Military Technical College
Kobry El-Kobba,
Cairo, Egypt

11-th International Conference
on Aerospace Sciences &
Aviation Technology

AN EMBEDDING ALGORITHM FOR DATA HIDING

Alaa Fahmy[1]

ABSTRACT

Data hiding is frequently termed steganography, which is very close to cryptography. The purpose of cryptography is to make messages unintelligible so that only the intended receiver who posses the secret keys can recover the messages. Sometimes, it may be desirable to achieve security and privacy by masking the messages instead of encrypting it. This problem is addressed by steganography, which has been used in spy applications. If Alice (spy in a foreign country) wants to send messages abroad, she will use a local communication channels. On the assumption that the communication channel is monitored all the time hence, sending an encrypted messages would raise suspicion in her. Therefore, Alice would prefer to use steganographic technique rather than a cryptographic one. An embedding algorithm, based on elliptic curve has been presented to provide both data encryption and data hiding in digital images.

KEYWORDS

Steganography, and Data hiding.

1. INTRODUCTION

Data hiding is frequently termed steganography [1], which is very close to cryptography. The purpose of cryptography is to make messages unintelligible so that only the intended receiver who posses the secret keys can recover the messages. Sometimes, it may be desirable to achieve security and privacy by masking the messages instead of encrypting it. This problem is addressed by steganography. Dating back, the first steganographic techniques included invisible writing by using special inks.

Today, the binary files are used to hide data. Digital images, and videos are suitable for this purpose. The hidden message may have no relationship to the carrier image in which it is embedded in contrast to secure communication. In this paper, we introduce an embedding algorithm based on elliptic curve cipher [2] to provide both data encryption and data hiding in digital images. The rest of the paper includes the

[1] Assoc. Prof. Dept. of Electrical Engineering, Military Technical College, Cairo, Egypt.

following: section 2 presents an overview of data hiding. Section 3 introduces the basic concepts of elliptic curves. Section 4 introduces the idea of the embedding algorithm to hide data. Section 5 concludes the paper.

2.Data Hiding Overview

Data hiding technique consists of the embedding algorithm and detector function. The embedding algorithm is used to hide secret messages inside a carrier document (it could be text/image). The embedding process is protected by a key so that only those who posses the secret key can access the hidden message. The detector function is applied to a carrier to get back the hidden secret message as shown in Fig.1 [3]. The most important properties of data hiding schemes are robustness, undetectability, invisibility, security, and complexity, which are defined below [4].

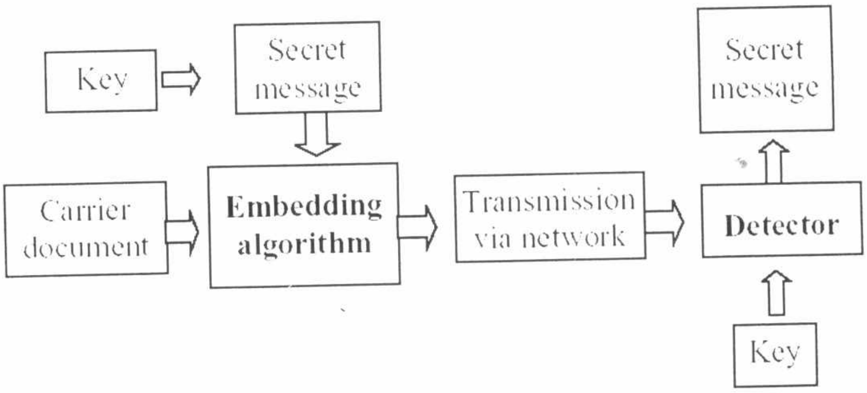


Fig.1 A Simplified Block Diagram For Data Hiding

Robustness
The embedded information is said to be robust if its presence can be reliably detected after the image has been modified but not destroyed beyond recognition. Robustness means resistance to "blind", non-targeted modifications, or common image operations.

