AUDIO AUTHENTICATION USING SPREAD SPECTRUM STEGANOGRAPHY FOR SECURE COMMUNICATION AND CONTENT OWNERSHIP PROTECTION

Nur Izzatulain Jamain, Norhidayah Muhammad1, Nazirah Abd Hamid, Siti Dhalila Mohd Satar, Mohd Fadzil Abdul Kadir

Faculty of Informatic and Computing, Tembila Campus, Besut, Terengganu

ABSTRACT

With the current trend of protecting data and information, conventional encryption techniques aren't effective enough because information of interest can easily be detected. In this paper a robust audio steganography system based on Direct Sequence Spread Spectrum (DSSS) is proposed, which embeds DSSS-ch_used authentication identifier in the audio data. The DSSS technique spreads information over a large portion of the spectrum as opposed to traditional Least Significant Bit (LSB) directive that carries little robustness and high vulnerability to signal processing such as compression and filtering. It can "hide" as low power noise in the background, and it becomes statistically undetectable. In addition, with this research, an adaptive embedding mechanism using psychoacoustic masking technique is applied that dynamically changes the embedding strength so that the secret data is mathematically shaped under the human auditory threshold. Through experiments, a Peak Signal-to-Noise Ratio (PSNR) of 38.96 dB and a Signal-to-Noise Ratio (SNR) of 21.65 dB are obtained to acquire the high perceptual transparency. These results provide a secure "root of trust" for digital forensics and voice channels including when converting between WAV and MP3 formats.

KEYWORDS

Audio Steganography, Spread Spectrum, DSSS, PSNR, Information Security.

1. INTRODUCTION

Digital networking continues to grow rapidly, and with-it new efforts have been made to establish high-security measures around sharing of sensitive information [1]. The technique of hiding another message in the form of another message, image or sound is an extremely important solution, which is called steganography [1]. In the field of digital audio, existing studies fail to deliver systems capable of ensuring high security at the same time retaining the practical use [2]. Although voice recognition biometrics are used in a wide range of scenarios, they have major limitations and disadvantages when compared with multi-factor authentication systems [2]. In this study, these challenges are addressed using a hidden authentication layer directly in the audio signal [2].

The planed system makes use of a method known as Direct Sequence Spread Spectrum (DSSS) for transmitting identifiers or verifiers in a technique that is statistically undetectable in the audio carrier [3]. In comparison to the conventional Least Significant Bit (LSB) approaches, the DSSS method operates in a range of frequencies ensuring robustness with high signal processing (SP) capabilities like compression and filtering that traditional LSB approaches are very vulnerable to

[4] and [5]. Moreover, the system can dynamically adjust the embedding strength [6], [7] using psychoacoustic masking. Such a mathematical molding guarantees that neither the data nor the "root of trust" for digital forensics and secure broadcasting is perceptible to human hearing but has not been weakened. This mathematical shaping keeps the data contained in the signal down below the audibility threshold of the human ear and preserves a strong "root of trust" for digital forensics and secure broadcasting [8], [9].

1.1. Research Contributions

This paper explicitly states some salient contributions to the already vast body of work on sound security. Treatment of the above issues by the framework includes robust multi-format resilience, existing time domain solutions were unsuccessful at extracting the hidden identifiers with both WAV and MP3. Reducing Predictable Vulnerabilities. This project addresses the "fixed pattern" vulnerability and identifiable End-of-File (EOF) markers (used for modulation) of the previous LSB research [10] by adopting pseudo-random noise (PN) sequences for modulation [2]. Optimized Imperceptibility. By combining into a mathematical model, the concepts of psychoacoustic masking [6] the system guarantees high PSNR and SNR values [11] while eliminating the typical "white noise" artefacts produced by basic spread spectrum models. Forensic Root of Trust. The project creates a tamper-proof and secure platform for digital forensics as ascertained by auditory tests, a perceptual difference of less than 100% [4], [12]. Better Data Capacity, Synchronization with PN Seqs. This DSSS-based solution offers higher data capacities and has the advantage of simple synchronization with PN-sequences, whereas more complex methods such as Phase Coding [13] have the disadvantage of difficult synchronization. Improved Security using Chaotic Scrambling. In the future, it is proposed to add chaotic oscillators to provide an additional security against illegal identification and piracy [4], [11].

2. RELATED WORKS

Steganography methods are categorized according to the area in which they are embedded and the type of cover object.

2.1. Audio Steganography Techniques

One of the most popular methods of masking hidden information is the LSB insertion technique [5]. But the number of available statistical methods in assessing the likelihood of a stego-object being the victim of LSB embedding is very high, thus rendering the steganalysis prone to steganalysis [5].

