Anti-forensics: the image asymmetry key and single layer perceptron for digital data security

Anti-forensics: the image asymmetry key and single layer perceptron for digital data security

D Mualfah[1*],Y Fatma[1], and R A Ramadhan[2]

[1] Universitas Muhammadiyah Riau, Pekanbaru, Indonesia
[2] Universitas Islam Riau, Pekanbaru, Indonesia

*Email: destimualfah@umri.ac.id

Abstract. Accessing or theft of public network information is very vulnerable to data transmission on public networks, this has an impact on the loss of those who have the information. In the digital forensic world, securing data is a technique for anti-forensic that aims to keep data safe. One way to secure information when sending is done is to encrypt the data and information. Encryption techniques play an important role in cryptography to hide text or messages into an image as a public key. The use of images as key objects is very effective for data security and the information cannot be read or unknown by unauthorized persons. In this case the asymmetric key cryptographic technique uses images and the single layer perceptron, asymmetric cryptography through the encryption and description process can produce data and information security keys that are difficult to guess. The use of keys is done by extracting image features using the GLCM (grey level co-occurrence matrix) to produce random symbols perfectly through the training process using a single layer perceptron encryption asymmetry key and description that can be accounted for integrity.

1. Introduction

Various criminal and criminal acts at the present time involve directly or indirectly information technology and data security on internet media, one of which is the sending of emails, websites and other computer uses that can invite various parties to commit information technology-based crimes freely. So, it is very prone to tapping both passive and active information by irresponsible parties [1].

Therefore, nowadays the development of digital forensics science is needed and is used by law enforcers in their efforts to disclose criminal events through the disclosure of entity-based evidence or digital and electronic devices used for security [2]. In data and information security, there are two techniques that can be used to secure information and data security, namely cryptography and steganography.

In this case the steganography technique is used for the concept of anti-forensic [3] which is used to hide a message in another message in the form of digital media [4], by inserting a text message into the media image. Image media is widely used because the image becomes a digital object that is the most difficult to distinguish between original images without text insertion and images that have been inserted text in the form of secret messages.

To keep the data safe one of them is by implementing cryptographic techniques. The oldest cryptographic method is symmetry key cryptography, where the encryption and decryption of messages use the same key [5]. The unequal distribution of symmetry keys through communication channels gave rise to asymmetric cryptography in 1976 which arose the inequality of symmetry key distribution through communication channels so that symmetry key cryptography was produced, where asymmetric key cryptographic processing was done through the process of encrypting and describing public and private keys.

Furthermore, the use of artificial intelligence can be combined in cryptography to be used as an anti-forensic technique to protect data and information to design a symmetry key cryptography using artificial neural networks (ANN) backpropagation [6]. Backpropagation ANN is obtained from the training process as a public key and private key through the process of encryption and description of data that results in different ciphertext. This research can produce asymmetric key cryptography from backpropagation ANN images for data security.

2. Basic Theory

2.1. Cryptography

Cryptography is a branch of science that uses mathematical equations to carry out the process of encryption and decryption of data. This technique is used to convert or convert data into certain codes, with the aim that information stored or transmitted over insecure networks such as the internet, cannot be read by anyone except by an authorized person [7].

In the Oxford English dictionary explains the understanding of cryptography as a secret technique in writing, with special characters, by using letters and characters outside their original form, or by other methods that can only be understood by those who process keys, as well as all other things. So, it can generally be interpreted as the art of writing or solving ciphers [8].

While the technique used to decrypt messages without knowing knowledge about the details of encryption is called cryptanalysis. Cryptography and cryptanalysis hereinafter referred to as cryptology. The cryptographic system consists of 5 parts, namely:

1. Plaintext (M) Plaintext is a message or data input whose original form can be understood. Plaintext is input for encryption.
2. Secret key (K)
3. The secret key is also an input for encryption. The secret key is a free value that will later be imposed on the original text and affect the output.
4. Ciphertext (C) Ciphertext is the result of an encryption algorithm that no longer has any true meaning.
5. Encryption algorithm (Ek) Has two data inputs namely plaintext and secret key. The encryption algorithm transforms the plaintext so that it produces a message that has been encrypted. $Ek{:}M \to C$, where $k \in K$.

2.2. Image Processing

Imagery is another term for image. The word image is used more in material relating to conceptual and technical aspects, while the word image is used when referring to the object being discussed in daily life. The image can be defined as a function (x,), where x and y are flat plane coordinates, and the value of the function f in each coordinate pair (x,) is called the intensity or grey level of the image at that point [9]. Digital imagery contains a number of basic elements. These basic elements are manipulated in image processing and further exploited in computer vision.

