CHAPTER 1

INTRODUCTION

1.1 Research Background

Audio steganography is a method that camouflages a secret message inside an audio file (Gera et al., 2021; Hingmire et al., 2023; Shashi Raj et al., 2022). The most common audio file formats utilised for audio steganography are .wav, .mid, .mp3, and .aiff (C. Li et al., 2020; Mahmoud & Elshoush, 2022; Patil et al., 2022). Audio steganography has more potential in concealing information because audio files are larger than images and small changes in amplitude can store a huge amount of information and less sophisticated than video which requires consideration on noise on each image per frame and audio simultaneously (Singh, 2016).

For any audio steganography to be successfully implemented, four main characteristics of audio steganography are important to be considered which are: capacity, imperceptibility, robustness, and security (Elshoush & Mahmoud, 2023). This thesis defines a balanced performance as the highest possible performance among all the characteristics without being extremely biased toward certain characteristics (Bhowal et al., 2017). This bias can be measured by evaluating each characteristic objectively using their own metrics.

Capacity refers to the amount of secret data that can be embedded inside a cover file (Somani & Madhu, 2015; Srivastava & Rafiq, 2012). Imperceptibility means that the furtive messages should not be detected by the human despite differences exist between cover file and stego-file audio (Somani & Madhu, 2015; Srivastava & Rafiq, 2012; Zumchak, 2016). Although security and robustness are quite similar, they are two distinct definitions. Robustness refers to the ability to protect the stego-file from the attack that aims to increase the error of the secret message or stego-file whether intentionally or not (Agbaje et al., 2017; Mohajon et al., 2018). Security is defined as the ability to protect the stego-file from attack that aim to gain information about the steganography method itself, such as size of the secret message or embedding location (Agbaje et al., 2017; Hemalatha & Ramathmika, 2020). Regarding the security characteristic, various types of research have been conducted on determining the security level of steganography by proposing various ways to detect the information related to the steganography used. Some of these methods focus on differentiating between the altered and unaltered parts of audio samples in stego-file to extract the related information (Farid, 2002; Geetha et al., 2006; Ye et al., 2012). Since there are various approaches to determining the security level, dynamic security characteristic is derived from the security characteristic introduced to accurately measure a part of security aspect and avoid overgeneralising the security characteristic and mixing all the aspect need to be covered in one plate. It is referred as the embedding method's ability to distribute the secret message effectively throughout cover audio in any situation to avoid a visual and statistical attack, hence avoiding differentiation between the altered and unaltered part of audio samples in stego-file (Alsabhany, 2019). However, these five characteristics have four trade-offs: 1) capacity-imperceptibility trade-off, 2)

imperceptibility-robustness trade-off, 3) robustness-capacity trade-off and 4) imperceptibility – dynamic security trade-off.

There are various techniques of audio steganography proposed in the past to improve these characteristics, such as least significant bit (LSB) modification, phase coding, echo hiding, parity coding, spread spectrum, and wavelet domain (Mostafa et al., 2019). One of the most prominent steganography techniques now in use is the LSB embedding system. Although it is a traditional approach, researchers have developed novel modifications to enhance the audio steganography properties (Bansal et al., 2020).

Audio steganography techniques typically involve the user selecting a cover to embed a secret message. However, there are alternative approaches that involve selecting cover from the cover database and synthesising a new cover for the embedding process. Specifically, this research focuses on selecting cover from the cover database approach to develop an audio steganography algorithm. This approach can help the end user to indirectly improve these characteristics.

Despite having good potential in improving the steganography characteristics' performances, based on the author's knowledge, there is no cover selection research that has been proposed for audio file. However, other researchers such as J. Wang et al., (2019) and Z. Wang et al. (2020) have proposed cover selection used for other media types such as image and video. In addition, most of these cover selection methods only consider characteristic(s) individually, neglecting other characteristics. Hence, this research is inspired by the existing cover image selection methods that consider several characteristics. Nevertheless, existing cover selection based on image and video cannot be implemented directly into cover audio selection because each cover has unique qualities. As a result, the metrics used to quantify their quality are typically distinct. For example, the PSNR formulation is often used as the selection criteria. However, it has

different formulation used to measure an image, audio or video stego-file because: 1) it measures the differences between the pixels of stego-file and cover image in an image; 2) it measure the differences between the amplitudes of stego-files and cover audio in audio; and 3) it measures the differences between stego-files and cover video resolution of the specific videos and RGB colour components in a video (Chadha et al., 2013; Hemeida et al., 2019; Liu et al., 2019). In addition, there are other distinct criteria in determining the qualities between these three such as image contrast in image, amplitude level in audio and pixel per frame in video differences which cannot be used to determine other types of cover. In general, this research aims to design a cover audio selection method that considers the trade-off between dynamic security, capacity, imperceptibility and robustness.

