A Novel Scheme for Image Authentication and Secret Data Sharing

Auqib Hamid Lone, Ab Waheed Lone and Prof. Moin Uddin
Department of Computer Science & Engineering, Jamia Hamdard, New Delhi
E-mail: {auqib92, waheedlone9}@gmail.com, {Prof.moinuddin}@jamiahamdard.ac.in

Abstract—Privacy protection is of extreme importance especially in security sensitive environments. Neither cryptography nor steganography comes up with ultimate solution for privacy preservation in open systems. However, combination of steganography and cryptography can greatly increase the security of communication and is usually considered a good practice for securing security driven communication environments. In this paper we propose a novel scheme for image authentication and secret data sharing based on three level security model viz: Compression, steganography and cryptography. Compression optimizes the resource usage, steganography conceals the existence of message and cryptography is used to achieve confidentiality and authentication.

Index Terms—Authentication, Data Hiding, LSB method, Image encryption, Generalized RSA.

I. INTRODUCTION

In Today's digital world, communication is lifeline. Being sure about what we receive is from authentic source (Authentication) and authorized access of secret information (Confidentiality) are of prime concern, particularly in privacy and security sensitive communication environment. There are several alternatives to implement such goals among which cryptography and steganography are most general and widely prevalent techniques. Cryptography is a mathematical art of secret writing which scrambles the message making it unreadable while steganography hides the data making data unobservable. The combination of two greatly increases the security of digital assets like data, images, audios, videos, logos etc. Most prevailing crypto-steganography approaches use symmetric key encryption for securing the hidden data but major limitation of such approach is key management. Secret key is required to be shared between sender and receiver making it susceptible to various attacks. Most of the previous approaches focused more on confidentiality of data but less on verifying the originality of source.

With a high degree of redundancy in digital file formats (images, audios, videos) makes them known for being used for steganography in line with compression techniques. Using image compression reduces irrelevance and redundancy of the image data enabling to store or transmit data in efficient form. Stenographic techniques like LSB, Transform domain techniques, Statistical methods, and Distortion techniques enable one to hide a secret message in an image file. Two main classificatory schemes used by modern stenographic categories for the taxonomy of algorithms are distinguished algorithms based on file type and then embedding technique employed. Cryptography in wired and wireless networks is making a new way around for authentication, confidentiality, non-repudiation and integrity aspects of participants and messages.

In this paper we introduced hybrid approach based upon compression (Run-length encoding), steganography(LSB) and cryptography (public key cryptography (Generalized RSA-AA) [1]) providing multiple layers of security. GRSA-AA includes novel and secure algorithm for key generation making it more secure than different existing variants of RSA algorithm. Moreover, in GRSA-AA the public key component n can be multiplication of 2,4, 8,…,2^k-2 prime numbers thus give flexibility to users in terms of size of message or image that is being encrypted. The idea of implementing multiple layers of security is to make the communication in such a way that no one can detect secret message inside stego-object and stego-object itself confirming the originality of source of message removing the major limitation of previous approaches.

Further sections of this paper are organized in the following way: Section 2 presents brief literature review. Section 3 presents the proposed model. In section 4 we present implementation results and in last section conclusion of work is presented.

II. LITERATURE REVIEW

In order to have complete, clear cut, unbiased, and broader prospective many sources have been explored. The Literature Review has been carried out according to the guidelines proposed by Kitchenham [2]. The systematic literature review has been carried in the following databases 1. IEEE Xplore 2. ACM digital Library 3. Science Direct 4. Wiley online Library 5. Springer The reason behind exploring these databases is their rich library of journals with high impact factors. The review also takes into account conference proceedings. The search term was 'Image Authentication and Secret Data Sharing'. The search was filtered to include the papers and conferences of previous 5 years. This was done to limit the scope of research to the present trends

instead of exploring unverified and undeveloped techniques. Our Search results zero journals and conference papers from Springer Database. Systematic literature review is summarized in following paragraph:

K. Arya and A. Bandil [3] proposed a new authentication technique involving the use of random sequence based secret sharing scheme with a data repair capability. Advantages of the scheme include ensuring integrity of the image and data repair capability. Limitations of the scheme include complexity due to additional parameters for computing authentication signal for block of image.

Che-Wei Lee and Wen-Hsiang Tsai [4] proposed a new blind authentication method based on the secret sharing technique with a data repair capability. Advantages of the scheme involve protection of data hidden in the alpha channel and data repair capability available for each block of image. Limitations of the scheme is increased complexity due to computation of authentication signal for each block of image.

Ahmad et al [5] proposed methods to protect the medical data by using the shared secret mechanism and steganography for 2 and 1 bit LSB.Advantages include relatively good quality images are produced in terms of PSNR.

