Multi Level - Secret Data Embedding in Audio Steganography

S.Saratha^{1*}, Dr.V.Murugan², Dr.P.Arockia Jansi Rani³

¹Research Scholar, Department of Computer Science and Engineering, Manonmaniam Sundaranar University Abishekapatti, Tirunelveli, India – 627 012

²Assistant Professor, Department of Computer Science, Government College of Arts and Science, Kadayanallur, India
³Associate Professor, Department of Computer Science and Engineering, Manonmaniam Sundaranar University, Abishekapatti,
Tirunelyeli, India

*Corresponding Author: keethu.saratha@gmail.com

Abstract: This paper presents a novel methodology for the secure and efficient concealment of data in audio signals. The approach combines Spectrogram Pixel Value Modification (SPVM) with RSA encryption and compression using the Discrete Wavelet Transform (DWT). The ultimate goal of this proposed method is to augment data security and capacity while preserving the quality of the audio signal.

The proposed methodology initiates with RSA encryption, where the message to be obscured is encrypted using the recipient's public key, ensuring confidentiality during the data embedding process. The data is then compressed using DWT, which decomposes the signal into wavelet coefficients that provide both frequency and time localization.

To conceal the compressed and encrypted data, the wavelet coefficients are transformed into a spectrogram representation using an appropriate Fourier transform. The SPVM method is then employed to subtly modify the pixel values of the spectrogram by incorporating binary values derived from the compressed and encrypted data. This ensures seamless integration of hidden information without compromising the audio signal's integrity.

To retrieve the concealed message, the process is reversed. The SPVM method is reverted to extract the modified pixel values from the spectrogram. The extracted pixel values are converted back into wavelet coefficients, and the inverse DWT is applied to obtain the compressed and encrypted data. Finally, the recipient's private key is used to decrypt the data, revealing the original message.

The experimental results demonstrate the practicality and effectiveness of the proposed approach, showcasing its potential in secure data hiding applications, audio watermarking, and confidential communication. The combination of SPVM, RSA encryption, and DWT compression provides a comprehensive solution for robust and secure data embedding while ensuring high audio quality.

Key terms: Cryptography, Steganography, imperceptible format, RSA, DWT, Spectrogram Pixel Value Modification (SPVM), LSBs, MSE, PSNR.

1. INTRODUCTION

In the modern communication system, network security plays a crucial part in data transmission. To prevent unauthorized access to data, it is necessary to secure the data confidentially. Cryptography and steganography support secure data transmission by ciphering or concealing sensitive data from human interception. Cryptography alters the data (audio, video, images) structure for securing the message thereby it can be easily attackable. In Steganography, the message is concealed without data structure alteration and is also called invisible communication.

Steganography is a technique that conceals the secret information behind any data (image, audio, text) and converts the data into an imperceptible format. It is distinct from cryptography by embedding the data with a prevalent carrier and thereby protects its security and privacy. The Steganographic Model is depicted in Figure 1. The basic Steganographic Model is divided into two blocks such as sender and receiver side. On the sender side, the cover signal (C) and Secret Data (D) are fed into the Steganographic encoder as input which encodes the secret data into the cover signal with a key. Similar to the cover signal a stego Object is created without visible changes. Then at the receiver side, the receiver send private

key to the steganographic decoder where it decode the stego object. At last secret data is decoded from the receiver end [5]. There are different Steganographic techniques such as Text steganography, Image sego, Audio Steganography, video Steganography, and Network Steganography. By combing cryptography and Steganography [1], the author tried to combine the benefits of two techniques to secure sensitive data. RSA cryptography and audio-based Steganography act as a sequential layer which is concealing

the encrypted data in LSB. But it doesn't have a different symmetric algorithm for the crypto layer which affects the efficiency of the method. To overcome the breaching attack over the image database [2], Steganography and Secret Share Cryptography (SSC) are combined to upgrade the security level where medical images are taken as stego images. There didn't focus on the combination of text, data, and image steganography with multi-domain transform.

