ARTICLE

SECRET DATA HIDING WITH IMAGES USING DATA COMPRESSION AND EMBEDDING ALGORITHM

Suresh G.[1]* and K. A. Parthasarathy[2]

[1]*Department of Computer Science and Engineering, St. Peter's University, Chennai, Tamil Nadu, INDIA*

[2]*Akheyaa College of Engineering, Puludhivakkam, Kancheepuram, Tamil Nadu, INDIA*

ABSTRACT

Objective: *The objective of this paper is to embed secret data with enhanced and compressed embedding scheme to maintain tight security and efficient data transmission in un-trusted networks. There are several literature work done in steganography to hide secret from legitimate user with many kinds of embedding scheme namely: Least Significant Bit, Discrete Cosine Transformation, Lempel-Ziv-Welch (LZW), Integer Haar Wavelet Transform(IHWT), Bacterial Foraging Optimization (BFO), Multiple encryption (ME), and MKB etc. However, these approaches failed to maintained tight privacy and data compression during efficient data transmission in un-trusted networks. These approaches are unable to fulfill current requirements completely. Neither they solve compression issues nor secure data embedding and are unable to compress the original size of data during embedding. It also does not maintain efficient encryption scheme to encrypt and decrypt the data.* ***Method:*** *To alleviate these issues, proposed framework developed a data compression and embedding algorithm. This approach performs the embedding with two cover file namely master file and output file with secret data file. Hence, it asks to source to reduce the data size based on requirement. After selection of two cover, once file content swaps with another cover file to maintain the privacy.* ***Finding:*** *Its supports various formats of secret data file. It provides complete information about embedding to destination like name, formats, compression level in percentage.* ***Improvement:*** *Based on experimental results, it realized that proposed systems performs well compared to existing approaches in terms of root mean squared error (RMSE), peak signal noise ratio (PSNR), and structural similarity index measurement (SSIM). It reduces RMSE 0.17 and increase PSNR 3.54db and SSI 0.30.*

KEY WORDS

Data hiding, data cove file, data compression and embedding algorithm, root mean squared error, similarity index measurement, and peak signal noise

Received: 16 Aug 2016
Accepted: 2 Sept 2016
Published: 12 Sept 2016

***Corresponding Author**
Email:
sureshspu.phd@gmail.com

INTRODUCTION

Nowadays, data embedding scheme is popular to share the secret data from one host to another hosts. It keeps secrecy of original data and easily transfers to concerned hosts. This approach is divided into two parts namely: Data hiding which covers the secret message or data with cover file to maintain the privacy. Cryptography approach which assists to data hiding approach to encode and decode the original data. Generally, cryptography approach is used to provide additional privacy for original data. The main objective of this approach is to transmit the original file or message to desired destination with tight privacy and reduced data size for efficient data transmission.

In existing, digital steganography, electronic communications include steganography coding under transport layer, like a document file, image file, media file program which can hide confidential data (i.e. secret files) through embedding methods. However, there is no privacy guarantee, if embed file is comprised with attacker. A new steganography method designed based on gray-level modification for true color images using multiple encryption(ME) algorithms (bit XOR operation, bits shuffling, and stego key-based encryption) to hide the secret text in images cover files[1]. It maintains the good privacy. However, this method produces high level of error rate during data extraction. It also takes long time to extracts secret data from cover file. Lempel-Ziv-Welch compression technique is worked to hide and compressed the image data by using modified kekre's algorithm. It's capable for large volume of data without incurring perceptual distortion [2]. However, this technique is only applicable for image data embedding. Integer Haar Wavelet Transform (IHWT) approach is designed through a lifting scheme to work on frequency of image [3]. It converts integer pixel values of an image into the integer wavelet coefficients and vice versa. However, this approach works based on some preconditioning methods and this approach has high complexity. Bacterial Foraging Optimization (BFO) approach developed for watermarking digital image to maintain the privacy. This method works based on bacterial movement [4]. However, this method produce high volume of error during extraction of data and embedding process is quite slow.

