Unicode and Colours Mapping for Cryptography and Steganography using Discrete Wavelet Transform

Madhusmita Sahu
Asst. Professor, Dept. of Computer Sc. &IT, ITER
S'O'A university
Bhubaneswar,Odisha,India

Soumya Ranjan Pradhan
Scholar Mtech, Dept of Computer
Sc. and IT, ITER
S'O'A university
Bhubaneswar,Odisha,India

Madhusmita Das
Asst. Professor, Dept. Of Computer Sc. &IT, ITER
S'O'A university
Bhubaneswar,Odisha,India

ABSTRACT

With the rapid growth of communication/transaction over internet and continuous efforts by eavesdroppers or Cyber attackers, there is a continuous need for new strong alternative algorithm. Techniques for information hiding are becoming increasingly more sophisticated and widespread. The primary goal, for designing such algorithms, is to provide security. However, performance and implementation ease and execution time are other factors that make an algorithm more usable. An improved encryption model, which is an extension of an existing algorithm based on Unicode and colours map can be used & Steganography using indiscrete wavelet transform (HCSSD) is proposed. The cipher text generated by this method can be varying in size as the plaintext and will suitable for practical use in the secure transmission of confidential information over the internet.

Keywords

Public key Crypto system, Unicode, *Stego Image, Wavelet Fusion*

1. INTRODUCTION

Internet, due to its wide availability and cheaper cost has attracted most business and financial applications. Use of internet has several advantages like rapid transactions, easy delivery, improved supply chain and customer services and strong social networking etc. Internet has become an integral part of our life. Despite its advantages, it faces challenges like maintaining privacy and confidentiality of messages which most users are scared of. In order to make data more secure over internet, encryption and decryption is a common approach. Cryptography is the area that deals with encoding (encryption) of data by applying transformation algorithms and then decoding (decryption) by intended user [1]. Cryptographic algorithm does not restrict to a domain, we are of opinion that producing unbreakable cipher text is an art not technology.

Steganography is the art and science of communicating in a way which hides the existence of the communication. The steganography make the presence of secret data appear invisible to eaves droppers such as key loggers or harmful tracking cookies, where the users' keystroke is monitored while entering password and personal information. For intelligent and tech-savvy generation, with the availability of high end processing tools, there is a continuous need to develop strong cipher text. This paper suggests an improvement over cryptographic algorithm suggested by Maram Balajee [2] on the basis of Unicode and Computer Supported Colours & a strong steganographic approach using Discrete Wavelet Transform (HCSSD) suggested by H S Majunatha Reddy, & K B Raja [3] to form Stego image.

1.1 Asymmetric cryptography

In general asymmetric key algorithm is used two different key for encryption and decryption known as private and public where public key is used for encryption and private key is used for decryption. RSA is the example of asymmetric key cryptography technique [4, 5, and 6].

a. The need for alternative to private key encryption

b. Early Development of Alternative Key Exchange Mechanism

c. Development of Asymmetrical Key Encryption

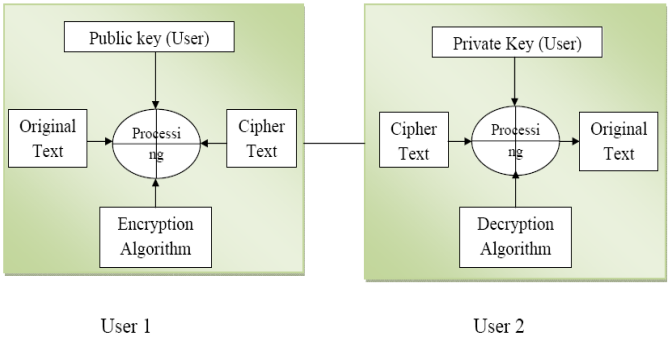

Figure 1: public key cryptography

2. UNICODE

Owing to limitation of traditional coding system for representing characters in computer, which forced users to create their own encoding schemes in order to represent other characters, a consistent encoding system was developed in the form of Unicode [7]. Unicode is most widespread and popular encoding scheme that has led to internationalization and localization of computer software. It implemented successfully in most recent technologies like xml, java programming language and Microsoft .net framework. Unicode was developed in conjunction with Universal Character Set standard. The latest version of Unicode i.e. Unicode 6.1 includes 110,000 characters and 100 scripts [7]. Unicode has several character encoding forms represented as:

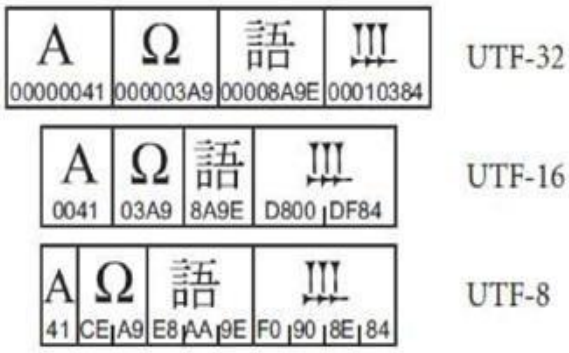

Figure 2: Unicode encoding forms

UTF-8: Uses one byte (8 bits) to encode English characters. It is widely used in HTML and similar protocols over Internet.
UTF-16: Uses two bytes (16 bits) to encode the most commonly used characters. Preferred usage in the case of storage optimization due to compact nature.
UTF-32: It is the simplest encoding form, which is capable of representing every character as one number. It is a preferred encoding form for Unix Platforms.

2.1 Code point

In general the meaning of code point is a particular integer that is being used to code the abstract character [7]. The standard understanding of code points in the Unicode Standard is to refer code point as their numeric value assigned in hexadecimal, with a "U+" prefix. For the example to encode the character in Latin, "P" is U+0050, "p" is U+0070, "6" is U+0036 and "?" is U+003F [7].

3. IMPACT OF COLOURS IN CRYPTOGRAPHY

A very wide range of colours is supported by the computer system. Further, the colours are greatly affected by viewing conditions. These conditions may vary based on the parameters including type of illumination, the amount of illumination, and the presence of other colours in the background [2, 8]. The colours that a system can display depend on the colour capability of the monitor and the graphic card used by the system. A form of colour encoding accepted among display devices manages colours in RGB format. A 24-bit RGB also called true colour encoding, can support a total of 16,777,216 different colours [9]. In addition to the components red (R), green (G) and blue (B) used in the RGB encoding, a transparency component alpha (A), can be added to increase this range of colours to a great extent. This colour encoding is known as ARGB encoding [10].These huge ranges of colours provide a large key domain and thus motivate the cryptographers to use colours for data hiding, steganography and encryption [11]. High resolution graphical images can easily hide data without much difference in the quality of image. A number of Image and visual cryptographic algorithms are developed but there are relatively fewer algorithms that transmit messages in form of colours. Motivated by the fact, we propose an algorithm that uses colour values in encrypted form to transmit messages and files.