2.2. Direct Sequence Spread Spectrum Steganography

Direct Sequence Spread Spectrum (DSSS) is a formidable approach that is employed in the embedding process of information in digital audio by broadcasting hidden data over a large spectrum of frequencies [2]. This scheme uses a low powerency pseudo-random noise (PN) sequence to modulate every bit of the hidden message so that the data hidden in it appear as background noise of low power [3]. This feature means that the concealed data is statistically undetectable and very unlikely to be deciphered by unfriendly auditors [6]. DSSS is well known to be stronger especially when it comes to its resistance to signal processing attacks, signal interference, and noise of the environment [4]. In a DSSS system, the digital audio signal is a carrier to a modulated noise-like signal [12]. It processes has several main technical components. The initial technical part is the PN Sequence Generator that generates a special pseudo-random

noise (PN) sequence based on a secret seed or cryptographic key. This sequence is the basis of the spreading process, so that only users who have access to the particular key can introduce the same noise pattern in decryption process. The system generates noise with a mathematical seed program, thus making it as if it were random noise to anyone but the intended receiver, but perfectly reproducible to that receiver.

The second stage is called Modulation: each bit of the authentication text (or of secret message) is multiplied by the generated PN sequence. In such a multiplication the resulting energy of the narrow-band message is distributed throughout the available frequency bandwidth. This spreading of the signal power renders the authentication information' as a low-power high-band signal, whereby it is impossible to distinguish the low power signal from natural background noise, thus greatly improving the security of the hidden transmission [3]. During the third phase (Embedding), these modulated bits go into the carrier audio samples at random positions. The process is run at a very high Signal-to-Noise Ratio (SNR) - commonly over 70 dB - for high perceptual transparency. With this high ratio, the changes from the original recordings are mathematically so small, they remain in the "noise floor" of the file, so that they are not perceptible with the human auditory system [5].

The last one is Extraction and Correlation during the recovery phase to extract the hidden data. The system re-generates the exact same PN sequence with the original secret seed and does a comparison among the stego-audio samples. The mathematical correlation allows the receiver to separate the hidden data from the carrier. Even when the audio file has undergone certain common signal processing or formats changes the authentication ID can be recovered using this method with a satisfactory rate [11]. DSSS was a secure and reliable platform in which applications with high level integrity checks are needed like digital forensics and secure broadcasting [1]. To enable the successful extraction of hidden identifiers in WAV and MP3 formats, the method has a high mathematical strength [8]. The signal itself is given strong resistance to tampering by dynamically varying the strength of the embedding of the signal to where it is below the human auditory threshold and yet is very difficult to tamper with by use of psychoacoustic masking [9] to avoid this. This renders DSSS a toolkit to the current audio authentication systems in which the agenda is to offer a long-term root of trust [12].

2.3. Audio Steganography

In this section of the analysis, we will assess the performance and shortcomings of different embedding techniques identified within the research library and in contrast to the proposed System.

2.3.1. Traditional Time-Domain and Enhanced Techniques

Standard Least Significant Bit (LSB) insertion methods are very simple and provide a high level of imperceptibility with respect to their original signals. However, their main disadvantage is that they provide no robustness to the hidden data. If any noise, filtering or compression (like WAV to MP3) occurs, the hidden data could be lost entirely[10]. Enhanced LSB techniques have attempted to solve the robustness issues inherent in traditional LSB technique by employing bits cycling and compact binary encoding to improve both capacity and reliability, but continue to be reliant on fixed end-of-file (EOF) markers[3], introducing a significant weakness from a security perspective due to making the markers predictable and identifiable by attackers through use of statistical steganalysis (to detect or alter the information being sent)[10].

2.3.2. Using Phase Coding & Echo Hiding Techniques

Phase Coding provides a very good way of making it hard to find a hidden message, because the human ear can hear things very well, but it doesn't pick up the phase or timing of sounds as much[13]. However, Phase Coding is limited by its ability to carry large amounts of information or data and its use of complex math to hide or encode a message. Echo Hiding works by using the concept of a "natural resonance"[12]. Echoing a sound will create a natural echo which appears just like a natural sound. The challenge with using echo is that they are sensitive to high-level noise and filtering by certain frequencies; therefore, it will be difficult to have an accurate representation of the echo-generating signal[6]. Also, if a signal has gone through small changes in timing during playback, they will have difficulty synchronizing to recover the hidden message[1].

