International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET

Real Time Speech Steganography for Secure Data Transmission

P Meghana Reddy[1], Sabah Samareen[1], Saniya Naaz[1], S Surekha[2]
B. Tech. Student[1], Assistant Professor[2]
ECE Department, Bhoj Reddy Engineering College for Women, Hyderabad, Telangana, India

ARTICLEINFO

Article History:
Accepted: 05 March 2023
Published: 24 March 2023

Publication Issue
Volume 10, Issue 2
March-April-2023

Page Number
134-140

ABSTRACT

Steganography is the art of hiding information into a cover object. The cover object could be any media like an image, audio or human speech. The essence of steganography lies in the ability to hide information in the cover object without degrading its quality and hence giving away hints of tampering. This project involves a spread spectrum representation-based speech steganography using discrete wavelet transform (DWT), which decomposes the cover speech signal into approximated and detail coefficients i.e., low frequency and high frequency components. Our proposed speech steganography provides enhanced imperceptibility since DWT reconstructs the decomposed information without degrading the quality of speech thus staying true to the nature of steganography. Our proposed approach is an extended version of existing Fast fourier transform (FFT) based steganography, where there is a lack of imperceptibility. Compared to existing FFT and any other existing methods it has been observed that there is a lower bit error rate and greater discrepancy. Apart from this we have noticed reduced effects of noise attacks.

Keywords : Steganography, Fourier Transform

I. INTRODUCTION

Communication between any two parties has always been subject to intentional or unintentional interception. This was earlier solved with cryptographic techniques. Cryptography maintained secrecy by making the information unintelligible, only the authorized user could acquire it with the appropriate knowledge of the suitable key, thus guaranteeing exclusive access to them. However, the blurred appearance of a secret message makes unauthorized users suspicious which is a drawback for cryptography. This drawback is covered by steganography which is a covert form of communication that leaves or shows no trace of communication whatsoever. It is made up of two words ‘stego’ meaning cover and ‘graphy’ meaning writing in Greek. Steganography these days makes use of cover objects like audio, video, images etc. but, this technique has been around several years before these media existed. Physical methods of steganography includes some interesting examples from ancient Greece and Rome some of which were, message tattoos on scalps of slaves on which hair were allowed to grow and they were sent across regions for communication, usage of invisible ink, messages under wax of writing tablets etc.

Steganography is a way of communicating such that the mere existance of communication is hidden. It is called as "covered writing", because of the presence of a "cover" used for sending any important secret message. It is a means for private and communication. The very presence of communication is hidden by embedding the secret message into the bland looking cover media objects, such as images. Steganography is a powerful tool which increases security in data transferring. In the steganographic scenario, the secret data is first concealed within another object which is called "cover object", to form "stego object" and then this new object can be transmitted or saved. Using different techniques, we can send secret data in the form of an image, a music file or even a video file by embedding it into the carrier, forming a stego-signal. At the receiver's end, the secret data can be recovered from the stego-signal using different algorithms.

Steganography is referred to hiding information or any secret message, this prevents unauthorized access. To obtain an effective steganography, one must need the following:

- Cover media object
- Secret message
- Embedding procedure
- Extraction process

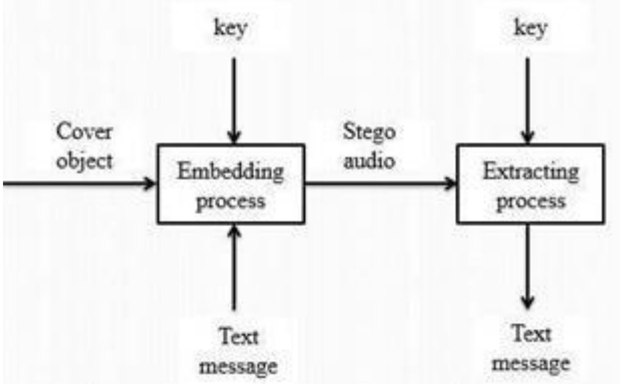


Fig 1 Components of Steganography

Figure 1 describes the entire procedure of steganography. The block diagram comprises of cover object, secret message, secure key, embedding and extraction procedures. The procedure starts off with obtaining the secret text message which will be embedded into a cover audio or speech with the help of key and an embedding procedure. This results in the output which is the stego speech being obtained. Then at the recipient end, the secret message will get extracted using the symmetric key and an extraction procedure.

