Kerbala Journal for Engineering Science

https://kjes.uokerbala.edu.iq/

A Review and Comparison for Audio Steganography Techniques Based on Voice over Internet Protocol

Baneen Qasem [a], Haider Ismael Shahadi [a]*, Muayad S. Kod [a], Hameed R. Farhan [a]

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

* Corresponding author: haider_almayaly@uokerbala.edu.iq

Received: 05 September 2021; Revised: 17 November 2021; Accepted: 21 November 2021

Abstract:

Cryptography and steganography are the major role approaches for covert communications. While cryptography enciphers the information in such a way to be non-understandable for unauthorized persons, steganography conceals secure information in ordinary cover data without making perceptible variations in cover data to make data attackers in illusion and they consider the transmitted data is just ordinary data. Therefore, steganography is considered superior compared to cryptography because the encrypted attract the attention of the data attackers. However, it requires bandwidth or saving storage much greater than the hidden secure data. Steganography can be used in different modes such as offline and online, also it can be utilized with different data cover such as image, video, audio, text, and data frames. This study focuses on the most challenging mode in steganography, which is online (i.e. real-time steganography). In addition, it focuses on the most important kind of cover media, audio signals, which should be processed in real-time. Accordingly, this study presents a focused review on data hidden in the Voice over Internet Protocol (VoIP) network where VoIP represents the standard of real-time communication-based on delivery voice over Internet. In addition to a brief survey for the existing real-time steganography approaches, this study also highlights the strengths and weaknesses of each approach. The rubric for assessing the approaches is based on some contradictory requirements for successful steganography approaches. The most important requirements include data embedding rate, imperceptibility, security, robustness, and algorithm complexity which specify the effectiveness of any steganographic

system. This study concludes that most of the existing approaches still lacked some features in one or more of the main requirements for real-time security-based steganography.

Keywords: Audio steganography, real-time covert communication, voice over internet protocol, audio compression, VoIP standards Codecs.

1. Introduction

Nowadays, shared data over the Internet is widely used. This sharing requires a secure manner, that is in continuous challenges and threats in the field of security and electronic criminals [1]-[3]. The traditional solution for this issue is considering the cryptographic system, which transforms the information into ambiguous data for unauthorized people [4]-[6]. However, data encryption gives a clear indicator to the attackers about how encrypted data is important for the transmitter person [7]. On the other hand, steganography is also an important data security approach that is recently spread. It embeds secure message in ordinary data (cover signal) without effect quality of the cover signal [8][9]. The efficient steganographic system doesn't make any suspicion about the hidden data inside cover signal [10]. Steganography can be categorized according to the type of the cover signals such as image, audio, and videos [11]-[13], or according to the type of processing into online, which requires processing in real-time, and offline processing [14][15]. There are many constraints for real-time processing according to the data type, the data compression type during communication, and its allowable perceptual time delay. The challenges in real-time steganography will increase in case of using an audio signal as a cover for hiding data. This is due to the sensitivity of the human auditory system (HAS),that can sense very small changes in audio quality [16][17]. Moreover, HAS can percept small delay periods in processing time that are equal to or above 30 ms for some people [18]. Another important challenge that is the standard voice communication protocols use data compression during send voice data, whether in mobile communication such as in Global System for Mobile communications (GSM), or Internet voice communication such as in VoIP. The compression directly reflects to cause problems in two important issues in steganography, robustness and embedding data rate. Some or most hidden data may be lost after data compression [19]. To increase the robustness, this requires embedding in higher energy locations in the cover signal as a result the embedding rate is extremely decreased.

This paper reviews the algorithms and schemes that have been proposed to achieve real-time steganography-based VoIP. VoIP provides global and free services, it is attached in several applications such as Skype and WhatsApp. So this study focuses on the most challenging mode in steganography, which requires real-time processing under the data compression condition. This study

also presents a brief and important comparison in terms of advantages and drawbacks for each reviewed scheme and how the scheme is suitable for real-time steganography-based VoIP, and which type of hidden data can be used as a secure message in order to be proper the requirements of real-time processing.

In order to reduce the effort from the reader, in Section Two, the main steganography requirements are defined and explained with specify some important performance indicators to show how these requirements are measured. In Section Three, the VoIP standard is overviewed and shows how it is operating for real-time communications. Section Four overviews the existing methods for real-time audio steganography-based VoIP under two categories, hiding secret data in cover before compression and hiding secret data in compressed cover. At the end of each category, a brief comparison beside our recommendation for what is the successful application for the reviewed schemes. Finally, Section five concludes the whole study in this paper.

2. Steganography requirements

For an efficient steganographic system, there are numerous important requirements that are used to measure the performance of any embedding algorithm, which are [20][7]: data embedding rate, imperceptibility, security, robustness, and algorithm complexity. These features are contradictory to each other, such that the increase in data embedding rate leads to degradation in the robustness of secret data and imperceptibility of stego-file [7][21]. Therefore, most researcher works tried to create a balancing trade-off between these requirements.