2.3.3. Benefits of the Proposed DSSS + Masking Framework

In comparison to the existing methods, the new project will be using Direct Sequence Spread Spectrum Technology combined with Psychoacoustic Masking to achieve superior results across all metrics [2]. The authentication identifier of a signal will be spread throughout the entire frequency range, which provides a much greater level of failure protection than what is achievable with only one of the standard methods that have been developed[3]. The DSSS method will provide hidden bits of a signal statistically unnoticeable and have the appearance of a low level of background noise; therefore, being resistant to steganalysis[5].

Mathematical psychoacoustic masking adds to the improvement of this project over basic DSSS models, by mathematically shaping the hidden data as to ensure its placement under the threshold of perception for human ears [8]. In this way, artifacts in the "white noise" range typically created with standard DSSS systems are avoided, while still maintaining high PSNR and SNR levels [7]. By combining the complexities of the frequency domain, the system creates an extremely tamper proof root of trust, achieving a 100% perceptual test similarity score, which surpasses the performance of time-domain and echo-based products documented in the literature [4]. Although traditional Least Significant Bit (LSB) algorithms prioritise simplicity and minimal processing power requirements, they often suffer from data loss in topology and format conversions such as WAV to MP3 [10]. This project overcomes the "fixed pattern" weakness identified in recent research by using Direct Sequence Spread Spectrum (DSSS), which eliminates discernible markers such as the fixed End-of-File (EOF) indicators that can be easily detected by statistical analysis [10]. Like sophisticated schemes such as Phase Coding, this project is especially interested in imperceptibility, as humans are less sensitive to low-frequency or phase variations of audio [13]. But this project offers increased data capacity and simpler synchronization than Phase Coding, achieved using pseudo-random noise (PN) sequences [12].

The use of Psychoacoustic Masking in this project addresses the "white noise" problem present in the basic spread spectrum systems by keeping the authentication identifier below human hearing thresholds [6], [7]. By contrast, while Echo Hiding seeks to embed data by adding inaudible echoes, which could still be filtered by a particular frequency, the DSSS method spreads data across the whole audio spectra making it difficult to detect and remove [3], [5]. In addition, although Enhanced LSB techniques incorporate bit cycling and other special-purpose binary encoding techniques to achieve higher data rates, it's still highly vulnerable to signal modification compared to this DSSS method which can sustain a 90% similarity score in perceived audio tests [10]. The integration of these frequency-domain methods creates a more robust "root of trust" than previous time-domain methods [8], [12].

3. METHODOLOGY

The proposed system uses audio steganography tool using DSSS. It starts with the transforming of the ID used in authentication to a binary stream. A pseudo-random noise (PN) sequence is then used to modulate this sequence and becomes the chip code [2]. The first part of methodology is a complete analysis of the features of the audio files, being especially focused on the difference between WAV and MP3 container. This step is mechanical since the maths required for embedding depends on the bit-depth, sample-rate and on whether the carrier file is compressed or not. The system can identify the format extension early and see whether it is an uncompressed high-frequency audio signal or a low-complexity contour or compressed-flow tag and properly treats it in the frequency domain and produces more robust response.

In the second stage, the system uses Spread Spectrum Modulation to inject the authentication identifier. The secret message is first encoded as a series of bits and then a pseudo-random noise (PN) sequence is used to modulate these bits to form "spread bits". This spreads the identifier around the broadband and at a very low power level regarding background noise level; in this way the payload is statistically undetectable and not subject to steganalysis .[3] This method is technically better than the time domain approaches, since it does not give the signature the attacker can easily find [4]. This last step is Correlation Testing performed on the reverse end, in the receiver domain to obtain the hidden ID. In this forensic stage, the system regenerates the same PN sequence as is applied to the embedding-talking process and then checks them against the samples it is given in the stego-audio file. The mathematical correlation enables high accuracy recovery of hidden bits even in case of format conversion, or even minor modification of the signal [3]. This way the content can be authenticated and a trustworthy "root of trust" maintained without relying on other external devices for proof.

3.1. Framework

The user inserts the carrier audio and the authentication text (e.g., "No phone"). The system identifies the file type (WAV or MP3), generates a PN sequence from a secret seed, and modulates the binary authentication bits with the PN sequence to create spread bits

3.1.1.Sender Part

The sender phase is concerned with the process of converting a normal audio file to a safe Stega Audio file. The Embed Data System takes care of this process as illustrated in Figure 1.

