ISSN (Online): 2230-7893
www.IJCEM.org

A Data Hiding Model for Image Steganography Using Primes: Towards Data Security

Mamatha.T
Assistant Professor, Department of Computer Science & Engineering
Maulana Azad College of Engineering & Technology,
Patna, Bihar, INDIA.
mamta.macet@gmail.com

Abstract
Steganography is the art and science of communicating in such a way that the presence of a message cannot be detected. It includes the concealment of information within computer files. This paper considers an information theoretic model for steganography with a passive adversary being proposed. The adversary's task of distinguishing between an innocent cover message C and a modified message S containing hidden information is interpreted as a testing problem. The message is the composition of some character. Every character of the message can be represented as an ASCII value which is either even or odd. Depending on this evenness, the character is encrypted differently. This paper describes how an even-odd encryption based on ASCII value is applied and how encrypted message is converted by using Gray code and embedding with picture can secure the message and thus makes cryptanalyst's job difficult.

Keywords: *Cryptography, Least Significant Bit (LSB), Most Significant Bit (MSB), Red Green Blue (RGB), Steganography, Rivest Shamir Adleman (RSA)*

1. Introduction

Today networks are seriously threatened by network attacks. Cryptography may be used at different levels of a security model. This paper presents an approach of ASCII based cryptography with LSB based image stenography for security purpose of data transaction in the network and the internet. Here, encryption is applied to the even or odd ASCII value of the character which represents the data. A character in the plain text is always changed to the ASCII value and adding a key value with it gets back the cipher text. This value is then converted to the equivalent binary number. Substitute these bits in the LSB position in each pixel which describes the image. The receiver collects these bits from the image and converting them in equivalent decimal number which is the cipher text and subtracting the key value from it, we get the ASCII value of the plain text. Converting this ASCII value to the equivalent character representation, we get the original text. A cryptanalyst can normally find the key but in this approach a combination of two prime numbers is used for encryption.

2. Literature Review

2.1 Cryptography

Cryptography refers to the art of protecting transmitted information from unauthorized interception or tampering. The other side of the coin, cryptanalysis, is the art of breaking such secret ciphers and reading the information, or perhaps replacing it with different information. It uses mathematical & logical principles to secure information. Encryption means the change of original information (plain text) into another form by some operations (algorithm) and decryption means the techniques of getting the original information by some operation (algorithm) from the encrypted data (cipher text). In private key cryptography, the encryption and decryption on plaintext is done with the same key and key is known to the sender and receiver. In the public key cryptography, two different keys called private key and public key are used. The public key is known to all authorized users, but the private key is known to one person- its owner. Cryptography and steganography are well known and widely used techniques that manipulate information in order to cipher or hide their existence.

2.2 Merkle-Hellman Knapsacks

The well known cryptosystem was first described by Merkle and Hellman [1] in 1978.The basic idea behind the Merkle-Hellman encryption scheme is to create a subset problem which can be solved easily and then to hide the super increasing nature by modular multiplication and permutation. The transformed vector forms the encrypted message and the original super increasing vector forms the private key and is used to decipher the message.

2.2.1 Mathematical Explanation

1. The first step is to choose a super increasing sequence of numbers of positive integers. A superincreasing sequence

is one where every number is greater than the sum of all proceeding numbers.

$$S= (s_1.s_2,s_3,\ldots\ldots\ldots s_n)$$

2. The second step is to convert all the characters of the message into binary. The binary sequence is represented by the variable b.
3. Third step is to choose two numbers: an integer 'a' which is greater than the sum of all numbers in the sequence 's' and its co-prime 'r'. The sequence's' and the numbers 'a' and 'r' collectively form the private key of the cryptosystem. All the elements of 's' are multiplied with the number 'r' and the modulus of the multiple is taken by dividing with the number 'a'.

$$\text{i.e } p_i = r*s_i \bmod(a)$$

All the elements p_1,p_2,p_3 ,……p_n of the sequence p are multiplied with the corresponding elements of the binary sequence b. The numbers are then added to create the encrypted message M_i.The sequence M = (M_1, M_2, M_3 ,……M_n) forms the public key of the cryptosystem.