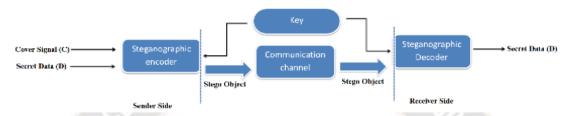


Figure 1: Basic Steganographic model

To hide a large amount of text inside the audio [3][17][18], the author used a 16bit WAV and 8-bit WAV audio file using the LSB algorithm. But the system won't support 32bit WAV audio files. Without affecting the normal encoding process of Huffman [4] data is embedded with the audio signal where Advanced Audio Coding (ACC) is used. But this method causes a storage problem due to ACC files which get increase their size during the embedding process. The LSB-based image steganography is utilized to hide the data inside the image using a symmetric key [5]. They chose bits that provide minimum resolution among the stego image and original image. But this method suffers from high computational time. The LSB and XOR combined and conceal data into an image using DCT [6]. Compare to LSB the DCT provides good performance but it suffers from low efficiency. Discrete Wavelet Transform (DWT) [10][11] is used to hide the message into the cover image which provides a better imperceptibility performance. To achieve high order security the author used audio steganography which is a good carrier medium for steganography technique. Several iterations take place to embed the secret text data into the MP3 file. This method utilized only 15MB audio files as a cover signal which can't be utilized for huge data transactions. The overall view of audio steganography [8][19] is analyzed with its algorithm where the author put forth the existing audio steganography techniques have less security, robustness, and low capacity of hiding secret data which will be improved by combining different methods of audio steganography. The author [15] uses the Goldbach code algorithm to secure the text data inside the images using the LSB technique. But it is prone to numerous attacks like noise disturbances. We analyze numerous research

papers on existing Steganographic methods thereby examined their pros and cons which will be utilized to implement our novel audio Steganographic method. Based on the above analysis, it is clear that audio steganography can hold more information than other Steganographic techniques. Because audio Steganographic poses high redundancy, efficiency, low-latency, and high transmission rate which makes it to uses as the cover signal. Due to this reason, we took the cover signal as audio and implement it in a proposed novel audio Steganography method that conceals secret data into an audio cover signal. By the motivation [8], we understood that to obtain a high perpetual embedding need to combine different techniques and algorithm sequentially. The main contribution of this paper is to develop a versatile method for audio steganography with high embedding capacity and fully recover hidden data than the existing method [15].To achieve the objective the proposed system is implemented with nearly five algorithms as Altered RSA, Altered Genetic algorithm, Altered DWT, Altered LWT with LSBs Substitution [13]. There are nearly six phases for concealing and retrieving the secret data as Encryption, Compression, Embedding, Extraction, Decompression, and Decryption. The rest of the paper is organized as follows: Section 2, Methodology of the proposed method, Section 3, Parametric Analysis, Section 4, and Conclusion.

2. METHODOLOGY

Audio Steganography has proven to be a competitive model for securing secret data in recent research. Typically, the proposed audio steganography contains two parts: transmitter and receiver side which is depicted in Figure 2.

On the transmitter side, three consecutive steps are followed to embed the secret data (image, audio, text) into cover

audio signals.

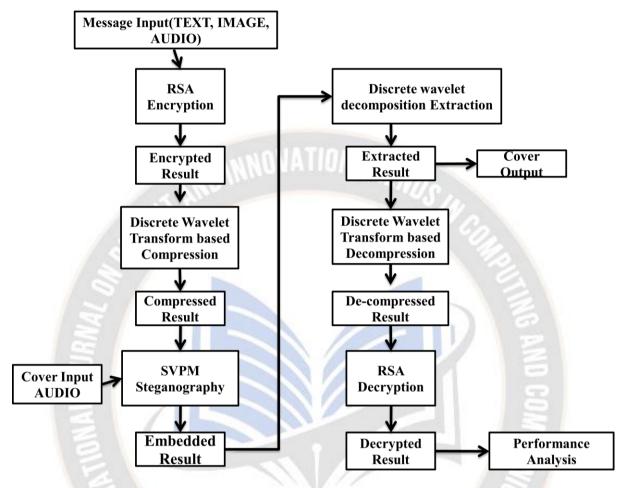


Figure 2: Architecture of Proposed Audio Steganographic Method

The proposed methodology aims to achieve secure and efficient data concealment in audio signals by utilizing a combination of three fundamental techniques: Spectrogram Pixel Value Modification (SPVM), RSA encryption, and Discrete Wavelet Transform (DWT) compression.

The procedure commences with RSA encryption, whereby the information to be hidden is encrypted utilizing the recipient's public key. This guarantees that only the intended recipient, with the corresponding private key, can later decrypt and access the concealed information. This step ensures confidentiality during the data embedding process.

Following the RSA encryption step, the encrypted data undergoes compression using the Discrete Wavelet Transform (DWT). This potent technique disintegrates the signal into various wavelet coefficients, capturing both frequency and time localization. The wavelet coefficients

constitute the compressed and encrypted data, offering an efficient representation with reduced redundancy.

