BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **19**, No 1
Sofia • 2019 Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.2478/cait-2019-0004

VoIP Steganography Methods, a Survey

S. Deepikaa, R. Saravanan

Vellore Institute of Technology, Vellore, India
E-mails: deepikaa.s@vit.ac.in rsaravanan@vit.ac.in

***Abstract*:** *Achieving secured data transmission is not always an easy job. Secret data sharing requires confidentiality and Undetectability. Steganography is preferred than cryptography to achieve undetectability. Steganography hides the secret data inside the other file such as text, audio, video, so that the existence of the secret data is completely hidden. Recent research focuses much on utilizing Voice over Internet Protocol* (*VoIP*) *calls as a carrier for data hiding. VoIP calls are much preferred among internet users for its wide availability, dynamic time limit and low cost. This paper focuses on data hiding methods that uses VoIP as a carrier. The paper also analyzes the performance of the algorithms using the three metrics undetectability, bandwidth and robustness.*

***Keywords*:** *VoIP, LSB, PCM, QIM.*

1. Introduction

Steganography is the art of hiding the secret data in other files such as image, audio, video as a cover. Network Steganography is the advanced version of Steganography that uses network packets and its protocols headers as a cover for data hiding. The next version of network Steganography is Voice over Internet Protocol (VoIP) Steganography. The VoIP calls are utilized as the cover to hide the secret data. VoIP, is the technology which is most preferred by the internet users than to traditional telephonic communication, for its wide availability of internet and low cost. The dynamic nature of the communication gives less chance to detect the presence of secret data. And also seems to be challenging on data embedding.

Fig. 1 depicts the VoIP communication model where the S is the secret data to be hidden and P represents the voice packets which acts as a cover. Steg is the output of the embedding process; it carries voice and secret data. Em is the embedding function that hides the Secret data (S) inside the voice Packets (P) by using any data hiding algorithm. Ex is the reverse process, where the secret data is extracted from the Steg.

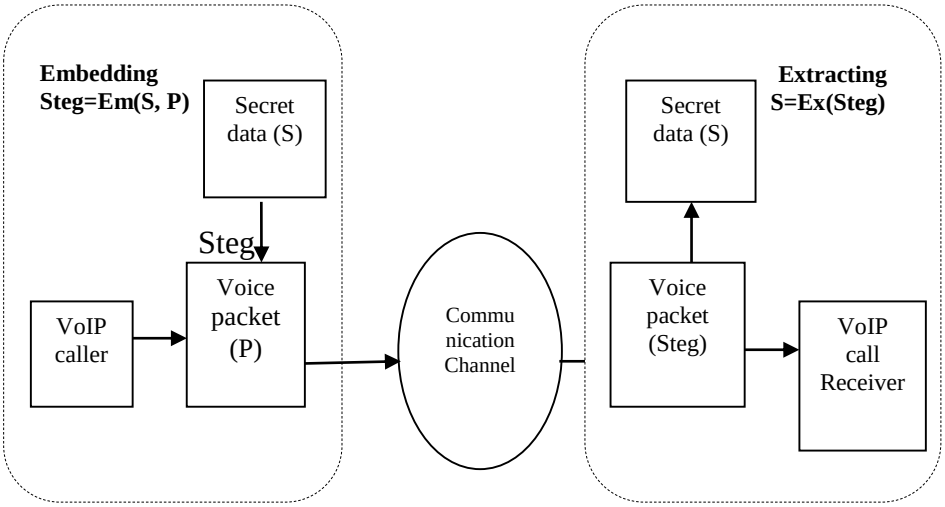

Fig. 1. VoIP Communication model

The secret data is extracted from the cover speech packet (steg). The original speech cover cannot be separated from the secret data, as the conversation is dynamic and does not affect the human auditory system. The conversation need not be recorded for any future purpose. VoIP steganography prefers a blind data hiding scheme. Generally there are two schemes such as blind and non blind data hiding. Blind data hiding does not require the original cover for extraction, but the non blind data hiding requires the original cover for extraction. This type of non blind data hiding is commonly used in the watermarking techniques to hide the digital signatures in the logo for authentication and added security. VoIP prefers the blind data hiding which does not require original speech samples. More over some data hiding methods perform the embedding process before converting them into packets during encoding and decoding also. So it is impossible to regenerate the speech sample from modified cover.

Some bandwidth will be wasted if original cover has to be sent to the receiver for extraction.

The main factors that measures the efficiency of the method is its bandwidth, Undetectability and robustness. Bandwidth refers to the capacity, the amount of secret data hidden per packet or unit time. Undetectability refers to the security, whether the presence of secret data is revealed or extracted. Robustness refers to the amount of changes the cover can withstand without destroying the secret data.

