Universal steganography model for low bit-rate speech codec

Universal Steganography Model for Low Bit-Rate Speech Codec

Shanyu Tang[1], Qing Chen[1], Wei Zhang[1], Yongfeng Huang[2], *Senior Member IEEE*

[1] *School of Computer Science, China University of Geosciences, Wuhan, 430074, P. R. China*

[2] *Department of Electronic Engineering, Tsinghua University, Beijing, 100084, P.R. China*

Abstract

Low bit-rate speech codec offers so many advantages over other codecs that it has become increasingly popular in audio communications such as mobile and VoIP (Voice over Internet Protocol) communications, and thus researching steganography in low bit-rate speech codec is of important significance. In this study, we proposed a universal VoIP steganography model for low bit-rate speech codec that uses the PESQ deterioration rate and the decoding error to automatically choose a data embedding algorithm for each VoIP bitstream, which enables ones to achieve covert communications using a low bit-rate speech codec efficiently and securely. Since no or little attention has been paid to steganography in iSAC (Internet Speech Audio Codec), it was chosen as the test codec to verify the effectiveness, security, and practicability of the proposed steganography model. The experimental results show that, with the proposed steganography model, it achieved the average PESQ deterioration rate of 4.04% (less than 5%, indicating strong imperceptibility) and a high data hiding capacity up to 12 bits/frame (400 bits/second, three times larger than other methods), and the proposed steganography model could effectively resist the latest steganalysis.

Keywords: Steganography model, low bit-rate speech codec, VoIP, information hiding, iSAC.

1. Introduction

With the rapid development and wide application of the Internet, information distribution and data communications over the Internet have become more and more

simple and efficient, so people are becoming increasingly concerned about the security of the private information transmitted over the Internet. Cryptography has been playing an important role in the traditional field of information security, but it is considered to be insufficient to meet the requirements of information security with new application demand and ever-increasing computing power.

Steganography, as an art and science of covert communications that conceal the existence of secret information embedded in cover media over an insecure network, has drawn a great deal of attention. Compared with cryptology, steganography not only hides the content of the secret information but also conceals the covert communication itself. Steganography enables ones to hide the secret information in vast quantities of information carriers. If an attacker wants to steal the secret information or just prevent the receiver from acquiring the secret information, she or he has to distinguish the real information carrier from the vast quantities of redundant information carriers, which is unlikely to achieve. Therefore, steganography is thought to be more secure than cryptography in some extent.

A large number of information hiding methods have been proposed for multimedia carriers, such as plaintext [1], video and audio files [2] [3], and images with BMP or JPEG format [4]. The steganography methods mentioned above are based on static storage media carriers. In comparison with these static storage media carriers, Voice over Internet Protocol (VoIP)-based communication as a real-time application, which allows users to make telephone calls via an Internet connection, has drawn progressively attention. Researchers such as Wojciech Mazurczyk, C. Kratzer, J. Dittmann, C.Y. Wang, et al. have made efforts on VoIP steganography; their studies have shown that VoIP is an excellent multimedia carrier in which steganography can be applied to [5-8].

So far the information hiding methods for VoIP steganography are mainly based on bit load embedding covert communication. These methods are used to achieve the goal of

covert communications by modifying the redundant information of bit load of streaming media. Huang, Tang, et al. [9] suggested a VoIP steganographic algorithm of embedding secret bits by replacing part of inactive frames of G.723.1 codec, which has the limitation of not being able to be applied to active frames. Tian et al. [10] presented a covert communication model based on least significant bits (LSB) steganography in VoIP. Liu et al. [11] proposed the Least Significant Digit (LSD) method to enhance the hiding capacity of the traditional LSB substitution steganography by exploiting more embedding states. The above three methods are mainly used for encoded parameters LSB embedding, capable of achieving high data hiding capacity, but both have negative impacts on the quality of synthetic speech and thus they are vulnerable to steganalysis based on statistical analysis methods.

Based on graph theory, Xiao et al. [12] proposed an algorithm called CNV (Complementary Neighbor Vertex）to optimize the group of codebook and used the QIM (Quantization Index Modulation) method to embed secret bits, which has strong imperceptibility but the data hiding capacity is quite low. Tian et al. [13] suggested a secure QIM method (Sec-QIM) to further improve the security of the QIM method by introducing random position selection. Liu et al. [14] transformed the data of the carrier into the transform domain and hid the secret information by modifying some parameters in the transform domain, resulting in a moderate data hiding capacity and imperceptibility. Huang et al. [15] suggested a new algorithm for steganography in low bit-rate VoIP audio streams by integrating information hiding into the process of speech encoding; although the modified steganography method offers strong imperceptibility (less than 5% PESQ deterioration rate) but the data hiding capacity is as low as 4 bits/frame (133.3 bits/second), which means that its practical applications are limited.

