IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 29, 2019, accepted July 18, 2019, date of publication July 22, 2019, date of current version August 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2930133

An Enhanced Steganographic Code and Its Application in Voice-Over-IP Steganography

LEJUN ZHANG[1], XIAOYAN HU[1], WAQAS RASHEED[1], TIANWEN HUANG[1], AND CHUNHUI ZHAO[1,2]

[1]College of Information Engineering, Yangzhou University, Yangzhou 225127, China
[2]College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

Corresponding author: Chunhui Zhao (zhaochunhui@hrbeu.edu.cn)

This work was supported in part by the Natural Science Foundation of Heilongjiang Province of China under Grant LC2016024, and in part by the Natural Science Foundation of the Jiangsu Higher Education Institutions under Grant 17KJB520044.

ABSTRACT Voice-over IP (VoIP) technology is a kind of digital transmission technology based on IP network. It is one of the important methods to use voice service in VoIP as steganographic carrier to ensure secure transmission. However, the traditional steganographic code has some problems, such as low embedding efficiency and weak concealment, which cannot meet the requirements of VoIP streaming media information hiding for the security of secret information. Therefore, a steganographic algorithm combining F5 and simplified wet paper code (SWPC) algorithm is proposed. The main idea is to embed secret messages in each row of the carrier matrix using the F5 algorithm, and then the SWPC algorithm is used to embed the columns according to the wet and dry characteristics of the wet paper code without affecting the results before row embedding. We use the VoIP streams encoded by the ITU-T G.729a codec as a carrier to verify the proposed scheme. The experimental results demonstrate that the proposed scheme can achieve relatively better IP speech data steganographic transparency and that it can outperform F5-WPC and SWPC approaches.

INDEX TERMS F5 algorithm, steganography, streaming media, voice over IP, wet paper code.

I. INTRODUCTION

Information hiding is a new type of secure communication technology that can hide secret information in a seemingly ordinary carrier to achieve the purpose of secrecy and security. It conceals the existence of hidden communication and in many cases provides better security than traditional communication methods. Therefore, it has been developing rapidly in recent years [1]–[3]. At present, the cover carrier in information hiding has been transformed from the original image to the streaming media. It is worth noting that IP voice technology has attracted extensive attention from many researchers [4]–[6]. The main reason is that the secret information embedded in voice over IP (VoIP) is dynamic and not easy to be discovered by illegal attackers, and VoIP voice stream can provide better steganography performance and embedding capacity. The research on VoIP streaming media information hiding can be divided into two directions: one is the information-hiding method based on voice carrier, whereas the other is the information-hiding method based on network protocol. One is the information-hiding method based on the network protocol involved in the process of IP voice transmission, and the other is the information-hiding method based on the voice carrier. For example, Huang *et al.* [7], [8], Banai *et al.* [9], and Jin *et al.* [10] exploited redundancy from the perspective of speech coding principles to determine the hidden locations in the speech stream. Xu and Yang [11] and Miao and Huang [12] studied the method of information hiding based on the parity of matrix coding and iLBC codebook index respectively, and obtained a good hiding effect, but the performance of hiding capacity was not high. Mazurczyk *et al.* [13] proposed transcoding the speech payload to compress the size and make room for hiding information, of which the LSB method is the most widely used, which has the advantages of having a large embedded capacity and low computational complexity. However, much of the LSB approach depends on the need for the sender and receiver to use the consensus overlay bit [14]. Therefore, security considerations remain.

