Journal of Infrastructure, Policy and Development 2025, 9(1), 10020.
https://doi.org/10.24294/jipd10020

Article

Proposal of an optimum method of audio steganography to secure data transfer

Mst Jannatun Ferdaush[1], **Touhid Bhuiyan**[2,*]

[1] Faculty of Business and Society, University of South Wales, Cardiff CF24 2FN, UK
[2] School of Information Technology, Washington University of Science and Technology, Alexandria VA 22314, USA
*** Corresponding author:** Touhid Bhuiyan, touhid.bhuiyan@wust.edu

CITATION

Ferdaush MJ, Bhuiyan T. (2025). Proposal of an optimum method of audio steganography to secure data transfer. Journal of Infrastructure, Policy and Development. 9(1): 10020. https://doi.org/10.24294/jipd10020

ARTICLE INFO

Received: 31 October 2024
Accepted: 3 January 2025
Available online: 13 January 2025

Abstract: This paper presents an effective method for performing audio steganography, which would help in improving the security of information transmission. Audio steganography is one of the techniques for hiding secret messages within an audio file to maintain communication secrecy from unwanted listeners. Most of these conventional methods have several drawbacks related to distortion, detectability, and inefficiency. To mitigate these issues, a new scheme is presented which combines the techniques of image interpolation with LSB encoding. It selects non-seed pixels to allow reversibility and diminish distortion in medical images. Our technique also embeds a fragile watermarking scheme to identify any breach during transmission to recover data securely and reliably. A magic rectangle has also been used for encryption to enhance data security. This paper proposes a robust, low-distortion audio steganography technique that provides high data integrity with undetectability and will have wide applications in sectors like e-healthcare, corporate data security, and forensic applications. In the future, this approach will be refined for broader audio formats and overall system robustness.

Keywords: secure data; audio steganography; optimum method; information security; visual cryptography e-healthcare; forensic application

1. Introduction

Steganography is the practice of embedding secret messages within public communications in such a way that, while an eavesdropper may intercept the communication, they cannot detect even the presence of the hidden message. Despite extensive research proposing new steganographic protocols and analyzing vulnerabilities in existing methods, there has been little effort toward formalizing the underlying concepts of steganographic security. Second, rigorous and comprehensive proofs of security within non-trivial models remain largely unexplored (Abiodun et al., 2023). Being both an art and a science, steganography ensures that only the sender and the intended recipient are aware of the hidden message. This technology has diverse applications, ranging from transmitting personal messages to securely transferring sensitive data between locations. By making transmissions appear indistinguishable from normal communications, steganography enables covert and undetectable information sharing.

This technology can also be utilized in network topologies, as explained in this page. This way is very handy for hiding communication between botnets and other systems controlled by hackers. Because some procedural packets are just very common and commonly overlooked, it might potentially be utilized to more disguise the source and destination of data. Finding when and how a system was infected via a packet dump might take a well-trained malware analyst hour to weeks. A well-designed network steganographic program could sustain longer periods.

Electronics are advancing in all aspects of our lives. Additionally, e-healthcare is a godsend, which is not only in terms of quality but also in terms of security and data transfer, ensuring that all patients receive the most effective therapies. The stated goal is difficult to fulfill, but if we can overcome difficulties such as electronic patient record (EPR) verification, payloads of medical images, secure transmissions, and precise recovery of sent data for correct diagnosis, we will be rewarded with the greatest healthcare available. Cryptography can solve security and content authentication problems. Data concealing including inserting EPR, logos, and other significant data in medical photos, may be utilized to improve safety and evade undesirable attention to data (Divya and Reddy, 2012; Jayaram et al., 2011). Using traditional hiding schemes, this data embedding could cause medical images to be distorted (Djebbar et al., 2011; Kaur and Behal, 2014). Such distortions, particularly in medical photographs (Mehta and Sihag, 2014), cannot be overlooked (Agrawal et al., 2023). We must utilize new data-concealing strategies to ensure that such circumstances do not happen.

2. Historical background

Steganography, derived from the Greek words “steganos” (covered) and “graphia” (writing), is the art of concealing a message within seemingly innocuous communication. This ancient practice, dating back to 440 BC in ancient Greece, has been employed throughout history to keep information hidden. Early techniques included writing messages on wood and coating them with wax, using invisible inks that were readable by heating or exposing them to light, and reducing documents to “microdots” in the times of World War II. Other ingenious techniques include null ciphers, where unencrypted messages are hidden within regular text. Nowadays, so many different modern methods and tools are in use for keeping information private.

One point that must be pointed out about steganography is the differentiation between cryptography. Though both are engaged in a process of making information unavailable, they stand differently. While cryptography deals with repressing data into an unreadable form through algorithm and key, steganography seeks to hide, rather, the very existence of the message, making it common and ordinary communication as usual. Most of the existing approaches in audio steganography suffer from problems such as perceptible distortion, low data capacity, and vulnerability to detection by attackers. For example, the conventional LSB-based methods, though simple and widely used, are prone to high distortion and are less robust against malicious attacks, as will be discussed in our related work section.

Our proposed approach overcomes these limitations by including a new combination of image interpolation techniques, fragile watermarking, and magic rectangle encryption. This technique will not only reduce distortion but also enhance robustness and integrity. In particular, fragile watermarking allows detecting unauthorized alterations during transmission, while the use of non-seed pixels ensures reversibility and minimal impact on the quality of the carrier file. These innovations thereby present a very valuable approach compared to other traditional techniques.