The steganography methods mentioned above represent three types of approaches to covert communications over VoIP streams. They all have their own advantages and weaknesses, and some methods can only be used under certain conditions. In view of

the situation, this study proposed a universal steganography model for low bit-rate speech codec based on VoIP, aiming at flexibly applying a steganography algorithm to a low bit-rate speech codec to meet different needs of imperceptibility and data hiding capacity. In the proposed steganography model, we first inferred the composition of the bitstream of low bit-rate speech codec from the process of encoding and decoding; then we used a specially designed analysis module to analyze different parts of the bitstream and find out which parts could be used for information hiding and which steganography algorithm should be adopted for better imperceptibility and higher data hiding capacity. In addition, we evaluated the proposed steganography model with low bit-rate speech codec iSAC, i.e. using iSAC to verify the effectiveness, security, and practicability of the proposed steganography model. The iSAC was chosen as the test codec because there was no published research paper detailing steganography in iSAC according to our knowledge. The evaluation test demonstrated that the proposed steganography model was effective in terms of strong imperceptibility (low PESQ deterioration rate) and high data hiding capacity.

The rest of the paper is organized as follows. Section 2 presents the proposed VoIP steganography model for covert communications over low bit-rate VoIP audio streams. Covert VoIP Communications in iSAC are discussed in Section 3. Section 4 details performance evaluation and steganalysis of the proposed steganography model. Finally, the paper concludes with a summary and directions for future work.

2. Proposed VoIP Steganography Model

Before designing an effective method for covert VoIP communication, a new steganography model is to be established for data hiding within audio streams encoded by low bit-rate speech codec. Figure 1 illustrates the proposed universal steganography model for low bit-rate speech codec which enables ones to efficiently and securely achieve covert VoIP communications in a low bit-rate speech codec. The proposed model describes how Encoder Module, Embedding Module, Analysis

Module, Decoder Module, and Extraction Module interact each other, to realize covert VoIP communication via steganography.

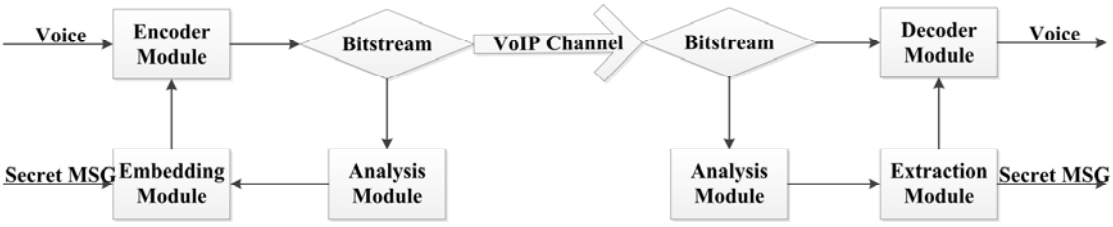

Fig. 1 Universal VoIP Steganography Model

Assuming that a sender wants to transmit a secret message (MSG) to a receiver, they pretend that they are talking about some inconspicuous topics over a VoIP communication channel. Firstly, they choose a low bit-rate speech codec, such as iSAC or iLBC (internet Low Bit Rate Codec). They then infer the composition of the bitstream of low bit-rate speech codec from the process of encoding and decoding. After that they use the Analysis Module (AM) to analyze different parts of the bitstream and find out which parts can be used for information hiding and which steganography algorithm should be adopted for better imperceptibility and data hiding capacity. Finally, the secret message that has been encrypted by a shared key previously is sent to the Encoder Module (E_nM) through the Embedding Module (E_mM). Thus the receiver retrieves the encrypted secret message from the Decoder Module (DM) through the Extraction Module (E_xM). Having used the shared key to decrypt the encrypted secret message, the receiver extracts the original secret message that the sender transmits over VoIP communication.

The Extraction Module E_nM is made up of the encoding process of low bit-rate speech codec and an additional interface which allows the bitstream of secret message to be embedded into the bitstream of low bit-rate speech codec. Correspondingly, the Decoder Module DM is composed of the decoding process of low bit-rate speech codec and the additional interface which allows the bitstream of secret message to be

extracted from the bitstream of low bit-rate speech codec.

