Key Based Data Embedding Technique in Image Steganography

Vinaykumar M Kolli*(Author)*
Dept. of CSE *(7th Semester)*
R V College of Engineering
Bangalore-560098
vikolli@cisco.com

Vaishakh B N *(Author)*
Dept. of CSE *(7th Semester)*
R V College of Engineering
Bangalore-560098
Vaishakh.bargurreddy@gmail.com

Abstract— The Project Key Based Data Embedding Technique in Image Steganography", is aimed at developing a key based approach for image steganography. Steganography is the art and science of writing hidden messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message, a form of security through obscurity.

The project presents a novel method of data embedding in images by making use of a key. The key is used for internally embedding data in the image. An additional level of security is added by using the Advanced Encryption Standard (AES)for encrypting the data before embedding it.

Index Terms—Encryption, Decryption, Key, Pixel, Steganography .

1. Introduction

Since the rise of the Internet one of the most important factors of information technology and communication has been the security of information. Cryptography was created as a technique for securing the secrecy of communication and many different methods have been developed to encrypt and decrypt data in order to keep the message secret.

Unfortunately it is sometimes not enough to keep the contents of a message secret, it may also be necessary to keep the existence of the message secret. The technique used to implement this, is called steganography.

Steganography is the art and science of invisible communication. This is accomplished through hiding information in other information, thus hiding the existence of the communicated information. The word steganography is derived from the Greek words “stegos” meaning cover and “grafia” meaning writing defining it as covered writing. In image steganography the information is hidden exclusively in images. Almost all digital file formats can be used for steganography, but the formats that are more suitable are those with a high degree of redundancy. Redundancy can be defined as the bits of an object that provide accuracy far greater than necessary for the objects use and display. The redundant bits of an object are those bits that can be altered without the alteration being detected easily. Image and audio files especially comply with this requirement .

Given the proliferation of digital images, especially on the Internet, and given the large amount of redundant bits present in the digital representation of an image, images are the most popular cover objects for steganography.

The basic structure of Steganography is made up of three components: the carrier, the message, and the key. The carrier can be a painting, a digital image, an mp3, even a TCP/IP packet among other things. It is the object that will carry the hidden message.

A key is used to decode/decipher/discover the hidden message. In this project, the focus is only on Image Steganography.

2. Evaluation Criteria

In order to reasonably evaluate the performance of various kinds of steganography it is necessary to define some criteria acceptable to the majority. Moreover, the evaluation criteria may also lead us to the right direction to improve the techniques. Three common requirements security, capacity, and imperceptibility, may be used to rate the performance of steganographic techniques.

- **Security**

Steganography may super from many active or passive attacks, correspondingly in the prisoner's problem when Wendy acts as an active or passive warden. If the existence of the secret message can only be estimated with a probability not higher than random guessing in the presence of some steganalytic systems, steganography may be considered secure under such steganalytic systems. Otherwise it may be claimed as insecure.

- **Capacity**

To be useful in conveying secret message, the hiding capacity provided by steganography should be as high as possible, which may be given in absolute measurement (such as the size of secret message), or in relative value (called data embedding rate, such bits per pixel, bits per non-zero discrete cosine transform coefficient, or the ratio of the secret message to the cover medium, etc.).

- **Imperceptibility**

Stego images should not have severe visual artifacts. Under the same level of security and capacity, the higher the fidelity of the stego image, the better. If the resultant stego image appears innocuous enough, one can believe this requirement to be satisfied well for the warden not having the original cover image to compare.

3. Review of Literature

In this section, a review of some of the existing and in use methods for steganography is explained along with their advantages and pitfalls.

3.1. LSB Substitution Method :

The most well-known steganographic technique in the data hiding held is least-significant bits (LSBs) substitution[1]. This method embeds the fixed-length secret bits in the same fixed length LSBs of pixels. Although this technique is simple, it generally causes noticeable distortion when the number of embedded bits for each pixel exceeds three. Several adaptive methods for steganography have been proposed to reduce the distortion caused by LSBs substitution. For example, adaptive methods vary the number of embedded bits in each pixel, and they possess better image quality than other methods using only simple LSBs substitution. However, this is achieved at the cost of a reduction in the embedding capacity.

3.2. Inverted Pattern Approach :

This inverted pattern (IP) LSB substitution approach[2] uses the idea of processing secret messages prior to embedding. In this method each section of secret images is determined to be inverted or not inverted before it is embedded. In addition, the bits which are used to record the transformation are treated as secret keys or extra data to be re-embedded.