Furthermore, integration of the magic rectangle encryption increases the security of the embedded data manifold, adding more complexity and making unauthorized decryption nearly impossible. This provides confidentiality for sensitive information, even in highly monitored environments. The hybrid approach increases the versatility of audio steganography to be applied in high-security domains such as healthcare, corporate data protection, and forensics. Our approach overcomes the weaknesses of the existing methods and provides a more holistic approach toward secure communication. Future work will focus on optimizing the method for real-time applications and extending compatibility to a broader range of audio formats, ensuring greater practicality and adaptability in diverse scenarios.

3. Materials and methods

3.1. Problem definition

A basic steganographic model is depicted in the **Figure 1** below. Steganography model—steganography tutorial—Edureka.

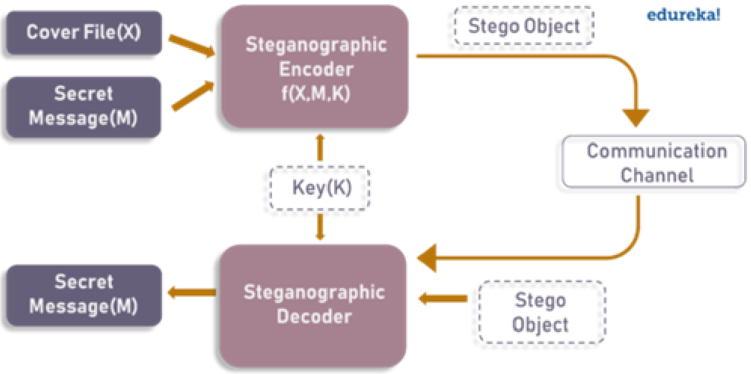

Figure 1. Problem definition flow.

3.2. Related work

Digital steganography has received a lot of attention. In 1996, the inaugural International Workshop on Information Hiding took place, followed by five further workshops and even a book on the subject. Surprisingly, nothing has been done to formalize steganography, and most of the research is based on heuristic approaches, such as steganography with digital photos, steganography with video systems, and so on. Information theoretic models for steganography have been proposed in a few articles (Balgurgi and Jagtap, 2012; Divya and Reddy, 2012; Kaur and Behal, 2014; Nehru and Dhar, 2012), but they are constrained in the same manner that Information theoretic cryptography is what it’s called.

3.3. Research contribution

The main contribution of this research could also be summarized as follows: -

Audio steganography is especially helpful in cases where covert communication is required, since it has some inherent advantages over image and video steganography:

Healthcare: Audio files find extensive use in telemedicine or patient monitoring systems for delivering critical information like doctor-patient consultations or medical advice. Audio steganography secures this communication imperceptibly, which plays an important role in guarding patients' privacy and adhering to healthcare regulatory compliance.

Forensics: Audio recordings are very common in forensics. Steganography can embed additional metadata, like timestamps or case identifiers, directly into the audio file without changing their perceptible quality. The capability for ensuring the integrity and authenticity of evidence in such a way that it keeps its admissibility in court.

Compared to image and video steganography, audio steganography enjoys the masking effect of the human auditory system, whereby slight changes to the audio signals are well below the threshold of human listeners. Moreover, the size of an audio file is smaller compared to video files, hence easier to transmit covertly with still ample embedding capacity for secure information. The suggested method allows us to create data concealment optimally. Not only does the proposed method allow us to incorporate EPR, but it also allows us to validate the material using a brittle watermark and hiding. A new picture interpolation technique has been presented to ensure the ability of medical imaging to be reversed. The image for the cover is created by interpolating the cover image's dimensions. Only the non-seed pixels of something like the cover photo are used to embed EPR to assure reversibility. Aside from the EPR, a fragile watermark is included to identify any infringement during transmissions. This watermark guarantees safe delivery. When the watermark placed at the receiving end is not retrieved, any attack on the image is confirmed. If the extracted watermark differs from the sent one, suggesting infringement, a request for retransmission is given for saving time; if the implanted and received watermarks are identical, we can extract the EPR. We encrypt EPR To ensure its safety and security utilizing a magic rectangle. The substitution of the least significant bit (LSB) is used for embedding.

3.4. Scope

The scope of this research is immense and critical in many fields. In business, audio data hiding could be used to protect confidential information, such as the formula for a certain chemical or the blueprint for an invention. This technology is also used in the non-commercial world, where individuals can keep personal information private. It is also subject to abuse by terrorist groups who can hide their communications and plan attacks using this technique. In addition, audio data hiding plays an important role in government agencies for information system security, covert communications, and the protection of copyrighted digital assets. It has very important forensic applications that allow verifying spoken words and other sounds by embedding hidden data in audio recordings. Moreover, it can be used in the music industry for tracking song broadcasts. This research encompasses the investigation of audio data hiding techniques in the framework of the ARTUS1 project, which is focused on the integration of animation characteristics into audio and video. The study focuses on the pre-processing step where a random bit pattern, obtained from a True

Random Number Generator, replaces the least significant bits of the audio samples. This paper does not go into the detailed implementation of the TRNG.

4. Literature review

4.1. Introduction

Audio steganography involves hiding information in an audio signal (**Figure 2**), taking advantage of the weak sides of human hearing. It is different from image steganography because it is not easily detected by the listener. However, this may affect the sound quality. However, due to the masking effect of HAS, audio steganography has a natural advantage. The masking effect, where higher-frequency sounds mask the lower-frequency ones, allows for subtle modifications of the audio signal, undistinguishable by the human ear, thus easily allowing the stego file to pass the "listening test" (**Figure 3**).

