AN ENCRYPTO- STEGO TECHNIQUE BASED SECURE DATA TRANSMISSION SYSTEM

Neha Sharma
M.E – EPDT
PEC,Chandigarh
Nehu_82s@yahoo.com

Mr.J.S.Bhatia
Director - CDAC
Mohali

Dr(Mrs)Neena Gupta
Asst. Professor
PEC,Chandigarh
ng65@rediffmail.com

Abstract - Digital communication has become an essential part of infrastructure nowadays, a lot of applications are Internet-based and in some cases it is desired that communication be made secret. Consequently, the security of information has become a fundamental issue. Two techniques are available to achieve this goal: Encryption and steganography is one of them. Using cryptography, the data is transformed into some other gibberish form and then the encrypted data is transmitted. In steganography, the data is embedded in an image file and the image file is transmitted. This paper proposed a system that combines the effect of these two methods to enhance the security of the data. This proposed system encrypts the data with a crypto algorithm and then embeds the encrypted text in an image file. The embedding process is done with help of stego-key, and the detection or reading of embedded information is possible only having this key. The stego key (user-specified or default) is used not only to facilitate random selection of bytes for hiding message file bits but also is used to encrypt the message file. The encryption method is based on XORing the message bytes with random numbers generated by a pseudo-random number generator whose seed is derived from the stego key. Here we also calculate the message digest of image and embed into image file to check integrity of message contents.

I. Introduction

The growing possibilities of modern communications need the special means of security especially on computer network. The network security is becoming more important as the amount of data being exchanged on the Internet is increasing. Security requirements are necessary both at the final user level and at the enterprise level, especially since the massive utilization of personal computers, networks, and the Internet with its global availability. Throughout time, computational security needs have been focused on different features: secrecy or confidentiality, identification, verification, non repudiation, integrity control and availability. 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 alterative solution in hiding information. The copyright of digital media such as audio, video and other media available in digital form may lead to large-scale unauthorized copying. The problem of unauthorized copying is of great concern especially to the music, film, book and software publishing industries. 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].

There are a number of ways for securing data. One is cryptography, where the sender uses an encryption key to scramble the message, this scrambled message is transmitted through the insecure public channel, and the reconstruction of the original, unencrypted message is possible only if the receiver has the appropriate decryption key. The second method is steganography, where the secret message is embedded in another message, thus the existence of message is unknown.

II. 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 [2].

1. Methodology for transforming plain text to cipher text.

All encryption algorithms are based on two general principles**:** substitution, in which each element in the plaintext is mapped into another element, and transposition, in which elements in the plaintext are rearranged. The fundamental requirement is that no information be lost.

2. Methodology for number of keys used.

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 symmetric**,** two-keys, or public-key encryption.

3. Methodology for processing plain text.

A block cipher processes the input one block of elements at a time, producing an output block for each input block. A stream cipher processes the input elements continuously, producing output one element at a time, as it goes along. The proposed algorithm uses a substitution cipher method. It is a symmetric key algorithm using the technique of stream cipher.

III. Steganography

Steganography derived from Greek word literally means covered writing. It includes vast array of secret communication method that conceals message very existence. Computer based steganography allows changes to be made to what are known as digital carriers such as images or sounds. Digital images, videos, sound files, and other computer files that contain perceptually irrelevant or redundant information can be used as “covers” or carriers to hide secret messages. After embedding a secret message into the cover-image, a so-called stegoimage is obtained [3].

The basic model of steganography consists of Carrier, Message, Embedding

algorithm and Stego key. The model for steganography is shown in Figure 1. Carrier is also known as a cover-object, which embeds the message and serves to hide its presence. The suitable carriers that can be used as cover object are Network Protocols such as TCP, IP and UDP, Audio that use digital audio formats such as wav, midi, avi, mpeg, mpi and voc, File and Disk that can hide and append files by using the slack space, Text files such as html and java, Image files such as bmp, gif and jpg, where they can be both color and gray-scale [5]. Message is the data that the sender wishes to remain it confidential. It can be plain text, ciphertext, other image, or anything that can be embedded in a hit stream such as a copyright mark, a covert communication, or a serial number. Password is known as a stego-key, which ensures that only the recipient who knows the message from a cover-object. The cover-object with the secretly embedded message is then called the stego-object.This stego object is then transferred to other end, there we have detector algorithm which extract the message from cover object [4].

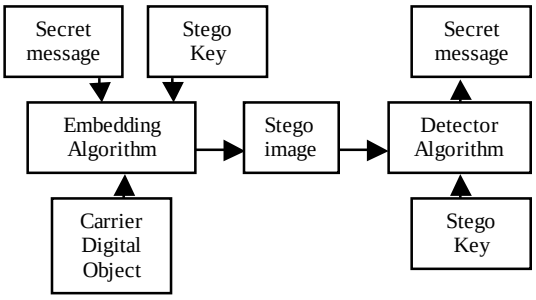


Fig. 1 Model of steganography

There are three different aspects in information-hiding systems contend with each other: capacity, security and robustness. Capacity refers to the amount of information that can be hidden in the cover medium, security to an eavesdropper's inability to detect hidden information and robustness to the amount of modification the stego medium can withstand before an adversary can destroy the hidden information.