2.1. Data embedding rate

Data embedding rate or embedding capacity represents the number of the secret bits that can be embedded into a cover file per unit of time or a maximum amount of secret data that the cover file can carry it without influence on the stego file quality, it is measured by bit per second (bps). Also, it can be evaluated as a ratio of the secret data size to the size of cover data [20][22].

2.2. Imperceptibility

In imperceptibility or perceptual quality, the amount of changes in the cover file after the embedding process must be inaudible by the human ears. In other words, it refers to the amount of changes that the cover can withstand without effect on the stego data quality [23]. To measure this requirement, there are two ways, either by hearing the two signals (cover and stego) by depending on the evaluation of a number of persons whose have perfect hearing or by using performance

parameters as mathematical measurements which are Signal to Noise Ratio (SNR) as in Eq.1[24][25], Mean Opinion Score (MOS), Perceptual Evaluation of Speech Quality (PESQ) and etc.

$$SNR = 10 \times log_{10} \frac{\sum_{i=1}^{n} C^2}{\sum_{i=1}^{n} [\hat{C}_i^2 - C_i^2]} \quad \dots (1)$$

where: C refers to the original cover file, $\hat{C}$ refers to the modified cover file (stego file), and n indicates the number of samples in the cover or stego file.

MOS is an objective measurement method depend on a mathematical model that simulates the human auditory system with predefined values between 1 and 5, where 1 is expressed for the worst sound quality [19]. However, MOS is complex and high expensive to perform [26]. Therefore, PESQ score can be used, which is one of the subjective method to measure the sound quality [27] that depend on a group of people they have very good hearing (sometimes they name golden ears) [19].

2.3.Robustness

The resistance of stego-file to overcome different attacks or the capability of the steganographic system to retrieve the secret data from the stego-file that passed through a noisy channel with a minimum error [28][20]. The performance parameters use to measure the steganographic system robustness are Normalized Cross-Correlation (NCC) as in Eq.2 [7] and Bit Error Rate (BER) as in Eq.3[29].

$$NCC = \frac{\sum_{i=1}^{m} s_i \times \hat{s}_i}{\sqrt{\sum_{i=1}^{m} s_i^2} \times \sqrt{\sum_{i=1}^{m} \hat{s}_i^2}} \quad \dots (2)$$

$$BER = \frac{\sum_{i=1}^{m} \begin{cases} 0, & \hat{s}_i = s_i \\ 1, & \hat{s}_i = s_i \end{cases}}{m} \times 100\% \dots (3)$$

where: S is the original secret data, $\hat{S}$ is the recovered secret data, n is the total number of secret speech samples, and m is the total number of secret speech bits.

There are different types of attacks, the stego-system should be robust against them such as statistical analysis, brute-force, LSB removal, amplification, addition of noise, re-sampling, and compression [30][25][31].

2.4. Security

An embedded secret message should be undetectable by any eavesdropper on the network this refers to the security [32]. The steganographic algorithm complexity against various attacks or

the amount of time requires to destroy the hiding algorithm and extracting the secure data can be used to measure security.

In general, most stego-algorithm employ simple cryptography with secret or private keys in order to increase the system security [25][33][34]. However, the complex crypto-algorithm may effect to the processing time performance of the overall system. Therefore, the suitable crypto-system should be chosen carefully according to the application of the stego-system.

2.5. Algorithm Complexity

To meet real-time communication requirements, the complexity of the hiding algorithm is an important issue because it proportionally associates with the processing time which specifies whether the used algorithm is suitable for VoIP requirements or not [4]. Consequently, the hiding algorithm must be not complex and fast. This requirement can be evaluated by measuring the consumed time from the sender to the receiver (as a processing time).

3. VoIP Communication Model

VoIP or Internet protocol (IP) telephony is the most commonly used service on the Internet, it depends on a group of techniques and protocols to allow users to transmit voice or any multimedia content over the Internet protocol network (IP network) [35][36]. VoIP adopts on packet switching strategy instead of circuit switching mode (which adopts in traditional telephone networks) to deliver voice communications over the Internet [15]. The basic VoIP requirements are signaling, encoding, transport, and gateway control. The purpose of the signaling phase is to establish and manage the connection between two users [18]. After the conversion is established, the transmitted analog signals must be encoded into digital data using one type of VoIP standard codecs [35]. Then, based on the packet switching strategy, the compressed data are splitting into packets.

When signaling and encoding occur, the voice packets will move using two protocols: Real-time Transport Protocol (RTP) and Real-time Control Protocol (RTCP) [15]. After that, gateway control is involved to convert the voice packets traffic into IP packets to be travel to the receiving end. At the reception end, the received voice packets are combined and decoded to the original form (analog voice signals) [36]. Figure 1 shows a general structure for VoIP network [18]. There are several types of VoIP standard codecs such as Internet Low Bit Rate Codec (iLBC), Global System for Mobile communications (GSM), G.711, G.722, G.723.1, G.726, G.728, and G.729 [37] [38].

The main VoIP metrics are flexible, reliable and inexpensive due to the use of Internet protocols for delivering data instead of landlines with minimum time delay (latency) [39]. Also, it

processes voice packets in a separate manner where if one or two packets are dropped during the transmission, it does not affect the entire data [39]. However, the basic drawback of VoIP service is totally dependent on the Internet connection so, it must have high quality with sufficient bandwidth available [35].

