Data Masking using Cryptographic Techniques in Steganography

Karthikeyan B, Narla Sai Teja, Kolisetty Sarath Chandra

Abstract: ***In today's world with the rapid development in technology we've entered the time where tremendous amounts of data are at constant risk and the security of data must be given utmost importance, we have cryptographic as well as steganographic techniques to achieve the same. In this paper, we have developed a method involving both these techniques to gain better results. Initially, the secret message to be transmitted is converted to ASCII values and is encrypted with the RSA asymmetric key cryptosystem and the obtained encrypted message which is in decimal format is converted to octal format and then to binary format by considering it as octal. Next, the binary bits attained are hidden inside any digital image by Least Significant Bits substitution to obtain the new image with secret message called as stego image is sent to the receiver, who performs inverse operations to obtain the secret message. Mean Squared Error and Peak Signal to Noise Ratio values between the original and stego images are computed and the results obtained by this method are found to be quite better than many other schemes.***

Keywords: ***Cryptography, LSB, MSE, PSNR, RSA, Steganography.***

I. INTRODUCTION

Providing a secure channel for communication has become a strenuous task in today's technologically advancing world. As the threats to privacy are increasing there is a great demand for security. Cryptography and Steganography are two vital techniques in providing security, but with the increase in the intruders across the network these techniques must be evolved timely to avoid these threats.

A. Cryptography

It is the process of converting the plaintext into an unintelligible ciphertext by sender and decoding the ciphertext back to original text by receiver with the help of secret keys for a secure transmission. Role of Cryptography is to provide Confidentiality, Integrity and Authenticity to the data. We have two categories in secret key cryptography besides the old traditional methods, Symmetric and asymmetric cryptographies.

Plaintext -> Ciphertext is Encryption

Ciphertext -> Plaintext is Decryption

B. Steganography

This is the process of concealing secret data by embedding it into digital media like Images, audio files, video files, HTMLs etc. These digital media act as a cover to secret data and helps in transmission. In this paper our focus is mainly on Image steganography where the cover medium used is an image.Both the techniques are very significant but are better when used together. When cryptography is used alone, the intruder knows the cipher text and if proper cryptanalysis is made the security can be compromised. Steganography when used alone, hides the data but once found by the intruder then decoding the data wouldn't be a big task. Hence the combined approach is encouraged as it provides confidentiality, integrity and authenticity as well as secrecy to the data. In combined approach the data is hidden from the intruder and even if found, decoding would be difficult, thus providing a double layer of security.In Steganography it is to be noted that the stego image formed after the secret message concealing must not be distorted much and must be analogous to the original image. If the distortions are noticeable steganographic security is compromised, MSE and PSNR values are calculated to know the amount of distortions. In this paper we have proposed a method involving asymmetric or public key cryptography and image steganography achieving better MSE and PSNR values.

II. LITERATURE SURVEY

Cryptography is the study of encryption of data which is the transformation of normal text to cipher text for the secure transmission of data and cryptanalysis is the decryption process of data without the knowledge of encryption schemes, various types of schemes used for enciphering and deciphering are studied in William Stallings's [8].Steganography is concealing of secret data in a cover medium, Image steganography is a kind in which cover medium is an image. Image Steganography can be achieved by quite a few techniques, one such method is LSB substitution where the original image is converted into stego image by manipulating the image's least significant bits with the message bits. The basic idea behind LSB substitution is proposed in [6] and the technique proposed in [7] gives a clear picture, in this technique the original message is divided into two parts and is processed, part one for message embedding and part two for showing changes done on part one.In image Steganography, sometimes stego images and normal images are distinguishable. So only limited data can be hidden in the cover image as discussed in [9].

Revised Manuscript Received on June 07, 2019.

Dr.Karthikeyan B SAP, School of Computing, SASTRA Deemed University, Thanjavur, India.

Narla Sai Teja Student, School of Computing, SASTRA Deemed University, Thanjavur, India.