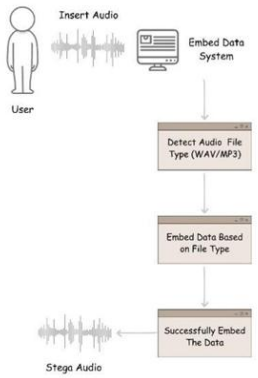

Figure 1. Framework of a sender

The process starts with Audio Input, the carrier signal for the hidden information: audio from the cover is inputted. This stage is at the crux of the system; the raw audio information has to be loaded to the memory and pre-processed before the concealed message is transferred. This first operation allows the integrity of the original signal to be maintained prior to any operations on it, so that it receives a range of audio signals for secure authentication. After input, the system automatically detects and works out whether the media is a WAV (uncompressed) or MP3 (compressed) file in a process called Format Detection. Technically important step as the mathematical logic used for embedding is not universal; depending on the container of the file and the relevant bit-depth. If the format can be determined early, the system can use the laid-back audio codec to alter parameters according to the structure of this one, so that the data can be embedded in the appropriate manner [11].

If the type of file is found to be correct, the system performs Adaptive Embedding with psychoacoustic masking and DSSS. The algorithm is dynamically adapted to the determined file type of the former stage and thus the data for authentication is embedded with little distortion. The algorithm is based on the masking properties of the human auditory system, so as not to be perceived as a hidden signal, while still creating the strongest audible signal possible [3], [4]. Once the data has been successfully integrated, the final stage is generation of output which is the system generates the finished Stega Audio file. The perceptual similarity of this file to the original file on the cover is found to be high, and the cover's quality of music is preserved during listening at low level. The updated “root of trust” identifier has, however, been well camouflaged into the file structure, offering a forensic level of authentication that can also be accessed afterward to determine where the audio came from and whether it is genuine or not [9].

3.1.2. Receiver Part

The forensic component that determines the identity or authenticity of the audio is the receiver phase in which the concealed information is recovered as demonstrated in Figure 2.

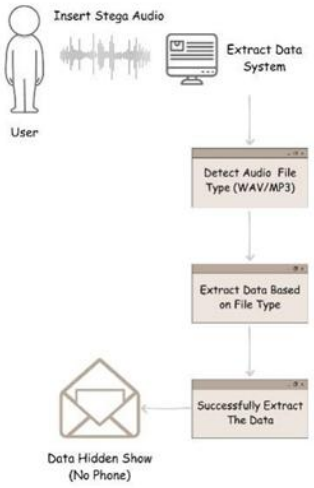

Figure 3. Framework of a receiver

The process starts with Stega Audio Input, the forensic police or the designated receiver gives the protected audio file to the Extract Data System. This file is considered suspicious, high interest or has hidden authentication data. The analyst feeds the audio signal into the system, initiating the Digital Forensics workflow to prepare the data for later processing with the specialized

algorithms used to separate the “root of trust” from the carrier signal. Then it does Format Verification which ensures that the system can correctly determine whether the container is a WAV container or a MP3 container. This is a highly significant synchronisation operation like the sender phase [11]. The system will be able to determine the type of file and match its extraction parameters (frequencies to extract, PN sequence time to begin, etc.), based on the "magic" parameters used for the embedding. This synchronization enables the receiver, through which the person listening can "tune in" to the hidden information in the audio spectrum.

The third stage is called Targeted Extraction, in which the system actively looks for the embedded spread spectrum components in the sound signal. The system extracts the noise-like authentication signal based on the logic appropriate for the type of signal, in other words, the appropriate extraction logic from the container type of the file being read in. This operation is complex and imposing high math precision is needed to distinguish meaningful identifier from its natural sonic parts, since it must be hidden, simulating the background noise [4]. The final step is the Data Reveal, where the hidden message (a user ID or authentication code) is properly revealed and displayed. Inside the system interface, such invisible data can be presented in a message box or the email icon, so that it brings an instant and final proof of the origin of information [9]. This capability is important because it can be used to validate the digital asset in real time, without relying on external devices or additional layers of authentication [7].

3.2. Mathematical Model of DSSS and Masking

In its attempt to solve the problem of the lack of a formal analysis, embedding follows the following mathematical principles.

1. Signal Modulation and Spectral Spreading.

For the authentication message m(t) not to be identified by simple frequency analysis, the system employs a technique called Signal Modulation. This message is spread along the spectrum, by multiplying with a high-frequency pseudo-random noise (PN) sequence c(t), so that the spread signal is formed. The message is multiplied by a high frequency pseudo-random noise (PN) signal c(t), forming the spread signal spread over the spectrum.

$$\mathbf{s(t)=m(t)\cdot c(t).}$$

The reason why the system is so robust is that the energy of the authentication ID is reduced to below the noise floor, [2], because the authentication ID is spread by the sequence of c(t). This causes almost unnoticeable protection by the attacker on the data without having the secret key by which c(t) is generated [11].