II. EXISTING METHODOLOGY

A. Fast Fourier Transform

This segment sheds light on the existing FFT-based spread spectrum representation for speech steganographic system. The initial step of this procedure is to ensure that the cover speech is transformed into frequency space using FFT, which allows the computation of the discrete fourier transform (DFT) of a speech signal with reduced number of computations. The next move is to ensure that the message info which is to be embedded into the frequency domain signal of a cover speech is converted into binary format by utilizing ASCII codes. Spread spectrum technique allows the use of bits -1, 0, 1; this binary message info can now be spread over the channel using pseudo noise, chip rate as key and embedding gain factor alpha The obtained output of the above procedure Wt will be combined with the FFT signal to get the final output i.e. the stego speech. To obtain the message inverse FFT has to be performed. On this output obtained, the inverse embedding procedure is applied to extract the hidden message from the stego speech. Figure 2 depicts this entire procedure.

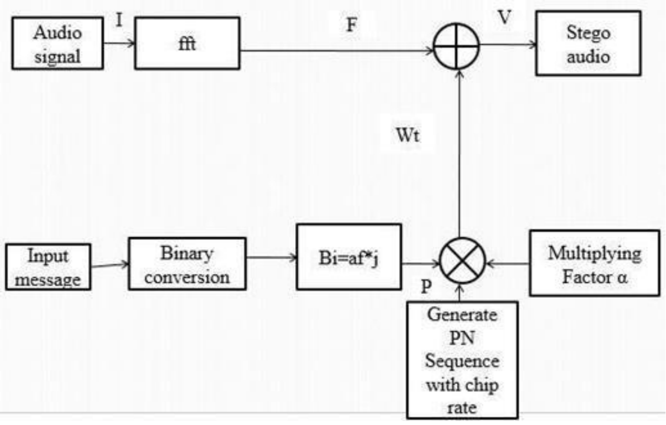


Fig 2 Speech Steganography using FFT

B. Other Existing Methods

Another technique for data hiding is based on phase coding. The first part of an audio segment is substituted by a reference phase that represents the data. In order to conserve the relative phase between segments, an adjustment is made in the phase of the succeeding segment. The series of steps of phase coding is as follows:

i. The audio signal to be used as cover is decomposed into smaller segments equal to the size of the message to be encoded

ii. A discrete Fourier transform (DFT) is then applied to obtain phase matrix of each segment

iii. The phase difference between each segment is calculated

iv. Identify the phase shifts between the consecutive segments. the relative phase differences between the

v. consecutive segments must remain unchanged

vi. Create a new phase matrix by using the new phase of the first segment and the set of original phase differences

vii. Use inverse DFT to regenerate the audio signal and then connect the audio segments together. The receiver determines the length of the secret message, then applies a DFT and extract the hidden message.

It has a low data transmission rate as the secret data are encoded only in the first segment of the cover signal. Any enhancement in the length of the segment may result in shifting the phase relations among the frequency elements of the segment, thus causing easier detection of the existence of a secret message. Thus, this algorithm is more efficient when hiding small amounts of data only.

Another technique is the Spread Spectrum (SS) coding method. Unlike the LSB coding technique, the SS coding scheme spreads the secret message using a code independent from the concrete cover signal across the frequency spectrum of the cover audio signal. The SS coding technique out performs LSB technique by offering a good quality for medium data transmission rates and is quite robust against steganalysis. However, the SS method may introduce noise to the stego audio just like the LSB technique.

Echo hiding is another steganography technique where an echo is introduced into the cover signal to embed data. For the process, three fundamental parameters need to be changed from the original cover signal: decay rate, time delay, and amplitude. These three parameters are defined and located below the human audible threshold limit which is different from the echo. The time delay is altered for the binary message to be hidden. The first and the second offsets/ time delays represent a one and a zero, respectively. Echo hiding provides better data transmission rate and is robust.