To incorporate the compressed and encrypted data into the audio signal, the wavelet coefficients are transformed into a spectrogram representation using a suitable Fourier transform. The SPVM method is then employed, carefully modifying the pixel values of the spectrogram based on binary values derived from the compressed and encrypted data. This delicate modification ensures that the hidden information is seamlessly integrated into the audio signal without causing noticeable degradation.

The extraction and decoding process allows the recipient to retrieve the hidden message. By reversing the SPVM method, the modified pixel values are extracted from the spectrogram. These values are subsequently converted back into wavelet coefficients, and the inverse DWT is applied to reconstruct the compressed and encrypted data. Finally,

decryption using the recipient's private key reveals the original message, completing the data extraction process.

Various experiments are conducted, employing different audio signals and messages for embedding, to evaluate the performance of the proposed approach. The data hiding capacity, robustness against attacks and noise, and computational complexity are meticulously measured and analyzed. The proposed approach can be evaluated by comparing it with existing techniques to identify its strengths and limitations.

A comprehensive security analysis assesses the resistance of the embedded data against unauthorized extraction attempts and different attacks. The practical application of the proposed methodology is discussed, emphasizing its potential usage in secure communication, audio watermarking, and data authentication. The algorithm utilized for encryption, compression, embedding, Extraction, Decompression, and Decryption are explained below,

Step 1: RSA Encryption Inputs: Message M to be hidden, Recipient's public key (e, n) Output: Encrypted data C

- 1. Generate the message M to be hidden.
- 2. Obtain the recipient's public key (e, n), where 'e' is the encryption exponent and 'n' is the modulus.
- 3. Convert the message M into an integer representation m, suitable for encryption.
- 4. Apply RSA encryption using the public key (e, n): C = m^e mod n
- 5. The resulting C is the encrypted data to be embedded.

The RSA's security is built on the mathematical difficulty of factoring large composite numbers into their prime factors. The use of large prime numbers in the key generation step makes the factorization of 'n' computationally infeasible with current technology. As a result, even with knowledge of the public key (e, n), it is extremely challenging to determine the private key (d) required for decryption without knowing the prime factors 'p' and 'q'. This property ensures the confidentiality and integrity of encrypted data in RSA-based communication and data protection systems.

Step 2: Compression using DWT Input: Encrypted data C Output: Compressed and encrypted wavelet coefficients Coeff

- 1. Choose a suitable DWT algorithm (e.g., Haar, Daubechies, or Symlet).
- 2. Convert the encrypted data C into a binary representation (C_bin).

- 3. Apply the selected DWT to C_bin, decomposing it into wavelet coefficients: Coeff = DWT(C_bin)
- 4. The resulting Coeff represents the compressed and encrypted data. Top of Form

The process receives encrypted data 'C' as input, which denotes the encrypted information that requires secure storage or transmission. For data compression without compromising security, the initial step is to convert the encrypted data 'C' into binary representation 'C_bin'. The binary data 'C_bin' then undergoes the Discrete Wavelet Transform using a selected DWT algorithm, such as Haar, Daubechies, or Symlet.

The DWT algorithm dissects the binary data 'C_bin' into wavelet coefficients 'Coeff'. These coefficients capture both the low-frequency components (approximation coefficients) and the high-frequency components (detail coefficients) of the encrypted data. Consequently, 'C' information gets represented more efficiently with fewer coefficients.

The resulting output 'Coeff' signifies the compressed and encrypted data, which is ready for storage or transmission. The reversible nature of DWT allows for easy reconstruction of the initial binary data 'C_bin' from 'Coeff'. This leads to decryption and recovery of the initial encrypted data 'C' when necessary. The combination of DWT-based compression and encryption offers a powerful approach for secure data storage and communication, where data size is reduced without compromising information security.

Step 3: SPVM Embedding Inputs: Compressed and encrypted wavelet coefficients Coeff, Audio signal S (spectrogram representation) Output: Modified audio signal S mod with hidden data

- 1. Convert the wavelet coefficients Coeff into a spectrogram representation using Fast Fourier Transform: Spectrogram = InverseDWT(Coeff)
- 2. Divide the spectrogram into smaller non-overlapping blocks (e.g., 8x8 pixels) for SPVM embedding.
- 3. For each block B in the spectrogram:
- Extract the corresponding binary values (binary_B) from Coeff.
- Convert binary_B into a decimal value dec_B.
- Get the pixel values P of block B from the audio signal
- Modify the pixel values P subtly based on dec_B using SPVM.
- Update the modified pixel values in the audio signal S mod.