T. Tuncer and E. Avci [6] proposed a new data hiding method based on secret sharing scheme with DNA-XOR operator for color images. Advantages include better protection of secret data as data is divided into 3 secret shares (red, green, blue).

G. Ulutas et al [7] proposed a new secret image sharing method by selecting the number of authentication bits proportional to block size. Advantages include improved authentication for increased block size and can authenticate individual stego-images as well. Limitations include calculation of vlock size and taking some extra bits for image authentication.

Nguyen et al [8] proposed a new reversible hiding scheme based on optimal exploiting modification direction matrix to embed large amount of secrets without causing distortion to images. Advantages of the scheme includes increased secret embedding capacity and prevention of image quality. Limitation of scheme is the complexity involved in selecting regions and usage of optional EMD table.

Y. Tsai et al [9] proposed a reversible data hiding scheme based on neighboring pixel differences. Advantages of the scheme is better quality than histogram shifting based typical schemes. Limitation of the scheme is smaller hiding capacity.

T. Nguyen and C. Chang [10] proposed a new reversible data hiding scheme based on the Sudoku technique. Advantage of the scheme is increased embedding capacity and consistency in maintaining good quality of stego-image. Limitation of the scheme is complexity involved in building reference matrix.

H. Sun, et al [11] proposed a data hiding system by means of flexible exploiting modification directions to achieve safer message concealments in image. Advantage of the system is higher embedding capacity and robustness against the blind steganalyzers. Limitations of the scheme is direct relation between n-cover pixels and secret to be embedded.

M. Deng et al [12] proposed an efficient scheme based on buyer-seller watermarking protocol utilizing homomorphic public key cryptosystem and composite signal representation. Advantages of the scheme include reduction of both computation overhead and bandwidth utilization. Limitation of the scheme is that it is very difficult to play in encrypted domain.

W. Hu et al [13] proposed a novel scheme based on Shamir's (t, n) threshold scheme and Galois Field GF(2^n) using dynamic embedding and least significant bit construction. Advantages of the scheme include lossless restoration of both secret image and cover image and satisfaction performance of dynamic embedding.

III. PROPOSED MODEL

The proposed scheme is a threefold model. In the first phase sender has the choice to select among various data/image compression techniques. Second phase involves Data hiding with sender having the choice to select among various data hiding techniques. For practical purposes we used least significant bit (LSB) technique [14], sparse space is created to accommodate some additional data (secret data) and using data hiding key data is embedded into the image. In third phase we use Generalized scheme over RSA using 2^k prime numbers [1] for encrypting image and data hiding key:

➢ Image is being encrypted using senders private key and can be only decrypted using senders public key, thus confirming authenticity of the sender.
➢ Data hiding key is being encrypted by receiver's public key and can only be decrypted using receivers private key, thus allowing only the receiver to decrypt the key and extract secret data.
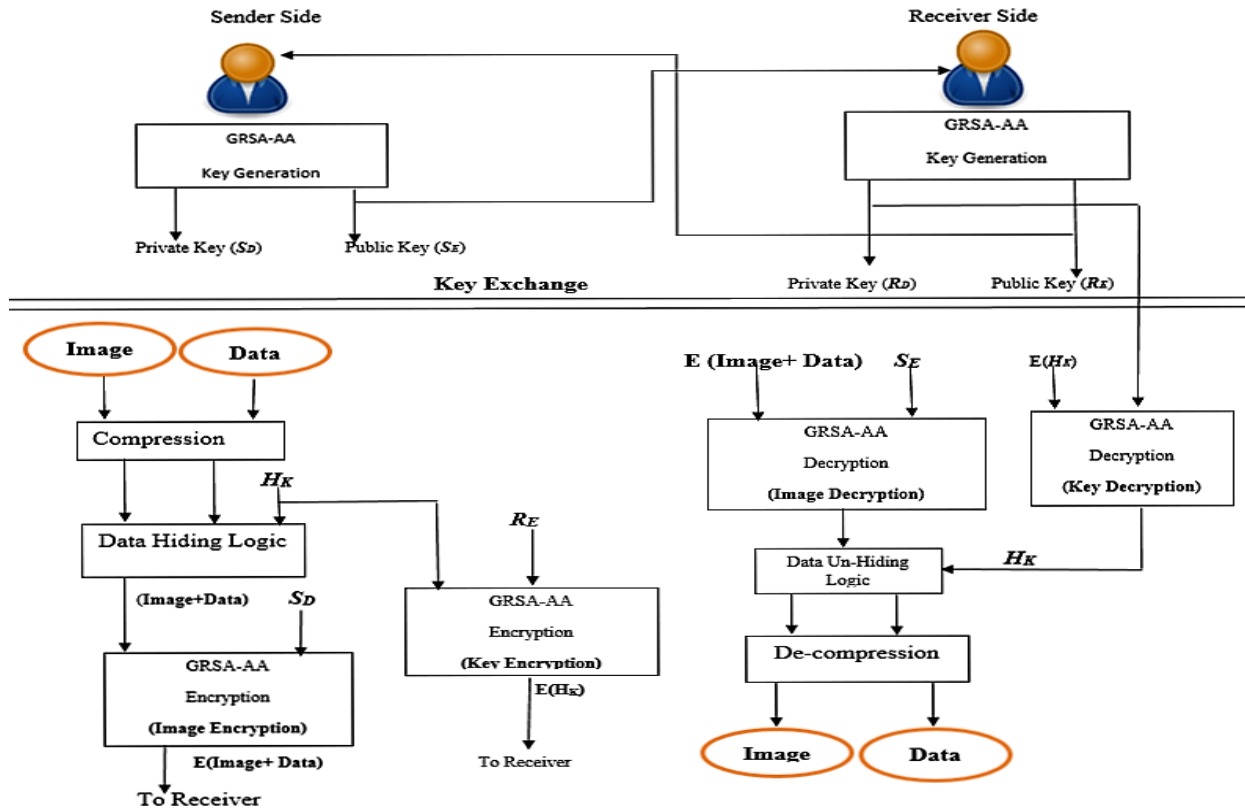
Fig.1. Novel Scheme for Image Authentication and Secret Data Sharing.

