International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 11, No.1 (Feb-2022)

https://dx.doi.org/10.12785/ijcds/110167

Covert VoIP Communication based on Audio Steganography

Baneen Q. Abd Ali[1], Haider I. Shahadi[1], Muayad S. Kod[1] and Hameed R. Farhan[1]

[1]*Department of Electrical and Electronic Engineering, University of Kerbala, Karbala, Iraq*

Received 1 Jul. 2021, Revised 31 Oct. 2021, Accepted 28 Jan. 2022, Published 15 Feb. 2022

Abstract: Voice over Internet Protocol (VoIP) is a popular and important internet protocol for real-time voice calling. It is used in several software applications such as Skype, WhatsApp, and Google Talk. However, communications over the internet can easily be exposed to hacking and eavesdropping. In secret communication, a private channel with robust encryption can provide the required protection for the communications. But private channels have more constraints than the public channel, such as specific communication devices and specific regions to cover. On the other hand, public communications such as VoIP can be done from anywhere without extra costs. This paper proposes a real-time covert communication approach that generates a secret speech channel inside a public speech channel based on audio steganography. The process requires high embedding capacity in a real-time manner. Therefore, the proposed approach encodes the secret speech using Internet Low Bit Rate Codec (iLBC) that can be used as a VoIP source Codec to compress the secret input speech with auto-correction for speech quality degradation over a lossy channel. Also, the cover speech is compressed by G.711 encoder at synchronous time periods (for each 20 ms) with respect to the secret speech to meet the requirements of the VoIP. Subsequently, the secret data hides in compressed cover speech data in real-time. The embedding procedure employs the only high strength data to maintain the quality of the cover data after embedding at greater than 40 dB in terms of signal-to-noise ratio (SNR). The secret data is 100% retrieved without error in the case of a lossless channel. This approach can be used in secure communications such as military and government communications.

Keywords: Audio steganography, Voice over Internet Protocol (VoIP), High embedding capacity, Internet Low Bit Rate Codec (iLBC)

1. Introduction

In the last few years, steganography has become widely used in different fields and applications; essentially, those require higher security levels such as military communication systems, online bank transactions, online payment, and medical services. In general, steganography is the process of hiding secret data in cover data, where several digital media can be used as a cover, such as images, text, audio, and video files [1]. Steganography in digital audio is more challenging due to the sensitivity of the human auditory system (HAS) compared to the human visual system (HVS) [2], [3]. This challenge is increased in the case of hiding audio within another audio in real-time.

The key features of the efficient audio steganographic system are; high hiding rate, imperceptibility, robustness, multi-levels of security, low complexity, and real-time communication capability. These features are contradictory to each other, such that increasing the hiding capacity leads to degradation in the robustness of secret data and imperceptibility of stego-file. Therefore, this trade-off is a complicated task. With the rapid implementation of the Voice over Internet Protocol (VoIP) applications over the Internet, most of the researchers are mainly focusing on using VoIP services in steganography by embedding the confident data in VoIP streams. This process provides an additional level of security, where it does not give the warden sufficient time to discover the real-time embedded secret data [4]. However, to realize real-time communication requirements on one side and to protect the secret embedded data on another side, there is a difficulty in introducing the hiding algorithm with more secure operations. Therefore, this paper proposes a scheme that meets real-time covert communications by creating a secret channel for embedded voice communication in a public channel (as a carrier or cover voice).

The rest of the paper is organized as follows: Section 2 presents the related work achieved in audio steganography for real-time communications. Section 3 illustrates the proposed scheme in this paper. Section 4 discusses the proposed system's results and compares our work with some related work. Finally, section 5 introduces the conclusion of the entire paper.

2. Related work

Recently, research on real-time covert communication-based VoIP has attracted the attention of many researchers.

E-mail address: baneen.q@s.uokerbala.edu.iq, haider_almayaly@uokerbala.edu.iq, muayad.kod@uokerbala.edu.iq, hameed.r.f@uokerbala.edu.iq

In the literature, two-hybrid approaches based on VoIP were introduced. One hides the encrypted secret message (encrypted with m-sequence encryption technique) into the cover speech using the least significant bits (LSBs) algorithm [5]. This algorithm provides adequate capacity, security, and low latency suitable for VoIP requirements. The other proposes the notion of partial similarity value (PSV) for matching the similarity between the LSBs of cover speech and secret message to set the proper value for the threshold PSV [6]. This approach can achieve a good balance between steganographic transparency and embedding capacity. Both above hybrid schemes are less sensitive for additive noise due to the use LSB technique.