A high data rate yet simple algorithm is a one consisting of detecting the silence intervals of a cover signal and the corresponding length of these intervals and changing them with the secret messages. However, this method isn’t very robust.

C. Limitations of Existing Methodology

Another approach is the least significant bit (LSB) approach, which embeds the message info into the speech data based on LSB approach. It was however noticed that, at the reconstruction end it was difficult to extract message data accurately. This method however, is subject to small modifications that occurs due to format conversion or compression. Even though this is a simple algorithm with high data rate, there are several algorithms that challenge the robustness of this method.

The FFT based method provides enhanced performance over conventional speech steganographic systems. However, it is unable to extract the original embedded message due to the reconstruction issue of FFT algorithm and visibly the stego output and the original signal are quite distinct which gives a clear indication of some sort of tampering, this isn’t expected in practical applications. The proposed wavelet based speech steganographic system is an extension of FFTbased approach. Due to the higher bandwidth efficiency of decimated wavelet decomposition, there will be a lossless reconstruction of the stego audio i.e. the cover audio containing the hidden message.

III. PROPOSED METHODOLOGY

During the development of a high performance speech steganography system, the following attributes must be accounted for

- The amount of information that can be hidden- hiding capacity
- As the core of steganography is being indefinite, it is necessary to ensure that no degradation of the cover media occurs
- When the procedure of reconstruction is complex unauthorized access becomes difficult
- Accuracy
- Efficient techniques

Speech steganography has progressed greatly due advancements in speech compression and data hiding. The concept of signal decomposition into lower and higher frequencies has its application in speech steganography due to the fact that speech or human audio signal can be split into low and high frequency components. The main content of the speech is present in the lower frequency component whereas the noise, signal information and other stray components are present in the high frequency components. Thus, it would be sensible to hide the message in the high frequency component of the signal to avoid any detection. Speech is a baseband signal in which most of the relevant frequency components are confined to a bandwidth of 4 and 7 kHz for narrowband and wideband speech, respectively. The distribution of the first three speech formants represents the primary cues to English vowels. Figure 3(a) and 3(b) shows the broadband speech spectrum for both a voiced frame and an unvoiced fricative frame. For all vowels and most voiced consonants, i.e. that make up most of the speech, the magnitude spectrum has very weak components at high frequencies, this shows that the majority of the speech content is in the low frequency components.

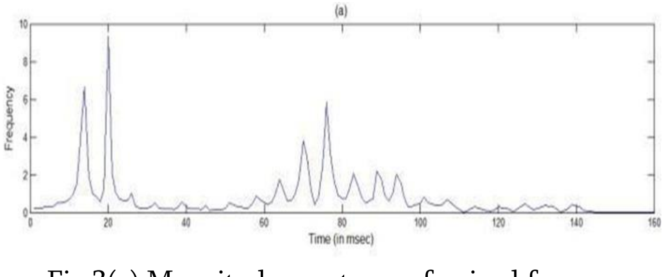


Fig 3(a) Magnitude spectrum of voiced frame

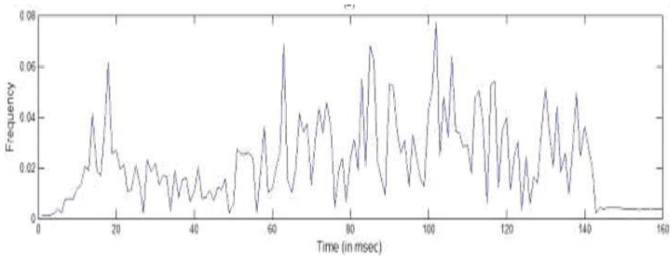


Fig 3(b) Magnitude Spectrum of unvoiced frame

A. Discrete Wavelet Transform

The wavelet transform transforms the signal from the time domain to the wavelet domain. This new domain contains more complicated basis functions called wavelets, mother wavelets or analyzing wavelets. Wavelet analysis is capable of enlightening aspects of data that other signal analysis techniques are unable to perform, aspects like trends, discontinuities in higher derivatives, breakdown points and self-similarity. The wavelet analysis enables splitting a signal in two parts, usually the high frequencies and the low frequencies part. This process is called decomposition. The edge components of the signal are largely limited to the high frequencies part. The signal goes through series of high pass filters to analyze the high frequencies, and goes through series of low pass filters to analyze the low frequencies. Filters of different cutoff frequencies are used to analyze the signal at different resolutions. The DWT involves choosing scales and positions based on powers of two, so called dyadic scales and positions.