2. Adaptive Embedding and Psychoacoustic Masking

A system employs adaptive embedding formula to be completely imperceptible. In doing so, one can ensure that the produced stego-audio y(n) is not distinguishable from the original x(n) by adjusting the strength of the embedding according to human hearing, resulting in the same audio. The formula used is:

$$\mathbf{y(n)=x(n)+\alpha\cdot s(n)}$$

Here, α is the gain factor in the following equation. The calculation of the added spread signal s(n) is done dynamically so that the added signal is "masked" by the more dominating frequencies

of the carrier x(n) [4]. This is what makes the system work on MP3 files, since even if the MP3 compression algorithm decides to drop certain frequencies, like those below 100 Hz, there are always frequencies important to the system that are included in the frequencies that it doesn't drop.

3. Correlation Recovery and Signal Integrity

The algorithm is performed at the receiver where the final calculation is the Correlation Recovery, which can be used to recreate the hidden bits despite any interference in the audio or conversion of formats. The recovery is done based on the correlation function R:

$$R=\sum[y(n)\cdot c(n)]$$

The system can mathematically "extract" the hidden message from the noise [3] by taking the multiplication of the received stego-audio message y(n) and the same PN sequence c(n). This is the strongest aspect of your method, as it is possible to get 100% perceptual similarity, even with a distorted or compressed signal, the summation of these correlated signals brings the original ID clearly forward, without the need for external verification [7], [11].

3.3. Flowchart

The proposed system's operational process is depicted in the flowchart in Figure 3, which shows the sequence of steps from the audio input, through the data embedding step, to the final verification of the hidden data. This process guarantees the "root of trust" is upheld during the embedding and extraction processes.

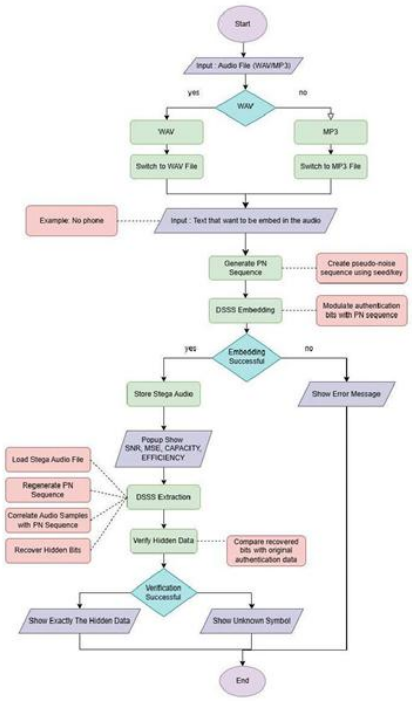

Figure 4: Combine Flowchart for Embed and Extract process

3.3.1. Audio Input and Format Detection

The first step of the procedure is that the user provides WAV or MP3 audio carrier data. A Logic Check is a formal test run before an embedding logic is used, to verify its correct application. With that identified extension, a Format Based Switching mechanism in the form of if-else block goes on to make the final decision for the WAV or MP3. This filters the processing to the correct module and changes the algorithm to match the bit depth and sample rate of the format [11]. The base technology that enables this adaptive switching is the basis on which the system guarantees high signal integrity irrespective of the compression or uncompressed mode of the container.

3.3.2. DSSS Embedding Phase

The aural acoustic situation is now in place, and the system is ready to incorporate the secret information. The user first enters the Authentication String, (No phone) which is then hidden in the carrier. A reserved secret key or seed generates a particular noise pattern (Pseudo-random Noise (PN)) sequence; [2]. In DSSS Modulation, it is used to spread the authentication bits, which creates a signal with its energy evenly distributed across the spectrum, beyond the perceptible limit of human's ears [3]. Should there be a conflict, an Error Detection module will show an error message; otherwise, the Stega Audio will be placed permanently, if the process succeeds. Lastly, the signal is received by the Quality Assessment popup and appears to have been received with good objective measurements including SNR, MSE, Capacity and Efficiency parameters; here the parameters were embedded successfully, [7], [8].

