ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 1, January 2018

Information hiding in low bi-rate speech codec: A Review

[1] Ruchi Patel, [2] Nikunj V. Tahilramani, [3] Ninad Bhatt Chandani, [4] D. Maheshwari
[1] PG student, [2] Assistant Prof. & Head, [3] Professor & Head, [4] Assistant Prof
Department of E&C
Silver Oak College of Engineering & Technology Ahmedabad, Gujarat, India.

Abstract: **- The purpose of this review paper is to provide concrete information about information hiding in low bit-rate speech codec. In recent years, due to the uninterrupted escalation of network bandwidth and the augmentation of network convergence gravitating, Voice over Internet Protocol (VoIP) is extensively used. This paper addresses the literature survey of information hiding in low bit-rate speech codec in time, frequency and wavelet and spectrum domain. The technique reviewed here are least significant bits(LSB), vector quantization, phase coding(PE), analysis by synthesis(ABS) based, quantization index modulation(QIM) which deals with information hiding in low bit rate speech codec. Information hiding is deliberate in terms of embedding capacity, signal to watermark ratio, hiding capacity.**

Key words: **- Information hiding; speech steganography; stego-signal; embedding; VoIP.**

I. INTRODUCTION

In information society, one of the most important features is information communication. New ap¬proaches of secure speech communication are planned to transmit security informa¬tion via Voice over Internet Protocol (VoIP). In past few years, voice over IP (VoIP) has become a very popular Internet streaming media communication service. VoIP is suitable to enable hidden communication throughout IP networks. The main idea of information hiding derives from early steganography, more strictly called "information cover," which literally means "covered writing". Information hiding can also be called "information concealing". Information hiding technology integrates several research subjects and technologies in various fields, such as communication, cryptography, network signal processing, and voice and image coding. Information hiding in low-bit rate speech codec is done in time, frequency and wavelet and spectrum domain. Information hiding in the low bit-rate speech stream is distributed into three classifications according to their hiding positions. In first classification information are hided in the compressed speech stream directly changing the value of some code element. In second classification information are hide in the prediction step of the short-term predictor (STP) of the speech codec. In third classification information are hide in the long-term predictor (LTP), For example, When vector quantization (VQ) process of linear prediction coefficients are done at that time the quantization-index modulation(QIM) technique hide the secret information. In addition, in many codecs the linear predictive analysis-by-synthesis coding model is mostly used, that reduce the deformation by decoding the encoded signal. The technique analyzed here are least significant bits(LSB),vector quantization, phase coding(PE),analysis by Synthesis (ABS) based, quantization index modulation(QIM), DWT FFT Technique which deals with information hiding in low bit-rate speech codec. This paper presents different information hiding techniques and explores their potential and limitations to insure secure communication.

II. INFORMATION HIDING MODEL

Information hiding has two specifications. The first specification is transparency that needs data set $\hat{X}$ to be extremely close to X. The second specification is robustness that needs that no matter what process are performed on X, in which secret information M must be kept hidden[21].

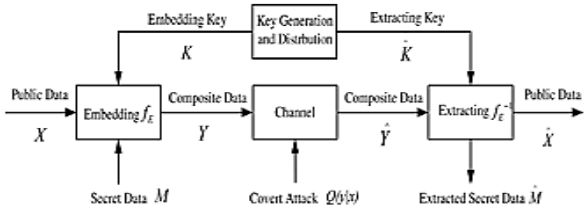

Figure:1 General model for information hiding application[21]

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 1, January 2018

Public data X is input of model which is carrier for secret data and that contain plain text or original data. Secret data M requires be secretly transmitting and embedding into original data. In embedding operation a mathematical function fE is used to embed secret data into original data using embedding key K. In original data the embedding secret data formed composite data Y and Ŷ. Y is transmitted to channel under covert attack Q(y/x) and then composite data Y is formed. In extracting operation a mathematical function fE-1 is used to extract secret data from composite data using extracting key K. The aim of covert attack Q(y/x) is digging out secret data. At receiver secret data M̂ are extracted.

III. DIFFERENT HIDING TECHNIQUES

1.LSB coding technique

LSB coding is one of the most widely used technique in information hiding in speech. LSB coding technique is based on replacing the LSB of binary string of each sample of digitized speech through binary identical of secret message [1]. For example in digitized speech the letter X (binary identical 1011000) is hidden where each of sample is constituted with 16 bits, then of 7 successive samples is replaced with each bit of binary equivalent of the latter X as illustrated in Table.

