Secure Data Transfer using RSA and Steganography

Zo Nun Khuma
Department of Information Technology,
West Yangon Technological University
Myanmar

Abstract: Today, transferring sensitive information via secure methods and protecting information privacy are becoming critically important, the two ways for securing and transferring messages and data are cryptography and steganography. Steganography is used for hiding messages in innocuous media (carriers) such as text, image, audio, video and protocol. Amongst these different carriers, digital images are the most popular because of their frequency on the internet. To be more robust in security, steganography can be combined with cryptographic techniques. In this paper, we propose Secure Data Transfer using RSA and Steganography in order to obtain a secure system. The system will be implemented by Java programming language.

Keywords: Cryptography, RSA, Secure Data Transfer, Image Carriers, Steganography

1. INTRODUCTION

Since the rise of the Internet one of the most important factors of communication and information technology has been the security of information. It is necessary to protect this information while communicated over insecure channels. Thus, a need exists for developing technology that will help protect the integrity of digital content and secure the intellectual property rights of owners. This has resulted in an explosive growth of the field of information hiding. In addition, the rapid growth of publishing and broadcasting technology also requires an alternative solution in hiding information. To overcome this problem, some invisible information can be embedded in the digital media in such a way that it could not be easily extracted without a specialized technique [1]. One and most widely used for securing data is Steganography. The word steganography is derived from the Greek words "*stegos*" meaning "cover" and "*grafia*" meaning "writing" defining it as "covered writing" [2]. Today, steganography is mostly used on computers with digital data being the carriers and networks being the high speed delivery channels. It is a performance that inserts secret messages into a cover file, so that the existence of the messages is not apparent. Research in information hiding has tremendous increased during the past decade with commercial interests driving the field [3].

For security enhancement, this system can be made by combining with another method for securing data, called Cryptography. 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.

2. RELATED WORK

Recently, there is a growing interest in the information security to the problems in the current data communication domain. Steganography is the idea of hiding private or sensitive data or information within something that appears to be nothing out of the ordinary.

The paper [4] pointed out that "some specific image based steganography techniques and show that an observer can indeed distinguish between images carrying a hidden message and images which do not carry a message." And the paper [5] also figures out that "given an image is believed to contain a secret message, identify where the message is hidden. We treat this problem as outlier detection". Moreover, the paper [6] pointed out that "The strength of steganography can be thus amplified by combining it with cryptography".

On the other hand, applying encryption algorithm on the message can enhance the security. And it is an important tool to protect information from attackers. To overcome this problem, the need is to add security mechanism in our system. To improve the performance of this proposed system, steganographic and cryptographic techniques are combined to implement the system.

3. STEGANOGRAPHY

3.1 Different Kinds of Steganography

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. Figure 1 shows the categories of steganography.

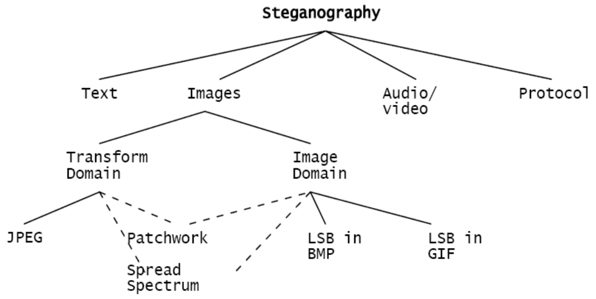


Figure 1 Categories of Steganography

Hiding information in text is historically the most important method of steganography. Text steganography using digital files is not used very often since text files have a very small amount of redundant data. The large amount of redundant bits presents in the digital representation of an image, so images are the most popular cover objects for steganography.

To hide information in audio files similar techniques are used as for image files. One different technique unique to audio steganography is masking, which exploits the properties of the human ear to hide information unnoticeably. A faint, but audible, sound becomes inaudible in the presence of another louder audible sound [3]. Although nearly equal to images in steganographic potential, the larger size of meaningful audio files makes them less popular to use than images [7].