Generally speaking, in the case of the same embedding rate (ER), the less the steganographic algorithm modifies the carrier, the less the possibility of hidden information being

The associate editor coordinating the review of this manuscript and approving it for publication was Malik Najmus Saqib.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

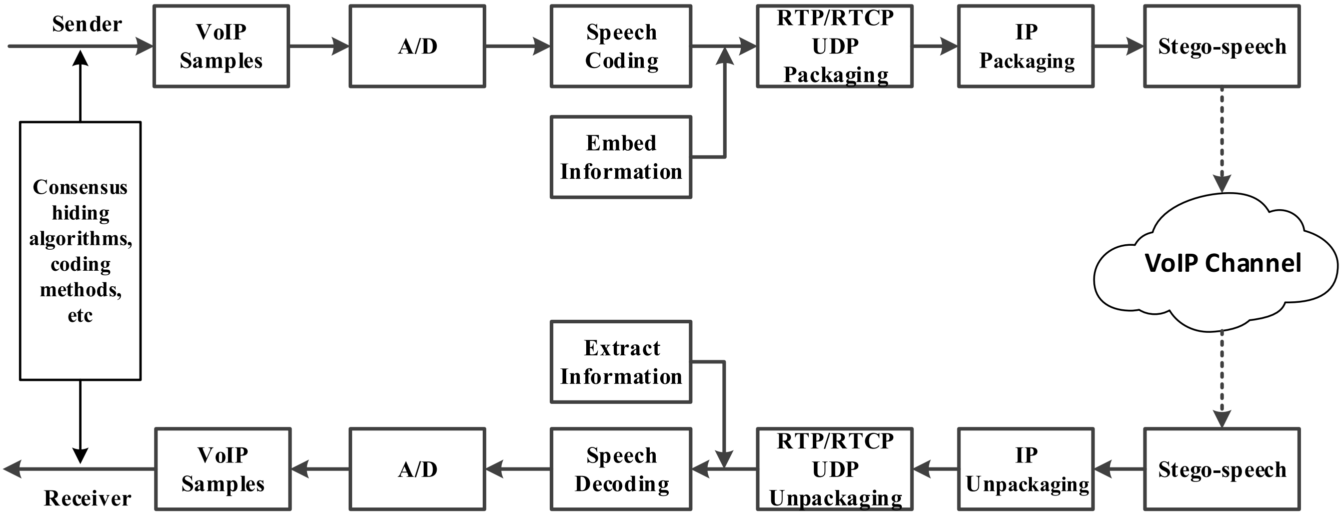


FIGURE 1. Covert communication over VoIP.

detected, that is, the higher the security [15], [16]. In order to further improve the efficiency and security of steganography, the information coding theory was applied to the information embedding process, and a variety of steganography codes were proposed, including matrix coding [17] and Wet Paper Code (WPC). Among them, matrix encoding was first proposed by Crandall [18] as a steganographic scheme to improve coding efficiency [19], [20]. This technology can reduce the modification of carrier data and improve the embedding efficiency at the cost of more carrier data. Matrix encoding was first applied to F5 algorithm, which can embed secret information when the carrier change is up to 1 bit. In 2005, Fridrich *et al.* [21] proposed the WPC, which determines which positions in the image can be modified through a selection rule known only to the sender, enabling the sender to embed information without modifying the sensitive area of the carrier, thus improving the security of steganography.

For matrix coding, in many cases, the carrier length is limited, which cannot meet this requirement, resulting in matrix coding not being able to be applied. While the higher conventional wet paper steganographic code calculated the solving process complexity, there may also be a case where no solution seriously affecting the efficiency of the secret messages is embedded. In view of the problem, this article proposes a steganographic code for enhanced VoIP steganographic method. It takes advantage of VoIP voice services as the steganographic carrier combines the F5 coding algorithm and Simplified Wet Paper Code (SWPC) coding algorithm and embeds secret messages into the VoIP streams. Compared with the traditional approaches, it can have a higher EE while meeting the real-time requirements of IP speech.

II. RELATED WORKS

A. COVERT COMMUNICATION OVER VoIP

The realization of VoIP streaming media information hiding is based on the principle of VoIP streaming media communication, and the purpose of information hiding is to embed the hidden information into the IP voice streaming carrier or the least important bit involved in the network protocol in the process of IP voice transmission [22]. The VoIP streaming media information hiding communication process is shown in Fig. 1.