Undetectability
It is required for secure covert communication. We say that the embedded information is undetectable if the image with the embedded message is consistent with a model of the source frorn which images are drawn. The concept of undetectability is inherently tied to the statistical model of the image source. If an attacker has a more detailed model of the source, he may be able to detect the presence of a hidden message. Meanwhile, the ability to detect the presence does not imply the ability to read the hidden message.

Invisibility
It is based on the properties of the human visual system. The embedded information is imperceptible if an average human subject is unable to distinguish between carriers that do contain hidden information and those that do not. A commonly

accepted experimental arrangement, which is called blind test, frequently used in psycho-visual experiments is based on randomly presenting a large number of carriers with and without hidden information and asking the subjects to identify which carriers contain hidden information. Success ratio close to 50% demonstrates that the subjects cannot distinguish carriers with hidden information.

Security
The embedding algorithm is said to be secure if the embedded information cannot be removed beyond reliable detection by targeted attacks based on a full knowledge of the embedding algorithm and the detector (except the secret key), and the knowledge of at least one carrier with hidden message. The concept of security also includes procedural attacks, such as the IBM attack [5], or attacks based on a partial knowledge of the carrier modifications due to message embedding [6].

Secure black-box public detector
It is a message detector implemented in hardware. It is assumed that the box cannot be reverse-engineered. The secret key used to read the hidden messages is wired-in the black box and cannot be recovered. The availability of the black box should not enable an attacker to recover the secret key or remove the hidden information from the carrier (again, we assume that the attacker has a full knowledge of the embedding algorithm and the inner workings of the detection function).

Secure public detector
It is an even stronger concept for which all details of the detector are publicly known. If such a detector is ever built, it would find tremendous applications since it can be implemented in software rather than tamper-proof hardware. It would enable building intelligent Internet browsers capable of filtering images containing certain marks.

Conflicting requirements
The above requirements are mutually competitive and cannot be clearly optimized at the same time. If we want to hide a large message inside an image, we cannot require at the same time absolute undetectability and large robustness. A reasonable compromise is always a necessity. On the other hand, if robustness to large distortion is an issue, the message that can be reliably hidden cannot be too long.

3. Elliptic Curves

A plane curve is defined to be the set of points satisfying an equation F(x,y)=0. The simple curves are lines and conic sections (degree 2 in x and y). The next simplest are cubic curves (degree 3). These include elliptic curves, so called because they arose historically from the problem of computing the circumference of an ellipse [7].

3.1 Weierstrass Equation

The most common equation to define the elliptic curves are known as Weierstrass equations [8]. For the prime field GF(p) with p>3, the Weierstrass equation is:

$$Y^2 = X^3 + aX + b$$

Where a, and b are integers modulo p for which $4a^3 + 27b^2 \neq 0 \pmod{p}$. For the binary finite fields $GF(2^m)$, the Weierstrass equation is:

$$Y^2 + XY = X^3 + aX^2 + b$$

Where a, and b are elements of $GF(2^m)$ with $b \neq 0$. The elliptic curve E consists of the solutions (x,y) over GF(q) to the defining equation, along with an additional point called the point at infinity (denoted O). The points other than O are called finite field points. The number of points on E (including O) is called the order of the curve E and denoted by #E(GF(q)).

3.2 Operation On Elliptic Curves

There are two basic operations , namely addition, and multiplication.

3.2.1 Addition Operation

Define the inverse of the point P =(x,y) to be:
-P = (x, -y) if q =p prime,
= (x, x+y) if $q = 2^m$.
Then, the sum P + Q of the points P and Q is the point R, with P, Q, and –R lie on a curve, with the property P + O =P, and P + (-P)=O, for all points P. To illustrate the addition operation on E over Z_p, let $P=(x_1,y_1)$, and $Q=(x_2,y_2)$ are points on E. If $x_2=x_1$, $y_2=-y_1$, then P+Q=O. Otherwise $P+Q=(x_3,y_3)$ where $x_3=\lambda^2-x_1-x_2$, $y_3=\lambda(x_1-x_3)-y_1$,
$\lambda =(y_2-y_1)/(x_2-x_1)$ if $P\neq Q$
$=(3x_1^2 +a)/2y_1$ if P=Q