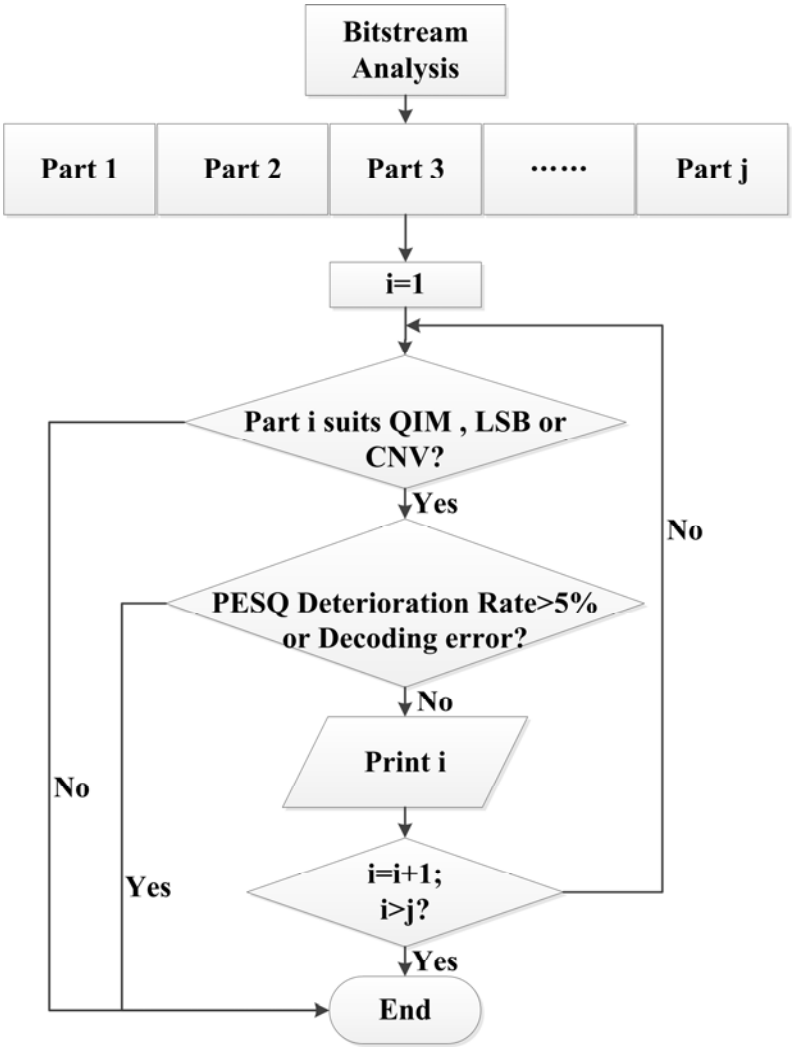

Fig. 2 Flowchart of Analysis Module

Figure 2 depicts how the Analysis Module AM works, i.e. how to use the PESQ deterioration rate and the decoding error to automatically choose a data embedding algorithm for each VoIP bitstream. The Analysis Module AM contains a number of commonly used information hiding algorithms. The analysis process matches the different parts of the bitstream of low bit-rate speech codec with the commonly used information hiding algorithms. After analysis, AM discovers which parts of the bitstream of low bit-rate speech codec could be used for information hiding and which information hiding algorithms could be adopted to realize covert VoIP communication

for better imperceptibility and high data hiding capacity.

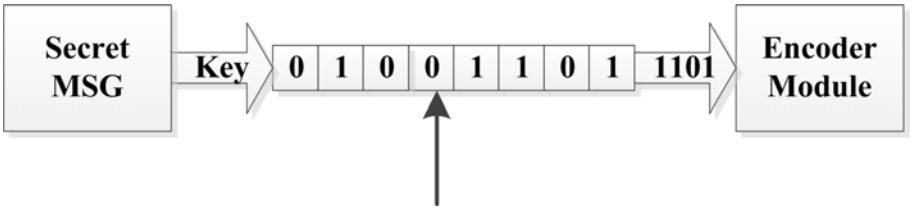

Fig. 3 Illustration of Embedding Module

Figure 3 depicts how the Embedding Module E_mM works. The secret message (MSG) may be a text or a picture, but whatever the secret message is, when it is read into memory after encrypted with the shared key, it only has 0 or 1. E_mM works as a container of 0 or 1 with a pointer pointing to the first data of 0 or 1. During the embedding process, when the E_nM asks for one bit through the additional interface, the E_mM gives one bit which is pointed by the pointer, and the pointer moves to the next bit which has not been used.

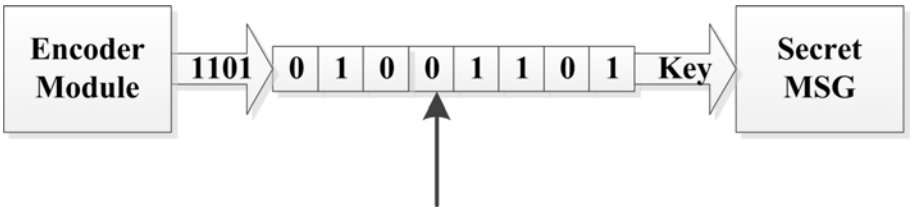

Fig. 4 Illustration of Extraction Module

Figure 4 describes how the Extraction Module E_xM works. E_xM works as a collection container of 0 or 1 with a pointer pointing to the last data of 0 or 1. During the extracting process, when the E_nM gives one bit through the additional interface, the E_xM collects one bit which is pointed by the pointer, and the pointer moves to the next bit which has not been used for collection.

3. Covert VoIP Communications over iSAC

Waveform-based codecs such as G.711 and G.726 make use of previous speech to

extrapolate the lost packets, so they need to store a large amount of previous samples. Low bit rate codecs of G.729, G.723.1, iLBC, and iSAC preserve previous coded parameters like LPC, pitch, and excitation, and the decoder keeps preserving the parameters during good frames, which are reused for predicting the lost speech.

The iSAC is a wideband speech codec, developed by Global IP Solutions (GIPS) which was acquired by Google Inc. in 2011. This codec is suitable for VoIP applications and streaming audio. The encoded blocks are encapsulated in an appropriate protocol for transportation, such as RTP (Real-Time Protocol). It is one of the codecs widely used by popular Internet applications such as AIM Triton, Gizmo5, QQ, and Google Talk. It was formerly a proprietary codec licensed by Global IP Solutions. As of June 2011, it is part of the open source WebRTC project, which includes a royalty-free license for iSAC when using the WebRTC codebase [16].