The protocol steganography refers to the technique of embedding information within messages and network control protocols used in network transmission [8]. In the layers of the OSI network model, there exist covert channels where steganography can be used [9].

3.2 Image Steganography

Images are the most popular cover objects used for steganography. For the different image file formats, different steganographic algorithms exist. For these different image file formats different steganography algorithm exist.

Image steganography techniques can be divided into two groups: those in the Image Domain and those in the Transform Domain [10]. Image – also known as spatial – domain techniques embed messages in the intensity of the pixels directly, while for transform – also known as frequency – domain, images are first transformed and then the message is embedded in the image [11]. Image domain techniques encompass bit-wise methods that apply bit insertion and noise manipulation and are sometimes characterized as “simple systems”. The image formats that are most suitable for image domain steganography are lossless and the techniques are typically dependent on the image format.

Steganography in the transform domain involves the manipulation of the algorithms and image transforms. These methods hide messages in more significant areas of the cover image, making it more robust. Many transform domain methods are independent of the image format and the embedded message may survive conversion between lossy and lossless compression.

3.3 Embedding data into an image

To a computer, an image is an array of numbers that represent light intensities at various points, or pixels. In digital, images are represented with the numerical values of each pixel where the value represents the color and intensity of the pixel. These pixels make up the image’s raster data. Digital images are typically stored in either 24-bit or 8-bit per pixel files. 24-bit images are known as true colour images. Obviously, a 24-bit image provides more space for hiding information as compared to 8 bit image [12].

3.4 Least significant bit insertion

The least significant bit insertion method is probably the most well-known image steganography technique. It is a common, simple approach to embed information in a graphical image file. This is the most common method used in this the data to be hidden is inserted into the least significant bits of the pixel information In digital, images are represented with the numerical values of each pixel where the value represents the color and intensity of the pixel. In 24 bit image we can embed 3 bits in each pixel while in 8-bit we can embed only 1 bit in each pixel. To hide an image in the LSBs of each byte of the 24- bit image, one can store 3 bits in each pixel. A 1024 x 768 image has the potential to hide a total of 2,359,296 bits of information. For e.g., the letter A can be hidden in three pixels. The binary value of A is 10000011.

The original raster data of 3 pixels may be:

(00100111 11101001 11001000)

(00100111 11001000 11101001)

(11001000 00100111 11101001).

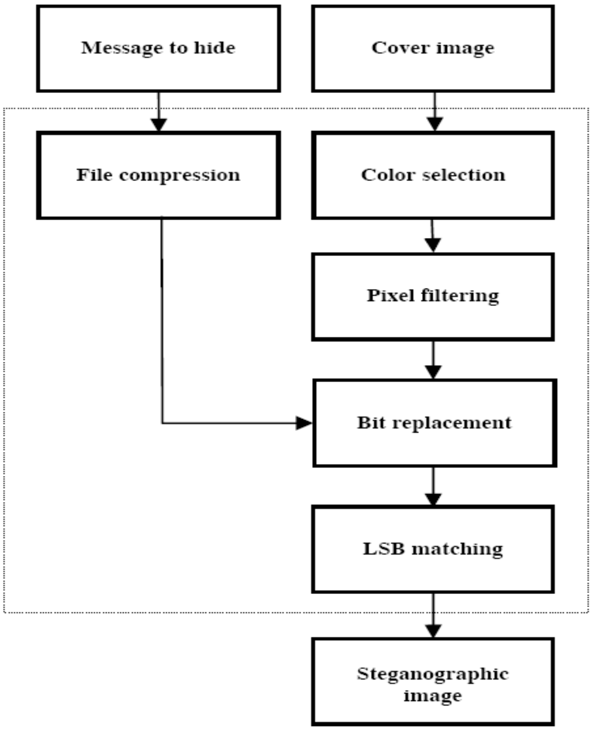


Figure 2 Structure of the Algorithm SLSB

After inserting the binary value for A.

