A Multi-Layer Steganographic method Based on Audio Time Domain Segmented and Network Steganography

Peng-Fei Xue[a], Han-Lin Liu, Jing-Song Hu, Rong-Gui Hu
Electronic Engineering Institute, National University of Defense Technology, Hefei 230037, China
[a]leorick092182@163.com

***Abstract*—Both audio steganography and network steganography are belong to modern steganography. Audio steganography has a large capacity. Network steganography is difficult to detect or track. In this paper, a multi-layer steganographic method based on the collaboration of them (MLS-ATDSS&NS) is proposed. MLS-ATDSS&NS is realized in two covert layers (audio steganography layer and network steganography layer) by two steps. A new audio time domain segmented steganography (ATDSS) method is proposed in step 1, and the collaboration method of ATDSS and NS is proposed in step 2. The experimental results showed that the advantage of MLS-ATDSS&NS over others is better trade-off between capacity, anti-detectability and robustness, that means higher steganographic capacity, better anti-detectability and stronger robustness.**

Keywords: Information Hiding; Audio Steganography; Network Steganography; Multi-Layer Steganography

I. INTRODUCTION

With the rapid development of information technology, there is an urgent problem to be solved, that is how to transmit information securely and effectively. The traditional secure communication technology is to encrypt the information. Its security mainly depends on the complexity of the encryption algorithm and the length of the key. But with the continuous improvement of computing power, especially the emergence of supercomputing clusters, it fails to guarantee the security only by increasing the key's length. As a novel way of covert communication, information hiding technology provides an effective way.

In the field of information hiding, there is a very important aspect, called steganography. Steganography is a technique that embeds secret information into a carrier for the purpose of covert communication. Up to now, modern steganography mainly include several types such as Network steganography, digital media steganography, etc.

Among these types, digital media steganography appears earlier, which is an important branch of modern steganography. Classifing digital media steganography by the carrier, it mainly includes image steganography, audio steganography and video steganography. The goal of image steganography is to deceive the Human Visual System (HVS). Similarly, the goal of audio steganography is to deceive the Human Auditory System (HAS). The HAS is more sensitive than the HVS in distinguishing minor distortions. Therefore, it is more difficult to hide the secret information in the audio signal. However, due to the large redundant of an audio file, the steganographic capacity of audio steganography is higher than that of image steganography. While steganography is developing, steganalysis is evolving too. Until now, there are so many steganalysis techniques to detect audio steganography. Yang proposed a detection method based on HCF Statistical Features which is called MIDI steganalysis [1] to detect audio LSB steganography. Xie proposed a detection method based on PN sequence estimation [2] to analyse audio spread spectrum steganography. A blind detection algorithm [3] proposed by Wang is to detect audio transform domain steganography. All these detection methods make the audio-based covert communication unsafe if only a single audio steganographic method is used.

Another important branch of modern steganography is network steganography. Network steganography arose since about 20 years ago. Compared with digital media steganography, the superiority of network steganography is stronger anti-detectability. But the disadvantage of it is lower capacity.

In order to improve the security of covert communication, a new multi-layer steganographic method (MLS-ATDSS&NS) is proposed in this paper, which is based on the collaboration of audio time domain segmented steganography (ATDSS) and network steganography (NS). The new audio steganography which called ATDSS is proposed firstly. Then, the collaboration method of ATDSS and NS is described. At last, several experiments is conducted. The experimental results showed that the proposed method MLS-ATDSS&NS gives a good performance in capacity, anti-detectability and robustness.

This paper is organized as follows. The related work is introduced in Section II, including audio steganography, network steganography, and multi-layer steganography. In section III, the proposed methods are presented. Experimental results and analysis are presented in section IV, followed by a conclusion in section V.

II. Related Works

A. Audio Steganography

Bender's research in 1996 showed that the HAS has a large spectral range, and its differential range is quite small [4]. Therefore, the high volume sound will cover the low volume sound. Apart from this, the HAS can senses the relative phase, but it can't perceive the absolute phase. These two limitations are the basis of audio steganography. In addition to the above two limitations, the advantage of taking an audio file as a carrier is that there are a large data redundancy, which means it can be embedded in a large amount of information. Some studies showed that all the works of Shakespeare could be embedded in an 8 minutes song. The audio been embedded in secret information is called stego audio. Audio steganography mainly includes the time domain steganography and the transform domain steganography according to the embedded domain. Time domain steganography contains many methods, such as the least significant bit (LSB) method, echo concealment (EC) method, phase encoding (PE) method and spread spectrum (SS) method, etc.

The LSB method is a simple but most widely used steganographic method. It directly replaces one least significant bit of a carrier with one secret information bit. The advantages of the LSB method are large embedding capacity, implemental ease, and strong anti-detectability. But it defect is weak robustness. The stego audio will be severely damaged after a simple signal processing, so that it is difficult to extract the complete secret information. But if the robustness of the LSB method is improved, its anti-detectability will be reduced. Many researchers were constantly looking for a trade-off between robustness and anti-detectability [4, 5, 6, 7, 8]. But there was no effective way to solve this problem.

The EC method utilizes the time domain masking effect of the HAS. It embeds the data covertly by changing the initial amplitude, decay rate and offset of the echo [4, 9, 10, 11]. The PE method utilizes the characteristic that HAS is sensitive to relative phase but not sensitive to absolute phase [13, 14]. The SS method is to spread the secret information in spectrum as many as possible [15]. This method has stronger robustness and better anti-detectability than the other three time domain steganographic methods.

The advantages of the time domain steganography are implemental ease and large capacity. But its flaw is weak robustness so that it can't resist the attack effectively on the carrier. In order to improve the robustness of the time domain steganography, many researchers began to pay attention to the transform domain steganography. There are 3 commonly used transformations, that is discrete Fourier transform (DFT), discrete cosine transform (DCT), and discrete wavelet transform (DWT).

DFT is a classical and effective mathematical tool which can be used in steganography [18, 19]. DCT has a good energy aggregation characteristic, and it is easy to achieve in the digital signal processor [20, 21]. But now, DWT has become one of the most important research directions of digital media steganographic method, especially in audio steganography [23, 24]. It gives a good localized performance both on the time domain and the frequency domain.