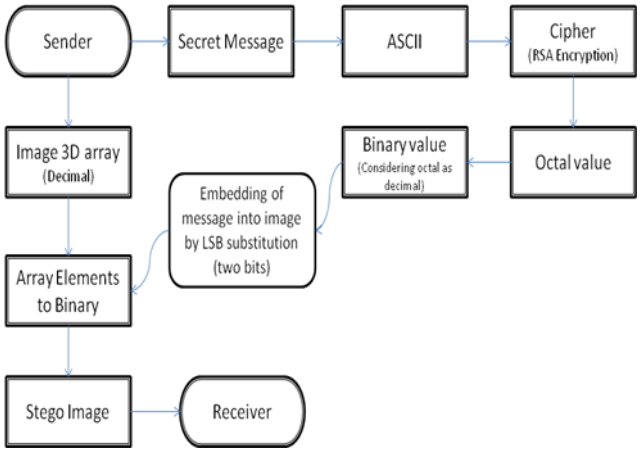
Kolisetty Sarath Chandra Student, School of Computing, SASTRA Deemed University, Thanjavur, India.

In this, an analysis is done on LSB based techniques for image steganography on the embedding capacity of cover images such that stego images are indistinguishable.By using LSB substitution based methods a large number of bits have to be modified. A method is proposed in [11] in which a bit mapping strategy is employed where bits from cover medium and secret data are grouped logically into pairs, these groups consisting of four bits of secret data are mapped with four most significant bits of cover medium and mapping status is maintained by employing two-bit LSB substitution.An improved version of the LSB substitution is inverted LSB technique as proposed in [3] where the secret message is complemented before applying inverted LSB method on randomly selected pixels of the image. This gives a better PSNR value because in this number of changes in bits of pixels of the original image are lesser than that in simple LSB techniques.Various steganographic approaches for data hiding in images in JPEG format and in spatial representation with an emphasis on the Imperceptibility, robustness, and hiding capacity and corresponding methods for steganalysis are studied with reference to [4] and [5].For the security of data we have cryptographic and steganographic techniques but both techniques have problems of their own, hence a combined technique as proposed in [1] where modified AES for encryption and enhancing PVD image steganography proposed in [2] for data hiding are used for improving data security and embedding capacity of images.Another combined approach cryptographic and steganographic techniques can be seen in [10] where the secret file to be transmitted is initially compressed and then operated by the AES algorithm to obtain ciphertext. The ciphertext is then hidden in the cover medium, for pixel selection Genetic algorithm is used so that the secret file is more secured. The genetic algorithm is a search based on the survival of the fittest theory.In [12] a dynamic procedure is adapted where both crypto and steganographic methods are used and the cryptosystem used is the RSA algorithm. In this the secret data is acted upon by the RSA cryptosystem, the result and cover medium are organized into blocks and an adaptive LSB substitution method is employed dynamically to assign cipher blocks in circular queues.The RSA cryptosystem and its applications were studied in [8] and the implementation of RSA Public key cryptosystem even for very long messages and the techniques to enhance the RSA cryptosystems are discussed in [13]. This includes a study on the performance of the RSA algorithm.

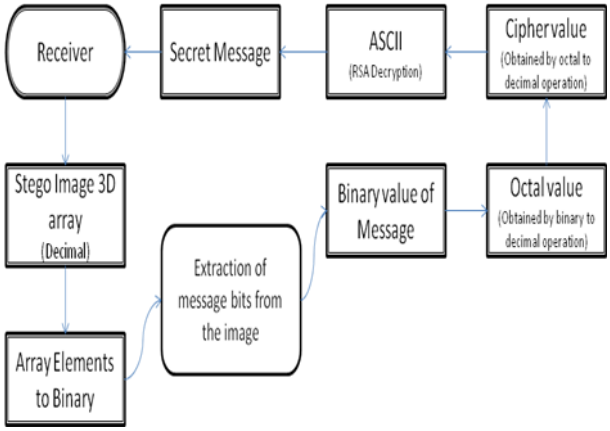
III. PROPOSED METHOD

Our proposed method of message security is a merger of Cryptography and Steganography. The sequence of steps followed is shown in the flow diagram.