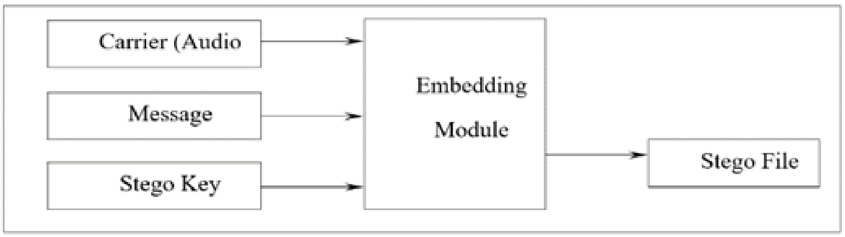

Figure 2. Visual diagram of audio steganographic model.

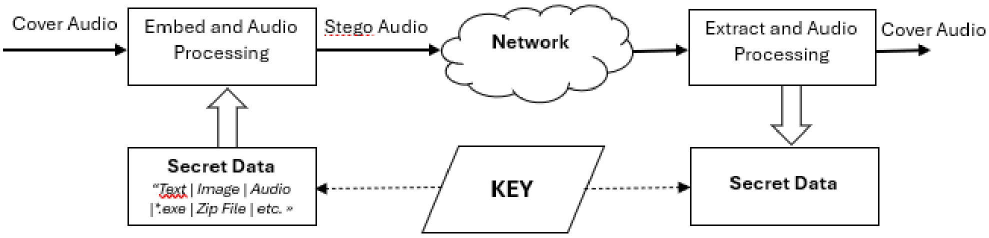
Cover Audio
Embed and Audio Processing
Stego Audio
Network
Extract and Audio Processing
Cover Audio
Secret Data
"Text | Image | Audio |*.exe | Zip File | etc. »
KEY
Secret Data

Figure 3. Basic audio steganographic model (Jayaram et al., 2011).

Various researchers have advanced the state-of-the-art in the audio-steganography field; however, most existing approaches still grapple with various problems about distortion and detectability, compounded with issues regarding capacity. The following is the relative presentation of representative methods that address the corresponding limitations:

LSB coding: LSB is the most straightforward technique used for embedding secret messages. In this method, the least significant bits of the audio samples are modified, which causes minimal variations that usually remain unnoticed. On the other hand, this technique may face serious distortion when large volumes of data are hidden. In addition, the method is susceptible to steganalysis tools, which detect and damage the hidden messages with much ease. This approach ensures minimum distortion with reversibility by adopting the non-seed pixels for data embedding.

Parity coding: Parity coding gains robustness by dividing the signal into blocks and embedding bits into the parity values. Although it offers better resistance to noise than LSB, its low data transmission rate prevents its use in applications that demand high-capacity data embedding. Our proposed method overcomes this deficiency by optimizing data embedding efficiency using a hybrid technique that employs interpolation and fragile watermarking.

Echo hiding: In Echo hiding, information is embedded by introducing echo signals in the audio. It has a good quality of imperceptibility and poor capacity with high computation complexity. Further, the echoes generated sometimes degrade the perceived audio quality, which may be unacceptable in certain critical applications such as medical or forensic areas.

Spread spectrum: Spread spectrum techniques spread the hidden information in the frequency spectrum of the audio signal. These methods are highly robust but result in noticeable distortions, especially in environments with low noise tolerance. Such distortions are not acceptable in applications like e-healthcare, where the integrity of audio files is critical for diagnosis and communication.

Phase coding: Phase coding is another technique that is blind, since it modifies the phase of the audio signal without affecting its amplitude. However, its payload is severely limited, since only the first portion of the signal can carry the secret data. This reduces its applicability in scenarios where large-scale data embedding is required (Lopez and Kumar, 2023).

Genetic algorithm-based techniques: Advanced techniques using genetic algorithms in combination with LSB or other methods enhance data capacity and robustness. However, most of the advanced techniques involve complex computation with higher time cost, making them unsuitable for real-time or resource-constrained environments.

How our method overcomes these weaknesses: Our proposed method overcomes the above-mentioned limitations by incorporating robust encryption through a magic rectangle, reversible data embedding through image interpolation, and fragile watermarking for tamper detection. This approach achieves:

- Reduced distortion: Embedding data in non-seed pixels minimizes distortion, preserving the audio file’s perceptual quality.
- Enhanced robustness: Fragile watermarking provides a layer of security that detects unauthorized alterations during transmission.
- Higher capacity: Our hybrid approach optimizes data embedding, allowing for greater payloads without compromising imperceptibility or robustness.

This combination of features positions our method as a significant advancement in audio steganography, addressing critical weaknesses in existing approaches while catering to the demands of high-security applications like healthcare and forensics.

4.2. Relevant works

Any effective audio steganographic scheme should possess the following three characteristics: perceptual transparency or inaudibility of distortion, capacity or data rate, and robustness. These three requirements are often represented as the “magic triangle” of data hiding, indicating the trade-offs inherent in these characteristics (Kekre et al., 2010). Djebbar et al. (2011) presented a detailed review of some digital audio steganography techniques. Their study showed that, due to the increasing demand for security in digital information, new steganography methods have been developed rapidly in recent years. Because of its easy availability and popularity, audio has become one of the most used carrier media for hidden messages. The main aim of audio steganography is to hide messages in digital audio files without any perceptibility by human audition.