B. Network Steganography

The concept of network steganography was first proposed by Szczypiorski [26]. After that, steganographic carrier can no longer be digital media, but may be other types, such as computer network data packets, frames, etc. Nowadays, network steganography has become a new important branch of information hiding. Over the last decade, a large number of network steganographic methods appeared explosively. Most of them is on the basis of redundancy. According to classification based on steganography pattern proposed by Wendzel et al. [27], network steganography can be divided into four categories: storage method, timing method, hybrid method, and transform domain method. Among them, storage method and timing method are the most widely used methods.

The storage method can hide the secret information both in the user data (payload) and in the protocol field (non-payload), such as the structure of PDU is modified [28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40], and the structure of PDU is remained except the header fields [41, 42, 31, 43].

The timing method can encode the secret information by changing the sending rate of packet [44, 45, 46, 47], inter-packet delay [48, 49, 50, 51], and changing the order of packet [52, 53, 54, 55], etc.

Seo pointed out that network steganography has three advantages over digital media steganography [56]. Firstly, network steganography has better anti-detectability. So it is suitable as a covert channel. Secondly, the secret information life cycle is shorter, which means not easy to leave traces. Thirdly, the bandwidth of network steganography is more flexible and not limited to the size of the carrier file. Network steganography can take many different forms, at different layers of the Internet stack by adopting different embedding schemes. Seo argues that there is no clear countermeasure or a one-to-many solution that can effectively detect all network steganography covert channels [56]. Network steganography has become an effective way to transmit secret information securely and effectively.

C. Multi-Level Steganography

Multi-level steganography method was first proposed by Al-Najjar et al. [57]. They embedded a bait picture into the LSB of the carrier, while the real secret information was hidden in the LSB of the bait picture to achieve multi-level steganography. Fraczek pointed out that multi-level steganography is a new method of deep hiding technology [58]. Multi-level steganography uses features of an existing steganographic method (the upper-level method) to create a new one (the lower-level method). This method will enhance the anti- detectability of covert communication. Samir extended the concept of multi-level steganography and applied it to audio steganography [59]. He proposed a method that can transmit two kinds of secret information in a single carrier at the same time. In the first level, the first secret information was embedded in the carrier file. Then he got a file containing the first secret information. In the second level, the file generated in the first level was used as a carrier. The second secret information was embedded in it. At last, there was a file containing two kinds of secret information.

III. The Proposed Method

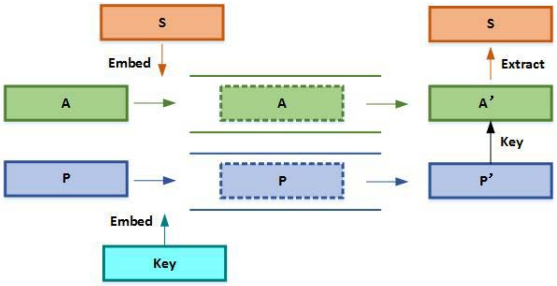


Figure 1. the basic idea of the proposed multi-layer steganographic method

As mentioned in section II, audio steganography has a large capacity while network steganography has a good anti-detectability and is difficult to be detected or tracked. Based on the idea of multi-level steganography and expanding it, a multi-layer steganographic method based on the collaboration of both the two steganography methods is proposed. As Fig. 1 shows, the proposed method transmit the secret information in two covert layers. One is in the audio steganography layer, the other is in the network steganography layer. The content being transmitted in the audio steganography layer is the audio file *A*' which contain a secret information *S*. Meanwhile the content being transmitted in the network steganography layer is data packet *P*' which contain key. The key is used to extract the secret information from *A*'.

A. Audio Time Domain Segmented Steganography (ATDSS)

As described in section II, two main branches of audio steganography are the time domain method and the transform domain method. Therefore, the segmented steganographic method should also contain the time domain segmented steganography and the transform domain segmented steganography. In this paper, we mainly focus on the audio time domain segmented steganography (ATDSS).

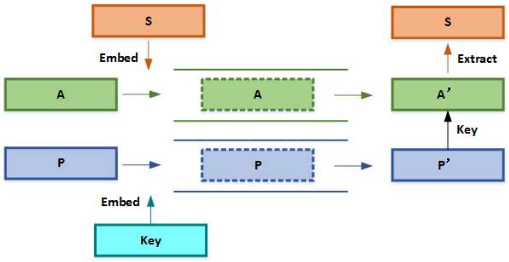


Figure 2. the basic idea of ATDSS

The basic idea of the ATDSS is shown in Fig. 2. The carrier is divided into several segments on the time domain and the information is also divided into several parts with the same amount. The information parts are embedded into carrier segments one by one and the imbedding methods are different from each other. As Fig. 2 shows, the carrier is divided into 3 segments. The embedding process can be described as follows.

1. Divide the carrier *A* into 3 segments, i.e. A_1, A_2, A_3
2. Divide the secret information *S* into 3 segments, i.e. S_1, S_2, S_3.
3. Embed S_1 into A_1 with method R_1, Embed S_2 into A_2 with method R_2, Embed S_3 into A_3 with method R_3. Then each segment contains one of the three parts of the secret information.
4. Reassemble the 3 segments together and get a new audio file *A*'.
5. Transmit this new audio file *A*' which contains all the secret information to the covert receiver

The covert sender and the covert receiver should agree on the segmentation point and the steganographic method used for each segment in advance. After the covert receiver successfully receiving the audio file *A*', he extracts the secret information from *A*' according to the agreement.

B. The Collaboration Method of ATDSS and NS

In Section III, the ATDSS method has been described. If the covert receiver wants to extract the secret information successfully, he must know the segmentation point and the steganographic method used for each segment exactly. Here the segmentation point and the steganographic method used for each segment is called the key. It is a critical issue to study that how to transmit the key from the covert sender to the covert receiver. One way is to embed the key into the same carrier which the steganography has been embedded. But it will reduce the effective steganographic capacity. There is another way, that is, both sides of the covert channel guarantee that they hold the same key before each communication. But it will bring additional unnecessary overhead and increase the risk of key disclosure.

