JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY
https://e-journal.uum.edu.my/index.php/jict

How to cite this article:
Noor Azam, M. H., Ridzuan, F., & Mohd Sayuti, M. N. S. (2023). Optimised cover selection for audio steganography using multi-objective evolutionary algorithm. *Journal of Information and Communication Technology, 22*(2), 255-282. https://doi.org/10.32890/jict2023.22.2.5

Optimised Cover Selection for Audio Steganography Using Multi-Objective Evolutionary Algorithm

*[1]Muhammad Harith Noor Azam, [2]Farida Ridzuan
& [3]M Norazizi Sham Mohd Sayuti
[1,2]Faculty of Science and Technology,
Universiti Sains Islam Malaysia, Malaysia
[3]Faculty of Engineering and Built Environment,
Universiti Sains Islam Malaysia, Malaysia

*[1]mh.noorazam@gmail.com
[2]farida@usim.edu.my
[3]azizi@usim.edu.my
*Corresponding author

Received: 14/4/2022 Revised: 25/1/2023 Accepted: 9/2/2023 Published: 3/4/2023

ABSTRACT

Existing embedding techniques depend on cover audio selected by users. Unknowingly, users may make a poor cover audio selection that is not optimised in its capacity or imperceptibility features, which could reduce the effectiveness of any embedding technique. As a trade-off exists between capacity and imperceptibility, producing a method focused on optimising both features is crucial. One of the search methods commonly used to find solutions for the trade-off problem in various fields is the Multi-Objective Evolutionary

Algorithm (MOEA). Therefore, this research proposed a new method for optimising cover audio selection for audio steganography using the Non-dominated Sorting Genetic Algorithm-II (NSGA-II), which falls under the MOEA Pareto dominance paradigm. The proposed method provided suggestions for cover audio to users based on imperceptibility and capacity features. The sample difference calculation was initially formulated to determine the maximum capacity for each cover audio defined in the cover audio database. Next, NSGA-II was implemented to determine the optimised solutions based on the parameters provided by each chromosome. The experimental results demonstrated the effectiveness of the proposed method as it managed to dominate the solutions from the previous method selected based on one criterion only. In addition, the proposed method considered that the trade-off managed to select the solution as the highest priority compared to the previous method, which put the same solution as low as 71 in the priority ranking. In conclusion, the method optimised the cover audio selected, thus, improving the effectiveness of the audio steganography used. It can be a response to help people whose computers and mobile devices continue to be unfamiliar with audio steganography in an age where information security is crucial.

Keywords: Audio Steganography, Cover Audio Selection, Multi-objective Optimisation Problem, Trade-off.

INTRODUCTION

Audio steganography aims to transmit data secretly by camouflaging a secret message inside an audio file (Dutta et al., 2019; Khairullah, 2019; Setiadi, 2022; Ye et al., 2019). Several audio steganography techniques have been proposed in previous studies, including low-bit embedding, parity coding, echo hiding, phase coding, spread spectrum, and wavelet domain (Ahani et al., 2015; Somani & Madhu, 2015). A successful audio steganography technique must comply with capacity, imperceptibility, and robustness conditions (Alsabhany et al., 2019; Amirtharajan & Rayappan, 2013; Zumchak, 2016).

Capacity refers to the amount of secret message that can be embedded within cover audio, while robustness is the ability of a secret message to withstand detections or attacks (Ballesteros & Moreno, 2012; Srivastava & Rafiq, 2012; Vimal, 2014; Zumchak, 2016). Lastly, imperceptibility is a degree of the ability to embed hidden code without distorting the audio signal perceived by human hearing (Zumchak, 2016). These three trade-offs exist in audio steganography: (1) between imperceptibility and capacity (Rustad et al., 2022), (2) between capacity and robustness (Durafe and Patidar, 2022), and (3) between robustness and imperceptibility (Kaur et al., 2017).

In the trade-off between imperceptibility and capacity, if the capacity is increased, the imperceptibility will decrease significantly, along with more obvious noise and vice versa (Indrayani, 2020). Similarly, in the trade-off between capacity and robustness, if the capacity is increased, robustness will decrease significantly and vice versa (Ali et al., 2017; Djebbar et al., 2012). Finally, in the trade-off between imperceptibility and robustness, if the imperceptibility is increased, robustness will decrease and vice versa (Gopalan & Shi, 2010).

A new audio steganography technique is commonly created by examining specific characteristics and parameters, proposing a solution, and evaluating the method (Wakiyama et al., 2010). However, despite the numerous enhancements provided by researchers, the quality of cover audio used in the audio steganography technique is often neglected. As a result, the end user is expected to choose any audio and input a secret text message without considering these characteristics. The cover audio needs to be selected cautiously to ensure efficient audio steganography. Poor cover audio selection will decrease the quality of a stego file (i.e., cover audio embedded with a secret message), specifically in its capacity or imperceptibility. Therefore, to tackle this quality problem, the current approach is to focus either on the audio steganography method's capacity or imperceptibility. Nevertheless, based on the review of existing studies, it can be concluded that research to address the optimal solution of this specific trade-off is still lacking. This trade-off is called multi-objective optimisation problem (MOP), as improving either capacity or imperceptibility leads to the degradation of the other characteristic. MOP can usually be solved by using Multi-Objective Evolutionary Algorithm (MOEA) (Yasear & Ku-Mahamud, 2021). Therefore, this research implemented MOEA to fill the research gap and maximise the imperceptibility and capacity features of the stego file by finding the optimal solution, which is a well-rounded performance and not just skewed heavily to one characteristic for cover audio selection. MOEA is a method to optimise the objective functions involved and process appropriate trade-offs from the decision-maker perspective (Chiandussi et al., 2012).

Problems with multiple objectives exist in numerous disciplines, where researchers are challenged to find solutions. MOEA is used for problem-solving of such a nature because of its ability to generate a list of solutions in a single run (Zhou et al., 2011). Nevertheless, it is currently seldom used in steganography due to a lack of research combining these two domains.

This research proposes a new method for optimised cover audio selection in audio steganography based on the imperceptibility and capacity features. The chromosome, i.e., the set of parameters defining the solution to the trade-off, was used to provide the optimised solution based on capacity calculation and imperceptibility ranking. NSGA-II was implemented to search and optimise solutions within the parameters defined. With the selected cover, the suggestion could be guaranteed as optimised cover audio for capacity and imperceptibility features. The contributions of this research are as follows:

- A new framework for the cover audio selection based on MOEA.
- Formulation of the fitness function for capacity and imperceptibility characteristics of audio steganography.
- An optimisation method based on the Least Significant Bit (LSB) that adapts MOEA for cover audio selection.

The rest of this paper is organised as follows. In the Related Studies section, LSB, MOEA, and cover audio selection are explained in detail. The research approach is elaborated in the Research Methodology section. The proposed cover selection method is described in the Proposed Work section. Experimental results and analysis performance are presented in the Results and Performance Analysis section. Finally, the Conclusion section concludes the whole paper and suggests future research.

RELATED STUDIES

LSB is one of the famous methods used for embedding in steganography and can be enhanced further by implementing cover audio selection.

The cover selection method can choose optimised audio based on some metrics that evaluate the audio steganography characteristic performances, thus, further improving the characteristic. As these characteristics have a trade-off between them, this research focuses more on MOEA, which is a searching technique famously used to find the optimised solution. This section briefly explains LSB, MOEA, and cover audio selection.