(00100111 11101000 11001000)

(00100110 11001000 11101000)

(11001000 00100111 11101001)

The highlighted bits are the only three actually changed in the 9 bytes of data. One can hide data in the least and second least significant bits and still the human eye would not be able to discern it. In this proposed algorithm we generate random number initialized with a stego-key and its output is combined with the input data, and this is embedded to a cover image. The usage of a stego-key is important, because the security of a protection system should not be based on the secrecy of the algorithm itself, instead of the choice of a secret key.

4. CRYPTOGRAPHY

Cryptography is an important element of any strategy to address message transmission security requirements. Cryptography is the study of methods of sending messages in disguised form so that only the intended recipients can remove the disguise and read the message. It is the practical art of converting messages or data into a different form, such that no-one can read them without having access to the 'key'. The message may be converted using a 'code' (in which case each character or group of characters is substituted by an alternative one), or a 'cypher' or 'cipher' (in which case the message as a whole is converted, rather than individual characters).Cryptology is the science underlying cryptography. Cryptanalysis is the science of 'breaking' or 'cracking' encryption schemes, i.e. discovering the decryption key. Cryptographic systems are generically classified along three independent dimensions [13]. If both sender and receiver use the same key, the system is referred to as symmetric, single-key, secret-key, or conventional encryption. If the sender and receiver each use a different key, the system is referred to as asymmetric, two keys, or public-key encryption.

4.1 Asymmetric Encryption

Asymmetric algorithm also called as public-key algorithm rely on one key (public key) for encryption and a different by related key (private key) for decryption as shown in Figure 3. Each recipient has a private key that is kept secret and a public key that is published for everyone. The sender looks up or is sent to the recipient's public key and uses it to encrypt the message. The recipient uses the private key to decrypt the message and never publishes or transmits the private key to anyone. Thus, the private key is never in transit and remains invulnerable [14].

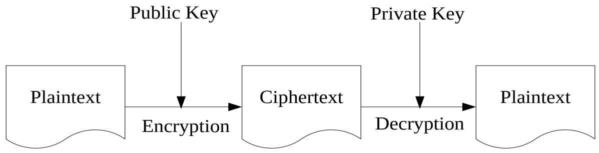


Figure 3 Public-Key Encryption Scheme

4.2 RSA ALGORITHM

The RSA algorithm is named after Ron Rivest, Adi Shamir and Len Adleman, who invented it in 1977[15]. The basic technique was first discovered in 1973 by Clifford Cocks [16] of CESG (part of the British GCHQ) but this was a secret until 1997. The patent taken out by RSA Labs has expired.

The RSA cryptosystem is the most widely-used public key cryptography algorithm in the world. It can be used to encrypt a message without the need to exchange a secret key separately.

The RSA algorithm can be used for both public key encryption and digital signatures. Its security is based on the difficulty of factoring large integers.

Party A can send an encrypted message to party B without any prior exchange of secret keys. A just uses B's public key to encrypt the message and B decrypts it using the private key, which only he knows. RSA can also be used to sign a message, so A can sign a message using their private key and B can verify it using A's public key.

4.2.1 Detailed Description of RSA Encryption Algorithm

In the encryption and decryption algorithm, converting plaintext to ciphertext the cryptographic algorithm is called encryption and converting ciphertext back to plaintext using cryptographic algorithm is called decryption. Each encryption scheme can be divided into three stages.

- Key Generation Stage
- Encryption Stage
- Decryption Stage

The security of any public-key algorithm, whether based on the difficulty of factoring large the numbers or the difficult of taking discrete logarithms of large numbers, depends on the size of those large numbers. If the number is small enough, we have no security. If the number is large enough, we have security all the computing power in the world working from now until the sun goes nova. We have a choice of key length, and it is an important choice. What follows is an analysis of different key lengths and their susceptibility to factoring, both today and in the near future.