Fig. 1 depicts the VoIP streaming media information hiding a general model of communication process, before the VoIP voice and data streaming media transmission, and consistency of communication that both sides needed to achieve the communication mode, including the use of hidden coding algorithm and specific way of speech coding, and so on, to ensure that the voice and data under the condition are not damaged , and success to get to the hidden information in the VoIP streams. First, the sender embedded the secret information into the IP voice carrier by consensus. Then, after receiving the IP voice data containing the secret information, the receiver uses the prearranged way to extract the secret information. For the third-party illegal attacker, although he/she can intercept the IP voice data, he/she cannot determine whether the carrier data are really hidden secret messages, to achieve the secret communication, information-hiding security, and low-latency requirements of VoIP services.

B. F5 ALGORITHM

Matrix encoding is a steganographic scheme to improve the encoding efficiency. Its principle is to reduce the modification of carrier data to improve the embedding efficiency at the cost of more carrier data. Matrix coding was first applied to F5 algorithm, which can embed x bit secret information by modifying 1 bit at most on n ($n = 2^x - 1$) bit carrier, greatly increasing the embedding efficiency of steganography, and reducing the distortion brought by embedding to carrier, thus achieving a qualitative leap in non-aggressiveness [23].

Assume that the original carrier object l is composed of n elements $\{l_i\}$, $l = \{l_1, l_2, \ldots, l_n\}$, which stands the carrier with n modifiable elements after replacement, m is secret messages containing x bits, l' is the embedded carrier object obtained by embedding the secret information m into l. F5 algorithm is the inverse process of the embedding algorithm, and the following is only the introduction of F5 embedding algorithm, whose process is roughly as follows:

1) Define the function f that can embed the secret information m of x bits into the carrier l, which is denoted as follows:

$$f(l) = \bigoplus_{i=1}^{n} l_i \cdot i \tag{1}$$

2) Using the XOR operation of the function f and secret messages m to find the position in l that needs to be modified, we can get:

$$v = f(l) \oplus m \tag{2}$$

3) Modify the rules in l as follows:

$$l' = \begin{cases} l & v = 0 \\ (l_1, l_2, \ldots, l_i, \ldots, l_n) & v = i \end{cases} \tag{3}$$

4) Repeat Steps 1 through 3 until secret messages are completely embedded.

Accordingly, the extraction function of F5 algorithm is defined as follows:

$$f(l) = f(m, l) = \left(\bigoplus_{i=1}^{n} l_i \cdot i \right) \oplus m \tag{4}$$

Accordingly, the extraction function of F5 algorithm can be defined as follows:

$$f(x') = f(0, x') = f(m, x) \oplus f(0, x) = m \tag{5}$$

C. THE PROPERTIES OF HAMMING CODE

Four important properties of hamming codes are given below, which have been proven in [24] and [25].

Lemma 1: Let C be a hamming code of length $2^x - 1$. Let $i, j \in \{1, \ldots, 2^x - 1\}$. Then, there exists a unique coordinate $g_1(i, j) \in \{1, \ldots, 2^x - 1\}$, such that the vector with support $\{i, j, g_1(i, j)\}$ belongs to C.

Lemma 2: Let C be a hamming code of length $2^x - 1$. Let $i, j \in \{1, \ldots, 2^x - 1\}$. It is always possible to take a parity check matrix such that for any coordinate $1 \leq i \leq 2^{x\text{-}1} - 1$ there exist two coordinates, specifically the $(2^{x\text{-}1} - 1 + i)$th and the $(2^x - 1)$th, such that vector v with support $g_1(i, j) \in \{1, \ldots, 2^x - 1\}$ belongs to code C, where $g_2(i) = 2^{t\text{-}1} - 1 + i$.

Lemma 3: Let C be a hamming code of length $2^x - 1$ with parity check matrix H. Let $i, j, r \in \{1, \ldots, 2^{x-1} - 1\}$ and $i < j < r$ such that vector u with support i, j, r belongs to code C. Then, the vector v with support $\{j, g_2(i), g_2(r)\}$ belongs to code C and $g_2(i), g_2(r) \in \{2^x - 1, \ldots, 2^x - 2\}$.