LSB

Low-bit embedding, famously known as LSB, was first proposed by Bender (1996). Since then, many researchers have improved the technique by enhancing its capacity, imperceptibility, and robustness. In the LSB embedding technique, both cover audio and secret messages are converted into their stream of binary value (Hameed et al., 2019; Nursalman et al., 2018; Solak, 2020). Then, the LSBs of audio samples are replaced with the sequence of bits of the secret message (Jayapandiyan et al., 2020; Sahu & Swain, 2022).

There are two categories of LSB, namely direct embedding and selective embedding (Tabares-Soto et al., 2019). The direct LSB embedding technique embeds data into the first sample without skipping any audio samples throughout the process until all data are successfully embedded, such as proposed by Bazyar and Sudirman (2015), Bender (1996), and Cvejic and Seppänen (2002). On the other hand, the selective LSB embedding technique only selects samples that pass the embedding condition before embedding the data until it produces a favourable outcome, as proposed by Alsabhany et al. (2020) and Wakiyama et al. (2010).

Additionally, a common way of developing a new method is by first considering certain features and parameters, then proposing a solution, and finally, evaluating the method's capacity and imperceptibility (Wakiyama et al., 2010). Despite many improvements proposed by researchers, the capacity and imperceptibility can be improved by choosing good cover audio. When using existing algorithms, the common practice of the end-user choosing audio and inserting a secret message without considering the capacity and imperceptibility features may lead to poor cover audio selection. Therefore, cover audio selection implementation can further improve this practice.

MOEA

MOEA is a widely recognised searching technique because of its practicality in research and real-world applications. It has been applied to various areas, such as engineering, security, and military, especially in the last two decades (Deb, 2000; Li et al., 2015; Mashwani et al., 2016). Furthermore, the technique does not require any derivative information and is simple to implement (Deb, 2000). In the past few years, several MOEAs have been developed to optimise real-world tasks, such as Non-dominated Sorting Genetic Algorithm (NSGA) and NSGA-II.

NSGA was introduced in 1994 by Siinivas and Deb (1994). This algorithm sorts the population by assigning rank zero (0) to the current non-dominated subset of the population and temporarily removing that subset from consideration. The remaining population is analysed to determine another non-dominated sub-set, which is then given rank one (1) and removed from consideration. The process is carried on until the entire population has been ranked. On the other hand, Deb et al. (2002) introduced NSGA-II, which is currently one of the most popular MOEAs. However, it does not have many similarities compared to the original NSGA. The authors kept the name NSGA-II to emphasise its genesis and place of origin (Deb, 2001). In addition, NSGA-II overcomes the drawbacks of NSGA, such as high computational complexity on non-dominated sorting, lack of elitism, and a need for specifying the sharing parameter (Kunkle, 2005). Considering these advantages, MOEA is also used in steganography to optimise features, either capacity, imperceptibility, or robustness. For instance, a recent study by Hajduk and Levický (2018) combined both MOEA and steganography for image steganography and implemented NSGA to optimise the stego file. The study proposed a new method of embedding a secret message within the discrete wavelet transform using singular value decomposition (SVD) and NSGA. Based on the algorithm, SVD was applied to a high-frequency band to acquire singular values that were later replaced by singular values of the secret message (image). The embedding capacity and the image quality were improved by increasing the image's imperceptibility feature. Therefore, the practicality of MOEA as a searching method can be expanded to find optimised cover audio.

Cover Selection

Most research focuses on finding different methods to hide a secret message. However, the result may not always be good since it depends on the cover file and the secret message (Shah & Bichkar, 2020). The main idea is to find the optimal cover file from its database using a specific search method for the specific secret message (Hajduk & Levický, 2018). If the criterion of a good cover file is met based on comparison parameters, the relevant cover file will be used as a carrier for the secret message. For image steganography, many studies have been introduced for cover image selection (Abdul Sattar & Talib Gaata, 2017; Hajduk & Levický, 2018; Sajedi & Jamzad, 2008; Shah & Bichkar, 2020; Wang et al., 2020; Wang & Zhang, 2019). The same goes for cover video selection in video steganography (Wang et al., 2019). Nevertheless, to the authors' knowledge, research on cover audio selection had never been conducted before. Therefore, this research was mainly motivated to fill this gap by finding a feasible method for cover audio selection to cater to the trade-off between characteristics in audio steganography.

RESEARCH METHODOLOGY

This section explains the research methodology used in this paper. In audio steganography, a similar research methodology has been used. This research methodology consists of three stages: identifying the problem, proposing the solution, and evaluating the proposed solution.

For the first stage, this research identified the research problem by reviewing previous literature in the Related Studies section. This research intended to solve the trade-off issue between characteristics in audio steganography using a cover selection method that is yet non-existent.

Next, in the second stage, this research proposed a new solution for the trade-off issue by optimising the trade-off between capacity and imperceptibility during the cover audio selection process. This proposed solution is explained further in the Proposed Work section.

Lastly, the proposed solution was evaluated in the final stage. Three evaluations were conducted, which were explained in detail, including the parameters used. All the evaluations and their results are presented along with their analysis in the Results and Performance Analysis section.

PROPOSED WORK

This section explains the proposed work on the optimised cover audio selection. The method aims to obtain audio steganography with high capacity and imperceptibility by optimising the trade-off between capacity and imperceptibility during cover audio selection. The proposed work is explained further in the subsections below.

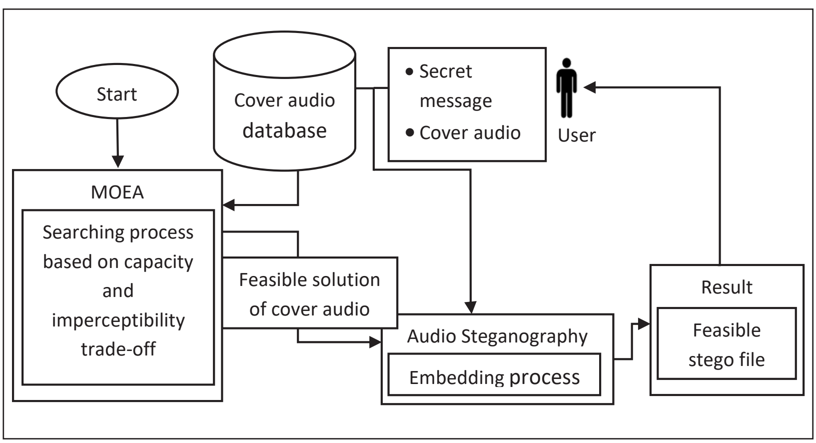
General Framework

A general framework for cover audio selection optimisation was proposed to fulfil the stated goal, as shown in Figure 1. This framework served as a basis for cover audio selection optimisation, involving a multi-objective problem based on capacity and imperceptibility characteristics while considering its trade-off. For the value of a good trade-off between imperceptibility and capacity, this framework considered all cover audios that achieved 30 dB of Signal-to-Noise Ratio (SNR) following Singh (2014). Meanwhile, the capacity was fixed based on user input. The following two objectives were considered in this research:

- Maximising the capacity using selected cover audio.
- Maximising the imperceptibility of selected cover audio.

Figure 1

General Framework of Cover Audio Selection



Based on the general framework shown in Figure 1, two fitness functions were used to rank the solution in terms of capacity and imperceptibility. Fitness function is also known as selection criteria.