B. Signal Decomposition

Starting with a discrete input speech signal, the DWT algorithm involves decomposing the signal into sets of coefficients. The approximation coefficients cA1 and the detail coefficients cD1 are there by obtained. In order to obtain the coefficient vectors, the signal goes through the low-pass filter LoD and through the highpass filter Hi_D for details. A down sampling by a factor of 2 or a dyadic decimation is then applied to obtain the approximation coefficients.

Mathematically the two-channel filtering of the discrete signal can be represented by:

$$cA1=\sum kcks2i-k \qquad (1)$$

$$cD1=\sum kgks2i-k \qquad (2)$$

These equations implement a convolution with a down sampling by a factor 2, then transfer the forward discrete wavelet transform. If the length of the initial stego-signal s is equal to n, and if the length of all filter is equivalent to 2N, then the equivalent lengths of the coefficients cA1 and cD1 are calculated by:

$$Floor((n-1)/2)+N \qquad (3)$$

This shows that the total length of the wavelet coefficients vector is always slightly greater than the length of the initial signal due to the filtering process used. Wavelet decomposition tree can be constructed by following an iterative decomposition process with successive approximations. Thus, the input stego signal is broken down in several subordinate resolution components.

C. Algorithm

This segment discusses the algorithm along with the required equations. The cover object to be used in this project will be a speech file in the WAV format. No, the information that is in the byte sequence format will be converted to ASCII and from that using spread spectrum method it will be converted to bits ranging from -1 to 1.

$$A=\{ai|ai\in\{-1,1\}\} \qquad (4)$$

Next, to create pseudo noise(key) the WAV signal's amplitude is represented as 16 bit signed integer, the values range form 2^{15+1} to 2^{15-1}. These values will be divided by2^{15-1} to obtain value lying in the range -1 to 1 and then this data is converted to frequency domain using FFT. Now, we have a sequence of values from -1 to 1 with chip rate cr. For n information signals we require cr x n pseudo noise sequences

$$P=\{pi|pi \in \{-1,1\}\} \qquad (5)$$

Modulate each information signal with the PN sequence until cr times, by multiplying the value. It will produce a signal B which is the distributed signal of A and of course with length cr times its original length. Initially, spread the information in A to B as follows:

$$B=\{bi|bi=ai,\ j.cr<=i<(j+i).cr\} \qquad (6)$$

Equations (5) and (6) are multiplied bit by bit i.e. modulated and multiplied with factor alpha . Equation (7) represents the message signal with the pseudo noise, equation (8) represents the stego-audio as the output where, v is the cover signal, w is the message signal and v' is the stego audio.

$$wi=\alpha.bi.pi \qquad (7)$$

$$vi' = vi + wi \qquad (8)$$

D. Steganalysis

Steganalysis is the reverse process of steganography that is performed at the receiver's end. It first determines whether there is a message hidden in the media and then aims at retrieving it. The effectiveness of steganalysis is dependent on the degree of sophistication and personalization of the techniques used by an attacker.

For the information to be retrieved, the receiver must generate the same PN sequence at its end. Each cover object signal will be multiplied with the corresponding PN sequence, which is shown by the equation (9) below

$$\sum_{i=j.cr}^{(j+1).cr-1} (p_i{}' v_i = p_i v_i + \alpha b_i p_i^2) \qquad (9)$$