2.3. Grey Level co-occurrence Matrix (GLCM)

GLCM (grey level co-occurrence matrix) is a feature extraction method that is widely used in classification because it can provide detailed information about an image in terms of texture [10]. In the feature extraction process using GLCM, the image will be converted into grayscale format so that for each pixel in the image region there will only be 1 grey value.

2.4. Perceptron

Perceptron is a simple form of neural network. Perceptron is usually used to classify a certain type of pattern that is often known as linear separation. Basically, perceptron in a neural network with one layer has an adjustable weight. The algorithm used by the rules of this perception will set its free parameters through the learning process. This activation function is made in such a way that there are restrictions between positive and negative regions

2.5. Anti-Forensic

According to [3] anti forensics is the "Application of the scientific method to digital media in order to invalidate factual information for judicial review." If digital forensics focuses on the act of searching, maintaining, identifying, retrieving, documenting, and making anti-forensic reports is the opposite Anti-forensic aims to keep data secure so that it cannot be opened or even read by other parties except the owner of the data and information. The purpose of anti-forensics is to thwart investigative actions and all acts of searching data on electronic devices. Basically, the purpose of anti-forensics is:

(1) Making digital data inaccessible and accessible, for example by hiding, providing passwords, encrypting, deleting, changing data integrity.

(2) Make an effort to make digital evidence unfeasible with legal standards, due to changes in the integrity of the data found. Anti-forensic activities can prolong the work of investigators because of the level of difficulty that digital forensic investigators must against.

3. Methodology

The key generation scheme using the single layer perceptron can be used for data security with key generation consists of two parts, namely the process to get the value of the image and the training process on artificial neural networks (ANN). The process of obtaining values from images is done using the image feature extraction method, namely the grey level co-occurrence matrix (GLCM) and colour moments. The value of the image extraction then becomes input into the training process as a network weight, the training process is carried out using backpropagation ANN.

Input on the training process consists of training data and weights from the extraction of image values, training data consist of alphanumeric characters and punctuation marks totalling 95 characters. Thus, the activation function used is a binary sigmoid activation function whose output from the training process is the public and private key. Then the encryption and decryption process is done using the backpropagation ANN approach which is applied using the input layer architecture to the hidden layer, while the decryption process is done using the hidden layer architecture to the output layer.

Furthermore, for communication, each party, the sender and receiver of the message, first conducts a key generation process, where the public unci is distributed to communication partners and the private key which is confidential is kept by each key generating party. Security services used are confidentiality (confidentiality) and authentication (authentication). An overview of the system is shown in Figure 1.

By scenario, there are two parties who will communicate confidentially, namely the sender of the message (call it A) and the recipient of the message (call it B). A and B then together do the key generation process, the resulting key is in the form of a public key (Kpu) and a private key (Kpr). B then sends his public key (KpuB) to A, and vice versa A sends his public key (KpuA) to B. A encrypts secret messages using his private key (KprA) to fulfil the authentication security service that produces ciphertext1. A then encrypts ciphertext1 again using public key B (KpuB) for confidentiality security services that produce ciphertext2. A then sends ciphertext2 to B. By B ciphertext2 is decrypted using KprB which produces ciphertext1. Then B decrypts the ciphertext1 again using KpuA to produce a plaintext, which is the original message sent by A. The following illustration shows encryption and decryption based on its security service.

Furthermore, encryption and decryption of messages to maintain confidentiality as well as authentication to meet the security aspects of both requires a public and private key both owned by sender A and receiver B.