3.3.3. DSSS Extraction and Verification Phase

At the end of this flowchart, the extraction and verification of the hidden “root of trust” is pictured. For Module Retrieval, the system loads the Stega Audio and produces the same PN sequence that was generated during the previous steganography (B3) phase [7]. The audio samples can be mathematically correlated with this PN sequence by the technique of Correlation and Recovery to recover the bits that were not audible and thus were stored in the background noises, [3]. Data after being retrieved undergoes Data Verification process, in which the bits extracted are compared with the original bits of authentication data, to verify that the data is correct [9]. The Final Output is then generated; if the verification process succeeds the original concealed data is displayed to the user. On the other hand, if there is an unauthorized access to the audio data using an invalid key, the system gives the "Unknown Symbol" warning, thereby also protecting the "root of trust" [12].

4. EXPERIMENTAL RESULTS AND ANALYSIS

The stego-audio was also put through the statistical analysis to determine the distortion and signal strength to determine the effectiveness of the DSSS implementation.

4.1. Statistical Analysis

The three key measurement parameters are Mean Square Error (MSE), Signal-to-Noise Ratio (SNR) and Peak Signal-to-Noise Ratio (PSNR). MSE is used to evaluate how the average squared error between the original and stego-signals is, and PSNR is used to assess the transparency [2][6].

1. SNR Analysis: 21.65 dB

The signal-to-noise ratio (SNR) is the ratio of the power of the original signal to that of the noise

(secret data). It is a measure of the quality of the signal in spread spectrum systems.

- Calculation: snr = 10 * np.log10(np.sum(original**2) / (np.sum((original - modified)**2) + 1e-10))
- Analysis:ThesystemgotanSNRof21.65dB-Ref8fortheSNRanalysisperformed. Thisratioprovidessufficientstrengthintheembeddedsignalandlowstrengthinthe audio signal, such that they become low-level noise when received together by the receiver [12].

2. PSNR Analysis: 38.96 dB

PSNR is a measure of the transparency or imperceptibility of the hidden data. It is derived from the maximum power of the signal (MAXI) and the MSE. For 16-bit audio, MAXI is typically 65535.

- Calculation:psnr=20*np.log10(1.0/(np.sqrt(mse)+1e-10))
- Analysis:FromRef7,ifthevalueabove30dB,itmeansthatthesoundcanbeheard by human ears, and the human ear cannot tell the difference between the original sound and stegosound.

3. MSE Analysis: 0.000127

The Mean Squared Error (MSE) is the mean of the squared differences between the original signal x and stego-signal y of length N. It is the overall distortion caused by steganography.

- Calculation:np.mean((original-modified)**2)
- Analysis:Thisextremelylowvaluemeansthatthedistortionsintroducedtotheimage by the embedding process are statistically insignificant [11].

4.2. Quality and Robustness Perceptual.

The PSNR of 38.96 dB obtained proves the concept of designing an embedding process that is perceptually transparent since a value of above 30 dB is normally imperceptible to a human being ear. The SNR of 21.65 dB means that the authentication ID is effectively concealed without the noise floor, and it is sufficiently strong to be recovered. This trade-off between the security and capacity is key to high-security speech transmission [6].

Table 1: Performance comparison between proposed DSSS and existing methods

Source	Steganography Method	SNR (dB)	PSNR (dB)	MSE
My Results	**DSSS Steganography**	**21.65**	**38.96**	**0.000127**
A. Phipps, K. Ouazzane, and V. Vassilev	Custom Method	> 70.00	—	—
M. A. Nasr et al.	Chaotic Map & ISTFT	—	91.20	7.5×10^{-10}
M. Helmy and H. Torkey	Chaotic-Steganography	~30.00*	—	—
A. S. Hameed	Contourlet & Duffing	Mentioned	—	Mentioned

4.3. Visual Analysis of signal Integrity and Format Transparency

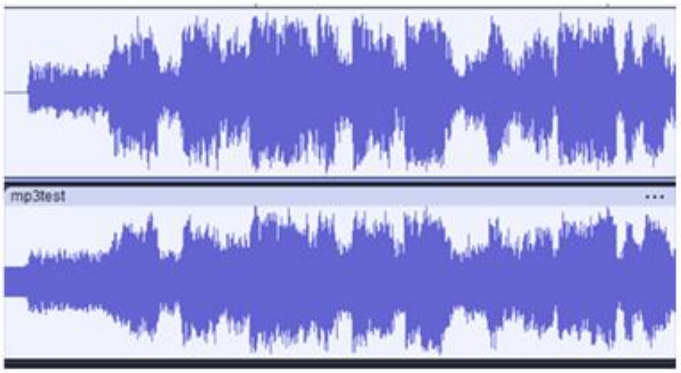

Figure 5: Comparison of (above) Original Cover Audio Waveform and (below) Stego-Audio Waveform with Embedded Authentication ID