3.2.2 Scalar Multiplication

Elliptic curve points can be added but not multiplied. However, it possible to perform scalar multiplication, which is another name for repeated addition of the same point. If n is a positive integer and P a point on E, then the scalar multiplication is nP (adding P n times), with the property OP =O, and (-n)P = n(-P). Meanwhile Meneze, Vanstone (MQV) [9], assume that the points P,Q, and –R could not lie on E.

4. The Embedding Algorithm

The embedding algorithm is based on the use of elliptic curve cryptosystem (ElGamal cryptosystem [10]). The hiding is applied to two different secret messages, RGB true color digital image and a text file. The embedding algorithm combines both the secure elliptic curves and a steganographic method similar to Least Significant Bit method (LSB). It encrypts the secret message before hiding in the carrier image.

4.1 Embedding Algorithm operation

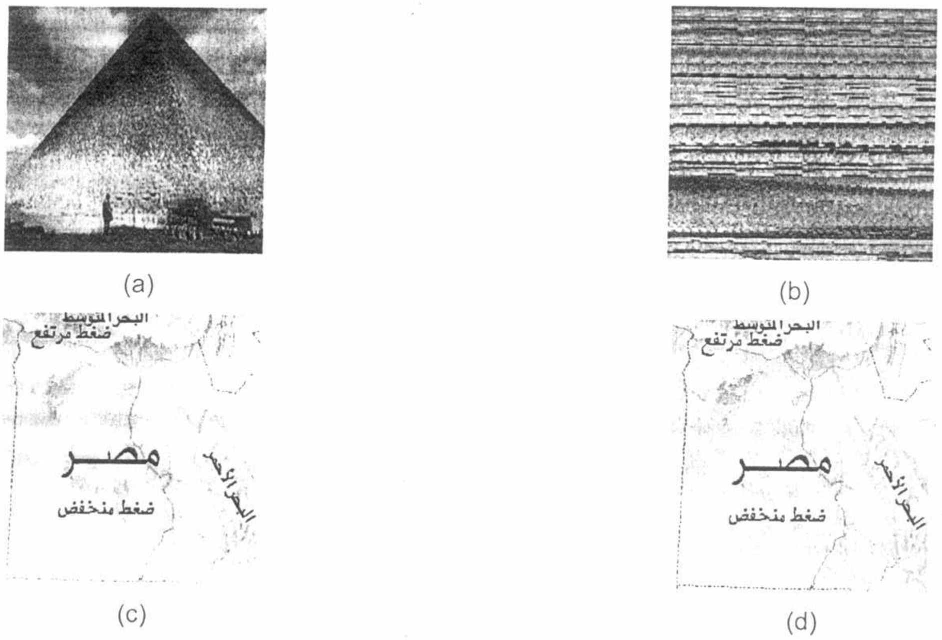
The encryption and hiding process have been done by software programming (Visual basic), and using an elliptic curve E such that:

$$Y^2+XY=X^3+a_2X^2+a_6 \text{ over } GF(2^{63}).$$

The procedures are done as follows:

(1) Read the secret message, which is considered as a digital RGB square image of 128 x 128 pixels, and 24 bits depth, where the red, green, and blue components are 8 bits each.
(2) Select an elliptic curve and test its non singularity, to be used for encryption.
(3) Pickup each two pixels of the secret image, which can be represented by the points P, and Q.
(4) Perform the addition/scalar operations for these two points/pixels, and follow the previous operations mentioned in section 3.2, to get a new point R=(x_3,y_3).
(5) Repeat steps 3 &4 to encrypt the secret image.
(6) Put the encrypted image in the LSB of the carrier image 256x256 pixels, which slightly modifying the red, green, and blue levels of the pixel's carrier image. The modifications will have the properties of a thermal, Gaussian noise commonly present in digitized images taken with an ordinary scanner [11]. The resulting image is a random collection of pixels with randomly distributed color levels without any spatial correlations.
(7) To extract the secret image, we follow the procedures of ElGamal cryptosystem based on elliptic curves [10].