Steganography helps us to hide the secret data in another cover medium such as text, audio, video. The stego cover file can be transferred from the sender to receiver through any form of digital transmission. There are many algorithms and methods proposed that hide the secret data. But the worst part is that, these methods are considered as static and the cover file can be captured and on application of various steganalysis algorithms available, the presence of the secret data is identified and may lead to the extraction of secret message.

Steganalysis is the method of extracting the secret data or identifying its presence or its related information that helps to extract the data. There are numerous

methods available to identify the presence of the secret message and extract them. For example, changes in the statistical information reveal the presence of secret data in the payload.

Network Steganography is the attractive research topic in recent days. The reason is that it provides high bandwidth when the data is embedded in the data packets depending on its length. It also allows us to use unused header fields for modification that facilitates data embedding and maintaining the secrecy. But again analysis of the network properties and on constant monitoring of network packets and comparison of the packet data and its header information at different checkpoints may reveal the presence of the secret data, Thereby application of steganalysis method leads to the extraction of secret message.

2. Overview of the existing methods

The paper gives a brief description of the following topics respectively, such as mechanism of VoIP, Data hiding methods, Analysis of various data hiding methods. This paper is an extension work done on the motivation of the survey work done by Mazurczyk [4].

2.1. Mechanism of VoIP

VoIP stands for Voice over Internet Protocol; VoIP does not connect the end users using the traditional circuit switching methods. The end users are assumed to have a good internet connection – either wired or wireless. The VoIP uses the packet switching methodology where the analog voice is coded to digital data. The digital data is then splitted into audio packets and these audio packets are sent to the receiver, where the audio packets are again assembled and decoded to human voice that reaches the end user.

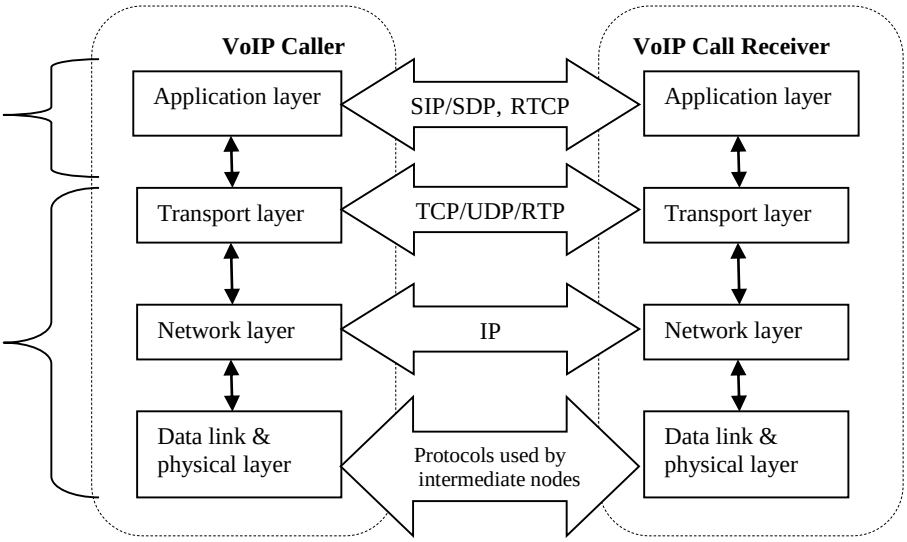

Fig. 2. VoIP Protocol stack

The VoIP call is made in two phases. Initially when the call is made, the connection is established between the two users with the signaling phase. The protocol, which handles signaling phase, is SIP (Session Initiation Protocol). Next phase is conversation phase, both users communicate among themselves either through audio or video calls. The voice data will be exchanged in terms of packets. Conversion phase is handled by RTP (Real Time Protocol) and RTCP (Real Time Control Protocol). Fig. 2 describes the various protocols used at each layer. This helps us find the possibilities of identifying cover position at each layer.

2.2. Data hiding methods in VoIP

The digital audio steganography were broadly classified into three domains [29], such as temporal, transform and coded domains. Temporal domain deals with time domain features, say audio speech recorded at different time intervals. The commonly used method would be LSB substitution, since the LSB features could be easy to extract and process, many authors prefer it. Transform domain deals with the frequency features such as spread spectrum. The frequency domain features proved more resistant to steganalysis than the temporal features in many cases. The coded domain deals with the encoder and decoder features of the codecs chosen. Either the data would be hidden in the speech and music using subband amplitude modulation or the data would be hidden in the bitstreams during the encoding.