A. Flow diagram of Encryption



B. Flow diagram of Decryption



C. Encryption

1. The secret message to be transmitted is converted into ASCII values. So we obtain the decimal values of the message. All the 256 characters of ASCII list can be used.
2. Then the values are encrypted with RSA public key cryptosystem using sender's private key and of receiver's public key to get the encrypted values which are in decimal format as well.
3. These encrypted values which are in decimal format are converted into octal format.
4. The values obtained in octal format are converted to binary format by assuming the encrypted values in octal format as decimal values.
5. Each character's binary values are padded with 0's as required to form a value with certain constant bits to make it easier for decryption and the image is read as a 3 dimensional array representing RGB values of the pixels.
6. The number of the characters of secret message is embedded into the first few bits of the first row of the image and then the message bits are embedded successively in next rows. The values in image array are in decimal format and are converted to binary format before embedding.

Embedding is done via LSB substitution. Two least significant bits of each element in image array are replaced with two-two bits of message successively.After the embedding of entire message into image to form a distorted image called as stego image, it is sent to the receiver.

D. Decryption

1. The receiver reads the stego image as a 3 dimensional array representing RGB values of the pixels and extracts the length of image from the first row. Extraction is done by first converting the elements of image array in decimal format to binary format and then last two bits of each element of array are taken and merged. Since each character's size is the constant chosen while encryption, characters can be extracted without any issues.
2. Next the message bits are extracted into set of bits with each set consisting of the number of constant bits chosen for each character while encryption.
3. Each set is converted into decimal format to obtain the values in octal format because during encryption, step [IV] of Encryption is done.
4. Then the values of each set obtained are converted from octal format to decimal format to obtain enciphered values.
5. These decimal values are decrypted with RSA algorithm using sender's public key and receiver's private key to obtain the original decimal values.
6. This decimal value in each set is the ASCII value of each character of secret message, so each of these decimal values are converted into strings to obtain the secret message.

E. RSA Algorithm

RSA is an asymmetric key cryptography system. In this the receiver has a public key and a private key which are generated using two different prime numbers. The former can be known to sender and the public without any security compromises but the latter must be kept obscured even from the sender.Message is enciphered using the public key by the sender and sent to the receiver.Deciphering is done by the receiver with the private key.Here two keys are used, hence called as asymmetric key cryptosystemThe algorithm with the generation of key, encryption and decryption is studied with reference to William Stallings [8]The distortions amidst the cover image and stego image are determined with the help of Mean squared error (MSE) and Peak signal to the noise ratio (PSNR) measures.

MSE and PSNR can be calculated as follows:

$$PSNR = 10\log_{10}\left(\frac{MAX_I^2}{MSE}\right)$$

$$= 20\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right)$$

$$= 20\log_{10}(MAX_I) - 10\log_{10}(MSE)$$

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[I(i,j) - k(i,j)\right]^2$$

Where I(I,j) is original image
K(I,j) is reconstructed image
m,n are dimensions of image
MAXI is maximum possible pixel value of image

Referred from [14].

IV. WORKING MODEL

Text to be hidden: 'hello world'
Length of the text: 11
Value of length is converted into binary
Binary value of length: '0000000001011'
Zeroes should be appended such that length is equal to 13.
This binary string is embedded in first row of matrix.
Text is converted to ASCII.
ASCII value of text :
[104,101,108,108,111,32,119,111,114,108,100]
These ASCII values are encrypted using RSA algorithm.
Prime numbers chosen in RSA algorithm are:
P=23
q=31
Cipher text obtained:
[292,188,271,271,567,280,491,567,321,271,679]
Cipher text is converted into octal.
Octal vector :
[444,274,417,417,1067,430,753,1067,501,417,1247]
Octal vector is converted to binary by considering it in decimal.
Binary string:
'0011011110000100010010001101000010011010000110000101011001101011100101111000110000101011001111101010011010000110011011111'

A. Appending length

Length is chosen to be a thirteen bit binary string, it is appended with 0's and embedded in first sufficient values in first row of matrix.The first seven values in first row of image array: [103, 103, 103, 103, 104, 104, 104….]
Binary string of length after appending 0's: 0000000001011
Last two bits in 103 are replaced by first two bits in the string, last two bits in 103 are replaced by next two bits and so on.
First row after embedding length of secret message is:
[100, 100, 100, 100, 105, 105, 106….]