Table: 1 Example of LSB coding:[1]

Sampled Speech Stream (16 bits)	'X' in binary	Speech stream with encoded message
1101 1100 1011 1000	1	1101 1100 1011 1001
1000 1100 0011 1011	0	1101 1100 1011 1000
1011 1100 1101 1000	1	1101 1100 1011 1001
1000 1100 1011 1100	1	1101 1100 1011 1001
1011 1100 0111 1101	0	1101 1100 1011 1000
1000 1100 0011 1100	0	1101 1100 1011 1000
1101 1100 0111 1111	0	1101 1100 1011 1000

In data hiding as the number of LSB's can be increased, the resulting noise also increased in speech. By discarding entire LSB plane an attacker can simply perceive the message[2]. For that other bits can be flipped to facilitate a new sample which is closer to original that reduced embedding error. For example, if the original sample value was 4 (equivalent binary 0100) and bit which is hidden into 4th LSB layer is 1, then LSB algorithm produced sample contain value 12(equivalent binary 1100) and result of planed algorithm produces sample which contain value 3(equivalent binary 0011), that is closer to original sample[7]. At receiving side, receiver desires admittance to the string of sample used in embedding method to remove a secret message from LSB encoded speech[4].

2. Vector Quantization Technique

The principle of vector quantization is block coding which used to compress media to make systematic use of network bandwidth and data storage space[5]. The hiding method for each watermark bit is accomplished by penetrating the best coordinated codeword for each LSP input and remaining LSP vector under the limitation that randomly selected bits are reliable with the watermark bits to be embedded. In each speech frame as a minimum 1 bits of watermark is hided. For hiding more information and maintaining the watermarked speech superiority, the location of hiding is chosen for keeping the SNRseg in a fixed range. In the watermark embedding system encoder penetrates for the codeword which gives the least deformation between the input LSP(remaining) vector and the recognized vector within the restriction of the index. Extremely adjacent reorganization vector may be get under the restriction, and the location prefering block may complete the predictable transaction between the quality and the payload[14]. Later than the replacement, the probability of distorted codeword (reorganization vector) best matches the input vector is extremely small. From Figure 1 it can be explicated better. If no watermark bits embedded, input vector quantized as c1 but using the perceptually insignificant bit replacement scheme, input vector quantized as c3 , c4 , c5 and so on. The probability as c2 should quantized extremely little. Since this is the best match penetrating method of vector quantization method, with this anticipated method c2 would be selected , and that is reduced the distortion after watermark embedding.

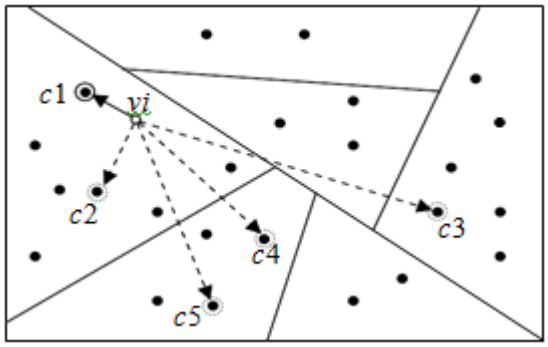

Figure: 2 Vector quantization [5]

3.Phase Coding Technique

In phase coding process the phase of an original audio segment is exchanged with a source phase that constitutes the information [7]. The phase of consequent segments is attuned with purpose of protect the comparative phase between segments. Phase coding is used as one of the most efficient coding process regarding the signal-to apparent noise ratio. In this process perceptible phase dispersion will arise due to phase change among each frequency component. On the other hand, given that the variation of

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 1, January 2018

phase is adequately small (adequately small depends on the viewer; proficient in broadcast radio can deduce variations that are minor to an average viewer), impossible to hear, that can be accomplished[8].

4. Analysis by synthesis algorithm

The key feature of ABS algorithm presents a speech synthesiser into the speech coder, relating with the speech analyser[9]. The speech coder produced synthesized speech and it would detect at decoder. Difference between the synthesised speech and original speech would be calculated and reduces the error between original speech and synthesized speech[10].

In this algorithm, speech coding data are employ as a carrier and a speech synthesizer is employ to hide confidential information into carrier through coding procedure. After hiding 1 bit of confidential speech into multi-codeword, the original speech and stego(mixed) speech are decoded separately. Difference between the original speech and synthesized speech is calculated as the minimum square error (MSE).

