Review Article

LSB-BASED AUDIO STEGANOGRAPHICAL FRAMEWORK FOR SECURING DATA IN TRANSIT

ABSTRACT

The digital age has emerged with a variety of benefits to organizations and individuals. Transmission of information over the public channels are becoming more widely used every day. The use of the internet as the major source of transmitting confidential data has resulted in the vulnerability of digital data to interception and unauthorized access and usage. Steganography is one of the ways by which data in transit can be secured without attracting unnecessary attention from intruders. In this paper, Least Bit Significant algorithm was used with an audio file for hiding information. The algorithm used in this research proves to be one of the simplest ways of securing data using audio steganography. The methods employed the LSB approaches by using audio files as the stego object for the implementation based in Java Programming Language. The experimental results also proved to be one of the best methods of implementing steganography. The accuracy of the stego objects shows high quality and similarity.

I. INTRODUCTION

Data security is a technique of protecting data from vicious forces and the unwanted actions of unapproved users. Since the less expensive way or approach of exchanging large amount

information confidentially is through the internet, hence, it is necessary to protect users and the data they transmit from one point to the other. The use of the internet as the major source of transmitting confidential data has resulted in the vulnerability of digital data to interception and unauthorized access and usage. This trend has resulted in huge losses to both content producers and owners. To ensure secure information on open channels, efforts to establish safety should be integrated into data communication systems over the web. The incorporation of safety measures into data communication systems is the surest way of protecting and safeguarding data transmission over public channels such as the Internet.

The need to communicate information as safely and as securely as feasible has been a subject of much debate for several years. Data is the fortune of any institution or organization. Due to this, security measures have become an issue of much concern to firms who deal with confidential data. Whichever system is chosen to make communication secure; the issue of major concern is the extent to which the system is safe. Steganography is the art of concealed or hidden communication. The reason for steganography is secret correspondence to conceal information from a third party.

Steganography is the method or technique of hiding a file, image, or message inside a different file, image, or message. Steganography is evolving digital media as it allows just the sender and the intended recipient to be able to identify the information transmitted through it. Steganography is regularly mistaken for cryptology because the two have some similarities as they are utilized to secure critical data. They vary because steganography consists of hiding data to create the impression that no message is covered at all

One major drawback with most of the information that are transmitted on the internet is that information is transmitted in a format which intruders can read and understand without difficulty.

After successfully acquiring the information illegally, intruders might divulge sensitive data such trade secrets to the public or other organizations, distort the information to malign a person or an organization or sometimes it is used to initiate attacks on these individuals and organizations. Steganography is one of the best methods that can be employed to curb this unpleasant and devastating act and trend.

With the current increase in usage of traffic security systems, military and other security organization secure their data by concealing the sender, the receiver and the content of the message using steganography. In digital elections, similar approaches are being proposed and adopted using mobile phone systems.

A few of the methods utilized as part of steganography are domain tools such as easy systems like Least Significant Bit (LSB) for embedding and noise manipulation, and transformation of domain which comprises manipulating algorithms and transforming images like discrete cosine transformation and wavelet transformation. Nonetheless, there are procedures that have both photo and domain tools like patchwork, pattern block encoding, spread spectrum techniques and concealing.

II. RELATED WORK

Data security is the process or the art of protecting data from vicious forces or users and the unsolicited activities of unapproved users. An enormous quantity of confidential data is transferred through the web or internet in public platforms as it is the cheapest and commonly available method. This technological growth and advancement have additionally rendered digital information highly susceptible to interception and then probable unapproved access and or use and have resulted major economic losses for content creators and rights holders.

In order to ensure that information available on open channels is secured, safety measures have to be integrated into data communication systems through the web [1].Steganography is part of the great technologies which aid in the attainment of the general target of secure transfer of information from senders to approved recipients. Steganography is method of hiding a file, image, or message inside a different file, image, or message. The term steganography has a Greek root which denotes "covered writing" or "concealed writing" [2] .Steganography is evolving the digital media as it allows just the sender and the intended recipient to be able to identify the information transmitted through it.

Although several universal methods are known for securing data in transit, they involve considerable overhead, making them impractical, especially compared to the format implored in their implementation. It is sometimes possible to devise data security techniques and methods that can secure data in transit without the use of format readable by human beings. Such techniques and methodology offer the benefits of securing data from an unauthorized usage without sacrificing efficiency. Steganography is the art and science of concealing information during communication so that it is not discovered [3] by a third party.

In the year 2015, Ayush Singhal et al [4] proposed that for cover objects, different types of digital media can be used and they used .wav audio as their cover file in the research work. They were able to hide the secret message inside the audio cover file.

In the year 2014, Rohit Tanwar and Monika Bisla [5] advised that one of the most important goal of any audio steganographic technique is that the process should be robust and the audio cover file generated must be resistant to malicious attacks as that is the main aim of the steganography process.

