An Approach for Data Quality Improvement in Audio Steganography

Parvathy.T
Student
Department of Computer Science
Pondicherry University
Parvathyt18@gmail.com

K.Vaitheki
Assistant Professor
Department of Computer Science
Pondicherry University
Karaikal

ABSTRACT

Covert communication by embedding a message or data file in a cover medium has been increasingly gaining importance in the all-encompassing field of information technology. Audio steganography is concerned with embedding information in a bland cover speech in a secure and robust manner. Communication, transmission security and robustness are essential for transmitting vital information to intended sources while denying access to unauthorized persons. By hiding the information using a cover or host audio, the information even exist is concealed during transmission.

The proposed work provides high bit rate improved LSB audio watermarking method. The improved quality of watermarked audio is found to be better on implementation of the proposed algorithm than in the standard LSB method.

Keywords

Audio Steganography, LSB, RSA, HAS.

1. INTRODUCTION

Steganography is the art and science of hiding secret messages .It establishes covert communication. Only the intended recipient knows the existence of the secret information. For this technique it has to satisfy two basic requirements. The first requirement is cover object which contains its own data that data may be audio, video, text or an image. The second is stego object which contains the secret data. The proposed work uses A Human Auditory System (HAS) to hide the data. Also concealing the data in audio is more difficult compared to other media. In steganography techniques audio provides good performance and oppurtunity to hide the information, because it provides good and high embedding capacity. These audio files are larger in size than the image file size.

2. LEAST SIGNIFICANT BIT METHOD

In this standard LSB approach, a binary stream of audio is replaced by binary pattern of secret message. This method uses water mark encoder for selecting the host audio subset samples which is choosen by a secret key. When extracting the concealed message, the decoder simply retrieves the binary values of samples from audio stego object. Letters with ASCII values and corresponding binary values:

Table 1.Binary values and related ASCII values for letters

Letter	ASCII value	Binary Value
A	065	01000001
D	100	01100100
G	071	01000111
I	105	01101001
U	117	01110101

From this table we can conclude that to embed the secret message into audio, and the corresponding eight bit binary values have to be embedded into the digitized audio file where each sample is represented with 16 bits.

Example

A letter ‘A’ is first converted into ASCII value then the equivalent binary value is going to embed in the audio file. The sample of audio is given below:

1001010**1** 0000110**1** 1100100**1** 1001011**0**

0000111**1** 1100101**1** 1001111**1** 0001000**0**

ASCII value of ‘A’ is ‘065’ and binary equivalent is is 01000001.These eight bits can be written to the LSB of each of the eight carrier byte as follows:

1001010**0** 0000110**1** 1100100**0** 1001011**0**

0000111**0** 1100101**0** 1001111**0** 0001000**1**

The main drawback of the approach specified above, hackers can easily guess the secret message before reaching the recipient.

3. TRANSMISSION PROCESS

The working flow is shown in the following figure:

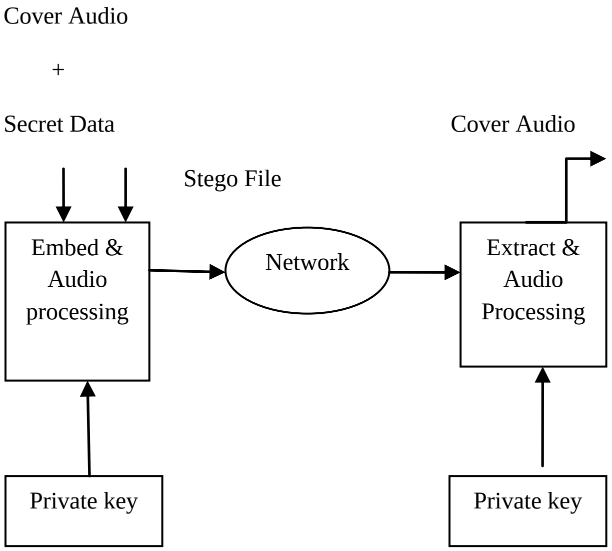


Fig 1: Transmission process

4. ALGORITHM FOR IMPROVING DATA QUALITY

The proposed algorithm facilitates some levels of protection. They are,

1. The first level of protection is choosing an embedding algorithm.

2. The second level of protection is choosing the key.

The high bit rate improvised LSB algorithm embeds the data that establishes a secure transmission of information and the second level of protection is provided by the private key that is used for both embedding and extracting the secret data. In order to increase the strength and confidentiality of transmission of data, both the sender and receiver should know the key .The data is initially encrypted by using asymmetric (RSA) Algorithm that facilitates a reliable communication , thereby reduces host audio distortion.

