文章编号:1001-5132（2008）04-0447-06

A Secure Speech Transmission System Based on Steganography

DU Cheng-tou, YAN Di-qun

(Faculty of Information Science and Technology, Ningbo University, Ningbo 315211, China)

Abstract: A digital encrypted speech transmission scheme based on audio-on-demand platform is proposed. Starting with G.729A coding algorithm being employed for secured speech, the bit stream from coder is embedded in open audio files using an improved LSB algorithm, and then output to the AOD network. The secured speech can be played back by the client from the system. Tests show that the audio quality is improved using the proposed algorithm, and the developed scheme has the better resistance to hostile attackers.

Key words: speech coding; steganography; audio on demand

CLC number: TP393 **Document code:** A

Digital speech communication has been applied in many fields. But communication between two parties over long distances has always been subject to interception. This led to the development of cryptography schemes. Cryptography schemes achieve security mainly through a process of making the speech unintelligible so that those who do not possess necessary keys cannot recover the speech. Though cryptography can hide the content of the speech, the existence of a cryptographic communication in progress can not be hidden from a third party. If the third party discovers the cryptographic communication, they might be able to decipher the speech. It can be seen that latent danger exists in cryptography schemes. The need to avoid this led to the development of steganography schemes which compensate cryptography by hiding the existence of a secret communication[1].

The secure speech transmission system based on steganography by embedding a secret speech file in a cover medium has been increasingly gaining importance in the field of information technology. By hiding the secret speech using a cover medium as a wrapper, the existence of the secret speech is concealed during transmission. Wang et al. proposes a covert communication scheme using image and text [2]. A brief survey of some representative techniques of steganography and steganalysis is presented. A novel algorithm for secure speech communication is designed in the literature[3]. First, MELP speech coding is used for secret speech. Then according to masking effect, the encoded bit-stream is embedded in the intermediate frequency point of DCT. The experimental results show that the speech which is embedded with information has transparent feature, and the proposed algorithm is strongly robust to many

Received date: 2007-09-05. JOURNAL OF NINGBO UNIVERSITY (NSEE): http://3xb.nbu.edu.cn
Foundation item: National Natural Science Foundation of China (60672070); Scientific Research Found of Ningbo University (XK0610032).
The first author's biography: DU Cheng-tou (1969-), male, Ninghai Zhejiang, senior lab master, research domain: data hiding, watermarking and speech coding. E-mail: duchengtou@nbu.edu.cn

attacks such as compression, filter and so on. But the practicality of the algorithm is not good because the original public speech is absolutely necessarily while extracting the secret information. Yang et al. develop a secure communication system using information hiding and encryption techniques [4, 5]. They apply Analysis By Synthesis (ABS) method and GSM speech coding for the secret speech. The encoded bit-stream is embedded in public speech and then transmitted through Public Switching Telephone Network (PSTN). The basic model of secret communication based on steganography has been introduced in the literature[6], and the characteristics of steganography in network environment are also analyzed. At last, the writer develops an application named voice substitution phone (VSP), but the realization of VSP is very complex.

Steganography, in general, relies on the imperfection of the human auditory and visual systems. There are many steganographic techniques for digital images as well as digital audio, and meanwhile, there are a lot of approaches to detect steganography in digital images. However, there are few methods published considering Audio-On-Demand (AOD) as a new field for applied steganography. The term "AOD" describes the transmission of audio data from a sender to a receiver using network. The receiver applies the reverse process and gets the reconstructed audio data. Many applications of AOD technology have been developed and are currently under development. For that reason, embedding secret speech in AOD communication is a very interesting task and may become subject of further studies. In this paper, we proposed a new method for secure speech transmission based on audio steganography via AOD communication.

1 Model of transmission system based on steganography

Models of transmission system based on cryptography and steganography are shown in Figure 1 and Figure 2. The main difference between two models lies in: The goal of cryptography is to maintain the secrecy of a speech, which is only designed to be read by certain people. It combines an encryption method with a secret key that determines the details of each code. The meaning of encrypted data is always obscure and this will attract the attention of attackers through its mere existence. Therefore, cryptography does not always provide safe communication. While cryptography is about concealing the content of a speech, steganography is about to concealing its existence. Concealing the transmission of a speech automatically enhances its security since third parties don't even realize that the communication has taken place. Thus it can be seen that steganography can provide more protection than cryptography.