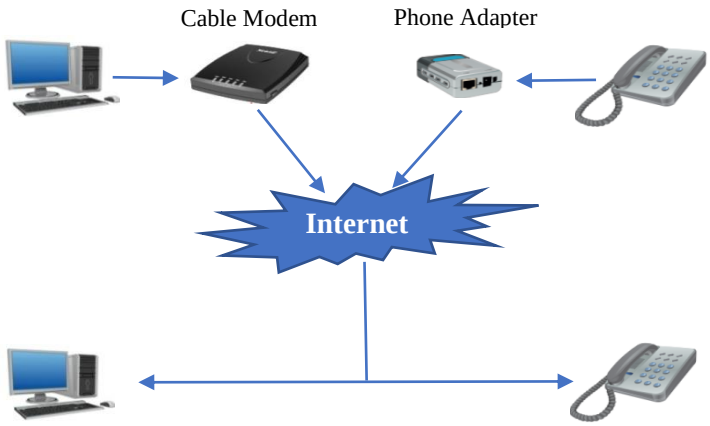

Figure 1 VoIP network structure [18].

4. Data Hiding in VoIP

Recently, a real-time covert communication based VoIP have attracted the attention of research community. However, it aims to create a secure communication channel for the transmitting secret data within VoIP network. As shown in Figure.2, the call between two persons in VoIP network can be used as a carrier for the secret data. The secret data may be text, image, speech, video or etc. in offline (previously saved or recorded) or in online (real-time). In the hiding approach, the secret data is embedded into the sender speech using an embedding algorithm which must meet real-time communication requirements. At the VoIP network, the stego data should not cause any suspicion for data attackers and also, they don't have any permission to access to the embedded secret data. After applying the recovery algorithm on the stego-speech, the receiver can retrieve the secret data.

Several hiding algorithms have been proposed in the literatures. These algorithms attempt to introduce new features to increase the steganographic system bandwidth, imperceptibility, robustness and/or security. To perform an embedding process based VoIP. We can classify these algorithms into two categories, performing the embedding process in uncompressed cover signal and in compressed cover signal. According to these categorizes, some of the existing real-time hiding techniques are reviewed in the following sub-sections.

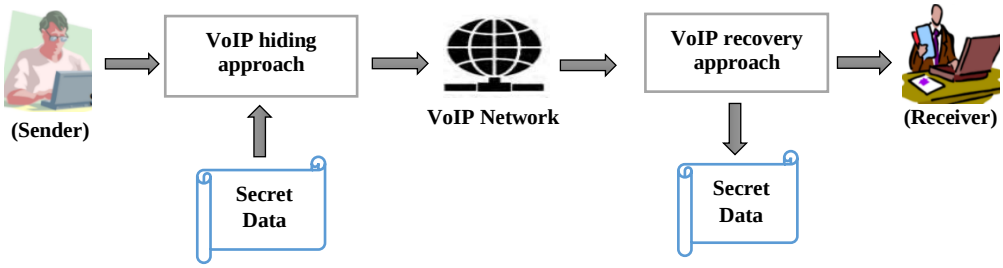

Figure 2 General structure for VoIP steganography [19].

4.1.First category

When the embedding stage is implemented before the compression stage of VoIP, the available hiding space (or embedding rate) in the ordinary cover data will be greater than the available hiding space in the compressed cover data. This is due to the fact that the amount of cover data before the compression is greater than the amount of compressed cover data. At the increasing of the embedding rate, the system robustness is degraded against codec compression. Therefore, the embedding algorithm must have significant robustness to face the compression process which supplies on the modified cover signal (cover signal with the hidden secret data) in VoIP channel and thereby saving the hidden secret data from lost. Also, the quality of the stego data must be acceptable or natural and does not cause any suspicion for eavesdroppers. Moreover, the processing time of the embedding algorithm should be proportional to the real-time communication requirements.

In [7], a lossless audio steganography scheme has been introduced based on integer lifting wavelet transform (Int2Int -LWT) and the LSB algorithm. It employs integer values to eliminate any error that may occur in the retrieved secret data in the rounding process using non-integer values.

A novel approach has been presented in [30] , it is based on an audio steganography in temporal domain. To perform the embedding process, the secret signal is split into an amplitude bits and sign bits. The amplitude bits are embedded into the four LSBs while the sign bits are embedded into the 5th to 8th bits of a cover signal. Research results show that the content of the audio signal does not get much affected if there are slight changes in amplitudes only, while slight changes in the ordering of the signs of amplitudes highly effect the content.

A real time steganography scheme was proposed [40], it based on revealing the silence intervals of uncompressed cover speech to embed the confident data. An embedding process is achieved by altering the number of the samples for the revealed silence intervals. The scheme has

efficient robustness to MPEG-1 layer III (MP3) compression and noise but the embedding rate is low. For real-time requirements, the proposed system is suitable for embedding off-line secret message into real-time cover signal.