An example of a secret image, and a carrier image with an embedded secret image by using Baker map [4] is shown in Fig.2. The secret image (128x128) pixels of 192 Kilobyte is encrypted using the technique mentioned in [10], then hiding the encrypted secret image in 256 x 256 pixels of 209 Kbytes true color image.



(a) (b)

(c) (d)

Fig.2 Secret image hiding in image by using Baker map, (a) Secret image 128 x 128 pixels, (b) Encrypted image using the improved technique, (c) Original carrier image 256 x 256 pixels, (d) Carrier image with hidden secret image

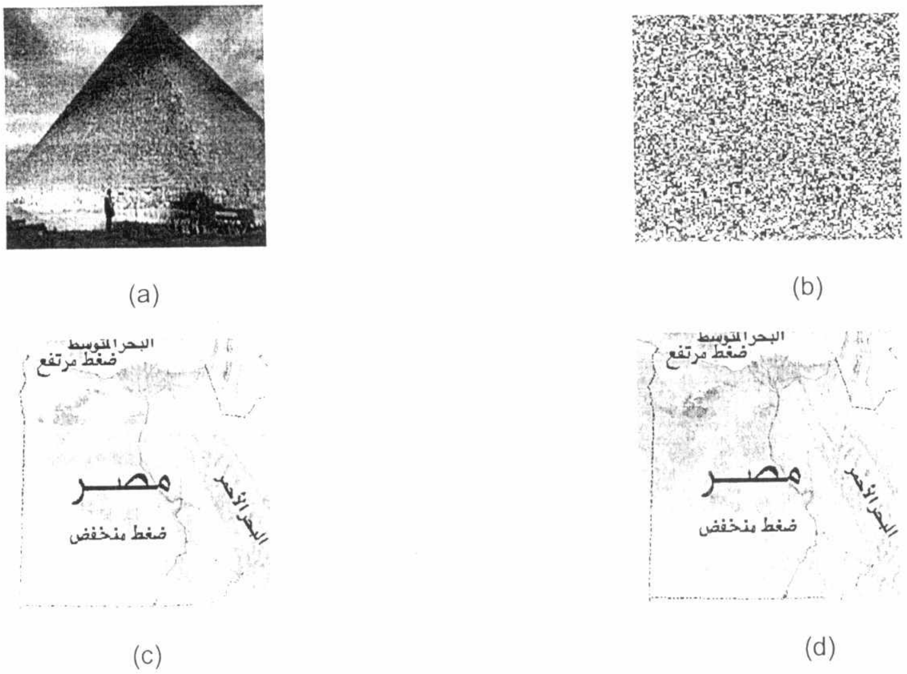


Fig.3 Secret image hiding in image by using Elliptic curves, (a) Secret image 128 x
(b) Encrypted image using the improved technique, (c) Original carrier 128 pixels,
image 256 x 256 pixels, (d) Carrier image with hidden secret image

Figure 3 shows, the same secret image and carrier image have been used with an embedded algorithm based on elliptic curve. It is clear that in both cases the hiding has been achieved without greatly affecting the quality of the image, but in the first example the encrypted image (48 Kbytes) still has the same features of the original image. While in the second example, which uses an elliptic curve , the encrypted image (48 Kbytes) resembles a Gaussian noise.

Both the sender and the recipient need the original unmodified carrier image and a secret key for encrypting the secret image. Even if an eavesdropper gets hold of the original image, the secret image is still protected by the cryptosystem. Thus, the scheme for hiding images provides a high degree of security and does not raise a suspicion that any secret information is being sent. The used cryptosystem has two purposes in the whole scheme. First, it increases the security of the scheme. Second, it converts the secret image into an uncorrelated, random looking image which, when encoded into the carrier, resembles a thermal Gaussian noise commonly present in digital images [12]. The appearance of the carrier with a hidden secret message makes an eavesdropper not even suspects that secret information is being sent.