In 2014, Kazem Qazanfari and Reza Safabakhsh [6] proposed an improved version of LSB++ approach. In this improved LSB++ they make distinction between sensitive pixels and allow protecting them from the embedding of extra bits, which results in the lower distortion in co-occurrence matrices.

In the year 2012, M. Baritha Begum and Y. Venkataramani [7] proposed an algorithm that included compression that reduces the redundancy of data. In their audio steganographic technique, dictionary based compression bits were hidden in the least significant bit of audio signals and the signal to noise ratio (SNR) was calculated. This audio Steganography was used to conduct for various compression algorithms with dictionary-based compression.

In the year 2009, S. Channalli and A. Jadhav [8] proposed a new LSB based method in which common bit pattern is used to hide data which can be used in audio steganography as well while using the bit patterns with different frequencies of audio signal

The major objective of steganography is to ensure secure communication in a totally untraceable method [9] and to prevent drawing attention to the concealed information being exchanged [10]. Its purpose is not to prevent unauthorized people from decoding the concealed information, but rather to prevent them from perceiving that its existence. If a steganography technique makes somebody to be suspicious of the carrier medium, then the technique is not successful [11]. Until recently, steganography has not received much attention as compared to cryptography. This situation has however changed rapidly and can be attributed to following reasons [12]. First and foremost, the interest of publishing and broadcasting firms in hiding encrypted copyright marks and serial numbers in digital files have increased tremendously. Secondly regulations by successive governments to restrict the availability of encryption services have motivated researchers to study methods by which private messages can be embedded in seemingly innocuous cover messages.

Figure 1 shows a basic steganography model consisting of Carrier, Message and Password proposed by Cachin [13]. Carrier is also known as *cover-object*, which the message is embedded and serves to hide the presence of the message. This model presented the technical details of steganography however not practical implementation was given by Cachin or any other researcher, thereby making the model not to be practically proven. According to the theoretical implementation of the model, message is the data that the sender wishes to remain as confidential, and this can be in any digital readable format [14]. Password is known as *stego-key*, which ensures that only recipient who know the corresponding decoding key will be able to extract the message from a *cover-object*. The *cover-object* with the secretly embedded message is then called the *stego-object*.

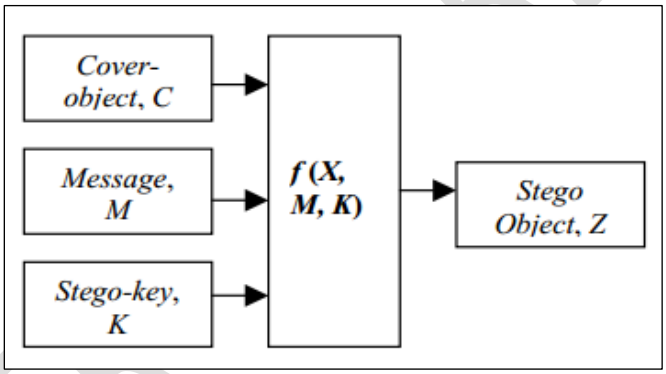

Figure 1:Basic Steganography Model

There are several suitable media that can be used as cover-objects such as network protocols, audio, a text file, video and image files [15].

Cryptography And Steganography

For a steganographic algorithm having a stego-key, given any cover object the embedding process generates a stego object. The extraction process takes the stego object and using the shared key applies the inverse algorithm to extract the hidden message.

Basically, the purpose of cryptography and steganography is to provide secret communication. However, steganography is not the same as cryptography. Cryptography hides the contents of a secret message from a malicious people, whereas steganography even conceals the existence of the message. According to Kessler, "The goal of cryptography is to make data unreadable by a third party, the goal of steganography is to hide the data from a third party" [16]. The most important requirement of any steganographic system is that it should be impossible for an eavesdropper to distinguish between ordinary objects and objects that contain secret data [17].

Table 1:Features of Steganography and Cryptography

Steganography	**Cryptography**
The passing of messages is unknown	The passing of message is known
Little known technology	Common technology
Technology still being developed for certain formats	Most algorithms known to government departments
Once detected, message is known	Strong algorithm are currently resistant to brute force attack Large expensive computing power required for cracking. Technology increase reduces strength
Many Carrier formats	

Steganography is often thought of only as a tool for a malicious user to subvert a security policy, but there are three fundamental classes of applying steganography. These includes subliminal communication [18], integrity and authentication, and illicit exfiltration of data [19].

Steganography Techniques

Over the past few years, numerous steganography techniques that embed hidden messages in multimedia objects have been proposed [10].

1. Least Significant Bits

Least significant bits (LSB) insertion is a simple approach to embedding information in a file. The simplest steganographic techniques embed the bits of the message directly into least significant bit plane of the cover-o in a deterministic sequence. Modulating the least significant bit does not result in human perceptible difference because the amplitude of the change is small.