4. WAVELET TRANSFORM

Wavelet transform is used to convert a spatial domain into frequency domain. The use of wavelet in image stenographic model lies in the fact that the wavelet transform clearly separates the high frequency and low frequency information on a pixel by pixel basis. Discrete Wavelet Transform (DWT) is preferred over Discrete Cosine Transforms (DCT) because image in low frequency at various levels can offer corresponding resolution needed. A one dimensional DWT is a repeated filter bank algorithm, and the input is convolved with high pass filter and a low pass filter. The result of latter convolution is smoothed version of the input, while the high frequency part is captured by the first convolution. The reconstruction involves a convolution with the synthesis filter and the results of this convolution are added. In two dimensional transform, first apply one step of the one dimensional transform to all rows and then repeat to all columns. This decomposition results into four classes or band coefficients [12]. The Haar Wavelet Transform is the simplest of all wavelet transform. In this the low frequency wavelet coefficient are generated by averaging the two pixel values and high frequency coefficients are generated by taking half of the difference of the same two pixels. The four bands obtained are approximate band (LL), Vertical Band (LH), Horizontal band (HL), and diagonal detail band (HH). The approximation band consists of low frequency wavelet coefficients, which contain significant part of the spatial domain image. The other bands also called as detail bands consists of high frequency coefficients, which contain the edge details of the spatial domain image. Research into human perception indicates that the retina of the eye splits an image into several frequency channels, each spanning a bandwidth of approximately one octave. The single in these channels is processed independently. Similarly in a multilevel decomposition, the image is separated into bands of approximately equal bandwidth on a logarithmic scale. It is therefore expected that use of the DWT will allow independent processing of the resulting components without significant perceptible interaction between them, and hence makes the process imperceptibility marking more effective [13]. For this reason the wavelet decompositions is commonly used for the fusion of images. Fusion technique include the simple method of pixel averaging to more complicated methods such as principal component analysis and wavelet transform fusion. Several approaches to image fusion can be distinguished; depending on whether the image is fused in the spatial domain or any other domains, and their transform fused. Image fusion is a process that produces a single image from a set of input images. The fused image contains more complete information, than any individual input. Since this is a sensor-compresses information problem, it follows that wavelets, classically useful for human visual processing, data compression and reconstruction are useful for such merging [14].Here the cover and payload are normalized and the wavelet coefficient is obtained by applying discrete wavelet transform. The approximation band coefficient of payload and wavelet coefficient of cover image are fused based on strength parameters alpha and beta. The capacity of the proposed algorithm is increased as the only approximation band of payload is considered.Other important applications of the fusion of images include medical imaging, microscopic imaging, remote sensing, computer vision & Robotics.

5. PROPOSED ALGORITHM

The proposed algorithm is based on the asymmetric key cryptography, mapping and fusion technique. The asymmetric key is produced at both ends which assigns security to original message. The mapping methods reduce the processing time of the algorithm and the mapping technique is used to perform the substitution, which substitutes a character with hexadecimal value of colour and overlapping of colour produces a coloured image which further fused with an image

to form an embedded fused image & further it is encoded to generate stego-image.so to hack the original message by eavesdroppers quite ambitious due to several keys generated in entire encryption as it replaced previous methodologies.