IV. Steganography vs Cryptography

Basically, the purpose of cryptography and steganography is to provide secret communication. Many people lump Steganography with cryptography, and while they are in many cases means to the same ends they are not the same thing. According to Dictionary.com: Steganography is:" Hiding a secret message within a larger one in such a way that others can not discern the presence or contents of the hidden message" and Cryptography is "The process or skill of communicating in, or deciphering secret writing or ciphers." Steganography can be used to cloak hidden messages in image, audio and even text files. It has until recently been the poor cousin of cryptography. Now, it is gaining new popularity with the current industry demands for digital watermarking and fingerprinting of audio and video. Steganography must not be confused with cryptography, where we transform the message so as to make its meaning obscure to malicious people who intercept it. Therefore, the definition of breaking the system is different. In cryptography, the system is broken when the attacker can read the secret message. Breaking a steganographic system needs the attacker to detect that steganography has been used and he is able to read the embedded message [6]. In addition, the security of classical steganography system relies on secrecy of the data encoding system. Once the encoding system is known, the

steganography system is defeated. The distinction between cryptography and steganography is an important one, and is summarized by the following:[7]

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

Table 1: Comparison

V. Proposed system

Data hiding techniques have been widely used to transmission of hiding secret message for long time. Ensuring data security is a big challenge for computer users. Businessmen, professionals, and home users all have some important data that they want to secure from others. In this proposed system we have the software for data encryption and then embed the cipher text in an image with help of stego key [9]. This algorithm combines the effect of these two methods to enhance the security of the data. The proposed algorithm encrypts the data with a crypto algorithm and then embeds the encrypted text in an image file. This algorithm improves the security of the data by embedding the encrypted text and not the plain text in an image. It also calculates message digest of cover file to check integrity of message. This system will satisfy four requirements which we must require for secure data transmission. These are' Confidentiality', or Message Content Security, Integrity of Message Content, Authentication, Security in an open system.

To **conceal** a message

Plain text - encryption - concealment of text

To **extract** a message

Concealed text - decryption - plain text

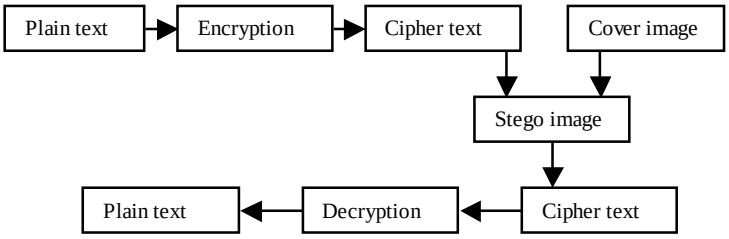


Fig. 2: Combination of steganograophy and cryptography

i. Encryption algorithm

The encryption algorithm built in is a shared key stream cipher algorithm which requires a secure exchange of a shared key that is outside the specification.The algorithm is used identically for encryption and decryption as the data stream is simply XORed with the generated key sequence. The algorithm is serial as it requires successive exchanges of state entries based on the key sequence. The algorithm features are

- Uses a variable length key from 1 to 256 bytes to initialize a 256-byte state table. The state table is used for subsequent generation of pseudo-random bytes and then to generate a pseudo-random stream which is XORed with the plaintext to give the ciphertext. Each element in the state

table is swapped at least once. This key is generated from stego key.

- The algorithm works in two phases, key setup and ciphering. Key setup is the first and most difficult phase of this algorithm. During a N-bit key setup (N being your key length), the encryption key is used to generate an encrypting variable using two arrays, state and key, and N-number of mixing operations. These mixing operations consist of swapping bytes, modulo operations, and other formulas.
- Once the encrypting variable is produced from the key setup, it enters the ciphering phase, where it is XORed with the plain text message to create an encrypted message..Once the receiver gets the encrypted message, he decrypts it by XORing the encrypted message with the same encrypting variable.

ii. Hashing algorithm

MD5 is a hashing algorithm that translates an indefinite length of data into a fixed length unique hash. It is useful in digitally sealing a message since it is next to impossible that two different messages can have the same "fingerprint" (digest) hence given a message fingerprint you can be sure of the message integrity In information hiding, the integrity of some data is necessary. The nature of these data depends on the application domain. For steganographic purpose, we need the hidden message to conserve its integrity. If the message is modified, the recipient must detect it. A simple solution is to hide in addition to the message, its hashed version. The MD5 message digest algorithm produces a 128-Bit hash-value of the given input data A message digest algorithm represents the functionality of an one-way hash function for computing a fixed sized data value (message digest, hash) from input data of arbitrary size. The length of the resulting hash value usually is shorter than the length of the input data. In this proposed system we embed this hash value in image file so that integrity of message contents are checked.

iii. Embedding data in 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 [10].

