IMPLEMENTATION OF DATA HIDING BY NEURAL NETWORK AND RETRIEVAL OR ENCRYPTED MULTIMEDIA DATA FILES

Komal Tahiliani
(Research Scholar)

Dr.N.K.Tiwari
(Director,(BIST,Bhopal))

Abstract: Hiding Messages in image data, which is generally known as Steganography is used for both illegal and legal scenarios. In such applications various file formats are used as cover-object which contains confidential data. As the new research field the techniques introduces Artificial Neural Network. It is an efficient method to solve complex problems. This newly developed technique uses Multilayer Perceptron algorithm of Neural Network for data security. Results are observed through different Media file as cover-objects. In the proposed method, MLP algorithm is implemented with traditional substitution method to obtain high embedding capacity with no visibility. The system provides confidentiality and integrity to the data during communication through open channels.

Key words: Data Hiding, Neural Network, MLP

1. Introduction

Today Internet[1] became an important part of our daily life, everybody wants to send and share information through internet. Therefore, everybody wants our secret data should be transferring in secure manner i.e. no manipulations, no hacking and modifications and its integrity should be maintained. Data security is the major issue for secure data transfer. There are several techniques used in this respect some of the well known techniques are:

1. Cryptography: Cryptography changes the contents of the file into unreadable form. Sender and receiver have the key to unlock the contents.
2. Watermarking: All objects are marked in the same way. Shows ownership of digital data.
3. Steganography or Data Hiding: Hides the existence of secret message into cover-object.
4. Digital Signature: Allows authorship of a document to be asserted[2].

Steganography and Cryptography are both used to ensure data confidentiality, security and data integrity. In cryptography message is encrypted into unreadable form, however the main difference is that within encryption anybody can see that both parties are communicating in secret, whereas Steganography hides the existence of a secret message in such a way that nobody can see that both parties are communicating in secret.

Encryption allows secure communication requiring a key to read the information. An attacker cannot remove the encryption but it is relatively easy to modify the file, making it unreadable for the intended recipient. Digital signatures allow authorship of a document to be asserted. The signature can be removed easily but any changes made will invalidate the signature, therefore integrity is maintained.Data hiding is a method of hiding secret messages into a cover-media such that an unintended observer will not be aware of the existence of the hidden messages. In this paper, 8-bit grayscale images are selected as the cover media. These images are called cover-images. Cover-images with the secret messages embedded in them are called stego-images. For data hiding methods, the image quality refers to the quality of the stego-images. the paper is mainly designed for providing security for the data. In this, the sender encrypts the data to some form. While encrypting the data in to some form, the key file is entered by the sender.

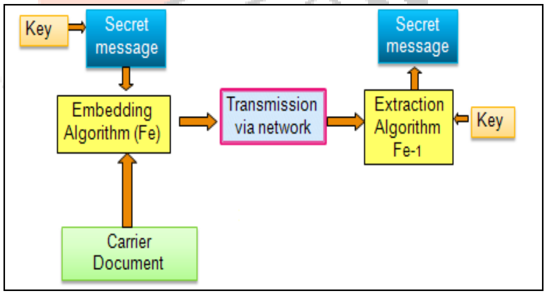


Fig.1. Data Hiding Process

The purpose of the key file[13,14] is to provide security to the system as it is known only to the sender and the receiver. Since the actual processing of the data takes place on the remote client the data has to be transported over the network, which requires a secured format of the transfer method. Present day transactions are considered to be "un-trusted" in terms of security, i.e., they are relatively easy to be hacked. And also we have to consider the transfer the large amount of data through the network will give errors while transferring. Nevertheless, sensitive data transfer is to be carried out even if there is lack of an alternative. Network security in the existing system is the motivation factor for a new system with higher-level security standards for the information exchange. This paper proposes the following features. It provides flexibility to the user to secure the data very easily. In this system the data is also hided inside the JPEG Image and Video file. The user who received the file will do the operations like de-embedding, and decryption in their level of hierarchy.