2. Masking and Filtering

Masking and filtering techniques, usually restricted to 24 bits and gray scale images, hide information by marking an image, in a manner similar to paper watermarks. The techniques perform analysis of the image, thus embed the information in significant areas so that the hidden message is more integral to the cover image than just hiding it in the noise level.

3. Transforms Techniques

Transform techniques embed the message by modulating coefficients in a transform domain, such as the Discrete Cosine Transform (DCT) used in JPEG compression, Discrete Fourier Transform, or Wavelet Transform. These methods hide messages in significant areas of the cover-object, which make them more robust to attack. Transformations can be applied over the entire object, to block throughout the object, or other variants.

Categories of Steganography

There are a lot of digital file format currently in used today. All these digital formats are suitable for the implementation of steganography, however those digital formats with high degree of redundancy is more prefer and suitable than those with low degree of redundancy. For a file to be of high degree of redundancy implies that the bits of that file can be changed without detecting the change easily. Example of such objects is video, audio and image files. With this, image,

video, and audio files are more suitable objects for the implementation of steganography. Figure 2 shows the various categories of file formats that can be used for steganography.

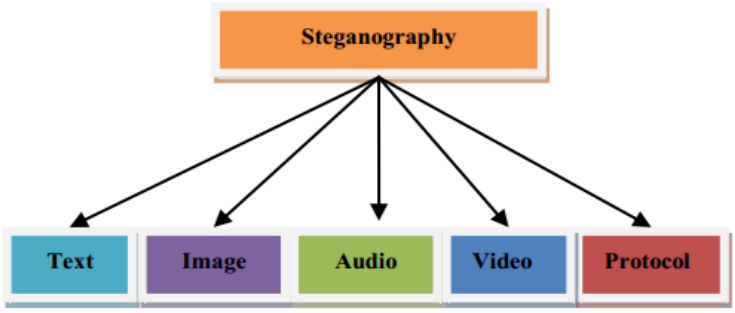

Figure 2: Categories of Steganography

Currently, most of the steganographic systems uses objects like video, image, and audio to implement data hiding Systems. This is because of the tendency at which digital images, audio and video are transmitted over the Internet in the form of emails. From Figure 2, these are the most widely used objects apart from the text.

Protocol steganography is receiving much attention in recent years due to the emergence of social media platforms for transmitting messages. The term protocol steganography refers to the technique of embedding data within messages and network control protocols used in network transmission. In the layers of the OSI network model there exist hidden channels where steganography can be used. An example of where information can be hidden is in the header of a TCP/ IP packet in some fields that are either optional or are never used.

It is worth noting that, steganographic systems can also be classified according to the cover modification applied in the embedding process. This classification scheme can be divided into the following categories.

- **Substitution system** replace unneeded parts of a cover with a secrete data.

- **Transform domain techniques** embed secrete message in a transform space of the signal (e.g., in frequency domain).
- **Spread spectrum techniques** implement ideas from spread spectrum communication.
- **Statistical methods** encode data by changing several statistical properties of a cover and use assumption testing in the extraction process.
- **Distortion methods** accumulate data by signal alteration and measure the deviation from the original cover in the decoding step.
- **Cover generation schemes** encode data in the approach a cover for secrete communication is created.

Properties of Steganography

According to [20], there are few key properties that need must be taken into consideration when creating a digital data hiding system.

- *Imperceptibility*: The goal of steganography is that object should appear identical before and after hiding.
- *Embedding Capacity:* It is the capacity of steganographic algorithm based on the quantum of message it can secretly transmit. Capacity is one of the challenging case in steganography.
- *Robustness:* Robustness refers to the degree of difficulty required to tear down embedded information without destroying the cover object itself.
- *Undetectability:* This property is as important as imperceptibility. It is the rate and accuracy at which a media containing an embedded data cannot be detected using statistical or technological means.

III. SYSTEM DESIGN AND METHODOLOGY

In this study, the researcher considers the Least Bit Square approaches to implementing audio steganography for securing data. The scope of the study is limited to audio steganography as a result of its availability and memory usage utilization in shared memory systems.

The Least-Significant Bit (LSB) Audio Steganography Implementation

This techniques implementation involves all kinds of audio irrespective of the number of channels the audio possessed. This technology involves the hiding of data in audio files. The first bits of every audio sample of sixteen bits (16bits) is either a plus or minus and the rest of the fifteen bits (15 bits) are divided into two groups. The first division has 7bits known MSB while the other division includes 8bits known as LSB. In this way the signals are interrupted, and data cannot be conveyed secure. For proper and secure conveyance, the payload is increased, and signals are improved.in the proposed audio steganography algorithm, an audio file will be considered as a cover object the message or text file is referred to as the secret message to be hidden in cover object.