Parameter	Meaning
H_K	Data Hiding Key
S_D	Senders Private Key
S_E	Senders Public Key
R_D	Receivers Private Key
R_E	Receivers Public Key
E (Image + Data)	Encrypted Image + Data
$E(H_K)$	Encrypted Data Hiding Key

Proposed model consists of following phases:

1. Data/Image compression and Decompression phase:
   In the first phase sender has the choice to select
   among various data/image compression techniques.
   Main purpose is to reduce irrelevance and redundancy
   of the image data in order to be able to store or
   transmit data in an efficient form. Data/image
   compression is used to reduce the resource usage.
   Compressing the data will also involve decompressing
   the data/image at the receiver side.

2. Data Hiding Phase: Data hiding phase starts by taking
   an input image. After reading the image and finding
   its pixel values and their corresponding binary values.
   Similarly, users input message is taken and binary
   value of each character is calculated. Algorithm in
   Data Hiding logic takes following parameters as input:
   number of rows, columns, zero matrix of size rows-
   $2 \times$columns-3.and Data Hiding key. Output of the
   algorithm produces an index matrix which specifies
   the points in image where to make replacement for

secret data points. This index matrix will house the
data hiding points. The chosen pixels are then used to
hide the secret message. For embedding data into an
image we use LSB (Least Significant Algorithm)
algorithm.

Algorithm 1: A LSB-based Embedding Algorithm

Input -: cover C
for i = 1 to Length(c), do
$S_j \leftarrow C_j$
for i = 1 to Length(m), do
Compute index ji where to store the ith message
bit of m
$S_{ji} \leftarrow LSB(C_{ji}) = m_i$
End for
Output -: Stego image S
---

3. Image and Data hiding key Encryption: For this we
   use generalized scheme over RSA algorithm using 2^k
   prime numbers that is Generalized RSA – AA
   (GRSA-AA) [1]. GRSA-AA uses 2^k prime numbers
   with secured key generation. This scheme is based on
   2^k (k>=2) prime numbers and depending upon the
   value of k we can have various variants with 4, 8, 16,
   32… prime numbers thus offering great resistance to
   direct attacks which earlier variants were suffering
   from. GRSA-AA includes novel and secure algorithm
   for key generation making it more secure than
   different existing variants of RSA algorithm. The
   advantage of using this algorithm that public key and
   private key components are generated by making use
   of N, where N is a function of 2^k prime numbers.

When an attacker gets public key component n out of {E, n} by using factorization techniques he can only get two initial prime numbers (since, n is product of first two prime numbers). However, finding remaining prime numbers is computationally infeasible as no relevant information is available to the attacker. Hence, it is difficult for the attacker to determine the private key component D out of {D, n} knowing public key component {E, n}. Thus, it is practically impossible to break system using brute force attack.

Image is being encrypted using senders private key and can be only decrypted using senders public key, thus confirming authenticity of the sender. Data hiding key is being encrypted by receiver's public key and can only be decrypted using receivers private key, thus allowing only the receiver to decrypt the key and extract secret data.

Algorithm 2: GRSA-AA Key Generation for 8 Prime Numbers.