The secret message is embedded into the audio by replacing the audio sample with its stream of binary values. Therefore, the number of the cover audio sample ($size_{sample}$) was calculated based on the audio duration (*d*) and sample rate (*sr)*. The $size_{sample}$ was modelled as in Equation 1:

$$size_{sample} = d * sr \tag{1}$$

In achieving a maximum potential capacity of audio steganography, $size_{sample}$ needs to be fully utilised. Therefore, the difference between the used sample and the unused sample (dif_{sample}) need to be kept at a minimum. However, since the objective is to maximise the value, the (dif_{sample}) value needs to be inversed to convert it from the minimum to the maximum value. This difference was modelled as in Equation 2:

$$dif_{sample} = \frac{1}{size_{sample} - size_{sampleused}} \tag{2}$$

where $size_{sampleused}$ was calculated based on the size of the binary message ($size_{sample}$) and bit embedded per sample (*bps)*. It was modelled as in Equation 3:

$$size_{sampleused} = size_{binmsg} - bps \tag{3}$$

On the other hand, imperceptibility was commonly measured using SNR. It was calculated using Equation 4:

$$\text{SNR} = 10 * log_{10} \frac{\sum_1^n x^2}{n * MSE} \tag{4}$$

Where *MSE* was calculated based on the original cover audio sample (*x*), stego file sample (*y*), and the number of the total sample (*n*), as computed using Equation 5:

$$\text{MSE} = \frac{\sum_1^n (x-y)^2}{n} \tag{5}$$

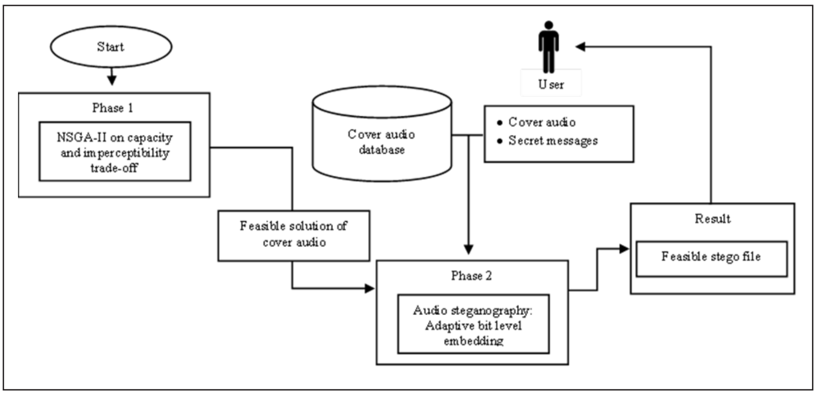
Based on the general framework and fitness function developed, this research proposed a method for cover audio selection optimisation in audio steganography using NSGA-II and LSB embedding techniques based on the number of bit levels.

Proposed Method

The proposed method is an optimisation technique based on LSB and adapted MOEA, consisting of two main phases: (1) multi-objective problem optimisation using NSGA-II, and (2) adaptive LSB embedding based on *bps*. In particular, the proposed method included an adapted NSGA-II and a new embedding technique. Figure 2 illustrates the proposed method, and the following briefly describes the two phases involved.

Figure 2

Framework of the Proposed Cover Audio Selection Optimisation in Audio Steganography Using NSGA-II and Adaptive LSB Embedding Technique Based on the Bit Level

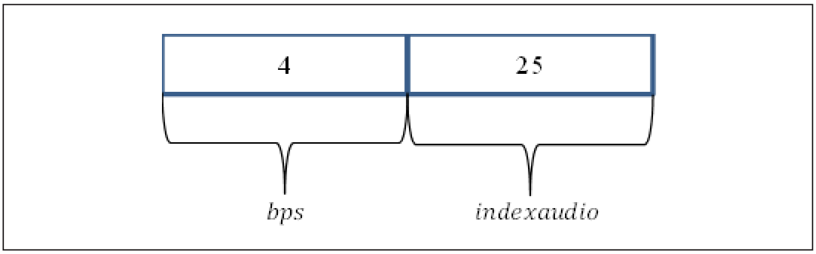


Phase 1: In this phase, all parameters, including probability for crossover and mutation, number of generations, size of chromosome in the population, and number of alleles (i.e., specified set of alternatives of parameters), in each chromosome were tuned specifically for this method. The *bps* value was set from 1 to 8, and the cover audio (*indexaudio*) value was set from 1 to 30. The size of the chromosome population was fixed at 30, and the number of generations was set to 15. Real value representation was used to map the space of the solution available for the cover selection problem. Therefore, the number of alleles used was 2: the first allele was employed for embedding

parameters based on *bps*, followed by the second allele, which was utilised to determine cover audio based on its index number. Figure 3 shows an example of parameters in the form of alleles for one chromosome.

Figure 3

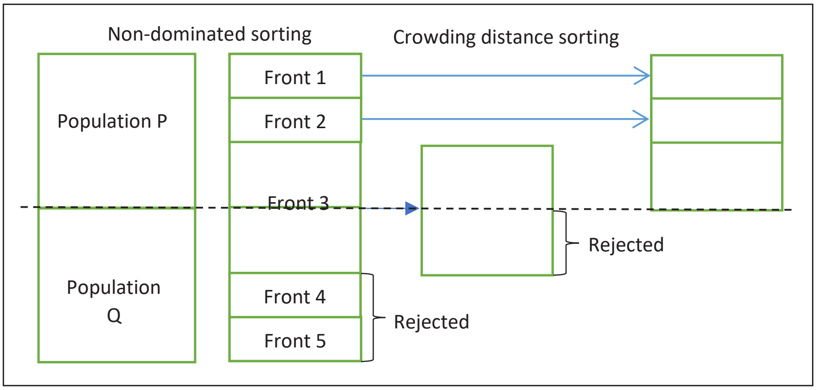
Example of Parameters in Alleles Form



The solutions underwent several processes, such as fast non-dominated sorting, selection, crossover, mutation, and fitness function evaluation. The crossover probability and mutation probability were set to 0.8 and 0.05. The fitness function, a particular objective function, characterised the problem and measured how close a given solution could achieve the target while considering all problem constraints defined in the framework. The fitness function evaluation was carried out using Equations 2 and 4. Meanwhile, the fast non-dominated sorting procedure sorted all individuals according to the level of non-domination to enhance convergence properties. Selection, crossover, and mutation were then applied to create new populations, which were remeasured in terms of their closeness to the target using the constraints defined in the framework. Then, all solutions (parents and child) were sorted to ensure convergence. The individual selection procedure was employed to create a mating pool by combining parent and child populations and selecting the best solutions (in terms of spread and fitness). This method adopted a suitable automatic mechanism based on the crowding distance (a measure of the density of solutions in a neighbourhood) to guarantee the diversity and spread of solutions. This method also produced the potential solution for selecting cover audio used in the next phase. Figures 4 and 5 further illustrate the NSGA-II operation.