2. Data Hiding Various Streams

2.1 Data hiding in still images

Data hiding [11]in still images presents a variety of challenges that arise due to the way the human visual system (HVS) works and the typical modifications that images undergo. Additionally, still images provide a relatively small host signal in which to hide data. A fairly typical 8-bit picture of 200 □ □ 200 pixels provides approximately 40 kilobytes (kB) of data space in which to work. This is equivalent to only around 5 seconds of telephone-quality audio or less than a single frame of NTSC television. Also, it is reasonable to expect that still images will be subject to operations ranging from simple affine transforms to nonlinear transforms such as cropping, blurring, filtering, and lossy compression [3].

2.2 Data hiding in audio

Data hiding [4] in audio signals is especially challenging, because the human auditory system (HAS) operates over a wide dynamic range. The HAS perceives over a range of power greater than one billion to one and a range of frequencies greater than one thousand to one, Sensitivity to additive random noise is also acute[12].

2.3 Data hiding in text

Soft-copy text [5,6] is in many ways the most difficult place to hide data. (Hard-copy text can be treated as a highly structured image and is readily amenable to a variety of techniques such as slight variations in letter forms, kerning, baseline, etc.). Data hiding in text is an exercise in the discovery of modifications that are not noticed by readers.

2.4 Data Hiding in Video

A video data embedding scheme [7,8] in which the embedded signature data is reconstructed without knowing the original host video. The proposed method enables high rate of data embedding and is robust to motion compensated coding, such as MPEG-2. Embedding is based on texture masking and utilizes a multi-dimensional lattice structure for encoding signature information. Signature data is embedded in individual .

3. Related Work

3.1 Traditional Data Hiding Methods:

Data hiding techniques [9][10] are classified according to mechanism and changes takes place during embedding as well as the media used for hiding.

a) Techniques according to the mechanism used are:

- Substitution system.
- Transform Domain Technique.
- Spread Spectrum Technique.
- Statistical Methods.
- Distortion Techniques.
- Cover Generation Technique

b) Data hiding are classified according to the file formats used in several Medias like text, image, audio and video.

3.2 Current state of art

A lot of research works have been carried out in the literature for data hiding and some of them have motivated us to take up this research. Brief reviews of some of those recent significant researches are presented below:

In the literature, many techniques for data hiding have been proposed [2-5]. One of the common techniques is based on manipulation the least significant bit (LSB) plans. A LSB substitution method replaces some LSB of the cover-image with the secret data [1,6-8,13,14].

Wang et al. [6] proposed to embed secret messages in the moderately significant bit of the cover-image. A genetic algorithm is developed to find an optimal substitution matrix for the embedding of the secret messages. They also proposed the use of local pixel adjustment process (LPAP) to improve the image quality of the stego-image. Recently, Wang et al. [8] proposed a novel method to embed data inside the host image. The method based on simple LSB substitution data hiding. They also developed the optimal *k* LSB substitution method to solve the problem when *k* is large.

Chang et al [7] proposed a method of finding the optimal LSB in image hiding by dynamic programming strategy. The proposed method finds the optimal LSB substitution that Wang [8] found of approximate OLSB, the method reduces the computation time too Safy et al., [4] proposed an adaptive steganographic technique in which the bits of the payload are hidden in the integer wavelet coefficients of the cover image adaptively along with optimum pixel adjustment algorithm. Hassan Mathkour et al., [5] compared the strengths and weaknesses of the existing techniques of steganography and implemented a new steganographic wizard based tool, which have been examined against several other tools like F5, S-Tools etc., for more robust and secure steganography. Vijayalakshmi et al.,[6] proposed modulo based image steganography algorithm. The method combines samples of LSB bits using addition modulo to get the value which is compared to the part of the payload. If these two values are equal, no change is made in sample otherwise add the difference of these two values to the sample.