4. Repeat Step 3 for all blocks in the spectrogram, embedding the hidden data throughout the audio signal.

Following the conversion of wavelet coefficients back into the spectrogram representation, the third step of the SPVM Embedding process involves dividing the spectrogram into smaller blocks. For each block, binary values corresponding to the wavelet coefficients are extracted and subsequently converted into decimal form. The pixel values of the block in the original audio signal are then subtly modified based on the decimal value, utilizing the SPVM method. This procedure is repeated for all blocks in the spectrogram, resulting in the modified audio signal (S_mod) with the hidden data seamlessly embedded. The SPVM method ensures that the modifications are imperceptible to human ears, thereby preserving audio quality while securely concealing the data within the audio signal.

Step 4: Extraction and Decoding Input: Modified audio signal S_mod with hidden data, Recipient's private key (d, n) Output: Decrypted message M_extracted

- 1. For each block B in the modified audio signal S_mod:
- Extract the modified pixel values P mod from block B.
- ❖ Convert P_mod into binary values binary_B_mod.
- ❖ Convert binary_B_mod into a decimal value dec B mod.
- * Replace the decimal value dec_B_mod in Coeff.
- 2. Apply the inverse Fourier transform to Coeff to obtain the updated encrypted data C_extracted.
- 3. Apply RSA decryption using the recipient's private key (d, n): M_extracted = C_extracted^d mod n
- Convert the decrypted integer M_extracted back into the original message M.

The objective of Algorithm 5 is to retrieve the concealed data from the audio signal that was previously embedded using the Spectrogram Pixel Value Modification (SPVM) method. To achieve this, the lifting wavelet decomposition is executed to segregate the audio signal's approximation and detail coefficients. The specific sections wherein the binary text is embedded are decoded from the detail coefficients to extract the binary text. This binary text is then extracted and transformed into decimal format, which results in fully recovered hidden data. The final output, ext_data, contains the recovered hidden information in decimal format, representing the original data securely concealed within the audio signal using the SPVM method. The reversibility of the SPVM embedding technique is demonstrated by this extraction procedure, as the original

data can be successfully retrieved from the altered audio signal.

3. PARAMETRIC ANALYSIS

The proposed novel audio steganography is implemented in the MATLAB platform the GUI of the output is depicted in Figure 3. This section displays the output of different phases (Encryption, Compression, Embedding, Extraction, Decompression, and Decryption). To evaluate the performance of the proposed audio steganography the performance metrics are calculated which is explained below,

a) Mean Square Error (MSE),

MSE is defined as the means square of the ratio between the actual and estimated values where the equation is represented below,

$$MSE = \frac{1}{N} * \sum_{i=1}^{N} \left[(actual - forecast)^2 \right]$$
 (1)

N =number of data,

actual= original of observed value,

forecast=value from the regression

b) Peak Signal to Noise Ratio (PSNR),

PSNR is a ratio of maximum value and MSE value that affects the fidelity of the representation. The equation (2) is represented below,

$$PSNR = 20(\frac{MAX_f}{\sqrt{MSE}}) \tag{2}$$

Where MSE= Mean Square Error,

 $MAX_f = \text{maximum value}$

c) Root MSE (RMSE),

RMSE is the square root of MSE which is represented in equation (3),

$$RMSE = \sqrt{MSE} \tag{3}$$

Where MSE= Mean Square Error

d) Normalized Absolute Error (NAE),

It measures the difference between the processed data and original data. If the result value is near zero then there is a similarity with the original data or else it has a poor quality prediction.

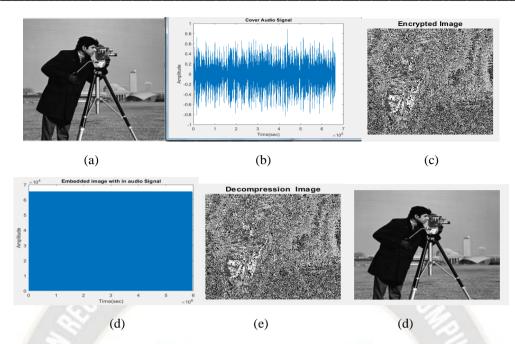


Figure 3: Audio Steganography by using image (a)Input Image (b)Cover Audio signal (c)Encrypted image (d) Embedded image with the audio signal (e) Decompression image (f) Exctracted image.