GRSA-AA_KeyGen ()
INPUT:
Eight prime numbers: $p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8$.
OUTPUT:
Public key components {E, n}
Private key components {D, n}
PROCEDURE:
$n \leftarrow p_1 * p_2$
$m \leftarrow p_3 * p_4$
$o \leftarrow p_5 * p_6$
$p \leftarrow p_7 * p_8$
$N_1 \leftarrow n * m$
$N_2 \leftarrow o * p$
$N \leftarrow N_1 * N_2$
/*Compute Euler phi values of n, m, o and p */
$\Phi(n) \leftarrow (p_1-1) * (p_2-1)$
$\Phi(m) \leftarrow (p_3-1) * (p_4-1)$
$\Phi(o) \leftarrow (p_5-1) * (p_6-1)$
$\Phi(p) \leftarrow (p_7-1) * (p_8-1)$
/*Compute Euler phi values of N */
$\Phi(N) \leftarrow \Phi(n) * \Phi(m) * \Phi(o) * \Phi(p)$
Find a random number e_1, satisfying $1 < e_1 < \Phi(n)$ and $gcd (e_1, \Phi(n)) = 1$
Find a random number e_2, satisfying $1 < e_2 < \Phi(m)$ and $gcd (e_2, \Phi(m)) = 1$
Find a random number e_3, satisfying $1 < e_3 < \Phi(o)$ and $gcd (e_3, \Phi(o)) = 1$
Find a random number e_4, satisfying $1 < e_4 < \Phi(p)$ and $gcd (e_4, \Phi(p)) = 1$
Compute $A_1 \leftarrow e_1^{e_2} \mod N_1$
Compute $A_2 \leftarrow e_3^{e_4} \mod N_2$
Compute $E' \leftarrow A_1^{A_2} \mod N$
Find a random number E, satisfying $1 < E < \Phi(n) * E'$ and $gcd (E, \Phi(n) * E') = 1$
Compute a random number D, such that,
$D \leftarrow E^{-1} \mod(\Phi(N) * E')$

Algorithm 3: GRSA-AA Encryption

GRSA-AA_Encrypt()
Input:
Plain text, Message (< n)
Public Key Components {E, n}
Output:
Cipher Text, C
Procedure:
$C \leftarrow M^E \mod n$

4. Image and Data hiding key Decryption: For this we use GRSA-AA Decryption function. First we decrypt image by Senders public key thus gives confirmation about authenticity of image source. Similarly, GRSA-AA decryption is performed to recover data hiding key, but decryption process is governed by receiver's private key thus will be sole entity having data hiding key.

Algorithm 4: GRSA-AA Decryption

GRSA-AA_Decrypt ()
Input:
Cipher Text Message, C
Private Key Components: {D, n}
Output:
$M \leftarrow C^D \mod n$

5. Data Extraction: In this phase the image with hidden data is taken and again pixel values are calculated. These pixel values are compared with the zero matrix (already created using the same procedure as discussed in data embedding) and pixels holding the additional data are nominated and decoding procedure is performed on these pixels to retrieve the message and original image pixel values.

Algorithm 5: A LSB-based Extracting Algorithm

Input -: Secret image s
for i = 1 to Length (m), do
Compute index ji where to store the ith
message bit of m
$m_{ji} \leftarrow LSB(C_{ji})$
End for

IV. IMPLEMENTATION AND RESULTS

For simulation purposes proposed model is implemented in MATLAB.Following screen shots depict working of proposed scheme:

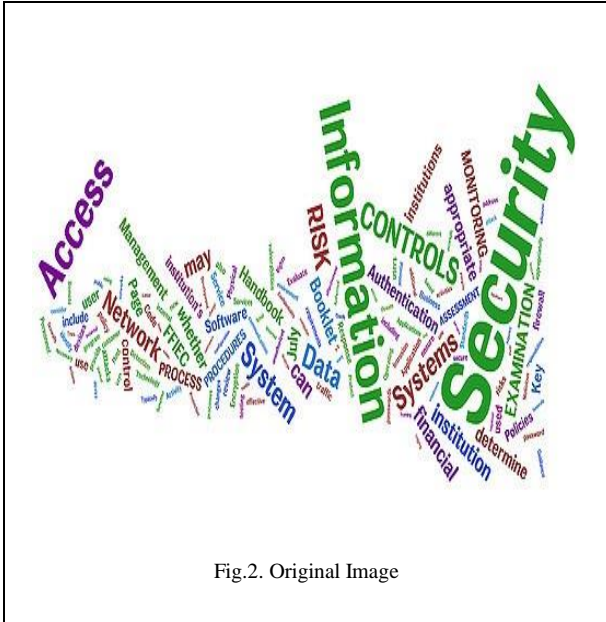*Processing at Senders Side*

Fig.2. Original Image

A Novel Scheme for Image Authentication and Secret Data sharing
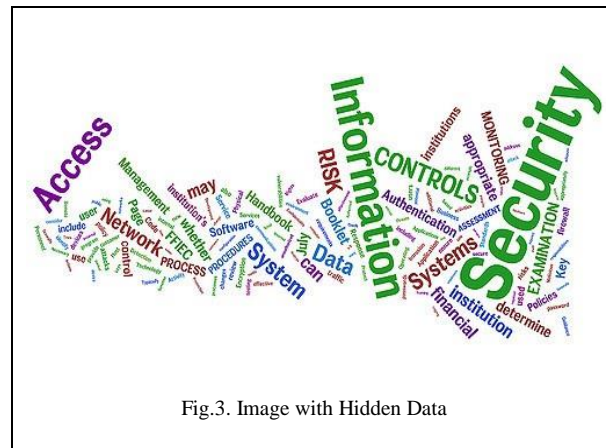
Original Message
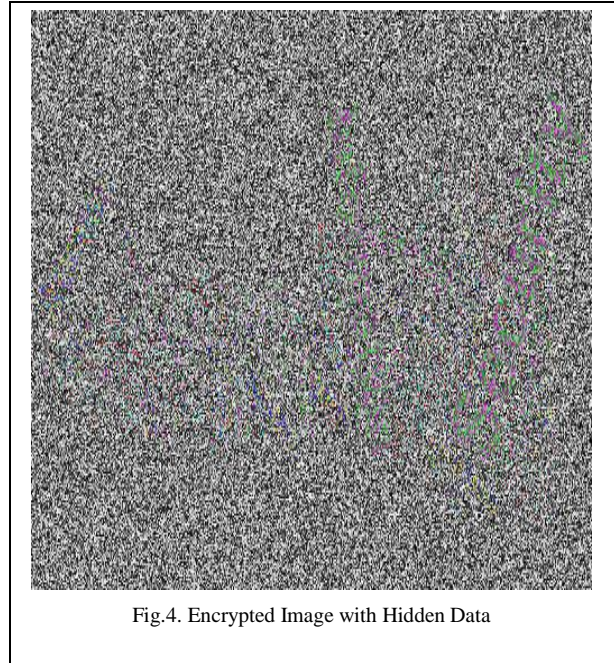