Figure 4

Flow Diagram of NSGA-II

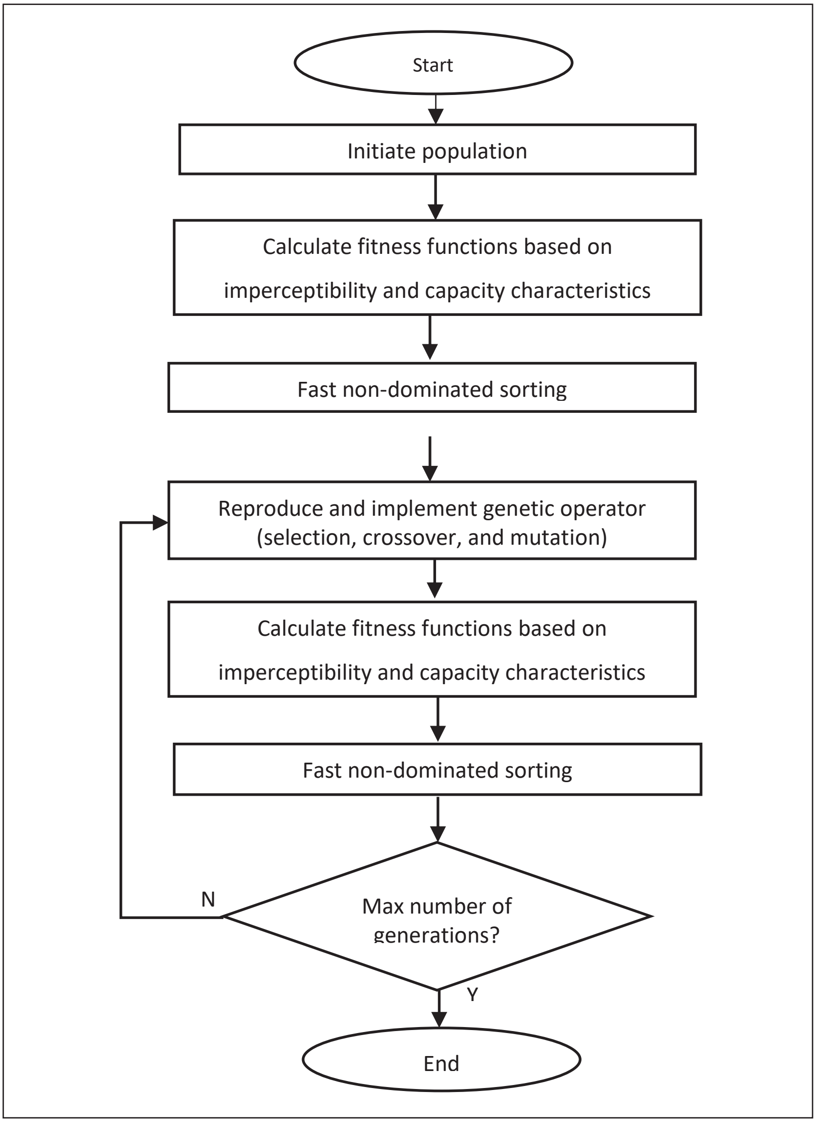


As shown in Figure 4, the algorithm first built a population of competing individuals, ranked, and sorted each individual according to the non-domination level, and applied evolutionary operations (EVOPs) to create a new pool of offspring. Then, the parents and offspring were combined (to create a new combined pool) before being partitioned into fronts. A niching process was then conducted by adding a crowding distance to each member of the fronts. This distance was used in the selection operator for keeping a diverse front by ensuring each member stayed a crowding distance apart. This approach maintained the population diversity and helped the optimisation process to explore the fitness landscape.

As seen from Figure 5, after the population was initialised, the ranking based on the capacity and imperceptibility was determined by calculating the fitness function. In non-dominated sorting, chromosome A was said to dominate another chromosome B if and only if A had no objective worse than B's and at least one better objective. The discovered non-dominated solutions were put in the first ranking, or the Pareto front, where the rest was put in the fronts 2, 3, and so on. Then, a new generation was produced. The fast non-dominated sorting was repeated until all solutions were found. These solutions were passed to the next phase to determine the suitable solution based on cover audio in the database and the secret message from the user. Next, a new cover audio was added to the audio database, while the secret message was changed into a stream of binary bits, and Phase 2 started.

Figure 5

Flow Chart of NSGA-II

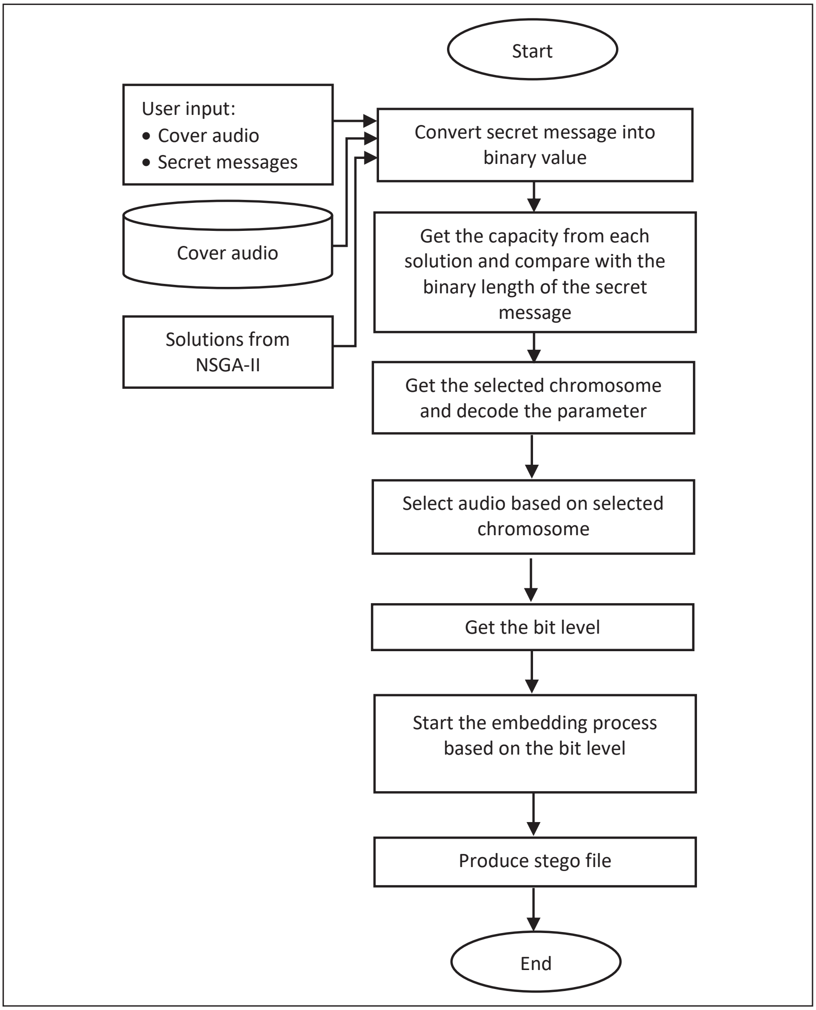


Phase 2: The embedding process was implemented based on the solution obtained from the previous phase. Figure 6 shows the flow chart of the proposed embedding technique. First, the secret message was converted into a stream of binary bits. Then, all solutions were

evaluated, and the capacity for each solution was compared with the secret message's binary stream length. The solution with the nearest capacity value to the secret message's length was selected. Next, the selected solution underwent decoding to gain the bit level and audio duration. The cover audio with the same duration as the solution was selected. Then, the secret message was embedded based on the bit level from the solution. Finally, a stego file was produced.