Lemma 4: Let C be a hamming code of length $2^x - 1$. Let $k = (k_1, k_2, \ldots, k_n) \in GF(2^n)$ and $supp(k) = \{i_1, i_2, i_3\}$. According to the F5 embedding function, there is a vector k' that has a forth component i_4 supporting $supp(k') = \{i_1, i_2, i_3, i_4\}$, where $i_4 = i_1 \oplus i_2 \oplus i_3$.

D. IMPROVED WPC USING SIMPLIFIED HAMMING PARITY-CHECK MATRIX (SWPC)

The traditional design principle of WPC is to divide the original carrier into two categories [26]. The sender divides the cover bits into wet and dry parts independently. When users need to embed secret messages in the original carrier, all the secrets of the information can be embedded only in a dry place. It is a process similar to the positions on a blank sheet of paper to write, only in a dry place to write, but cannot write information in wet places. When the original carrier, according to the division of predefined wet and dry basis, selects several position, realizes the hidden information embedded, and contains the secret after the carrier sent information to the receiver, the receiver receives hidden information contained by the carrier; the carrier no longer distinguishes between information regarding which position was dry and what place is wet and directly contains the secret information received by the carrier of information extraction. Therefore, it ensures that the wet and dry division basis selected by the sender will not be leaked in the process of information transmission, thus improving the security of information hiding.

Assume that the cover object L has n elements, denoted by $L = \{l_1, l_2, \ldots, l_n\}$. The sender could discretionarily choose k elements l_j in L to hide information, denoted by $L' = \{l_1', l_2', \ldots, l_t'\}, j \in J \in \{1, 2, \ldots, n\}, |J|{=}k$. The receiver can extract embedded information in the following way:

$$DL' = m \tag{6}$$

where D is the $p \times n$ random binary matrix. Subtracting DL from both sides, we can get the following:

$$D(L' - L) = m - DL \tag{7}$$

As in Equation(7), the sender can determine the bits that need to be modified in L according to the difference between the embedded information m and DL. Denoting $v = (L' - L)$, we have the following:

$$Hv = m - DL \tag{8}$$

where H is the submatrix of k column vectors of D. Obviously, the non-zero element in v corresponds to the element for which the sender has to modify L during the embedding process.

Different from the traditional WPC construction process, the WPC of the check matrix is constructed from the reverse direction, and the expandable matrix is generated according to the parameters. Each steganographic SWPC algorithm to the carrier for not more than 2 bits of modification can achieve rapid coding while ensuring that the encoding scheme always has a solution.

III. AN ENHANCED STEGANOGRAPHIC CODE CONSTRUCTION METHOD FOR VoIP

In this paper, we present a construction method for IP speech streams, combining F5 and SWPC codecs. First, the IP speech