Djebbar et al. performed a comparative study of different audio steganography techniques that have so far been developed. It critically analyzed the strengths and weaknesses of various approaches to help identify the most appropriate techniques for use in a given application. Its results suggested that frequency domain techniques outperform temporal domain techniques in terms of both capacity and imperceptibility, as well as resistance to detection (Kumar and Singh, 2022). Mehta and Sihag (2014) propose a technique that increases security by requiring all files for data decryption; otherwise, information will not be able to be discovered. The LSB (Least Significant Bit) approach is used to hide information in a media file, which replaces the LSB of audio with binary data (**Figure 4**). In the geographic realm, steganography techniques based on the LSB are highly common. The Steganographic technique is shown as follows:

Audio steganography has become a vital area of interest in recent years, with different methods being developed by researchers to increase the robustness of data hiding in an audio signal. This paper reviews five outstanding recent studies, explaining their methodologies and strengths and weaknesses. Nasr (2024) introduces a new scheme that will merge image interpolation techniques with fragile watermarking for embedding data into audio signals. It henceforth intends to provide reduced distortion and more robustness against several attacks. The presented method seems to be quite promising for enhancement in terms of imperceptibility and security, though it may suffer from complexity issues while being utilized for real-time applications (Nasr et al., 2024).

Aslantas and Hanilci present a comparative study of five audio steganography techniques: LSB, echo hiding, wavelet coding, spread spectrum, and kepstrum. The results of each method concerning imperceptibility, robustness, and capacity are discussed. The results show that while some techniques result in high capacity, their distortion is higher, or the robustness is low. This proves the trade-off inherent in the methods of audio steganography (Aslantas and Hanilci, 2022). Saranya and Reddy's (2024) research presents a MATLAB-based audio steganography method that integrates LSB coding with RC7 encryption and chaotic decryption. The combined approach aims to enhance the security and robustness of the hidden data. However, the reliance on LSB coding may still render the method susceptible to certain steganalysis attacks, and the computational complexity could be a concern for large-scale implementations (Saranya and Reddy, 2024).

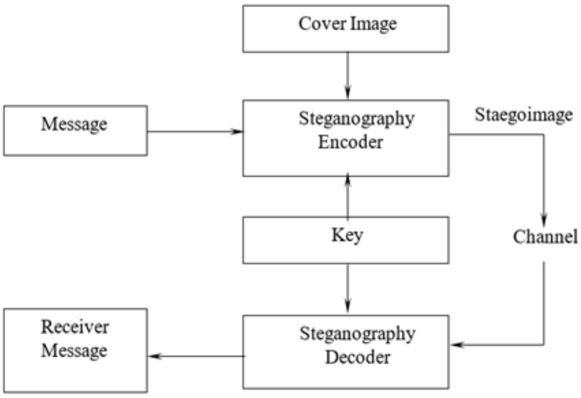

Figure 4. Steganography scenario.

Li et al. introduce a coverless audio steganography model using generative adversarial networks to hide secret audio without changing any existing cover audio (**Figure 5**). This method synthesizes the stego-audio directly, hence enhancing security by avoiding detectable changes in cover audio. While innovative, this approach has effectiveness highly reliant on the quality of the GAN model, and the training may be resource-consuming (Li et al., 2023). Marszalek and Bilski's research appraises the efficiency of different commercial off-the-shelf software in conducting audio steganography. The analysis focused on the trade-off between the perceptual transparency, robustness, and capacity of embedded messages. The results found that while COTS tools were quite user-friendly, the efficiency of these tools on several steganographic criteria was not optimal, hinting at the need to proceed with more advanced and tunable solutions (Marszalek and Bilski, 2022).

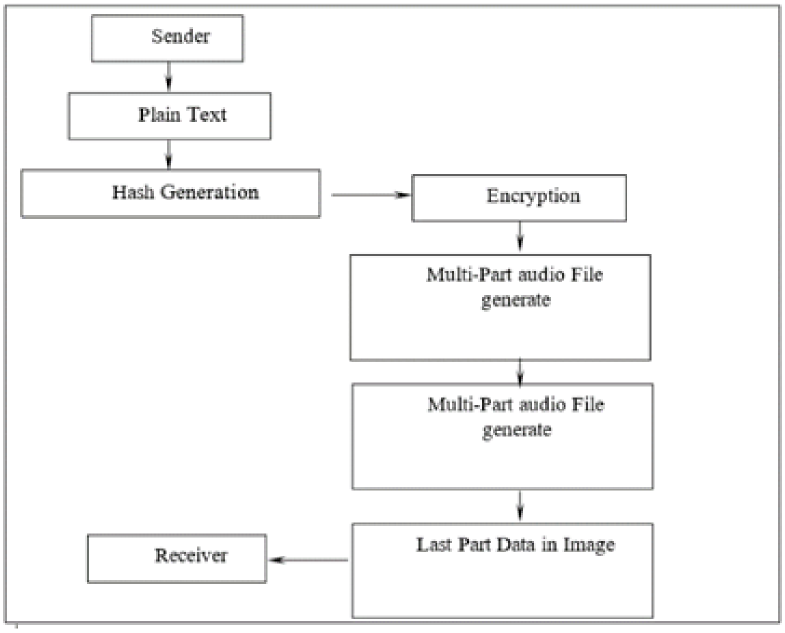

Figure 5. Flowchart.