Mazurczyk [4] have given a brief discussion on many VoIP Steganography techniques. The survey helped to understand the fundamental concepts and many pre-existing methods from 2003-2012. According to his survey, there were three main possibilities to hide the secret data in a VoIP packet; the first and commonly used method was to embed the secret data in the voice payload or to hide the data in the redundant fields or unused fields of the header. The other popular method was to modify the payload's time relations and embed the secret message. There were methods that used both (hybrid) the packets time relations and payload together to achieve higher efficiency. Search results showed that, many authors preferred payload modification for data hiding than the time relations. The main advantage in payload modification was large embedding capacity in voice payload packets than the other methods. Practical implementation was carried out easier with the payload packets than the other methods. Unused header modification again required a lot of changes in the header like CRC and various other related fields and also embedding capacity was less comparatively.

Mazurczyk, Szaga and Szczypiorski [5] in 2012 also proposed *Transcoding steganography*, in which the author used low bit codec to perform compression of cover data in the voice payload packet and left the voice payload size unaltered after compression, and change of codec was intentionally left unmodified. So that the reduced space in the voice payload was intended to embed the secret data. Data embedding does not have any restrictions to any particular algorithm, the low bit codec was chosen to provide nearer voice quality with reduced payload size. Janicki, Mazurczyk and Szczypiorski [6] in 2014 analyzed the various possibilities of detecting the presence of secret data and has done experimental results to show the codecs that can withstand the detection. The author extended the work

[8] to analyze the covert and overt codec pairs under several classifiers to indentify the most undetectable pairs. The author also discussed about the selection of the codecs [28] based on the various parameters to achieve high performance in terms of cost and bandwidth.

From the classifications observed from [29] and [4]. Fig. 3 describes the application of these classifications with respect to VoIP.

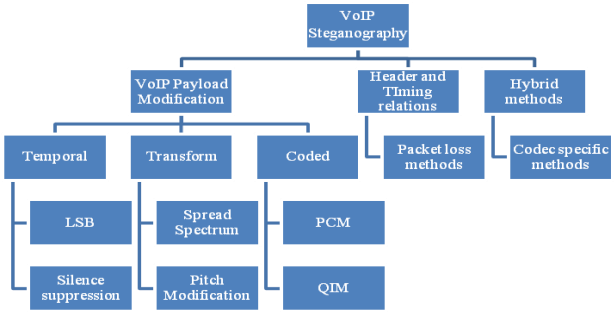

Fig. 3. Classification of VoIP Steganography methods

2.2.1. Least Significant Bit (LSB) modification methods

There has been many methods proposed under LSB modification method, hence it is considered the most commonly used substitution method in the field of Steganography (audio, video). Many authors have applied the LSB Modification in the VoIP packets for embedding secret data in the voice payload packets. At the sender side, the least significant bit of each sample will be usually substituted with the secret data bit, the process continues till all the secret data bits are inserted. The modified packet will be sent to the receiver without or with encryption for added security. At the receiver side, the secret data bits will be extracted from the LSB. The major advantage of LSB method was the high embedding capacity. The major disadvantage was the security. Many steganalysis techniques have been identified to detect the secret data. Statistical steganalysis are more proven to detect the presence of secret data and extract the secret information.

Many authors have proposed many new methods with the LSB, The difference were found in the selection of LSB bits to be embedded with secret data bit, as all the LSB bits do not carry secret bit. if each and every LSB bit carries a secret message the attacker can easily detect the presence of the secret message just by extracting the LSB bits and assemble it together. The difference in the embedding process lies in selection of LSB bits for modification and how the embedding is carried out. Some authors also combined existing methods with their proposed method to yield high bandwidth and security. There are some methods that can withstand the traditional steganalysis techniques and proved its security. Some works are highlighted here which work on different mechanisms for selecting the positions for embedding and some hybrid methods are also highlighted.

H u i et al. [2] in 2011 proposed two schemes, “Adaptive Partial Matching Steganography”, in which the Partial Similarity Value (PSV) was identified between

the covers and secret message was encrypted for added security. Two threshold values were set to decide the embedding process and a flag bit for each LSB part was set to recognize the cover that needs to be embedded. An improved APMS scheme was proposed in the same paper where the PSV was calculated between the covers and original messages with two flag bits, to know if the corresponding cover part was used for substitution and whether the secret message was encrypted or not. Comparing both schemes the latter one achieved higher Steganographic bandwidth and good transparency but increased the complexity of the algorithm.

Qin et al. [12] in 2015 worked on embedding the secret data by using random binary matrix of size 3×4. The scheme implementation took 3 bits of secret message in 4 bits cover with not more than 2 bits changed. The number of bits that needed modification was identified by using a Distortion function. The function helped to identify the cover bits with distortion value not exceeding 2 bits. The random binary matrix embedding scheme proved to have good embedding transparency and capacity.