The authors in [7] extended [5], [6] by proposing a steganography scheme based on a comprehensive adaptive partial matching steganography to measure the similarity between the cover and embedded message. The balance between the steganographic transparency and bandwidth is achieved by employing two thresholds of PSV and an m sequence. Also, to reduce the delay and realize real-time requirements, the encryption is integrated with the embedding process. The approaches in [5], [6], [7] were evaluated using ITU-T G.729a as VoIP codec for cover speech. By revealing the silence intervals of cover speech. Authors in [8] presented a real-time steganography approach to embedding the secret data into the cover signal by altering the number of the silence intervals. The approach is robust to MPEG-1 layer III (MP3) compression and noise, but the embedding capacity is low. Another real-time steganography technique is proposed in [9] that hides a secret message encrypted with the Advanced Encryption Standard (AES-128) algorithm into cover speech samples encoded by PCM Codec. The embedding algorithm adopted on LSB replacements with parameter R as the interval to hide each secret bit into one byte of cover audio. With increasing the value of R, the suggested model is more immune to simple statistical analysis, but the embedding capacity is reduced. In [10], an adaptive VoIP audio stream was presented. It enhanced the security of the traditional LSB algorithm, which is used to embed the secret message within VoIP audio streams by adopting three techniques: value-based multiple insertions (VAMI), voice damage offset (VODO), and voice activity detection dynamic insertion (VADDI). Also, it used G.711 as a cover audio codec to evaluate the system efficiency. The introduced system has better transparency, but the hiding rate is low (about 102.28 bps). A covert steganography system was suggested in [11], which divides the encrypted secret data with a block cipher into blocks for randomly embedding each block into VoIP streams using chaotic mapping. The system can protect the integrity of secret data by computing the message digest and sending it to the receiver. Also, it has sufficient security due to the use of key distribution. However, the embedding bit rate is low, which is between 0.5 and 8 kbps. The researchers in [12] studied a novel high capacity embedding algorithm to reveal the inactive frames of low-bitrate audio streams, which are more appropriate for hiding secret data than the active frames. This algorithm can obtain perfect imperceptibility, and high hiding rate. By employing the characteristic of the speech codec, paper [13] proposed a novel technique to hide sensitive information into the excitation pulse positions of the G.723.1 codec. This technique offers good security and efficiency as compared with existing similar work. In [14], speech steganography scheme based Fast Fourier Transform (FFT) speared spectrum representation was introduced. The aim was to embed a text message encrypted with pseudo-random sequences (PN) into the frequency space of cover speech with good performance and imperceptibility.

Several real-time covert VoIP steganography approaches are proposed, as discussed above. However, the embedding capacity for most of them is unsuitable for embedding real-time secret speech into real-time cover speech. Also, most of them did not have auto-correction for lost secret speech packets. Therefore, it is necessary to propose a lossless covert VoIP steganography scheme that has a high hiding rate of up to 25% of the cover speech size and excellent imperceptibility above 44 dB in terms of signal to noise ratio (SNR). Moreover, it has a packet loss correlation system and provides good robustness against additive noise due to the use of the hold bit factor (HB) to increase the depth of embedding. Also, a Pseudo-Random Generator (PRN) with a secret key is employed to improve the security level. The scheme is implemented using VoIP public channel as a carrier for secret communication. It is known that every channel has a limited capacity, so the great challenge within this work is the capacity of the public channel in addition to the proportional relationship between the processing time of the embedding algorithm and the real-time communication requirements.

3. The proposed real-time covert communication based VoIP

The proposed approach is designed to create a secure channel for the transferred secret data within the VoIP network. As shown in figure 1, the conversation between two ordinary persons in a public channel can be considered a carrier for the secret data exchanged between two essential persons. By employing a model operator with high energy cover samples, the secret speech of the first important person is embedded into the cover speech of the first ordinary person. There is no permission for any eavesdropper or even the second ordinary person at the receiver end to access the embedded secret data at the VoIP network. The second important person can retrieve the secret speech after applying the proposed recovery algorithm on the stego-speech (cover speech with the hidden confident data).

The proposed covert communication system compresses the input cover speech using G.711 Codec that most VoIP communications widely support. G.711 speech Codec is a narrow band Codec that is ordinarily designed for use in telephony. There are two versions of G.711, μ-law and A-law, which slightly differs in the number of bits per sample for the input speech that eventually converted into 8 bits

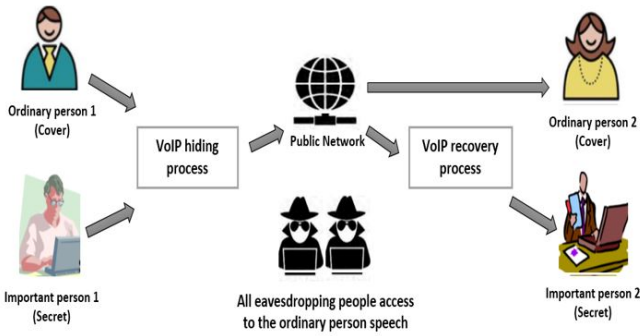

Figure 1. General scheme of the introduced real time covert communication system

in the output for both types. Since the VoIP network is a connection-less system and provides unreliable delivery service, which leads to loss or delay packets from the retrieved secret speech. Therefore, Internet Low Bit Rate Codec (iLBC) can be used as a VoIP source Codec to compress the secret input speech with auto-correction for speech quality degradation over a lossy channel.

iLBC is a free-narrow band speech codec that can be implemented with two transmission rates; 13.33 kbps and 15.20 kbps for an encoding packet every 30 ms and 20 ms, respectively. The input rate of iLBC is fixed and equal to 128 kbps, where the input signal to the encoder must be pulse code modulation speech signal sampled at 8 kHz with a resolution of 16 bits per sample. Therefore, about 10.4% (for 13.33 kbps mode) or 11.9% (for 15.20 kbps mode) compression rate compared to the original size of secret speech is achieved using a straightforward and fast process that is proportional to the real-time processing. The steps of the introduced embedding and recovery algorithms are demonstrated in the following sub-sections.