Fig.3. Image with Hidden Data

Fig.4. Encrypted Image with Hidden Data

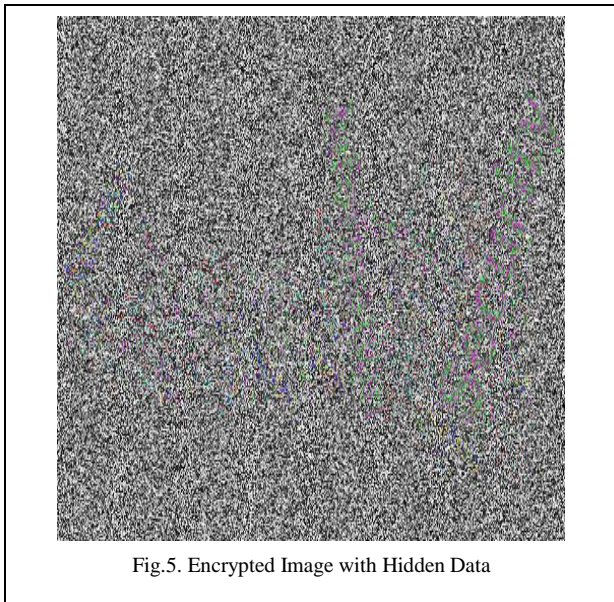*Processing at Receivers Side*

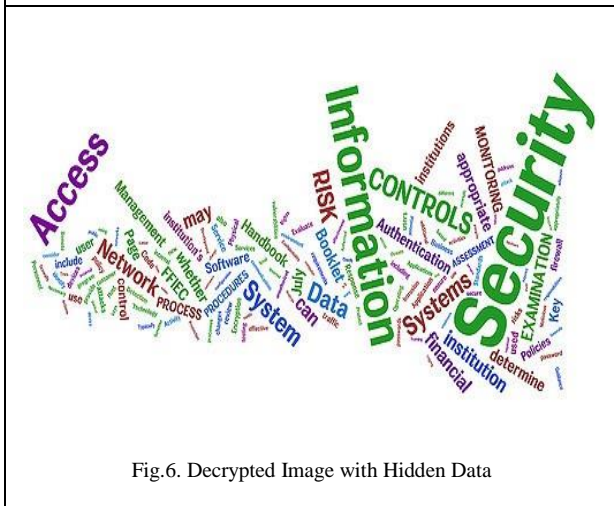Fig.5. Encrypted Image with Hidden Data

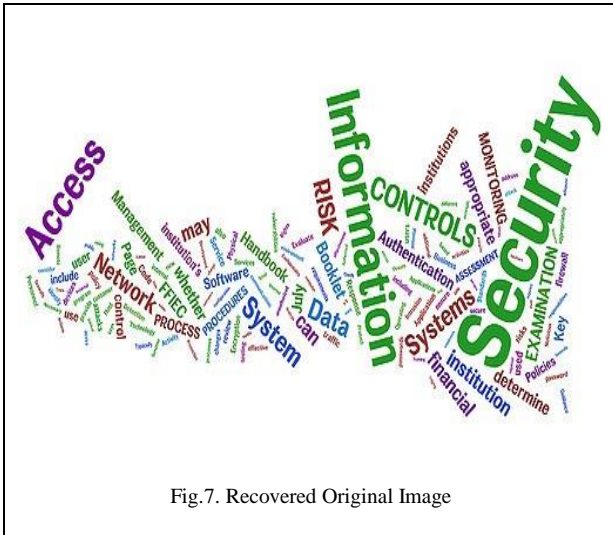Fig.6. Decrypted Image with Hidden Data

Fig.7. Recovered Original Image

A Novel Scheme for Image Authentication and Secret Data sharing

Recovered Original Message

Figure 2 is an image serving as a carrier for the Original message that needs to go through encryption before sent to the desired destination.

Figure 3 contains Original message enshrouded using LSB technique and Figure 4 represents encryption performed on the image shown in Figure 3 using Generalized RSA Algorithim.

Figure 5 is encrypted image received by the desired recipeint and Figure 6 is the image decrypted using Generalized RSA Algorithim by the recipeint.

Figure 7 is the decrypted Image with hidden data and contiguous to it is the original message Received.

The RGB image named Test image sized 125×125 was used as the original image for experimental results. After converting RGB image into Grayscale image, a zero matrix of size 123*122 was created. An algorithm taking inputs as zero matrix, dim1, dim2, rows, and columns, data hiding key (dim1=number of rows-2, dim2=number of columns-3) gives output an Index matrix which specifies the chosen data hiding points. A Message "A novel scheme for image authentication and secret data sharing" is hided into the LSB points specified by Index matrix. For Encryption of Image with Hidden data,

Generalized RSA with private key =123 of sender is used. Public key of receiver with value =111 was used to encrypt the Data hiding key. Figure 3 and figure 4 correspond to Image with hidden data and Encrypted image with hidden data at receiver side the public key of sender with value = 456, is used to decrypt the image containing secret data. Receivers own private key with value =567 is used to decrypt the Data hiding key to recover the secret data. Figure 6 represents the decrypted image with secret data and figure 7 represents the recovered original image.

Security Analysis of proposed scheme

The Security of proposed scheme solely relies on Security of GRSA-AA [1]. In this subsection we discuss the security of GRSA-AA.

Original RSA is vulnerable to various attacks including timing attack and factorization attack. The time to break RSA system is equivalent to the time for factorizing public key n. This simply requires finding prime factors of n. For this purpose, elliptic curve factorization (ECM) and General Number Field Sieve (GNFS). GNFS and ECM are the first and third fastest factoring methods respectively. ECM is commonly being used for small number factoring whereas GNFS is capable of factoring larger than 100 bits. However, the beauty of GRSA-AA lies in the fact that even if ECM or GNFS can be used to factor public key n, but this parameter is not sufficient to recover private key D (which is function of N not only n) to break the system. The above factoring methods can be used to find prime numbers namely p_1 and p_2 but other prime numbers (remaining 6 in case of GRSA-AA using 8 primes and remaining 14 primes in case of GRSA-AA 16) can be found only a brute force attack which is computationally infeasible for large prime numbers. Thus, time required to break the system is defined as:

$t_{system} = t_{p1,p2} + t_{bruteforce}$
$t_{system} =$ Time taken to break the system
$t_{p1,p2} =$ Time taken to find p_1 and p_2 using GNFS or ECM
$t_{bruteforce} =$ Time taken for brute force attack