Table 1. Summary of methods on Temporal domain

VoIP Methods	Technique	Metrics improved		
		Undetectability	Bandwidth	Robustness
LSB Modification	Random binary matrix embedding		✓	
	Partial similarity value		✓	
	PSV and matrix embedding		✓	
	Adjustable guide matrix		✓	
	Multi dimensional spatial model with retransmission for packet loss.	✓	✓	
	Chaotic maps and message digest	✓		
	Linear Predictive Encoding with matrix embedding	✓	✓	
	Adaptive bitrate modulation and hamming matrix encoding strategy	✓		
Silence suppression methods	Introduce fake silence packets with secret messages	✓	✓	

Tian et al. [10] in 2015 proposed an “Improved adaptive partial matching Steganography for VoIP” which was an extended work of the APMS discussed prior. This method proved to be a more effective approach than the previous one as it balances the Steganographic transparency and bandwidth. The prior version performed embedding depending on the PSV only. The new method performed embedding which was guided with increased unequal probabilities instead of identical probabilities. The author also made a good attempt to make use of covers considered unsuitable. “Matrix Encoding strategy” is applied with not more than one bit changed. This helped the author to achieve higher bandwidth than the prior work. Last step of this method uses two flag bits with one more encrypted form of secret message.

Tian et al. [11] in 2015 worked on embedding the secret data into audio packets by dividing them into small parts of different length. An adjustable guide matrix was generated to fit the variable cover length. The author has gained the optimal embedding performance compared to the existing methods.

Jiang and Tang [19] in 2017 have proposed a method to embed the secret message using chaotic maps and message digest. Message digest was employed to encrypt the secret data using block cipher technique, which is then embedded in the voice packets using chaotic maps. The chaotic maps were designed using the logistic mapping using an equation. Chaotic maps stay resistant to steganalysis, since the voice packets to be embedded are chosen randomly. The results proved to withstand statistical steganalysis.

Liu, Li and Wang [21] in 2017 used the linear predictive encoding method with matrix embedding technique to improve the embedding efficiency with better security. The author used the LPC, four methods to generate the codewords based on minimum distance of linear predictive coefficient vectors. The position, template for embedding and cover portion to be embedded was selected with the help of private key. The selected speech data frames were partially embedded and the code words for embedding was obtained from mapping table and implemented by matrix embedding technique. The selected code words were replaced with the cover to obtain the steganogram. The paper provides better security with less speech distortion and better performance.

Tian et al. [24] in 2017 worked to improve the embedding security with the adaptive bitrate modulation techniques. The method first splitted the frames into groups and used different bitrate modulation for each group to embed the secret messages that matched the bitrate vector and the second method also used hamming matrix encoding strategy for those frames where the bitrate was not identical with the previous speech frame. The author achieved higher transparency and embedding efficiency with good security to resist traditional steganalysis, however the bandwidth could not be achieved.

2.2.2. Spread spectrum techniques

Spread spectrum methods are used in telecommunication for spreading a signal across the frequency domain to achieve wider bandwidth. Later it was extended for various purposes such as secure communication, noise jamming, etc. Frequency Hopping Spread Spectrum (FHSS) and Direct Sequence Spread Spectrum (DSSS) were the methods employed in signal spreading. These techniques used pseudorandom sequences to frame the pattern for spreading across the bandwidth.

Kohls et al. [16] in 2016 used DSSS based Steganography for data hiding. Hiding was carried out by modulating information bits using pseudo random, orthogonal noise sequences and repeating the spreading operation. Parameterized modulation and acknowledgment schemes were adapted to the characteristics of a specific deployment scenario that provided a more performant censorship circumvention System. Data rate and robustness were improved. The scheme aimed more on Undetectability by not altering the statistical characteristics.

Table 2. Methods of transform domain

VoIP Methods	Technique	Metrics improved		
		Undetectability	Bandwidth	Robustness
Spread spectrum	DSSS	✓	✓	✓
Pitch modification	Hide F0		✓	

2.2.3. Pulse code modulation technique

Pulse code modulation deals with the process of representing the amplitude of quantized sample of analog waves into digital form. The original analog signal is achieved back using the reverse process called demodulation.

Rui and Huang [27] experimentally worked on embedding the secret data depending upon the smoothness of the speech block. The speech stream was divided into blocks. The scheme selected less number of cover bits in the flat blocks, since flat blocks are more sensitive that may lead to degradation in the speech affecting the quality. On the other part, more number of cover bits was selected in the sharp blocks that are less sensitive to voice degradation. The proposed method proved to withstand RS detection and achieved good embedding capacity with less degradation of voice.

Yijing et al. [26] in 2013 worked on embedding the secret data in the real time audio packets in VoIP communication with smart grids. Secret message was encrypted with AES and then different algorithms are used in embedding the secret message. The author concludes that the method can withstand statistical steganalysis and each packet was of different capacity for secret data insertion.