A. *Hiding Algorithm Stages*

The main stages of the hiding algorithm are shown in figure 3, include the following:

- An input cover speech (C) sampled at 32 kHz with a resolution of 16 bits is segmented into N segments without overlap. Each segment has 640 samples to realize synchronization between the cover and secret speech signals. Also, a sampled input secret speech (S) at 8 kHz and 16 bits of resolution is segmented without overlap into N segments of 160 samples per segment. A period for each cover segment (C_i) or secret segment (S_i) is 20 mile-second, where i= 1,2.…, N.
- By employing G.711 Codec, each cover segment is compressed to produce (C_{i_com}) with 512 samples, where each compressed sample is represented with 8 bits ($b_7b_6b_5b_4b_3b_2b_1b_0$). Also, each secret segment is compressed using iLBC Codec at a compression ratio of 11.9% to obtain (S_{i_com}) with 38 samples and a resolution of 8 bits/sample.
- According to the PRN output, the resulting compressed secret segment (S_{i_com}) is randomly permuted. The output of PRN depends on a specific secret key (SK) that is fed by the user.
- Subsequently, each secret sample (S_{ij_com}) is divided into four digits of two bits to obtain ($S_{ijd\_com\_p}$), where d is the digit order which may be 1, 2, 3, or 4.
- In the hiding stage, high energy cover samples from (C_{i_com}) are selected concerning a certain input threshold (Thr), where ($C_{ij} > Thr$) . So, only the selected cover samples, as seen in 1, will be employed for hiding the secret digit ($S_{ijd\_com\_p}$).

$$C_{ik\_Selected} = C_{ij} \tag{1}$$

where k=1,2,…N, N <= 640

- Now, to determine the first position of hiding from the LSB direction in $C_{ik\_Selected}$, HB is used, which has values from 0 to 3, where the depth of embedding is increased with increasing the value of HB. For example, if HB equals 2, the positions of embedding will be 3^{rd} and 4^{th} bits in $C_{ik\_Selected}$.
- After that, extracting the hold bit/s from the selected cover sample to obtain $C_{ik\_Selected\_HB}$ to be used for hiding the secret digit $S_{ijd\_com\_p}$. Equations 2 and 3 perform the hiding process. Figure 2 represents the flowchart of equations 2 and 3.

$$R = C_{ik\_Selected\_HB} \bmod 4 \tag{2}$$

$$\widehat{C}_{ik\_HB} = C_{ik\_Selected\_HB} - (R - S_{ijd\_com\_p}) \tag{3}$$

If the remainder (R) is similar to the determined secret digit ($S_{ijd\_com\_P}$), then the stego sample ($\hat{C}_{ik\_HB}$) is exactly the same as the candidate cover sample ($C_{ik\_Selected\_HB}$) without any change. Otherwise, there is a modification that has a maximum value of (3) in the stego-sample.

- Subsequently, combining the hold bit/s (in case of HB >0) with the $\hat{C}_{ik\_HB}$ to get $\hat{C}_{ik}$. Then, the resulting stego samples are rebuilt into one segment ($\hat{C}_i$ with a length of 640 samples) carries the secret segment (S_i with a length of 38 samples).
- Finally, the N stego-segments are combined into stego-data ($\hat{C}$), which is transmitted via the VoIP network.

Any person in VoIP network can access the stego data and apply G.711 decoder to hear a speech signal similar to the cover without perceptible difference compared to the original one.