To overcome these issues, Data Compression and Embedding Algorithm are implemented to maintain tight privacy secret data file from malicious or external threat. This approach's objective is to design efficient data embedding framework to transmit the secure data to destination without revealing data privacy and without affecting the quality of cover file. This approach does not only consider on privacy in steganography but it also considers the size of original data for reliable data transmission to source. It works on both sides of source and destination. In detail, it assists data covering with encoding process to embed the original data with two cover files namely master file and output file to build the secure embedding process. Once, file selection is completed then it performs the file compression methods to reduce the original size of file. Hence, it proceeds for data encoding process with AES algorithms. Finally, it embeds the original data with cover files. After successful data covering process, master file content replace with output file content to conflict the unauthorized users. The paper contribution is as follows:

IIOAB JOURNAL
ISSN: 0976-3104

1. To build efficient embedding framework for contributing secure data to source without compromising the quality of cover file
2. Combine the compression approach with embedding frame work to maintain data privacy of original data and reliable data transmission in un-trusted networks.
3. Make strong embedding framework capabilities to supports various types of cover file and secure data file to utilize embedding framework.
4. Improve the Root Mean Squared Error (RMSE), Peak Signal Noise Ratio (PSNR) and Structural Similarity Index Measurement (SSIM) of proposed algorithm when compared to existing approaches

The organization of rest of paper is as follows: Section 2 introduces the relevant or closest work to proposed algorithm. Section 3 elaborates the system methodology, implemented framework details with mathematical derivation. Section 4 expresses details about implemented algorithm with performance. Finally, section 5 summarizes the overall work with future enhancements.

RELATED WORK

The [5] developed a novel multicarrier/ signature iterative generalized least-squares (M-IGLS) methods to find out unknown data hidden in hosts via multicarrier spread-spectrum embedding. [6] Designed blindly extraction technique is considered to the host signal, via multicarrier embedding. The hidden data is extracted from media like audio, video or image. In [7] introduced a new high bit rate LSB Picture information concealing system. The fundamental concept of the LSB computation is information installing which causes negligible implanting contortion of the host picture. [8] Developed reduced distortion algorithm for LSB image steganography. The main concept of this technique is data hiding bit embedding that causes minimal embedding distortion of the host image. [9] implemented Direct Sequence Spread Spectrum method for steganography with audio data. This method can be applied to embed messages in audio data. The information to be embedded must first modulated using the pseudo- random key sequence to increase the security and robustness of the system.

The [10] introduced Multicarrier Least Square (MLS) algorithm to extract unknown data hiding in image via multicarrier SS embedding. The data is covered in the image via DCT multicarrier DSSS (Direct Sequence Spread Spectrum). In [11] worked on spread spectrum (SS) image watermarking schemes using discrete wavelet transform (DWT), bio-orthogonal DWT and M-band wavelets coupled with various modulations, multiplexing and signaling methods for performance improvement in spread spectrum image watermarking. In [12] developed Half-Tone Pixel Swapping approach from carrier stego image to enhance the privacy as well as the image embedding capability. In [13] implemented H264 /AVC, video encryption, data embedding, data extraction. In[14] designed efficient data hiding method using audio steganography to convey safely in a totally imperceptible way and to abstain from attracting suspicion to the transmission of data.

In [15] focused on the data security approach which combined with encryption and steganographic techniques for secret communication by hiding inside the multimedia files. The files composed of insignificant bits which can be used for overwriting of other data. In [16] introduced advanced image visual cryptography by multilevel decomposition to maintain privacy and certainty of pictures. A content owner encrypts the original image by using an encoding key, and an information-hider can embed data into encrypted image. With an encrypted image, a receiver could initially decode with the encoding key, then extract the embedded information and recover the original image. In[17] designed difference expansion algorithm to embed data into the host image without causing overflow /underflow and distortion problems. In[18] studied about steganography and their applied technique. This paper also studied data hiding in audio signals and utilized LSP method for audio data hiding. In [19] introduced new steganography technique for audio data hiding. This approach embeds secret data into image hence image is embedded into the audio.