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 X768 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)
.After inserting the binary value for A.
(00100111 11101000 11001000)
(00100110 11001000 11101000)
(11001000 00100111 11101001)
The underlined bits are the only three actually changed in the 8 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.

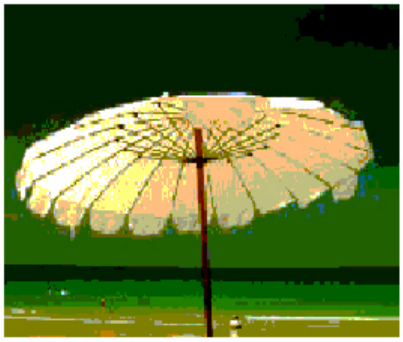
iv.Multi-level securities proposed in the algorithm:

To hide a text into the image:

1.Apply encryption algorithm on the text with a strong stream cipher mechanism The stego key is used in encryption of data.
2. The cipher text file is embedded into the stego medium.
3. The password or stego-key is also encrypted and embedded in image file.
4. The stego key is used to select random pixels for embedding data. The stego key (user-specified or default) is used not only to facilitate random selection of bytes for hiding message file bits but also is used to encrypt the message file. The encryption method is based on XORing the message bytes with random numbers generated by a pseudo-random number generator whose seed is derived from the stego key.
5. The message digest of image file is also embedded to check integrity of message

To extract a text from the image:

1.The stego key is required to extract the message.
2.Message digest is calculated to check the integrity of file
3 By extracting the LSBs from the stego image, a file containing cipher text is obtained.
4 This file is decrypted using encryption algorithm to get the original file
If either stego key or message digest does not match with embedded stego key and hash value of file then it will not be able to extract the message. An example of steganography system is as shown below.



VI. Results of Proposed system

An information hiding system, Secure Information Hiding System (SIHS) has been developed to provide confidentiality security service. SIHS employs an image file as a carrier to hide a message and focuses on Least Significant Bit (LSB) as one of the steganography techniques. The stego file does not reveal any difference in attributes like size, content etc from that of the original file. Hence it is difficult for someone to find out that this image contains a message.This is because the amplitude of the change is small, and therefore modulating the LSB does not result in a human- perceptible difference. Thus allowing high perceptual transparency of LSB**.**In the experiment, we found that the size of information to he hidden relatively depends on the size of the cover-image**.** The message size must be smaller than the image. On an average LSB requires only half of bits in an image to be changed. 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. Here we embed the password, message as well as message digests of file. The password so embedded is encrypted and while extracting message one need the correct password to get the correct message. Even if one character is altered in the password the original message can't be obtained

Hence a hacker must know the following in order to extract the embedded message from the image file.

1. Algorithm to extract the message from the image. (stego algorithm)
2. Encryption algorithm.
3. Correct password for algorithm.
4. Message digest algorithm for check integrity of message

With these increased levels of protection using encryption algorithm, the proposed system for steganography is more strong from attacks than any other existing system.

VII. Conclusion and Comments

Until recently, 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 confidentiality of information and provides a means of communicating privately. Here message is first encrypted and then embed in image file with help of steganographic system. The system security is further enhance by using a password to embed the message. 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.

Refernces

[1] Z. Hrytskiv, S**.** Voloshynovskiy & Y. Rytsar "Cryptography of Video InformationIn Modem communica -tion", Electronics And Energefics**,** vol. 11, pp. 115-125, 1998

[2] Stinsown,D."Cryptography:Theory and practice"

[3] C. Cachin, "An Information-theoretic Model for steganography", in proceeding 2nd Information Hiding Workshop**,** vol.1525, pp.306-318, 1998

[4] Neil F. Johnson, Zoran uric,Sushil.

Jajodia, " Information Hiding: steganography and Watermarking – Attacks and Countermeasures", Kluwer Academic Press, Norwrll, MA,New York, 2000

[5] R A Isbell, "Steganography: Hidden Menace or Hidden Saviour", steganography White Paper,IO May 2002

[6] J. Zollner, H. Federrath, H. Klimant, et al.,"Modeling the Security of Systems", Steganographic in 2nd Workshop on Informafion Hiding**,** Portland, April 1998, pp. 345-355. proceeding of IEEE**,** pp. 1062-1078, July1999.

[7] W. Bender, D. Gruhl, N. Morimoto and A.Lu, "Techniques for Data Hiding", Sysfems Journal**,** vol. 35, 1996

[8] M. M Amin, M. Salleh, S. Ibrahim, M .R. Katmin, and M. Z. I. Shamsuddin," Information Hiding using Steganography", IEEE 0-7803-7773-March 7,2003

[9] N. Provos, P. Honeyman, "Detecting Steganography Content on the Internet".Transformation", ZEICE Tram.

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

[11] Robert keern,"Steganography and implementation and detection", Jan 21,2004.

[12] Alain C. Brainos II East Carolina University ,"Study of steganography and The Art of Hiding of informatiom ", November 13,2003.