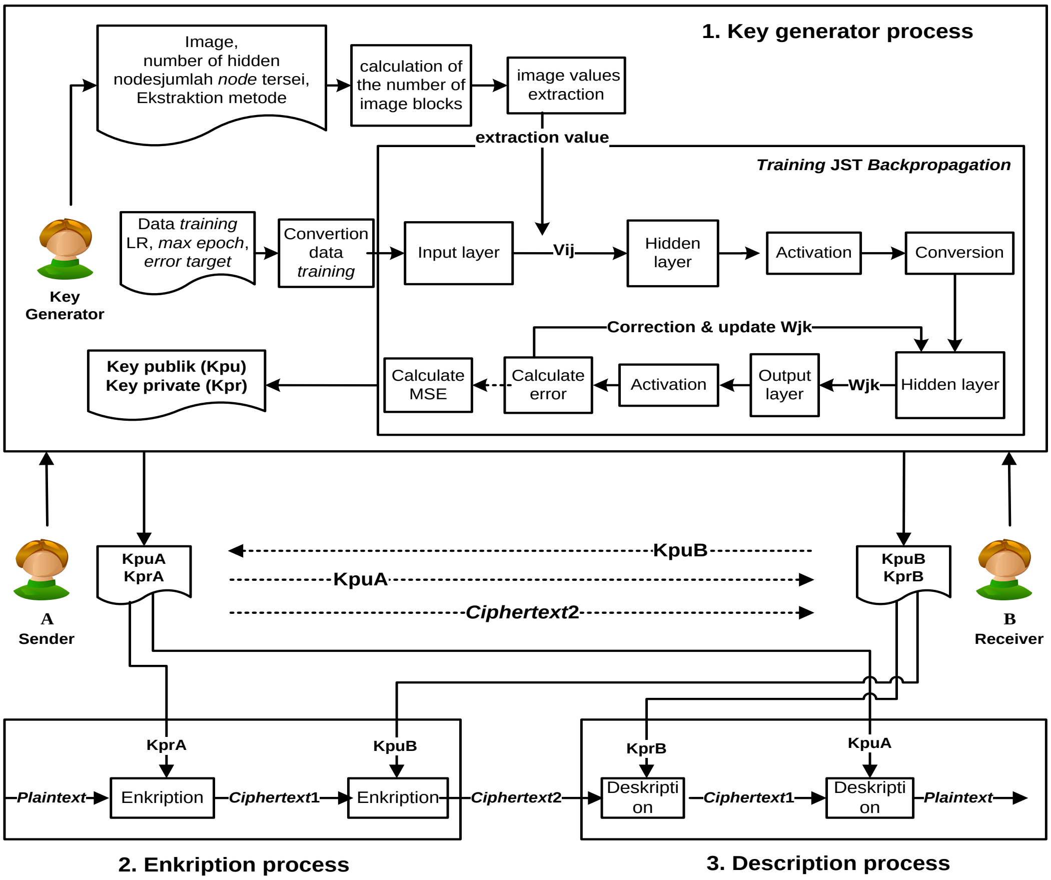

Figure 1. Key generator and process encryption description

4. Key Deployment Process

The key generation process consists of two parts, namely the process to get the value of the image and the training process on artificial neural networks (ANN):

4.1. Extract Values from the Image

The feature extraction process is a process in which to obtain numerical values which will be used as weights in the ANN training process. That acts as a weight v on the network that is between the input layer and the hidden layer. The extraction process of image values is carried out using the GLCM method and colour moments. This stage utilizes a colour image chosen by the party who wants to generate the key. Before extracting the image value, the number of image blocks will be calculated first. The image will be divided into sections, this is done to meet the required value length.

4.2. Training of JST backpropagation

After extracting the image value, the training process is then carried out on the ANN. The extraction value from the image will be the weight of V_ij which is in the input layer to the hidden layer. During the training process the weight of the W_jk between the hidden layer and the output layer is continuously updated with the conditions until it reaches the specified max epoch or error tolerance. The flow of the ANN training process.

4.3. Training Process Encryption and Description

The encryption process is the process of encoding a message so that the message cannot be read by another party, the sequence of the encryption process. The decryption process is carried out by the recipient of the message. The decryption process is the process of returning encrypted messages (ciphertext) in the form of original messages (plaintext).

4.4. Extraction Image

The feature extraction process is carried out to obtain the value that will be used as a weight at the backpropagation ANN training stage. Where this weight acts as Vij which is located between the input layer and the hidden layer. To perform the feature extraction process from the image, the GLCM texture feature extraction method and the colour feature extraction method are colour moments. The pictures used in the testing process are 3 pictures as shown in table 1.

Table 1. Picture used in the testing process

No.	Image	File Name	Ratio	Size
1		Bunga.jpg	1021 x 1029	73,3 KB
2		Kopi.jpg	1024 x 670	240 KB
3		Mobil.jpg	1024 x 768	214 KB

4.5. Key Development Process (training JST)

The key generation process is carried out by training using backpropagation ANN. The input needed in the training process is the training data, the number of hidden nodes, and the value generated from the image feature extraction process. The extraction value of the image will act as a weight (Vij) that is between the input layer and the hidden layer. The output of the training process is the weight (Vij) and the weight (Wjk) which will be used by the public and private key. Tests show that the system can generate a key. The formation of public and private keys is shown in Figure 2.

```
    13
    -8,24734362780311     32,8838988303268     22,8478992704122     -54,3168782721366
46,9403048631096  -69,46005615538  5,23638895836391  3,71358363953214
    -6,07735180473919  -54,6362339138164  -25,0876811509  ...  35,5767875088922
```

Figure 2. Public Key A (KpuA)

Tests carried out show the encryption process has been successfully carried out, the message can be encoded into an unreadable form. A then sends the ciphertext message to B. To find out the contents of the message, B must decrypt the message using his private key (KprB). Based on tests conducted, the system has been able to encrypt and decrypt messages by maintaining data confidentiality. Proses encryption and decryption for authentication.