ABS algorithm model:

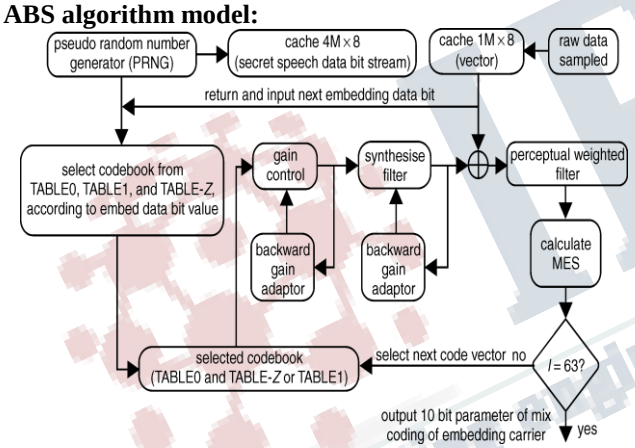

Figure: 3 ABS algorithm block diagram[9]

Two speech are recorded first, one is original speech and second is secret speech. The original speech carrier are used in embedding process to embedded the whole secret speech into it. The original speech carrier x(n) generating a vector or sub frame every five successive samples. An optimal code book's is length is 10bts. To minimize complexity of codebook search, it divided into two vector table in which one is shape codebook which contains 128 independent code vector and second is gain codebook that contain eight unit value that have zero symmetry. The shape codebook vector is selected as embedding parameters. An embedding encoder chooses a code vector depends on an embedded bit: if embedded bit is zero, it select code vector from TABEL-0 and TABEL-Z and if embedded bit is 1, it select code vector from TABEL-1.

In gain controller select vector is amplified as an exciter of synthesis filter, forming a local decoded signal. The frequency weighted MSE between the local signal & original signal is computed to maintain minimum error. Using last quantize gain of excitation signal the excitation gain is extract and update one-by-one vector. The frequency weighted MSE is calculated[9].

$$MSE = \|x(n) - \tilde{x}_{ij}\| = \sigma^2(n)\|\hat{x}(n) - g_i H y_j\|$$

Here, H stands for response function of synthesis filter cascade with perceptual weighted filter. gi ith gain value of TABLE-Z. yj is jth code vector in TABLE-0 or TABLE-1[9].

$$\hat{x}(n) = \frac{x(n)}{\sigma(n)}$$

5. QIM (Quantization index modulation) method

In Quantization index modulation method one confidential bit is hided in each index[10]. The aim of this scheme is to split the entire codebook in two parts and allocate a label '0' or '1' to all codeword. Simply the equivalent part of the codebook is used while a secret bit is embedded. At receiver, the hidden bit is removed by examine where codeword belongs to in entire codebook. The receiver is capable to remove the message directly from the compressed speech stream when the channel is reliable. At same instance instead of searches whole codebook the embedding algorithm searches only in half of codebook, so additional delay does not appear. The distortion is augmented as the number of code words used in quantization is reduced. To reduced distortion, the most important duty is to discover an ideal codebook divider system. The Complementary Neighbor Vertices (CNV) algorithm is used for that[11].

6. DWT FFT Technique

In this method secret speech signal is hide into coefficients in the wavelet domain. DWT divides the original speech into low and high frequency components in which low frequency is most important for speech realization and high-frequency component crashes flavor or nicety to the signal[12]. The speech signal are diverged in approximation and details using wavelet analysis in which details are low scale and approximation are high frequency components. To assist the hiding process secret speech and original speech are exposed to pre-processing. The original speech is divided into L-ms frames. Then DWT of each frames are calculated to decay into high and low frequency. After that FFT are applied to high frequency wavelets that generating a spectrum. That spectrum is then decayed into magnitude and phase spectrum. The hiding process contains constituting last L element of acquired spectra by the LPC parameters of secret speech. Each frame of secret message is hided in low amplitude high frequency section of magnitude spectrum of original signal[13].

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 1, January 2018

Evaluation of hiding techniques

In digital audio and speech it is easiest way to hide information. LSB method is simple and effortless to execute, embed and extracts information quickly and has a elevated hiding capacity. LSB information hiding approach is used in G.711 and G.729 speech codec[1],[2],[4].

In vector quantization process, the caliber of watermarked speech signal can be controlled. In vector quantization minimum 3 bits are hided in each speech frame. In [5] the G.729 codec test the effectiveness of this process.

When we have small quantity of information that needs to be hided at that time phase coding method is applied. In phase coding the audible distortion is extremely condensed, as phase changed slowly and transitioning between phase changes. Channel capacity in phase coding diverse from 8bps to 32 bps depending on contain of sound[6],[7],[8].