In [20] focused on new reversible data hiding method through serving room before encryption with a traditional Reversible Data Hiding algorithm. It used which enables to reversibly embed data in the encrypted image. In[21] developed a new technique for sending secret messages securely, using steganographic technique. This system worked for multiple level of security for data hiding. In [22] developed versatile audio steganographic methods to obtain secure and robust high rate of secret data. It also worked on digital audio steganography, which has emerged as a prominent source of data hiding across novel telecommunication technologies. In[23] reviewed the literature of digital audio steganographic techniques. It also explored their potentials and limitations to ensure secure communication. In[24] investigated analyzed various kinds of existing methods of steganography along with some common standards and guidelines drawn from the literature.
In [25] introduced TIMG algorithm to achieve the different goals of privacy for secure transmission of data. In [26] implemented Byte Rotation Algorithm which contains two techniques; one is random key generation technique & parallel encryption and decryption technique to perform encryption and decryption with minimal time. In [27] developed random scan algorithm to make the data more secure encryption before

IIOAB JOURNAL
ISSN: 0976-3104

data embedding. In [28] designed Efficient and secure cloud data migration (ESCDM) algorithm to migrate owner data from one cloud to another cloud. It migrate the data from one cloud to another after the confirmation from data owner side. In [29] introduced video steganography with digital watermarking techniques for data protection. This system hides large volume of data within a video but it has limitation in perceivable distortion during processing. In [30] worked a data hiding method using pixel value difference (PVD) steganography for digital image steganography to achieve more edges and to enhance the capacity of PVD. In [31] designed Pixel Value Ordering (PVO) method based embedding for image based Reversible Data Hiding (RDH) to plant secret data in minimal and maximum pixel values of each block of image.

SYSTEM DESIGN AND IMPLEMENTATION

This section expresses the brief knowledge about system design and implementation of proposed methodology. In details, proposed methodology works both side namely as source and destination. Here, source will select and secret file along with two cover data to embed the content with tight privacy. This system does not only maintain privacy but it also reduce the original size of data to enhance the embedding scheme. The workflow of proposed system is explained [Fig. 1] in details. The proposed system is divided into following module namely: Source, Data Covering and Encoding, Destination, Data Extraction & Decoding and data compression and embedding algorithm.