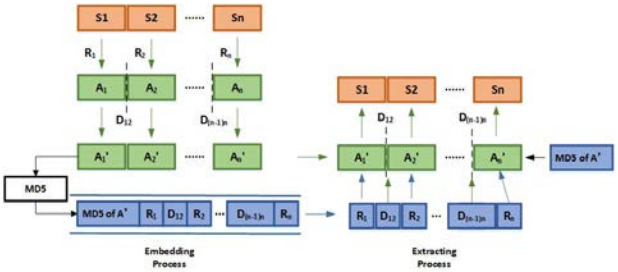


Figure 3. the basic idea of the collaboration method of ATDSS and network steganography

To solve the problem of key delivery, we propose a multi-layer steganographic method based on the collaboration of audio steganography and network steganography. The secret information is transmitted in audio steganography layer. Meanwhile the key is delivered in network steganography layer.

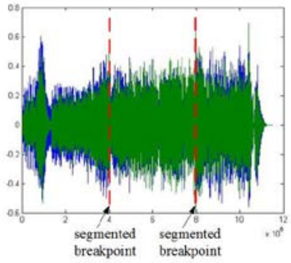


Figure 4. example of dividing the carrier directly

Suppose that 3 audio steganographic methods, e.g. M_1, M_2, M_3 is selected. The basic idea of the proposed method is shown in Fig. 3. Since the secret information and the key are transmitted in 2 channels, the covert receiver may receive some *A*'s and some keys at the same time. If he wants to extract the secret information successfully, he must find the right key to match an *A*'. Here the MD5 message digest is used as a unique identifier to mark the *A*'s. It's worth noting that this article is not studying some kind of audio steganographic method, but focus on the collaboration mechanism between audio steganography and network steganography. So the 3 selected audio steganographic methods has been proposed by some other reseachers.

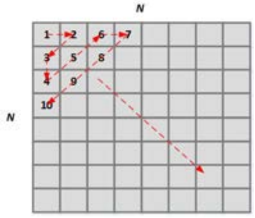


Figure 5. Zigzag Scan

The proposed method has two processes. One is embedding process, the other is extracting process (shown in Fig. 3). Suppose that the size of the secret information is S_V, and the capacity of the carrier is *C*. Generally, $S_V < C$. The Multi-Layer Steganographic Method Based on Collaboration of Audio Time Domain Segmented Steganography and Network Steganography (MLS-ATDSS&NS) are described in Algorithm 1.

Algorithm 1. MLS-ATDSS&NS

```
Input: the audio file $A$ as a carrier, the secret information $S$, 3 kinds of steganographic method $M_1$, $M_2$, $M_3$
Output: the audio file $A$' containing the secret, the key to extract secret information from $A$
1: Divide the carrier $A$ into n segments with equal size. Mark each segment as $A_i$ ($i$=1, 2,...$n$). Record the segmented breakpoint location as $D_{12}$, $D_{23}$, $D_{34}$,..., $D_{(n-1)n}$. Divide the secret information $S$ into n segments with equal size. Mark each segment as $S_i$ ($i$=1, 2,...$n$).
2: Embed the $i$th segment of secret information $S_i$ into the ith segment of carrier $A_i$ with method $R_i$. $R_i$ is defined by $R_i$ = random($M_1$, $M_2$, $M_3$). Then, get $n$ segments of the audio file. Mark each segment as $A'_i$ ($i$=1, 2,...$n$).
3: Reassemble the $n$ segments of the audio file and get $A$' containing the whole secret information $S$. Send $A$' to the covert receiver.
4: Calculate the MD5 message digest of $A$'.
5: Generate the key. The key is a sequence defined as (MD5, $R_1$, $D_{12}$, $R_2$, $D_{23}$,..., $R_{n-1}$, $D_{(n-1)n}$, $R_n$).
6: Covertly transmit the key to the covert receiver using network steganography.
7: The covert receiver gets the key and $A$'. At first, extract the MD5 meaasge digest from the key and use it to match $A$'. Secondly, divide $A$' into n segments again with the segmented breakpoint location $D_{12}$, $D_{23}$, $D_{34}$,..., $D_{(n-1)n}$. Thirdly, extract $i$th part of secret information $S_i$ from $i$th segement $A'_i$ according the steganographic method $R_i$. At last, reassemble all the $S_i$ and get the secret information $S$.
```

It is important to note that the segmentation of the carrier on the time domain does not mean to segment the audio file directly (shown in Fig. 4). The correct way is to convert it into an array and then divide the array into several segments. The advantage of this segmentation method is that the segmented breakpoint location can be accurately recorded and the secret information can be extracted easily.

Generally speaking, the secret information is a text or an image. If we need covertly transmit a text, we convert it into 1-dimensional array directly. If we need to covertly transmit an image, we should firstly transform it from 2-dimensional data to 1-dimensional data. In this paper, the Zigzag Scan is used to perform it (as shown in Fig. 5).

In Algorithm 1, the key is a sequence defined as

$$(\mathrm{MD5}, R_1, D_{12}, R_2, D_{23}, ..., R_{n-1}, D_{(n-1)n}, R_n) \quad (1)$$

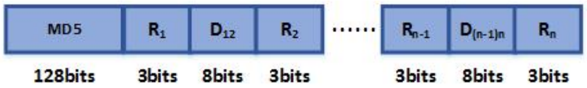


Figure 3. the format of key

The size of MD5 message digest is 128 bits. We use 8 bits to record the segmented breakpoint location and 3 bits to represent the steganographic method. Then the format of key is shown in Fig. 6. As the number of segments increases, the length of key increases.

IV. Experimental Results and Analyses

The proposed method has been achieved mainly by using Matlab platform and Scapy, which will be introduced in the following part. First, a feasibility verification is conducted. Second, the capacity, anti-detectability, and robustness of the proposed method is analyzed.

A. Experiment Environment

The experiment environment is as follows.

a. Network environment: 100Mb/s switched Ethernet, LAN, two hosts (One is the covert sender, which IP address is 192.168.1.17. The other is the covert receiver, which IP address is 192.168.1.18).

b. Host configuration: Windows 7 64-bit systems, Intel Core i5-2300 CPU, 8G memory, 100Mbps NIC.

c. Tools: Microsoft Visual C++ 6.0, Matlab R2013b , WinPcap development kit, Python 2.6.3, Scapy2.3.1.

d. The secret information: 255*255 grayscale image, the name is 'lena.bmp'.

e. The carrier: 4 minutes and 20 seconds of a two-channel stereo music, sampling frequency 44.1KHz, 16bit quantication, the name is 'music.wav'.

f. Evaluation indictors selection:

Bit Error Rate (BER): BER is used to measure the percentage of the number of different bits between the extracting secret information and the original secret information after the carrier having been attacked. The smaller the BER is, the better the robustness will be.