In a nutshell, recent advances in audio steganography have presented new methods towards enhancing security, robustness, and capacity. However, there is still some imbalance in these factors, and further research is necessary to develop methods that will be effective and practical in real-world applications.

Singh (2016) provided a comprehensive overview of various concepts and recent techniques in audio steganography, offering a detailed examination of methodologies like Parity Coding, Least Significant Bit (LSB) coding, Phase Coding, Echo Data Hiding, and Spread Spectrum. Each technique was analyzed in terms of its strengths, weaknesses, and capacity to conceal information effectively. For instance, Parity Coding divides audio signals into blocks and embeds secret data into their parity bits, ensuring robustness but often at the cost of limited data capacity. Similarly, the LSB approach, while straightforward and capable of embedding significant amounts of data, suffers from low resistance to steganalysis and perceptible distortion when used directly.

Phase Coding offers high imperceptibility by altering the phase of the audio signal, though its capacity remains limited, making it more suitable for smaller data payloads. Echo Data Hiding leverages echo signals to embed information, achieving high robustness but introducing perceptible distortions that may compromise audio quality. Spread Spectrum techniques, meanwhile, distribute hidden data across a wide frequency range, offering strong resistance to detection but often requiring significant computational resources. Singh also highlighted recent advancements in the field, such as the application of genetic algorithms to audio steganography. This approach enhances data concealment and robustness by dynamically adapting to the characteristics of the audio signal. However, the use of genetic algorithms introduces computational complexity, which may limit their practicality in real-time scenarios.

Coding in the LSB format: It is possible to use the least significant bit because most alterations do not result in perceptible changes to the sounds (**Figure 6** and **7**). Another approach is to make use of human limitations.

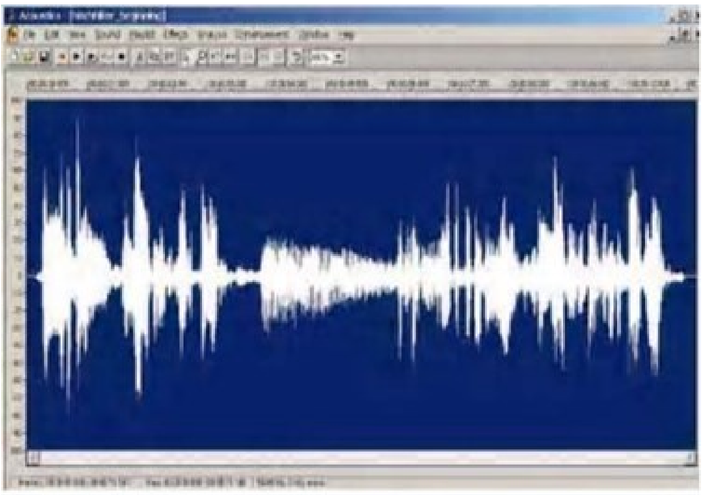

Figure 6. Comparisons between a WAV carrier file before the LSB coding.

Figure 7. Comparisons between a WAV carrier file after the LSB coding.

Communications can be encrypted using unique, undetectable, and human-audible frequencies. By embedding messages within sound files at frequencies exceeding 20,000 Hz (ultrasound), they become imperceptible to human hearing and thus evade detection through typical human inspection (Bender et al., 1996). Divya and Reddy (2012) enhanced traditional LSB steganography by implementing multiple LSB substitution techniques. This approach significantly increased the capacity of cover audio for data embedding by modifying multiple, varying LSBs within each audio sample. These techniques, utilizing up to 7 LSBs for data embedding, demonstrated a 35–70% increase in data hiding capacity compared to conventional methods that typically employ 4 LSBs. Cryptographic security was further strengthened through the integration of the RSA algorithm.

Asad (2012) proposed a novel three-layered scheme for audio steganography based on LSB replacement. This approach involves passing the secret message through two intermediate layers before embedding it within the cover audio in the third layer. Balgurgi and Jagtap (2012) integrated two fundamental aspects of network security: cryptography and steganography. They developed a two-level encryption technique for user data by combining the LSB approach with the XORing method, providing an additional layer of security. Rosaline and Raj (2013) addressed the limitations of traditional LSB steganography, such as increased distortion and reduced embedding efficiency. They proposed a novel method for hiding secret messages within images using Adaptive Pixel Pair Matching. This innovative approach, known as the Optimum Pixel Adjustment Process (OP AP), effectively overcomes the drawbacks of the conventional LSB method.

Rakshit and Ganguly have proposed a new two-layer security system that adds significant strength to the traditional cryptographic systems. This system embeds digital watermarking or fingerprinting with digital watermarking, effectively integrating an identity symbol or signature data into the original information. Their approach leverages a combination of Visual Cryptography and Audio Steganography encoding algorithms. This primarily follows the intensity-based visual cryptographic scheme, where the secret image's intensity of the pixels is distributed across all basis matrices. These bases eventually construct eight different share images in this approach. The generated shared pieces make re-creation of the exact original secret image possible via superimposition. Meanwhile, the audio steganography technique in the time domain performs. Nehru and Dhar (2012) conducted an in-depth study of audio steganography techniques, focusing on two prominent approaches: the Least Significant Bit (LSB) method and the genetic algorithm. Their research comprehensively investigated the characteristics and performance of these audio steganography techniques, providing valuable insights into their strengths and weaknesses.