2.2.4. Quantization Index Modulation (QIM)

Quantization is involved in converting the sampled analog values into discrete digital values with the help of quantized table. QIM works on encoding the quantized values into bit streams based on the codebook. The codebook is a form of tables, which consists of some patterns and its matching 3-digit bit. The authors focus on framing the codebook, such that the secret message will be embedded along with codes. Since embedding takes place along with the encoding process, the presence of secret data seems indistinguishable and improves undetectability.

Tian, Liu and Li [7] in 2014 designed a key based codebook division strategy that follows Kirchoff principle to provide higher security. With added security, the author also introduced selecting random positions dynamically and used matrix embedding technique for embedding which proved to have higher transparency and security.

Liu et al. [18] in 2016 proposed a neighbor index division, in which the neighbor indexed codewords were divided into identical sub codebooks with suitable stegocoding strategy and in turn each codebook has k subcodebooks, $K \geq 3$. Multi-ary coding strategy was introduced for odd numbered division. Even though the experimental results didn't outperform the CNV algorithm in reducing the distortion, the author have tried a simpler implementation and flexible partitioning method that was suitable for practical implementation.

Table 3. Methods on coded domain

VoIP Methods	Technique	Metrics improved		
		Undetectability	Bandwidth	Robustness
Pulse code modulation	Smoothness of speech block. (flat or sharp block)	✓	✓	
	Variable packet size with variable embedding using smart grids.	✓		
Quantization index modulation	Codebook division strategy on Kirchhoff principle	✓		
	Neighbor index division	✓		
	QIM graph model	✓		✓
	Quantization index set (LPC)	✓	✓	

Huang et al. [22] in 2017 used the QIM, that created a graph model for codebook space of quanitzer, based on which QIM controlled algorithm for partitioning the codebook space was preferred. The codebook division was mapped with the secret keys depending upon the codeword partition balance and partition diversity. With the help of these secret keys the author proved the proposed method has improved undetectability and robustness than the original QIM method.

Liu, Li and Wang [23] in 2017 proposed a QIM Steganography method that replaced the quantization index set in linear predictive coding where one quantization index was changed to hide at most three binary bits. The author tried to reduce the number of alterations to be done, which paved the way to improve undetectability. The quantization index division was carried upon based on genetic algorithm parameters to avoid the additional distortions caused due to embedding. Good resistance and Steganographic capacity was achieved through this method compared to existing LPC method.

2.2.5. Pitch modification method

Janicki et al. [9] in 2016 proposed a method Hide F0, where the F0 parameter (termed as FP) in the paper. Linear approximation of FP was done based on the Mean square error value calculated from the last frame FP and the Current frame FP. If the error was above threshold, the approximation was not done and the approximation flag was set to 0. If the error was equal or below the threshold then the approximation was done and the flag was set to 1 to specify that the current frame has the secret message inside. Decoding was done only to the frames with the approximation flag value set to one.

Janicki [14] in 2016 improved a method HideF0 which uses fine pitch (frequency F0) parameter that acts as a channel for data hiding. The maximum bandwidth achieved using speex codec in mode 5 is 220 bps with no Steganographic cost, which was comparatively high than the original HideF0 methods. But the improved version suffers a slight decrease in the speech quality due to an approximation flag in the packet header.

2.2.6. Packet loss methods

The speech samples are segmented as data packets of same size or of different size and transmitted over the IP network. The protocols involved in the transmission of data packets are TCP, UDP. The TCP protocol is connection oriented and guarantees the delivery between the end users; TCP holds the logical connection between the client and server throughout the data transfer. TCP attempts for retransmission of lost packets, if ack packet is not received. So delivery of all packets will be guaranteed. UDP is connectionless protocol and does not guarantee the delivery of all packets. UDP does not care about the lost packets. Normally the VoIP prefers the UDP than the TCP, as TCP's retransmission mechanism causes an unnecessary delay in the real time voice communication, which frustrates the end users. In this type real time communication. The small number of packet loss creates only a minor impact on speech quality. The small degradation of speech is acceptable by the human auditory system.

Due to constant mobility, VoIP conversation may suffer from packet loss. Due to the packets loss, speech quality is affected. Another problem may even lead to a serious threat of losing the secret data embedded in it.

LACK [1] method, which used the packet loss as an advantage and created space in hidden communication channels by forcing packet loss and introduces redundant frames with secret messages embedded in it.

Qi, Peng and Sharif [13] in 2016 worked on the technique LACK proposed by Mazurczyk and Lubacz [1], The original method embedded the secret message by forced loss of VoIP frames and introduced redundant frames with secret message embedded in it. As a result, the number of redundant frames was increased. The improved LACK technique used Discrete Spring Transform (DST) technique. Applying DST reduced the perceptual redundancy, which occurred in the original paper. Good quality of voice was achieved with reduced redundancy and also provides perceptual possibilities of how multimedia channel can be exploited.