B. Appending 'hello world'

The given text 'hello world' is embedded from second row of matrix.The second row of matrix is: [103, 103, 103, 103, 104, 104....]The binary string of cipher text of the first letter 'h' is embedded in the following way:First the cipher value of h is appended with 0's to form constant number of bits.Binary string: '00110111100' (No. of bits to be formed which is constant for every character is taken as 11)Now the first value in matrix is 103 and its binary value is 1100111

Now least two significant values in above binary value are replaced with first two bits in the binary string.Binary value of first number in matrix is changed to 1100100.Now first value of matrix is 100.The next value in matrix is 103 and its binary value is 1100111.Now least two significant values in above binary value are replaced with next two bits in the binary string.Binary value of next number in matrix is changed to 1100111.Now second value of matrix is 103.
The next value in matrix is 103 and its binary value is 1100111.Now least two significant values in above binary value are replaced with next two bits in the binary string.Binary value of next number in matrix is changed to 1100101.Now third value of matrix is 101.And so on till every character is covered.

C. Decryption

The first line of matrix consists of length of string hidden in it. Length is thirteen bit binary length which is obtained by grouping least two significant bits in the values of first line until length of gathered bits is 13. (If only one bit is present in value of matrix, it has to be considered as least significant bit)
The first row of matrix is: [100, 100, 100, 100, 105, 105, 106…….]Least two significant bits gathered are: 0000000001011Decimal value of above string is: 11 (length is 11)Initialize a counter to 0.Get first eleven bits from least significant bits in values from second row of the matrix.The eleven bits are: '00110111100'The decimal value of binary string is: 444The decimal value of above octal number is: 292
Above value is decrypted using RSA decryption. The plain text after decryption is: 104Above value is ASCII value. It represents the letter: 'h'Counter gets incremented.Next 11 its are obtained, above steps are repeated, character obtained is appended to 'h' and counter gets incremented.These steps are repeated until counter gets equal to length.Finally the secret message is obtained when counter equals the length.

Fig. 1 Cover Image

Fig. 2 Stego image after insertion of "hello world"

The mean-squared error is 0.000176000The Peak-SNR value is 85.6757

V. EXPERIMENTAL RESULTS

The method proposed in this study is simulated in MATLAB and the following results are achieved.Three different digital images are taken and secret messages of length 1000 are encrypted and embedded in each of the images to obtain the stego images.
Size of cover image of Earth - 1.31 MB
Size of cover image of Windows - 627 KB
Size of cover image of Tajmahal - 747 KB