Wein Hong et al., [7] proposed a lossless steganography technique wherein the secret information is hidden inside the compressed Absolute Moment Block Truncation Coding image. Naji et al., [8] analysed different steganographic techniques and weaknesses in the respective techniques.

Hassan Shirali-Shahreza and Mohammad Shirali- Shahreza [9] proposed a synonym text steganographic technique in which the words in American English are substituted by the words having different terms in British English and vice-versa.

4 Proposed Method

4.1 Neural Networks:

Proposed method uses a multi-layered perceptron model[15,16] for neural network model. Multi-layered perceptron basically has a synaptic link structure with neurons between the layers but no synaptic link among the layer itself. Signals given to the input layer will propagate forwardly according to the synaptic weight of the neurons connected from the input values and reaches to the output layer as shown in Figure1.:

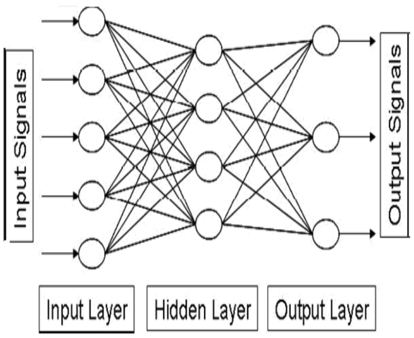


Fig.1 Neural Networks Structure

Normally sigmoid function[17] is used for this model and is expressed as follows

$$f(x) = \frac{1}{1+e^{-x}}$$

Each synaptic link has a network weight. The network weight from unit *i* to unit *j* is expressed as *wij* and the output value for unit *i* is expressed as *Oi*. The output values for the unit is determined by the network weight and the input signal. Consequently, to change the output value to a desired value, adjustment of these network weights are needed.

In proposed method, we use back propagation learning as learning method. Back propagation learning is a supervised learning. This method tries to lower the difference between the teacher signal and the output signal by changing the network weight. Changes of the network weight according to the difference in the upper layer propagate backward to the lower layer. This difference between the teacher signal values are called as error and often expressed as δ . When teacher signal *tk* is given to the unit *k*
of output layer, the error *δk* will be calculated by following function:

$$\delta_k = (t_k - O_k) \cdot f'(O_k)$$

To calculate the error value *δj* for hidden unit, error value *δk* of the output unit is used. The function to calculate the error value *δj* for hidden unit *j* is as follows:

$$\delta_j = (\sum_k \delta_k w_{jk}) \cdot f'(O_j)$$

After calculating the error values for all units in all layers, then network can change its network weight. The network weight is changed by using following function:

$$\Delta w_{ij} = \eta \delta_j O_i$$

η in this function is called learning rate. Learning rate is a constant which normally has a value between 0 and 1 and generally represents the speed of learning process.

5. Experimental Evaluation

5.1 Performance Evaluation Parameter

Peak Signal to Noise Ratio (PSNR)[12] is used to evaluate the quality of the stego after embedding the secret message in the cover image. PSNR measures the quality of the image by comparing the original image or cover image with the stego-image. It is expressed in terms of logarithmic decibel scale(dB).Higher the value of PSNR the more quality the stego have. The acceptance range of PSNR is ,typical values for the PSNR in lossy image and video compression are between 30 and 50 dB, where higher is better, but acceptable values for wireless transmission quality loss are considered be about 20 dB to 25 dB.It is most easily defined via the mean squared error (*MSE*). A cover image C (i, j) that contains N by N pixels and a stego image S(i, j) where S is generated by embedding /mapping the message bit stream , *MSE* is defined as:

The PSNR is defined as:

$$\text{PSNR} = 10 \log_{10} 255^2 / \text{MSE}$$

5.2 Implementation

The proposed application is developed in java programming and applied on different file format of different size and observed possibilities of data hiding and retrieving as follows:

Cover Object	**File format**	**Hide**	**Retrieve**
Text	docx	√	√
Presentation	ppt	√	√
Html	htm	√	√
Image	jpg	√	√
Image	bmp	√	√
Image	gif	√	√
Audio	mp3	√	√
Audio	wav	√	√
Audio	wma	√	√
Video	mpeg	√	√
Video	avi	√	√
Video	wmv	√	√

Table1. Different File Formats as Cover

5.3Experimental simulation

To calculate the simulation results, the four groups’ of Multimedia containers are prepared. Each group of suspect data with varying file format is combined to form a common set of Text, Audio, Video and Image that means the embedded information was a randomly multimedia data. To perform MLP Algorithm in the implementation a Self set was defined as a collection of 100 clean, 512x512-pixel JPEG, 512 x512 BMP images, 5 MB Audio File, 50 MB Video File, 1 MB Text File.

5.3. 1Data Hiding in Image

In our simulation we had tried to hide text contained inside different size and format of the image file like BMP, JPEG and GIF. PSNR of the file are calculated which is shown by table 2.

Fig.7.Original image Fig. 8.Stego image

Observed PSNRs for the respective image file formats are given by the table.2

S No.	Image File Formats	Size (in KB)	PSNR (in dB)
1	JPG	6.98	44.3
2	GIF	144	41.2
3	BMP	768	37.6

Table 2: Image File Formats and its PSNR

5.3.2 Data Hiding in Text

The following fig.9 HTML text frame can be used as container for embedding information where as table .3 represents the PSNR for the text formats.

```
<html>
<body>
………
<p>Image  Example:</p>
                                   <a
href="http://www.quackit.com/travel/new_zealan
                       d/milford_sound">
<img
src="/pix/milford_sound/milford_sound_t.jpg"
style="max-width:100%" alt="Milford Sound in
New Zealand" /></a>
```

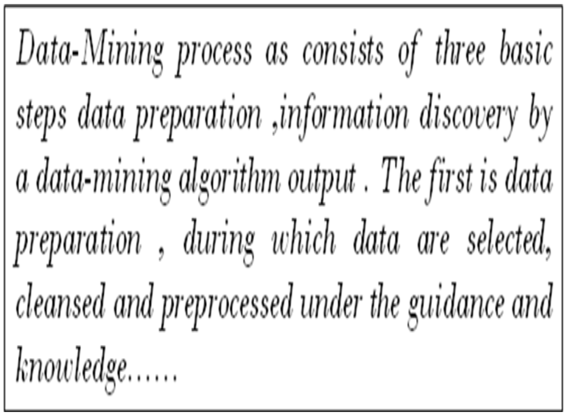
S No.	Text File Format	Size (in KB)	PSNR (in dB)
1	DOC	27.1	25.4
2	PPT	578	29.2
3	HTML	20	34.3

Table 3: Text File Formats and its PSNR

5.3.3Data Hiding in Audio and Video

Artificial Neural Network's data hiding method is used to embed information in different audio (table.5) and video file (table.4) and we have calculated their respective PSNR which indicates at the better stego quality image.

Fig.9. Video Frame



Data-Mining process as consists of three basic steps data preparation ,information discovery by a data-mining algorithm output . The first is data preparation , during which data are selected, cleansed and preprocessed under the guidance and knowledge……

Fig. 10.Message to be embedded into Video Frame

S No.	Video File Formats	Size (in MB)	PSNR (in dB)
1	MPEG	3.2	31.3
2	AVI	1.6	38.7
3	WMV	2.8	41.8

Table 4.Video File Formats and its PSNR

S No.	Audio File Format	Size (in MB)	PSNR (in dB)
1	MP3	2.58	45.3
2	WAV	2.03	41.2
3	WMA	2.5	47.4

Table 5.Audio File Formats and its PSNR

6. Conclusion

The security of data is of extreme importance in today's information-based society, including the fields of military, diplomacy, corporation, medicine, and even the individual, the information have to be safeguarded to avoid the unauthorized or illegal accesses and prevent the misuses and abuses. Any system or technique that deals with, processing information (data), wants to put this data in shapes or forms of media under the condition that it must be not visible in its new form for human observer. All such systems are called hiding systems for information. In this paper, we had technique as possible methods for embedding data in host text, image, video and audio signals .While we have had some degree of success Proposed method uses MLP Algorithm for classifying the input patterns to corresponding hidden signals.

