DATA HIDING USING CRYPTOGRAPHY AND STEGANOGRAPHY TECHNIQUES

Mr. B.V. Sathish Kumar [1]
Manideepika Bathena [2], Ratna kumari Chilaka [3], Lavanya Valli Damayanthi Bondada [4], Bhavya Lakshmi Maheswari Cherukumalli [5]

[1] *Assistant Professor, Department of ECE, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India*
[2-5] *Undergraduate Students, Department of ECE, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, Andhra Pradesh, India*

ABSTRACT

In this world of keeping information safe and secure, we need strong methods. This paper introduces an innovative method, RSA-LSB, that combines RSA cryptography and LSB (Least Significant Bit) steganography to enhance data security. RSA cryptography is used to lock up the data, making it secure with special key-based system. This extra layer of protection stops unauthorized access and makes sure the hidden info stays secret. RSA's complexity, relying on big prime numbers makes the encryption strong. At the same time, LSB steganography hides the encrypted data in images, audio, or videos without any changes in the audio or the visuals. By replacing the least important bits of pixels or audio samples, the secret info gets encoded, while the original look and sound stay intact. This process also includes the compression of the cover image with lossy compression techniques to reduce the image information. Combining RSA cryptography and LSB steganography not only makes hidden data more secure but also opens new possibilities for exploring cryptographic steganography.

Keyword: - *RSA(Rivest-Shamir-Adelman) cryptography, LSB (Least Significant Bit) Steganography, Stego image, Ciphertext, Encryption, Decryption, Embedding, Extracting, DWT compression, MSE, PSNR, SSIM.*

1. INTRODUCTION

When it comes to digital security, the combination of RSA cryptography with LSB steganography in the MATLAB environment offers a powerful way to hide private information from pictures. Information security is achieved using asymmetric key pairs in RSA cryptography, which is well known for its strong encryption powers. To ensure secrecy and integrity, the data is encrypted using the public key, which can only be unlocked with the matching private key. In addition, LSB steganography hides the information while slightly changing the appearance of the image by inserting encrypted data into the least significant pixels. This hybrid method offers a potent mechanism for safe data concealing by fusing the strength of RSA encryption with the subtlety of LSB steganography.

RSA cryptography and LSB steganography combined with Discrete Wavelet Transform (DWT) image compression create a powerful method, by breaking down an image into its frequency components. DWT allows for effective image compression that reduces data storage needs without significantly sacrificing visual quality. This compressed image becomes a perfect carrier for hiding sensitive data in the context of data hiding. MATLAB's extensive image processing and cryptography functionalities make it the perfect platform for implementing this strategy. The product is visually unchanged photographs that hide encrypted data within their pixels, making it difficult for third parties to access or discover. In response to the growing requirement for strong information security in today's digital landscape, this integration of RSA cryptography and LSB steganography provides a flexible solution for secure communication, digital watermarking, and covert data transmission.

2. EXISTING SYSTEM CLASSIFICATION METHODS

The art of concealing data through encryption and revealing it through decryption is known as cryptography. Information confidentiality, integrity, and authenticity are all maintained using cryptography. The Greek word for steganography is "covered writing." The practice of hiding information's existence within seemingly innocent carriers is known as steganography.

There exist different techniques to perform data hiding in image processing and some works that already exist are “Improved File Security System Using Multiple Image Steganography”, “Implementation of AES Algorithm on Text and Image using MATLAB”, “Research and implementation of RSA algorithm for encryption and decryption”, “Bit-Plane Slicing Algorithm for Crime Data Security using Fusion Technologies” and many known algorithms can be utilized in cryptography. They include Advanced Encryption Standard (AES), Blow fish, Data Encryption Standard (DES), RC4 Rivest-Shamir-Adleman (RSA).

2.1 LIMITATIONS OF EXISTING SYSTEMS

The existing systems each has few of the drawbacks in each case. In the case of AES algorithm, it is more secure, faster with higher efficiency but complex in nature. When we consider blowfish algorithm it is only applicable where the key does not change frequently, like a communication link or an automatic file encryptor.
Watermarking provides image security at different level and has high embedding capacity where as its entropy value is not very high in general. The image stitching technique is involved of noise, FAST based approach has poor accuracy rate.

LSB based steganography, slicing method is limited only to 24-bit color depth file format. K-LSB based techniques, local entropy filter combinedly has a less probability of distortion and loss of information but it has a disadvantage as if you make slight modification then it can destroy the hidden information.
The combination of RSA cryptography and LSB steganography along with DWT image compression can improve the security and it has good embedding capacity so that it does not show major changes in the image before and after embedding the message.