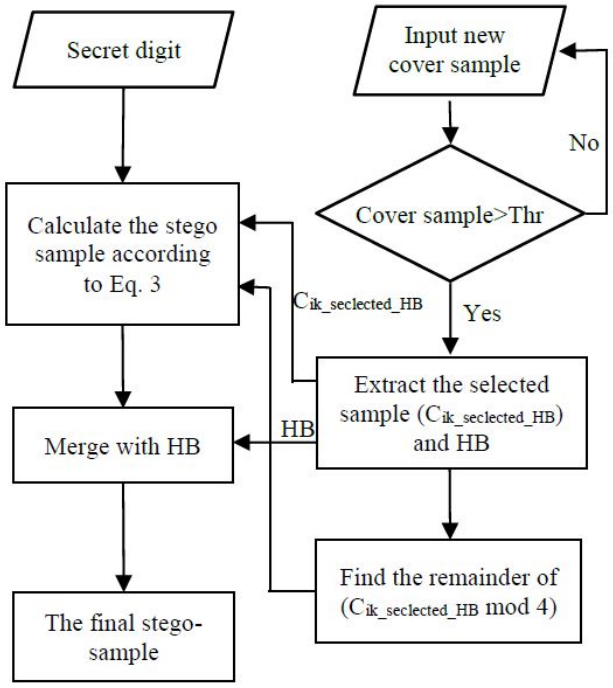

Figure 2. Flowchart of the embedding process for a single secret digit

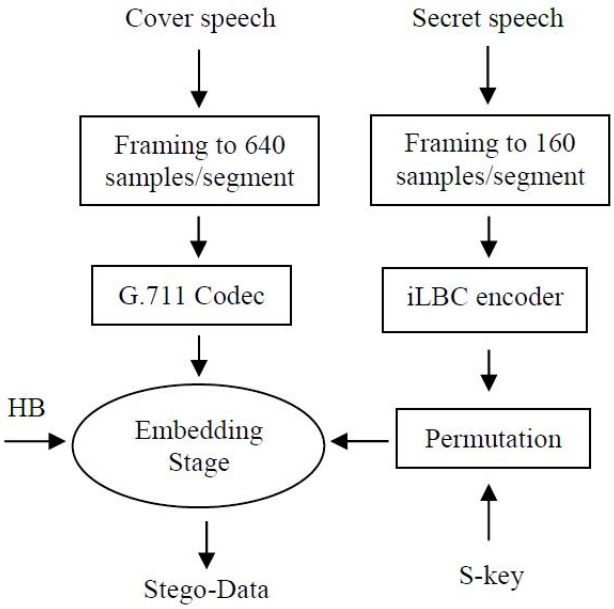

Figure 3. General block diagram for the embedding scheme of the proposed approach

B. Recovery Algorithm Stages

The main steps of the recovery algorithm are shown in figure 4, and explained as in the following:

- An input stego-data ($\hat{C}$) is segmented without overlap into N segments of 640 samples/segment.
- Selecting high energy stego samples from ($\hat{C}_i$) which are greater than Thr. Then the secret digits are sequentially extracted after removing the hold bit/s from each selected stego sample (in case of HB>0) to get $\hat{C}_{ik\_HB}$. The following formula is used to retrieve the secret data.

$$S_{ijd\_com\_p} = \hat{C}_{ik\_HB} \bmod 4 \tag{4}$$

- Successively, every four digits are combined into one secret sample, then rebuilding the segment i from the retrieved secret data ($S_{i\_com\_p}$)) with length of 38 samples.
- The inverse of the permutation that employed in the hiding process at the sender side is used to decrypt the resulted secret segment to obtain S_{i_com}.
- After that, decompress S_{i_com} using iLBC decoder to get the retrieved secret speech segment S_i with 160 samples/segment.
- Finally, the retrieved secret speech segments are sequentially combined into the final secret speech (S).

Retrieved secret speech is an error-free in the case of lossless channel due to use robust method has invertible operations without losing any information from the hidden data.

4. RESULTS AND DISCUSSION

This section describes the experimental results of the proposed scheme as well as a comparison with the related works. All the presented tests are based on employing a compressed cover speech of 256000 bps (512000 bps before the compression) to hide 128000 bps secret speech. We have achieved these tests for different real-time input cover speech signals with different real-time input secret speech signals. To evaluate the performance of this work, all the simulations have been done in MATLAB (2017a) environment.

A. Perceptual quality and embedding capacity tests

The perceptual quality or imperceptibility for the proposed approach can be measured mathematically by calculating signal to noise power ratio (SNR) using equation [15] 5 and perceptual evaluation of speech quality ($PESQ$).

$$SNR = 10 \times \log_{10} \frac{\sum_{i=1}^{n} C_i^2}{\sum_{i=1}^{n} (\hat{C}_i^2 - C_i^2)} \tag{5}$$

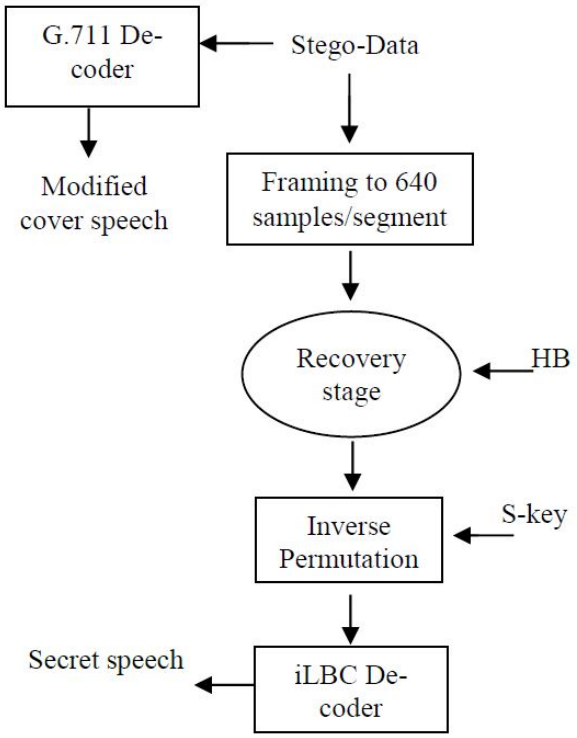

Figure 4. General block diagram for the recovery scheme of the proposed approach

Where n is the number of samples in cover (C) or stego ($\hat{C}$) speech.

Basically, the recommended value of SNR is 25 dB or more. According to the ITU-T recommendation P.862, the range of $PESQ$ is -0.5 to 4.5 which is considered as 1 instead of -0.5 and 5 instead of 4.5 for more convenience in the calculation where 1 indicates the worst sound quality [16].