For authentication purposes, a private key A (KprA) is required which is described in 3 for message encryption. For example, the plaintext used can be seen in Figure 3.

> *Authentication is a service related to identification. This function is applied to both entities with the information itself. The two communicating parties must introduce themselves to one another.*

Figure 3. Plaintext Data

A will send a message to B, so A must encrypt the message with his private key (KprA) which will produce a ciphertext, just call it ciphertext1. This encryption aims to provide authentication for messages sent. Ciphertext1 is shown in Figure 4.

```
eseohrzh`gzstyqes}webo}}xwsywxwx{w}}}zyyqes}webo}}xweOeOeZUYdzd^^|yz{}{}}z~~|}yqes}
webo}}xw||wz~~tu}~~}{y|w|}||x}~~}|}yzy~}y}y~~}}yqes}webo}}xw|z|w{}~}y}~}}eOeOeZUYdzd^^|y
w|~yt}~~~z{yqes}webo}}xw|z|w{}~}y}~}}||zx~~}uv~~~}eOeOeZUYdzd^^ ... |mu|}}x|x{|
```

Figure 4. Ciphertext1 Data

Based on the results of research conducted on public key generation using images and backpropagation neural networks, the following conclusions are obtained:

a) The system built using image feature extraction and backpropagation ANN with binary sigmoid activation function has successfully encoded the message and returned it to its original shape.
b) Successes encrypting and decrypting messages is affected by the accuracy achieved during the training process. The success of the training process is influenced by the number of hidden nodes and initial weight (Vij) used.
c) The training process carried out using weights (Vij) GLCM extraction results obtained 8 of 10 images reached 100% accuracy with the number of hidden nodes 15 to 30 nodes. Whereas the colour moment feature extraction results achieved a low accuracy of 29.47%.
d) A system built with a key length of ± 112 values takes 1.4322E + 189 years if brute force is keyed.
e) The system built has fulfilled the security aspects of confidentiality (confidentiality) and authentication

5. Conclusion

After doing a number of things regarding public key generation using images and backpropagation neural networks, the following conclusions are obtained:

1. Success in encrypting and decrypting messages is affected by the accuracy achieved during the training process. The success of the training process is influenced by the number of hidden nodes and initial weight (Vij) used.
2. The training process carried out using weights (Vij) GLCM extraction results obtained 8 of 10 images reached 100% accuracy with the number of hidden nodes 15 to 30 nodes. Whereas the colour moment feature extraction results achieved a low accuracy of 29.47%.
3. A system built with a key length of ± 112 values takes 1.4322E + 189 years if brute force is keyed.
4. The key that was built has fulfilled the security aspects of confidentiality (confidentiality) and authentication (authentication).

References

[1] Prayudi. (2015). Anti forensic techniques are being used extensively for cybercrime. Retrived from https://catatanforensikadigital.wordpress.com/category/steganography/.

[2] Lakespecialist, J. (2019). What is steganography and how does it differ from cryptography. Retrived from http://www.crime-research.org/articles/Stegano26/.

[3] Haryanto, E. 2016. Implementation Teknik Steganography For Anti Forensic Citra. University of Janabadra.

[4] Mualfah, D., & Riadi, I. (2017). Network Forensics For Detecting Flooding Attack On Web Server. IJCSIS) International Journal of Computer Science and Information Security, 15(2).

[5] Fatma, Y., Dkk. (2018). Implementasi Steganografi Pada Teks Terenkripsi dengan Algoritma RSA. Jurnal Fasilkom.

[6] Menezes, A. J., Orschot, P.C. dan Vanstone, S.A., 1996, Handbook of Applied Cryptography, Electrical Engineering and Computer Science, Massachusetts Institute of Technology More references.

[7] Sadikin, R., 2012, Kriptografi untuk Keamanan Jaringan, Penerbit ANDI, Yogyakarta.

[8] Supriyanto, A. (2017), File To Image Encryption (FTIE) Menggunakan Algoritma Randomized

Text Dan Arnold Cat Map (ACM) Untuk Keamanan Transmisi Data Digital. H@dfex.

[9] Gonzalez, R., C., Woods, R., E., 2002, Digital Image Processing, Second Edition. Prentice Hall International, New Jersey.

[10] Gadkari, D., 2004, Image Quality Analysis using GLCM, Thesis, Science in Modeling and Simulation, University of Cetral Florida, Florida.