In [41], speech steganography scheme based DWT transformation was presented to embed a secret message encrypted with pseudo-random sequences (PN) into the frequency space of real-time cover speech. This scheme has good performance and imperceptibility with low complexity due to use not complex equations.

Based on Int2Int- LWT algorithm, a novel audio steganography was introduced [42]. It used random positions for embedding the secret message blocks based on adaptive selection. To improve the imperceptibility, weighted block matching (WBM) process was employed to compute the similarity between each message block and cover blocks. Lossless massage retrieved because it used integer numbers (int2intHaar LWT) instead of conventional DWT thereby it canceled the approximation in data. The proposed algorithm cannot be used for real time communications because it has high complexity and consumes more processing time. Also, it needs scanning during matching process.

Datta et. al, presented an audio steganography scheme [43] based module operator to embed the secret data. The target string as secret data is converted into an equivalent hexadecimal where only four bits was taken at a time. This scheme offers high embedding capacity with minimum distortion in the recovered secret data.

In [44], introduced image in audio embedding approach in order to improve embedding capacity of the audio steganographic system. However, the secret data was split into blocks and based on chaotic technique, each block bits XORed with various random sequence to enhance security and robustness against different attacks.

A brief summary for the above reviewing literatures is shown in table.1.

Table 1 Summary for the selected literatures.

Reference	Techniques	Strengths	Weaknesses
[7]	Int2Int-LWT, LSB technique	Simple, fast, and has adequate security	It is not suitable for VoIP because the embedding before compress speech, so that the hidden speech in LSBs of cover may lose during the compression process in VoIP

[30]	Amplitude and sign bits, LSB technique.	- High data embedding rate about 16kbps. - Enhance security and quality MOS=4.8. - More robust to LSB, AWGN, and resampling attacks.	Less robust to additive noise.
[40]	Silence intervals	- High robustness to MP3 compression, and LSB attacks. - Enhance imperceptibility MOS=4.1 - Low complexity. - Real time cover speech.	- Low data embedding rate (32.1bps). - Secret data in offline. - Less robust to re-sampling attacks.
[41]	Spread spectrum representation, DWT.	- Enhance robustness (BER=0.0000001 without noise and BER=0.000452 in a noisy channel) - Low computational complexity.	- Higher storage capacity. - Unsuitable for real time communications. - An error may occur in the retrieved data.
[42]	Adaptive embedding positions, Int2Int–LWT, WBM	-Enhance each of robustness (NC=0.95 in a noisy channel) and perceptual quality (SNR=35db). - High data embedding rate equal to 300kbps.	Unsuitable for real time communications.
[43]	Module operator, hexadecimal.	- Less error in the recovered secret data. - Enhances robustness	It cannot be used for real time communications.
[44]	DWT, XOR, and chaotic map techniques.	Improve data embedding rate (320kbps) and security.	- Consumes more processing time. - Cannot be used for real time communications.

4.2. Second Category

To make use from the compression process provided by VoIP service, the secret data can be embedded into the compressed cover signal (compress with one type of VoIP standard Codec). Then, there will be a limitation in the embedding rate or embedding capacity so that any change in the compressed data will cause distortion in the decompression process. Therefore, the applications in this field is also limited by hiding simple static image, text, or offline speech. Furthermore, the resulted setgo data must have undoubted quality for any attacker on the network with an acceptable robustness against removal or distortion of an embedded secret data.

In literatures, paper [2] presented a real time hiding algorithm to create a secure communication channel into VoIP network. A Mix-excitation linear prediction (MELP) secret speech of low bitrate was embedded into online compressed cover speech (compress with G.729 Codec) by combining matrix coding and interleaving method. The proposed approach performed with an embedding capacity equal to 30 % of the compressed cover speech size.

In [4], a real time information hiding based VoIP scheme was introduced. The proposed scheme is suitable for embedding online secret speech into online cover speech. This scheme did not include any level of security and has less robustness to statistical analysis and additive noise. The

hiding process in this approach was implemented by using traditional LSB replacements to embed the compressed secret speech (compress with Speex Codec) into the LSBs of compressed cover speech (compress with G.711 Codec). For real time requirements, Skype, WhatsApp, Linphone can be used to implement this approach. Skype, WhatsApp, Linphone can be used to implement this approach.

Tang et. al proposed a VoIP steganography technique [14] to embed secret message encrypted using Advanced Encryption Standard (AES) method and symmetric key within the cover audio coded with PCM Codec. Different hiding location intervals for LSB insertions are used to obtain variable hiding capacities. However, the embedding rate is suitable for embedding offline secret information.

In [19], an adaptive VoIP audio stream was presented. It enhanced the security of traditional LSB algorithm which used to embed the secret message within VoIP audio streams by adopting three techniques: value-based multiple insertion (VAMI), voice damage offset (VODO) and voice activity detection dynamic insertion (VADDI). Also, it used G.711 as cover audio codec to evaluate the system efficiency. The introduced system has better transparent but the hiding rate is low around 102.28bps.