Normalized Coefficient (NC): NC is used to measure the similarity between the extracting secret information and the original secret information after the carrier has been attacked. The bigger the NC is, the better the robustness will be.

Peak Signal to Noise Ratio (PSNR): PSNR is used to measure the anti-detectability of the carrier. The higher the PSNR is, the better the anti-detectability of the carrier will be.

B. *Feasibility Verification*

Feasibility verification of the proposed method is carried out first. In the audio steganography layer, 3 audio steganographic methods have been chosen as alternative methods, namely simple LSB method, DWT-based method and RLE & IWT-based method. As described in section 2.1, simple LSB method is the earliest and most commonly used steganographic method. The advantages of this method are its large capacity, implemental ease, and good anti-detectability. But its robustness is poor. DWT-based method improve simple LSB method by replacing LSB of wavelet coefficients with secret information. The robustness of DWT-based method is better than that of simple LSB method. RLE & IWT-based method was a new audio steganographic method proposed by Liu in 2017 [60]. In the network steganography layer, we select TMPCNSM as network steganographic method. This method was proposed by Xue in 2017 [61]. Its steganographic bandwidth is 3.79 bits/packet.

Two hosts in the LAN are used to do the experiment for 100 times. The result is recorded in Table I. In 50 experiments among them, the carrier is divided into 2 segments on the time domain (n = 2). We use simple LSB method and DWT-based method respectively to embed part of secret information in 2 segments. In the other 50 experiments, the carrier is divided into 3 segments on the time domain (n = 3). We use 3 methods (DWT-based method, RLE & IWT-based method, simple LSB method) respectively to embed part of secret information in 3 segments.

TABLE I. THE RESULT OF FEASIBILITY VERIFICATION

	Times	**Methods**	**Average Traffic**	**Extraction BER(%)**
n = 2	50	simple LSB, DWT	142 bits	0
n = 3	50	DWT, RLE & IWT, simple LSB	153 bits	0

Among the 100 experiments, the secret information can be successfully embedded in the carrier. Regardless of the factors such as external attack or interference, the BER of the extracted secret information is 0.

Through the feasibility verification, it is proved that the proposed method can realize the ATDSS in audio steganography layer and covert transmission in network steganography layer. The covert receiver can successfully extract the secret information after receiving the contents transmitted from the audio covert channel and the network covert channel respectively. Among the 100 experiments, the secret information can be successfully embedded in the carrier. Regardless of the factors such as external attack or interference, the BER of the extracted secret information is 0.

C. *Capacity Analyses*

The steganographic capacity is also referred to as data embedding payload, which measures the amount of secret information that can be hidden in carrier per unit length. It is usually represented by bit rate (bit per-second, bps), that is, the bits number of the secret information that can be embedded in carrier per second.

The audio steganographic method of modifying the LSB on the audio time domain is mainly to embed the secret information in the LSB of the time domain signal or other bits that do not affect the audio anti-detectability. While audio steganography based on wavelet domain mostly embeds the secret information into the wavelet coefficients in a one-to-one relationship. The number of wavelet coefficients determines the embedding capacity. Generally speaking, the steganographic capacity of previous method is greater than that of the latter one. As shown in Table II, the steganographic capacity of simple LSB method is greater than that of all other methods.

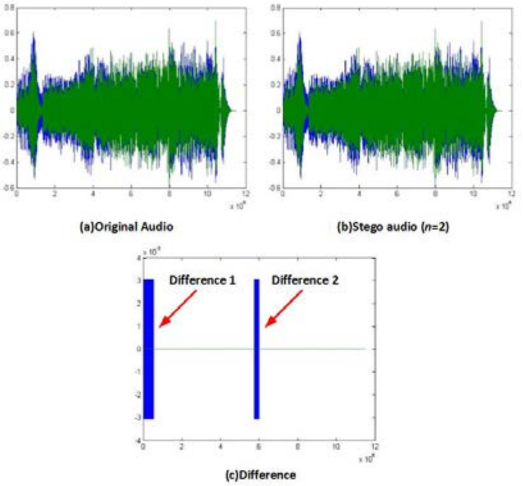
There are 3 important aspects to evaluate a steganographic method, i.e. the steganographic capacity, the anti-detectability and the robustness. There is a trade-off between these three aspects. High capacity leas to poor anti-detectability. Therefore, the anti-detectability of the transform domain based methods is superior to that of the time domain based methods. As Table II indicates, no matter n = 2 or n = 3 is used, the steganographic capacity of the proposed method is greater than that of DWT-based method and RLE & IWT-based method, but less than that of simple LSB method. It means that the proposed method has a certain capacity. Meanwhile, its anti-detectability keeps on a good level. More analysis of anti-detectability will be described in the next section.

TABLE II. STEGANOGRAPHIC CAPACITY

	simple LSB	**DWT**	**RLE & IWT**	**Our method (n = 2)**	**Our method (n = 3)**
Capacity (bps)	86767	19500	22940	53133	43069

In network steganography layer, the steganographic bandwidth of TMPCNSM method is 3.79bits/packet, which is higher than that of most other network steganographic methods. More information about TMPCNSM can be referred in [61].

D. Feasibility Verification



(a)Original Audio

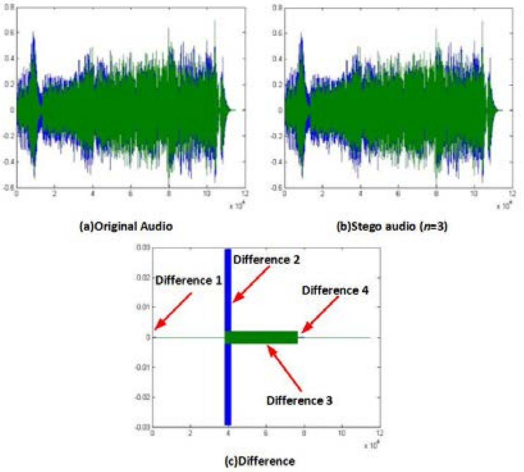
(b)Stego audio (*n*=2)

(c)Difference

Figure 4. original audio compare with stego audio(*n* = 2)

In the image steganography, the change of the image can be observed to intuitively estimate the anti-detectability of the steganographic method. Similarly, in the audio steganography, the waveforms before and after steganography can be compared for anti-detectability evaluation with the assistance of HAS. Fig. 7 and Fig. 8 show the difference between the original audio and the stego audio (audio embedded secret information).