The encryption key consists of the pair of integers (e, n), and the decryption key is (d, n) as shown in Figure 4. The starting point in finding keys for this RSA algorithm is to select a value of n. The value of n should be quite large, a product of two primes p and q. Both p and q should be large themselves. Typically p and q are approximately 100 digits each, so that n is approximately 200 digits long. This length effectively inhibits factoring n to infer p and q.

Next a relatively large integer e is chosen so that e is relatively primed to (p-1)*(q-1). (Recall that "relatively prime" means that e has no factors in common with (p-1)*(q-1)). An easy way to guarantee that e is relatively prime to (p-1)*(q-1) is to choose e as prime that is larger than both (p-1)*(q-1).

Finally, it is needed to select d such that e*d≡1 mod (p-1)*(q-1) as illustrated in Figure 5.

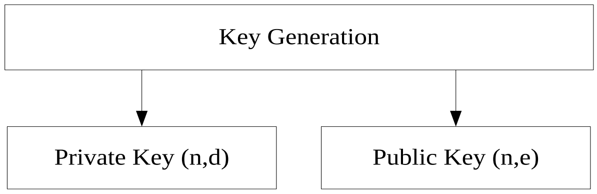


Figure 4 Block Diagram of Key Generation

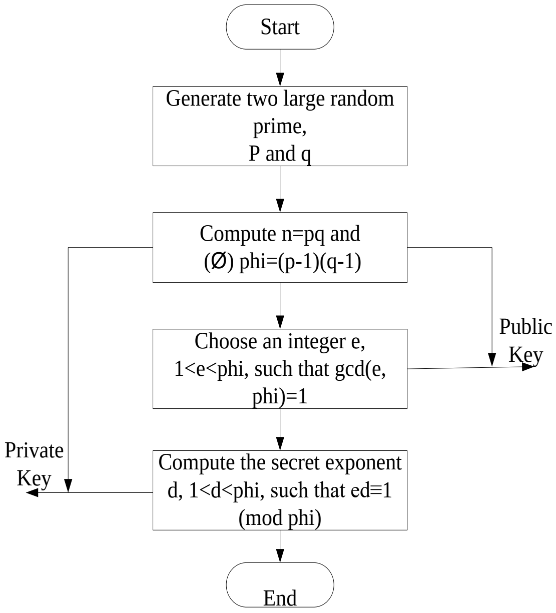

Start
Generate two large random prime, P and q
Compute n=pq and (Ø) phi=(p-1)(q-1)
Choose an integer e, 1<e<phi, such that gcd(e, phi)=1
Public Key
Private Key
Compute the secret exponent d, 1<d<phi, such that ed≡1 (mod phi)
End

Figure 5 Block Diagram of RSA Key Generation Algorithm

4.2.2 Small Example of RSA Method

Consider the case p = 53, q = 61, n = p * q = 53 * 61 = 3233 and d = 791 then ϕ(3233) = 52 * 60 = 3120 and e can be computed as e = inv (791, 3233). Therefore, e = 71, the multiplicative inverse (mod 3233) of d = 791.

With n = 3233 can be encoded two letters per block, substitution a two-digit numbers for each letter:

blank = 00, A = 01, B = 02, C = 03, D = 04, E = 05, F = 06,…….., Z = 26

Thus, the message M = RENAISSANCE

Becomes RE | NA | IS | SA | NC | E |

is encoded: 1704 1800 0818 1800 1302 0426

The first block, (M = 1704) is enciphered …

M71 = 3106 (mod 3233)

The whole message is enciphered as:

3106 0100 0931 2691 1984 2927

The reader can check those deciphering words:

3016791 = 704 (mod 3233), etc….

To encrypt and decrypt efficiently, use the fast exponentiation algorithm [17]. It is called "exponentiation by repeated squaring and multiplication".

5. PROPOSED SYSTEM

As shown in Figure 6, two ways of securing method are used in this proposed system: Steganography and Cryptography.