Figure 6

Flow Chart of the Proposed Embedding Technique



RESULTS AND PERFORMANCE ANALYSIS

This section presents the analysis of the performance of the proposed method. Three experiments were conducted: (1) fine-tuning the size of the population used in the proposed cover selection method, (2) fine-tuning the number of generations used in the proposed cover selection method, and (3) a comparison analysis between the selection criteria. Fine-tuning experiments were performed on the NSGA-II parameters to ensure all optimal solutions could be found, while a comparative analysis was conducted to investigate the performance of the proposed method. The audio format selected for all these experiments was a mono wave file with a 44.1kHz sampling rate and 16-bit per sample. These audios were selected based on audio length from the combination of audio from speech and music. They were downloaded from a free online audio library, www.freesound.org. The selection criteria used for the comparative analysis were adapted from Xin and Jiaojiao (2018), which employed selection criteria based on the difference between used and unused pixels for embedding the image and the distortion value of the image after the embedding process. Therefore, the difference criterion in this research was implemented using Equation 2, and the distortion value was computed using Equation 4.

Fine-Tuning the Number of Generations Used

This experiment aimed to determine the best number of generations used to find the optimised solution. The setup used for this fine-tuning is shown in Table 1, while Figure 7 indicates the results of the experiment.

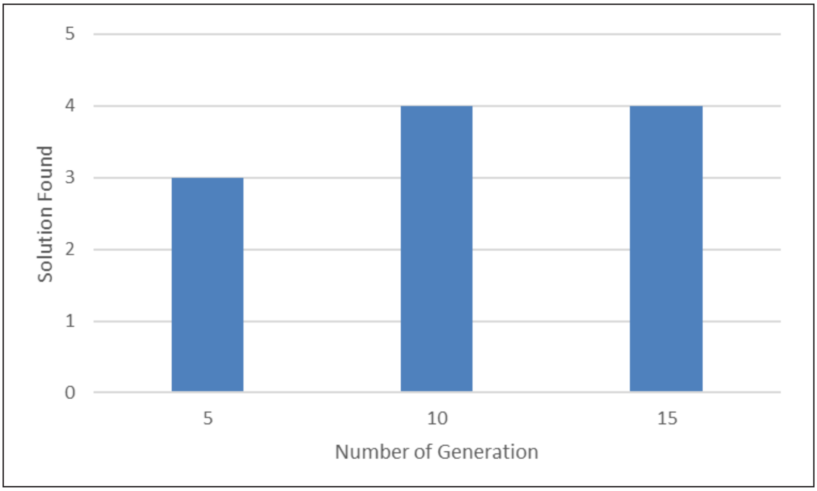
Table 1

Parameters Setup for Tuning the Number of Generations Used

Parameters	Value
Number of audios in database	30
Bps	1-8
Crossover probability	0.8
Mutation probability	0.05
Number of generations	5, 10 and 15
Number of populations	30

Figure 7

Number of Solutions Found Based on Number of Generations



The results in Figure 7 demonstrated that the number of solutions found increased as the number of generations increased. The lowest number of solutions was found when five generations were used, whereas the maximum number of solutions was found when ten and fifteen generations were used. Due to the same result produced by the ten and fifteen generations, it was better to implement ten as the number of generations to reduce the execution time for the selection process. Overall, the results revealed that the best number of chromosomes to use were ten and fifteen generations, with the difference in execution time only.

Fine-Tuning the Number of Chromosomes Used

The objective of this experiment was to ascertain the best number of chromosomes per population used to find the optimised solution. The setup used for this fine-tuning is shown in Table 2, while Figure 8 presents the results of the experiment.

Table 2

Parameters Setup for Tuning the Number of Chromosomes used

Parameters	Value
Number of audios in database	30
Bps	1-8
Crossover probability	0.8
Mutation probability	0.05
Number of generations	10
Number of populations	10, 20 and 30

Figure 8

Number of Solutions Found Based on the Number of Chromosomes

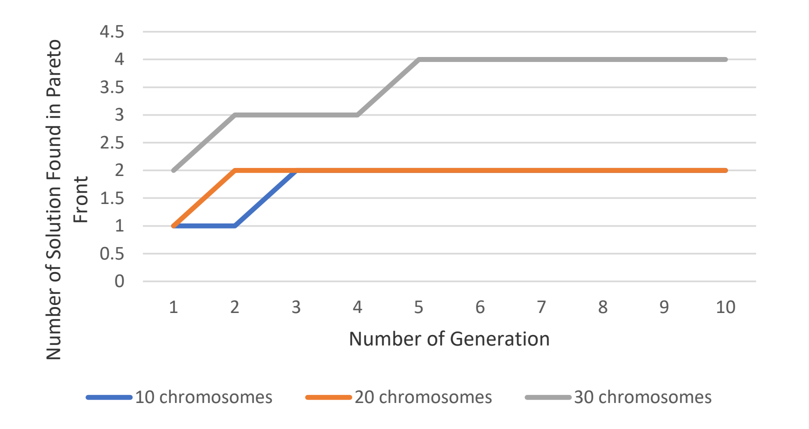


Figure 8 illustrates that the number of solutions found had a positive linear correlation with the number of chromosomes. The lowest number of solutions was observed when using ten chromosomes and twenty chromosomes per generation, whereas using 30 chromosomes resulted in finding the maximum number of solutions. Furthermore, the increasing trend results from each number of chromosomes setup indicated that they managed to find new solutions that outperformed previous solutions, becoming new solutions in the Pareto front. Overall, the results revealed that thirty chromosomes were the best number to use. This quantity had a higher tendency to visit all the searching space area, which was 240 solutions from combining thirty audio files and eight different *bps*.

Comparison Analysis between the Selection Criteria

This analysis aimed to determine the best selection criteria based on the solutions selected. The nearest implementation of the cover selection method, which covered several characteristics during the selection process, was the method by Rashid (2020). However, since this method focused on the image cover selection, a direct comparison was impossible as the metric used was unsuitable for audio. Therefore, for the comparison purpose, three selection criteria were considered in this analysis: 1) SNR and dif_{sample}, 2) SNR only, and 3) dif_{sample} only. Two selection criteria, SNR only and dif_{sample} only, focused on one characteristic method by Rashid (2020), which assessed based on different criteria separately. Meanwhile, SNR and dif_{sample} represented the proposed method that evaluated each criterion simultaneously. The setup used for this tuning is shown in Table 3. The results revealed the top ten solutions for 12.5KB embedding and 5KB embedding. The top ten solutions from the sorted result of 12.5KB embedding are shown in Table 4, while the top ten sorted result from 5KB embedding is displayed in Table 5. For a better view, Figures 9 and 10 illustrate the top four and top three solutions found based on each criterion from 12.5KB and 5KB, respectively. The results were sorted based on the SNR and sample difference, dif_{sample}, values simultaneously. Tables 4 and 5 also presented the corresponding ranking for the solution where SNR only and dif_{sample} only were used as its selection criteria. For sorting based on SNR and dif_{sample} only, higher SNR and dif_{sample} indicated better performance, thus a higher ranking was given to that specific solution. On the other hand, when sorting solutions based on SNR and dif_{sample} simultaneously, the solution was awarded a better ranking if either one or both its SNR and dif_{sample} values were greater than the others, and only if neither SNR nor dif_{sample} was lower than the others.