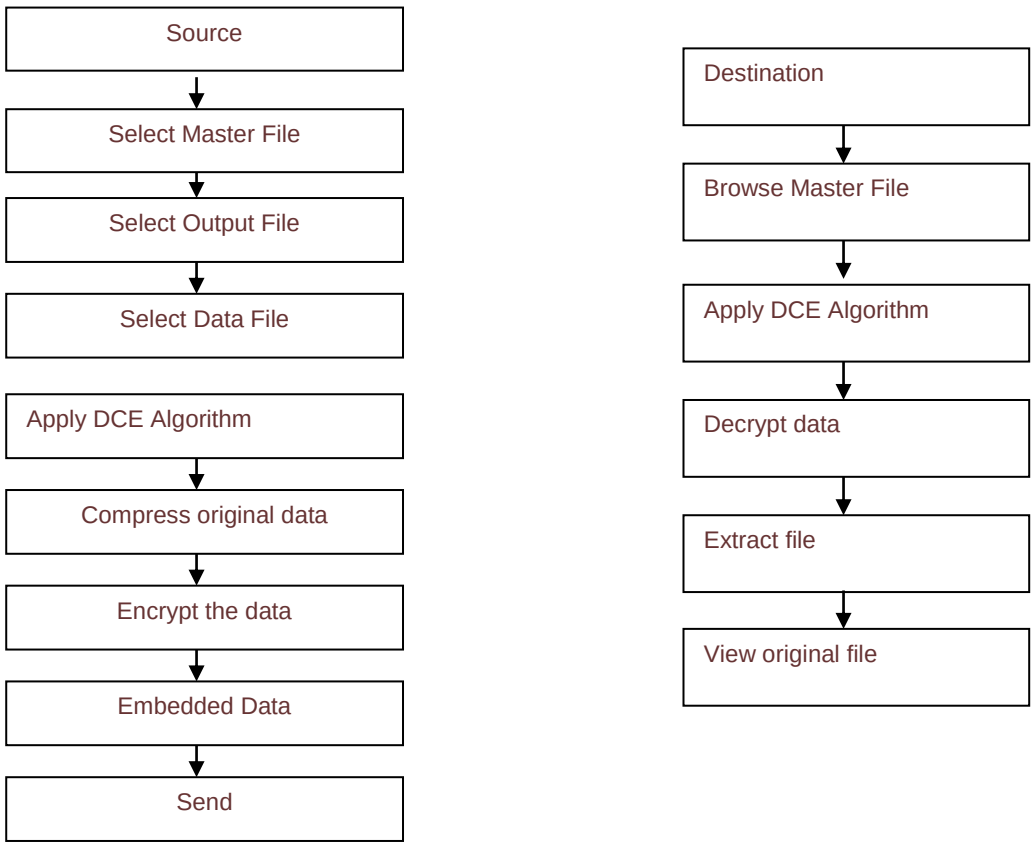

Fig.1: Workflow diagram for Data Compression and Embedding Algorithm

Here, source acts like a data owner who has large volume of secret data. Source wants to share the secret data with embedding format to maintain tight security in efficient way with minimal cost. Source can send secret message and as well secret file.

Data Covering and Encoding

Data covering is a hiding process to cover the secret file to enhance the embedding scheme. This scheme supports many types of data formats for covering and extraction in secure and efficient manner. This system select three file during embedding namely: master file, output file and secret data file. After selection of file, it reduced the size by using compression techniques. Hence, it encodes data using AES algorithm and cover the data to transmit to desired destination.

IIOAB JOURNAL
ISSN: 0976-3104

Destination

Destination extracts and decodes the original data with valid key from receive data covering file. In details, proposed algorithm works extract the error free data and explore in original size.

Data Extraction & Decoding

Data extraction is process to extract the secret data from cover file and after decoding process. This process main objective is to extract the secret data along with original size without affecting quality covering file. Here, AES algorithm is utilized for decoding the data. It extracts many types of data in various formats like audio, video and Document, PDF, Text etc.

Data Compression and Embedding Algorithm

Data Compression and Embedding Algorithm is implemented to maintain tight privacy secret data file from malicious or external threat. This approach's objective is to design efficient data embedding framework to transmit the secure data to destination without revealing data privacy and without affecting the quality of cover file or data. This approach does not only consider privacy in steganography but it also considers the size of original data for reliable data transmission to source in un-trusted. It works on both sides as well source and destination. In details, it assists data covering with encoding process to embed the original data with two cover file namely as master file and output file to build the secure embedding process. Once, file selection is completed then it performs file compression methods to reduce the original size of file. Hence, it proceeds for data encoding process with AES algorithms. Finally, it embeds the original data with cover files. After successful of data covering process, master file content replace with output file content to conflict the unauthorized users. This system always tries to avoid data file privacy revealing. Destination side, this approach co-operates to extract the cover file with original size without affecting the quality of cover file. After extraction of cover file, it performs decoding process with valid key using AES algorithm. This algorithm also displays the compression data percentage (%) along with cover file details.

This algorithm contains two process namely as a data covering & encoding and Data extraction & decoding process which are explained in details in above modules. The mathematical expression of proposed algorithm is described below in details.