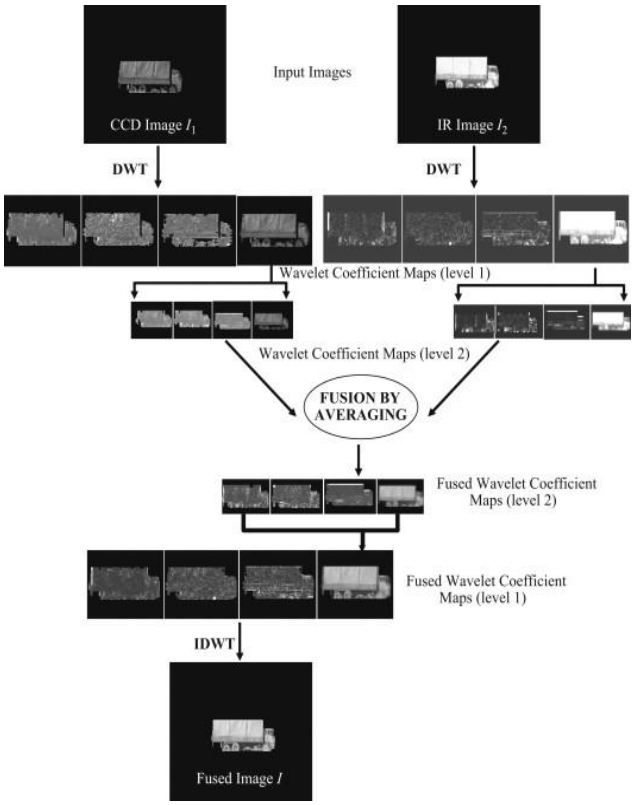

Fig 3: Discrete wavelet fusion using images

5.1 Pseudocode1

Generating Unicode values & mixes corresponding Colours
Input: Secret Message & two Images
Output: Final Stego object which contains Image as carrier & hidden fused image with Overlapped Colours & their Unicode Values.
Steps:
a. Input the secret message to be form cipher text.
b. Extract each character from the message.
c. Embed the character in unicode format.
d. Display the equivalent unicode Value.
e. Find out the Corresponding Colures from the Matching table.
f. Mix the colour with a Key which is in bit format.
f. Then generate a compressed fused image by using wavelet fusion technique with a new image.
g. Generate final stego object with a steganographic key, cover Image & the existing fused Object.

5.2 Pseudocode2

Input: stego object
Output: secret message in readable form
Steps:
a. Select the stego object.
b. extract the compressed fused image by the stego key.
c. Convert the compressed Image to a normal fused image.
d. Defuse it to get the encrypted colour Image.
e. The colour image goes for reverse mixing with the key provided by receiver side.
f. Read the unicode value from the mapping table.
g. For each unicode value, find the equivalent characters by decrypting the cipher text.
h. Repeat until all the cipher texts are converted into characters.
i. Accumulated characters to form the secret message.

Char	UNICODE	COLOR	Char	UNICODE	COLOR
a	U+0061		n	U+006E	
b	U+0062		o	U+006F	
c	U+0063		p	U+0070	
d	U+0064		q	U+0071	
e	U+0065		r	U+0072	
f	U+0066		s	U+0073	
g	U+0067		t	U+0074	
h	U+0068		u	U+0075	
i	U+0069		v	U+0076	
j	U+006A		w	U+0077	
k	U+006B		x	U+0078	
l	U+006C		y	U+0079	
m	U+006D		z	U+007A	