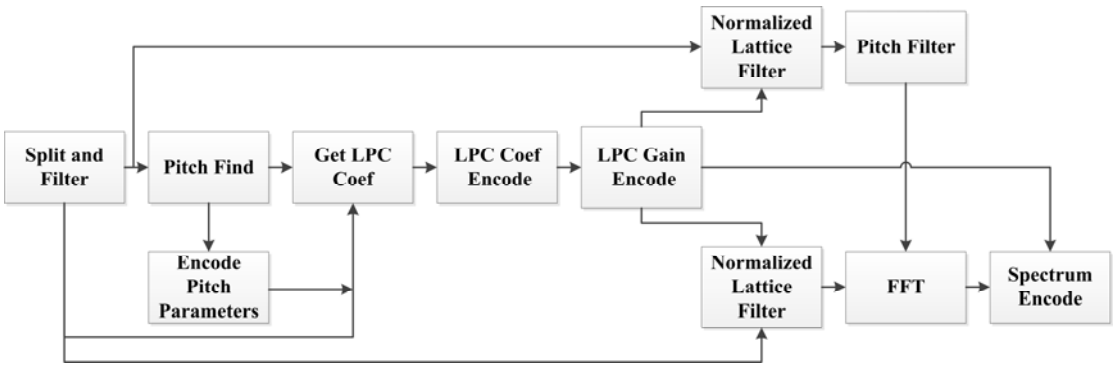

Fig. 5 iSAC low band encoder

The iSAC adopts the traditional CELP (Code Excited Linear Prediction) dual exciting analysis model, with the frame length of 30 ms or 60 ms. After analyzing and extracting all kinds of parameters which are used for reconstituting voice, the encoder sends these parameters to the decoder via the Internet. The decoder decodes these parameters and reconstitutes voice through the adaptive codebook and the fixed codebook.

The iSAC uses the KLT (Karhuner-Loeve Transform) algorithm before quantizing Pitch Lag and Pitch Gain, which is different from the traditional speech codec. Besides, in the process of the residual error coding, iSAC uses noise spectrum approximation for the real part and the imaginary part in the frequency domain after the FFT (Fast Fourier Transformation) transformation rather than using noise spectrum approximation in the time domain, which is very rare in the traditional codec. The iSAC divides the encoding process into two parts, low band encoder (Fig. 5) and high band encoder (Fig. 6). The low band encoder deals with the low band of the speech signal, and the high band encoder deals with the high band of the speech signal.

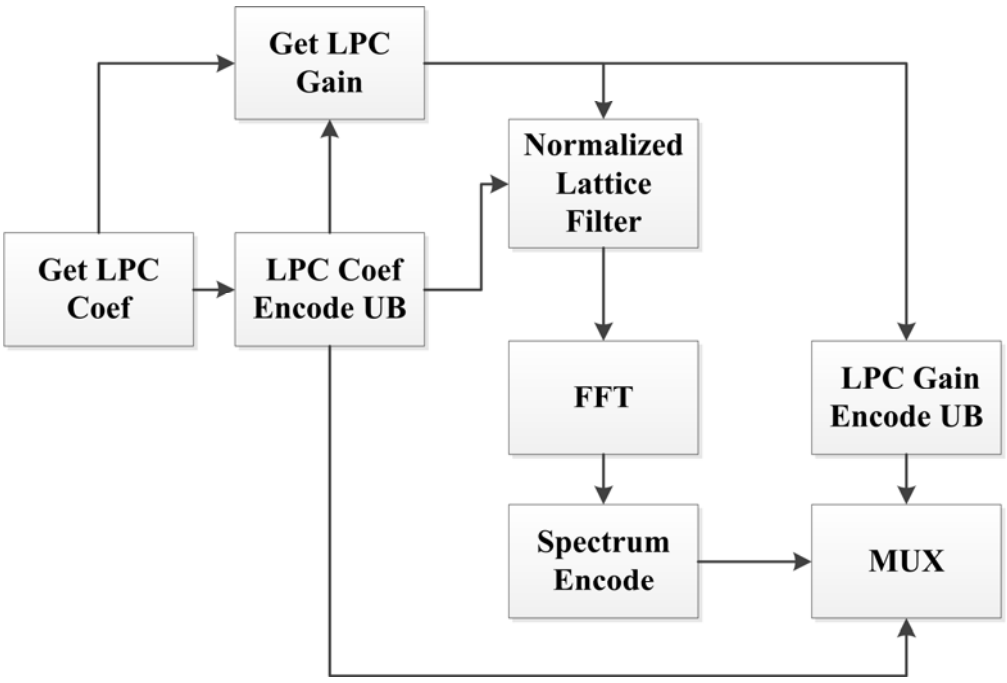

Fig. 6 iSAC high band encoder

We chose iSAC as the test codec because there was no publication detailing steganography in iSAC according to our knowledge. After downloading the iSAC codec from the web site [17], we first inferred the composition of the bitstream of iSAC codec from the process of encoding and decoding.

Fig. 7 Composition of the Bitstream of iSAC

Figure 7 depicts the composition of the bitstream of iSAC. Having used the Analysis Module (AM, Fig. 2) to analyze different parts of the bitstream, we found that Pitch Gain, LPC Coef, and LPC Gain could be used for information hiding, and LSB and QIM data embedding algorithms could be adopted for better imperceptibility and high data hiding capacity.