Shahadi et. al proposed a real time stegnography approach [34] based on embedding a real time compressed secret speech (compress with Int2Int –LWT algorithm) into a real time compressed cover speech (compressed with G.711 Codec) using module operator to ensue minimum error in the retrieved secret data. To maintain stego speech quality, the proposed approach adopted on high energy cover samples for embedding the secret data. It has high embedding capacity about 6kbps with effective robustness in noisy channel because it varies the depth of embedding based on hold factor (HF).

Two hybrid approaches based VoIP were introduced. One hides the encrypted secret message (encrypted with m-sequence encryption technique) into the compressed cover speech using least significant bits (LSBs) algorithm [45]. This algorithm provides acceptable 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 compresed cover speech and secret message to set proper value for the threshold PSV [46]. This approach can achieve good balance between steganographic transparency and embedding capacity. Both above hybrid schemes are less sensitive for additive noise due to use LSB technique.

Tian et. al in [47] extended [45][46] by proposing steganography scheme based on a comprehensive adaptive partial matching steganography to measure the similarity between 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 delay and realizing real time requirements, the encryption is integrated with the embedding process. The approaches [45],[46]and [47] use ITU-T G.729a as VoIP codec for cover speech.

A covert steganography system was suggested in [48] which divides the encrypted secret data (encrypt 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 use key distribution. However, the embedding bit rate is low between 0.5 and 8kbps. In this approach, the secret data may be offline string, image or speech while the cover speech is in real-time. Any one of VoIP applications can be used to implement this approach.

In [49], a novel embedding algorithm has been proposed based on revealing the inactive frames of low-bit rate audio streams which are more appropriate for hiding secret data than the active frames. To detect the inactive frames, a developed voice activity detection (VAD) technique was employed. This approach can obtain high embedding rate with acceptable distortion in the stego-speech quality.

Based on the characteristic of the speech codec, paper [50] proposed a novel technique by altering the codes of excitation pulse positions of G.723.1 speech codec to hide the sensitive information. This technique offers good security and efficiency as compared with existing similar works.

A brief summary for the above reviewing literatures is shown in Table 2.

Table 2 Summary for the selected literatures.

Reference	Technique	Strengths	Weaknesses
[2]	Matrix coding, interleaving method, Mix-excitation linear prediction (MELP), G.729.	- Better data embedding rate (2400bps). - Low complexity.	Offline secret data.
[4]	It used G.711 codec to compress the secret speech and then embedded it into cover speech using the LSB technique.	- It is suitable for real time VoIP communication. - High hiding capacity of about 128kbps.	- Less robustness to noise. - Does not contain any level of security.

[14]	LSB technique, AES encryption, PCM Codec,	- Enhance security. - Real time cover speech. - High embedding capacity of about 3968bps.	- Less robust to LSB attacks. - Simple offline secret data.
[19]	LSB technique, three approaches VAMI, VODO and VADDI, G.711.	Improve security and imperceptibility (MOS=4.3).	Low data embedding rate 102.28bps.
[34]	Int2Int-LWT, G.711, module operator.	-High embedding capacity (64kbps). -Full recovery of secret data.	Imperceptibility in a noisy channel needs to be improved.
[45]	LSB substitution, partial similarity value (PSV), M-sequence strategy, RSA algorithm, G.729a	Provide balancing between security, latency, and transparency (MOS=4).	Less immunity to additive noise.
[46]	LSB substitution, partial similarity value (PSV), G.729a.	Provides efficient performance with an acceptable balance between the hiding capacity and imperceptibility (MOS=4).	- Less immunity to additive noise. - Hiding capacity is limited.
[47]	An adaptive partial-matching strategy between the message bits and the LSBs of the cover, triple M-sequences, G.729a.	Better balance between bandwidth and imperceptibility (MOS=3.8).	- Less immunity to additive noise. - Hiding capacity is limited.
[48]	Chaotic maps, message digest, AES encryption, PCM.	- High resistance to statistical steganalysis. - Efficient stego data quality (SNR=38 and PESQ=4.5).	- Data embedding rate is not limited. - Embed simple secret data in offline.
[49]	Active frames, inactive frames, VAD, G.723.1.	High data embedding rate (6300bps) with acceptable imperceptibility (MOS=3.8)	Stego-speech quality needs to be modified.
[50]	Pulse code positions, G.723.1.	Offer better security and imperceptibility (PESQ=3.2).	Low data embedding rate of about 566bps.

5. Conclusion

Real-time communications based on VoIP play a major role in steganography essentially, with the increased utilization of VoIP applications. This is due to the fact that VoIP stream is a good carrier for transmitting the secret data in a secure manner where the eavesdroppers on the network do not have sufficient time to guess if there are embedded secret data in this stream or not. In this study, delivery speech about Internet protocols (VoIP technique) is discussed with a brief review on some of the existing real-time embedding techniques under two categories: embedding the secret data at the cover signal before the compression process provided by VoIP and embedding the secret data at the compressed cover data. The first category has high data embedding rate but the robustness needs to be adjustable to face the VoIP compression. In the second category, the compression provided by VoIP is exploited to embed the secret data but there is a limitation in the data embedding rate and applications. However, to extract the secret data with minimum error, the real-time hiding system must have an acceptable quality with good immunity against data attackers and noise. Moreover, the processing time of the hiding technique should be suitable for real-time requirements.

