki, A. (2009). Robust Audio Steganography via Genetic Algorithm. In *Third International Conference on Information & Communication Technologies ICICT2009* (pp. 149–153), 15–16, Karachi, Pakistan.
4. Abdullah, S.M., Manaf, A.A., & Zamani, M. (2010). Capacity and quality improvement in reversible image watermarking approach. in *6th International Conference on Networked Computing and Advanced Information Management* (pp. 16–18), Seoul, Korea.
5. Abdullah, S. M., Manaf, A. A., & Zamani, M. (2010). Recursive reversible image watermarking using enhancement of difference expansion techniques. *Journal of Information Security Research*, *1*(2), 64–70.
6. Zamani, M., Manaf, A.A., Ahmad, R., & Zeki, A. (2009). An approach to improve the robustness of substitution techniques of audio steganography. in *2nd IEEE International Conference on Computer Science and Information Technology 2009, Volume 2* (pp. 5–9), 8–11, Beijing, China.
7. Zamani, M., Manaf, A. A., Ahmad, R., Jaryani, F., Taherdoost, H., Chaeikar, S. S., et al. (2010). A novel approach for genetic audio watermarking. *Journal of Information Assurance and Security*, *5*, 102–111.
8. Zamani, M., Manaf, A. A., Ahmad, R., Jaryani, F., Taherdoost, Hamed, Chaeikar, S. S., et al. (2010). Genetic audio steganography. *International Journal on Recent Trends in Engineering & Technology [IJRTET]*, *3*(2), 89–91.
9. Zamani, M., Manaf, A.A., Ahmad, R., Jaryani, F., Taherdoost, H., & Zeki, A. (2009). A secure audio steganography approach. In *The 4th International Conference for Internet Technology and Secured Transactions* (pp. 501–506), 9–12, London, UK.
10. Zamani, M., Manaf, A.Bt.A., & Abdullah, S.M. (2012). Efficient embedding for audio steganography. In *2nd International Conference on Environment, Economics, Energy, Devices, Systems, Communications, Computers, Mathematics (EDSCM '13)*, France, April 2–4.
11. Zamani, M., Manaf, A.Bt.A., Abdullah, S.M., & Chaeikar, S.S. (2012). Correlation between PSNR and bit per sample rate in audio steganography. in *11th International Conference on Signal Processing (SIP '12)*, Saint Malo & Mont Saint-Michel, France, April 2–4.
12. Zamani, M., Manaf, A.Bt.A., & Abdullah, S.M. (2012). Correlation between PSNR and size ratio in audio steganography. In: *11th International Conference on Telecommunications and Informatics (TELE-INFO '12)*. Saint Malo & Mont Saint-Michel, France, April 2–4.
13. Zamani, M., Manaf, A Bt A, & Abdullah, S. M. (2012). An overview on audio steganography techniques. *International Journal of Digital Content Technology and its Applications (JDCTA)*, *6*, 107–122.
14. Zamani, M., Manaf, A.A., & Daruis, R. (2012). Azizah technique for efficiency measurement in steganography. In *8th International Conference on Digital Content, Multimedia Technology and its Applications*, June 26–28, Jeju, Korea.
15. Zamani, M., Manaf, A Bt A, Abdullah, S. M., & Chaeikar, S. S. (2012). Mazdak technique for PSNR estimation in audio steganography. *Applied Mechanics and Materials*, *229—-231*, 2798–2803.
16. Oludele, A. (2010). Artificial intelligence and security. *International Journal of Computational Intelligence and Information Security 1*(3).
17. Zeki, A. M. (2011). A robust watermark embedding in smooth areas. *Research Journal of Information Technology*, *3*(2), 123–131. doi:10.3923/rjit.2011.123.131.
18. Zamani, M., Manaf, A.A., & Ahmad, R. (2009) Current problems of substitution techniques of audio steganography. In *The 2009 International Conference on Artificial Intelligence and Pattern Recognition* (pp. 154–160), 13-16, Orlando, Florida, USA.
19. Cvejic, N. (2004). *Algorithms for audio watermarking and steganography*. Oulu: University of Oulu.
20. Wang, R.-Z., Lin, C.-F., & Lin, J.-C. (2001). Image hiding by optimal LSB substitution and genetic algorithm. *Pattern Recognition*, *34*(3), 671–683.