As Fig. 7 shows, when *n* = 2, there are 2 differences (Fig. 7(c) shows). It is mainly due to the fact that LSB method is used to embed the secret information at the beginning of each segment causing a change in the sampling point. The rest part of carrier is not embedded secret information, so the sampling point does not change, and there is no difference.



(a)Original Audio

(b)Stego audio (*n*=3)

(c)Difference

Figure 5. original audio compare with stego audio(*n* = 3)

In Fig. 8(c), there are 4 differences between original audio and stego audio. The first and forth tiny difference are due to embed secret information using LSB. The second difference is caused by DWT-based method. The third difference is due to the integer normalization of the coefficients before the integer wavelet transformation.

TABLE III. PSNR OF CARRIER

	simple LSB	DWT	RLE & IWT	Our method (*n* = 2)	Our method (*n* = 3)
PSNR	106.59	103.59	48.83	104.83	53.45

In addition to the use of subjective way for anti-detectability evaluation, some indicators can also be used to analyze it objectively. There are many indicators for evaluating anti-detectability, such as MOS, AB method, and SNR, etc. In this paper, PSNR is used as evaluating indicator. When the carrier embeds the secret information, the higher the PSNR of the carrier is, the better the anti- detectability of the steganographic method will be. Table III shows the result of PSNR. When *n* = 2, the PSNR of the proposed method is between simple LSB method and DWT-based method. Combined with Table II, it shows that the anti-detectability of the proposed method does not significantly decline, but the capacity increases. When *n* = 3, the PSNR of the proposed method is higher than that of RLE & IWT-based method. The capacity of the proposed method is also higher than that of RLE & IWT-based method. Comparing to the RLE & IWT-based method [60], both robustness and anti-detection of the proposed method are improved.

Figure 6. the comparison of the original and extracted information without attack

E. Robustness Analyses

Robustness reflects the ability to resist attack. Usually, BER and NC are used to measure robustness. The lower the BER is, the stronger the robustness of the steganographic method will be. The greater the NC is, the stronger the robustness of the steganographic method will be. In addition to using BER and NC to measure robustness, the PSNR of secret information extracted from the carrier is calculated, which can also evaluate robustness. When the stego audio is not attacked, the extracted information is exactly the same as the original information (Fig. 9). The BER is 0, and the NC is 1. Then, 5 common attacks are selected to test the robustness of the steganographic method, i.e. the lift quantization, the reduced quantization, the up-sampling, the down-sampling, and the noise attack.

1) Attack 1: Lift Quantization

The lift quantization of the 16-bit stego audio is carried out, that is, lift it to 32-bit and back to 16-bit. When *n* = 2, the extracted information is shown in Fig. 10(b). When *n* = 3, the extracted information is shown in Fig. 10(c). Table IV shows the results of indicators after the lift quantization.

TABLE IV. THE RESULT AFTER LIFT QUANTIZATION

	simple LSB	**DWT**	**RLE & IWT**	**Our method (*n* = 2)**	**Our method (*n* = 3)**
BER(%)	0	0	0	0	0
NC	1	1	1	1	1
PSNR	96.15	96.15	96.15	96.15	96.15

(a)Original Secret Information
(b)Extracted Information (n=2)
(c)Extracted Information (n=3)

Figure 7. the comparison of the original and extracted information after lift quantization

As shown in Fig. 10 and Table IV, the proposed method can resist the lift quantization attack. No matter *n* = 2 or *n* = 3, the extracted information can be correctly recognized.

2) Attack 2: Reduced Quantization

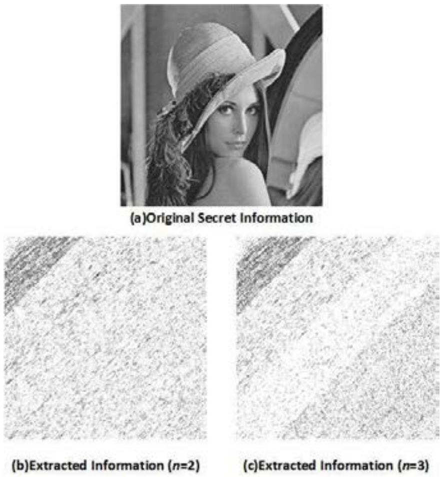

(a)Original Secret Information
(b)Extracted Information (n=2)
(c)Extracted Information (n=3)

Figure 8. the comparison of the original and extracted information after reduced quantization

The reduced quantization of the 16-bit stego audio is carried out, that is, reduce it to 8-bit and back to 16-bit. When *n* = 2, the extracted information is shown in Fig. 11(b). When *n* = 3, the extracted information is shown in Fig. 11(c). Table V shows the results of indicators after reduced quantization.

TABLE V. THE RESULT AFTER REDUCED QUANTIZATION

	simple LSB	DWT	RLE & IWT	Our method (*n* = 2)	Our method (*n* = 3)
BER(%)	55.23	54.78	53.45	50.87	53.86
NC	0.8734	0.8832	0.9046	0.9127	0.9008
PSNR	5.7237	5.7681	5.842	5.8472	5.8876

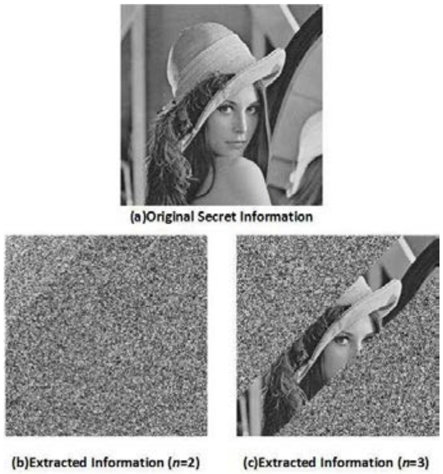
As shown in Fig. 11, the proposed method can't resist the reduced quantization attack. No matter *n* = 2 or *n* = 3, the extracted information cannot be recognized. But in Table V, the BER of the proposed method is lower than that of other method. Meanwhile the NC and PSNR is higher than others. It proves that the proposed method increases robustness.