From Figure 3, the secret data is taken as audio signal (SA1.wav) (a) which is concealed in the cover audio signal (b). The RSA encrypted algorithm encrypts the audio signal (c). Then DWT compressed the encrypted data and produced the compressed data (d). SPVM embed the audio within the audio signal (e). Then the decompressed audio signal is generated (f). Finally, the secret audio data is decrypted by the receiver based on the private key.

The equation (4) is represented below,

$$NAE = \frac{\sum_{m=1}^{M} \square \sum_{n=1}^{N} \square [X(m,n) - Y(m,n)]}{\sum_{m=1}^{M} \square \sum_{n=1}^{N} \square |X(m,n)|}$$
(4)

Where X -original data

Y-processed data

e) Laplacian Mean Square Error (LMSE)

LMSE is a ratio of the square root of differences between the processed and original data to the original data. The equation (5) is represented below,

$$LMSE = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} [O(m,n) - O(Y(m,n))]^{2}}{\sum_{m=1}^{M} \sum_{n=1}^{N} [O(X(m,n))]^{2}} (5)$$

Where X -original data

Y-processed data

In Figure 3, the secret data is chosen as an image (cameramen. tiff) where the RGB image is converted into Grey-scale image (a), and then RSA encryption is proceeded over the input image and constructed a encrypted image (c). DWT algorithm compressed the data and then embedding takes place through Spectrum Pixel Value Modification where the input image is encoded inside the cover audio signal (d). Then based on the receiver private key extraction happen which split the cover signal and secret image in it. Then the decompressed image (e) is constructed and then RSA decrypts the decompressed image and produces a decrypted image (f).

Secret File	PSNR	LMSE	RMSE	GAE	Average Difference
SA1.wav	70.1580	1.0605	0.0278	1.0001	0.0017
SA2.wav	66.5141	0.9263	0.1205	1.0001	0.0029
Lena.Tiff	71.2169	1.5650e-06	0	1.0001	0.0016
Cameraman.Tiff	79.3489	0.0575	0.0245	1.0000	0.0015
Textinput.txt	49.8765	1.6026e-05	0.1554	1.0179	0.1083
Txtinput.txt	53.668	5.1872e-05	0.5290	1.0021	0.0560

Table 1: File size of input data in each step

Table 1 represents the size of the different files used as secret data in proposed audio steganography where file size is varied for each step. For example, consider the sender send secret data as an image (Lena. tiff) whose file size is

4.19 MB. The file size is compressed to 2.09 MB. After the encryption, the file size is changed to 4.19 MB. The cover audio signal size is 42.35 MB which is embedded with the secret input image and size increased to 381.9MB.

METHODS	AUDIO		IMAGE		TEXT	
PAPER	AVERAGE	AVERAGE	AVERAGE	AVERAGE	AVERAGE	AVERAGE
COMPARISON PROPOSED	PSNR	RMSE	PSNR	RMSE	PSNR	RMSE
METHOD	69.1680	0.0178	69.3059	0	67.9535	0.2554
[16]	-	ONN	41.73	0.0095	-	-
[17]	7.0	Mr.		"LIUS III	64.24	2.03
[18]	(SE)		-	- "6	16	-
[19]	47.1	1.3	-	-		-

The recovery image quality is similar to the original secret image in our proposed model hereby enhancing image quality. The performance of the proposed method is analysed and compared with the existing methods based on metric parameters (Average Peak Signal to Noise Ratio(APSNR), Average Root Mean Square Error(ARMS)) for different secret data(Text, Image and Audio) Which is tabulated in the above table. The result states that the proposed method provides a higher PSNR value and lower RMSE than the existing methods[16 = 19]. The average PSNR value is calculated for secret data it represents that the proposed method extracts data without distortion from original embedded data.

4. CONCLUSION

We have proposed a novel audio steganography method based on RSA, DWT, Spectrum Pixel Value Modification algorithms. The proposed method provides fully recovered hidden data and provides high perceptual quality in embedding capacity which is nearly 25% for cover signal. There are three consecutive algorithms RSA, DWT, SPVM are utilized to secure the secrete data in an imperceptible position. The PSNR value of secret data is evaluated and represents the quality of data which higher than the existing steganography methods. MSSIM value is higher for secret data and ensures the efficiency of the proposed method. The Average PSNR and Average RMSE value is calculated for the proposed method along with existing methods which represent that the proposed method provides a higher recovery quality, imperceptibility of hiding data, and better

robustness, thereby successfully achieved the objective of research analysis of the proposed method.