Fig. 3 Windows cover image

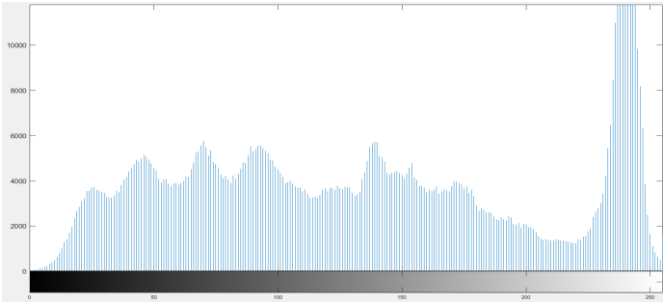


Fig. 4 Windows cover image histogram

Fig. 5 Windows stego image

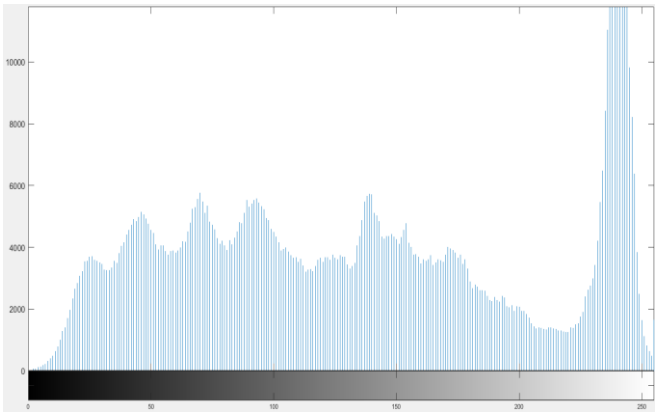


Fig. 6 Windows stego image histogram

Fig. 7 Tajmahal Cover image

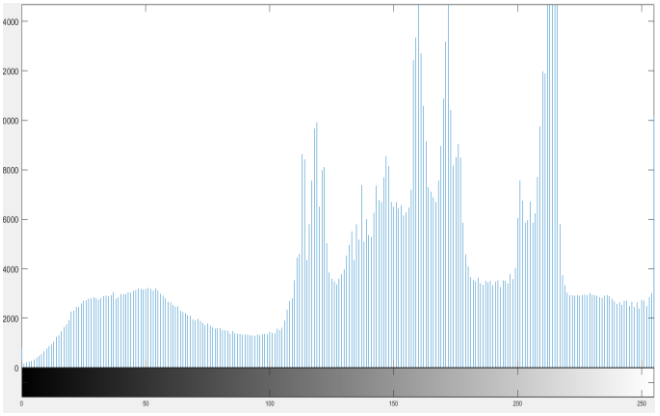


Fig. 8 Tajmahal Cover image histogram

Fig. 9 Tajmahal Stego image

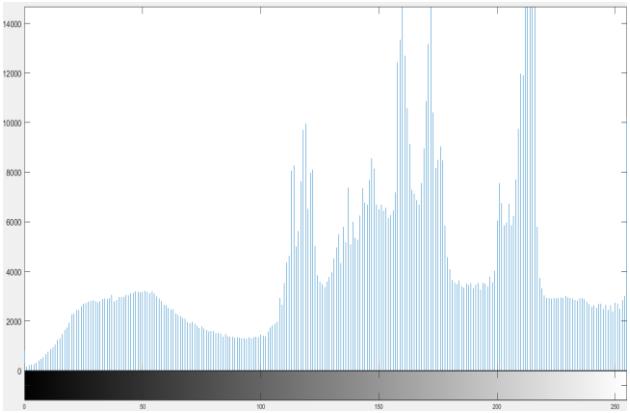


Fig. 10 Tajmahal stego image histogram

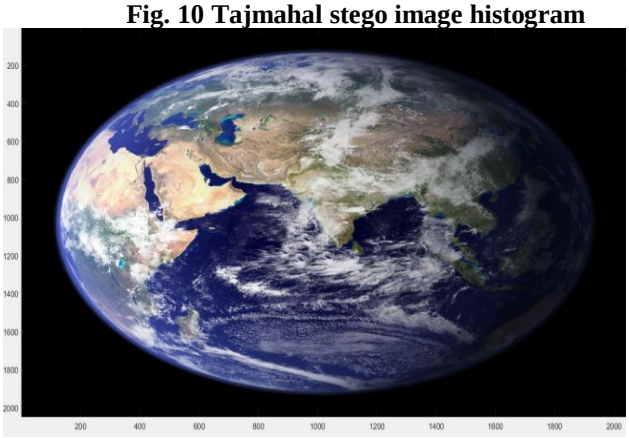


Fig. 11 Earth Cover image

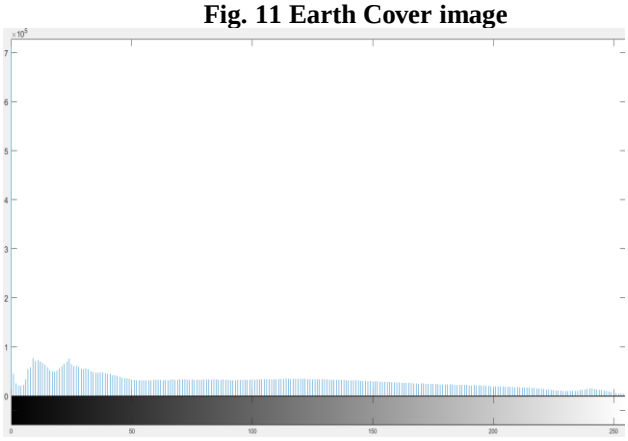


Fig. 12 Earth Cover image histogram

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Fig. 13 Earth Stego image

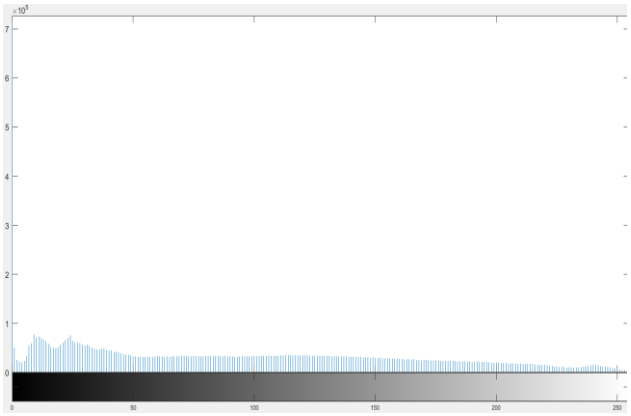


Fig. 14 Earth Stego image histogram

It is obvious that the distortions between the original digital image and the message embedded stego image are not noticeable with even the histograms looking quite similar. Hence the intruders cannot find the difference between normal images and images carrying secret data.MSE and PSNR measures are computed for secret messages of various lengths embedded within these images and are tabulated as follows:

Table: I MSE, PSNR and SNR values

Image	Size of text	MSE	PSNR	SNR
Earth	50	2.48E-05	94.187	85.6912
Earth	100	4.98E-05	91.1559	82.66
Earth	500	0.000249	84.1607	75.6648
Earth	1000	0.000499	81.149	72.6531
Windows	50	0.000902	78.5795	74.2699
Windows	100	0.001774	75.6406	71.331
Windows	500	0.007987	69.1071	64.7975
Windows	1000	0.016803	65.8768	61.5672
Tajmahal	50	0.00082	78.995	75.2134
Tajmahal	100	0.001324	76.9105	73.1288
Tajmahal	500	0.00558	70.6648	66.8832
Tajmahal	1000	0.010878	67.7652	63.9836

The graphs of MSE and PSNR for secret messages of various sizes of three images are plotted as per the data in table.