LSB algorithm is a classic Steganography method used to conceal the existence of secret data inside a “public” cover. The LSB or “Least Significant Bit”, in computing terms, represents the bit at the unit’s place in the binary representation of a number. For example, we can represent the decimal number 170 in binary notation as 10101010. The least significant bit, in this case, is 0.

In the simplistic form, LSB algorithm replaces the LSB of each byte in the “carrier” data with one bit from the “secret” message [21].

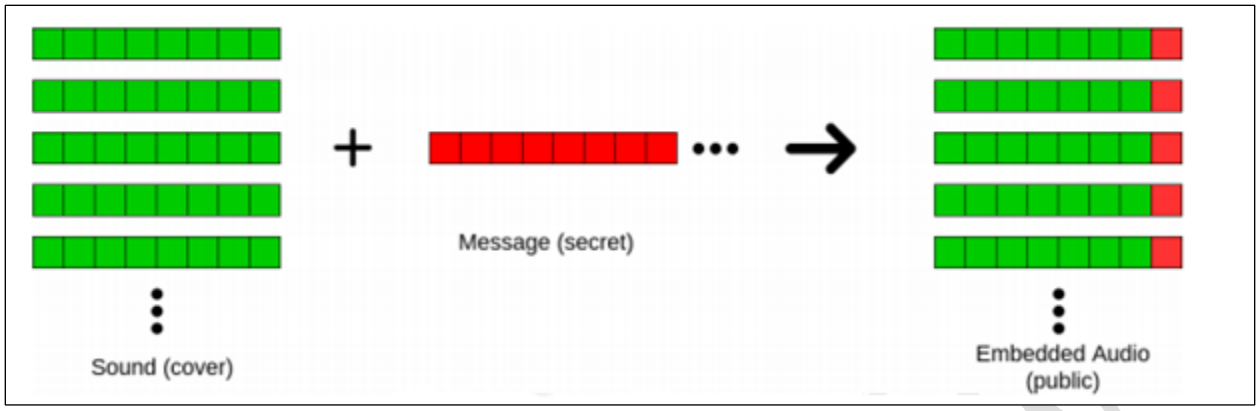

Figure 3:Encryption process

The sender performs “embedding” of the bits of secret messages onto the carrier data byte-by-byte. Whereas the receiver performs the “extraction” procedure by reading LSB bits of each byte of received data, this way the receiver reconstructs the secret message.

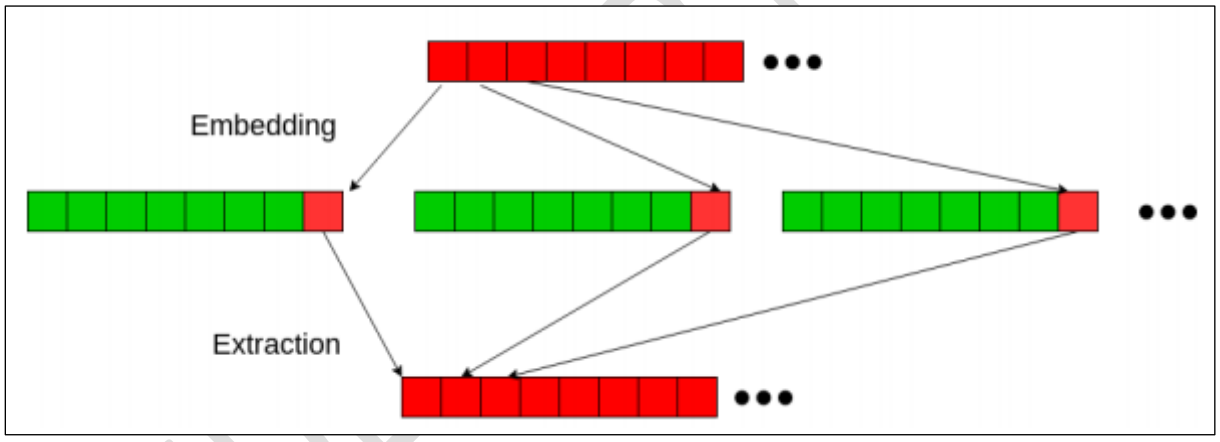

Figure 4:Embedding and Extraction process

The advantage of the LSB techniques lies in its ease of implementation and simplicity. The LSB method allows high embedding capacity and uses different frequency levels for more security. Hiding the secret data using audio lowers the chances of the secret data being detected. This techniques for audio files work smoothly for all audio format as implemented in Java. Using these algorithms for encoding and decoding, one can retrieve the secrete message exactly as the original data.

IV. RESULTS AND IMPLEMENTATION

The purpose of this study is the implementation of steganography using Least Significant Bit methods. This section seeks to present the result of the study by analyzing and interpreting the data collected, methods and techniques used in conducting the study. Different approaches were put in place in order to have better and deeper representation of the results by implementing LSB technique for hiding data in audio objects. For the implementation of the systems, the above stated scenario was considered and implemented using Java Programming Language. In all, testing was done through the normal viewing using the human senses to distinguish the original and the resultant object. The implementation of the Secure Transit Data System (STDS) was implemented in two folds, that is, encoding and decoding Audio Steganography presented using the LSB processes.