Therefore, uneasy for researchers to propose an embedding technique with more secure operations to hide and protect confidential data. The strengths and weaknesses for each reviewed work are presented which can be utilized for future enhancements. Also, how the reviewed work is suitable for real-time covert communication and what type of hidden secret data is appropriate for it all is explained. Furthermore, the most important contradictory requirements for an efficient steganographic system are offered with the performance measurements used to evaluate these requirements. The main finding for this study is that most of the existing works still lack one or more of the main real-time steganography requirements.

For future works, the embedding in the compressed cover audio signals must be covered by introducing an efficient and robust real time steganographic approach based on VoIP that verified high data embedding rate, acceptable perceptual quality, and high level of security.

References

[1] H. I. Shahadi and R. Jidin, High capacity and inaudibility audio steganography scheme, Proc. 2011 7th Int. Conf. Inf. Assur. Secur. IAS 2011, pp. 104–109, 2011.

[2] Zhijun W, Haijuan C, and Douzhe L, "An approach of steganography in G.729 bitstream based on matrix coding and interleaving," Chinese J. Electron., vol. 24 no.1, pp. 157–165, 2015.

[3] R. Setiawan and T. Ahmad, "Improving the Capacity of Data Hiding by Modifying the Interpolation of Audio Samples," 2020 8th Int. Conf. Inf. Commun. Technol. ICoICT 2020, pp. 7–12, 2020.

[4] C. Wang and Q. Wu, "Information hiding in real-time VoIP streams," Proc. - 9th IEEE Int. Symp. Multimedia, ISM 2007, pp. 255–262, 2007.

[5] A. H. Ali, M. R. Mokhtar, and L. E. George, "Recent approaches for VoIP steganography," Indian J. Sci. Technol., vol. 9, no. 38, 2016.

[6] N. Shetty, "Steganography for Secure Data Transmission," Int. J. Comput. Intell. Res., vol. 13, no. 10, pp. 2289–2295, 2017.

[7] Shahadi H I, Jidin R, and Way W H, "Lossless audio steganography based on lifting wavelet transform and dynamic stego key," Indian J. Sci. Technol., vol. 7 (3), pp. 323–334, 2014.

[8] Jayaram, Ranganatha, and Anupama, "Information Hiding Using Audio Steganography - A Survey," Int. J. Multimed. Its Appl., vol. 3, no. 3, pp. 86–96, 2011.

[9] A. Khan, A. Siddiqa, S. Munib, and S. A. Malik, "A recent survey of reversible watermarking techniques," Inf. Sci. (Ny)., vol. 279, no. June 2019, pp. 251–272, 2014.

[10] Shahad H I, Jidin R, and Way W H, "Concurrent hardware architecture for dual-mode audio steganography processor-based FPGA," Comput. Electr. Eng., vol. 49, pp. 95–116, 2016.

[11] D. E. Skopin, I. M. M. El-Emary, R. J. Rasras, and R. S. Diab, “Advanced algorithms in audio steganography for hiding human speech signal,” Proc. - 2nd IEEE Int. Conf. Adv. Comput. Control. ICACC 2010, vol. 3, pp. 29–32, 2010.

[12] S. Rekik, D. Guerchi, S. A. Selouani, and H. Hamam, “Speech steganography using wavelet and Fourier transforms,” Eurasip J. Audio, Speech, Music Process., vol. 2012, no. 1, pp. 1–14, 2012.

[13] M. A. Ahmed, L. M. Kiah, B. B. Zaidan, and A. A. Zaidan, “A novel embedding method to increase capacity and robustness of low-bit encoding audio steganography technique using noise gate software logic algorithm,” Journal of Applied Sciences, vol. 10, no. 1. pp. 59–64, 2010.

[14] Tang S Y, Jiang Y J, Zhang L P, and Z. B. Zhou, “Audio steganography with AES for real-time covert voice over internet protocol communications,” Sci. China Inf. Sci., vol. 57, no. 3, pp. 1–14, 2014.

[15] S. Deepikaa and R. Saravanan, “VoIP steganography methods, a survey,” Cybern. Inf. Technol., vol. 19, no. 1, pp. 73–87, 2019.

[16] El-Khamy S E, Korany N O, and El-Sherif M H, “A security enhanced robust audio steganography algorithm for image hiding using sample comparison in discrete wavelet transform domain and RSA encryption,” Multimed. Tools Appl., vol. 76, no. 22, pp. 24091–24106, 2016.

[17] Tan D, Lu Y, Yan X, and Wang X, “A simple review of audio steganography,” Proc. 2019 IEEE 3rd Inf. Technol. Networking, Electron. Autom. Control Conf. ITNEC 2019, no. Itnec, pp. 1409–1413, 2019.

[18] R. Roselinkiruba and R. Balakirshnan, “Secure steganography in audio using inactive frames of VoIP streams,” 2013 IEEE Conf. Inf. Commun. Technol. ICT 2013, no. Ict, pp. 491–495, 2013.