Yu et al. (2001) presented a spatial-domain color image watermarking technique. In their method, a binary watermark was encoded, and a neural network was trained by using a set of training patterns. These patterns were obtained from the differences of intensity between the blue component of the central pixel and its neighboring pixels in a window of pre-specified size. Each training pattern consisted of nine input vectors and one output vector. The trained neural network was utilized to perform the extraction of the watermark from images belonging to legitimate users.

Wang et al. presented a new image data watermarking system that combines neural networks with DWT. PRNG was used to choose a set of coordinates from the DWT decomposition of the image. A training set was formed for the neural network, in which each training pattern was made up of eight input vectors and four predicted outputs. The trained network was then used to embed the watermark into the image (Wang et al., 2023).

The shortcomings of some previous methodologies in audio steganography will be presented in this section and compared with the proposed one. Most classic techniques, including echo hiding, spread spectrum, and parity coding, usually have defects of poor robustness since the human ear is sensitive to any subtle added noise into the audio. Phase coding is effective, but it has a very low data transmission rate because the secret message is encoded only in the initial segment of the signal; it is applicable only for small volumes of data. One of the commonly used methods is the Least Significant Bit method, in which each bit of the secret message replaces the least significant bit of each audio sample. Though the LSB method provides a high data embedding capacity, it may introduce audible distortions in some cases, affecting the quality of the stego-audio. One of the commonly used methods is the Least Significant Bit method, in which each bit of the secret message replaces the least significant bit of each audio sample. Though the LSB method provides a high data embedding capacity, it may introduce audible distortions in some cases, affecting the quality of the stego-audio. **Table 1** compares different steganography methods based on their strengths, weaknesses, embedding methods, and hiding rates.

This proposal tends to eliminate the said limitations through presenting a much more secure and efficient technique for hiding secret information within audio. It highlights the advantages in key improvement on security against unauthorized access, with an improvement on resilience against noise and distortion; high amount of data can be embedded without degradation of audio. Besides, the proposed method will not change the file size of the original file and supports several audio file formats; therefore, it is an effective and versatile technique in covert communication.

Table 1. Summary and evaluation of audio steganography methods.

Hiding Domain	Methods	Strengths	Weaknesses	Hiding Rate
Temporal Domain	Low bit encoding	Simple and Easy way of hiding information with a high bit rate	Easy to extract and destroy	16 kbps
	Echo hiding	Resilient to lossy data compression algorithm	Low security and capacity	40–50 bps
Frequency Domain	Magnitude spectrum	Longer messages to hide and less likely to be affected by errors during transmission	Low recovery quality	3 kbps
	Tone insertion	Imperceptibility and concealment of embedded data	Lack of transparency and security	250 bps
	Phase spectrum	Robust against signal processing manipulations and data retrieval needs the original signal	Low capacity	333 bps
	Spread spectrum	Provide better robustness	Vulnerable to time-scale modification	20 bps
	Cepstral domain	Robust against signal processing operations.	Perceptible signal distortions and low robustness	54 bps
Wavelet Domain	Wavelets coefficients	Provide high embedding capacity	Lossy data retrieval	200 kbps
Codecs Domain	Codebook modification	High robustness	Low embedded rate	2 kbps
	Bitstream hiding	High robustness	Low embedded rate	400 bps

5. Methodology

5.1. Magic rectangle encryption

Magic rectangle encryption plays a vital role in securing embedded data. In this technique, data is arranged into a matrix or rectangle such that each row and column will satisfy one pre-defined cryptographic condition, thereby increasing randomness and making it computationally impossible for unauthorized decryption.

5.2. Example of magic rectangle encryption

Consider embedding a secret message, "DATA," where the ASCII values of each character are arranged into a 2 × 2 magic rectangle:

68 (D)	**65 (A)**
84 (T)	65 (A)

Using a predefined cryptographic key, the values are shuffled and encoded to generate a secure pattern. At the receiver's end, the same key is used to reconstruct the original matrix and retrieve the data. This encryption ensures that even if the carrier file is intercepted, the hidden message will remain incomprehensible without the decryption key.

In **Figure 8**, the diagram illustrates the setup of the magic rectangle and its transformation process.

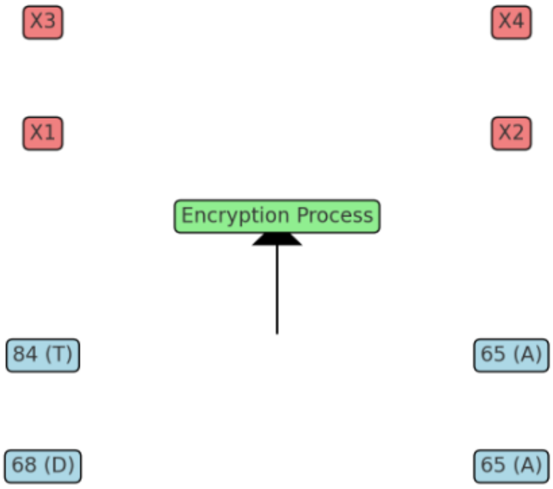

Figure 8. The setup of the magic rectangle and its transformation process.

- Top section: Shows the initial magic rectangle with ASCII values of the secret message arranged in a matrix.
- Arrow and middle section: Represents the encryption process that transforms the original values into encrypted placeholders (X1, X2, etc.).
- Bottom section: Displays the resulting encrypted magic rectangle.

This visualization demonstrates how the encryption secures the data.