3. PROPOSED SYSTEM

This paper proposes the combined form of RSA cryptography with LSB steganography which includes DWT image compression for easy transmission of data. Reducing the portions of an image that the user is not primarily interested in is the fundamental goal of image compression. This makes the image more suitable for storing and sending data over a variety of communication channels while also decreasing the size of the image by decreasing the number of pixels. The fundamental concept of the suggested algorithm is depicted in the above Figure and demonstrates how confidential information can be sent from the sender to the recipient.

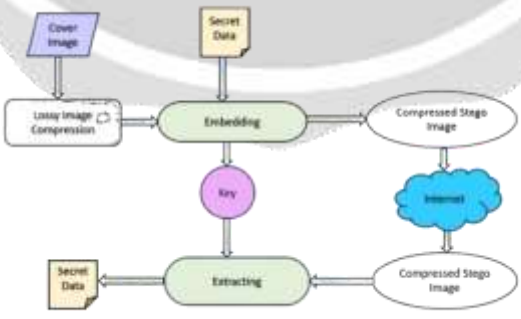

Fig-1: Basic block diagram

3.1 BASIC MODEL

The secret message is encrypted using RSA cryptography, the cover-image is compressed by the DWT algorithm and then the cover-image is combined with the secret message via LSB and sent them over the Internet to the destination as a compressed file. These encoded streams (bits) are then sent to a decoder that decodes these streams (bits) and the final output image is retrieved as a decoding file output. Lossy and lossless image decompression is used to reduce the number of bits required to represent an image.

3.2 RSA CRYPTOGRAPHY

This process includes the encryption of the secret message before embedding in the image. This RSA encryption uses a key pair (private and public keys) to encrypt the message to be hidden. Cryptographic algorithms are broadly classified into two types, namely, symmetric key algorithms and asymmetric key algorithms. Symmetric encryption uses the key for both encryption and decryption. In the domain of asymmetric key encryption, the bedrock of secure communication is established through a pair of distinct keys: the public key, widely shared with potential senders, and the private key, a closely guarded secret known exclusively to the designated recipient. This two-key model defines asymmetric systems, where data encrypted using the public key can only be unraveled by its corresponding private key.

One of the most prominent implementations of asymmetric key cryptography is the RSA algorithm. RSA, standing for Rivest-Shamir-Adleman, is a widely used asymmetric key encryption standard that plays a crucial role in securing data where confidentiality is paramount. Its security is founded on the intricate analysis of many compounds and the complex computation of the unit of moral roots for a specified odd integer (e). The RSA public key is represented by an integer pair (n, e), where 'n' is typically the result of multiplying two prime numbers. For this key pair generation, we use the Extended Euclidean Algorithm.
Here is an overview of the RSA key pair generation process, emphasizing the use of the Extended Euclidean Algorithm:

1. **Key Pair Generation**:
   a. Select Two Large Prime Numbers p and q, and calculate the modulus ($N = p * q$).
   b. To calculate Euler's Totient Function $\varphi(N)$, Compute $\varphi(N) = (p-1) * (q-1)$.
   c. Select a public exponent (e) such that $(1 < e < \varphi(N))$ and (e) is coprime to $(\varphi(N))$.
   d. Use the Extended Euclidean Algorithm to find (d), the modular multiplicative inverse of (e) modulo $(\varphi(N))$.
   The algorithm yields coefficients (x) and (y) such that $(ex + \varphi(N)y = 1)$, and (d) is the value of (x).

2. **Extended Euclidean Algorithm:**
   a. Initialization:
   - Start with the equation $(ex + \varphi(N)y = 1)$.
   - Initialize $(a = e)$, $(b = \varphi(N))$, $(x_0 = 1)$, $(y_0 = 0)$, $(x_1 = 0)$, $(y_1 = 1)$.
   b. Iterative Steps:
   - Use the Euclidean Algorithm to find the greatest common divisor GCD of a and b.
   - Update q as the quotient and r as the remainder: $(a = bq + r)$.
   - Update a, b, x_0, y_0, x_1, and y_1 accordingly.
   c. Termination:
   - Continue the iterations until b becomes 0.
   - The coefficients x and y from the last non-zero remainder step provide the solution to $ex + \varphi(N)y = 1$, and x is the modular multiplicative inverse of e modulo $(\varphi(N))$.
   d. Private Exponent Calculation, $d = x \bmod \varphi(N)$.

3. **Public and Private Key Pairs:**
   - The public key is (N, e) and is distributed openly.
   - The private key is (N, d) and must be kept secret.

The use of the Extended Euclidean Algorithm ensures the efficient and accurate calculation of the private exponent, which is a critical component of the RSA cryptosystem.