Fig 4: Sample mapping of char/symbol/digit unicode and colour

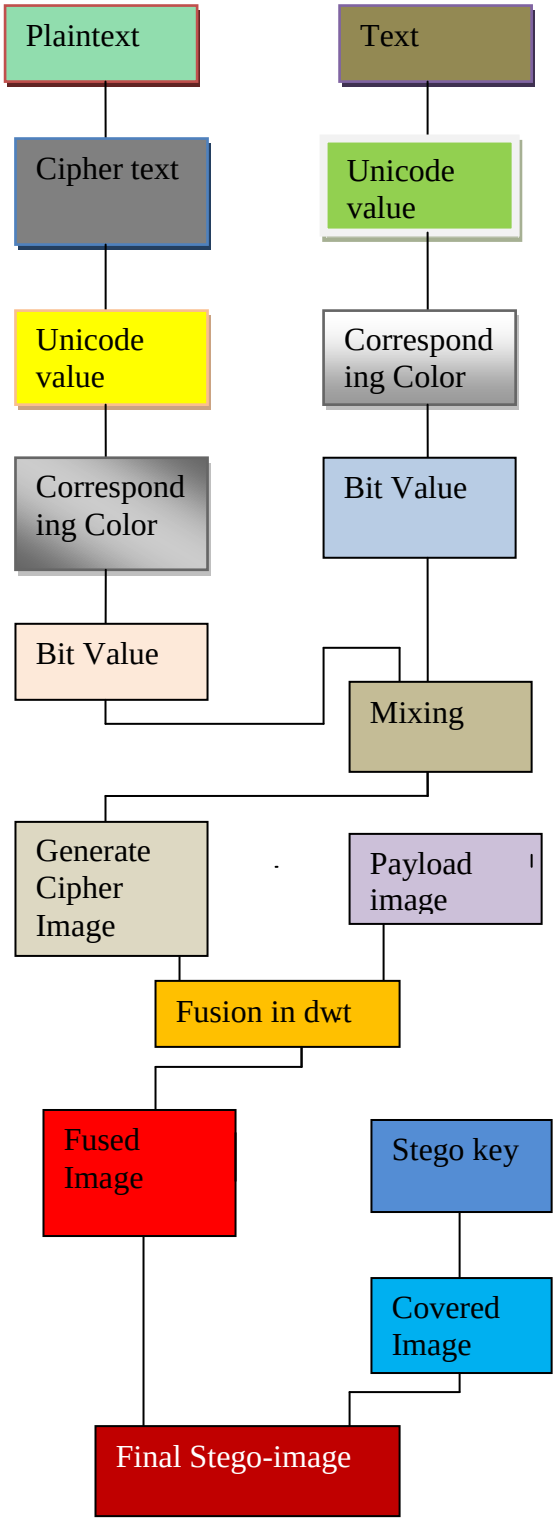

Fig 5: Encryption model

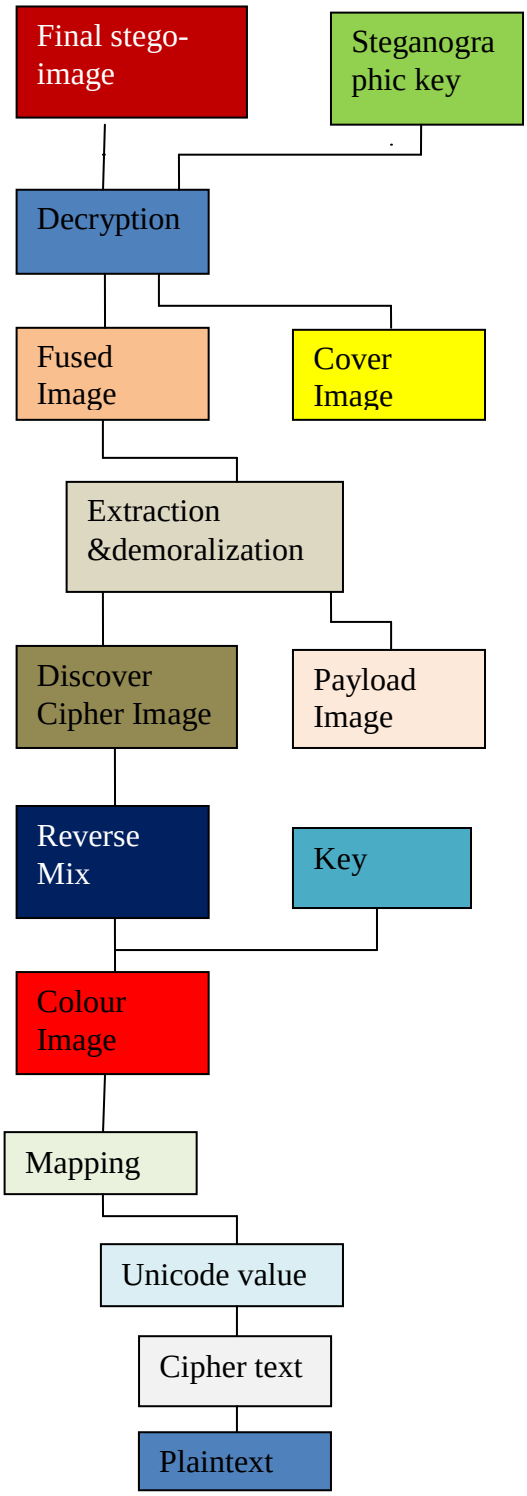

Fig. 6: Decryption model

5.3 Encryption

The proposed algorithm takes a file as input & it converted to cipher text & then it picks each character in the cipher text and finds the hexadecimal value of the colour assigned to the character by looking into the mapping table. The hexadecimal value is converted to binary value, further the binary value goes through mixing process using simple XOR and circular shift operations. After a certain number of iterations the resulting binary value is converted to ASCII value which is further overlapped with a key & the overlapped image further normalized with a image to form the fused image is hidden in Cover image with the help of steganographic key & forms a stego- image for increasing the complexity of security & it remains unintelligible to the unauthorized party.