Table 3

Parameters Setup for Comparison Analysis Used

Parameters	Value
Number of audios in the database	30
Bps	1-8
Crossover probability	0.8
Mutation probability	0.05

(continued)

Parameters	Value
Number of generations	10
Number of populations	10
Size of secret message	5KB and 12.5KB
Selection criteria	1. SNR and dif_{sample} 2. only SNR only dif_{sample}

Table 4

The Sorted Result from 12.5KB Embedding Based on the Ranking of SNR and and its Corresponding Ranking in SNR only and only

Solution [*indexaudio,bps*]	SNR	dif_{sample}	Ranking of the solution according to the selection criteria		
			SNR and dif_{sample}	SNR only	dif_{sample} only
[29,2]	84.6942	2.89E-05	1	1	67
[4,2]	68.9644	0.000592417	1	23	1
[3,3]	78.6229	9.10E-05	1	3	11
[7,2]	76.2750	0.000181851	1	4	4
[29,3]	79.64139	1.95E-05	2	2	145
[5,2]	66.0458	0.000310174	2	37	2
[8,2]	72.0167	0.000174886	2	12	5
[22,2]	75.2100	4.22E-05	2	7	31
[6,2]	66.9485	0.000199322	2	31	3
[11,2]	71.9595	0.000113417	3	13	8

Based on Table 4, it can be observed that solution [4,2] was ranked first using both SNR and dif_{sample} selection criteria, ranked 23rd based on SNR-only selection criterion, and ranked first in dif_{sample} selection criterion. While solution [29,2] was ranked first based on SNR and dif_{sample} selection criteria, ranked first based on SNR-only selection criterion and ranked 67th in dif_{sample} selection criterion. Next, it was identified that solution [7,2] was ranked first based on SNR and dif_{sample} selection criteria and ranked fourth in both SNR only and dif_{sample} only. Finally, solution [3,3] was ranked first based on SNR and dif_{sample} selection criteria, ranked third based on SNR-only selection criterion, and ranked eleventh using dif_{sample} selection criteria. All these four solutions were on equal ranking based on SNR and dif_{sample} as either one of their SNR or dif_{sample} value was greater than the others and neither SNR nor dif_{sample} was lower than the others.

Figure 9

Top Four Solutions Performance From 12.5KB Embedding Based on SNR and , SNR only, and only

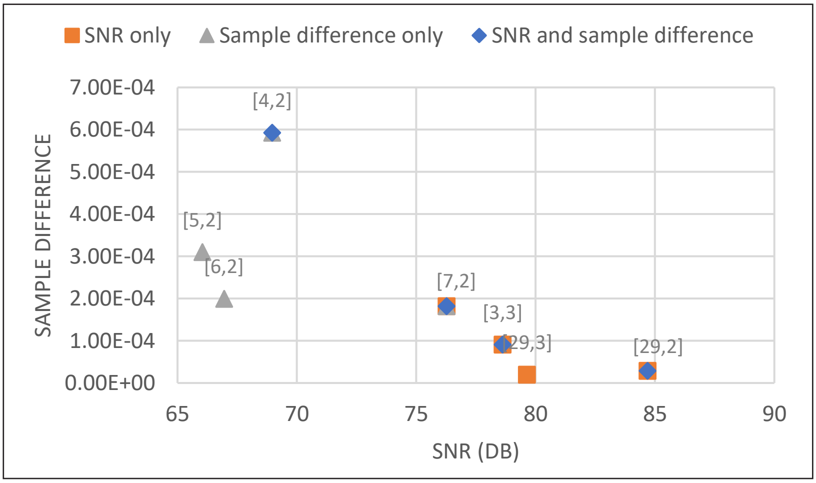


Figure 9 displays the top four solutions for each selection criterion. Several solutions were ranked in the top four in two or more selection criteria. Solution [29,2] was ranked in the top four when using SNR and dif_{sample} and SNR only simultaneously as its selection criteria, while solution [4,2] was ranked in the top four when using SNR and dif_{sample} simultaneously, and dif_{sample} only as its selection criteria. Next, solution [3,3] was ranked in the top four when using SNR and dif_{sample} simultaneously and SNR only as its selection criteria, while solution [7,2] was ranked top four for all selection criteria. The only solutions in the top four that were not ranked in the top ten for more than two selection criteria were [5,2], [6,2], and [29,3]. Solutions [5,2] and [6,2] were ranked based on dif_{sample} only as their selection criteria and solution [29,3] was based on SNR only. In this case, solutions [5,2] and [6,2] were dominated by solution [4,2], which was ranked higher as solution [4,2] performed better in terms of SNR and sample difference compared to solutions [5,2] and [6,2]. Solution [29,3] was dominated by solution [29,2] as the latter was ranked higher than the former because it scored better in terms of SNR and sample difference.

Table 5

The Sorted Result from 5KB Embedding Based on the Ranking of SNR and and its Corresponding Ranking in SNR only and only.

Solution [*indexaudio,bps*]	SNR	dif_{sample}	Ranking of the solution according to the selection criteria: SNR and dif_{sample}	SNR only	dif_{sample} only
[1,1]	70.6716	0.000246792	1	56	1
[3,1]	90.7609	0.000232992	1	2	3
[29,1]	92.6795	2.24E-05	1	1	71
[2,1]	68.7170	0.000246792	2	65	1
[4,1]	76.5718	8.57E-05	2	25	4
[7,1]	83.9147	6.46E-05	2	5	7
[3,2]	87.1155	4.11E-05	2	4	17
[29,2]	88.6598	1.55E-05	2	3	182
[6,1]	74.8631	6.67E-05	3	34	6
[29,3]	83.4760	1.40E-05	3	7	222

Table 5 shows the sorted result of the solutions for 5KB embedding based on the ranking of SNR and dif_{sample} criteria, and the corresponding ranking for SNR only and dif_{sample} only criteria. It was observed that solution [1,1] was ranked first when both SNR and dif_{sample} selection criteria were used, ranked 56th when just SNR selection criterion were used, and ranked first when dif_{sample} selection criterion was applied. On the other hand, solution [3,1] was ranked first using both SNR and dif_{sample} selection criteria, second using just SNR selection criterion, and third using dif_{sample} selection criterion. Next, solution [29,1] was ranked first using both SNR and dif_{sample} selection criteria, ranked first using just SNR selection criterion, and ranked third using dif_{sample} selection criterion. These three solutions were on equal ranking based on SNR and dif_{sample} since each of them had a higher SNR or dif_{sample} value than the others, and neither of them had a lower SNR or dif_{sample}.