3.3 DISCRETE WAVELET TRANSFORM (DWT LOSSY DATA COMPRESSION)

The discrete wavelet transform (DWT) stands out as a potent technique in image processing, utilizing wavelets to efficiently convert images into a series of wavelet coefficients, more space-efficient than traditional pixel blocks. In one dimension, signals undergo division into high and low frequencies using low-pass and high-pass filters for DWT. This method is crucial for image compression, ensuring minimal loss of information and maintaining a lossless image compression approach. DWT is founded on a time-scale representation, providing multiresolution capabilities, making it preferable for signal compression over conventional methods. It is recognized as a valuable computational tool applicable in diverse processing applications. Particularly in image processing, wavelet transformations, such as DWT, excel in minimizing noise and blurring.

In DWT, various filters can be employed, with the Haar filter being the simplest and widely used. The 2D implementation of DWT involves dividing the image into four sub-bands: LL–LH–HL–HH. The coefficients of each sub-band are calculated using the Haar filter equations:

$$LL(x, y) = \frac{[\,p(x, y) + p(x, y+1) + p(x+1, y) + p(x+1, y+1)]}{2}$$

$$LH(x, y) = \frac{\left[p(x, y) + p(x, y+1) - p(x+1, y) - p(x+1, y+1)\right]}{2}$$

$$HL(x, y) = \frac{\left[p(x, y) - p(x, y+1) + p(x+1, y) - p(x+1, y+1)\right]}{2}$$

$$HH(x, y) = \frac{\left[p(x, y) - p(x, y+1) - p(x+1, y) + p(x+1, y+1)\right]}{2}$$

Here, (p) represents the image pixel, (x) is the row number, and (y) is the column number. This process yields four sub-bands (LL1, LH1, HL1, and HH1) during the Level 1 decomposition. Subsequently, a similar procedure is repeated in the sub-bands, such as LL1, to obtain LL2, LH2, HL2, HH2, and so forth. This hierarchical decomposition enables effective representation and compression of image data.

3.4 LSB STEGANOGRAPHY

Steganography, a method of covert communication, encompasses four distinct domains, each employing unique techniques to conceal confidential information within cover media:

1. Spatial Domain:
In this domain, confidential message bits are directly embedded within the bits of the cover media, often utilizing simple algorithms like LSB (Least Significant Bit). While imperceptible to the human eye, LSB alterations may be detectable through statistical tests. However, LSB methods have limitations, including a small message size, susceptibility to damage during cover media compression, and vulnerability to information loss with minimal changes to the media.

2. Transform Domain:
Transformation techniques such as Discrete Cosine Transform (DCT), Discrete Wave Switching (DWT), and Discrete Fourier Transform (DFT) are utilized. DCT divides the cover media into 8×8 blocks, quantifies them with a quantization table, and embeds sensitive bits within each block. DWT divides the cover media into sub-bands, ensuring that embedding in the primary sub-band (LL) retains message integrity despite compression. DFT converts input signal points into output pairs, offering a robust concealment method.

3. Spread Spectrum:
This method embeds the secret message within the noise of the cover media generated during image acquisition. Utilizing a blind scheme, it offers a payload capacity while maintaining cover media fidelity.

4. Model-Based:
Cover media is divided into two parts, with the first part remaining unchanged during embedding. The secret message is embedded in the second part without altering the statistical properties of the cover media. Notably, this method boasts high modulation capacity and resistance against various attacks.

Each domain presents unique advantages and challenges. While transform domain techniques offer increased message size and resistance to attacks, they suffer from longer embedding and extraction times. Conversely, spatial domain methods like LSB are fast and simple but are limited in terms of message size and susceptibility to compression-induced damage. Ultimately, the selection of steganographic method depends on factors such as desired message size, security requirements, and computational resources.

4. RESULTS & DISCUSSIONS

In modern digital environments, ensuring the security and confidentiality of sensitive information is paramount. The model's approach to data protection represents a sophisticated synthesis of various techniques aimed at fortifying data security throughout its lifecycle. Beginning with the RSA encryption algorithm, data is rendered unintelligible to unauthorized parties, with a public key for encryption and a private key for decryption ensuring that only authorized users can access the original information. Following encryption, the application of lossy compression techniques through the Discrete Wavelet Transform (DWT) optimizes storage and transmission efficiency while preserving essential image details by decomposing the cover image into its frequency components.