Pseudo code: Data covering and Encoding
Input: Select Master File MF, Output File OPF, and Secret Data File SDF
Output: Get Data Cover File DCF
Procedure:
Start;
Browse MF, OPF and SDF;
Validate MF and OPF;
If MF does not contains any SDF already
Replace OPF content with MF content;
Encode the SDF;
Complete the embedding process;
Else
Display the message MF have already some other SDF;
Interrupt the embedding process;
End;

Pseudo code: Data Extraction and Decoding
Input: Select DCF
Output: View MF and SDF
Procedure:
Start;
Select MF;
Verify the MF;
If MF contains SDF
Display the complete cover file information with compression level;
Decode the SDF with Valid Key;
If Key is validated;
Decode & Display SDF;
Else
Display message to enter valid key;
Else
Display message there is no contents in this MF;
End;

ISSN: 0976-3104

Pseudo code for Data Compression and Embedding Algorithm

RESULTS AND DISCUSSION

Implementation Setup

In order to compares proposed mechanism with existing algorithm. The experiment is conducted on a laptop with Intel Dual Core processor (1.836 Hz), 2GB memory, and Window 7 Ultimate system. Here, this method implemented in JAVA with JDK 1.8 and NetBeans 8.0.

Performance Evaluation Parameter

In this phase, proposed scheme represent mathematical model to enhance privacy of secret data. In this scheme, security model work between data sender and data receiver. Even though, server is not trusted then also data sender secret message will be in safe during data transmissions. It displays following model separately such as Root Mean Squared Error (MSE), Peak Signal Noise Ratio (PSNR) and Structural Similarity Index Measurement (SSIM).
Root Mean Square Error (RMSE)
Root Mean Square Error (RMSE) is computed by getting the square root of the mean square error (MSE). The RMSE can be calculated as follows.

$$RMSE = \sqrt{MSE} \quad (1)$$

Peak Signal Noise Ratio (PSNR)

In this phase PSNR represents the mathematic model to perform the noise ratio of embedding images of secret during data embedding. The mathematical equation is expressed in equation (2).

$$PSNR = 10\log_{10}\frac{L^2}{MSE} \quad (2)$$

Where, PSNR is peak signal noise ratio of embedding, L is peak signal level for a grey scale of image frame it is taken as 255.

Structural Similarity Index Measurement (SSIM)

The SSIM metric is used to compute the similarity between two images. It designs for modeling of any image distortion as a combination of three factors that are loss of correlation, luminance distortion and contrast distortion. Mathematically, the SSIM is calculated in equation (3).

$$SSIM(x,y) = l(x,y).c(x,y).s(x,y) \quad (3)$$

$$l(x,y) = \frac{2\mu_x\mu_y + C1}{\mu_x^2 + \mu_y^2 + C1}$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + C2}{\sigma_x^2 + \sigma_y^2 + C2}$$

$$s(x,y) = \frac{\sigma_{xy} + C3}{\sigma_x\sigma_y + C3}$$

Where $\mu_x\mu_y$ is average value of x and y σ_x^2, σ_y^2 is variance x and y and σ_{xy} is the co-variance of x and y. l is dynamic range of pixel value and c contains three variable c1, c2, and c3 to stabilize division of weak denominator. X and y holds two windows size of images and varied pixel value.

[Table 1] expressed root means square error(RMSE), peak signal noise ratio(PNSR), and structural similarity index measurement(SSIM) for JPEG, PNG and BMP image formats with existing algorithms namely BFO[4], IHWT[3], LZW[2] and ME[1] approaches to evaluate the proposed algorithm.

IIOAB JOURNAL
ISSN: 0976-3104

Table 1. RMSE, PSNR and SSIM for JPEG, PNG and BMP formats Image Datasets

Learning Algorithm	Jpg			.png			.bmp		
	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	MSE	PSNR	SSIM
BFO	7.01	31.22	0.47	5.52	33.30	0.68	5.41	33.28	0.59
IHWT	1.78	51.37	1.00	1.45	47.89	0.99	1.21	51.44	1.00
LZW	1.72	43.50	0.99	1.69	43.58	0.97	0.04	43.47	0.97
ME	0.67	58.06	1.00	0.67	58.04	0.91	0.67	58.04	1.00
DCE	0.50	62.05	1.00	0.50	62.03	1.00	0.51	62.03	1.00