Figure 10

Top Three Solutions Performance From 5KB Embedding Based on SNR and , SNR only, and only

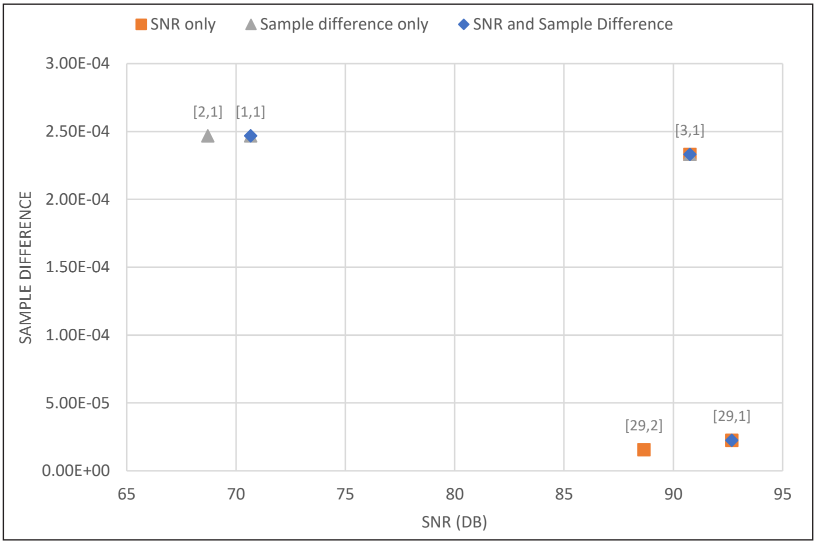


Figure 10 demonstrates the top three solutions performance for 12.5KB embedding based on each selection criteria. Some solutions were ranked in the top three in two or more different selection criteria. Solution [29,1] achieved the top three in SNR and dif_{sample} simultaneously and SNR only, while solution [1,1] achieved the top three in SNR and dif_{sample} simultaneously and dif_{sample} only. Next, solution [3,1] achieved the top three in all criteria. The only solutions that were not selected in more than two criteria were [2,1] and [29,2]. The two solutions' selection criteria were based on dif_{sample} only and one solution was based on SNR only. Solution [2,1] was dominated by solution [1,1], in which the latter obtained a higher SNR value than the former, although both solutions had equal values in the

dif_{sample} criterion. On the other hand, solution [29,2] was dominated by solution [29,1]. and solution [29,1] was ranked higher than solution [29,2] in both criteria.

In conclusion, using SNR and dif_{sample} simultaneously as the is highly recommended as it increased the performance of the cover selection compared with using these two criteria individually. Using these selection criteria separately resulted in inefficient cover selection and further lowered the overall stego file's performance.

CONCLUSION

This research proposed a novel cover audio selection to optimise the trade-off between capacity and imperceptibility during the selection, thus obtaining high capacity and imperceptibility of the audio steganography. This research demonstrated that high capacity and imperceptibility could be achieved by optimising the trade-off between capacity and imperceptibility using the NSGA-II algorithm and adaptive LSB embedding based on the bit level. The experimental result showed that the proposed method managed to optimise the cover and *bps* selection. It also succeeded in dominating the previous selection method, which considered only one selection criterion. Moreover, the proposed general framework can also be used as the foundation for cover audio selection purposes, as the MOEA and audio steganography components can be changed depending on the objective of interest.

This research only focused on optimising the trade-off between capacity and imperceptibility. Therefore, this research will enhance the audio steganography method to manage all the trade-offs between imperceptibility, capacity, and robustness features of audio steganography and improve the embedding method. As there is no similar existing method in audio steganography that concentrates on cover selection, this research also conducted a comparison study against another audio-based cover selection work by adapting existing research from the image or video steganography field.

ACKNOWLEDGMENT

This work was partly supported by the Ministry of Higher Education (MoHE) Malaysia under the FRGS/1/2020/ICT02/USIM/02/1 research grant. The authors would like to express their gratitude to Universiti Sains Islam Malaysia (USIM) and MoHE for the support and facilities provided.

REFERENCES

Abdul Sattar, I., & Talib Gaata, M. (2017, March). Image steganography technique based on adaptive random key generator with suitable cover selection. In *Annual Conference on New Trends in Information & Communications Technology Applications (NTICT'2017)* (pp. 208–212). IEEE.

Ahani, S., Ghaemmaghami, S., & Wang, Z. J. (2015). A sparse representation based wavelet domain speech steganography method. *IEEE/ACM Transactions on Audio Speech and Language Processing*, *23*(1), 80–91. https://doi.org/10.1109/TASLP.2014.2372313

Ali, A. H., Mokhtar, M. R., & George, L. E. (2017). Enhancing the hiding capacity of audio steganography based on block mapping. *Journal of Theoretical and Applied Information Technology*, *95*(7), 1441–1448.

Alsabhany, Ahmed A., Ridzuan, F., & Azni, A. H. (2020). The progressive multilevel embedding method for audio steganography. *Journal of Physics: Conference Series*, *1551*(1). https://doi.org/10.1088/1742-6596/1551/1/012011

Alsabhany, Ahmed Abduljabbar, Ridzuan, F., & Azni, A. H. (2019). The adaptive multi-level phase coding method in audio steganography. *IEEE Access*, *7*, 129291–129306. https://doi.org/10.1109/ACCESS.2019.2940640

Amirtharajan, R., & Rayappan, J. B. B. (2013). Steganography - Time to time: A review. *Research Journal of Information Technology*, *5*(2), 53–66.

Ballesteros, D. M. L., & Moreno, J. M. A. (2012). Highly transparent steganography model of speech signals using efficient wavelet masking. *Expert Systems with Applications*, *39*(10), 9141–9149. https://doi.org/10.1016/j.eswa.2012.02.066

Bazyar, M., & Sudirman, R. (2015). A new method to increase the capacity of audio steganography based on the LSB algorithm. *Jurnal Teknologi*, *6*, 49–53.

Bender, W., & Gruhl, D. (1996). Techniques for data hiding. *IBM Systems Journal*, *35.3.4*(1996), 313–336. https://doi.org/10.1147/sj.353.0313

Chiandussi, G., Codegone, M., Ferrero, S., & Varesio, F. E. (2012). Comparison of multi-objective optimization methodologies for engineering applications. *Computers and Mathematics with Applications*, *63*(5), 912–942. https://doi.org/10.1016/j.camwa.2011.11.057

Cvejic, N., & Seppänen, T. (2002, October). A wavelet domain LSB insertion algorithm for high capacity audio steganography. In *Proceedings of 2002 IEEE 10th Digital Signal Processing Workshop, 2002 and the 2nd Signal Processing Education Workshop* (pp. 53–55). IEEE. https://doi.org/10.1109/DSPWS.2002.1231075

Deb, K, Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, *6*(2), 182–197.

Deb, Kalyanmoy. (2000). Multi-objective optimization using evolutionary algorithms: An introduction. *Multi-Objective Evolutionary Optimisation for Product Design and Manufacturing*, *6*(3), 1–24. https://doi.org/2011003

Djebbar, F., Ayad, B., Meraim, K. A., & Hamam, H. (2012). Comparative study of digital audio steganography techniques. *EURASIP Journal on Audio, Speech, and Music Processing*, *2012*(25), 1–16. https://doi.org/10.1186/1687-4722-2012-25

Durafe, A., & Patidar, V. (2022). Development and analysis of IWT-SVD and DWT-SVD steganography using fractal cover. *Journal of King Saud University - Computer and Information Sciences*, *34*(7), 4483–4498. https://doi.org/10.1016/j.jksuci.2020.10.008

Dutta, H., Das, R. K., Nandi, S., & Prasanna, S. R. M. (2019). An overview of digital audio steganography. *IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India)*, *37*(6), 1–19. https://doi.org/10.1080/02564602.2019.1699454

Gopalan, K., & Shi, Q. (2010, August). Audio steganography using bit modification – A Tradeoff on perceptibility and data robustness for large payload audio embedding. In *2010 Proceedings of 19th International Conference on Computer Communications and Networks* (pp. 1–6). IEEE.