Practically launching a brute force attack would be difficult task with large prime numbers of 1024 bits or greater, so it is almost practically impossible to launch a brute force attack against such a system and giving unbreakable cryptographic system to its users
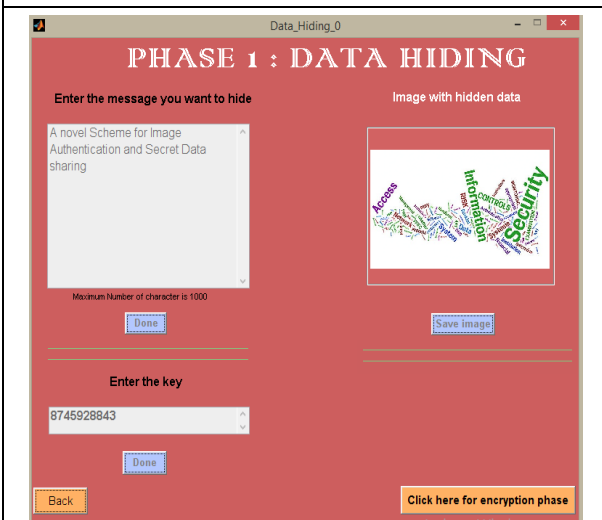
Snap Shots of Working Model

For practical purposes we created working demo application of the proposed scheme. Following are the snap shots of running working model of the proposed scheme. Every snap shot has commentary given depicting its core functionality.
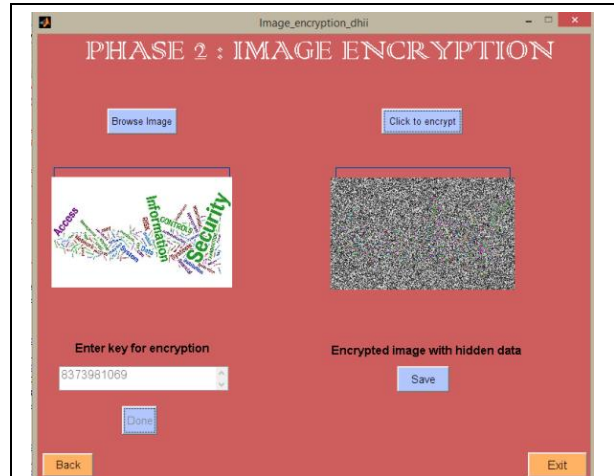
Home page of Application

Following interface of the application allows user either to work as a sender or a receiver. Depending upon the choice selected, user can act as sender or receiver and perform necessary operations either as sender or receiver.

Interface for Data Hiding

Following interface allows users

1. To write secret message that needs to be send to desired destination.
2. Select the image in which secret data is to embedded.
3. Entering the hiding key for embedding secret data inside the image.

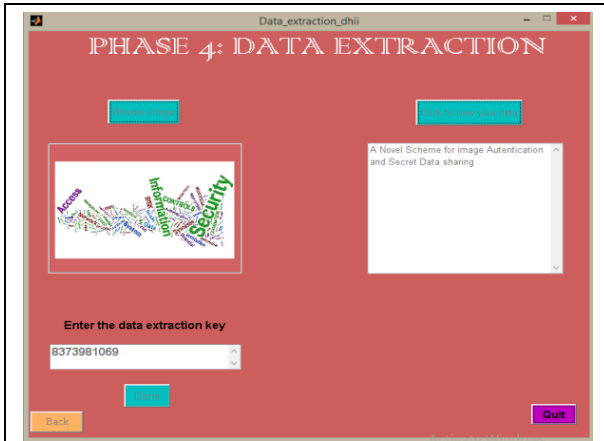On clicking Done button on the interface secret data is embedded inside the image.

Interface for Image (with Hidden Data) Encryption
Following interface allows to select image with hidden data and entering encryption key. On clicking click to encrypt button on interface image with hidden data is encrypted with generalized RSA algorithm.

Interface for Encrypted Image (with hidden data) Decryption

Following Interface allows selecting the encrypted image with hidden data and entering the decryption key. On clicking click to decrypt button on interface encrypted image is decrypted using generalized RSA algorithm and plain image with hidden data is recovered.

Interface for Data Extraction

Following interface allows to select image with hidden data amd entering the hiding key . Clicking done button first and then click to get data button on the interface allows application users to recover original secret data.

Note: Encryption and Decryption were carried using Algorithm 3 and 4 respectively. And all Keys were generated using algorithm 2

V. Conclusion

In this paper, we presented a novel approach for image authentication and secret data sharing by employing LSB as data hiding logic and GRSA-AA algorithm for encryption and decryption purposes. GRSA-AA uses 2^k large prime numbers thereby increasing the time required to find the private key components (D, n) in which D is a function of N. An attacker factorizes the n\to determine p_1 and p_2, however the value of D depends on N (not n) which is a product of 2^k prime numbers Public key component E of (E, n) in GRSA-AA is computed using multiple iterative modular and Euler functions complying Fermat Little Theorem. Hence, it is computationally infeasible to determine value of D based on known factorization of n. (computationally infeasible to determine D) and thus making GRSA-AA stronger than other available RSA variants. We have provided working model and implementation results for the proposed scheme in earlier sections. Based upon implementation results and conclusion the proposed scheme seems to be simple but effective in terms of security and privacy it offers. Moreover, the proposed model is generic in nature any data hiding technique and encryption mechanism can be employed depending upon the level of security required.