REFERENCE

- 1. Al-Juaid, Nouf, and Adnan Gutub, "Combining RSA and audio steganography on personal computers for enhancing security", published in *SN Applied Sciences* 1.8:830, (2019).
- A. Sivasankari, S. Krishnaveni, "Optimal Wavelet Coefficients Based Steganography for Image Security with Secret Sharing Cryptography Model", published in Cybersecurity and Secure Information Systems, (2019).
- K.P. Adhiya, Swati A. Patil, "Hiding Text in Audio Using LSB Based Steganography", published in Information and Knowledge Management, ISSN 2224-5758, (2012).
- Jie Zhu, Rang-Ding Wang, Juan Li and Di-Qun Yan, "A Huffman Coding Section-based Steganography for AAC Audio", published in Information Technology Journal, Volume: 10 | Issue: 10, (2011).
- Dr.Amarendra K, Venkata Naresh Mandhala, B.Chetangupta, G.Geetha Sudheshna, V.Venkata Anusha, "Image Steganography Using LSB", published in International journal of scientific & technology research volume 8, issue 12, (2019).
- Gayatri G. Bobade, A. G. Patil, "Testing of Image Steganography with use of LSB and DCT Techniques", published in International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-10, (2019).

- _____
- Mohammed J. Alhaddad, Monagi H. Alkinani , Mohammed Salem Atoum and AlaaAbdulsalmAlarood, "Evolutionary Detection Accuracy of Secret Data in Audio Steganography for Securing 5G-Enabled Internet of Things", published in Symmetry, (2020).
- 8. HrishikeshDutta, Rohan Kumar Das, Sukumar Nandi and S. R. MahadevaPrasanna, "An Overview of Digital Audio Steganography", published in IETE Technical review, (2019).
- Sazeen T. Abdulrazzaq, Mohammed M. Siddeq, Marcos A. Rodrigues, "A Novel Steganography Approach for Audio Files", published in SN Computer Science, (2020).
- 10. Manal K. Oudah, Aqeela N. Abed, Rula S. Khudhair, Saad M. Kaleefah, "Improvement of Image Steganography Using Discrete Wavelet Transform", published in Engineering and Technology Journal, Vol. 38, (2020).
- 11. Guru Vimal Kumar Murugan&RagupathyUthandipalayamSubramaniyam, "Performance analysis of image steganography using wavelet transform for safe and secured transaction", published in Multimedia Tools and Applications, (2019).
- 12. Ying Li, Brian K. Via, Yaoxiang Li, "Lifting wavelet transform for Vis-NIR spectral data optimization to predict wood density", published in SpectrochimicaActa Part A: Molecular and Biomolecular Spectroscopy, (2020).
- 13. SanchitaGesu, ShivamVasudeva, Snehil Bhatt, Santhosh B,"A Novel Technique for Secure Data Transmission using Audio/Video Files", published in International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181,(2015).
- 14. S. Goel, S. Gupta, and N. Kaushik, "Image Steganography -- Least Significant Bit with Multiple Progressions," published in Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA), (2015).
- 15. Jan Carlo T. Arroyo, AllemarJhone P. Delima, "LSB Image Steganography with Data Compression Technique Using Goldbach G0 Code Algorithm", published in International Journal of Emerging Trends in Engineering Research, Volume 8,(2020).
- 16. Said E. El-Khamy, Noha O. Korany & marwa H. El-Sherif, "A Security enchanced robust audio Stedio Steganography algorithm for image hiding using sample comparison in discrete wavelet transform domain and RSA encryption", published in Multimedia tools and Applications, (2016).

- 17. Hinal Somani, K. Madhu, "Robust and High Capacity Audio Steganography Modified Dual Randomness LSB Method", published in International Journal of Advance Research and Innovative Ideas in Education, (2016).
- 18. M. Parthasarathi and Shreekala, "Secured Data Hiding in Audio Files using Audio Steganography Algorithm", published in International journal of Pure and Applied Mathematices, (2017).
- 19. Ahamed Hussian Ali, Loay Edwar George, A. A. Zaidan % Mohd Rosmadi Mokhtar, "High capacity transparent secure audio steganography model based on fractal coding and chaotic map in temporal domain", published in multimedia tools and Applications, (2018).