www.Matlabi.ir

21. Wu, M.-N., Lin, M.-H., & Chang, C.-C. (2004). A LSB substitution oriented image hiding strategy using genetic algorithms. *Lecture Notes in Computer Science*, *3309*, 219–229.
22. Luo, W., Huang, F., & Huang, J. (2010). Edge adaptive image steganography based on LSB matching revisited. *IEEE Transactions on Information Forensics and Security*, *5*(2), 201–214.
23. Ji, R., Yao, H., Liu, S., & Wang, L. (2006). Genetic algorithm based optimal block mapping method for LSB substitution. In *International Conference on Intelligent Information Hiding and Multimedia Signal Processing* (pp. 215–218).
24. Wang, R., Lin, C., & Lin, J. (2001). Image hiding by optimal LSB substitution and genetic algorithm. *Pattern Recognition*, *34*(3), 671–683.
25. Liu, S., Chen, T., Yao, H., & Gao, W. (2004). A variable depth LSB data hiding technique in images. *International Conference on Machine Learning and Cybernetics*, *7*, 3990–3994.
26. Kalker, T., & Haitsma, J. (2000). Efficient detection of a spatial spread-spectrum watermark in mpeg video streams. In *Proceedings of IEEE International Conference on Image Processing* Vancouver, BC (pp. 407–410).
27. Pickholtz, R. L., Schilling, D. L., & Milstein, L. B. (1982). Theory of spread-spectrum communications: A tutorial. *IEEE Transactions on Communications*, *30*(5), 855–884.
28. Zamani, M., Manaf, A.A., & Ahmad, R. (2009). Knots of substitution techniques of audio steganography. In *The 2009 International Conference on Telecom Technology and Applications* (pp. 415-419), June 6–8, Manila, Philippines.
29. Zamani, M., Manaf, A. A., Zeidanloo, H. R., & Chaeikar, S. S. (2011). Genetic subtitution-based audio steganography for high-capacity applications. *International Journal for Internet Technology and Secured Transactions (IJITST)*, *3*(1), 97–110. doi:10.1504/IJITST.2011.039681.
30. Zamani, M., Manaf, A.A., Ahmad, R., Zeki, A., & Abdullah, S. (2009) Genetic algorithm as an approach to resolve the problems of substitution techniques of audio steganography. In *The 2009 International Conference on Genetic and Evolutionary Methods* (pp. 170-175), 13–16 Las Vegas, Nevada, USA.
31. Zamani, M., Manaf, A.A., Ahmad, R., Zeki, A., & Abdullah, S. (2009) A genetic-algorithm-based approach for audio steganography. In *International Conference on Communities and Communications. World Academy of Science, Engineering and Technology 54* (pp. 359–363), 24–26 June 2009, Paris, France.
32. Zamani, M., Manaf, A.A., Ahmad, R., Zeki, A., & Magalingam, P. (2009). A novel approach for audio watermarking. In: *5th International Conference on Information Assurance and Security* (pp. 83-86), 18–20 Xi'an, China.
33. Zeki, A.M., Manaf, A.A., & Zamani, M. (2010). Bit-plane model: Theory and implementation. In *Engineering Conference 2010 (EnCon2010)*, April 14–16. Kuching, Sarawak, Malaysia.
34. Zamani, M., & Manaf, A.A. (2010). Azizah's formula to measure the efficiency of steganography techniques. In *2nd International Conference on Information and Multimedia Technology (ICIMT 2010)*. 28–30 December, Hong Kong, China.
35. Zamani, M., & Manaf, A.A. (2010). Mazdak's method to estimate the PSNR of audio steganography techniques. In *International Conference on Computer and Computational Intelligence (ICCCI 2010)*. 25–26 December, Nanning, China.
36. Darsana, R., & Vijayan, A. (2011). Audio steganography using modified LSB and PVD. Trends in network and communications. *Communications in Computer and Information Science*, *197*(1), 11–20.
37. Wang, S., Song, X., & Niu, X. (2012). An affine transformation based image steganography approach. *International Journal of Digital Content Technology and its Applications*, *6*(1), 14–85.