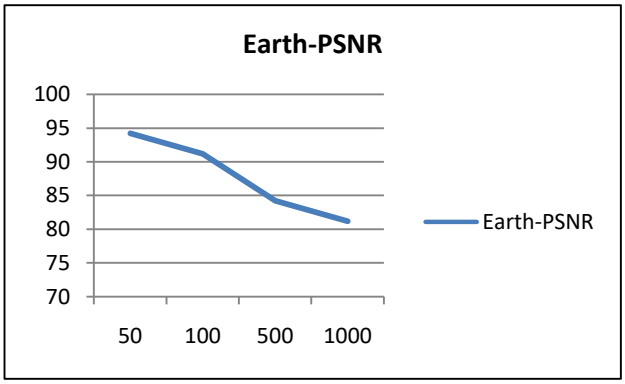


Fig. 15 PSNR of Earth image

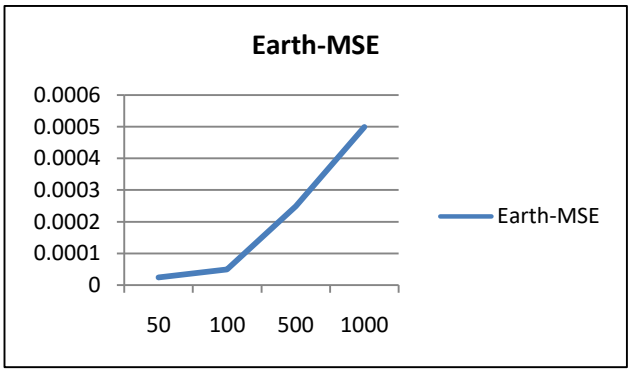


Fig. 16 MSE of Earth image

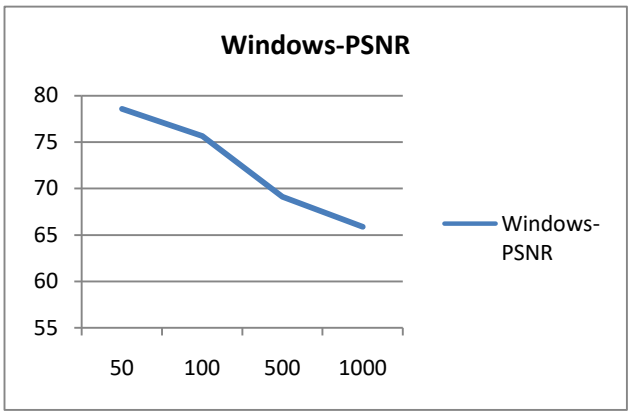


Fig. 17 PSNR of Windows image

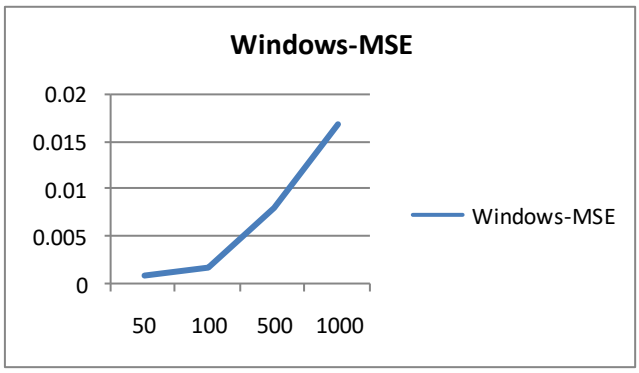


Fig. 18 MSE of Windows image

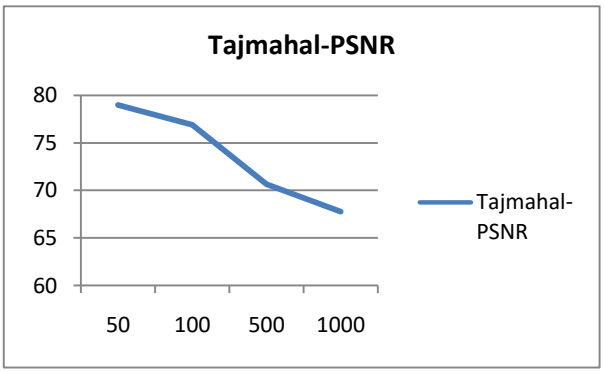


Fig: 19 PSNR of Tajmahal image

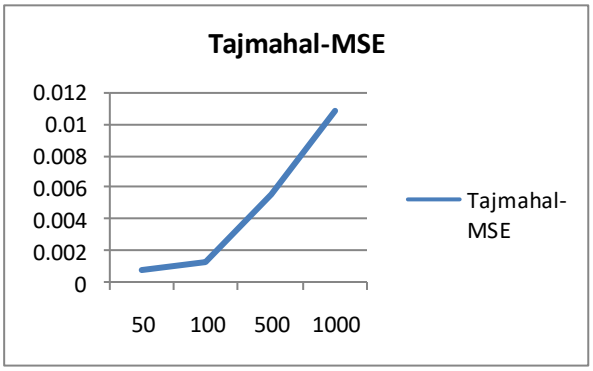


Fig. 20 MSE of Tajmahal image

VI. CONCLUSION

In this study a novel secure communication approach by merging the cryptographic techniques and steganographic techniques has been proposed, thus providing a double layer of security to the data. For encryption of data RSA asymmetric key cryptosystem is used and for embedding of encrypted secret message into the image LSB method is used. This method is simulated in MATLAB and the results are obtained. MSE and PSNR values which are the measures of distortions between the cover images used and the stego images obtained are computed.Higher the PSNR values better the final image obtained and lower the MSE values lesser the distortions. The image with more size has better PSNR values. The results are found to be quite better than other techniques. Thus our proposed method increases the confidentiality and integrity of message with the use of cryptography as well as hides the message exceptionally well with less distortions.