3.3. The Modulo Method :

In this method embedding is done by subtracting any remainder obtained by dividing with 10 and adding the data to be hidden. This is demonstrated by let the pixel be p and data be less than 10 say d then new pixel formed is p1=p- remainder (p/10) +d.

Similarly mod100 method divides the pixel by 100, remainder obtained is subtracted and the data is added to it to get the new pixel. The data hidden will simply be equal to the remainder obtained by dividing the new pixel by 10 or 100 accordingly.

3.4. IP Method using Relative Entropy :

Relative entropy measures the information discrepancy between two different sources with an optimal threshold obtained by minimizing relative entropy [2]. In this method instead of finding the mean square error for inverted pattern approach, the relative entropy is calculated to decide whether S or S suites the pixel. In probability theory and information theory, the Kullback Leibler divergence (also information divergence, information gain, or relative entropy) is a non-symmetric measure of the difference between two probability distributions P and Q.

3.5. DCT (Discrete Cosine Transform)

In this method, a transform domain technique, DCT[3] is used to hide messages in significant areas of the cover image. Here pixels are split into 88 blocks. Then, all blocks are DCT transformed each block encodes exactly one secret message bit.

4. Overview of Key Based Image Steganography.

In a recent investigation [4], the key is consisting of 256 numbers (0 to 255) arranged in a random manner. A given plaintext, converted into decimal numbers by using EBCDIC code, is permuted by using the key. Further this is modified by XORing with the key. The plaintext obtained in this manner is hidden in the image. In this analysis, the process of concealment is totally guided by the key and the modified plaintext under consideration is placed in different columns of the image depending upon the key. In this case, the plaintext under consideration occupies completely four consecutive rows of the image. In this process, the last two binary bits of each pixel value, in the columns, are replaced by an appropriate pair of binary bits of the numbers corresponding to the modified plaintext. As shown in the analysis, this process can be applied for Steganography of 64 plaintexts at the most.

Using the literature surveyed above as the basis, in this project, the approach uses a key based approach for Image Steganography. We first generate a key by using Diffie Hellman Key Exchange. The data file is encrypted using Advanced Encryption Algorithm (AES).

The key is used to determine at which position the data has to be embedded. The key determines the bits which are to be replaced. This is the embedding technique. On the receivers end, the key is generated again by using the same prime number and hence, the same key is generated. This key is used to perform the inverse of the process done on the sender's end.

5. High Level Design of the Project

This section provides an overview of how the functionality and the working of the data embedding techniques in image steganography using a key based approach. The overall functionality of the application is divided into different modules in an efficient way. The system architecture is shown in Figure 1.

The user chooses a cover image and the file to be hidden. Based on the key generated by using DH Key Exchange mechanism, steganography takes place. The altered pixels are embedded into the result, the morphed image. This is transmitted across a network.

In the receiver side, the same key is generated and the hidden file is recovered.

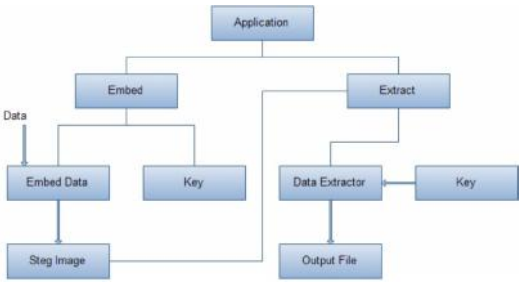

Fig 1.High Level Design

The entire program starts with user entering his choice of input as to either embed or extract. Based on this, the GUI module responds appropriately with success or failure. The digital media chosen by user is taken as input for embedding phase. The digital media is encrypted using AES algorithm before embedding to stego image. This encrypted data is embedded onto image using embedding algorithm to get the stego image containing secret data. A key is generated using Diffe Hellman Key exchange which acts as the input for the embedding algorithm.

At the receiver end, the Diffie Hellman key is generated once again. From the stego images, secret data is extracted. It will be in encrypted form. The original input data is extracted by using decompression module of AES algorithm. This gives the hidden file's original content.