38. Qin, J. (2011). An improved method to secret message length estimation in ±k embedding steganography. *Journal of Convergence Information Technology*, *6*(5), 339–346.
39. Xiang, L., Sun, X., Luo, G., & Xia, B. (2011). Steganalysis of syntactic transformation based steganography. *International Journal of Digital Content Technology and its Applications*, *5*(5), 320–330.
40. Verma, H., Kaur, R., & Kumar, R. (2009) Random sample audio watermarking algorithm for compressed wave files. *IJCSNS International Journal of Computer Science and Network Security, 9*(11).
41. Agaian, S.S., Akopia, D., & D'Souza, S.A. (2005) Two Algorithms in digital audio steganography using quantized frequency domain embedding and reversible integer transforms. Non-linear Signal Processing Lab, University of Texas at San Antonio, 6900 North Loop 1604 West, San Antonio, Texas 78249, USA.
42. Foo, S.W., & Dong, Q. (2009). Robustness of one bit per sample audio watermarking. In: *Source: Proceedings of IEEE International Symposium on Circuits and Systems ISCAS* (pp. 932–935) 2009.
43. Liu, S., Yao, H., Gao, W., & Yang, D. (2007). Minimizing the distortion spatial data hiding based on equivalence class. Source: Lecture Notes in Computer Science, v 4681 LNCS, pp. 667–678, 2007. In *Advanced Intelligent Computing Theories and Applications: With Aspects of Theoretical and Methodological Issues Proceedings of 3rd International Conference on Intelligent Computing, ICIC*.
44. Cvejic, N., & Seppanen, T. (2002). Increasing the capacity of LSB-based audio steganography. In *IEEE Workshop on Multimedia Signal Processing* (pp. 336–338).
45. Cvejic, N., & Seppäne., T. (2005). Increasing robustness of LSB audio steganography by reduced distortion LSB coding. *Journal of Universal Computer Science*, *11*(1), 56–65.
46. Delforouzi, A., & Pooyan, M. (2008). Adaptive digital audio steganography based on integer wavelet transform. *Circuits, Systems, and Signal Processing*, *27*(2), 247–259.
47. Zhang, X., & Wang, S. (2012). Efficient data hiding with histogram-preserving property. *Telecommunication Systems*, *49*(2), 179–185. doi:10.1007/s11235-010-9364-5.
48. Mazurczyk, W., & Szczypiorski, K. (2012). Evaluation of steganographic methods for oversized IP packets. *Telecommunication Systems*, *49*(2), 207–217. doi:10.1007/s11235-010-9362-7.
49. Szczypiorski, K. (2012). A performance analysis of HICCUPS-a steganographic system for WLAN. *Telecommunication Systems*, *49*(2), 255–259. doi:10.1007/s11235-010-9363-6.
50. Mazurczyk, W., & Lubacz, J. (2010). LACK-a VoIP steganographic method. *Telecommunication Systems*, *45*(2–3), 153–163. doi:10.1007/s11235-009-9245-y. Special Issue: SI.
51. Huang, H.-C., & Fang, W.-C. (2010). Techniques and applications of intelligent multimedia data hiding. *Telecommunication Systems*, *44*(3–4), 241–251. doi:10.1007/s11235-009-9262-x. Special Issue: SI.
52. Maity, S. P., Maity, S., & Sil, J. (2012). Multicarrier spread spectrum watermarking for secure error concealment in fading channel. *Telecommunication Systems*, *49*(2), 219–229. doi:10.1007/s11235-010-9369-0.
53. He, H., & Zhang, J. (2012). Cryptanalysis on majority-voting based self-recovery watermarking scheme. *Telecommunication Systems*, *49*(2), 231–238. doi:10.1007/s11235-010-9380-5.
54. Zhang, L., & Zhou, P.-P. (2010). Localized affine transform resistant watermarking in region-of-interest. *Telecommunication Systems*, *44*(3–4), 205–220. doi:10.1007/s11235-009-9260-z. Special Issue: SI.
55. Singh, R., Vatsa, M., Singh, S. K., & Upadhyay, S. (2009). Integrating SVM classification with SVD watermarking for intelligent video authentication. *Telecommunication Systems*, *40*(1–2), 5–15. doi:10.1007/s11235-008-9141-x.