Actually, in the case of color images, there is even more room for hiding messages because each pixel is a triple of red, green, and blue. Again, replacing two or more least significant bits of each pixel increases the capacity of the scheme but at the same time the risk of making statistically detectable changes also increases. Therefore, it is important to study the security of each specific steganographic technique and argue why it is secure. Even the simple least significant bit encoding may under certain circumstances introduce detectable changes. Aura [13] suggests to change only a small fraction of the carrier bits. For example, modify each hundredth pixel in the carrier by one gray level. Depending on the image noise, these changes will hopefully be compatible with the uncertainties involved with any statistical model of the image. Before any secret message hiding technique can be claimed as secure, we need to carefully investigate the carrier images and their statistical properties. The noise component may not be uniform within the image but may depend on the pixel position in the image. For example, pixels corresponding to a bright white color may be saturated at 255 even though the overall model of the noise can be Gaussian with a non-zero variance.

5. Conclusion

Data hiding is frequently termed steganography, which is very close to cryptography. If Alice, which is a spy in a foreign country, wants to send messages abroad, she will use a local communication channels. On the assumption that the communication channel is monitored all the time hence, sending an encrypted messages would raise suspicion in her. Therefore, Alice would prefer to use steganographic technique rather than a cryptographic one. An embedding algorithm, based on elliptic curve has been presented to provide both data encryption and data hiding in digital images. A carrier image should be selected for a proper and secure hiding. Therefore, the appearance of the carrier with a hidden secret message makes an eavesdropper not even suspects that secret information is being sent.

References

[1] Bruce Schneier, "Applied Cryptography", 2nd edition, John Wiley & Sons, Inc,1996.

[2] IEEE Standard Specifications for public key cryptography, IEEE std 1363-2000.

[3] W. Bender, " Techniques for Data Hiding ", IBM System Journal , Vol.35, No.3-4, pp.313-336, 1996.

[4] Jiri Fridrich, "Applications of data hiding in digital images", ISPACS'98 Conference in Melbourne, Australia November 4-6, 1998

[5] S. Craver, N. Memon, B. Yeo, and M. Yeung, "Can Invisible Watermarks Resolve Rightful Ownerships?" Technical Report RC 20509, IBM Research Division, July 1996.

[6] J. Fridrich, "Robust digital watermarking based on key-dependent basis functions", The 2nd Information Hiding Workshop in Portland, Oregon, April 15–17, 1998.

[7] Silverman, J. "The Arithmetic of Elliptic Curves", Springer_Verlag, 1986.

[8] Meneze, A., Okamoto, T., and Vanstone, S. "Reducing Elliptic Curve Logarithms to Logarithms in a Finite Field", IEEE Transactions on Information Theory 39 (1993), pp. 1639-1646.

[9] Donglasr Stinson," Cryptography Theory and Practice", 2nd edition, Chapman & Hall/CRC, 2002.

[10]T. ElGamal, "A public key cryptosystem nd signature scheme based on discrete logarithms", IEEE Trans. On information theory,IT-31, no.4, pp.469-472, 1985.

[11] J. Fridrich, "Secure Image Ciphering Based on Chaos", Final Technical Report RL-TR-97-155, Rome Laboratory, New York, March 1997.

[12] J. Fridrich, " Symmetric Ciphers Based on Two-Dimensional Chaotic maps", International Journal of Bifurcation and Chaos, Vol. 8, No. 6, June 1998.

[13] T. Aura, "Invisible communication", Proc. of the HUT Seminar on Network Security '95, Espoo, Finland, Nov 1995. Telecommunications Software and Multimedia Laboratory, Helsinki University of Technology.