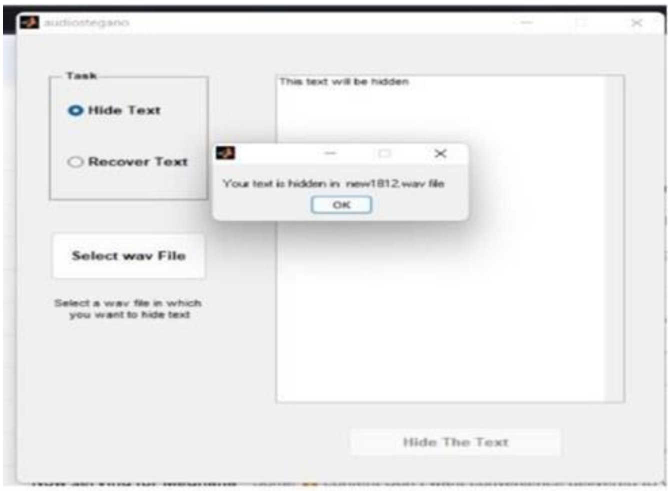
The value of the first term will be close to 0 for a large chip rate. This is because the random value of PN sequence causes the sum of the signal approaching 0. While the second term has interesting properties. Because the PN sequence has value 1 or - 1, then the result of pi^2 is 1. Thus, the term can follows:

$$\sum_{i=j.cr}^{(j+1).cr-1} \alpha b_i \qquad (10)$$

Form Equation (10) we can observe that we have obtained the bi term which is the message signal containing the pseudo noise. Because we have defined bi has a value 1 or 1, then we simply conclude that if the term exceeds the value of zero, we assume that the information retrieved is 1 and if the value is less than zero, we assume that the information retrieved is 0. This is the reason we choose the domain of B and P. From the previous explanation, we can conclude that the value of bi must exceed a certain threshold value in information retrieval.

IV. RESULTS AND DISCUSSIONS

This section describes the experimental analysis of the proposed speech steganography. All simulations were performed in the MATLAB 2018a environment. We tested the proposed and the existing method for a speech sample from. The user has to type the secret message (8 to 10 characters) to be embedded in the cover speech. Figure 4 shows the secret message entered into the text box. Figure 4(a) and 4(b) show that the waveforms of both the cover signal and stego-audio are quite visually similar hence maintaining the essence of steganography by remaining imperceptible. The bit error rate values (BER) are quite high, as shown in Table 1. Figure 4 shows the performance of the proposed real-time speech steganography, which reveals the similarity between cover and stego speech and leads to higher imperceptibility compared to the existing FFT-based speech steganography.



audiostegano

This text will be hidden

Your text is hidden in new1812.wav file

Select a wav file in which you want to hide text

Figure 4 Output window

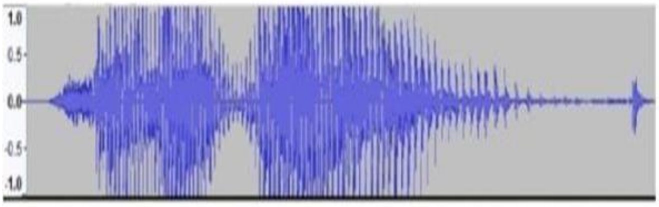


Fig 4(a) Waveform of human speech signal

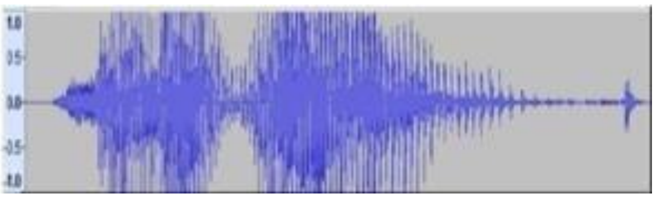


Fig 4(b) Waveform of stego-audio

Table1 BER values of existing and proposed speech steganography

Parameter	FFT-based steganography	DWT-based steganography(proposed method)
BER without noise	0.00145	0.0000001
BER with noise	4.25	0.000452

V. REFERENCES

[1]. G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. (references)

[2]. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[3]. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4]. K. Elissa, “Title of paper if known,” unpublished.

[5]. Interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[6]. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magnetooptical media and plastic substrate Annual Conf.
Magnetics Japan, p. 301, 1982].

[7]. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

Cite this article as :

P Meghana Reddy, Sabah Samareen, Saniya Naaz, S Surekha, "Real Time Speech Steganography for Secure Data Transmission", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 23951990, Volume 10 Issue 2, pp. 134-141, March-April 2023.
Journal URL : https://ijsrset.com/IJSRSET201551