3) Attack 3: Up-Sampling

The stego audio is up-sampled from original sampling rate 44.1kHz to 88.2kHz and then down-sampled it to 44.1kHz again. When *n* = 2, the extracted information is shown in Fig. 12(b). When *n* = 3, the extracted information is shown in Fig. 12(c). Table VI shows the results of indicators after up-sampling.

TABLE VI. THE RESULT AFTER UP-SAMPLING

	simple LSB	DWT	RLE & IWT	Our method (*n* = 2)	Our method (*n* = 3)
BER(%)	50.07	50.12	9.7	49.9	34.38
NC	0.806	0.8121	0.954	0.8061	0.7768
PSNR	9.01	9.03	35.17	9.05	18.06



(a)Original Secret Information

(b)Extracted Information (*n*=2)

(c)Extracted Information (*n*=3)

Figure 9. the comparison of the original and extracted information after up-sampling

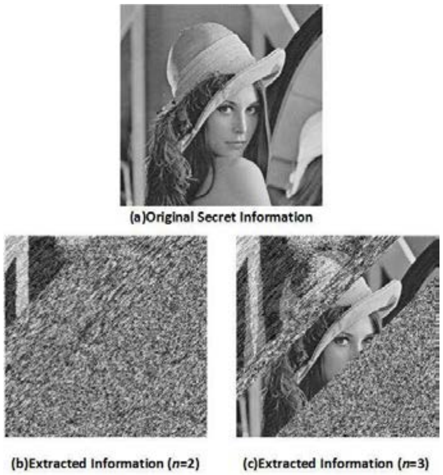
As shown in Fig. 12, when *n* = 2, the proposed method can't resist the up-sampling attack. When *n* = 3, the proposed method can partially resist it. That is, a part of the secret information can be extracted from the stego audio after being attacked. As shown in Table VI, with *n* = 3, the robustness of the proposed method is better than that with *n* = 2.

4) Attack 4: Down-Sampling

The stego audio is down-sampled from original sampling rate 44.1kHz to 22.05kHz and then up-sampled it to 44.1kHz again. When *n* = 2, the extracted information is shown in Fig. 13(b). When *n* = 3, the extracted information is shown in Fig. 13(c). Table VII shows the results of indicators after down-sampling.

TABLE VII. THE RESULT AFTER DOWN-SAMPLING

	simple LSB	**DWT**	**RLE & IWT**	**Our method (*n* = 2)**	**Our method (*n* = 3)**
BER(%)	49.17	49.45	0	49.31	32.47
NC	0.8057	0.8106	1	0.8071	0.8813
PSNR	9.1171	9.3114	96.1588	9.2245	11.6252



(a)Original Secret Information

(b)Extracted Information (*n*=2)

(c)Extracted Information (*n*=3)

Figure 10. the comparison of the original and extracted information after down-sampling

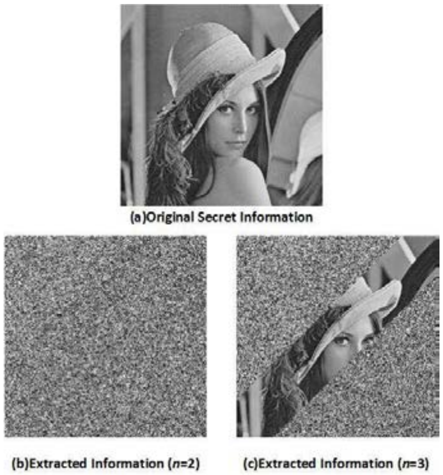
As shown in Fig. 13, when *n* = 2, the proposed method can't resist the up-sampling attack. When *n* = 3, the proposed method can partially resist it. That is, a part of the secret information can be extracted from the stego audio after being attacked. As shown in Table VII, when *n* = 3, the robustness of the proposed method is better than that when *n* = 2. By comparing Table VI and Table VII, the BER after down-sampling attack is lower than that after up-sampling attack. And the NC after down-sampling attack is higher than that after up-sampling attack. It means that the proposed method can resist down-sampling attack better than up-sampling attack.

5) Attack 5: Noise Attack (SNR = 55dB)

A white noise with 55dB is added to stego audio. When *n* = 2, the extracted information is shown in Fig. 14(b). When *n* = 3, the extracted information is shown in Fig. 14(c). Table VIII shows the results of indicators after noise attack.

TABLE VIII. THE RESULT AFTER NOISE ATTACK

	simple LSB	**DWT**	**RLE & IWT**	**Our method (*n* = 2)**	**Our method (*n* = 3)**
BER(%)	49.86	50.09	0	49.87	33.45
NC	0.8047	0.8078	1	0.8063	0.873
PSNR	9.0184	9.0486	96.1588	9.0375	11.1018



(a)Original Secret Information

(b)Extracted Information (*n*=2) (c)Extracted Information (*n*=3)

Figure 11. the comparison of the original and extracted information after noise attack

As shown in Fig. 14, when *n* = 2, the proposed method can't resist the noise attack. When *n* = 3, the proposed method can partially resist it. That is, a part of the secret information can be extracted from the stego audio after being attacked. As shown in Table VIII, with *n* = 3, the robustness of the proposed method is better than that with *n* = 2.

In summary, the robustness of the proposed method is related to the robustness of the steganographic method used in each segment. If the steganographic method used in each segment is robust, then the proposed method is robust too. Vice versa. This phenomenon reflects the difference between segments. Using this phenomenon, we can reorganize the secret information properly. For the important part of the secret information, we should use a stronger robust method for embedding. Even the stego audio is attacked, the covert receiver can also extract some useful information. For other unimportant parts, we can use a less robust method for embedding. These parts can even become false information to confuse the eavesdropper.

V. CONCLUSIONS

The capacity, anti-detectability and robustness are three important elements of steganography. For a long period, many scholars have been studying on how to enhance the capacity or purely improve the anti-detectability. But this promotion is at the expense of reducing robustness. We propose a multi-layer method based on the collaboration of audio steganography layer and network steganography layer (called MLS-ATDSS&NS). In audio steganography layer, the secret information is embedded by using the proposed ATDSS method. In network steganography layer, the key is transmitted covertly by using network steganographic method. Through analyses of experimental results, it is found that the proposed method can achieve a higher steganographic capacity than others. At the same time, owe to the collaboration of multi-layer, the proposed method has better anti-detectability and robustness. In the future, we will develop segmented steganography on transform domain.