Jiang et al. [15] in 2016 worked on the VoIP communication scheme based on UDP and the secret data was first encrypted using block cipher method. Each cipher block was then embedded into the audio signal encoded by Pulse Code Modulation (PCM) codec. The author used the prediction technique based on Gilbert packet loss model, which took the real time network traffic statistics data of number of packets sent and lost. This model predicts whether the VoIP packets sent through the network will be lost or not. If the prediction result declared, that the packet may not be lost, the corresponding packet was chosen for embedding there by efficiently reducing the risk of packet loss after embedding. PESQ and SNR of the stego samples and non stego samples revealed the fact that, the increase in the packet loss level lead to high speech detoriation rate.

Table 4. Packet loss methods

VoIP Methods	Technique	Metrics improved		
		Undetectability	Bandwidth	Robustness
Packet loss methods	Lack	✓		
	LACK with DST		✓	
	Gilbert packet loss model		✓	

Huang and Tang [17] in 2016 have proposed a multi dimensional spatial steganography model which aimed to improve the embedding capacity and security. The author identified various data hiding features (orthogonal hiding features) to form n- dimensional hiding space. Separate hiding algorithms were indulged for n elements. The author attempted to handle the packet loss by retransmitting the lost packets on receiving three unordered ack packets and before the expiry of the timer. Using Mel Frequency Cepstral Coefficients (MFCC) as a steganalysis method, only secret message embedded by LSB method was detected, claiming that security was improved, such that among multi dimensional data hidden, only part of the secret message was extracted and remaining message remains undetectable.

2.2.7. Silence suppression methods

The VoIP communication in general has some silence intervals in the middle of the speech for listening. Silence packets are intended to save the bandwidth. But these silence packets are utilized in the favor of covert communications by the researchers.

Schmidt et al. [20] in 2017 worked on the hybrid approach where the header fields of the RTP packets were modified to introduce fake packets containing silence with secret messages embedded in chunks, so as the original voice traffic doesn't have any change and does not affect the voice quality except a few degradations. In spite of not using the VoIP payload packets for modification, the presence of secret data is revealed in steganalysis.

Table 5. Codec specific methods

VoIP Methods	Technique	Metrics improved		
		Undetectability	Bandwidth	Robustness
Codec specific methods	Linear spectral frequencies (iLBC)	✓		
	Linear predictive encoding with matrix embedding (G.723.1)	✓	✓	
	Fixed codebook transposing (G.729)	✓	✓	

2.2.8. Codec specific methods

Codecs are the algorithms that are intended to perform encoding and decoding of the digital data stream for transmission. Each and every codec have some unique features apart from baseline system. These features make the difference in choosing the codecs for various purposes. For example, Low bit rate codecs are used in VoIP communications to save bandwidth as it requires less than 4Kb/s for transmission and storage. But low bit rate codecs have less redundant data, which is a challenge for the VoIP steganography to hide secret message. Researchers have chosen specific codecs for embedding secret message as follows.

Calpouzos and Varol [3] in 2013 worked on the iLBC which is an open source low bit rate codec. Author makes use of LSF (Linear Spectral Frequencies) for hiding the secret data and the embedding process is done during the compression itself so the RTP payload requires no modification and the secret data remains unidentified separately, but produces some mild distortions.

T i a n, L i u and L i [7] in 2014 worked on G.723.1 speech codec. The author has taken the frame rates 6.3 Kb/s and 5.3 Kb/s of the specific codec for secret data embedding. Bit rate Downgrading (BD) and Bit rate Switching (BS) are the two methods proposed. The first one converts the speech frames with high bitrate to low bitrate to create more space for embedding. Second method switches the bitrates to encode the secret message into the speech frames. Experimental results prove that the two methods embed the secret message with less voice degradation and negligible delays.

S h u f a n, T a n g and C h e n [25] in 2016 worked on G.729 speech codec. Fixed codebook parameter of G.729 codec is taken as cover for data hiding. The previous works on this parameter does not yield imperceptibility, as direct replacement in the LSB of the fixed codebook resulted in large displacements. Thus the author has done encoding by transposing the locations of adjacent pulses in the fixed codebook vector which resulted in smaller displacements. The location relationship between the adjacent pulses and parity value is encoded with the secret message. The experimental results also yielded high imperceptibility.

2.3. Evaluation metrics

Apart from the metrics specified above, there are some other metrics which measures the quality. The metrics used in the above mentioned methods are listed here for easy reference.

Embedding Rate (ER) or Hiding Rate (HR) = Number of *n* secret bits embedded in *m* cover bits (*n*/*m*).

Bit Change Rate (BCR) = Number of *r* cover bits changed/the total number of *m* cover bits (*r*/*m*).

Embedding Efficiency (EE) = Number of secret bits embedded at a single embedding change.

Codeword Change Rate (CCR) = Number of *k* codewords changed in *m* cover codewords (*k*/*m*).