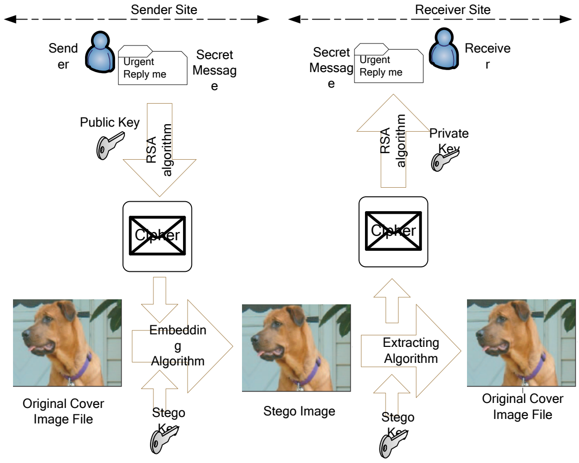


Figure 6 Proposed System Design

In the first step, at the sender site, the plaintext (secret message) is encrypted with the help of RSA, popular algorithm and then the cipher text is produced as in Figure 7. Because it is an asymmetric key algorithm, the Public Key is applied for encryption and Private Key for decryption.

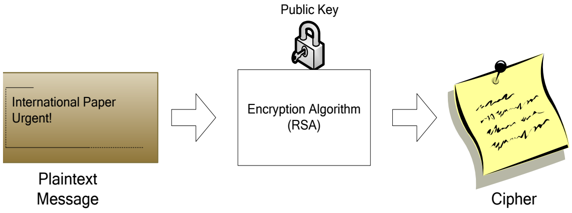


Figure 7 Encrypting the Message

Here we embed the message (Cipher) into the carrier image by using Least Significant Bit (LSB) insertion method. After embedding a secret message into the cover-image, a so-called stego image is obtained. The size of information to be hidden relatively depends on the size of the cover-image. Any image file can hide the message of size of one – eight the size of original cover file, e.g if cover image is 128 bytes then it hide 16 bytes of message without any distortion.

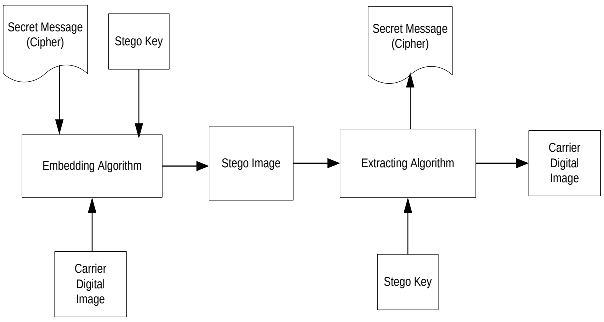


Figure 8 Embedding and Extracting the Message (Cipher)

After this process, at the receiver site, the embedded message is extracted from the stego image file and obtained the cipher text again as depicted in Figure 8.

The resulting cipher text is decrypted by using RSA algorithm with the receiver's Private Key. Finally, the receiver obtains the sent message (plaintext) as shown in Figure 9.