5.4 Decryption

The hidden message from the stego object is extracted using the reverse process of the technique explained in Data Embedding. The stego object is given as an input to the software that allows simple recovery of the fused image and it demoralized to generate overlapped image. The receiver uses the shared key to generate the mapping table. finds the hexadecimal value of the colour assigned to the character by looking into the mapping table .The encrypted text at the receiver side goes through a reverse mixing process to generate original binary value; the binary value is then converted to equivalent hexadecimal value. The hexadecimal value is looked up in the mapping table to perform reverse mapping and generate the plaintext message.

6. EXPLANATION WITH EXAMPLE

Suppose Shyam has to send a message "**hello ram**" to Ram and the private key shared by them is "500001". Shyam creates a mapping table using this value. The colour generated by the value 500001 is assigned to 1st Unicode character, colour produced by 500002 to 2nd Unicode character and so on till the last Unicode character. The message "**hello ram** "is translated using mapping table as:

Character	Unicode	Color	Hexadecimal Code
h	U+0068		#C61333
e	U+0065		#C61330
l	U+006C		#C61337
l	U+006C		#C61337
o	U+006F		#C6133A
r	U+0072		#C6133D
a	U+006F		#C6132C
m	U+006D		#C61338

Fig 7: Conversion from unicode to colour

The Cipher text"einndkpo" &convert into Unicode values & map with the corresponding colour &the hexadecimal values of colours are converted to the corresponding binary values over which logical and mathematical functions are performed to transform the values to some other binary value. The ASCII code equivalent of each binary value is obtained and the cipher text "**#@%^*&!)**" & the Cipher image obtained & DWT & the stego-image generated using a key. In the Receiver side it performs reverse process to obtain Fused Image & then performs decryption to obtain binary value of colour. The binary value is converted to hexadecimal value which is searched in mapping table and colour, Unicode and character is obtained.

Fig 8: Carrier image

Fig 9: Stego image

7. EXPECTED OUTCOME

In this section, the complexity of algorithm is analysed. The proposed algorithm is efficient in terms of high security and transmission time [15]. The algorithm uses four levels of security. In the first level it assigns a colour to each Unicode character and maintains the information in the mapping table. In the second level it encrypts the colour value to special encrypted text. These substitutions and transformations provide security necessary to meet the requirements of a strong cryptographic algorithm .In the third level it fuse the cipher image into another image (key) so that it adds more security to the algorithm .In the fourth level it uses steganographic key with the cover image for higher complexity of the algorithm which will remain remains unintelligible to the third party. In context of transmission time, the algorithm convert the 32 bit ARGB value to 8 bit ASCII character which reduce the size of cipher text and thus reduces the transmission time over the network. Other advantages include use of simple operations that take less processing time; large range of colours used in the algorithm which prevent the algorithm from weak key attacks & provide better image quality with no pixel distortion [16]. An

evaluation model can also be developed to compare the existing algorithms with the proposed algorithm [17, 18, and 19]. The evaluation model should necessarily take care of key length used in each algorithm. The time consumption of algorithm is also an important factor and hence a feature of time consumption calculation should be incorporated in evaluation model [20]. The strength of the proposed model resides in the new concept of key image. Involving two values (the cover image and the key value) in place of only one (the cover) probably we will be able to change the cover coefficients randomly at the time of implementation of the proposed model. This opportunity does not give a steganoanalytic tool the chance of searching for a predictable set of modifications. The proposed approach has many applications in hiding and coding messages within standard Medias, such as images or videos. The Proposed encryption model will be suitable for many different applications:

a. Bulk encryption: The proposed model will be efficient in encrypting data files or a continuous data stream.

b. Random bit generation: The proposed model will be efficient in producing single random bits.

c. Packet encryption: The proposed model will be efficient in encrypting packet-sized data. It should implementable in an application where successive packets may be encrypted or decrypted with different keys.