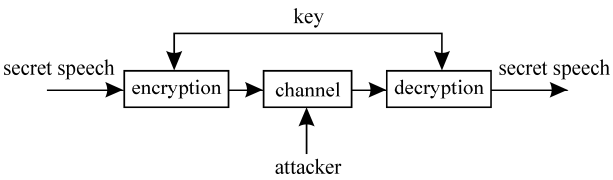


Fig.1 Model of secret speech communication based on cryptography

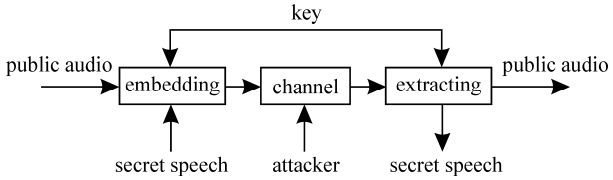


Fig.2 Model of secret speech communication based on steganography

The fundamental requirement of a secure speech transmission system based on steganography is capacity and security. For capacity, it is required to keep the amount of secret speech as much as can be embedded in a given cover audio. To obtain higher capacity, one can make use of compression techniques. Data compression is the process of encoding information using few bits than a more obvious representation would use. Data can either be lossless or loss compressed. The original data can be perfectly recovered from the lossless-compressed data. Some amount of data is lost in the loss compression and the recovered information is an

approximate version of the original data. For the security requirement, one can take be fulfilled in general implementation by either the encryption manipulation using private-key/public-key system or through pseudo-random generation using a seed key.

In this paper, we have developed a framework implemented on AOD platform for a secure speech transmission system which allows us to embed and extract a secret speech that is first compressed and then hidden into a cover audio.

2 System architecture

The proposed AOD system employs Customer/ Server (C/S) architecture to deliver real-time audio data over the network, as shown in Figure 3.

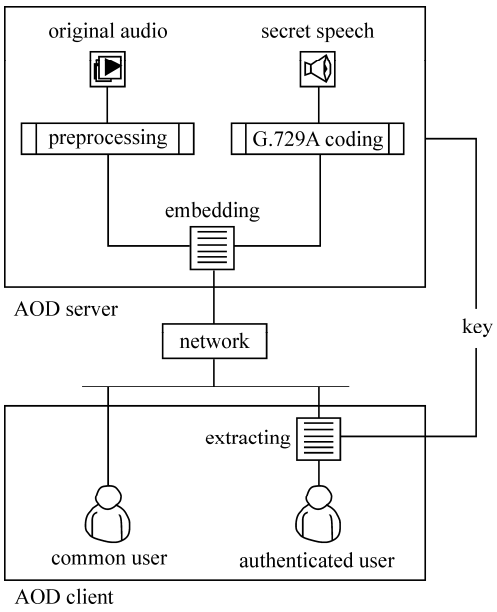


Fig.3 System architecture

There are three main components involved in this system:

(1) The main task of Sever is to fulfill preprocessing and hiding for the secret speech. And it should be responsible for delivering the audio data, controlling the flow of data, and other tasks are related to performance monitoring and control.

(2) Client contains an audio player that parses the transmitted data into a format that can be played back on the client's computer. It means that Sever and Client must have a matching codec in order to successfully transfer the audio data.

(3) As we known, TCP protocol can deliver data stream reliably in a correct order. However it costs long delays due to large packet headers and retransmissions. UDP protocol has higher efficiency because of its smaller overhead. So we make use of TCP to transmit control commands and UDP to be responsible for the data transmission between Sever and Client.

Although common users may play back the audio program from the list provided by Sever, the secret speech which has been embedded in public audio is unperceptive to them. For authenticated users, they can not only play back the public audio, but also the secret speech recovered from the public audio.