5.3. Fragile watermarking for tamper detection

In a fragile watermarking scheme, on the other hand, there is an intention to provide detection against unauthorized signal changes at some level of transmission. A technique includes embedding the watermark (perhaps a checksum or hash value) within selected parts of an audio signal.

5.4. How tamper detection works

Embedding phase: During data embedding, the watermark value calculated from the original signal and secret message is embedded in carrier audio. Detection phase: The watermark is extracted at the receiver's end from the received signal and recalculated. Any mismatch between the embedded and recalculated watermark will indicate tampering.

5.5. Ensuring data integrity

In the case of tampering, the system flags the corrupted data for retransmission. Therefore, only verified and unchanged information is processed and retrieved to retain the integrity of the hidden data.

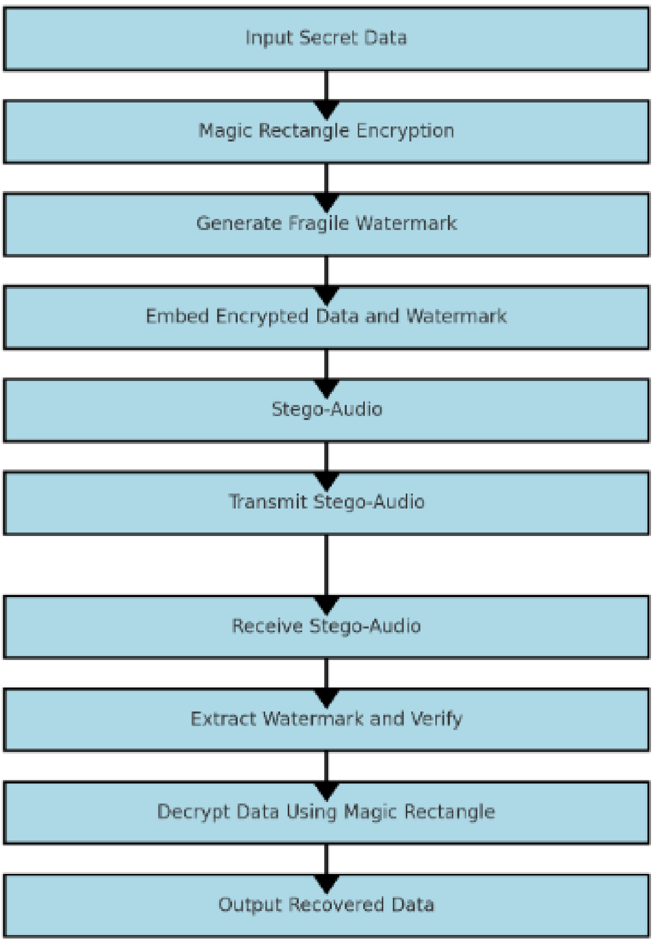
Input Secret Data
Magic Rectangle Encryption
Generate Fragile Watermark
Embed Encrypted Data and Watermark
Stego-Audio
Transmit Stego-Audio
Receive Stego-Audio
Extract Watermark and Verify
Decrypt Data Using Magic Rectangle
Output Recovered Data

Figure 9. Flowchart of embedding and retrieval process.

In **Figure 9**, the flowchart illustrates the embedding and retrieval process:

- Input Secret Data: The process begins with the input of the data to be hidden.
- Magic Rectangle Encryption: The data is encrypted using magic rectangle technique.
- Generate Fragile Watermark: A fragile watermark is created to ensure data integrity during transmission.
- Embed Encrypted Data and Watermark: The encrypted data and watermark are embedded into the audio carrier.
- Stego-Audio: The resulting stego-audio contains hidden data and is ready for transmission.
- Transmit Stego-Audio: The stego-audio is sent over the communication channel.
- Receive Stego-Audio: The receiver obtains the stego-audio for processing.
- Extract Watermark and Verify: The watermark is extracted and verified to detect any tampering.
- Decrypt Data Using Magic Rectangle: The encrypted data is decrypted to retrieve the hidden information.
- Output Recovered Data: The original secret data is output after successful verification and decryption.

6. Conclusion and future scope

6.1. Conclusion

Researchers have developed different steganography techniques to increase the security of information transmitted over networks. This paper highlights the significance of steganography in the modern digital world. Audio steganography is one of the better methods that has been found to hide messages in audio. The combination of the LSB approach with other cryptographic methods, such as encryption and decryption, offers a highly effective and secure communication framework. However, direct application of the LSB method can introduce distortions to the audio cover file. While the parity approach exhibits high SNR, it often entails a large data rate. Also, the XORing approach, while able to expand data embedding capacity by exploiting multiple layers of LSB, may not be fully functional with certain file formats like.wav.

Future strides in steganography must overcome these challenges by allowing more file formats and paving the way for new, emerging security threats. This paper proposes a new approach that offers more enhanced data integrity, a reduction in distortion, and increased applicability of audio steganography. With the “magic rectangle” integrating various new ideas on encryption, added with fragile watermarking, it paves a wider way and is much stronger in security-related communication. As technology continues to evolve, these methods can be extended for a broader spectrum of applications including real-time secure communication, IoT devices, and multimedia content distribution, ensuring steganography remains a significant tool in safeguarding sensitive information in the digital age.