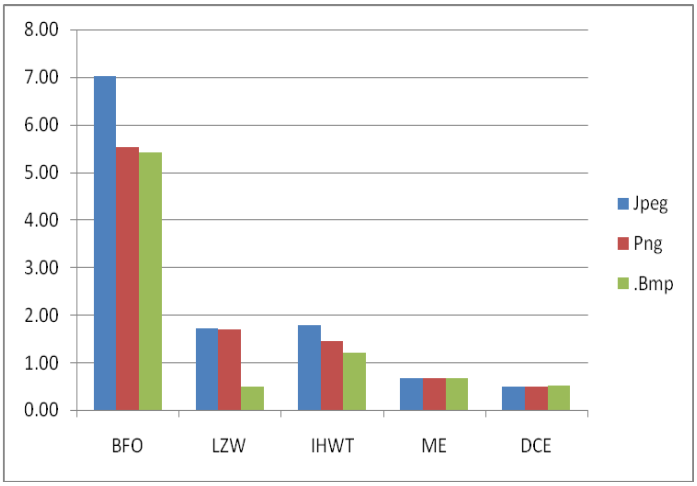

Fig. 2: Root Mean Square Error for .Jpeg, .Png and .Bmp Datasets

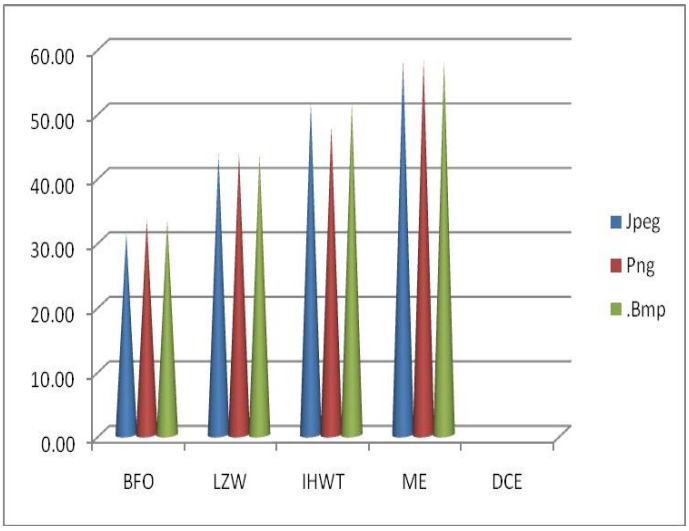

Fig. 3: Peak Signal Noise Ration for .Jpeg, .Png and .Bmp Datasets

IIOAB JOURNAL
ISSN: 0976-3104

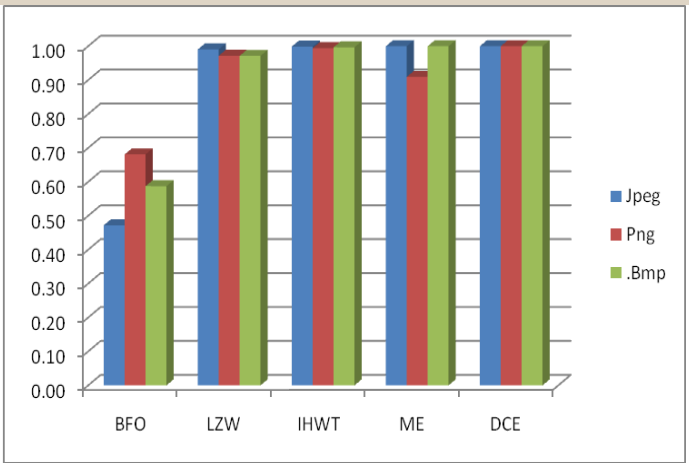

Fig. 4: Structural Similarity Index Measurement .Jpeg, .Png and .Bmp Dataset.