[19] Z. Wei, B. Zhao, B. Liu, J. Su, L. Xu, and E. Xu, “A novel steganography approach for voice over IP,” J. Ambient Intell. Humaniz. Comput., vol. 5, no. 4, pp. 601–610, 2014.

[20] N. Tiwari, D. Madhu, and D. Meenu, “Spatial Domain Image Steganography based on Security and Randomization,” Int. J. Adv. Comput. Sci. Appl., vol. 5, no. 1, pp. 156–159, 2014.

[21] H. Shahadi and R. Jidin, “High Capacity and Resistance to Additive Noise Audio Steganography Algorithm,” IJCSI Int. J. Comput. Sci. Issues, vol. 8, no. 5, pp. 176–184, 2011.

[22] S. Shirali-Shahreza and M. Shirali-Shahreza, “Steganography in silence intervals of speech,” Proc. - 2008 4th Int. Conf. Intell. Inf. Hiding Multimed. Signal Process. IIH-MSP 2008, pp.

605–607, 2008.

[23] M. H. N. Azam, F. Ridzuan, M. N. S. M. Sayuti, and A. A. Alsabhany, “Balancing the Trade-Off between Capacity and Imperceptibility for Least Significant Bit Audio Steganography Method: A New Parameter,” 2019 IEEE Conf. Appl. Inf. Netw. Secur. AINS 2019, pp. 48–53, 2019.

[24] M. H. A. Al-Hooti, T. Ahmad, and S. Djanali, “Audio Data Hiding Using Octal Modulus Function Based Unsigned Integer Sample Values,” 2018 Int. Conf. Comput. Eng. Netw. Intell. Multimedia, CENIM 2018 - Proceeding, pp. 48–53, 2018.

[25] A. H. Ali, L. E. George, A. A. Zaidan, and M. R. Mokhtar, “High capacity, transparent and secure audio steganography model based on fractal coding and chaotic map in temporal domain,” Multimed. Tools Appl., vol. 77, no. 23, pp. 31487–31516, 2018.

[26] Z. Wu and W. Yang, “G . 711-Based Adaptive Speech Information Hiding,” pp. 1139–1144, 2006.

[27] N. Aoki, “A semi-lossless steganography technique for G.711 telephony speech,” Proc. Sixth Int. Conf. Intell. Inf. Hiding Multimed. Signal Process. Darmstadt, Ger., pp. 534–537, 2010.

[28] A. H. Ali, M. R. Mokhtar, and L. E. George, “Enhancing the hiding capacity of audio steganography based on block mapping,” J. Theor. Appl. Inf. Technol., vol. 95, no. 7, pp. 1441–1448, 2017.

[29] Anjana K A, Satheesh C C, Kamal S, and Supriya M H, “Spread spectrum based encrypted audio steganographic system with improved security,” Proceeding Second Int. Conf. Circuits, Control. Commun. IEEE, pp. 109–114, 2017.

[30] S. S. Bharti, M. Gupta, and S. Agarwal, “A novel approach for audio steganography by processing of amplitudes and signs of secret audio separately,” Multimed. Tools Appl., vol. 78, no. 16, pp. 23179–23201, 2019.

[31] S. Gupta and N. Dhanda, “Audio Steganography Using Discrete Wavelet Transformation (DWT) & Discrete Cosine Transformation (DCT),” IOSR J. Comput. Eng., vol. 17, no. 2, pp. 2278–661, 2015.

[32] S. Hemalatha, U. D. Acharya, and A. Renuka, “Wavelet transform based steganography technique to hide audio signals in image,” Procedia Comput. Sci., vol. 47, no. C, pp. 272–281, 2015.

[33] S. S. Verma, “A Novel Technique for Data Hiding in Audio Carrier by Using Sample Comparison in DWT Domain,” pp. 0–4, 2014.

[34] H. I. Shahadi, M. S. Kod, B. Qasem, and R. Hameed, “Real-Time Scheme for Covert Communication Based VoIP Real-Time Scheme for Covert Communication Based VoIP,”

2021.

[35] Z. Wu, J. Guo, C. Zhang, and C. Li, “Steganography and steganalysis in voice over ip: A review,” Sensors (Switzerland), vol. 21, no. 4, pp. 1–31, 2021.

[36] W. Mazurczyk, “VoIP steganography and its detection-a survey,” ACM Comput. Surv., vol. 46, no. 2, 2013.

[37] F. M. Guo, A. Talevski, and E. Chang, “Voice over Internet protocol on mobile devices,” Proc. - 6th IEEE/ACIS Int. Conf. Comput. Inf. Sci. ICIS 2007; 1st IEEE/ACIS Int. Work. e-Activity, IWEA 2007, no. Icis, pp. 163–168, 2007.

[38] G. S. Miliefsky, “Securing Voice Over Internet Protocol (VoIP),” Hakin9 Pract. Prot., vol. 90, no. 9, p. 51, 2010.

[39] M. DeSantis, “Understanding Voice over Internet Protocol (VoIP),” US Cert, pp. 1–5, 2008.