Improving the speech superiority with confidential speech data can be accomplished by ABS algorithm. In this algorithm information rate at 3.2 kbps can be hided into carrier speech[9].

Quantized index modulation accomplishes good rate-distortion-robustness execution[11]. As a outcome, these methods have very constructive performance characteristics in terms of their attainable transaction along with the toughness of embedding, the degradation to the host signal effected by hiding quantity of information hided. In [10] G.723.1 verify the efficiency of this technique.

DWT-FFT technique utilize high frequencies using a DWT, and then utilize the low-pass spectral characteristics of speech magnitude spectrum for hiding another speech signal in the low-amplitude high-frequencies region of original speech signal. This process accomplishes 3.68 and 4.14 PESQ average for DWT-FFT and FFT algorithm respectively[12].

IV.PARAMETRES SIGNIFYING AMOUNT OF INFORMATION EMBEDDING

The signifying amount of information embedding parameters are listed below[12]:

a) Data Hiding capacity (DHC): it is the ratio between the total summation of number of variable length frames to the number of variable length frames in the unvoiced periods of the speech signal. Data hiding capacity is associated with the number of data keys or a particular pattern signifying the secret data pattern for a particular look up table of different entities to show the steganographic data information to the receiver.

The hiding capacity is defined as:

DHC=

$$\frac{\text{Summation of number of variable frame length}}{\text{Total number of variable length frames in silence periods}}$$

b) Data embedding to signal ratio: This specific parameter indicates the amount of secret data which can be fetched at the receiver in the form of variable frame of silence speech signal.

Data embedding to signal ratio can be computed as follows:

$$\frac{\text{Number of frames utilized for information embedding in speech}}{\text{Number of total frames in the speech}}$$

c) Data embedding utility factor: it is defined as the ratio of no of frames utilized for data embedding to the total number of frames of silence passage. This parameter is associated with the summation of total variable length frame number incorporated in modulo mathematical operation to hide a specific secret data and total number of silence speech samples in the speech wav file. The variable frame length number is always greater than the size of minimum silence interval which is equal to one frame size of the speech It is calculated as following:

$$\frac{\text{Number of frames utilized in data embedding}}{\text{Total number of frames of silence intervals of speech signal}}$$

V. SUBJECTIVE AND OBJECTIVE MEASURES

In subjective estimation, MOS analysis is being shown[12]. MOS (Mean opinion score) that is used to conclude the superiority of the quantized index modulated speech at the output of transmitter which is nothing but the hybrid steganographic or watermarked speech. The quality of output speech is requested to observe by randomly 5 to 10 persons. They requested to rate the quality of the speech signal according to the options accessible in Table 2 by playing the speech in noiseless environment.

Sr No	Choice	MOS
1	Excellent	1
2	Good	2
3	Fair	3
4	Poor	4
5	Unacceptable	5

To analyses the concert of these all techniques, the indiscernibly and detecting rate of hidden information are estimated. Two objectives are used to evaluate difference between original speech and secret speech which are: PESQ (perceptual evaluation of speech quality) and SNRseg (average subsection signal-to-noise ratio).
For anticipating the subjective quality of speech PESQ is used[8]. Perceptual based study is performed by analysis of perceptual evaluation of speech quality (PESQ). PESQ is intended to analyze particular parameters of audio, including time warping, variable delays, transcoding, and noise. Higher the MOS (mean opinion score) value better the subjective speech quality. The rate of SNRseg represent the deformation quantity induced by the embedded information in the original speech Sc (m,n). In speech signal noisy signal's SNR is below 20dB as well SNR of 30 dB and above denoted that speech signal is conserved. SNR rate is obtain by following equation[8]:

$$SNR_{dB} = 10\log_{10}\left(\frac{\sum_{n=1}^{N}|s_c(m,n)|^2}{\sum_{n=1}^{N}|s_c(m,n)-s_s(m,n)|^2}\right)$$

ss(m, n) is the stego-audio signal where: m = 1, ...M and n = 1, . . .N. M is the number of frames in milliseconds(ms) and N is the number of samples in each frame.

VI. CONCLUSION

In recent years due to widely use of internet streaming, security against digitized audio or speech is required. This paper presents the information hiding techniques to achieve high efficiency and best security. These techniques have their own potential and limitation to insure secure communication. Also, differences between reviewed techniques are based on their application in which they used their hiding capacity and hided information security level.