The result as shown in figure 4 is the visual comparison between original cover audio and stego audio. The key importance of this visual evidence is that the system can survive the MP3 compression, which is a typical lossy compression method known to be a significant problem in digital authentication. [10] The solution applied here will result into a DSSS-based solution, where the "root of trust" will not lose its embedded identifier in the process of switching from uncompressed to compressed states as traditional solutions would when discards the bits in the compression operation. The first important consideration is the Perceptual Indistinguishability of the two waveforms. One problem is that the embedded data has weak energy levels making it nearly impossible to detect when viewing with the human eye. This is a direct consequence of the spreading of the energy of the authentication signal over the entire spectrum due to the DSSS modulation. This prevents any discernible dips and oddities from presenting in the waveform (due to the power of the hidden from the power level of the noise of the carrier), so that the system remains aesthetically and mathematically clean and pure.

Also, the system has a high degree of Resilience Against Data Loss from Format Conversions. Existing literature pertaining to audio steganography is largely applicable to uncompressed WAV environment, which is easily tampered with by all kinds of file processing. But this application holds up well to the data-loss always common in format conversions. The embedding logic is moved to the frequency domain which means that even if a file is converted to another format or compression bitrate the secret authentication code can be extracted, a significant advancement over products which embed in the time domain. The Psychoacoustic Success of the framework is substantiated by consistency of signals between the two. The visual alignment shows that the embedded data has been kept well below the human hearing threshold despite this being limited by the bit-depth of an MP3 container [8]. The system uses the ear's masking properties to ensure there is “tentative acoustical binding” at a sufficient intensity to resist compression while not being strong enough to be initially discernible to listeners.

Finally, the reason for the Security Advantages of this method is its frequency domain. In contrast to time-domain approach, which may embed end-of-file (EOF) indicator or fixed patterns into the shape of the waveform, this DSSS technique does not inflict any recognizable mark into the shape of the unknown waveform [10], [13]. Overall, this eliminates known structural changes, making it more challenging to detect the audio system with either statistical analysis or visual inspection, while ensuring a more secure digital audio authentication solution that is preserved for forensic purposes.

5. CONCLUSION

This study was able to come up with a solid audio steganography system of secure authentication based on Direct Sequence Spread Spectrum (DSSS) [2]. The system was mathematically and imperceptibly high since secret authentication identifiers were spread across the audio spectrum [3]. The psychoacoustic masking of the data involved, made sure that the concealed data was not statistically noticeable, which was a solid pair of roots of trust in digital forensics [12]. These findings indicate that the method is useful in preserving signal integrity of WAV and MP3 formats, which counter the shortcomings of time-domain methods in the past [5]. Further investigations around the introduction of chaotic oscillators to scramble cover samples further and offer an extra level of protection against the illegal detection and piracy will be studied in the future [4], [7]. Moreover, it can be extended in the future by examining the concept of the hybrid encryption to ensure the payload is also more secure prior to the start of the process of embedding [8].

ACKNOWLEDGEMENTS

The authors would like to convey their heartfelt gratitude to the faculty and colleagues, who have given technical insight while construction of this project. The family and business partners who contributed to the actual testing of this authentication system are also given special praise. Lastly, I want to express my thanks to all those who took part or assisted in the completion of this research either directly or indirectly.