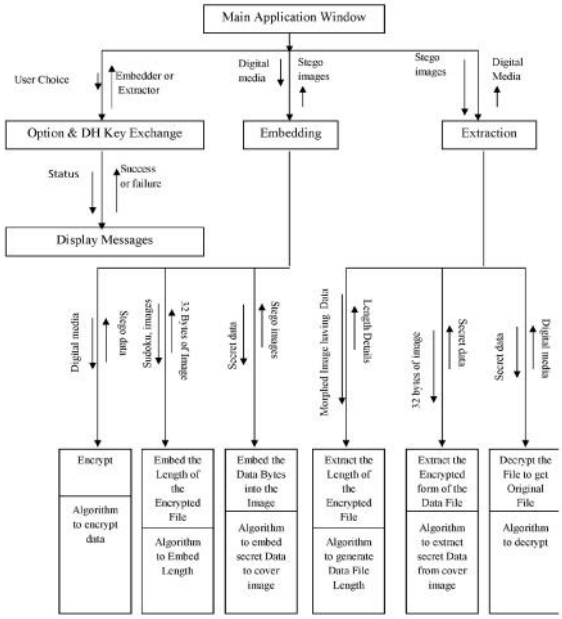

Fig 2. Structure Chart

5.1. Session Key Generation Module :

Definition :
This module generates a session key used for embedding using the principles of Diffie Hellman protocol.

Input Resources :
Takes two global parameters Q and alpha
Output Resources :
It returns a session key k.
Functionality :
Using the parameters Q and alpha, calculations are performed on these and finally a key is generated which is used for embedding the data.