Audio Steganography Implementation

Audio signals have a characteristic redundancy and unpredictable nature that make them ideal to be used as a cover to hide secret information. Like image, audio files may be modified in such a way that it can contain some secret information using the LSB. In the case of audio or sound files, each sampling point of the file is substituted with the least significant bit. With this approach, large amount of data can easily be encoded onto the audio file. The redundancy of bits that exist in the binary coding of numbers, and alphabets forms the basis of this approach.

Looking at the binary code of numbers from 0 to 9, and from A (a) to P (p) for both casing, it can be observed that, these characters are only different in their respective last 4 bits. Thus, their first 4 bit are similar, thereby implying that, any number or alphabet can easily be represented by the last 4 bits and adding either 0 or 1 at its first position. To differentiate whether the character is

number, uppercase alphabet or lowercase alphabet control symbols are used which is of the same type as that of number or alphabet.

For special symbols like !, “ , # , $, %, & , (, ,) , *, + , ‘, - , . , / is also observed and these special symbols can also be embedded in WAV file. When embedding the textual information in any audio file, first the audio signal is converted into bits. Then the message to be embedded is encrypted and converted. By applying LSB algorithm, the message is embedded into 16 bits or 8 bits audio sample.

Audio Steganography Encoding Process

The underlying technology the encoding process is the LSB. In summary, the encoding algorithm takes in a text to be embedded as an input, convert the text into a 5-bit code by checking the redundancy in the binary coding structure of the characters involved. The next is to the read the audio file as the cover object. The selected audio file or the cover object is then used to hide the converted 5-bit code of text using the proposed methodology. This process is repeated until the entire message is embedded successfully into the audio file.

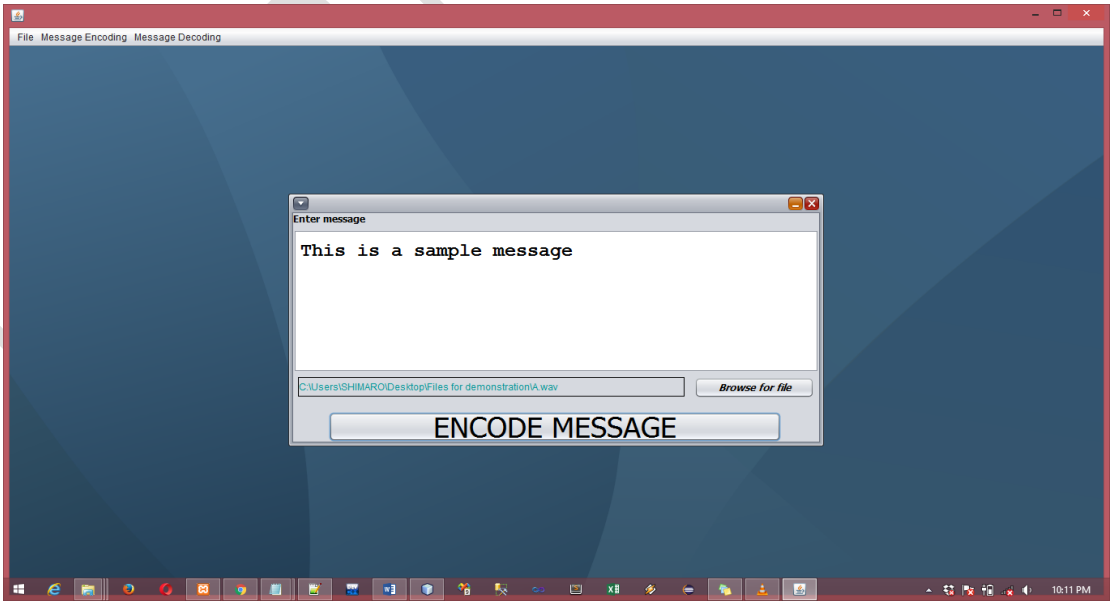

Figure 5: Audio Embedding user interface

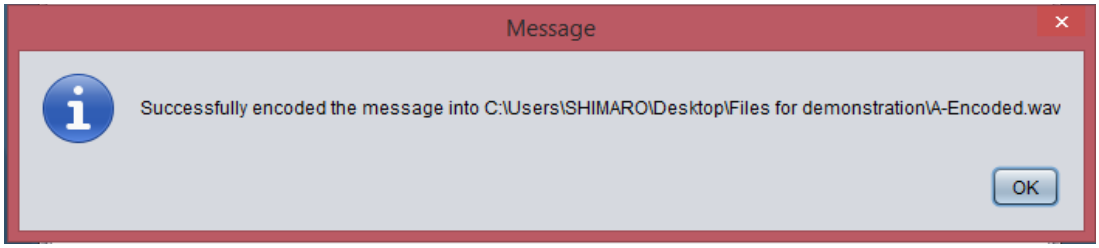

Figure 6:Audio Encoding Status dialog

Audio Steganography Decoding Process

The decoding process is the reverse of the encoding process described above. The stego-object thus the cover audio that has the encoded message is read as an input. The message embedded is then extracted by reading the control symbols in samples using LSB. All the selected samples are stored with their LSB positions. The resultant array is then subjected to some minimal operation of division using the number of rows and columns leading to the final extraction of the messages.