Considering the limitation of the paper length, we here only show the test results of LPC Gain using QIM data embedding algorithm. In the process of LPC Gain quantization, we divided the quantization index collection into two parts based on parity. When 0 is to be hidden, we searched for the optimal quantization index in the even part of the quantization index collection; when 1 is to be hidden, we searched for the optimal quantization index in the odd part of the quantization index collection. In the process of LPC Gain dequantization, if a quantization index was even, it meant that 0 had been hidden in this quantization index, and if a quantization index was odd, it meant that 1 had been hidden in this quantization index. It is anticipated that if one quantization index is even, then the adjacent quantization index must be odd, and if one quantization index is odd, then the adjacent quantization index must be even. So the quantization index calculated by the QIM data embedding algorithm is the optimal quantization index or the quantization index second to the optimal quantization index. Thus the QIM algorithm mentioned above has a negligible impact on the quality of speech. In addition, with the QIM data embedding algorithm, the data embedding rate of the secret message achieved 12 bits / frame (400 bits / second).

Suppose the secret bitstream after being encrypted by the shared key is denoted by

$$S = [S(0), S(1), S(2), S(i), \ldots S(L-1)]$$

where L is the length of the secret bitstream, $S(i) \in \{0,1\}$, $0 \leqslant i < L$, index[j] is the array used for storing the quantization index (since each frame has 12 LPC Gain coefficients, it has $0 \leqslant j < 12$), i = 0, j = 0, N = 0, and U is the quantizer.

The secret information embedding algorithm for each frame is described as follows:

Step 1: U is adjusted to U' based on s(i), $U' = \{x \mid x \in U, x \equiv S(i) \text{ Mod}(2)\}$. After the adjustment, the element in U' has the same parity with S(i), then search for the optimal quantization index[j] (index[j]∈U') in U', i = i + 1, j = j + 1.

Step 2: If i = L, the secret information hiding completes, otherwise turn to **Step 3**.

Step 3: If j = 12, the secret information hiding in the current frame completes, then start a new frame, j is initialized to 0 again; otherwise, turn to **Step 1**.

Suppose the secret bitstream to be extracted is denoted by

$$S=[S(0), S(1), S(2), S(i), \ldots S(L-1)]$$

where L is the length of the secret bitstream to be extracted, $S(i) \in \{0,1\}$, $0 \leqslant i < L$, index[j] is the array used for storing the quantization index (since each frame has 12 LPC Gain coefficients, it has $0 \leqslant j < 12$), i = 0, and j = 0. According to the secret information embedding algorithm, index[j] has the same parity (0 or 1) with S(i), thus S(i) can be extracted based on the parity of index[j].

The secret information extraction algorithm for each frame is described as follows:

Step 1: The secret bitstream encrypted by the shared key is calculated through the equation S(i) = index[j] Mod(2), i = i + 1, j = j + 1.

Step 2: If i = L, the secret information extracting completes; otherwise turn to **Step 3**.

Step 3: If j = 12, the secret information extracting from the current frame finishes, then start a new frame, j is initialized to 0 again; otherwise, turn to **Step 1**.

4. Performance Evaluation and Steganalysis

Since no or little attention has been paid to VoIP steganography in iSAC (Internet Speech Audio Codec), it was chosen as the test codec to verify the effectiveness, security, and practicability of the proposed steganography model.

4.1 Test Samples

In order to verify universality of the proposed steganography model described above, we chose speech sample pools consisting of a large number of VoIP audio streams / clips taken from different people. The sample pools included four types of audio stream pools, which are Chinese Man Speech (CM), Chinese Woman Speech (CW), English Man Speech (EM), and English Woman Speech (EW). Each sample pool consisted of 500 pieces of VoIP stream samples with length of 20 seconds. These samples made up the twenty minutes sample library called Sample-20.

In the experiments, each frame was fully embedded based on the algorithm proposed above. Covert communication was achieved by concealing the existence of the secret information embedded in cover media over an insecure channel. Any changes to the cover media (audio samples) would be suspicious to the third party, thus the original audio samples as the cover media were not scrambled during the steganography experiments.

4.2 Perceptual Evaluation of Speech Quality

We used the perceptual evaluation speech quality (PESQ) value to assess the subjective quality of the stego audio samples (VoIP clips with data hiding). ITU-T P.862 PESQ recommendation was employed to measure the subjective quality of the stego audio samples (audio streams with hidden information), as shown in Fig. 8. The testing method is an objective method for predicting the subjective quality of narrowband speech codecs. As Fig. 8 shows, it uses the perceptual evaluation speech quality (PESQ) value to assess the subjective quality of the stego audio. As the PESQ is not well matched with mean opinion score (MOS), PESQ-listening

quality objective (LQO) is recommended to evaluate the quality of the stego audio. The PESQ is then mapped to the MOS-LQO value.