www.Matlabi.ir

56. Jabbar, Alaa A., Sahib, Shahrin Bin., & Zamani, Mazdak. (2013). Multimedia Data Hiding Evaluation Metrics. In *7th WSEAS International Conference on Computer Engineering and Applications (CEA '13)*. 9–11 January, Milan, Italy.
57. Jabbar, A.A., Sahib, S.B., & Zamani, M. (2013) An introduction to watermarking techniques. In *12th WSEAS International Conference on Applications of Computer Engineering (ACE '13)*. January 30 to February 1, Cambridge, MA, USA.
58. Zamani, M., Manaf, A.Bt.A., Abdullah, S.M., & Chaeikar, S.S. (2012). Correlation between PSNR and bit per sample rate in audio steganography. In *11th International Conference on Signal Processing (SIP '12)* (pp. 163–168). April 2–4, Saint Malo & Mont Saint-Michel, France.
59. Zamani, M., Manaf, A.Bt.A., & Abdullah, S.M. (2012) Correlation between PSNR and size ratio in audio steganography. In *11th International Conference on Telecommunications and Informatics (TELE-INFO '12)* (pp. 82–87). April 2–4, Saint Malo & Mont Saint-Michel. France.
60. Zamani, M., Manaf, A.Bt.A., & Abdullah, S.M. (2012) Efficient embedding for audio steganography. In *2nd International Conference on Environment, Economics, Energy, Devices, Systems, Communications, Computers, Mathematics (EDSCM '13)* (pp. 195–199). April 2–4, Saint Malo & Mont Saint-Michel, France.

Mazdak Zamani received his Ph.D. degree in 2010 on the topic of "Genetic based substitution techniques for audio steganography" from Universiti Teknologi Malaysia (UTM), Malaysia. He was then appointed as Visiting Lecture at UTM until 2012. He has been with UTM as Senior Lecturer since 2012. His main research interests include Multimedia Security, Wireless Security and Secure Architecture and Models

Azizah Bt Abdul Manaf (Ph.D.) is a Professor of Image Processing and Pattern Recognition at Universiti Teknologi Malaysia (UTM). She graduated with B.Eng. (Electrical-Communication & Control) in 1980, M.Sc. Computer Science (1985) and Ph.D. (Image Processing) in 1995. Her current research areas are in Image Processing and Pattern Recognition, Information Security, Watermarking, Steganography and Digital Computer Forensics. She has supervised and graduated a large number of postgraduate students at the Masters and PhD level in these areas and has also authored and co-authored a number of computer related books, written numerous articles in journals and has presented an extensive amount of research papers at national and international conferences on her area of expertise. She has also been invited as keynote speakers and academic talks on her research area at international and national conferences and seminars. Besides being member of ACM, IEEE Computer Society and SDIWC Advisory Board Member, Prof. Dr. Azizah is also currently the President of Malaysia IRRS (International Rough Set Society)-Malaysian Chapter.

www.Matlabi.ir