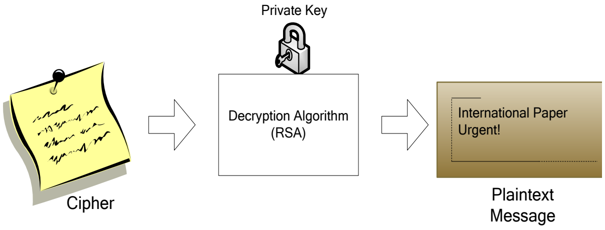


Figure 9 Decrypting the Message

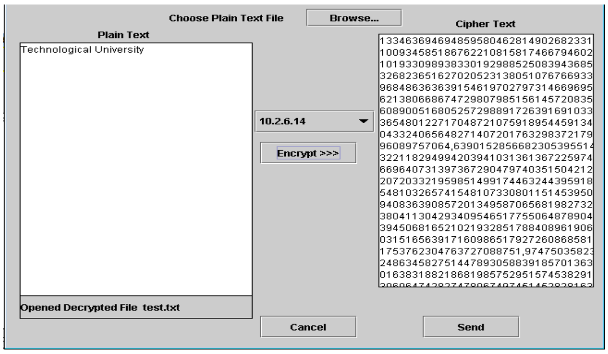
6. PROS AND CONS

Table 1 Steganography Vs Cryptography

Steganography	Cryptography
Unknown message passing	Known message passing
Steganography prevents discovery of the very existence of communication	Encryption prevents an unauthorized party from discovering the contents of a communication
Little known technology	Common technology
Technology still being developed for certain formats	Most of algorithm known by all
Once detected message is known	Strong current algorithms are currently resistant to attack, larger expensive computing power is required for cracking
Steganography does not alter the structure of the secret message	Cryptography alter the structure of the secret message

7. TEST AND RESULT

Firstly, the system is implemented by using StegoJava. User



Technological University

10.2.6.14

```
1334636946948595804628149026823331
1009345851867622108158174667946O2
1019330989383301929885250839436 85
3268236516270205231380510767669 33
9684863636391546197027973146696 95
6213806686747298079851561457208 35
6089005168052572988917263916910 33
3654801227170487210759189544591 34
0433240656482714072017632983721 79
96089757064,639015285668230539551 4
3221182949942039410313613672259 74
6696407313973672904797403515042 12
2072033219598514991744632443959 18
5481032657415481073308011514539 50
9408363908572013495870656819827 32
3804113042934095465177550648789 04
3945068165210219328517884089619 06
0315165639171609865179272608685 81
1753762304763727088751,974750358 23
2486345827514478930588391857013 63
0163831882186819857529515745382 91
```

Opened Decrypted File test.txt

Figure 10 Encrypting the Message with RSA Algorithm

need to enter the message or choose the text file from Browse first. Figure 10 shows the encryption process with RSA algorithm.

The resulting cipher message is fetched to the embedding mode and is embedded the GIF image which is used as a cover image as in Figure 11 by the help of LSB.

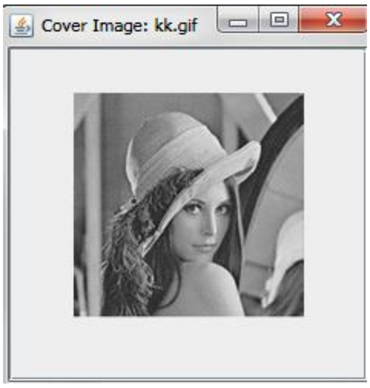


Figure 11 The Cover Image

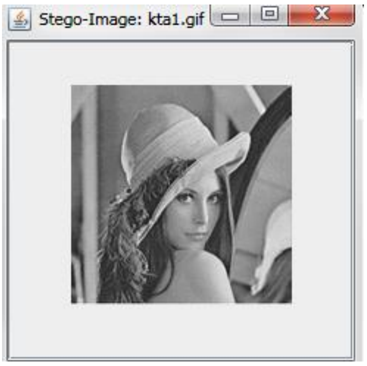
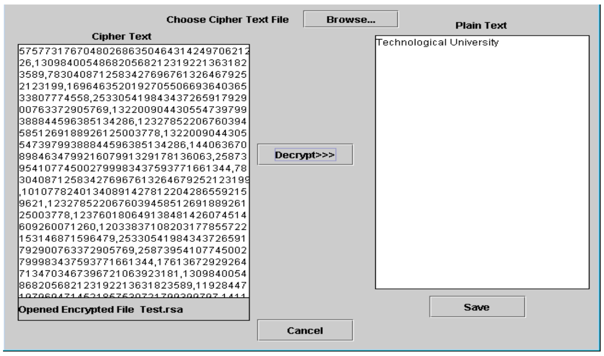