REFERENCES

1. Marwa E. Saleh , Abdelmgeid A. Aly and Fatma A. Omara, “Data Security Using Cryptography and Steganography Techniques”, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 6, 2016.
2. M. E. Saleh, A. A. Aly, and F. A. Omara, “Enhancing Pixel Value Difference (PVD) Image Steganography by Using Mobile Phone Keypad (MPK) Coding”, International Journal of Computer Science and Security (IJCSS), Volume (9), Issue (2), pp. 397 - 397, 2015.
3. Rupali Bhardwaj and Vaishali Sharma, “Image Steganography Based on Complemented Message and Inverted bit LSB Substitution”, 6th International Conference On Advances In Computing & Communications, ICACC 2016, 6-8 September 2016, Cochin, India.
4. Bin Li, Junhui H, Jiwu Huang and Yun Qing Shi, “A Survey on Image Steganography and Steganalysis”, Int. J. Inf. Hiding Multimedia Signal Process. 2 (2011) 142–172.
5. Zaid Al-Omari and Ahmad T. Al-Taani, “A Survey on Digital Image Steganography”, ICIT 2015 The 7th International Conference on Information Technology
6. Chi-K Chan and L.M. Cheng, “Hiding data in images by simple LSB substitution”, Pattern Recognition, vol. 37, no. 3, pp. 469474, 2004.
7. Marghny H. Mohamed and Loay M. Mohamed, (2016), “High Capacity Image Steganography Technique based on LSB Substitution Method”, Applied Mathematics & Information Sciences, Appl. Math. Inf. Sci. 10, No. 1, 2016, pp.: 259- 266.
8. W. Stallings, “Cryptography and Network Security: Principles and Practices, 6th edition”.
9. R. Chandramouli and N. Memon, "Analysis of LSB Based Image Steganography Techniques", Image Processing Proceedings. 2001 International Conference on, vol. 3, pp. 1019-1022.
10. P. Sethi and V. Kapoor, “A proposed novel architecture for information hiding in image steganography by using genetic algorithm and cryptography”, Procedia Computer Science 87 (2016) 61 – 66.
11. A. Zakaria, M. Hussain, A. Wahab, M. Idris, N. Abdullah and KH. Jung, “High-capacity image steganography with minimum modified bits based on data mapping and LSB substitution”, Appl Sci. 2018; 8(11):2199.
12. M. Jain, L. Saroj Kumar and Sunil Kumar V, “Adaptive circular queue image steganography with RSA cryptosystem”, Perspectives in Science Volume 8, September 2016, Pages 417-420.
13. M. Preetha and M. Nithya, “A study and performance analysis of RSA algorithm”, International Journal of Computer Science and Mobile Computing, Vol. 2, Issue. 6, June 2013, pg.126 – 139.
14. Andrzej Chybicki, "Applications of compression techniques for reducing the size of multibeam sonar records", 2008 1st International Conference on Information Technology, 05/2008.

AUTHORS PROFILE

Dr.Karthikeyan B SAP, School of Computing, SASTRA Deemed University, Thanjavur, India. M.Sc.,M.Tech.,Ph.D. bkarthikeyan@it.sastra.edu Ph.+91 8072134028.

Narla Sai Teja Student, School of Computing, SASTRA Deemed University, Thanjavur, India. Btech CSE-IV. narla.saiteja@gmail.com Ph.+91 9629607463.

Kolisetty Sarath Chandra Student, School of Computing, SASTRA Deemed University, Thanjavur, India. Btech CSE-IV. ksarathchandra98@gmail.com Ph.+91 9182585155.