1.2 Research Motivation

Since the outbreak of COVID-19, there has been a significant shift in the digital communication landscape. According to the Malaysian Communications and Multimedia Commission (2023), the subscription rates for fixed-broadband, mobile-broadband, and mobile cellular in 2022 were 47.6%, 131.0%, and 145.3%, respectively. In the first quarter of 2023, these rates increased to 48.6%, 132.0%, and 147.6%. Due to the pandemic, the public has become more reliant on the internet, which has become a necessity (Suruhanjaya Komunikasi dan Multimedia Malaysia et al., 2023). As a result, the Government of Malaysia has recognized communication as a public utility. Hence, audio steganography can be one from many implementations to secure the communication between the users. Audio steganography has been continuously improved over the years with the main objective of facilitating secure communication. As a result, many users used audio steganography to secure their communication.

However, many people are unaware of the concept behind steganography, leading them to choose inappropriate audio covers and ultimately making the process less secure. This research aims to help the non-expert to secure their communication by using audio steganography efficiently.

1.3 Research Problem

Characteristics of audio steganography are dependent not only on the technique but also on the secret message and cover audio. Audio steganography frequently generates subpar stego-files due to the user's insertion of improper cover audio. An underloading situation may occur when user uses improper cover audio with its secret messages. It is defined as a situation in which the audio steganography cannot disperse the secret message throughout the cover employed due to the hidden message's limited size (Alsabhany et al., 2019) which lead to the low dynamic security. In addition of that, current cover selection methods such as those proposed by Subhedar and Mankar (2013), Z. Wang et al. (2020), Shah and Bichkar (2020), Molato and Gerardo (2018) and Rashid (2020) are lack of consideration on the characteristics' trade-offs, produced unbalanced stego-file that are highly skewed toward one or two characteristics and significantly reduced in other characteristics. Therefore, based on the defined problems, this research will improve the audio steganography technique problem related with the improper selection of cover audio by introducing cover selection - based audio steganography. This research will design an algorithm to optimize the stego-file performance by finding the set of balanced solutions for cover audio selection that satisfy the trade-offs between characteristics. Hence, without having an in-depth knowledge of the technicality of audio steganography, this algorithm will indirectly help users to implement audio steganography to achieve balanced performance.

1.4 Research Questions

This research attempts to answer the following questions:

1. How to design a solution for audio steganography that could:
    a. improve the performance of the dynamic security characteristic.
    b. select audio based on the trade-off between all characteristics for audio steganography.
2. How to implement the algorithm of the solution for cover audio selection that considers the trade-off between characteristics on the cover audio selection and improve dynamic security?
3. How to evaluate and compare the performance of the proposed algorithm?

1.5 Research Objectives

This research aims to achieve the following objectives:

1. To design the solution for audio steganography by:
    a. formulating enhanced solution that improves the performance of dynamic security characteristic of audio steganography.
    b. formulating selection criteria for cover audio selection used in audio steganography that can select audio based on the trade-off between all characteristics.
2. To implement the solution in an algorithm that satisfies the trade-off between all characteristics on the cover audio selection and improve the dynamic security characteristic performance.
3. To evaluate the performance of the proposed algorithm using maximum size secret message can be embedded, Mean Square Error (MSE), Signal to Noise Ratio (SNR) and Peak Signal to Noise Ratio (PSNR), bit error rate

(BER), and segmental SNR (seg-SNR) and compare with the other algorithm.

1.6 Research Scopes

The scopes of this research are:

1. This research focuses on selecting audio for audio steganography.
2. In this research, the audio used for cover selection and the evaluation processes are:
    a. based on the unified audio format with a 44.1kHz sample rate, mono audio channel and 16-bit per sample.
    b. from the free online audio library (www.freesound.org).
3. In this research, the process of evaluation performance focuses only on the capacity, imperceptibility, robustness, and dynamic security characteristics.
4. In this research, the robustness characteristic is only measured in terms of Additive White Gaussian Noise (AWGN).
5. In this research, the dynamic security characteristic is only measured in terms of signal analysis.
6. In this research, only objective metrics are used in the evaluation process. For imperceptibility, MSE, SNR and PSNR are used. For capacity evaluation, maximum size of secret message embedded is used. For robustness evaluation, bit error rate (BER) is used for evaluating the message retrieval accuracy under an AWGN attack. For dynamic security evaluation, seg-SNR is used for evaluating the message distribution under signal analysis attack.

1.7 Thesis Organization

This thesis is organised into six (6) chapters. Chapter 1 discusses an important overview of this research including the research problems, questions, objectives, and scopes. Chapter 2 presents the review of the audio steganography and cover selection, the new proposed classification and a critical review of each classification and research gap. Next, the research methodology, evaluation environment, evaluation criteria and design of proposed solution are presented in Chapter 3. Chapter 4 focus on the implementation of proposed solution into cover selection -based audio steganography (CAS) and empirically presented using actual parameter values. The results and discussion of the experiments conducted were discussed in Chapter 5. Lastly, Chapter 6 revisits all the research objectives and discusses the research limitations and contributions.