REFERENCES

[1] B. Yang, L. Guo, Y. J. Wang, and C. P. Wang, MIDI Audio Steganalysis via HCF-based Statistical Features, Communications Technology, vol.9, pp.159–161, 2010.

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73. C. H. Xie, Y. M. Cheng, and Y. K. Chen, PN Sequence Estimation and Spread-Spectrum Steganalysis, Acta Electronica Sinica, vol.2, pp.255–259, 2011.

[3] Y. L.Wang, Research of Speech Steganalysis Method. University of Scicence and Technology of China, Hefei, 2008.

[4] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, Techniques for data hiding, IBM systems journal, vol.35, no.3, pp.313–336, 1996.

[5] K. T. R. Naidu, T. V. S. G. Prasad, and P. G. Mamatha, An Approach of Robust High Capacity Audio Steganography and Cryptography Using LSB Algorithms, International Journal of Engineering Trends and Technology (IJETT), vol.9, no.10, pp.521–524, 2014.

[6] J. Vimal, and A. M. Alex, Audio steganography using dual randomness LSB method, International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), pp.941–944. IEEE, USA, 2014.

[7] S. Roy, J. Parida, A. K. Singh, and A. S. Sairam, Audio steganography using LSB encoding technique with increased capacity and bit error rate optimization, International Conference on Computational Science, Engineering and Information Technology, pp.372–376. ACM, USA, 2012.

[8] H. Kumar, Enhanced LSB technique for audio steganography, International Conference on Computing Communication & Networking Technologies (ICCCNT), pp.1–4. IEEE, USA, 2012.

[9] B. S. Ko, R. Nishimura, and Y. Suzuki, Time-spread echo method for digital audio watermarking, IEEE Transactions on Multimedia, vol.7, no.2, pp.212–221, 2005.

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

[11] S. Mitra, and S. Manoharan, Experiments with and enhancements to echo hiding, International Conference on Systems and Networks Communications, pp.119–124. IEEE, USA, 2009.

[12] X. He, A. I. Illiev, and M. S. Scordilis, A high capacity watermarking technique for stereo audio, International Conference on Acoustics, Speech, and Signal Processing, pp.393-398. IEEE, USA, 2004.

[13] M. Tong, C. Y. Hao, X. J. Liu, and Y. P. Chen, An Audio Information Hiding Method Based on Fixed Additive Phase Amending, Computer Engineering, vol.32, no.1, pp.213–215, 2006.

[14] M. Tong, H. B. Ji, and X. J. Liu, Audio Information Hiding Method Based on Phase Coding, Computer Engineering, vol.9, no.34,, pp.7–9, 2008.

[15] H. Matsuoka, Spread spectrum audio steganography using sub-band phase shifting, International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp.3–6. IEEE, USA, 2006.

[16] M. G. Zou, and Z. T. Li, A wav-audio steganography algorithm based on amplitude modifying, International Conference on Computational Intelligence and Security (CIS), pp.489–493. IEEE, USA, 2014.

[17] D. Pal, and N. Ghoshal, A Robust Audio Steganographic Scheme in Time Domain (RASSTD), International Journal of computer Applications, vol.80, no.15, 2013.

[18] L. Xie, J. S. Zhang, and H. J. He, Robust audio watermarking scheme based on nonuniform discrete Fourier transform, International Conference on Engineering of Intelligent Systems, pp.1–5. IEEE, USA, 2006.

[19] A. Fan, and Q. Sun, The New Formulas of Simultaneously Calculating DFT and IDFT of Real Sequences and Its Application in Digital Audio Watermarking, JOURNAL-SICHUAN UNIVERSITY ENGINEERING SCIENCE EDITION, vol.40, no.2, pp.96–100, 2008.

[20] M. P. Jain, and V. Trivedi, Effective Audio Steganography by using Coefficient Comparison in DCT Domain, International Journal of Engineering Research and Technology, vol.2, no.8, pp.72–78, 2013.

[21] H. S. Chen, Y. Y.Wang, Q. Y. U, and M. Tan, Research for a DCT based Blind Audio Steganography Algorithm, Netinfo Security, vol.9, 60–63, 2013.

[22] N. Cvejic, Algorithms for audio watermarking and steganography, Oulun yliopisto, 2004.

[23] N. Gupta, and N. Sharma, DWT and LSB based Audio Steganography, International Conference on Optimization, Reliabilty, and Information Technology (ICROIT), pp.428–431, IEEE, 2014.

[24] A. M. Meligy, M. M. Nasef, and F. T. Eid, An Efficient Method to Audio Steganography based on Modification of Least Significant Bit Technique using Random Keys, International Journal of Computer Network and Information Security, vol.7, no.7, pp.24–29, 2015.

[25] H. I. Shahadi, R. Jidin, and W. H. Way, Lossless audio steganography based on lifting wavelet transform and dynamic stego key, Indian Journal of Science and Technology, vol.7, no.3, pp.323–334, 2014.

[26] K. Szczypiorski, Steganography in TCP/IP networks, State of the Art and a Proposal of a New System–HICCUPS, Institute of Telecommunications' seminar,Warsaw University of Technology, Poland, 2003.

[27] S. Wendzel, S. Zander, B. Fechner, and C. Herdin, Pattern-based survey and categorization of network covert channel techniques, ACM Computing Surveys (CSUR), vol.47, no.3, pp.50–51, 2015.

[28] S. J. Murdoch, and S. Lewis, Embedding covert channels into TCP/IP, International Workshop on Information Hiding, pp.247–261. Springer, Germany, 2005.

[29] C. G. Girling, Covert Channels in LAN's, IEEE Transactions on software engineering, vol.13, no.2, pp.292–296, 1987.

[30] A. Dyatlov, and S. Castro, Exploitation of data streams authorized by a network access control system for arbitrary data transfers: tunneling and covert channels over the HTTP protocol, Grayworld, USA, http://grayworld. net/projects/papers/html/covert paper. html, 2003.