2.3 RSA Cryptosystem

In 1978, Ronald L Rivest, A.Shamir and Leonard M. Adleman [2] proposed a method for realizing public key encryption as suggested by Deffie and Hellman [3]. RSA is a public key algorithm that is used for encryption, Signature and Key Agreement.. RSA typically uses the size of 1024 and 2048. The RSA standard is specified RFC 3447, RSA Cryptography Specifications Version 2.1 RSA cryptography system with public keys, based on modular exponentiation is considered as the most reliable cryptography system in the world. An overview of RSA is given below where a participant creates the public and private keys.

2.3.1 Parameter Generation

1. Select two large prime numbers p and q.
2. Find n=p*q, where n is the modulus that is made public. The length of n is considered as the RSA key length.
3. Choose a random number 'e' as a public key in the range $0<e< (p-1)(q-1)$ such that $\gcd(e,(p-1)(q-1))=1$
4. Find private key d such that $(e*d) \bmod (p-1)(q-1))=1$

2.3.2 Encryption

Consider the device A that needs to send a message M to B securely

5. Let e be B's public key. Since e is public, A has access to it.
6. To encrypt the message M, represent the message as an integer in the range $0<M<n$. .
7. Cipher text $C = M^e \bmod n$, where n is the modulus.

2.3.3 Decryption

8. Let C be the cipher text received from A.
9. Calculate Message $M = C^d \bmod n$, where d is B's private key and n is the modulus.

It is easy to generate large prime numbers and multiply them but it is extremely difficult to factor the product. The RSA technique is costly, relatively slow and thereby limiting the throughput rate.

3. Steganography

Steganography is the art and science of hiding information by embedding messages within other, seemingly harmless messages. Steganography means "covered writing" in Greek. As the goal of steganography is to hide the presence of a message and to create a covert channel, it can be seen as the complement of cryptography, whose goal is to hide the content of a message. The object may be an image, audio, video or text only. A famous illustration of steganography is Simmons' Prisoners' Problem [4][5]. An assumption can be made based on this model is that if both the sender and receiver share some common secret information then the corresponding steganography protocol is known as the secret key steganography where as pure steganography means that there is none prior information shared by sender and receiver. If the public key of the receiver is known to the sender, the steganographic protocol is called public key steganography [4-6]. Almost all digital file formats can be used for steganography, but the image and audio files are more suitable because of their high degree of redundancy [8]. Digital file formats can be used for steganography, but the formats that are more suitable are those with a high degree of redundancy. Redundancy can be defined as the bits of an object that provide accuracy for greater than necessary for the object's use and display [4]. The redundant bits are those bits that can be altered without the alteration being detected easily [5]. An obvious method was to hide a secret message in every n^{th} letter of every word of a text message. Digital representation of image has large amount of redundant bits and for that images are the most popular cover objects for steganography.

3.1 Image Steganography

Image steganography technique can be divided into two groups: those in the Image domain and those in the Transform Domain [7]. Image domain technique embed message in the intensity of the pixels directly, while for transform domain, images are first transformed and then the message is embedded in the image [8].

Image domain techniques encompasses bit-wise methods that apply bit insertion and noise manipulation and are sometimes characterized as simple system but in the transform domain involves the manipulation of the algorithms and image transforms [8].A block diagram of a

ISSN (Online): 2230-7893
www.IJCEM.org

generic image steganographic system is given in Fig.(1). A message is embedded in a digital image (cover image) through an embedding algorithm, with the help of a secret key. The resulting stego image is transmitted over a channel to the receiver where it is processed by the extraction algorithm using the same key. During transmission the stego image, it can be monitored by unauthenticated viewers who will only notice the transmission of an image without discovering the existence of the hidden message.

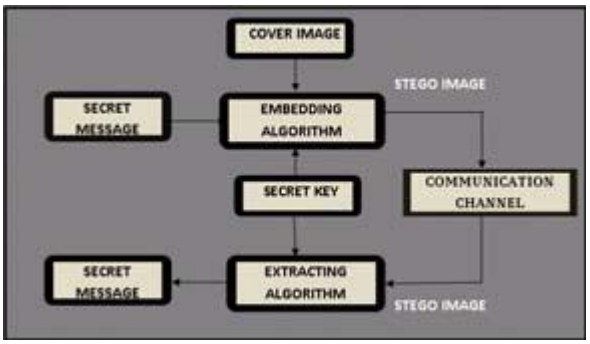

Fig.1 Generic form of image steganography

In this paper a specific secret-key image based data hiding model has been proposed which uses an image as the cover data and the secret information is embedded in the cover to form the stego image.