3 Principles

3.1 Secret speech coding strategy

In order to increase system's capacity, secret speech must be compressed firstly. G.729A is a low code rate speech compression standard defined by ITU-T in 1996. It is an algorithm for encoding speech signals at 8 kbit·s^{-1} using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP) technology[7]. Here are some considerations about selecting G.729A algorithm as preferred coding scheme: (1) With its low code rate, high capacity can be attained. (2) With excellent quality of the construction synthesis speech, secure communication quality can be guaranteed. (3) With strong robustness of the coding parameter, the ability to resist the attacks can be further enhanced.

3.2 Audio data hiding algorithm

Since the early 1990s, data hiding has become a hot topic attracting much attention from researchers in various fields. The appearance and development of audio data hiding has come being a new domain for information security. In recent years, the research on audio steganography which is applied to secret com-

munication system is often based on digital watermarking technology. However, most of digital watermarking algorithms cannot satisfy the requirements of secret communication system in high capacity and low complexity. For example, transform-based audio watermarking algorithm can obtain some satisfying results through taking advantage of the characteristic of human audio system (HAS)[8-10]. But high computation complexity always exists because of the inevitable transformation and inverse transformation. Meanwhile transform-based algorithm only revises partial transform coefficients to hide secret message so that the capacity of it is very small. So it is not the best choice for secure speech communication system.

Data hiding in the least significant bits (LSB) of cover samples in the time domain is one of the most common and simplest algorithms[11]. LSB method can achieve high capacity.

Chan et al proposed a data hiding scheme by optimal pixel adjustment process (OPAP)[12]. Using the proposed algorithm, the WMSE between the cover-image and the stego-image is shown to be less than 1/2 of that obtained by the simple LSB substitution method. This algorithm is designed for cover-image. In order to using it in cover-audio, the algorithm is modified and redesigned in this paper. The basic principle of the proposed audio data hiding algorithm is described as follows:

Let S be the original l-bit cover audio with N samples represented as

$$S = \{s_i \mid 1 \le i \le N, s_i \in \{0,1,\cdots,2^l - 1\}\}. \quad (1)$$

And $s_i^{'}$, $s_i^{''}$ are the corresponding sample values of the i sample in the stego-audio obtained by the direct replacement of the k least significant bits of s_i with k message bits and the refined cover-audio. Let $\delta_i = s_i^{'} - s_i$ be the embedding error between $s_i^{'}$ and s_i. Therefore,

$$0 < |\delta_i| < 2^k. \quad (2)$$

Based on the value of δ_i, the proposed algorithm can be described as follows:

Case 1 ($2^{k-1} < \delta_i < 2^k$): If $s_i^{'} \ge 2^k$, then $s_i^{''} = s_i^{'} - 2^k$; otherwise $s_i^{''} = s_i^{'}$;

Case 2 ($-2^{k-1} \le \delta_i \le 2^{k-1}$): $s_i^{''} = s_i^{'}$;

Case 3 ($-2^k < \delta_i < -2^{k-1}$): If $s_i^{'} < 2^l - 2^k$, then $s_i^{''} = s_i^{'} + 2^k$; otherwise $s_i^{''} = s_i^{'}$.

Let $\delta_i^{'} = s_i^{''} - s_i$ be the embedding error between $s_i^{''}$ and s_i. It can be seen that the absolute embedding error between audio samples in the cover-audio and the stego-audio obtained after the proposed algorithm is limited to

$$0 \le |\delta_i^{'}| \le 2^{k-1}. \quad (3)$$

So the audio quality of the stego-audio is greatly improved with low computational complexity.

4 Results

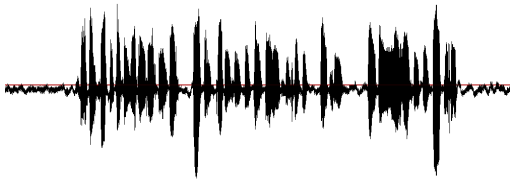
We have developed an application for a secure speech transmission system in Windows environment. The system includes a Sever and a Client. The Sever allows the encoding and hiding of secret speech in cover audio. The encoding and decoding algorithms for secret speech are G.729A algorithm. And audio data hiding algorithm is based on LSB algorithm with OPAP. The Client consists of a friendly user interface that allows the user to select and play back different audio file from the program list. For authenticated user, he can hear the secret speech which is recovered from public audio file. The cover-audio and stego-audio are shown in Figure.4. From Figure 4, it can be seen that the audio quality of stego-audio obtained by our algorithm is better than that by simple LSB substitution method. The waveforms of original and recovered secret speech are shown in Figure 5. From Figure 5, we can find that there are some differences between the original and recovered secret speech in the waveform's profile. This is because G.729A belongs to parametric audio coding algorithm. Subjective experiments show that the naturalness and intelligibility of the reconstructed speech are all acceptable.