To evaluate the perceptual quality for the proposed approach, several tests in terms of SNR and $PESQ$ are presented as shown in table I. However, the SNR value is measured between the compressed cover and stego data, while the values of $PESQ$ are evaluated between the original and decompressed cover speech. Three cover and secret speeches have been tested with threshold values, 64 and 128. The obtained results show that the proposed approach has an excellent perceptual quality for various input covers and secret speech. This is because the introduced embedding algorithm employs only the high-energy cover samples to embed the secret data. The cover speech quality is maintained by reducing the difference between the original and altered samples to minimum values. This yields a very high hiding capacity that is equal to 25% from the cover speech size, where it is calculated as the ratio of the compressed secret speech to the compressed cover speech. All the listed results in table I have been obtained with $HB = 0$. However, in the case of a noisy channel, it is necessary to increase the system robustness to avoid loss in the embedded data, this is achieved by increasing the value of HB factor, as will be presented in the next subsection.

B. Robustness tests against additive noise

To measure the similarity between the original and the retrieved secret speech, Normalized Cross-Correlation (NC) and Bit Error Rate (BER) are computed according to equation 6 [17], [18] and equation 7 [19], respectively.

$$NC = \frac{\sum_{i=1}^{m} S_i \times \hat{S}_i}{\sqrt{\sum_{i=1}^{m} S_i^2} \times \sqrt{\sum_{i=1}^{m} \hat{S}_i^2}} \tag{6}$$

$$BER = \sum_{i=1}^{k} \frac{b}{k} \times 100\% \tag{7}$$

Where: m is the total number of samples in the secret speech, k is the total number of bits in the secret speech, and $b = 0$ at $\hat{S}_i = S_i$, and $b = 1$ at $\hat{S}_i \neq S_i$

Table II shows several tests for the proposed system robustness in terms of NC and BER for both the embedding stage and the overall system (embedding and compression stages). These results show that the proposed system has good robustness with a lossless retrieved secret speech at normal case (without adding noise) for different input cover and secret signals. With adding noise, the system robustness is increased according to the embedding insertion depth that is controlled by the HB value. As seen in table III, the increase of HB value leads to an increase in the secret speech immunity against Adaptive Wight Gaussians Noise (AWGN) but simultaneously reduces the stego-speech quality at an acceptable range.

C. Perceptual quality in a noisy channel with different levels of packet losses

The quality of the proposed system has been tested in a lossy channel with different levels of packet losses according to the MATLAB model that is presented in figure 5. The packet losses can be occurred due to noise, routing, or/and latency. These losses lead to decreasing the quality of the transmitted data. In our system, we found that the increase in the percentage of packet loss from 2% to 10% will cause an acceptable degradation in the stego speech and hidden data quality, where the average values of SNR and NC are 35 dB and 0.88 respectively. Figure 6 shows the effecting of packet losses on the SNR, while figure 7 shows the effecting of packet losses on the NC. With percentage packet losses of more than 10%, the degradation in quality will appear clearly and cause an error in the retrieved secret speech.

D. System complexity

The system complexity is a critical issue for the real-time communications. In the proposed approach, there are two stages: compression and embedding. The compression

TABLE I. Perceptual quality tests

Secret signals	Cover signals	*Threshold*	*SNR*(*dB*)	*PESQ*
S_1	C_1	64	44.9076	4.12
		128	44.9504	4.17
	C_2	64	44.7308	4.01
		128	44.8176	4.10
	C_3	64	45.1437	4.22
		128	45.1516	4.26
S_2	C_1	64	44.9239	4.13
		128	44.9872	4.18
	C_2	64	44.7501	4.03
		128	44.8060	4.08
	C_3	64	45.1628	4.29
		128	45.2003	4.43
S_3	C_1	64	44.9017	4.11
		128	44.9308	4.15
	C_2	64	44.7591	4.04
		128	44.7927	4.06
	C_3	64	45.1620	4.28
		128	45.1701	4.30

TABLE II. Robustness tests

Secret signals	Cover signals	Overall system		Threshold	Embedding Stage	
		NC	*BER %*		*NC*	*BER %*
S_1	C_1	0.8801	0.0019	64	1	0
				128	1	0
	C_2			64	1	0
				128	1	0
	C_3			64	1	0
				128	1	0
S_2	C_1	0.9467	0.0012	64	1	0
				128	1	0
	C_2			64	1	0
				128	1	0
	C_3			64	1	0
				128	1	0
S_3	C_1	0.8744	0.0022	64	1	0
				128	1	0
	C_2			64	1	0
				128	1	0
	C_3			64	1	0
				128	1	0

TABLE III. Tests of secret speech immunity against AWGN for the introduced approach

HB	*SNR* (dB) for a stego-speech	*NC* for different AWGN in dB				
		50	40	30	20	10
0	42.47	0.985	0.936	0.898	0.822	0.741
1	39.52	0.987	0.962	0.925	0.825	0.772
2	36.37	0.991	0.977	0.936	0.857	0.801
3	32.61	1	0.983	0.959	0.915	0.841