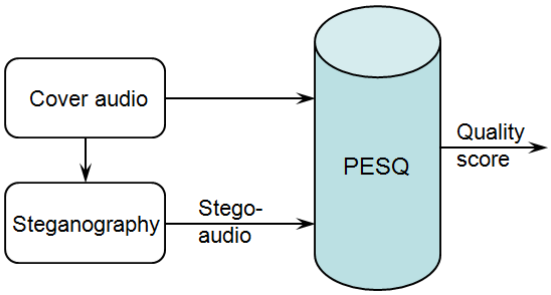

Fig. 8 PESQ measurements

4.3 Performance Measurements and Steganalysis

Figure 9 shows the PESQ values for the original VoIP stream samples (Chinese Man Speech) encoded by iSAC codec without any data embedding (indicated by 'No hiding'), and the PESQ values for the stego audio samples processed by the iSAC codec with data embedding by means of the proposed steganography algorithm (denoted by 'Hiding'), when the 20-second VoIP stream samples were used as cover media. As Fig. 9 shows, for the cover media of Chinese Man Speech, the variations in PESQ between the original speech samples and the stego speech samples were so small, which means the proposed steganography algorithm has little effect on the quality of the original speech.

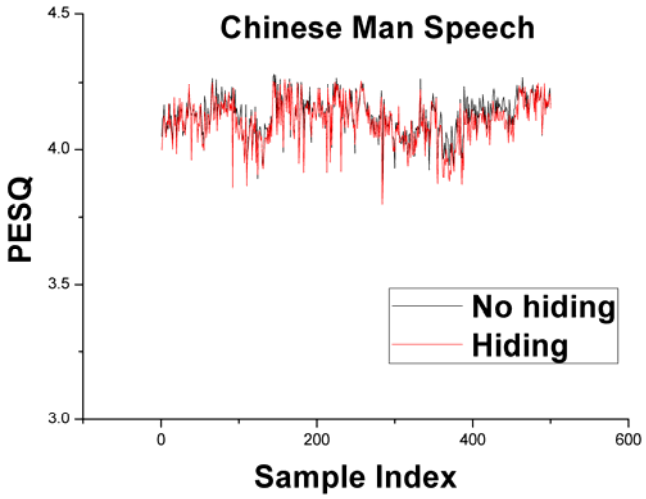

Fig. 9 PESQ values for 20-second CM samples using the proposed steganography algorithm

Figures 10, 11 and 12 show comparisons of PESQ values between the original VoIP stream samples and the stego audio samples processed by iSAC using the proposed steganography algorithm, for 20-second CW samples, 20-second EM samples, and 20-second EW samples, respectively. There were no obvious discrepancies in the PESQ value without (black curve: No hiding) and with data embedding (red curve: Hiding). As Figs. 10, 11 and 12 show, the variations in PESQ between the original VoIP stream samples and the stego audio samples were so small, indicating that the proposed steganography algorithm had no or very little impact on the quality of the synthesized speech.

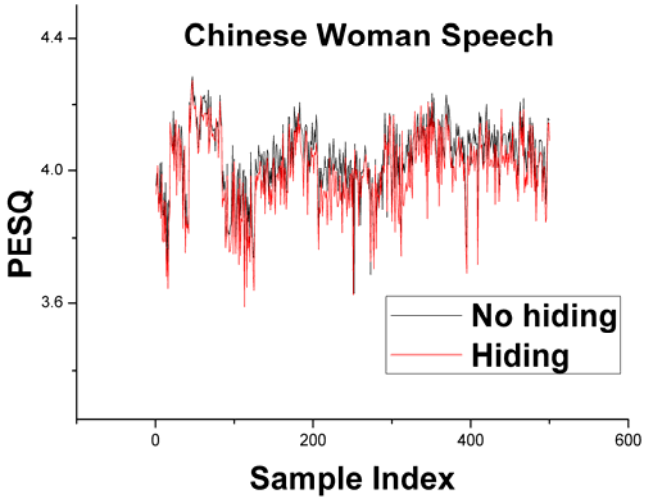

Fig. 10 PESQ values for 20-second CW samples using the proposed steganography algorithm

Table I shows the comparisons of PESQ (Perceptual Evaluation of Speech Quality) values between using the proposed steganography algorithm (Proposed Algorithm) and without using the steganography algorithm (Normal Codec). The negative changes of PESQ values indicated the deterioration of PESQ values while the positive changes of PESQ values indicated the optimization of PESQ values. The maximum deterioration rate of four types of audio samples (CM, CW, EM, EW) was 10.77%, and the average deterioration rate of these samples were 3.32%, 5.18%, 1.86%, and 5.79%, respectively. The total average deterioration rate of all

the samples was 4.04%, which is quite small and acceptable.

Table II shows the comparisons of changes in PESQ values and data hiding capacities between the proposed steganography algorithm and the steganography algorithm presented in [15]. The data hiding capacity was measured by the data embedding rate of the secret information in VoIP streams, as described in [15]. As Table II shows, both the average deterioration rate and the Standard Deviation of the proposed steganography algorithm are comparable to those of the steganography algorithm presented in [15]; with the proposed steganography algorithm using QIM, the data embedding rate of the secret information achieved 12 bits/frame (400 bits/second), which was three times higher than that of the steganography algorithm presented in [15].