7. References

[1] Arup Kumar Bhaumik, Minkyu Choi,Rosslin J.Robles and Maricel O.Balitanas. “Data Hiding in video” International journal of Database Theory and Application vol.2, No.2 June 2009.pp 9-14

[2] Jacob T. Jackson, Gregg H. Gunsch, Roger L. Claypoole, and Gary B.Lamont.:Novel Steganography Detection Using an Artificial Immune System Approach” IEEE 2010

[3] Poulami Dutta, Debnath Bhattacharyya and Tai-hoon Kim.”Data hiding in Audio Signal”.International Journal of Database theory and Application vol.2,No.2 June2009 pp.1-8

[4] Shan Wang,Bian Yang and Xiamu Niu .”A secure Steganography Method based on Genetic Algorithm” Journal of Information Hiding and Multimedia Signal Processing ,’Ubiquitous International’,ISSN2073-4212 ,Volume 1 Number 1 ,2010, pp:323-322.

[5] Encryption, Data Hiding, and Hostile Code. In Write Work.com. Retrieved 02:12, October 24, 2012

[6] Jammi Ashok, y.Raju k.Srinivas “Steganography: an overview” International Journal of Engineering Science and Technology Vol. 2(10), 2010pp: 5985-5992

[7] W.Bender,D.Gruhl,N.Morimoto,A.lu. “Techniques for data hiding” IBM System Journal, Vol 35,No. 3&4,1996 pp:313-336

[8] M.S Shashidhara,A.Pandurangan,Suhasini CHV “Data hiding and retrival of Encrypted file in Images and videos using ANN methods Int.J.Advanced Networking and Applications Vol.02,Issu:02,2010, pp:544-549

[9] Abbas Cheddad,Joan Condell,Kenvin Curran,Paul Mc Kevitt.”Digital image stegnogrphy:survey and analysis of current methods”.SIgnal processI 90 2010 pp: 727-752

[10] Jayaram P.,Ranganatha H.R,Anuppama H S .”Information Hiding using Audio Steganography”.The international Journal of Multimedia & Its Application (IJMA) Vol.3, No.3 August 2011,pp-86-96,

[11] I,Aveibas,N.Memon,B.Sankur.”Steganalysis based on image quality metrics”. Multimedia Signal Processing, 2001. IEEE IV workshop 2001

[12] El-sayed M.El-Alfy,Azzat A.Al-sadi “Pixel-Valued Differencing Steganogrphy:Attacks and Improvements” ICCIT,2012, pp:757-765

[13] Rosziati Ibrahim , Teoh suk kuan “ Steganography Algorithm to hide Secret message inside an image” Computer Technology and Application 2 ,2011, pp:102-108

[14] Imran Khan “An Efficient Neural Network based Algorithm of Steganography for image” International journal of Computer Technology and Electronics Engineering(IJCTEE) Vol.1,Issue 2, ISSN 2249-6343, 2011, pp:63 -67

[15] K.B Shiva Kumar K B Raja “Bit Length Replacement Steganography Based on DCT coefficients” International Journal of Engineering Science and Technology vol.2(8), 2010, pp:3561-3570

[16] Souvik Bhattacharya, Lalan Kumar ang gautam Sanyal “ A novel approach of data hiding using Pixel Mappng Methods” International journal of computer science and information security, Vol. 8, no. 4 july 2010 pp:207-214