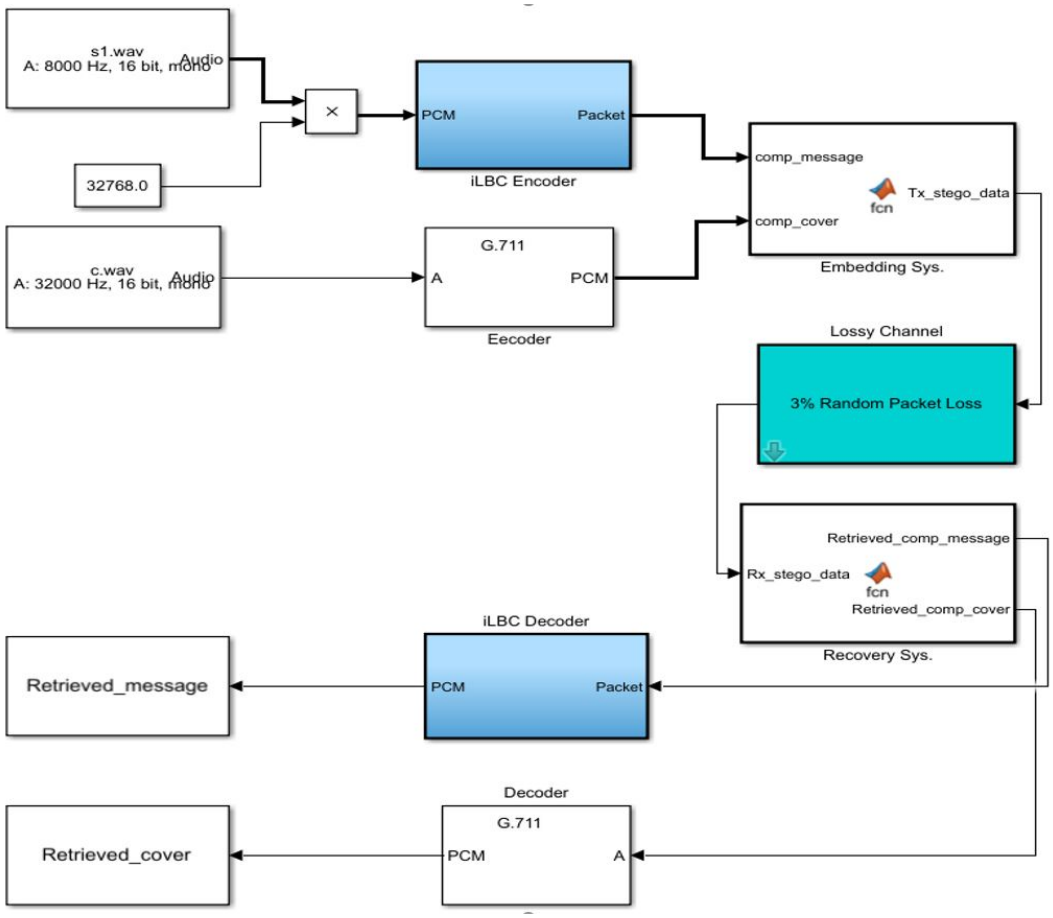

Figure 5. Simulation diagram to test the quality of the proposed system at a lossy channel with different packet loss rates

stage is a standard that already proves and meets real-time communications. To measure the complexity of the remainder stage (embedding stage) for a period of 20msec (it is the period of the processed speech in our system), we can compare our procedure with more complex other procedure such as Fast Fourier Transform (FFT) spectrum analyzer which operates in real time [20]. We find that it deals with complex operations. Traditionally, each complex addition or multiplication operation analogous to four times real addition or eight times real multiplication operation respectively. Our embedding algorithm deals with only real operations which equal to (1280) addition and (640) multiplication. Therefore, the complexity of our embedding stage is about 18.75% from the FFT complexity. This percentage confirms the suitability of the proposed system to be implemented in real time with low computational complexity and thereby low processing time within the human imperceptibility rate.

E. Comparison with related work

To confirm the effectiveness of the introduced approach, comparisons of the data embedding rate and quality performance between the proposed approach and other related works were conducted. However, most existing literature in VoIP steganography area is devoted on embedding offline