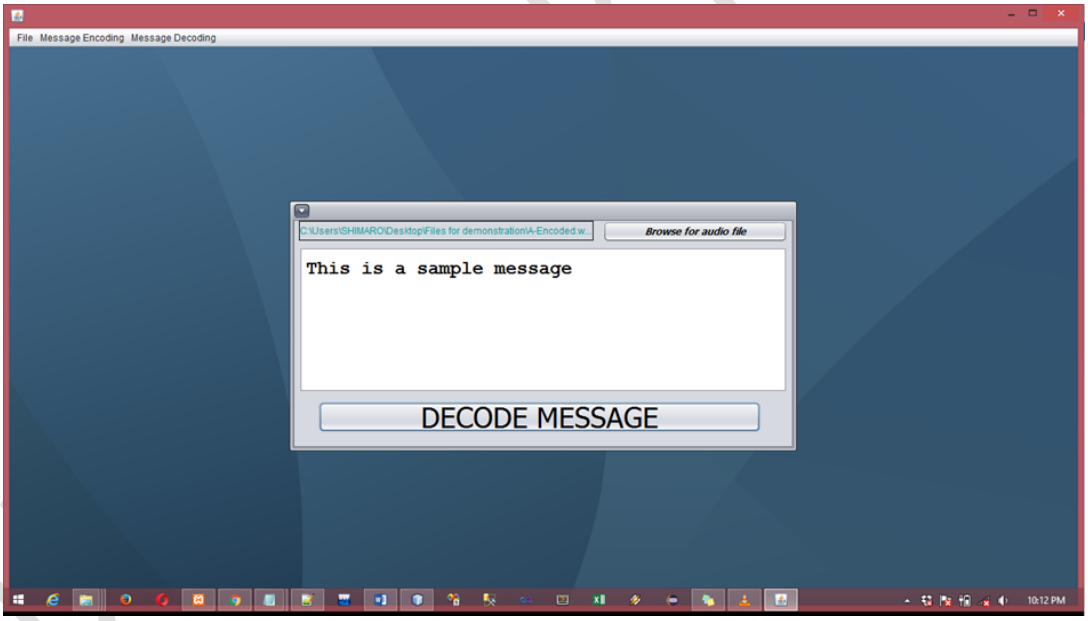

Figure 7:Audio Decoding User Interface

Experimental Result

After successful implementation of the embedding and the decoding process, a wave form was created from the two samples files. It can be observed from the figure below that, the encoded

and the original files have the same wave forms. This shows that the proposed technology does not distort the audio file, thereby not attracting attention.