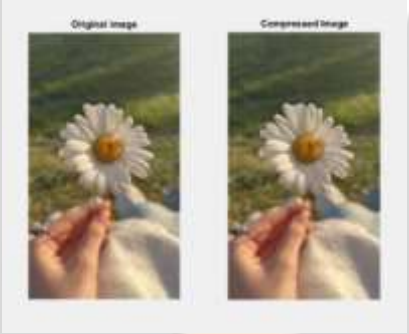

FIG-3: DWT Compressed Image

Subsequently, steganography techniques come into play, seamlessly embedding encrypted data within the cover image. Methods like Least Significant Bit (LSB) steganography subtly modify pixel values, ensuring the hidden information remains imperceptible to human observers while safeguarding its confidentiality and integrity. Once embedded, further compression using lossless methods is applied to the stego image, preserving its visual quality and fidelity while enhancing storage and transmission efficiency. This comprehensive process culminates in a robust security framework where authorized recipients can extract the encrypted data using steganography techniques and decrypt it using the corresponding private key, thereby restoring the original information to its readable format.

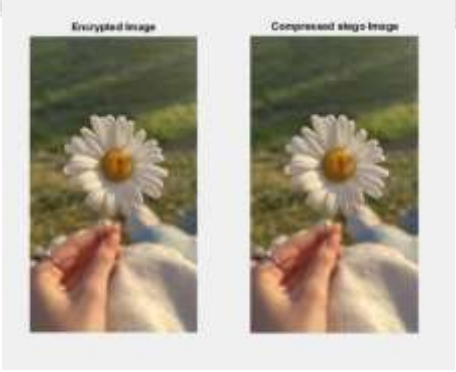

FIG-4: Compressed Stego Image

FIG-5: Output for RSA cryptography, MSE, PSNR, SSIM values

Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM) are all metrics commonly used to evaluate the quality of images. Each of them serves a specific purpose and provides different insights into the fidelity of the reconstructed or compressed image compared to the original.

Mean Squared Error (MSE):
MSE is a simple and widely used metric for measuring the average squared difference between the original image and its reconstructed version. It is calculated as the average of the squared differences between corresponding pixels in the original and the processed images.

$$MSN = \frac{1}{N}\sum_{x=0}^{W-1}\sum_{y=0}^{H-1}((I_1(x,y) - I_2(x,y))^2$$

Where $I_1(x,y)$ and $I_2(x,y)$ are the intensity values of the images, N is the total number of pixels in the image (W width multiplied by H height).

Peak Signal-to-Noise Ratio (PSNR):
PSNR is a metric that measures the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. In image processing, PSNR is used to measure the quality of the compressed image by comparing it to the original image.

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

Where MAX is the maximum possible pixel value of the images (typically 255 for 8-bit images), MSE is the mean square error between the two images.

Structural Similarity Index (SSIM):
SSIM is a metric that quantifies the similarity between two images. It considers luminance, contrast, and structure, and is more aligned with human perception compared to metrics like MSE or PSNR.
SSIM is calculated using several components, including luminance similarity, contrast similarity, and structural similarity. The SSIM index ranges from -1 to 1, where 1 indicates perfect similarity between images.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Where μ_x, μ_y are the mean values of x and y respectively,

σ_x^2, σ_y^2 are the variances of x and y respectively,

σ_{xy} is the co-variance between x and y,

c_1, c_2 are the constants to stabilize the division with weak denominator.

From the above outputs we can tell that the given original image is compressed and then the secret message is embedded into the compressed image through LSB steganography to give a compressed stego image as output. This obtained is like the original image with minor changes in it.

FIG-6: Compressed Stego image for a black & white picture

In summary, the model's approach integrates encryption, compression, steganography, and decryption techniques into a cohesive strategy that ensures comprehensive protection for sensitive data. By employing these methods in concert, the model not only secures data against unauthorized access but also maintains its integrity during storage and transmission, thereby instilling confidence in data handling processes across various domains.

REFERENCES

[1]. Elharrouss, Omar, Noor Almaadeed, and Somaya Al-Maadeed. "An image steganography approach based on k-least significant bits (k-LSB)." 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT). IEEE, 2020.

[2]. Ramya, G., P. P. Janarthanan, and D. Mohanapriya. "Steganography Based Data Hiding for Security Applications." 2018 International Conference on Intelligent Computing and Communication for Smart World (I2C2SW). IEEE, 2018.

[3]. Aparna, V. S., et al. "Implementation of AES Algorithm on Text and Image using MATLAB." 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI). IEEE, 2019.

[4]. Zhou, Xin, and Xiaofei Tang. "Research and implementation of RSA algorithm for encryption and decryption." Proceedings of 2011 6th international forum on strategic technology. Vol. 2. IEEE, 2011.

[5]. Al-Haj, Ali, and Hiba Abdel-Nabi. "Digital image security based on data hiding and cryptography." 2017 3rd International conference on information management (ICIM). IEEE, 2017.