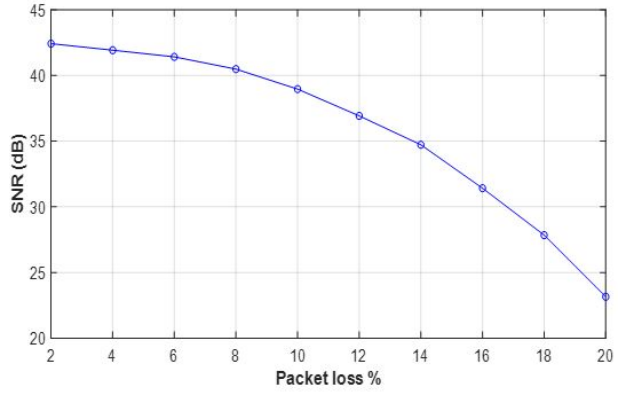

Figure 6. Effect of packet losses on the SNR

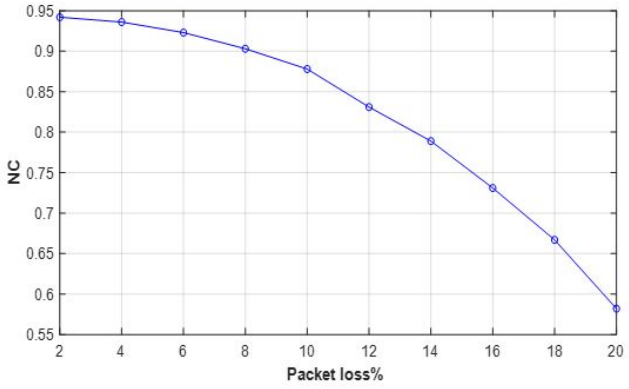

Figure 7. Effect of packet losses on the NC

secret data which may be text, image, or speech into real-time cover speech [21], [22]. Therefore, their embedding rates are unsuitable for hiding real-time secret speech. In this study, we hide 128kilo bits per second (as embedding rate) of real-time secret speech within real-time cover speech which is 53, 226, 32, and 59 times greater than [11], [12], [21], [23] respectively. As indicated in table IV, the comparison is based on averages values for SNR and $PESQ$ and embedding rate, where the proposed stego-data has an excellent perceptual quality that it is not less than 44.5dB of SNR and 4.16 of $PESQ$ as average values in the case of $HB = 0$.

In addition to the superiority of the suggested approach in terms of quality performance, it can also fully retrieve the secret speech in a lossless noisy channel. With adding noise (AWGN), the proposed approach has good or acceptable robustness due to employing higher embedding depth in hiding the confident data, whereas [5], [9], [21] and most of the existing techniques have used only LSB replacements to hide the secret data. Thus, these techniques are weak or very sensitive to additive noise. Also, several literatures embed the secret data before the compression stage [14], [15], [24], which makes it more vulnerable to damage or loss of the embedded secret data during the compression. Thus, there is a weakness in the robustness of their systems. To avoid this problem, we have been embedded the secret data in the compressed cover speech (after the compression process).

The proposed approach has a low computational complexity due to the utilization of a simple and fast embedding algorithm doesn't employ any complex operation as compared to [14]. This leads to consuming low processing time suitable for real time communications-based VoIP requirements.

5. CONCLUSION

In this study, we have proposed a real-time covert communication scheme based on public and cheap channel (VoIP) by embedding a real-time secret speech within a real-time cover speech. This is achieved via a secure channel that is created into the VoIP network. Two VoIP codecs are used which are G.711 and iLBC Codecs. In order to realize real-time communication bandwidth, G.711 encoder and decoder are employed to compress and decompress the cover speech and stego data respectively. While the secret speech is compressed and decompressed using iLBC encoder and decoder respectively. Thus, the hidden data size is reduced. The experimental results on speech quality show that the stego speech has excellent imperceptibility due to embedding the secret data at high-energy cover samples. This is achieved with an embedding algorithm that has low computational complexity and low processing time suitable for VoIP requirements. The embedding at higher positions (by using the HB factor) can improve the immunity of the embedded secret data against additive noise (AWGN). At a lossless channel, the proposed scheme is an error-free retrieved secret speech. Also, the security level is improved due to the utilization of random permutations adopted PRN and stego-key on secret data before the embedding. Moreover, a fixed high hiding rate of up to 128kbps (25% of the input cover speech size) with better perceptual quality is achieved.

REFERENCES

[1] Z. Wu, J. Guo, C. Zhang, and C. Li, "Steganography and steganalysis in voice over ip: A review," *Sensors*, vol. 21, no. 4, p. 1032, 2021.

[2] S. E. El-Khamy, N. O. Korany, and M. H. El-Sherif, "A security enhanced robust audio steganography algorithm for image hiding using sample comparison in discrete wavelet transform domain and rsa encryption," *Multimedia Tools and Applications*, vol. 76, no. 22, pp. 24 091–24 106, 2017.

[3] D. Tan, Y. Lu, X. Yan, and X. Wang, "A simple review of audio steganography," in *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*. IEEE, 2019, pp. 1409–1413.

[4] S. Deepikaa and R. Saravanan, "Voip steganography methods, a survey," *Cybern. Inf. Technol*, vol. 19, pp. 73–87, 2019.

[5] H. Tian, K. Zhou, H. Jiang, J. Liu, Y. Huang, and D. Feng, "An m-sequence based steganography model for voice over ip," in *2009*

TABLE IV. Comparative results of the perceptual quality and hiding rate with some related approaches

	Jiang [11]	Li [23]	Tang [21]	Lin [12]	The proposed scheme
SNR (dB)	38.7	Not reported	27.5	Not reported	44.9
$PESQ$	4.04	3.02	3.5	3.811	4.16
Hiding rate (bps)	8000	566	3968	2150	128000

IEEE International Conference on Communications. IEEE, 2009, pp. 1–5.

[6] H. Tian, K. Zhou, H. Jiang, Y. Huang, J. Liu, and D. Feng, "An adaptive steganography scheme for voice over ip," in *2009 IEEE International Symposium on Circuits and Systems*. IEEE, 2009, pp. 2922–2925.

[7] H. Tian, H. Jiang, K. Zhou, and D. Feng, "Adaptive partial-matching steganography for voice over ip using triple m sequences," *Computer Communications*, vol. 34, no. 18, pp. 2236–2247, 2011.

[8] M. H. Shirali-Shahreza and S. Shirali-Shahreza, "Real-time and mpeg-1 layer iii compression resistant steganography in speech," *IET Information security*, vol. 4, no. 1, pp. 1–7, 2010.

[9] Y. Jiang, L. Zhang, S. Tang, and Z. Zhou, "Real-time covert voip communications over smart grids by using aes-based audio steganography," in *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*. IEEE, 2013, pp. 2102–2107.