[31] R. Rios, J. A. Onieva, and J. Lopez, HIDE DHCP: Covert communications through network configuration messages, IFIP International Information Security Conference, pp.162–173. Springer, Germany, 2012.

[32] X. G. Zou, Q. Li, S. He. Sun, and X. M. Niu, The research on information hiding based on command sequence of FTP protocol, Knowledge-Based Intelligent Information and Engineering Systems, pp.178–178. Springer, Germany, 2005.

[33] Z. Trabelsi, and I. Jawhar, Covert file transfer protocol based on the IP record route option, Journal of Information Assurance and Security, vol.5, no.1, pp.64–73, 2010.

[34] T. Graf, Messaging over IPv6 destination options, 2003.

[35] S. Wendzel, B. Kahler, and T. Rist, Covert channels and their prevention in building automation protocols: A prototype exemplified using BACnet, International Conference on Green Computing and Communications (GreenCom), pp.731–736. IEEE, USA, 2012.

[36] N. B. Lucena, G. Lewandowski, and S. J. Chapin, Covert channels in IPv6, International Workshop on Privacy Enhancing Technologies, pp.147–166. Springer, Germany, 2005.

[37] Sebastian Zander, Grenville Armitage, and Philip Branch, Covert channels in the IP time to live field, pp.298–302, 2006.

[38] J. Giffin, R. Greenstadt, P. Litwack, and R. Tibbetts, Covert messaging through TCP timestamps, International Workshop on Privacy Enhancing Technologies, pp.194–208. Springer, Germany, 2002.

[39] T. Handel, and M. Sandford, Hiding data in the OSI network model, Information Hiding, pp.23–38. Springer, Germany, 1996.

[40] S. Wendzel, Protocol channels as a new design alternative of covert channels, http://cds.cern.ch/record/1126596, arXiv:0809.1949, 2008.

[41] M. Wolf, Covert channels in LAN protocols, Local Area Network Security, pp.89–101. Springer, Germany, 1989.

[42] M. Mehic, J. Slachta, and M. Voznak, Whispering through DDoS attack, Perspectives in Science, vol.7, pp.95–100, 2016.

[43] N. B. Lucena, J. Pease, P. Yadollahpour, and S. J. Chapin, Syntax and semantics-preserving application-layer protocol steganography, International Workshop on Information Hiding, pp.164–179. Springer, Germany, 2004.

[44] L. H. Yao, X. C. Zi, L. Pan, and J. H. Li, A study of on/off timing channel based on packet delay distribution, Computers & Security, vol.28, no.8, pp.785–794, 2009.

[45] S. Cabuk, C. E. Brodley, and C. Shields, IP covert timing channels: design and detection, Proceedings of the 11th ACM conference on Computer and communications security, pp.178–187. ACM, USA, 2004.

[46] M. A. Padlipsky, D. W. Snow, and P. A. Karger, Limitations of end-to-end encryption in secure computer networks, DTIC Document, 1978.

[47] W. Q. Li, and G. L. He, Towards a protocol for autonomic covert communication, International Conference on Autonomic and Trusted Computing, pp.106–117. Springer, Germany, 2011.

[48] S. H. Sellke, C. C. Wang, S. Bagchi, and N. B. Shroff, Covert tcp/ip timing channels: theory to implementation, Proceedings of the 28th Conference on computer communications, pp.2204–2212. ACM, USA, 2009.

[49] S. Gianvecchio, and H. Wang, Detecting covert timing channels: an entropy-based approach, Proceedings of the 14th ACM conference on Computer and communications security, pp.307–316. ACM, USA, 2007.

[50] V. Berk, A. Giani, G. Cybenko, and N. Hanover, Detection of covert channel encoding in network packet delays, Rapport technique TR536, de lUniversit´e de Dartmouth, vol.19, 2005.

[51] S. Gianvecchio, H. Wang, D. Wijesekera, and S. Jajodia, Model-based covert timing channels: Au- tomated modeling and evasion, International Workshop on Recent Advances in Intrusion Detection, pp.211–230. Springer, Germany, 2008.

[52] X. P. Luo, E. W. Chan, and R. K. Chang, Cloak: A ten-fold way for reliable covert communications, European Symposium on Research in Computer Security, pp.283–298. Springer, Germany, 2007.

[53] K. Ahsan, and D. Kundur, Practical data hiding in TCP/IP, International Workshop on Multimedia Security at ACM Multimedia, vol.2, no.7, 2002.

[54] D. Kundur, and K. Ahsan, Practical Internet steganography: data hiding in IP, International Workshop on Security of Information Systems, 2003.

[55] R. C. Chakinala, A. Kumarasubramanian, R. Manokaran, G. Noubir, P. Rangan, and R. Sundaram, Steganographic communication in ordered channels, International Workshop on Information Hiding, pp.42–57. Springer, Germany, 2006.

[56] J. O. Seo, Sathiamoorthy Manoharan, and Aniket Mahanti, A Discussion and Review of Network Steganography, International Conference on Big Data Intelligence and Computing and Cyber Science and Technology Congress, pp.384–391. IEEE, USA, 2016.

[57] A. J. Al-Najjar, The decoy: multi-level digital multimedia steganography model, International Conference on Mathematics and Computers in Science and Engineering, pp.445–450. World Scientific and Engineering Academy and Society, USA, 2008.

[58] W. Frkaczek, W. Mazurczyk, and . Szczypiorski, Multi-level steganography: Improving hidden communication in networks, Journal of Universal Computer Science (J. UCS), vol.18, no.14, pp.1967–1986, 2012.

[59] S. K. Bandyopadhyay, and B. G. Banik, Multi-Level Steganographic Algorithm for Audio Steganography using LSB Modification and Parity Encoding Technique, International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), vol.1, no.2, pp.71–74, 2012.

[60] H. L. Liu, J. J. Liu, X. H. Yan, and P. F. Xue, An Audio Steganography Method Based on Run Length Encoding and Integer Wavelet Transform, International Journal of Digital Crime and Forensics (IJDCF), vol.2, no.10, 2017.

[61] P. F. Xue, J. S. Hu, H. L. Liu, and R. G. Hu, A New Network Steganography Method Based on the Transverse Multi-Protocol Collaboration, Journal of Information Hiding and Multimedia Signal Processing, vol.8, no.2, pp.445–459, 2017.