Acknowledgement

First of all, we are very grateful to almighty Allah for giving us patience and perseverance for completing the work successfully. Further we would like to thank Mr. Aqeel Khalique (Assistant Prof. Jamia Hamdard) for his kind support. At Last but not least we would like to thank Dr. Sherin Zafar (Assistant Prof. Jamia Hamdard) for her valuable suggestions and Kind Supervision.

References

[1] A. Lone, "Generalized RSA Using 2^k Prime Numbers with Secure Key Generation", M. Tech, Jamia Hamdard (Hamdard University), 2016.

[2] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey and S. Linkman, "Systematic literature reviews in software engineering – A systematic literature review", *Information and Software Technology*, vol. 51, no. 1, pp. 7-15, 2009.

[3] K. Arya and A. Bandil, "An improved image authentication technique using random-sequence based secret-sharing scheme", *2014 9th International Conference on Industrial and Information Systems (ICIIS)*, 2014.

[4] Che-Wei Lee and Wen-Hsiang Tsai, "A Secret-Sharing-Based Method for Authentication of Grayscale Document Images via the Use of the PNG Image With a Data Repair Capability", *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 207-218, 2012.

[5] T. Ahmad, H. Studiawan, H. Ahmad, R. Ijtihadie and W. Wibisono, "Shared secret-based steganography for protecting medical data", *2014 International Conference on Computer, Control, Informatics and Its Applications (IC3INA)*, 2014.

[6] T. Tuncer and E. Avci, "A reversible data hiding algorithm based on probabilistic DNA-XOR secret sharing scheme for color images", *Displays*, vol. 41, pp. 1-8, 2016.

[7] G. Ulutas, M. Ulutas and V. Nabiyev, "Secret image sharing scheme with adaptive authentication strength", *Pattern Recognition Letters*, vol. 34, no. 3, pp. 283-291, 2013.

[8] T. Nguyen, C. Chang and N. Huynh, "A novel reversible data hiding scheme based on difference-histogram modification and optimal EMD algorithm", *Journal of Visual Communication and Image Representation*, vol. 33, pp. 389-397, 2015.

[9] Y. Tsai, D. Tsai and C. Liu, "Reversible data hiding scheme based on neighboring pixel differences", *Digital Signal Processing*, vol. 23, no. 3, pp. 919-927, 2013.

[10] T. Nguyen and C. Chang, "A reversible data hiding scheme based on the Sudoku technique", *Displays*, vol. 39, pp. 109-116, 2015.

[11] H. Sun, C. Weng, S. Wang and C. Yang, "Data embedding in image-media using weight-function on modulo operations", *TECS*, vol. 12, no. 2, pp. 1-12, 2013.

[12] M. Deng, T. Bianchi, A. Piva and B. Preneel, "An efficient buyer-seller watermarking protocol based on composite signal representation", *Proceedings of the 11th ACM workshop on Multimedia and security - MM&Sec '09*, 2009.

[13] W. Hu, M. Li, C. Guo and Y. Ren, "Reversible secret image sharing with steganography and dynamic embedding", *Security and Communication Networks*, 2012.

[14] P. S., B. S. and P. R., "A Novel Security Scheme for Secret Data using Cryptography and Steganography", *International Journal of Computer Network and Information Security*, vol. 4, no. 2, pp. 36-42, 2012.

Authors' Profiles

Auqib Hamid Lone received B.Tech degree in Information Technology and Telecommunication Engineering from Baba Ghulam Shah Badshah University, Rajouri, Jammu and Kashmir, India, in 2014, and is presently pursing MTech degree in Information Security and Cyber Forensics from Jamia Hamdard (Hamdard University), New Delhi, India. His research interest is in following fields: Cryptography, Network Security and Cyber Forensics.

Ab Waheed Lone received the B.Tech degree in Computer Science and engineering from University of Kashmir, India in 2014 and is presently pursuing M.Tech in Computer sciences and engineering degree from Jamia Hamdard (Hamdard University), New Delhi, India. His research interests are in Algorithm analysis, Network security, MATLAB.

Prof. Moin Uddin is currently serving as Dean and Professor in Department of Computer Science & Engineering, Faculty of Engineering and Technology Jamia Hamdard, New Delhi, India. He is Doctorate (Ph.D.) in the field of "ELECTRONICS AND COMPUTER ENGINEERING" from a well reputed Technical University (INDIA) and is having more than 25 years teaching and R & D experience in Engineering Colleges at various levels in India and Iraq.