Steganography Using Genetic Algorithm

Prachi
M.Tech(CSE), Shri Baba Mast Nath Engineering College , Rohtak, India.

Sunita
Department of Computer Science and Engineering, Shri Baba Mast Nath Engineering College, Rohtak, India.

Rajiv Sharma
Department of Computer Science and Engineering, Shri Baba Mast Nath Engineering College, Rohtak, India.

Abstract – Steganography provides security by hiding the cipher text into a seemingly invisible image or other formats. According to Johnson et al., (2001), "Steganography is the art of hiding and transmitting data through apparently innocuous carriers to conceal the existence of data". The level of visibility is decreased using many hiding techniques in „Image Modeling" like LSB „Manipulation", „Masking and filtering". These techniques are performed by different steganographic algorithms like F5, LSB, JSteg etc. and the act of detecting the information hidden through these algorithms is called "Steganalysis". "Cryptography" is the art of science used to achieve security by encoding the data to transform them into non-readable formats so that unauthorized users cannot gain access to it.

Index Terms – Steganography, Cipher, LSB, Cryptography.

1. INTRODUCTION

In the field of Data Communication, there is a fear that the data may get snooped at the time of sending it from sender to receiver. And, this is the reason that leads to the need of Information Security to carry out a secure communication over the network. The fast evolving Internet and the digital information revolution have made it feasible for consumers all over the world to exchange multimedia files. So, more robust methods are required to ensure a secure transfer and data security is now an inseparable part of Data Communication. And, Cryptography can help achieving this information security.

The advantage of Steganography, over Cryptography alone, is that messages do not attract attention to themselves. Cryptography takes a file and transforms it into a new file. Thus, the plain text is transformed to the cipher text through encryption algorithm and encryption key. And, on the receiver side, the cipher text is transformed back into the plain text using the decryption algorithm and the decryption key.

On the other hand, Cryptography hides a file into another file before sending it. Usually this is done by making changes to the bytes of the file in such a way that does not obviously change the file that much that the changes are visible.

Thus, where the objective of Cryptography is to make the message unreadable to anyone who does not have the key and the exact algorithm. The objective of Cryptography is to hide the existence of the message.

Since this research focuses on two of most important requirements of Cryptography which are payload and imperceptibility, the following objectives are attempted to be achieved in this study

1. To develop an efficient model in audio Cryptography following the genetic based substitution approach.
2. To implement the genetic based substitution technique of audio Cryptography techniques while maintaining imperceptibility.
3. To improve the imperceptibility of substitution techniques of audio Cryptography techniques preserving the payload.
4. To achieve basic requirements such as high capacity for hidden messages, high security & good invisibility.
5. To check the robustness of our method using PSNR & MSE.

2. RELATED WORK

The majority of today's steganographic systems uses multimedia objects like image, audio, video etc as cover media because people often transmit digital pictures over email and other Internet communication.[1] In modern approach, depending on the nature of cover object, Steganography can be divided into five types:

- Text Steganography
- Image Steganography
- Audio Steganography
- Video Steganography
- Protocol Steganography

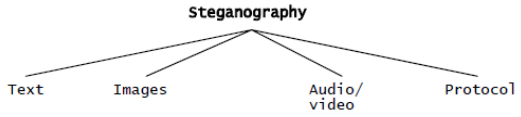


Figure.2.1 Categories of Steganography

Secrete data in video create "stego" video file which send to the receiver side. Proposed system introduces a novel and more secure method of video Cryptography.

Text Steganography involves hiding information in plain text. Many techniques involves the modification of the layout of a text, rules like using every n-th character.

Image Steganography involves hiding information in image. To hide information, straight message insertion may encode every bit of information in the image.

In a computer-based Audio Steganography system , secret messages are embedded in digital sound.

Video files are generally a collection of images and sounds, so most of the presented techniques on images and audio can be applied to video files too.