REFERENCES

[1] Liu, Lihua, et al. "Perceptually transparent information hiding in G. 729 bitstream." Intelligent Information Hiding and Multimedia Signal Processing, 2008. IIHMSP'08 International Conference on. IEEE, 2008.

[2] Ito, Akinori, and Yoiti Suzuki. "Information hiding for G. 711 speech based on substitution of least significant bits and estimation of tolerable distortion." IEICE transactions on fundamentals of electronics, communications and computer sciences 93.7 (2010): 1279-1286.

[3] R. Miao and Y.F. Huang, " An Approach of Covert Communication Based on the Adaptive Steganography Scheme on Voice over IP", Communications (ICC), IEEE International Conference on Kyoto ,5-9 , pp. 1- 5, June 2011.

[4] Jin L, Ke Z, Hui T (2012) Least-significant-digit steganography in low bitrate speech. In: IEEE International Conference on Communications (ICC), 2012. IEEE, pp 1133–1137.

[5] Liu, Ji-Xin, Zhe-Ming Lu, and Hao Luo. "A CELP-speech information hiding algorithm based on vector quantization." Information Assurance and Security, 2009. IAS'09. Fifth International Conference on. Vol. 2. IEEE, 2009.

[6] Bender, Walter, et al. "Techniques for data hiding." IBM systems journal 35.3.4 (1996): 313-336.

[7] Djebbar, Fatiha, et al. "Comparative study of digital audio steganography techniques." EURASIP Journal on Audio, Speech, and Music Processing 2012.1 (2012): 25.

[8] Kulkarni, Sheetal A., Patil SB Patil, and B. S. Patil. "A Optimized and Secure Audio Steganography for Hiding Secret Information-Review." Journal of Electronics and Communication Engineering (IOSR-JECE) (2012): 12-16.
[9] Wu, Zhi-jun, Wei Yang, and Yi-xian Yang. "ABS-based speech information hiding approach." Electronics Letters 39.22 (2003): 1617-1619.

[10] Xiao, Bo, Yongfeng Huang, and Shanyu Tang. "An approach to information hiding in low bit-rate speech stream." Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE. IEEE, 2008.

[11] Chen, Brian, and Gregory W. Wornell. "Quantization index modulation methods for digital watermarking and information embedding of multimedia." Journal of VLSI Signal Processing Systems 27.1-2 (2001): 7-33.

[12] Tahilramani, Nikunj V., and Ninad Bhatt. "A hybrid scheme of information hiding incorporating steganography as well as watermarking in the speech signal using Quantization index modulation (QIM)." Communication Systems, Computing and IT Applications (CSCITA), 2017 2nd International Conference on. IEEE, 2017.

[13] Rekik, Siwar, et al. "Speech steganography using wavelet and Fourier transforms." EURASIP Journal on Audio, Speech, and Music Processing 2012.1 (2012): 20.

[14] Patel, Shruhad Kumar J., and Nikunj V. Tahilramani. "Information Hiding Techniques: Watermarking, Steganography: A Review."

[15] Wu, Zhijun. Information Hiding in Speech Signals for Secure Communication. Syngress, 2014.

[16] Li, Song-bin, Huai-zhou Tao, and Yong-feng Huang. "Detection of quantization index modulation steganography in G. 723.1 bit stream based on quantization index sequence analysis." Journal of Zhejiang University-Science C 13.8 (2012): 624-634.

[17] Xu, Tingting, and Zhen Yang. "Simple and effective speech steganography in G. 723.1 low-rate codes." Wireless Communications & Signal Processing, 2009. WCSP 2009. International Conference on. IEEE, 2009.

[18] Lin, Rong-San. "An Imperceptible Information Hiding in Encoded Bits of Speech Signal." Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2015 International Conference on. IEEE, 2015.

[19] Huang, Yongfeng, et al. "Steganography integration into a low-bit rate speech codec." IEEE transactions on information forensics and security 7.6 (2012): 1865-1875.

[20] Liu, Peng, Songbin Li, and Haiqiang Wang. "Steganography integrated into linear predictive coding for low bit-rate speech codec." Multimedia Tools and Applications 76.2 (2017): 2837-2859.

[21] “Information Hiding in Speech Signals for Secure Communication”, Zhijun Wu, ISBN: 978-0-12-801328-1.

[22] Katzenbeisser, Stefan, and Fabien Petitcolas. Information hiding techniques for steganography and digital watermarking. Artech house, 2000.

[23] Bandyopadhyay, Samir K., et al. "A tutorial review on steganography." International conference on contemporary computing. Vol. 101. 2008.