Perceptual Evaluation of Speech Quality (PESQ) = It is a testing methodology to evaluate the speech quality. This is standardized as ITU-T recommendation P.862.

Mean Opinion Score-Listening Quality Objective (MOS-LQO) = Converted from PESQ the score ranges from 1(bad) up to 5(best).

3. Conclusion

Steganography has been an age old technique for hiding secret message, which evolved from time to time with various media as a cover. VoIP steganography has been one of the dynamic techniques for hiding secret messages in VoIP payload packets and other unused header fields. Upon analyzing various VoIP steganographic methods, some observations were made; these observations may help us identify the possibilities of various cover positions and also provides the drawbacks of the existing methods, which can be taken for future enhancements.

Table 1 summarizes the methods of temporal domain discussed with the metrics improved by the author. Traditional LSB methods provided us high embedding capacity but were identified using statistical steganalysis methods. Even though LSB techniques are less secure, many researchers propose new techniques or combined methods using LSB to provide a high security and capacity especially, Matrix embedding technique was used separately or combined with other techniques. The implementation of LSB methods with experimental results proved to be convenient and successful. In spite of improving the undetectability of the LSB methods, the robustness was not improved. Robustness depends on the speech quality. Due to high modification in the LSB methods, the speech quality degraded to a large extend

Table 2 summarizes the methods under transform domain. The methods of transform domain prove to be robust because of its masking effect. Unfortunately only a few promising works were done. Table 3 describes the methods under coded domain. Other VoIP Steganography methods also tend to provide security and bandwidth, but the implementation is more complex than the traditional LSB methods. Another domain, in which VoIP techniques achieve promising results, is QIM. More research is needed in this domain, since the implementation is not complex; once the codebook is framed then the embedding can be performed accordingly. Table 4 describes the packet loss methods, packet loss methods requires more modifications in the packet headers but bandwidth achieved would be less than the payload methods. Table 5 elaborates the codec specific methods, Even though they used combination of above said methods, the limitation to specific codec remains a drawback.

However, Bandwidth capacity and security (undetectability) are inversely proportional to each other. If size of secret bits embedded in a packet is more then the packets are more prone to steganalysis and vice versa happens – if the size of the secret bits to be embedded in a packet is less it is less prone to steganalysis.

References

1. Mazurczyk, W., J. Lubacz. LACK-a VoIP Steganographic Method – Telecommunication Systems, Vol. **45**, 2010, No 2-3, pp. 153-163.
2. Hui, T., H. Jiang, K. Zhou, D. Feng. Adaptive Partial-Matching Steganography for Voice over IP Using Triple M Sequences. – Computer Communications, Vol. **34**, 2011, No 18, pp. 2236-2247.
3. Calpouzos, G., C. Varol. GiLBCSteg: VoIP Steganography Utilizing the Internet Low Bit-Rate Codec. – Computer Science and Information Technology, Vol. **1**, 2013, No 2, pp. 153-158.
4. Mazurczyk, W. VoIP Steganography and Its Detection – a Survey. – ACM Computing Surveys (CSUR), Vol. **46**, 2013, No 2, p. 20.
5. Mazurczyk, W., P. Szaga, K. Szczypiorski. Using Transcoding for Hidden Communication in IP Telephony. – Multimedia Tools and Applications, Vol. **70**, 2014, No 3, pp. 2139-2165.
6. Janicki, A., W. Mazurczyk, K. Szczypiorski. Steganalysis of Transcoding Steganography. – Annals of telecommunications-annales des telecommunications, Vol. **69**, 2014, No 7, pp. 449-460.
7. Tian, H., J. Liu, S. Li. Improving Security of Quantization-Index-Modulation Steganography in Low Bit-Rate Speech Streams. – Multimedia Systems, Vol. **20**, 2014, No 2, pp. 143-154.