(a) Original cover-audio (44.1 kHz, 16 bits/sample)

(b) Stego-audio by simple LSB algorithm (4 LSBs, PSNR=74.01)

(c) Stego-audio by the proposed algorithm (4 LSBs, PSNR=76.98)

Fig.4 Waveform of cover-audio and stego-audio



(a) Original secret speech

(b) Recovered secret speech

Fig.5 Waveform of original and recovered secret speech

5 Conclusions

A novel secure speech communication system that is different from the traditional method has been designed and developed. The secret speech is compressed and embedded in a public audio, and then transmitted through AOD communication. It is very easy to escape the attack of illegal eavesdropper. The experimental results show that the proposed system is secure, practical and with high capacity.

We are currently doing more research on improving LSB data hiding algorithm which can effectively resist against various attacks. How to evaluate the security of the proposed system is also a challenge in our future work.

References:

[1] Bender W, Gruhl D, Morimoto N. Techniques for data hiding[J]. IBM Systems Journal, 1996, 35(3/4):313-336.

[2] Wang S Z, Chao C, Zhang X P. Undercover communication using image and text as disguise and countermeasures[J]. Journal of Shanghai University (English Edition), 2006, 10(1):33-34.

[3] Chen L, Zhang XW. Study of information hiding in security speech communication[J]. Journal of PLA University of Science and Technology: Natural Science, 2002, 3(6):1-5.

[4] Yang W, Yang Y X. Secrecy speech communication system based on steganography techniques[J]. Journal of China Institute of Communications, 2004, 25(2):75-81.

[5] Yue J Q, Niu X X, Yang Y X. Information hiding technology of voice encryption communication[J]. Journal of Beijing University of Posts and Telecommunications, 2002, 25(1):79-82.

[6] Wang J J, Zhang X M. Research on audio information hiding technology in network communication[J]. Journal of Beijing Institute of Petrochemical Technology, 2005, 13(4):54-59.

[7] ITU-T. Recommendation G.729 annex a, reduced complexity 8kbit/s CS-ACELP speech codec[S]. 1996.

[8] Wu S Q, Huang J W, Huang D R. DWT-based audio watermarking with self-synchronization[J]. Journal of Computer, 2004, 27(3):365-370.

[9] Wang Q S, Sun S H. A novel algorithm for embedding watermarks into digital audio signals[J]. ACTA Acustica, 2001, 26(5):464-467.

[10] Kirovski D, Malvar H. Spread-spectrum watermarking of audio signal[J]. IEEE Transactions on Signal Processing, 2003, 51(4):1 020-1 033.

[11] Amin M M, Salleh M, lbrahim S, et al. Information hiding using steganography[C]//In Proceedings of the 4th NCTT, 2003:21-25.

[12] Chan C K, Cheng L M. Hiding data in images by simple LSB substitution[J]. Pattern Recognition, 2004, 37(3): 469-474.

基于隐写算法的保密语音传输系统

杜呈透，严迪群

（宁波大学 信息科学与工程学院，浙江 宁波 315211）

摘要：提出了一种基于隐写算法的保密语音传输系统，首先采用 G.729A 编码算法对保密语音进行低码率压缩编码，然后通过改进 LSB 隐写算法将保密语音码流隐藏到公开音频中，并利用音频点播平台发布到网络上，最后通过客户端点播实现保密语音的提取和回放. 测试数据结果表明：通过该改进算法，载体音频的感知质量得到了显著的提高，并且系统对于恶意攻击者具有更好的隐秘性.

关键词：语音编码；隐写；音频点播

中图分类号：TP393 文献标识码：A

（责任编辑 章践立）