Secret message will be converted into ASCII value and after that it will be converted into binary pattern. Audio file also converted into binary pattern before concealing the message. There four methods to hide data in any audio file formats.

- Encoding
- Decoding
- Encryption
- Decryption

4.1 Encoding

It is a process of hiding the message in the audio.

4.2 Decoding

It is a process of retrieving the message from the audio.

4.3 Encryption

The user allowed entering the public key/shared key in any combination of numbers, symbols and characters. The key contains set of characters. All characters are converted to ASCII value and add all the ASCII value to get single number. And that single number is converted to bit pattern and by simple logical operation (XOR) you can get a single number less than 128.It is a new private key .It is added to the characters one by one in the message, before encoding.

4.4 Decryption

The LSB bits of consecutive eight 254 or 255 bytes are taken and subtracted with the key to get the original character.

Calculate the intervals for hiding message into the binary pattern of audio file such as below:

1. The audio file contains set of bytes. For example, take an audio file which play for 10secs It has more than 60,000 bytes. Each byte is received and checked if the received byte is 254 or 255.If it is byte 255 or 254, then encoding is done.

2. Inserting binary pattern of message into that intervals such as quality and audioable charecteristics could not alter.

3. For one character to encode we need eight 254 or 255 bytes. One character is hidden in consecutive eight 254 or 255 bytes.

4. Identifying the message pattern for last packet such as mark the end of message, the LSB bit of next eight consecutive 254 or 255 bytes which comes after all the message have encoded are replaced by 1.

5. The encoded file is decoded to get the message.

The main idea of high bit rate improvised LSB algorithm is watermark bit embedding which causes the minimum embedding distortion of host audio files. Also gives low computational complexity.

5. ADVANTAGES OF PROPOSED APPROACH

- Avoiding communication in well known forms of great reduces the risk of information being leaked in transit.
- When embedding the message within medium which does not alter the medium quality. The selection of audio file gives high capacity (as much as possible to hide the data).
- Consumption of time to encode and decode is reduced. Provision of sending the file to the destination is given so that after encoding the user can send the file by giving destination IP address.
- This proposed method provides greater security and it is an efficient method for hiding the secret information from hackers and sent to the destination in a safe and undetectable manner.
- This proposed system also ensures that the size of the file is not changed even after encoding.

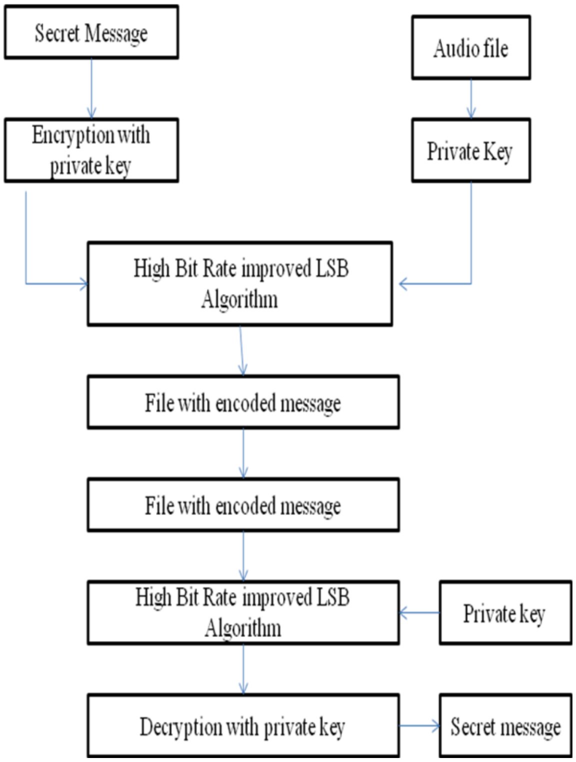
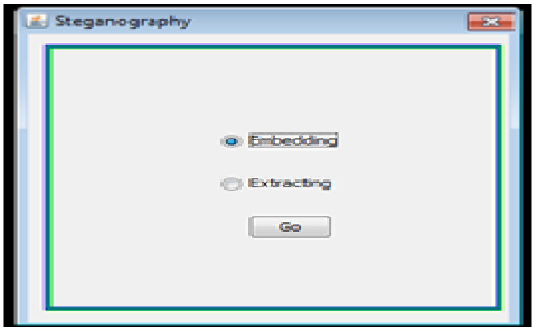