Figure 12 The Stego-image

Finally, the resulting cipher message extracting from the stego-image is decrypted to get the original message as shown in Figure 13.



```
575773176704802686350464314249706212
26,130984005486820568212319221363182
3589,78304087125834276967613264679 25
2123199,16964635201927055066936403 65
3380777455 8,253305419843437265917929
00763372905769,1322009044305547 39799
3888445963851 34286,1232785220676 0394
58512691889261250037 78,1322009044305
5473979938884459638513 4286,144063670
89846347992160799132917813606 3,25873
9541077450027999834375937716 61344,78
304087125834276967613264679252123199
,10107782401340891427812204286559215
9621,12327852206760394585126918 89261
25003778,12376018064913848142607 4514
609260071260,12033837108203177855722
1531468715964 79,253305419843437265 91
792900763372905769,258739541077 45002
7999834375937716613 44,1761367292926 4
71347034673967210639231 81,1309840054
868205682123192213631823589,11928447
```

Opened Encrypted File Test.rsa

Technological University

Figure 13 Decrypting the Cipher Message with Private Key

8. CONCLUSION

Till now, information hiding techniques received very much less attention from the research community and from industry than cryptography. Steganography has its place in security. It is not intended to replace cryptography but supplement it. In this paper we give an idea to enhance the security of system by combining the two techniques. It can enhance integrity and confidentiality of information and provides a means of communicating privately. Here message is first encrypted and then embed in image file with the help of steganographic system. There are infinite number of steganography applications for digital image including copyright protection, feature tagging, and secret communication. This paper explores a tiny fraction of the art of steganography. It goes well beyond simply embedding text in an image.

9. REFERENCES

[1] Dunbar, B. “Steganographic techniques and their use in an Open-Systems environment”, *SANS Institute*, January 2002

[2] Moerland, T. “Steganography and Steganalysis”, *Leiden Institute of Advanced Computing Science*,

www.liacs.nl/home/ tmoerl/privtech.pdf

[3] Silman, J. “Steganography and Steganalysis: An

Overview”, *SANS Institute*, 2001

[4] Chandramouli, R. & Nasir Memon, “Analysis of LSB Based Image Steganography techniques”, 2001 IEEE

[5] Ian Davidson & Goutam Paul, “Locating Secret Messages in Images”, *KDD’04*, Seattle, Washington, USA, August 22-25, 2004

[6] Morkel, T., JHP Eloff and MS Olivier, “An Overview of Image Steganography,” *in Proceeding of the fifth Annual Information Security South Africa Conference (ISSA 2005)*, Sandton, South Africa, June/July 2005 (Published electronically)

[7] Artz, D. “Digital Steganography: Hiding Data within Data”, *IEEE Internet Computing Journal*, June 2001

[8] Ahsan, K. & D Kundur, , “Practical Data hiding in TCP/IP”, *Proceedings of the Workshop on Multimedia Security at ACM Multimedia*, 2002

[9] Handel, T. & M. Sandford, “Hiding data in the OSI network model”, *Proceedings of the 1st International Workshop on Information Hiding*, June 1996

[10] Silman,J. “Steganography: and Steganalysis: An Overview”, *SANS Institute, 2001*

[11] Lee, Y.K. & L.H. Chen, “High capacity image steganographic model”, *Visual Image Signal Processing*, 147:03, June 2000

[12] Lin E.T. and E.J. Delp, “A Review of Data Hiding in Digital Images”, JUNE 2001

[13] Stinsown, D. “Cryptography:Theory and Practice”

[14] Cox, C., Killian, J., Leighton, T. and Shamoon, T., “Secure Spread Spectrum Communication for Multimedia”, N.E.C. Research Institute, 1995

[15] Rivest, R., Shamir, A. and Adleman, L., “ *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems”,* Communications of the ACM, 21 (2), pp. 120-126, February 1978

[16] Clifford Cocks, “*A Note on 'Non-Secret Encryption”,* CESG Research Report, 20 November 1973

[17] Johnston, P., “RSA Encryption Algorithm”, No Date