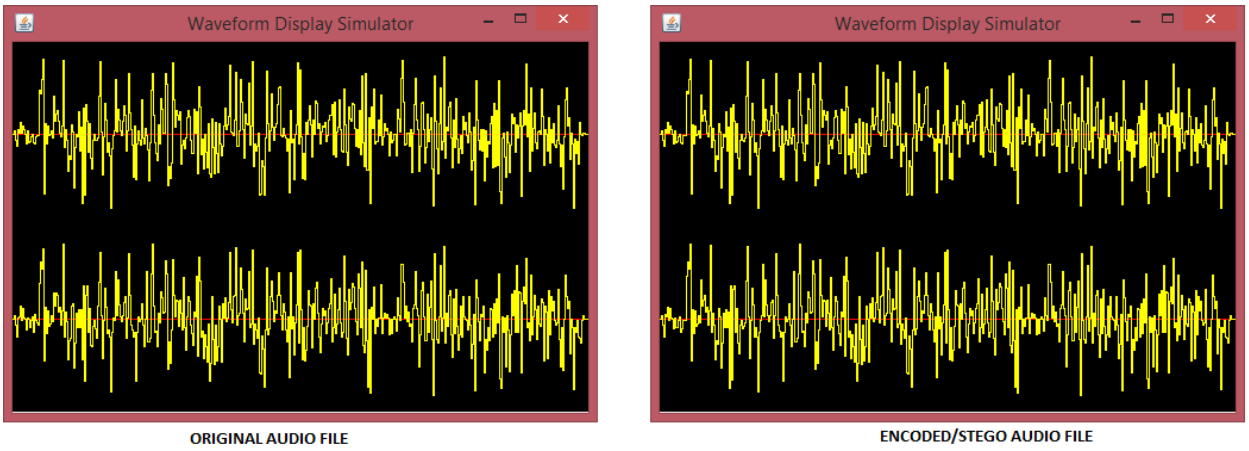

Figure 8:Audio file sample A waveform

SAMPLE B

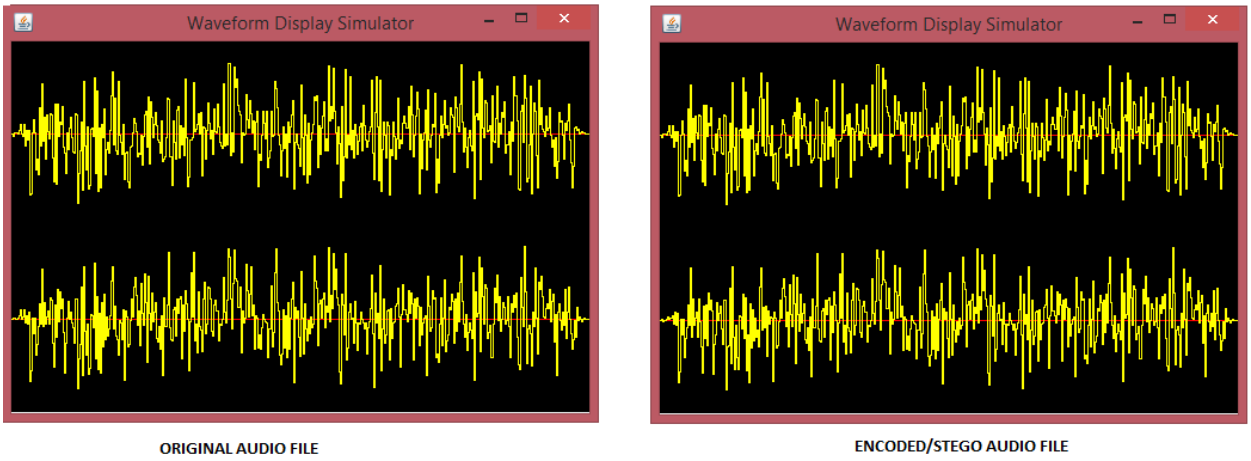

Figure 9:Audio file sample B waveform

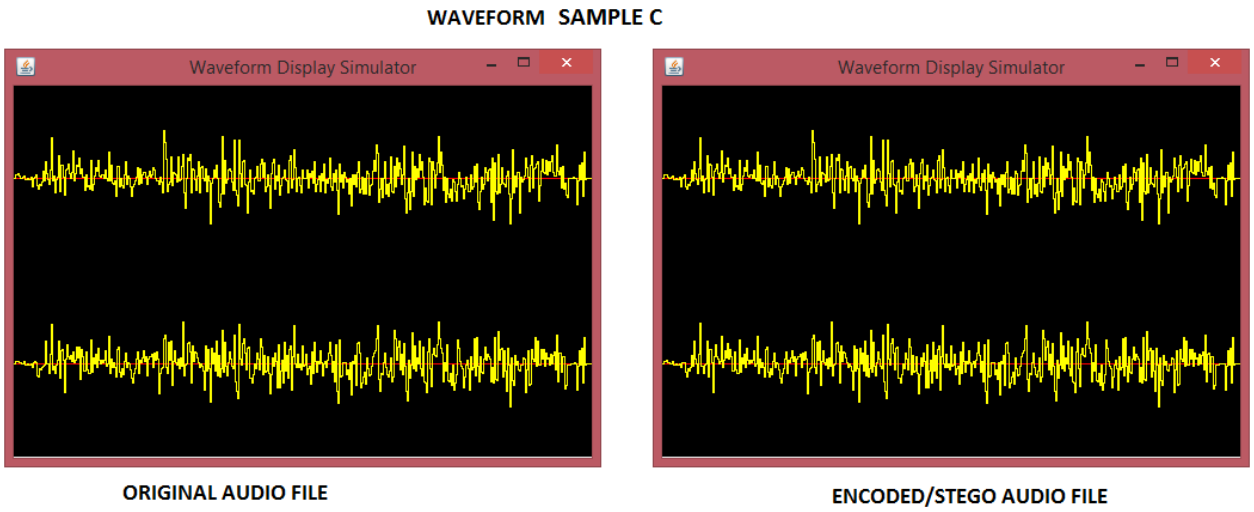

Figure 10:Audio file sample C waveform

V. CONCLUSION

This project was meant to secure data in transit using audio steganography. Steganography is one of the ways by which data in transit can be secured without attracting unnecessary attention from intruders. The algorithm used in this research proves to be one of the simplest ways of securing data using audio steganography. The methods employ the LSB approaches by using audio files as the stego object for the implementation based in Java Programming Language. The experimental results also proved to be one of the best methods of implementing steganography. The accuracy of the stego objects as compared to the original objects is of high quality and similarity.