Secret Message
Audio file
Encryption with private key
Private Key
High Bit Rate improved LSB Algorithm
File with encoded message
File with encoded message
High Bit Rate improved LSB Algorithm
Private key
Decryption with private key
Secret message

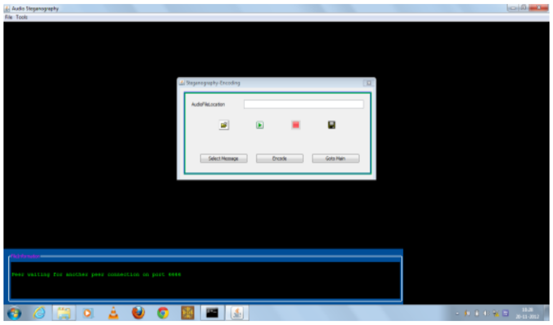
Figure 2 .Audio Steganography Model

6. IMPLEMENTATION AND RESULTS



6.1 Main Screen

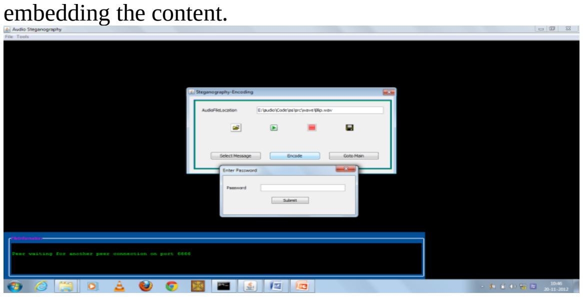
This screen contains two options, one is to encode and the other to decode. The user selects one of the two options and based on his choice the control is moved on to the encoding or decoding window.



Peer waiting for another peer connection on port 6666

6.2 Encoding

The actual encoding takes place. The window is provided with options to select the audio file which is the cover file and the text file which is the stego object. And also it provides the facility to play an audio . The audio is read in bit by bit form, each eighth bit of the image is used for embedding the content.



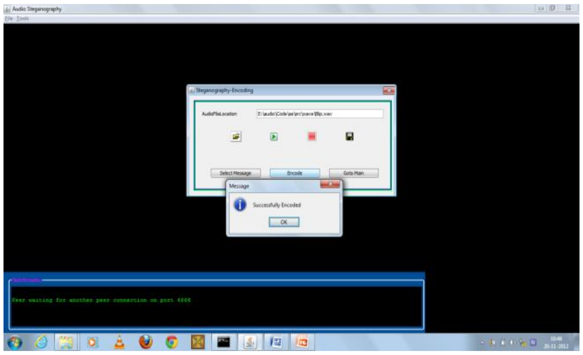
AudioFileLocation E:\audio\Code\as\src\wave\Blip.wav

Password

Peer waiting for another peer connection on port 6666

6.3 Password

To ensure protection of the file, an option to verify password is also provided. The password is encrypted using the same algorithm and hence secure.



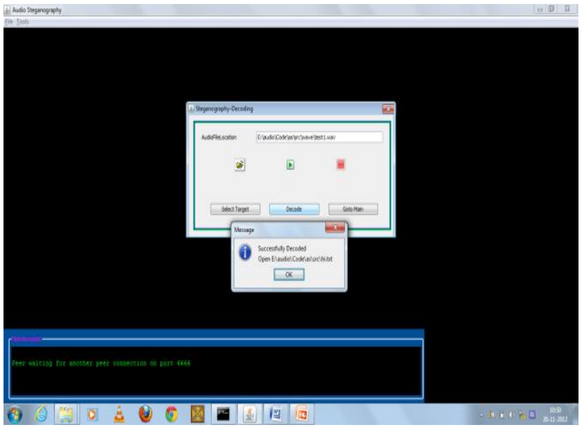
AudioFileLocation: E:\audio\Code\as\src\wave\Blip.wav

Successfully Encoded

```
Peer waiting for another peer connection on port 6666
```

6.4 Successful Encoding

From the above window , one can conclude that secret message is concealed within the audio file properly.