Based on [Fig. 2 to 4] result performance respect of root mean square error, peak signal noise ratio and Structural Similarity Index Measurement for overall images dataset, it realized that proposed approached performs well. In details, ME algorithm is closest approach to proposed method with respect of RMSE, PNSR and SSIM. However, ME algorithm performance is low compare than proposed DCE algorithm. DCE approach reduces RMS 0.17, increase PNSR 3.54 and SSIM 0.30 compare then ME algorithms. Hence, this paper states that proposed DCE algorithm is best approach on overall datasets.

CONCLUSION

This paper presents Data Compression and Embedding Algorithm to maintain tight privacy of secret data file from malicious or external threat. This approach's objective is to design efficient data embedding framework to transmit the secure data to destination without revealing data privacy and without affecting the quality of cover file. This approach does not only consider on privacy in steganography but it also considers the size of original data for reliable data transmission to source. It works on both sides as well source and destination. In details, it assists data covering with encoding process to embed the original data with two cover file namely as master file and output file to build the secure embedding process. Once, file selection is completed then it performs the file compression methods to reduce the original size of file. Hence, it proceeds for data encoding process with AES algorithms. Finally, it embeds the original data with cover files. After successful of data covering process, master file content replace with output file content to conflict the unauthorized users. Its supports various formats secret data file. Based on experimental result, it realized that proposed systems performs well compare than existing approaches in the terms of root mean squared error (RMSE), PSNR and SSIM. It enhances the RMSE 0.17 PSNR 3.54db and SSI 0.30. In future, this paper can be extended with vehicular ad-hoc network to establish reliable transmission between sources to destination.

CONFLICT OF INTEREST
There is no conflict of interest.

ACKNOWLEDGEMENTS
None

FINANCIAL DISCLOSURE
None.

REFERENCES

[1] Muhammad K, Ahmad J, Sajjad M., Zubair M. Secure image steganography using cryptography and image transposition, arXiv preprint arXiv:1510.04413, 2015, pp.1-22.
[2] Jassim FA. Increasing Compression Ratio in PNG Images by k-Modulus Method for Image Transformation, arXiv preprint arXiv:1307.0036, 2013, pp.1-10.
[3] Abu NA, Adi PW, Mohd O.[2014] Robust digital image steganography within coefficient difference on integer haar wavelet transform. International Journal of Video & Image Processing and Network Security, 14(2):1-8.
[4] Jain C, Chugh A.[2015] Design and Development of BFO Based Robust Watermarking Algorithm for Digital Image. International Journal of Computer Applications Innovations in Computing and Information Technology, pp. 14-18.
[5] Li M, Kulhandjian MK, Pados DA, Batalama SN, Medley MJ.[2013] Extracting spread-spectrum hidden data from digital media. IEEE Transactions on Information Forensics and Security, 8(7):1201-1210.
[6] Dhas YS, Abisha D. M-IGLS Based Extracting Hidden Data from Digital Media, 2014 Jan, 3(1), pp. 6544-6548.
[7] Devi RN, Ranjith B.[2014] Extracting Spread-Spectrum Hidden Data from Picture Representation. International Journal of Computer Engineering in Research Trends 1(6): 405-08.