The term protocol Steganography refers to the technique of embedding information within messages and network control protocols used in network transmission.

FUNDAMENTAL PROPERTIES

Steganographic algorithms can be characterized by a number of defining properties. Three of them, which are most important for audio steganographic algorithms are:

- Transparency,
- Robustness and
- Capacity.

And these can be explained as:

1. Transparency

Transparency evaluates the audible distortion due to signal modifications like the case of embedding the message or the case of message attacking. As in, most of the applications require the additional data to be added by the Cryptography algorithm affecting the perceptual quality of the audio host signal.[9] The fidelity of the Cryptography algorithm is usually defined as a perceptual similarity between the original and stego audio sequence. However, the quality of the stego audio is usually degraded, either intentionally by an adversary or unintentionally in the transmission process, before a person perceives it.

2. Robustness

Robustness measures the ability of embedded data or watermark to withstand against intentional and unintentional attacks.[9] Unintentional attacks generally include common data manipulations such as lossy compression, digital-to-analog conversion, re-sampling, re-quantization, etc. whereas intentional attacks cover a broad range of media degradations which include addition white and colored noise, rescaling, rotation (for image and video Cryptography schemes), resizing, cropping, random chopping, and filtering attacks.

Also, the robustness of the algorithm is defined as an ability of the data detector to extract the embedded message, the secret information after common signal processing manipulations.

3. Capacity

Capacity of an information hiding scheme can be defined as the amount of information that a data hiding scheme can successfully embed without introducing perceptual distortion in the marked media.

In the case of audio, it evaluates the amount of possible embedding of the secret message into the audio signal. The embedding capacity is the all included embedding capacity (not the payload) and can be measured in percent (%), bits per second or frame and bits per megabyte or kilo byte audio signal.

3. PORPOSED MODELLING

Genetic Algorithms are search algorithms that are based on concepts of natural selection and natural genetics. The biological background i.e. the basic genetics can be explained as:

- Each organism has a set of rules that define how the organism is built. And, all living organisms are composed of cells.
- In each set, there is composition of chromosomes. Chromosomes define the whole structure of organism.
- A chromosome consists of genes,blocks of DNA.
- Each gene encodes a particular protine that represents a trait (feature), i.e. color of eyes.
- Possible settings for a trait (eg. Blue,brown) are called alleles.
- Complete set of genetic material(all chromosome) is called a genome.
- Particular set of genes in a genome is called genotype.
- The physical expression of the genotype(the original organism itself after birth) is called phenotype ,its physical and mental characteristics like eyes color, intelligence etc.
- When the two organisms mate, they share their genes. The resultant off springs may end up having half the

genes from one parent and half from the other. This process is called Recombination (crossover).

- The new created off springs can then be mutated. Mutation means, that the elements of DNA are a bit changed. This change is mainly caused by errors in copying genes from parents.
- The fitness of an organism is measured by the success of organism in its life (survival).

Basic Terminology:

- Chromosome- a set of genes, a chromosome contains the solution in form of genes.
- Gene- a part of chromosome, a gene contains a part of the solution.
- Individual- same as chromosome.
- Allele- the value of individual variables.

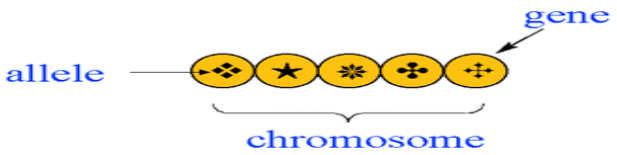


Figure 3.1. Chromosome

- Population- number of individuals
- Fitness- the value assigned to an individual, based on how far or close an individual is from the solution.
- Fitness Function- a function that assigns fitness values to the individuals. It is problem specific.
- Breeding- taking two fit individuals and then intermingling their chromosome to create two new individuals.
- Mutation- changing a random gene in the individual.
- Selection- selecting individuals for creating the next generation.