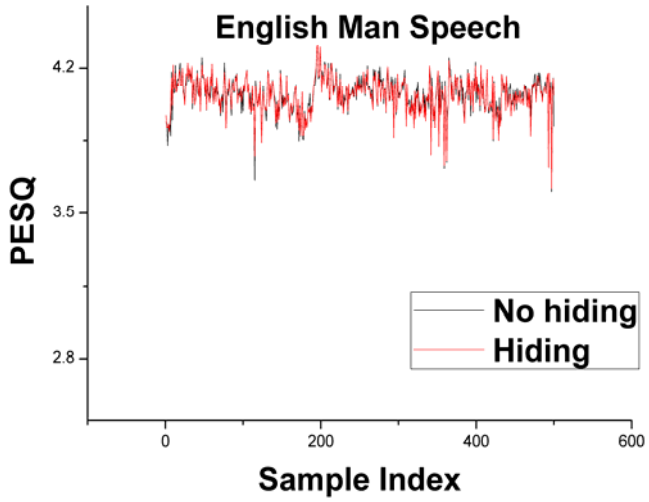

Fig. 11 PESQ values for 20-second EM samples using the proposed steganography algorithm

Different from the steganalysis methods described in references [18] and [19], we had developed an effective steganalysis method for VoIP steganography [20] [21]. For comparison purposes, we used the steganalysis method described in reference [20] to compute the PESQ detection rate in this study, i.e. DMFCC (Derivative Mel-Frequency Cepstral Coefficients) characteristic-based SVM (Support Vector Machine) detection method, to evaluate the security of the proposed steganography

algorithm. The DMFCC method can achieve a satisfying analysis result when it is used for LSB matching, Hid4PGP, and so on. SVM set RBF core function as its default parameter. The LIBSVM version which we used in the test was Version 3.0. In the SVM-scale of LIBSVM, the lower was -1, the upper was 1, and the other parameters used were set as default values. We used the DMFCC method to detect the proposed algorithm with the sample pool of Sample-20 in this study (Table III).

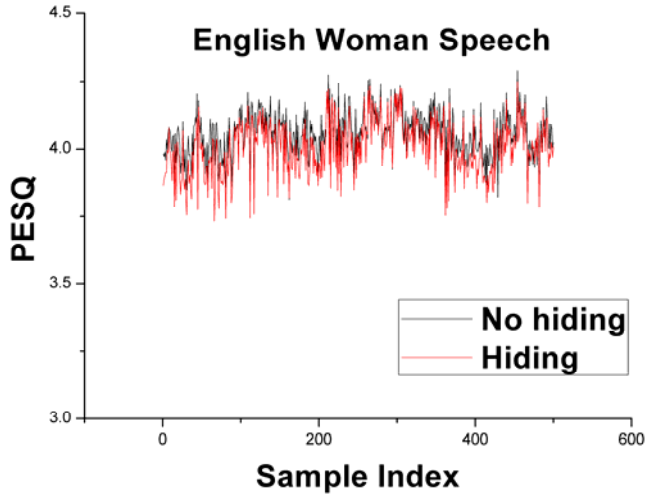

Fig. 12 PESQ values for 20-second EW samples using the proposed steganography algorithm

As Table III shows, the maximum correction detection rate of the four types of sample sets was 62.75%, and the minimum correction detection rate of the four types of sample sets was 49.70%, indicating that the proposed steganography algorithm could effectively resist the latest steganalysis, which is DMFCC characteristic-based SVM detection method [20].

5. Conclusion and Future Work

In this paper, we proposed a universal steganography model for low bit-rate speech codec to achieve covert VoIP communications in low bit-rate speech codec efficiently and securely. We used iSAC as the test codec to verify the effectiveness and practicability of the steganography model. The test results demonstrated that, with the

proposed model, VoIP steganography in iSAC had little impact on the PESQ of the speech and the average deterioration of PESQ was under 5%. The detection rate using DMFCC-SVM steganalysis method was under 62.75%, indicating that the proposed steganography algorithm could effectively prevent from being detected by the latest steganalysis. Moreover, with the proposed steganography model using QIM, the data embedding rate of the secret information achieved 12 bits/frame (400 bits/second), which was three times higher than other methods. The future work is to apply the proposed steganography model to other low bit-rate speech codecs, such as iLBC.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grant 61272469 and Grant 61271392, and the Wuhan Scientific Research Program under Grant 2013010501010144.

References

1. Zander S, Armitage G, Branch P. A survey of covert channels and countermeasures in computer network protocols. *IEEE Communications Surveys & Tutorials* 2007; **9** (3) : 44-57.
2. Badura S, Rymaszewski S. Transform domain steganography in DVD video and audio content. *Proceedings of the IEEE International Workshop on Imaging Systems and Techniques* 2007; pp. 1-5.
3. Yan D, Wang R, Zhang L. Quantization step parity-based steganography for MP3 audio. *Fundamenta Informaticae* 2009; **97** (1-2) : 1-14.
4. Fridrich J, Tevný T, Kodovský J. Statistically undetectable JPEG steganography: Dead ends, challenges, and opportunities. *Proceedings of the 9th ACM workshop on Multimedia & Security* 2007; pp. 3-14.
5. Kratzer C, Dittmann J, Vogel T. Design and evaluation of steganography for voice-over-IP. *Proceedings of IEEE International Symposium on Circuits and*

Systems 2006; pp. 2397-2340.