AudioFileLocation: E:\audio\Code\as\src\wave\test1.wav

Successfully Decoded
Open E:\audio\Code\as\src\hi.txt

```
Peer waiting for another peer connection on port 6666
```

6.5 Successful Decoding

In the decoding phase the embedded bits are just retrived. The LSB of the audio file are retrived and written in a text file from where it can be read. A password is provided to ensure secure transmission on the file. The secret message decoded from the audio file. This process done successfully by the same algorithm.The secret data and audio file is get separated finally at receiving end in a more secure manner.

7. CONCLUSION AND FUTURE WORK

The carrier or cover audio must be carefully selected as hiding information may introduce enough visible noise and the cover audio should contain some randomness. It is better for the steganographer to create own audios. It is a good, efficient method for improving the quality of the data by reduction of distortion And does not change the size of the file even after encoding and also suitable for any type of audio file format. Encryption and Decryption techniques have been used to make the security system robust. It is well modulated system but has been limited to certain restrictions. A better function can be determined to improve the ability of the technique. Combining still more steganography methods may improve the strength of the technique. The quality of sound depends on the size of the audio and length of the message selected by user. There are also other weighting algorithms like spread spectrum, echo data hiding etc., and those can be implemented and further the performance can be upgraded to higher levels practically. The algorithm could be extended for compressed audio file formats and very large audio files.

8. REFERENCES

[1] Cvejic, N. Seppanen, T. Increasing robustness of LSB Audio Steganography by Reduced Distortion LSB coding, Media Team Oulu Group, Oulu Univ.,Finland.

[2] Cvejic N. and Seppänen T. Increasing the capacity of LSB based audio steganography, Proc. 5th IEEE St. International Workshop on Multimedia Signal Processing,Thomas,VI,Dec 2002, pp. 336-338.

[3] F.A.P. Petitcolas, R.J. Anderson, and M.G. Kuh Information Hiding—A Survey Proc.IEEE, vol.87, no. 7, 1999, pp. 1062–1078.

[4] S.S.Divya , M. Ram Mohan Reddy, Hiding Text In Audio using multiple LSB steganography and provide security using steganography.

[5] Ming Chen; Ru Zhang; Xinxin Niu; Yixian Yan Intelligent Information Hiding and Multimedia signal processing,2006.IIH-MSP '06.International Conference on Dec 2006.

[6] Samir Kumar Bandyopadhyay,Biswajita Datta Higher LSB layer based Audio Steganography Technique. IJECT, Vol.2, issue 4.Oct- Dec2011.

[7] R.A.Santosa , P.Bao, Audio to image wavlet Transform based audio steganography,Proc.of 47th Int. Symposium ELMAR ,June 2005,pp. 209-212.

[8] C.C.Chang, T.S.Chen and H.S.Hsia,An Effective Image Steganographic Schema Based on Wavlet Transform and Pattern –Based Modification,IEEE Proceedings of the 2003 Int Conference On Computer Networks and Mobile computing,2003.

[9] K. Gopalan, Audio Steganography using Bit Modification ,Proc.IEEE, Int ,Accoustics,Speech, And Signal Processing,Vol.2,pp421-424,April 09.

[10] W.Bender ,D.Gruhl,N.Morimoto and A.Lu, Techniques for data hiding,IBM systems Journal,Vol.35,Issues 3&4,1996,pp313-336.

[11] Anderson, R., Petitcolas, F.On the limits of the Steganography ,IEEE Journal Selected Areas in Communications,16,4,474-481.

[12] Lee, Y., Chen, L., High capacity image steganographic model, IEE Proceedings on Vision, Image and Signal Processing, 147, 3, 288-294.