OUTLINE OF GENETIC APPROACH

1. [START] Generate random population of n chromosomes.

2. [FITNESS] Evaluvate fitness f(x) of each chromosome x in the population.

3. [NEW POPULATION] Create a new population by repeating following steps until the population is complete:

(a) [SELECTION] Select two parent chromosomes from the population according to their fitness.

(b) [CROSSOVER] With a crossover probability, crossover the parents to form new offsprings(children).

(c) [MUTATION] With mutation probability, mutate new offspring at locus (position in chromosome).

(d) Place new offspring in the population.

4. [REPLACE] Use new population for further run of the program.

5. [TEST] If the end condition is satisfied, stop.

6. [LOOP] Go to step 2

New algorithms keep emerging prompted by the performance of their ancestors,by the rapid development of information technology and by the need for an enhanced security system. The discovery of the LSB embedding mechanism is actually a big achievement. Although it is perfect in not deceiving the HAS, its weak resistance to attacks left researchers wondering where to apply it next until they successfully applied it with the GENETIC APPROACH.

4. RESULTS AND DISCUSSIONS

1. Object/Project Snapshot

At first, the cover audio file format that is chosen is WAVE audio file format because this format is original of all the formats.

Figure 1 Wave Audio

Now,initially the cover audio file is played.

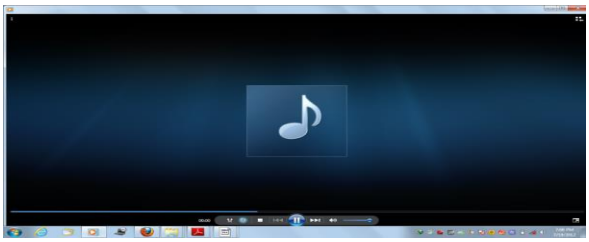


Figure 2 Input played

And the signal is analyzed in .NET using the function Waveread and the plot is obtained as:

Now, after embedding the secret message the Stego Audio is obtained in output and again the plot for the Audio, this time the Stego Audio is obtained. And, the output i.e. the Stego Audio is played.

2 Evaluation

PSNR Values for the case when the secret message, to embed is same and the size of cover audio is varying.

Wave Audio	Wave Size (Audio Size In Bytes)	Message Size (Image Size In Bytes)	PSNR
Bird Wave	315,392 bytes	12,288 bytes	86.358
Drum & Bass Wave	471,040 bytes	12,288 bytes	77.099
Express Wave	307,200 bytes	12,288 bytes	74.545
Funky Wave	208,896 bytes	12,288 bytes	85.514
Philtered Wave	212,992 bytes	12,288 bytes	85.985

Table 1 Message(Image) is same & the Wave Files size differ

And, the graph generated after observing the values given in table 1 is shown as:

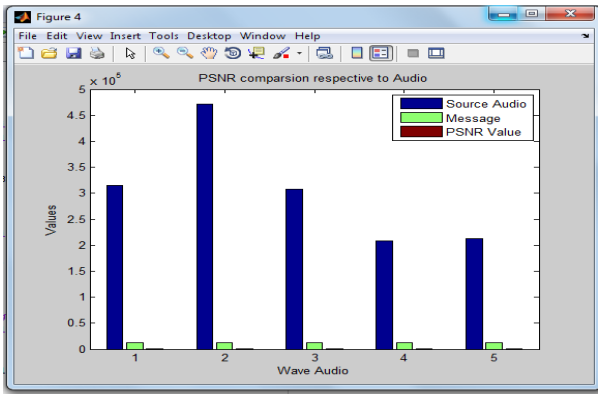
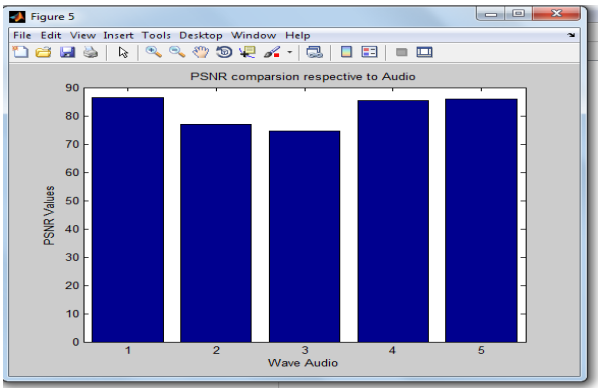