6. Dittmann J, Hesse D. Steganography and steganalysis in voice-over IP scenarios: Operational aspects and first experiences with a new steganalysis tool set. *Proceedings of SPIE* 2005; pp. 607-618.
7. Wang CY, Wu Q. Information hiding in real-time VoIP streams. *Proceedings of the 9th IEEE International Symposium on Multimedia* 2007; pp. 255-262.
8. Mazurczyk W, Kotulski Z. Covert channel for improving VoIP security. *Advances in Information Processing and Protection, Springer* 2007; pp. 271-280.
9. Huang YF, Tang S, Yuan J. Steganography in inactive frames of VoIP streams encoded by source codec. *IEEE Transactions on Information Forensics and Security* 2011; **6** (2) : 296-306.
10. Tian H, Zhou K, Huang YF. A covert communication model based on least significant bits steganography in voice over IP. *International Conference for Young Computer Scientists* 2008; pp. 647-652.
11. Liu J, Zhou K, Tian H. Least-significant-digit steganography in low bitrate speech. *2012 IEEE International Conference on Communications (ICC) 2012*; pp.1133-1137.
12. Xiao B, Huang, YF, Tang, S. An approach to information hiding in low bit-rate speech stream. *Proceedings of IEEE Global Telecommunications Conference* 2008; pp.1-5.
13. Tian H, Liu J, Li SB. Improving Security of quantization-index-modulation steganography in low bit-rate speech streams. *Multimedia Systems*; **20**(2) : 143-154.
14. Liu L, Li M, Li Q. Perceptually transparent information hiding in G.729 bitstream. *Proceedings of 4th International Conference for Intelligent Information Hiding and Multimedia Signal Processing* 2008; pp. 406-409.
15. Huang Y, Liu CH, Tang S. Steganography Integration into a Low-Bit Rate Speech Codec. *IEEE Transactions on Information Forensics and Security* 2012; **7**(6) : 1865-1875.
16. Web site. http://en.wikipedia.org/wiki/Internet_Speech_Audio_Codec, Sep. 2013.

17. Web site. https://code.google.com/p/webrtc/source/browse/#svn%2Ftrunk%2Fwebrtc%2Fmodules%2Faudio_coding%2Fcodecs%2Fisac, Sep. 2013.
18. Huang YF, Tang S, Bao C. Steganalysis of compressed speech to detect covert voice over Internet protocol channels. *IET Information Security* 2011; **5** (1) : 26-32.
19. Liu Q, Sung AH, Qiao M. Temporal derivative-based spectrum and mel-cepstrum audio steganalysis. *IEEE Transactions on Information Forensics and Security* 2009; **4** (3) : 359-368.
20. Huang YF, Tang S, Zhang Y. Detection of covert voice over Internet protocol communications using sliding window-based steganalysis. *IET Communications* 2011; **5** (7) : 929-936.
21. Huang YF, Yuan J. Key distribution in the covert communication based on VoIP. *Chinese Journal of Electronics* 2011; **20** (2) : 357-361.

TABLE I

PESQ STATISTICS AT FULL EMBEDDING BIT-RATE

Audio Clips	Proposed Algorithm				Normal Codec				% Change in PESQ			
	Average	Min	Max	Standard Deviation	Average	Min	Max	Standard Deviation	Average	Min	Max	Standard Deviation
CM	4.12	3.80	4.22	0.10	4.13	3.90	4.24	0.10	-3.32	-6.85	0.82	1.33
CW	3.99	3.62	4.22	0.12	4.04	3.71	4.23	0.12	-5.18	-10.76	0.61	1.67
EM	4.06	3.62	4.30	0.11	4.07	3.72	4.22	0.11	-1.86	-4.35	1.10	0.96
EW	4.00	3.74	4.22	0.10	4.05	3.67	4.29	0.11	-5.79	-10.77	0.28	1.89

TABLE II

COMPARISIONS OF CHANGES IN PESQ AND DATA HIDING CAPACITIES BETWEEN THE PROPOESD STEGANOGRAPHY ALGORITHM AND OTHER ALGORITHM [15]

Audio Clips	% Change in PESQ							
	Proposed Algorithm				Algorithm Presented in [15]			
	Average	Min	Max	Standard Deviation	Average	Min	Max	Standard Deviation
CM	-3.32	-6.85	0.82	1.33	-2.58	-6.90	0.88	1.24
CW	-5.18	-10.76	0.61	1.67	-4.36	-10.04	0.54	1.53
EM	-1.86	-4.35	1.10	0.96	-1.00	-3.35	1.02	0.80
EW	-5.79	-10.77	0.28	1.89	-5.20	-10.40	-0.33	1.78
	Data hiding capacity							
	12 bits/frame (400 bits/second)				4 bits/frame (133.3 bits/second)			

TABLE III

STEGANALYSIS RESULTS OF THE PROPOSED ALGORITHM USING DMFCC AT DIFFERENT DETECTION WINDOWS

Window Length (Frames)	CM (%)	CW (%)	EM (%)	EW (%)
10	49.70	52.36	50.45	49.89
20	51.38	52.24	51.29	51.77
40	52.66	54.72	55.41	53.32
80	51.64	55.28	52.97	55.14
160	51.80	55.74	53.27	54.71
320	58.29	59.51	51.92	58.18
480	59.40	60.78	57.62	60.80
640	57.65	61.54	58.26	61.73
800	58.17	60.82	59.93	59.26
1000	62.75	61.46	58.27	59.49
Average	55.33	57.45	54.94	56.42
Max	62.75	61.54	59.93	61.73
Min	49.70	52.24	50.45	49.89