Data is the backbone and the lifeline of every organization. Data security has become one of the major ways by which organization are committing their resources to. Therefore, there is the need to implement cheaper but robust and secure methods of securing data. The knowledge of this technology is still new to most practitioners in the area of Information Security.

In the future, more work should be carried out by technology and science-based institutions into the area of information hiding. It is the hope of the researcher that, future works can take two or more objects as input and embed the secret messages in them. Other quality metrics can also be used to analyze the performance of the proposed algorithms.

Finally, future researchers should try to include into their work how best this technology can be used in mobile phones and how best protocol steganography can be used to secure data on the Internet.

VI. REFERENCES

[1] M. A. Qadir and I. Ahmad, "Digital text watermarking: secure content delivery and data hiding in digital documents," in *Proceedings 39th Annual 2005 International Carnahan Conference on Security Technology*, 2005.

[2] J. M. a. S. Mangal, "An Overview of Image Steganography using LSB Technique," *IJCA Proceedings on National Conference on Advances in Computer Science and Applications (NCACSA 2012),* vol. 3, pp. 10-13, 2012.

[3] M. Ramkumar and A. N. Akansu, "Some design issues for robust data hiding systems," in *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers (Cat. No.CH37020)*, 1999.

[4] N. S. M. S. B. Ayush Singhal, "An Advanced Approach for Implementation of Audio Steganography," *International Journal For Science, Technology and Engineering,* vol. 1, no. 12, pp. 66-71, 2015.

[5] R. Tanwar and M. Bisla, "Audio Steganography," in *2014 International Conference on Reliability Optimization and Information Technology (ICROIT)*, 2014.

[6] R. S. Kazem Qazanfari, "A new steganography method which preserves histogram: Generalization of LSB++,," *Information Sciences,* vol. 277, pp. 90-101, 2014.

[7] Y. V. M. Baritha Begum, "LSB Based Steganography based on Text Compression," *Procedia Engineering ,* vol. 30, pp. 703-712, 2012.

[8] S. C. a. A. Jadhav, "Steganography an Art of Hiding Data," *International Journal on Computer Science and Engineering(IJCSE),* 2009.

[9] J. S. Johnson N.F., Steganalysis of Images Created Using Current Steganography Software. In: Aucsmith D. (eds) Information Hiding.IH 1998. Lecture Notes in Computer Science, vol 1525, Berlin: Springer, 1998.

[10] N. F. J. a. S. Jajodia, "Exploring steganography: Seeing the unseen,," *Computer,* vol. 31, no. 2, pp. 26-34, 1998.

[11] P. H. N. Provos, "Detecting Steganographic Content on the Internet," CITI Technical Report 01-11, Michigan, 2021.

[12] R. Anderson, "Analysis of LSB Based Image Steganography Techniques," *IEEE,* pp. 474-481, 1998.

[13] C. Cachin, "An Information-Theorectical Model for Steganography," in *In Proceeding of 2nd Information Hiding Workshop*, 1998.

[14] F. P. S. Katzenbeisser, "Defining security in Steganographic Systems," in *Proc. SPIE 4675, Security and Watermarking of Multimedia Contents IV,*, 2002.

[15] R. J. A. a. M. G. K. F. A. P. Petitcolas, "Information hiding-a survey," in *In Proceedings of the IEEE*, 1999.

[16] C. H. Gary C. Kessler, "Chapter 2- An Overview of Steganography," in *Advances in Computers*, vol. 83, Marvin V. Zelkowitz, Ed., Elsevier, 2011, pp. 51-107.

[17] J. F. T. H. Miroslav Goljan, "New blind steganalysis and its implications," in *Proceedings Volume 6072, Security, Steganography, and Watermarking of Multimedia Contents VIII;*, San Jose, 2006.

[18] M. Gasser, Building A secure Computer Systems, USA: Van Nostrand Reinhold Co, 1998.

[19] S. L. N. F. M. a. L. O. J. T. Brassil, "Electronic marking and identification techniques to discourage document copying,," *IEEE Journal on Selected Areas in Communications,,* vol. 13, no. 8, pp. 1495-1504, 1995.

[20] B. S. a. R. Shanthakumari, "Efficient Adaptive Steganography for Color Images Based on LSBMR Algorithm," *ICTACT Journal on Image and Video Processing,* vol. 02, no. 03, pp. 387-392, 2012.

[21] S. K. Arora, "Audio Steganography : The art of hiding secrets within earshot(part 2 of 2)," 17 June 2018. [Online]. Available: https://sumit-arora.medium.com/audio-steganography-the-art-of-hiding-secrets-within-earshot-part-2-of-2-c76b1be719b3. [Accessed 1 August 2021].