Figure 3 Comparision respective to different Audio covers where Source File, Message & PSNR measure is compared.



PSNR Values for the case when the secret message, to embed is of different size and the size of cover audio is same.

5. CONCLUSION

This paper is a short form of world Steganography. We have shown how the simplest methods work and how they can be explored. We have used symmetric encryption algo to provide more security. Research in this field has already begun. Next to Steganography, one of the most active fields of research is mass detection tools for hidden contents. The problems are really big. At first, known statistical tests are fragile and for many embedding schemes we still do not know which properties to test. At second, the today traffic in public networks is so overwhelming, that is too hard to rigorously Practically, asymmetric algorithms like RSA are used for the key exchange and symmetric algorithms are used for encryption / decryption. Further, general implementation limitations of cryptographic algorithm emphasis the selection between hardware and software cryptosystem, choosing among symmetric and Asymmetric key algorithm and the essential factors to be followed to have a secure key management.

REFERENCES

[1] A Tutorial Review on Steganography-Samir K Bandyopadhyay, Debnath Bhattacharyya1, Debashis Ganguly1, Swarnendu Mukherjee1 and Poulami Das1 http://www.jiit.ac.in/jiit/ic3/IC3_2008/IC3-2008/APP2_21.pdf

[2] SANS Institute InfoSec Reading Room Cryptography: A right way Steganography_1584.pdf

[3] Cryptography& Steganalysis:Different Approaches-Soumyendu Das,Subhendu Das,Bijoy Bandyopadhyay,Sugata Sanyal http://arxiv.org/ftp/arxiv/papers/1111/1111.3758.pdf

[4] A review of audio based Cryptographyand digital watermarking-M. L. Mat Kiah1, B. B. Zaidan2,3,4, A. A. Zaidan2,3,4*, A. Mohammed Ahmed1 and Sameer Hasan Al-bakri1 http://www.academicjournals.org/IJPS/abstracts/abstracts/abstract2011/18Aug/Kiah%20et%20al.htm

[5] A Detailed look of Audio Cryptography Techniques using LSB and Genetic Algorithm Approach Gunjan Nehru, Puja Dhar IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 2, January 2012 ISSN (Online): 1694-0814 www.IJCSI.org

[6] Algorithm For Audio Watermarking & Steganography-Nedeljko,Cvejic http://herkules.oulu.fi/isbn9514273842/isbn9514273842.pdf

[7] Data Hiding in Audio Signal: A Review-Poulami Dutta1, Debnath Bhattacharyya1, and Tai-hoon Kim2 http://www.sersc.org/journals/IJDTA/vol2_no2/1.pdf

[8] Efficient Method Of Audio Cryptographyby Modified LSB Algorithm & Strong Encryption Key With Enhanced Security-R Sridevi, DR. A Damodram, DR. SVL.Narasimham http://www.jatit.org/volumes/research-papers/Vol5No6/15Vol5No6.pdf

[9] A Genetic-Algorithm-Based Approach for Audio Steganography-Mazdak Zamani 1, Azizah A. Manaf 2, Rabiah B. Ahmad 3, Akram M. Zeki 4, and Shahidan Abdullah5 http://www.waset.org/journals/waset/v54/v54-63.pdf

[10] SANS Institute-InfoSec Reading Room Steganography: The Right Way http://www.sans.org/reading_room/whitepapers/stenganography.pdf