8. J a n i c k i, A., W. M a z u r c z y k, K. S z c z y p i o r s k i. On the Undetectability of Transcoding Steganography. – Security and Communication Networks, Vol. **8**, 2015, No 18, pp. 3804-3814.
9. J a n i c k i, A. Novel Method of Hiding Information in IP Telephony Using Pitch Approximation. – In: Proc. of 10th International Conference on Availability, Reliability and Security (ARES'15), IEEE, 2015, pp. 429-435.
10. T i a n, H., J. Q i n, S. G u o, Y. H u a n g, J. L i u, T. W a n g, Y. C a i. Improved Adaptive Partial-Matching Steganography for Voice over IP. – Computer Communications, Vol. **70**, 2015, pp. 95-108.
11. T i a n, H., J. Q i n, Y. H u a n g, Y. C h e n, T. W a n g, J. L i u, Y. C a i. Optimal Matrix Embedding for Voice-over-IP Steganography. – Signal Processing, Vol. **117**, 2015, pp. 33-43.
12. Q i n, J., H. T i a n, Y. H u a n g, J. L i u, Y. C h e n, T. W a n g, X. A. W a n g. An Efficient VoIP Steganography Based on Random Binary Matrix. – In: Proc. of 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC'15), IEEE, 2015, 2015, pp. 462-465.
13. Q i, Q., D. P e n g, H. S h a r i f. DST Approach to Enhance Audio Quality on Lost Audio Packet Steganography. – EURASIP Journal on Information Security, 2016, No 1, p. 20.
14. J a n i c k i, A. Pitch Based Steganography for Speex Voice Codec. – Security and Communication Networks, Vol. **9**, 2016, No 15, pp. 2923-2933.
15. J i a n g, Y., S. T a n g, L. Z h a n g, M. X i o n g, Y. J. Y i p. Covert Voice over Internet Protocol Communications with Packet Loss Based on Fractal Interpolation. – ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), Vol. **12**, 2016, No 4, p. 54.
16. K o h l s, K., T. H o l z, D. K o l o s s a, C. P ö p p e r. Skypeline: Robust Hidden Data Transmission for VoIP. – In: Proc. of 11th ACM on Asia Conference on Computer and Communications Security, ACM, 2016, pp. 877-888.
17. H u a n g, Y. F., S. Y. T a n g. Covert Voice over Internet Protocol Communications Based on Spatial Model. – Science China Technological Sciences, Vol. **59**, 2016, No 1, pp. 117-127.
18. L i u, J., H. T i a n, J. L u, Y. C h e n. Neighbor-Index-Division Steganography Based on QIM Method for G.723.1 Speech Streams. – Journal of Ambient Intelligence and Humanized Computing, Vol. **7**, 2016, No 1, pp. 139-147.
19. J i a n g, Y., S. T a n g. An Efficient and Secure VoIP Communication System with Chaotic Mapping and Message Digest. – Multimedia Systems, 2017, pp. 1-9.
20. S c h m i d t, S. S., W. M a z u r c z y k, J. K e l l e r, L. C a v i g l i o n e. A New Data-Hiding Approach for IP Telephony Applications with Silence Suppression. – In: Proc. of 12th International Conference on Availability, Reliability and Security, ACM, 2017, p. 83.
21. L i u, P., S. L i, H. W a n g. Steganography Integrated into Linear Predictive Coding for Low Bit-Rate Speech Codec. – Multimedia Tools and Applications, Vol. **76**, 2017, No 2, pp. 2837-2859.
22. H u a n g, Y., H. T a o, B. X i a o, C. C h a n g. Steganography in Low Bit-Rate Speech Streams Based on Quantization Index Modulation Controlled by Keys. – Science China Technological Sciences, Vol. **60**, 2017, No 10, pp. 1585-1596.
23. L i u, P., S. L i, H. W a n g. Steganography in Vector Quantization Process of Linear Predictive Coding for Low-Bit-Rate Speech Codec. – Multimedia Systems, Vol. **23**, 2017, No 4, pp. 485-497.
24. T i a n, H., J. S u n, C. C. C h a n g, J. Q i n, Y. C h e n. Hiding Information into Voice-over-IP Streams Using Adaptive Bitrate Modulation. – IEEE Communications Letters, Vol. **21**, 2017, No 4, pp. 749-752.
25. S h u f a n, Y., G. T a n g, Y. C h e n. Incorporating Data Hiding into G.729 Speech Codec. – Multimedia Tools and Applications, Vol. **75**, 2016, No 18, pp. 11493-11512.
26. Y i j i n g, J., L. Z h a n g, S. T a n g, Z. Z h o u. Real-Time Covert VoIP Communications over Smart Grids by Using AES-Based Audio Steganography. – In: IEEE International Conference on Green Computing and Communications (GreenCom), 2013 IEEE and Internet of Things (iThings/CPSCom), and IEEE Cyber, Physical and Social Computing, pp. 2102-2107.

27. R u i, M., Y H u a n g. An Approach of Covert Communication Based on the Adaptive Steganography Scheme on Voice Over IP. – In: IEEE International Conference on Communications (ICC), 2011, pp. 1-5.
28. J a n i c k i, A., W. M a z u r c z y k, K. S z c z y p i o r s k i. Influence of Speech Codecs Selection on Transcoding Steganography. – Telecommunication Systems, Vol. **59**, 2015, No 3, pp. 305-315.
29. D j e b b a r, F., B. A y a d, K. A. M e r a i m, H. H a m a m. Comparative Study of Digital Audio Steganography Techniques. – EURASIP Journal on Audio, Speech, and Music Processing, Vol. **1**, 2012, p. 25.

Received: 11.09.2018; Second Version: 21.12.2018; Accepted: 08.01.2019