6.2. Future scope

The results reveal that even with the changes made to the direct use of LSB, there is still room for development in highly secure data retrieval, better robustness, higher data rate, full recovery of cover audio, and compatibility with various audio formats. A complete system for LSB is being developed under our research team, which will be fully demonstrated in our next paper. The future scope of the research involves adapting the proposed method for real-time applications by optimizing algorithms for reduced computational overhead and leveraging parallel processing to enable secure live audio streams. The method will be extended to support multiple file formats, including compressed types like mp3 and aac, addressing challenges posed by lossy compression. The robustness against emerging threats will be enhanced by the inclusion of advanced encryption techniques, such as quantum-safe cryptography, and the development of adaptive embedding algorithms to minimize detectability. Integration with technologies like blockchain for secure logging, AI-based optimization for improved embedding, and lightweight adaptations for IoT devices will further enhance the method’s capabilities. The approach will be further extended for various applications of secure content distribution, education, and training; also, the development of user-friendly tools or APIs will make it more versatile and accessible in a number of domains, including healthcare, forensics, and the entertainment industry. This will help to set the approach as a robust and future-ready solution for secure audio steganography.

Author contributions: Conceptualization, TB; methodology, TB; software, MJF; validation, MJF, and TB; formal analysis, MJF; investigation, MJF; resources, TB; data curation, MJF; writing—original draft preparation, MJF; writing—review and editing, TB; visualization, MJF; supervision, TB; project administration, TB; funding acquisition, TB. All authors have read and agreed to the published version of the manuscript.

Conflict of interest: The authors declare no conflict of interest.

References

Abiodun, A. O., Mohamad, E. Z., & Oyewola, D. O. (2023). Towards robust image-in-audio deep steganography.

Agrawal, M., & Tyagi, S. (2023). StegoHound: A novel multi-approach method for efficient and effective identification and extraction of digital evidence masked by steganographic techniques in WAV and MP3 files.

Asad, M., Junaid Gilani & Adnan Khalid, (2012). Three Layered Model for Audio Steganography, International Conference on Emerging Technologies.

Aslantas, F. and Hanilci, C. (2022). Comparative Analysis of Audio Steganography Methods. Journal of Innovative Science and Engineering (JISE), 6(1).

Balgurgi, P.P. & Jagtap, S.K., (2012). Intelligent processing: An approach of audio steganography, Communication, Information & Computing Technology, 2012 International Conference on, pp.1-6, 19-20 Oct. 2012.

Bender, W., Gruhl, D., Morimoto, N. & Lu, A. (1996). Techniques for Data Hiding, IBM Systems Journal, 35(3), pp. 313-336, 1996

Divya, S.S. & Reddy, M.R.M. (2012). Hiding Text in Audio Using Multiple LSB Steganography and Provide Security Using Cryptography. International Journal of Scientific & Technology Research, Vol. 1, pp. 68-70, July 2012

Djebbar, F., Ayady, B., Habib, Meraim, HKA. (2011). A view on latest audio steganography, International Conference on Innovations in Information Technology, 2011, Pages 409-414

Jayaram, P., Ranganatha, H.R. & Anupama, H.S., (2011). Information Hiding using Audio Steganography- A Survey, The International Journal of Multimedia & Its Applications, 3(3), 86-96.

Kaur, N., & Behal, S. (2014). Audio steganography techniques-a survey. Int Journal of Engineering Research and Applications ISSN, 2248-9622.

Kekre, H.B., Athawale, A., Rao, S. & Athawale, U., (2010). Information Hiding in Audio Signals. International Journal of Computer Applications, 7(9), October 2010.

Kumar, R., & Singh, P. (2022). A comprehensive review on steganography techniques for text, images, and audio. IEEE Xplore.

Li, J., Wang, K., & Jia, X. (2023). A Coverless Audio Steganography Based on Generative Adversarial Networks. Electronics, 12(5), 1253.

Lopez, M. S., & Kumar, A. (2023). Information security for audio steganography using a phase coding algorithm. European Journal of Technological and Applied Sciences, 5(3), 56-64.

Marszalek, P. & Bilski, P. (2022). Steganography in Audio Files -COTS Software Analysis. International Journal of Electronics and Telecommunications, 69(1):121-126.

Mehta, U., & Sihag, D. (2014). Multi-Part Data Hiding in Audio Steganography. International Journal of Advanced Research in Computer and Communication Engineering, 3(9), 8037-8039.

Nasr, M.A., El-Shafai, W., El-Rabaie, ES.M. et al. (2024). A robust audio steganography technique based on image encryption using different chaotic maps. Scientific Reports 14, 22054.

Nehru, G. & Dhar, P., (2012). A Detailed Look of Audio Steganography Techniques Using LSB And Genetic Algorithm Approach, International Journal of Computer Science (IJCSI), Vol. 9, pp. 402-406, Jan. 2012

Rosaline, S. I. & Raj, M. A., (2013). Adaptive Pixel Pair Matching based Steganography for audio files, Emerging Trends in VLSI, Embedded System, Nano Electronics and Telecommunication System, International Conference on, pp.1-5, 7-9 Jan. 2013.

Saranya, S. S. & Reddy, P. L. C., Prasanth, K. (2024). Digital audio steganography using LSB and RC7 algorithms for security applications. AIP Conf. Proc. 3075, 020083.

Singh, P. (2016). A comparative study of audio steganography techniques. J Int Res J Eng Technol, 3(4), 581-585.

Wang, K., Chen, Y., & Lu, P. (2023). Video Steganography: Recent advances and challenges. Multimedia Tools and Applications.

Yu, P. T., Tsai, H. H. & Lin, J.S. (2001). Digital watermarking based on neural networks for color images, Signal Processing, 81, 663-671.