[8] Bijwe MKB, Bamnote GR.[2014] Extracting Spread-Spectrum Hidden Data from an Image, International Journal of Computing and Technology, 1(3):74-79.
[9] Malik H, Kang SS.[2013] Designing and Evaluation of Performance of a Spread Spectrum Technique for Audio Steganography, International Journal of Advanced Research in Computer Science and Software Engineering, 3(8):.37-45.
[10] Redekar S, Gunjan R. [2015] Extracting Spread Spectrum Data from Image Using MLS Algorithm, International Journal of Computer Science and Mobile Computing, 4(7): 253-262.
[11] Maity SP, Kundu MK.[2011] Performance improvement in spread spectrum image watermarking using wavelets. International Journal of Wavelets, Multi resolution and information processing, 9(1):1-33.
[12] Deshmukh SA, Sambhare, PB.[2015] An Authentication of Secretely Encrypted Message using Half-Tone Pixel Swapping from Carrier Stego Image, International Journal of Computer Science and Information Technologies,6(3) :2409-2414.
[13] Pawar YA, Sawant SD.[2015] Data Hiding by Code word Substitution (Encrypted H. 264/AVC Video Stream), International Journal of Engineering Sciences & Research Technology, 4(7):62-28.
[14] Kulkarni M, Phatak M, Rathod U, Prajapati S, Mujgond MS.[2016] Efficient Data Hiding Scheme using Audio Steganography, International Research Journal of Engineering and Technology, 3(3): 1701-1706.
[15] Shruti.[2015] Enhanced Data Security Using Digital media-Video Stegnography, An international journal of advanced computer technology, 4(7): 1934-37.
[16] Nagdive MPS, Raut AB.[2015] An Advanced Image Visual Cryptography By Multilevel Decomposition For Data Hiding, International Journal of Science, Engineering and Technology Research, 4(5):1536-42.
[17] Rao TVN, Govardhan A, Badashah SJ.[2010] Improved Lossless Embedding and Extraction-A Data Hiding Mechanism, International Journal of Computer Science and Information Technology, 2(2):75-86.
[18] Olanweraju RF, Khalifa OO. Increasing the hiding capacity of low-bit encoding audio steganography using a novel embedding technique, World Applied Sciences Journal, 2010 1(20), pp.79-83.
[19] Lavanya B, Smruthi Y, Elisala SR.[2013] Data hiding in audio by using image steganography technique, International Journal of Emerging Trends & Technology in Computer Science, 2(6):.27-30.
[20] Priya V, Sudharson D. [2014] Reversible Information Hiding in Videos, International Journal of Research in Computer Applications and Robotics, 2(2):35-40.
[21] Kumar PM, Shunmuganathan KL.[2010] A reversible high embedding capacity data hiding technique for hiding secret data in images, International Journal of Computer Science and Information Security, 7(3):109-115.
[22] Djebbar F, Ayad B, Meraim KA, Hamam H. [2016]Comparative study of digital audio steganography techniques. EURASIP Journal on Audio, Speech, and Music Processing, 2012(1):1-16,
[23] Kulkarni SA, Patil SB Patil, Patil BS. [2012]A Optimized and Secure Audio Steganography for Hiding Secret Information – Review, Journal of Electronics and Communication Engineering, pp.12-16.
[24] Cheddad A, Condell J, Curran K, MC Kevitt P.[2014] Digital image steganography: Survey and analysis of current methods, Signal processing,x, 90(3):727-752.
[25] More S, Arab MM.[2014] Efficient TIMG Algorithm for Secure Transmission of Data, IOSR Journal of Computer Science, pp43-46.
[26] Maitri PV, Waghole DS, Deshpande VS.[2015] Low latency for file encryption and decryption using BRA algorithm in network security, In 2015 IEEE International Conference on Pervasive Computing (ICPC), pp. 1-4.
[27] Gupta S, Saxena J, Singh S.[2015] Design of Random Scan Algorithm in Video Steganography for Security Purposes, IOSR Journal of Electronics and Communication Engineering, 10(5):14-20.
[28] Krishnan R, Mathivanan V.[2015] Privacy Proof of Data Transportability from one Cloud to another Cloud Service Provider, International Journal of Control Theory and Application, 8(2):381-389.
[29] Khosla S, Kaur P. [2014]Secure Data Hiding Technique using Video Steganography and Watermarking, International Journal of Computer Applications, 95(20): 7-12.
[30] Halder T, Karforma S, Mandal RA.[2015] Novel Data Hiding Approach by Pixel-Value-Difference Steganography and Optimal Adjustment to Secure E-Governance Documents, Indian Journal of Science and Technology, 8(16):1-7.
[31] Shastri S, Thanikaiselvan V. [2016] PVO based Reversible Data Hiding with Improved Embedding Capacity and Security. Indian Journal of Science and Technology, 9(5): 1-7.