[10] Z. Wei, B. Zhao, B. Liu, J. Su, L. Xu, and E. Xu, "A novel steganography approach for voice over ip," *Journal of Ambient Intelligence and Humanized Computing*, vol. 5, no. 4, pp. 601–610, 2014.

[11] Y. Jiang and S. Tang, "An efficient and secure voip communication system with chaotic mapping and message digest," *Multimedia Systems*, vol. 24, no. 3, pp. 355–363, 2018.

[12] R.-S. Lin, "A synchronization scheme for hiding information in encoded bitstream of inactive speech signal." *J. Inf. Hiding Multim. Signal Process.*, vol. 7, no. 5, pp. 916–929, 2016.

[13] F. Li, B. Li, L. Peng, W. Chen, L. Zheng, and K. Xu, "A steganographic method based on high bit rates speech codec of g. 723.1," in *International Conference on Cloud Computing and Security*. Springer, 2018, pp. 312–322.

[14] P. M. Kumar and K. Srinivas, "Real time implementation of speech steganography," in *2019 International conference on smart systems and inventive technology (ICSSIT)*. IEEE, 2019, pp. 365–369.

[15] H. I. Shahadi, R. Jidin, and W. H. Way, "Lossless audio steganography based on lifting wavelet transform and dynamic stego key," *Indian Journal of Science and Technology*, vol. 7, no. 3, p. 323, 2014.

[16] I.-T. Recommendation, "Perceptual evaluation of speech quality (pesq): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs," *Rec. ITU-T P. 862*, 2001.

[17] H. I. Shahadi, R. Jidin, and W. H. Way, "Concurrent hardware architecture for dual-mode audio steganography processor-based fpga," *Computers & Electrical Engineering*, vol. 49, pp. 95–116, 2016.

[18] H. I. Shahadi, "Covert communication model for speech signals based on an indirect and adaptive encryption technique," *Computers & Electrical Engineering*, vol. 68, pp. 425–436, 2018.

[19] A. A. Krishnan, C. S. Chandran, S. Kamal, and M. Supriya, "Spread spectrum based encrypted audio steganographic system with improved security," in *2017 International Conference on Circuits, Controls, and Communications (CCUBE)*. IEEE, 2017, pp. 109–114.

[20] W. Lv, C. Shen, F. Gui, Z. Tian, and D. Jiang, "Real-time spectrum analyzer based on all phase fft spectrum analysis," in *2013 Fourth International Conference on Digital Manufacturing & Automation*. IEEE, 2013, pp. 966–969.

[21] S. Tang, Y. Jiang, L. Zhang, and Z. Zhou, "Audio steganography with aes for real-time covert voice over internet protocol communications," *Science China Information Sciences*, vol. 57, no. 3, pp. 1–14, 2014.

[22] Z. Wu, H. Cao, and D. Li, "An approach of steganography in g. 729 bitstream based on matrix coding and interleaving," *Chinese Journal of Electronics*, vol. 24, no. 1, pp. 157–165, 2015.

[23] F. Li, B. Li, Y. Huang, Y. Feng, L. Peng, and N. Zhou, "Research on covert communication channel based on modulation of common compressed speech codec," *Neural Computing and Applications*, pp. 1–14, 2020.

[24] S. S. Bharti, M. Gupta, and S. Agarwal, "A novel approach for audio steganography by processing of amplitudes and signs of secret audio separately," *Multimedia Tools and Applications*, vol. 78, no. 16, pp. 23 179–23 201, 2019.

Baneen Q. Abed Ali received her B.Sc in electrical and electronic engineering from the University of Kerbala (UOK), Kerbala, Iraq in 2015. Since then she has been working in the electrical engineering field in Kerbala university. Currenly, she is Msc. Student in the electrical and electronic engineering dept., university of Kerbala. Her research interest is communication security.

Haider I. Shahadi received his B.Sc degree in information engineering from the University of Baghdad, Iraq in 2001, his master's degree in Electronic and Communication Engineering from the University of Baghdad-Iraq in 2004, and his Ph.D. in Electronic and Communication Engineering from the Tenaga National University, Malaysia in 2014. Currently, he is a professor at the University of Kerbala, Iraq. His research interests include digital signal and multimedia processing, data security, FPGA design and implementation and embedded systems, IOT systems, and smart systems.

Muayad S. Kod received his B.Sc. and M.Sc. degrees from the Department of Electronics and Communications, Al-Nahrain University-Iraq, in 2002 and 2005, respectively. He received the Ph.D. degree from the Department of Electrical Engineering and Electronics, University of Liverpool-UK, in 2016. Currently, he is a Lecturer at the Department of Electrical Engineering and Electronics, University of Kerbala, Iraq. His current research interests include wireless power transfer and telemetry to implantable medical devices, wearable and implantable.

Hameed R. Farhan received his B.Sc., M.Sc. and Ph.D. degrees in electronic engineering from University of Technology, Baghdad, Iraq, in 1986, 2011, and 2018, respectively. He is currently a lecturer at the Electrical and Electronic Engineering Department, College of Engineering, University of Kerbala. His significant interests include Digital Electronics, DSP, Image Processing, Pattern Recognition, and Computer Vision