[40] Shirali-Shahreza M H and Shirali-Shahreza S, “Real-time and MPEG-1 layer III compression resistant steganography in speech,” IET Inf. Secur., vol. 4, no. 1, pp. 1–7, 2010.

[41] P. M. Kumar and K. Srinivas, “Real Time Implementation of Speech Steganography,” Proc. 2nd Int. Conf. Smart Syst. Inven. Technol. ICSSIT 2019, vol. 23, no. 1, pp. 365–369, 2019.

[42] H. I. Shahadi, R. Jidin, and W. H. Way, “A novel and high capacity audio steganography algorithm based on adaptive data embedding positions,” Res. J. Appl. Sci. Eng. Technol., vol. 7, no. 11, pp. 2311–2323, 2014.

[43] B. Datta and S. Tat, “Robust High Capacity Audio Steganography using Modulo Operator,” 2015 IEEE, 2015.

[44] S. E. El-Khamy, N. O. Korany, and M. H. El-Sherif, “Robust image hiding in audio based on integer wavelet transform and Chaotic maps hopping,” Natl. Radio Sci. Conf. NRSC, Proc., no. Nrsc, pp. 205–212, 2017.

[45] H. Tian, K. Zhou, H. Jiang, J. Liu, Y. Huang, and D. Feng, “An M-sequence based steganography model for voice over IP,” IEEE Int. Conf. Commun., 2009.

[46] H. Tian, K. Zhou, H. Jiang, Y. Huang, J. Liu, and D. Feng, “An adaptive steganography scheme for voice over IP,” Proc. - IEEE Int. Symp. Circuits Syst., pp. 2922–2925, 2009.

[47] H. Tian, H. Jiang, K. Zhou, and D. Feng, “Adaptive partial-matching steganography for voice over IP using triple M sequences,” Comput. Commun., vol. 34, no. 18, pp. 2236–2247, 2011.

[48] Y. Jiang and S. Tang, “An efficient and secure VoIP communication system with chaotic mapping and message digest,” Multimed. Syst., vol. 24, no. 3, pp. 355–363, 2018.

[49] R. S. Lin, “A synchronization scheme for hiding information in encoded bitstream of inactive speech signal,” J. Inf. Hiding Multimed. Signal Process., vol. 7, no. 5, pp. 916–929, 2016.

[50] 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, vol. 11068 LNCS. 2018.

مراجعة ومقارنة تقنيات إخفاء الصوت على أساس بروتوكول الصوت عبر الإنترنت

الخلاصة: *يعد التشفير وإخفاء المعلومات من الطرق الرئيسية التي تلعب دورًا مهمًا في الاتصالات السرية. بينما يقوم التشفير بتشفير المعلومات بطريقة تجعلها غير مفهومة للأشخاص غير المصرح لهم ، فإن إخفاء المعلومات يخفي المعلومات الآمنة في بيانات الغلاف العادية دون إجراء اختلافات ملحوظة في بيانات الغلاف لجعل مهاجمي البيانات في وهم وهم يعتبرون البيانات المرسلة مجرد بيانات عادية. لذلك ، يعتبر علم إخفاء المعلومات أفضل مقارنة بالتشفير لأن المعلومات المشفرة تجذب انتباه مهاجمي البيانات. ومع ذلك ، فإنه يتطلب حزمة أتصال أو توفير مساحة تخزين أكبر بكثير من البيانات الآمنة المخفية. يمكن استخدام أخفاء البيانات في أوضاع مختلفة منها غير متصل بالإنترنت ومتصل عبر الإنترنت، كما يمكن استخدامها مع أغطية بيانات مختلفة مثل الصور والفيديو والصوت والنصوص المكتوبة وغيرها. تركز هذه الدراسة على الوضع الأكثر تحديًا في إخفاء المعلومات أو إخفاء المعلومات عبر الإنترنت أو في الوقت الفعلي. كذلك، يركز فقط على أهم أنواع أنواع بيانات الغلاف التي تتطلب معالجة في الوقت الفعلي ، وهي الإشارات الصوتية. وفقًا لذلك ، تقدم هذه الدراسة مراجعة مركزة على إخفاء المعلومات الذي يتعامل مع بروتوكول نقل الصوت عبر الإنترنت (VoIP) الذي يمثل معيار الصوت القائم على الاتصال في الوقت الفعلي عبر الإنترنت. بالإضافة إلى مسح موجز لأساليب إخفاء المعلومات في الوقت الفعلي الحالية ، تسلط هذه الدراسة الضوء أيضًا على نقاط القوة والضعف في كل نهج. يعتمد نموذج تقييم الأساليب على بعض المتطلبات المتناقضة لنُهج إخفاء المعلومات الناجحة. تشمل المتطلبات الأكثر أهمية معدل تضمين البيانات ، وعدم الإدراك ، والأمان ، والمتانة ، وتعقيد الخوارزمية التي تحدد فعالية أي نظام إخفاء المعلومات. خلصت هذه الدراسة إلى أن معظم الأساليب الحالية لا تزال تفتقر إلى بعض الميزات في واحد أو أكثر من المتطلبات الرئيسية لإخفاء المعلومات القائمة على الأمان في وقت الأتصال الفعلي.*