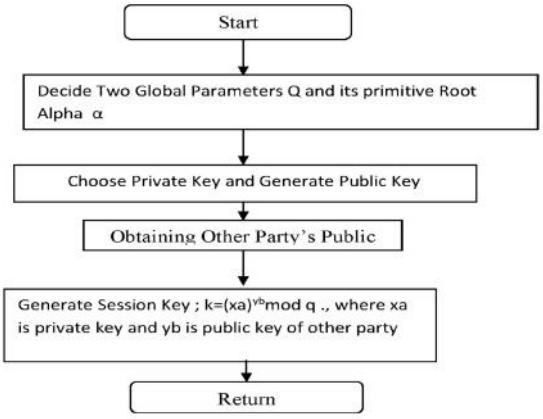

Fig 3. Session Key Generator Module

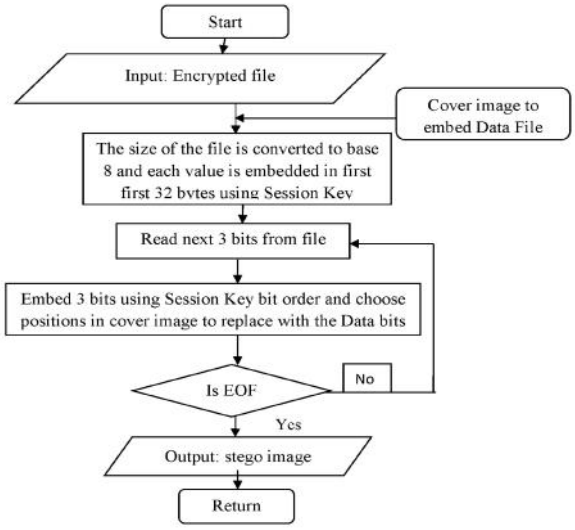

Fig 4. Embedding Algorithm Flow Chart

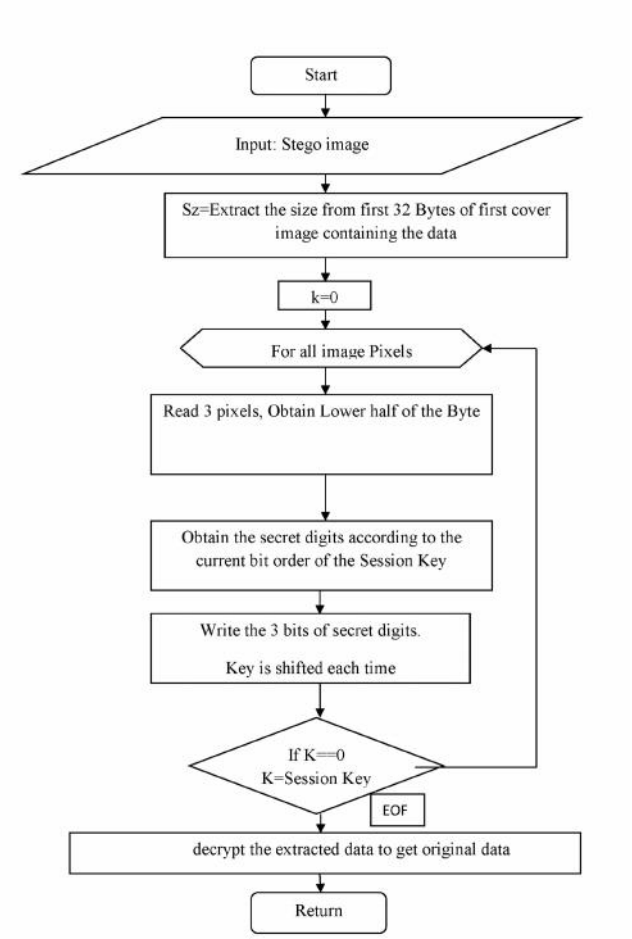

Fig 5. Data Extraction Module

6. DIFFICULTIES ENCOUNTERED AND STRATEGIES USED TO TACKLE

There were a number of challenges that were faced while implementing the steganography using Eclipse and SWT. Some challenges were challenging and ended up in helping us think innovatively and come up with efficient solutions. Some major problems that were encountered have been stated in brief along with their solutions.

- Initially the software could embed only text and other unformatted types of files.
  Solution : This Problem was solved later by using special ways to read the formatted files (pdf's, audio files).

- Initially the 10 digit key was being entered by the user could not be transmitted to the other end.
  Solution : Diffie Hellman key exchange module was later developed which helped in key distribution.

- Initially it was difficult to view both the morphed and the original image at the same time.
  Solution : This problem was later solved by developing a module in SWT to compare the two images side by side.

7. RESULTS AND ANALYSIS :

While designing image steganography, main factors to be considered are

- Embedding data securely in a cover image.
- Number bits to be embedded per pixel.
- Less distortion in embedded cover image compared to original cover image.

In the proposed system first condition is met by using AES encryption technique specific to input media _les. The bits embedded per pixel are decided based on the key. This is an improvement over LSB. Also, the distortion is less when compared to LSB.

7.1. PSNR Ratio for Images :

PSNR Ratio analysis is done to analyze the quality of images. The phrase peak Signal to-Noise Ratio, often abbreviated PSNR, is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale.

The PSNR is most commonly used as a measure of quality of reconstruction of lossy compression codecs (e.g., or image steganography). The signal in this case is the original data, and the noise is the error introduced by steganography. A higher PSNR value indicates a higher quality of the teganographed image. It is most easily defined via the mean squared error (MSE). Given a noise-free mn monochrome image I and its noisy approximation K, MSE is defined as: The PSNR is defined as:

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

Mean Square Error (MSE)

$$\begin{aligned} PSNR &= 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \\ &= 20 \cdot \log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \\ &= 20 \cdot \log_{10}(MAX_I) - 10 \cdot \log_{10}(MSE) \end{aligned}$$

Peak Signal to Noise Ratio (PSNR)

8. OUTPUT OBTAINED:


Before Steganography


After Steganography

9. CONCLUSION :

The project Key Based Data Embedding Technique in Image Steganography" is an information hiding method proposed to improve the robustness of image steganography by using a key based approach which is not likely to be LSB at all times.

Also, the data embedded is encrypted using AES encryption algorithm. The proposed method can hide any kind of data. The process supports hiding PDFs, text _les, formatted word documents and images.

9.1. Future Enhancements:

Any development is a continuous process, which does not terminate once an executable has been generated. Every project has certain limitations due to various reasons. Some of these limitations could be overcome, given sufficient time while others can be overcome with better technologies. Hence, some of the future enhancements that could be performed on these steganographic systems are as follows:

Special compression techniques for different types of files can be added to convert them internally before performing steganography. A web application maybe developed thus helping naive users to use it as a web service.

10. REFERENCES

[1] R. J. Anderson and Fabien A. P. Petitcolas, On the limits of steganography, IEEE Journal on Selected Areas in Communications, vol. 16, no. 4, pp. 474-481, 1998.

[2] Birgit P tzmann, Information hiding terminology-results of an informal plenary meeting and additional proposals, Proc. of the First International Workshop on Information Hiding, vol. 1174, pp.347-350. Springer, 1996.

[3] Huaiqing Wang and Shuozhong Wang, Cyber warfare: Steganography vs. steganalysis, Communications of the ACM, vol. 47, no. 10, pp. 76-82, 2004.

[4] V.U.K.Sastry, Ch.Samson., Key based Steganography in a gray level image involve ing permutation and modular arithmetic addition . Volume 2, No. 6, June 2011 Journal of Global Research in Computer Science.

[5] Niels Provos and Peter Honeyman, Hide and seek: An introduction to steganography, IEEE Security and Privacy, vol. 1, no.3, pp. 32-44, 2003.

[6] R. Chandramouli, M. Kharrazi, and N. Memon, Image steganography and steganalysis concepts andpractice, Proc. of IWDW'03, vol. 2939, pp. 35-49, Springer, 2003.