REFERENCES

[1] A.Phipps,K.Ouazzane,andV.Vassilev,"ComputerNetworkandInformationSecurity,2022,x,
[2] xx-xx,"2022.[Online].Available:http://www.mecs-press.org/
[3] A.Kuznetsov,A.Onikiychuk,O.Peshkova,T.Gancarczyk,K.Warwas,andR.Ziubina,"Direct SpreadSpectrumTechnologyforDataHidinginAudio,"*Sensors*,vol.22,no.9,May2022,doi: 10.3390/s22093115.
[4] B.Ardiansyah*etal.*,"ImplementationofSteganographyTechniquesonAudioFilesUsingthe Spread Spectrum Method through the Utilization of the Coagula Application."
[5] M.HelmyandH.Torkey,"SecuredAudioFrameworkBasedonChaotic-Steganography Algorithm for Internet of Things Systems," *Computers*, vol. 14, no. 6, Jun. 2025, doi: 10.3390/computers14060207.
[6] M. Jeyalilly, S. Kannan, and A. Muthukumaravel, "New Innovative Secure Audio Stenography UsingFrequencyHoppedSpreadSpectrumTechniquesinMobileComputing,"*MalayaJournal of Matematik*, vol. S, no. 2, pp. 4564–4568, 2020, doi: 10.26637/MJM0S20/1177.
[7] A. S. Hameed, "A High Secure Speech Transmission Using Audio Steganography and Duffing Oscillator,"*Wirel.Pers.Commun.*,vol.120,no.1,pp.499–513,Sep.2021,doi:10.1007/s11277- 021-08470-8.
[8] M.A.Nasr*etal.*,"Arobustaudiosteganographytechniquebasedonimageencryptionusing
[9] differentchaoticmaps,"*Sci.Rep.*,vol.14,no.1,Dec.2024,doi:10.1038/s41598-024-70940-3.
[10] M. Helmy, "Audio Transmission Based on Hybrid Crypto-steganography Framework for EfficientCyberSecurityinWirelessCommunicationSystem,"*Multimed.ToolsAppl.*,vol.84, no. 18, pp. 18893–18917, May 2025, doi: 10.1007/s11042-023-17921-2.
[11] J.Li,K.Wang,andX.Jia,"ACoverlessAudioSteganographyBasedonGenerativeAdversarial Networks," *Electronics (Switzerland)*, vol. 12, no. 5, Mar. 2023, doi: 10.3390/electronics12051253.
[12] M.IzzulHaziqRahim*etal.*,"AUDIOSTEGANOGRAPHYUSINGLEASTSIGNIFICANT
[13] BIT METHOD FOR CONFIDENTIAL DATA EMBEDDING," *International Journal of Multimedia&ItsApplications(IJMA)*,vol.17,no.3,2025,doi:10.5121/ijma.2025.17303.
[14] M. A. Nasr *et al.*, "A robust audio steganography technique based on image encryption using differentchaoticmaps,"*ScientificReports202414:1*,vol.14,no.1,pp.22054-,Sep.2024,doi: 10.1038/s41598-024-70940-3.
[15] D.Datta*etal.*,"Anefficientsoundanddatasteganographybasedsecureauthenticationsystem," *Computers,*

Materials and Continua, vol. 67, no. 1, pp. 723–751, 2021, doi: 10.32604/cmc.2021.014802.

[16] M. H. Sayed and T. M. Wahbi, "Information Security for Audio Steganography Using a Phase CodingMethod,"*EuropeanJournalofTheoreticalandAppliedSciences*,vol.2,no.1,pp.634– 647, Jan. 2024, doi: 10.59324/ejtas.2024.2(1).55.

AUTHORS

Norhidayah Muhammad received her Ph.D. from Universiti Teknologi Mara (UiTM), her M.Sc. from Universiti Malaysia Pahang (UMP), and her BIT from Universiti Malaysia Terengganu (UMT). She is a lecturer at Universiti Sultan Zainal Abidin, specializing in software project management, cryptography and cryptanalysis (including key management and random number generation), data hiding and steganography, data security and digital security and privacy, and data encryption.

Nazirah Abd Hamid is a lecturer at Universiti Sultan Zainal Abidin, Terengganu, Malaysia. She received her BIT from Universiti Utara Malaysia (UUM), her M.Sc. from Universiti Teknologi Malaysia (UTM), and her Ph.D. from Universiti Teknikal Malaysia Melaka (UTeM). She specializes in biometric security systems and security services, including digital forensics, steganography, network security, public key infrastructure, and biometrics.

Siti Dhalila Mohd Satar received her BIT from Universiti Kebangsaan Malaysia (UKM), her M.Sc. from Universiti Teknologi Malaysia (UTM), and her Ph.D. from Universiti Putra Malaysia (UPM). She is a lecturer and Academic Program Coordinator (PPA) for the Bachelor of Computer Science (Computer Network Security with Honours) program at Universiti Sultan Zainal Abidin. She specializes in authentication systems and security services, including digital forensics, steganography, network security, public key infrastructure, and biometrics.

Mohd Fadzil Abdul Kadir is a senior lecturer at Universiti Sultan Zainal Abidin. He received his Ph.D. in Engineering from Mie University, Japan, his M.Sc. from Universiti Utara Malaysia, and his B.Eng. also from Mie University, Japan. He specializes in image processing and pattern recognition, encryption systems, data hiding and steganography, and image recognition.

Nur Izzatulain Jamain is a final-year student at Universiti Sultan Zainal Abidin, Terengganu, Malaysia. She is currently pursuing her B.Sc. in Computer Network Security, focusing on proactive security measures and forensic analysis. She also has experience in network traffic analysis, infrastructure security, and IoT security.