8. CONCLUSION

In the field of security data hiding is the most important task. Cost of the security and efficiency will depend on the confidentiality and sensitivity of the data. So this type of data hiding, proposed model will be more secure. Some data strings e.g. password sending, a small information and costly data requires very much security, and will be use proposed model. When security, efficiency and cost is prime concerned rather others parameters then proposed model will be best suitable for use. Some images where all pixels intensities are equal, those will not good for uses of proposed model, so will be aware of it before use. Different type of statistics can be used to increase the capability of data hiding and correspondingly a level of security can be increase.

9. REFERENCES

[1] I Venkata Sai Manoj," Cryptography" AND Steganography" International Journal of Computer Applications, (0975-8887) Volume 1-No. 12©2010.

[2] Maram Balajee" Unicode & Color Integration Tool for Encryption & Decryption"International Journal on Computer Science and Engineering (IJCSE), ISSN: 0975, Vol.3No.3, Mar 2011.

[3] H S Majunatha Reddy, & K B Raja "High Capacity and Security Steganography Using Discrete Wavelet Transform", International Journal of Computer Science and Security (IJCSS), Volume (3), Issue (6).

[4] William stallings, ―Cryptography and Network Security: Principles & Practices‖, second edition.

[5] Luis von Ahn, Nicholas J. Hopper., Public-Key Steganography.

[6] Li Dongjiang, Wang Yandan, Chen Hong," The research on key generation in RSA public- key cryptosystem", Fourth International Conference on Computational and Information Sciences DOI 10.1109/ICCIS.348 2012.

[7] Web reference: http://www.unicode.org [As accessed on: 02-November- 2012]

[8] Misako Suwa," Color-Mixing Correction of Overlapped Colours in Scanner Images, "International Conference on Document Analysis and Recognition 1520-5363/11 $26.00 IEEE DOI10.1109/ICDAR.2011.52 © 2011.

[9] Web reference: http://cloford.com/resources/colours/500col.htm [As accessed on: 02-December-2012].

[10] Web reference: http://processing.org [As accessed on: 29-December- 2012].[11] Sujay Narayan and Gaurav Prasad," Two New Approaches for Secured Image Steganography Using Cryptographic Techniques and Type Conversions "Signal and Image Processing, An International Journal(SIPIJ),Vol.1,No.2,Dec 2010.

[12] Lisa M.Marvel and Charles T. Retter, "A Methodology for Data Hiding using Images," IEEE conference on Military communication, vol. 3, Issue. 18-21, pp. 1044-1047, 1998.

[13] LIU Tong, QIU Zheng-ding "A DWT-based colour Images Steganography Scheme" IEEE International Conference on Signal Processing, vol. 2, pp.1568-1571, 2002.

[14] Jessica Fridrich, Miroslav Goijan and David Soukal, "Higher-order statistical steganalysis of palette images" Proceeding of SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents, vol. 5020, pp. 178-190, 2003.

[15] Nameer N. EL-Emam,"Hiding a Large Amount of Data with High Security Using Steganography Algorithm, "Journal of Computer Science, ISSN 1549-3636 Vol.3 (4): 223-232, 2007.

[16] J. K. Mandal and Debashis Das," Colour Image Steganography Based on Pixel Value Differencing in Spatial Domain", International Journal of Information Sciences and Techniques (IJIST), Vol.2, No.4, July 2012.

[17] Aamar Nadeem and M Y Javed "A Performance Comparsion of data encryption algorithim,"IEEE Information and Communication Technologies, 2005 IEEE DOI 10.1109/ICICT.2005.1598 556.

[18]Neil F. Johnson and Sushi1 Jajodia Steganalysis,"The Investigation of Hidden Information," IEEE -7803-9914-5/98/$10.00 1998.

[19] Chi-Kwong Chan, L.M. Cheng,"Hiding data in images by simple LSB substitution,"Pattern Recognition Society. Published by Elsevier Ltd., doi:10.1016/j.patcog. 2003.

[20]Yan Wang and Ming Hu"Timing evaluation of the known Cryptographic algorithm", International Conferences on Computational Intelligence and Security 978-0-7695-7/09 $26.00©2009 IEEE DOI 10.1109/CIS.2009.81