4. Background

In the proposed approach an even and odd numbers representing the plain text (which are the ASCII values) are treated differently. We know that the sum of an even number and an odd number is odd and sum of two odd numbers is even. Considering this axioms, the two key values chosen should be odd and relatively prime. If two odd values are chosen, then, even and odd numbers representing the plaintext gets converted to odd and even respectively. So at the time of decryption little bit opposite task should be performed. If $g_n \ldots\ldots\ldots g_2 g_1 g_0$ denote a code word in the $(n+1)^{st}$-bit Gray code and $b_n \ldots\ldots\ldots b_2 b_1 b_0$ designate the corresponding binary number, where the subscripts 0 and n denote the least significant and most significant digits, respectively. Then the i^{th} digit g_i can be obtained from the corresponding binary number as follows:

$$g_n = b_n$$
$$g_i = b_i \oplus b_{i+1}, \quad 0 \le i \le n-1$$

To convert the Gray codes to binary number follow the process:

$$b_n = g_n$$
$$b_i = g_i \oplus b_{i+1}, \quad 0 \le i \le n-1$$

Least significant bit insertion is to embedding information in a cover image. The least significant bit of some or all of the bytes inside an image is changed to a bit of the secret message. When using a 24-bit image, a bit of each of the red, green and blue color components can be used, since they are each represented by a byte. In other words, one can store 3 in each pixel. On average, only half of the bits in an image will need to be modified to hide a secret message using the maximum cover size [10]. Since there are 256 possible intensities of each primary color, changing the LSB image steganogrpahy of a pixel results in small changes in the intensity of the colors. These changes cannot be perceived by the human eye. So message is successfully hidden.

5. Proposed Method

In this method a shared key pair ($Even_K$, Odd_K) which are odd and prime are chosen and the ASCII value representing the character is tested for evenness. If it is even, then $Even_K$ is added to the number and if it is not even then Odd_K is added to it. When we add odd number with an even number results to an odd number and in the same fashion an odd number added with an odd number results an even number. On the decrypting side each number is tested for evenness and if it is even then Odd_K is subtracted but if it is not even then $Even_K$ is subtracted.

5.1 Forward Process

1. Choose a pair of key value ($Even_K$, Odd_K) which are primes
2. M = ASCII value of a character in the message
3. If M mod 2 = 0 then
   $C = M + Even_K$
   Else
   $C = M + Odd_K$
4. Convert C into equivalent binary number
5. Convert the binary to the Gray code
6. Substitute this bit in the LSB position of the Image pixel
7. Send the image to the receiver.

In this algorithm, M is the ASCII value of plaintext and C is the number representing the cipher text. The two numbers $Even_K$ and Odd_K are the components of the shared key. The term "mod" indicates that the remainder obtained by dividing M by 2 is used for comparing evenness.

5.2 Backward Process

1. Extract the bits from the image and group by 8 bits.
2. Convert the Gray code to the equivalent binary number separately.
3. Convert binary number to equivalent decimal

number which is C

4. If C mod 2 = 0 then
   $M = C - Odd_K$
   Else
   $M = C - Even_K$
5. Convert M (which is the ASCII value) to the equivalent character.
6 Putting this character we shall get the original message

In this algorithm M and C are the same as before. The two numbers $Even_K$ and Odd_K are the components of the same shared key.

5.3 Example

Suppose that the shared key is the pair (11, 19). The sender needs to send the message "abcxyz". The ASCII values of (a, b, c, x, y, z) are (97, 98, 99, 120, 121, 122) respectively. Then encryption process calculates for each character as: 97 mod 2 = 1 so C = 97+ 19 = 116, for second character b as 98 mod 2 = 0; so C= 98+11=109. Similarly for rest of the characters as:
99 mod 2 =1 so C= 99+19 =118
120 mod 2 =0 so C= 120+11=131
121 mod 2 =1 so C= 121+19=140
122 mod 2 =0 so C= 122+11=133
These numbers (78, 91, 77, 195, 138, and 199) are the cipher text. Now converting this number to the equivalent binary number, we get (01110100, 01101101, 01110110, 10000011, 10001100, 10000101) respectively. Converting this number to gray code we get (01001110, 01011011, 01001101, 01011001, 10001010, 11000111) respectively. Each bit represents a pixel. Three 8-bit bytes, one byte for each of RGB, is called 24 bit color. Each 8 bit RGB component can have 256 possible values, ranging from 0 to 255. Now a grid of 16 pixels of a 24-bit image can hold all this gray code in the LSB position as:

01010000	01100111	01001100
01101000	01100011	01000101
00111001	01101110	01010100
00101011	01110010	01101001
00101001	01000010	01101111
00111011	01000110	01001101
01101000	01100010	01100101
01111011	01101010	01011101
01010011	01100111	01001100
01101000	01101110	01010010
00101011	01110010	01101101
01001110	01011010	01011010
01010111	01010110	01101011
01101010	01011011	01010111
01011000	01011000	01010010
01101011	01101111	01101001

Changing LSB in each value would allow minor variations in color and unnoticeable to human eye. This image is send to the receiver. In the receiving end, simply extract the appropriate LSB bits from the image and group it by 8 bits i.e. 01110100, 01101101, 01110110, 10000011, 10001100, 10000101. Now convert this into binary number and that is 01110100, 01101101, 01110110, 10000011, 10001100, and 10000101 respectively. Convert this binary number to equivalent decimal number which are 116,109,118,131, 140,133 respectively and it is the cipher text. The decryption process calculates 116 mod 2 as 0 so M = 116 - 19 = 97. Similarly

109 mod 2 as 1 so M=109-11 = 98
118 mod 2 = 0 so M=118-19 = 99
131 mod 2 = 1 so M=131-11 =120
140 mod 2=0 so M=140-19=121and similar for rest of the cipher text. These numbers are the ASCII value of the message. Now converting this value to the equivalent character we get the message "abcxyz" which is the original message.

6. Security and Concern

There is nothing common in between two numbers rather than both of them are odd prime number. If one number is known to the adversary, he cannot deduce the other number. In case of a 32 bit machine (long integer of 32 bits), each number can be 32 bits long. If one number is fixed, the other number can be any one of 232 possibilities and the first number can be one of 232 possibilities. So the number of possible alternatives becomes 232*232 = 264. Trying possible alternatives are not so easy. Further the steganography itself will hide the converted information in a secured way so that human eye cannot easily detect it.

Conclusion

Steganography is the art and science of writing hidden messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message, a form of security through obscurity. Cryptography is an essential ingredient in this revolution, and is necessary to preserve privacy from computerized censors capable of scanning millions of pages of documents for even one sensitive datum. Though the cipher text can be broken, this even-odd based cryptography differentiates the encryption scheme to be applied based on the evenness or oddness of the ASCII value of the character. Further this encrypted message converted by using Gray code embedding with picture can reduce the security tension.

References

[1] R.C. Merkle and M. Hellman, Hiding Information and Signatures in Trap Door Knapsacks, IEEE Trans. Inform. Theory, vol 24 1978,pp 525-530.

[2] R. L. Rivest, A Shamir and L. Adleman, A method for obtaining digital signatures and public-key cryptosystems, Communications of the Association for Computing Machinery, vol 21, no.2, pp 120-126.

[3] W. Diffie and M. E. Hellman, New direction in cryptography, IEEE Transactions on Information Theory, vol. IT- 22 ,no. 6,pp.644-654.K. Elissa,

[4] Souvik Bhattacharyya. and Gautam Sanyal. Study of secure Steganography model In Proceedings of International Conference on Advanced Computing and Communication Technologies (ICACCT-2008), Panipath , India 2008.

[5] G. Simmons, The prisoners problem and the subliminal channel, CRYPTO, 1983

[6] Souvik Bhattacharyya. and Gautam Sanyal. An Image Based Steganography Model for Promoting Global Cyber Security. In Proceedings of International Conference on Systemics, Cybernetics and Informatics, Hyderabad, India , 2009.

[7] K. Ahsan and D. Kundur, Practical Data hiding in TCP/IP, Proceedings of the workshop on Multimedia security at ACM Multimedia, 2002

[8] J.Silman, Steganography and Steganalysis: An Overview, SANS Institute, 2001

[9] Y.K. Lee and L.H. Chen, High capacity image steganographic model, Visual Image Signal Processing,147: 03, June 2000

[10] R. Krenn, Steganography and Steganalysis, www.krenn.nl/univ/cry/steg/article.pdf

Author

Mamatha.T is working as Assistant Professor in Computer Science & Engineering Department at Maulana Azad College of Engineering & Technology, Patna, The college is affiliated to Magadh University, Bodh Gaya, Bihar, INDIA and approved by All India Council of Technical Education (AICTE), New Delhi, INDIA. She has presented papers in many International and National Conferences. She is a life member of Computer Society of India.