Hajduk, V., & Levický, D. (2018, April). Cover selection steganography with intra-image scanning. In *2018 28th International Conference Radioelektronika (RADIOELEKTRONIKA)* (pp. 1–4). IEEE.

Hameed, M. A., Hassaballah, M., Aly, S., & Awad, A. I. (2019). An adaptive image steganography method based on histogram of oriented gradient and PVD-LSB techniques. *IEEE Access*, *7*, 185189–185204. https://doi.org/10.1109/ACCESS.2019.2960254

Indrayani, R. (2020, November). Modified LSB on audio steganography using WAV format. In *2020 3rd International Conference on Information and Communications Technology (ICOIACT 2020)* (pp. 466–470). IEEE. https://doi.org/10.1109/ICOIACT50329.2020.9332132

Jayapandiyan, J. R., Kavitha, C., & Sakthivel, K. (2020). Enhanced least significant bit replacement algorithm in spatial domain of steganography using character sequence optimization. *IEEE Access*, *8*, 136537–136545. https://doi.org/10.1109/ACCESS.2020.3009234

Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms*. Wiley.

Kaur, A., Dutta, M. K., Soni, K. M., & Taneja, N. (2017). Localized & self adaptive audio watermarking algorithm in the wavelet domain. *Journal of Information Security and Applications*, *33*, 1–15. https://doi.org/10.1016/j.jisa.2016.12.003

Khairullah, M. (2019). A novel steganography method using transliteration of Bengali text. *Journal of King Saud University - Computer and Information Sciences*, *31*(3), 348–366. https://doi.org/10.1016/j.jksuci.2018.01.008

Kunkle, D. (2005). *A summary and comparison of MOEA algorithms*. Northeastern University, Boston, Massachusetts.

Li, B., Li, J., Tang, K. E., & Yao, X. I. N. (2015). Many-objective evolutionary algorithms: A survey. *ACM Computing Surveys (CSUR)*, *48*(1), 1–37.

Mashwani, W. K., Salhi, A., Jan, M. A., Sulaiman, M., Adeeb Khanum, R., & Algarni, A. (2016). Evolutionary algorithms based on decomposition and indicator functions: State-of-the-art survey. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *7*(2), 583–593.

Nursalman, M., Rachman, J. R., & Sidik, F. (2018, October). Implementation of low bit coding algorithm and cipher block with electronic code book mode for data legality in audio steganographic streaming. In *2018 International Conference on Information Technology Systems and Innovation (ICITSI 2018)* (pp. 330–335). IEEE. https://doi.org/10.1109/ICITSI.2018.8695921

Rashid, R. D. (2020). Cover image selection for embedding based on different criteria. In *Mobile Multimedia/Image Processing, Security, and Applications 2020*, vol. 11399, pp. 179–188. SPIE, 2020. https://doi.org/10.1117/12.2560720

Rustad, S., Setiadi, D. R. I. M., Syukur, A., & Andono, P. N. (2022). Inverted LSB image steganography using adaptive pattern to improve imperceptibility. *Journal of King Saud University - Computer and Information Sciences*, *34*(6), 3559–3568. https://doi.org/10.1016/j.jksuci.2020.12.017

Sahu, A. K., & Swain, G. (2022). High fidelity based reversible data hiding using modified LSB matching and pixel difference. *Journal of King Saud University - Computer and Information Sciences*, *34*(4), 1395–1409. https://doi.org/10.1016/j.jksuci.2019.07.004

Sajedi, H., & Jamzad, M. (2008, July). Cover selection steganography method based on similarity of image blocks. In *Proceedings - 8th IEEE International Conference on Computer and Information Technology Workshops (CIT Workshops 2008*) (pp. 379–384). IEEE. https://doi.org/10.1109/CIT.2008.Workshops.34

Setiadi, D. R. I. M. (2022). Improved payload capacity in LSB image steganography uses dilated hybrid edge detection. *Journal of King Saud University - Computer and Information Sciences*, *34*(2), 104–114. https://doi.org/10.1016/j.jksuci.2019.12.007

Shah, P. D., & Bichkar, R. S. (2020, June). Genetic algorithm based approach to select suitable cover image for image steganography. In *2020 International Conference for Emerging Technology (INCET 2020)* (pp. 1–5). https://doi.org/10.1109/INCET49848.2020.9154032

Siinivas, N., & Deb, K. (1994). Multiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary Computation*, *2*(3), 221–248.

Singh, K. U. (2014). A survey on audio steganography approaches. *International Journal of Computer Applications*, *95*(14), 7–14.

Solak, S. (2020). High embedding capacity data hiding technique based on EMSD and LSB substitution algorithms. *IEEE Access*, *8*, 166513–166524. https://doi.org/10.1109/ACCESS.2020.3023197

Somani, H., & Madhu, K. M. (2015). A survey on digital audio steganography techniques used for secure transmission of data. *International Journal of Engineering Development and Research*, *3*(4), 236–239.

Srivastava, M., & Rafiq, M. Q. (2012). A novel approach to secure communication using audio steganography. *Advanced Materials Research*, *408*, 963–969.

Tabares-Soto, R., Raúl, R. P., & Gustavo, I. (2019). Deep learning applied to steganalysis of digital images: A systematic review. *IEEE Access*, *7*, 68970–68990. https://doi.org/10.1109/ACCESS.2019.2918086

Vimal, J. (2014). Literature review on audio steganographic techniques. *International Journal of Engineering Trends and Technology*, *11*(5), 246–248. http://www.ijettjournal.org

Wakiyama, M., Hidaka, Y., & Nozaki, K. (2010, October). An audio steganography by a low-bit coding method with wave files. In *2010 6th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP 2010)* (pp. 530–533). IEEE.

Wang, J., Jia, X., Kang, X., & Shi, Y. Q. (2019). A cover selection hevc video steganography based on intra prediction mode. *IEEE Access*, *7*, 119393–119402. https://doi.org/10.1109/ACCESS.2019.2936614

Wang, Z., & Zhang, X. (2019). Secure cover selection for steganography. *IEEE Access*, *7*, 57857–57867. https://doi.org/10.1109/ACCESS.2019.2914226

Wang, Z., Zhang, X., & Qian, Z. (2020). Practical cover selection for steganography. *IEEE Signal Processing Letters*, *27*(c), 71–75. https://doi.org/10.1109/LSP.2019.2956416

Xin, L., & Jiaojiao, Y. (2018, April). Two embedding strategies for payload distribution in multiple images steganography. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (pp. 1982–1986). IEEE.

Yasear, S. A., & Ku-Mahamud, K. R. (2021). Review of multi-objective swarm intelligence optimization algorithms. *Journal of Information and Communication Technology*, *20*(2), 171–211. https://doi.org/10.32890/jict2021.20.2.3

Ye, D., Jiang, S., & Huang, J. (2019). Heard more than heard: An audio steganography method based on GAN. *ArXiv Preprint ArXiv:1907.04986.*, 1–13. http://arxiv.org/abs/1907.04986

Zhou, A., Qu, B., Li, H., Zhao, S., & Nagaratnam, P. (2011). Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm and Evolutionary Computation*, *1*(1), 32–49. https://doi.org/10.1016/j.swevo.2011.03.001

Zumchak, S. M. (2016). *Audio steganography: A comparative study of techniques and tools*. [Doctoral dissertation, Utica College].