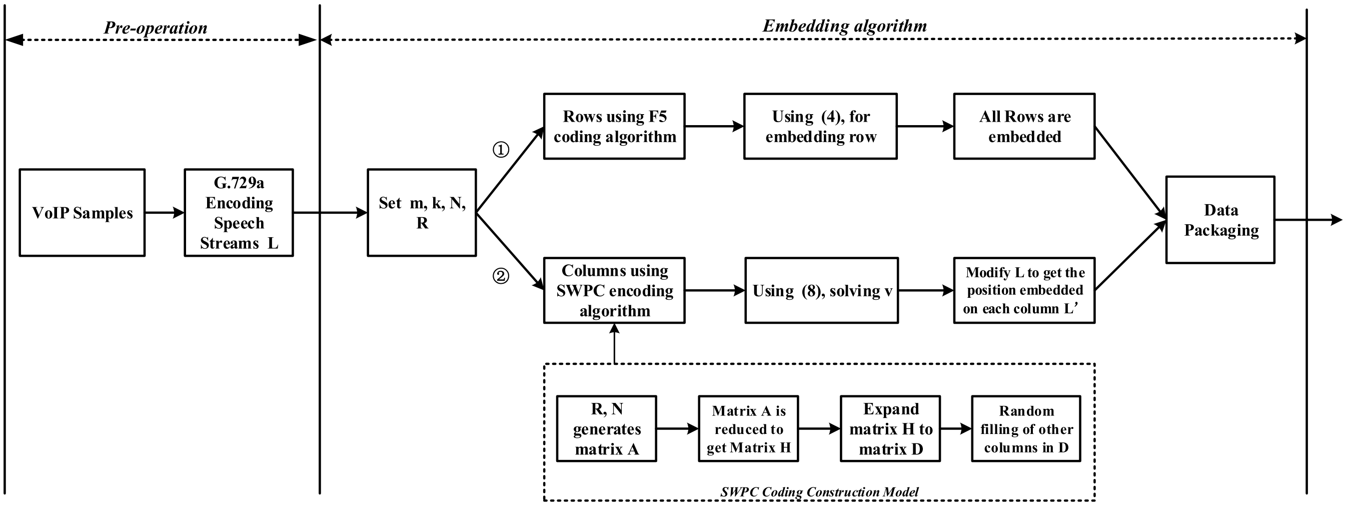

Pre-operation
Embedding algorithm
VoIP Samples
G.729a Encoding Speech Streams L
Set m, k, N, R
①
②
Rows using F5 coding algorithm
Using (4), for embedding row
All Rows are embedded
Columns using SWPC encoding algorithm
Using (8), solving v
Modify L to get the position embedded on each column L'
Data Packaging
R, N generates matrix A
Matrix A is reduced to get Matrix H
Expand matrix H to matrix D
Random filling of other columns in D
SWPC Coding Construction Model

FIGURE 2. The process of embedding algorithm.

streams are arranged in a matrix, and the F5 coding algorithm is used to perform the first embedding of the secret information on the matrix block, but the second embedding is performed on the column using the more efficient SWPC algorithm. Compared with F5-WPC algorithm and SWPC algorithm approaches, the proposed method maintains the steganographic transparency of the IP speech streams and improves the embedding efficiency of the secret information.

The design idea is that the IP speech streams L encoded by the ITU-T G.729a encoder are divided into several independent matrix blocks with the length of $N \times (2^x - 1)$, where N is the grouping length of the WPC selected by the sender in the SWPC algorithm, and the secret information m embedded in x bits is grouped according to R for each length.

First, scanning the matrix, the rows in the matrix use F5 algorithm to embed hidden information. Using the nature of F5 algorithm, can keep the result of row embedding, and does not need to introduce too much modification. In most cases, it only needs to modify 1 bit. For column embedding, we need to consider the proportion of 'dry' elements in the WPC and the position to be modified in case of row conflict. When 1th column is embedded, if row j is not modified, the j position in 1th column is recorded as 'wet', and then the proportion of 'dry' position is $(2^x - 1)/2^x$. Thus, the number of bits of secret information that can be embedded in 1th column is $Nx(2^x - 1)/2^x(2^x - 1)$, making $(2^x - N/2)N(2^x - 1)/(2^x - 1)$ changes on average. On this basis, it is assumed that the subscript of the modified carrier data in the row is r_i $(1 \leq i \leq N)$, and each column of the matrix is scanned. According to *Lemma 1* and *2*, scan only each column of the matrix element divided by the last column in the matrix , that is columns 1 through $(2^{x-1} - 1)$ embedded. Suppose that the subscript of the column index c_i $(1 \leq j \leq 2^{x-1} - 1)$ that needs to be modified in the matrix column conflicts with r_i. The modification process of the column index is as follows:

1) If either r_i or c_j needs to be modified or neither needs to be modified, in order not to affect the embedding result of row i, according to *Lemma 2*, change the subscripts of $2^{x-1} - 1 + j$ and $2^x - 1$. Further reduce the number of modifications, according to *Lemma 4*, modify $r_i \oplus \left(2^{x-1} - 1 + j\right) \oplus (2^x - 1)$ to replace $2^{x-1} - 1 + j$ and $2^x - 1$.
2) Coordinate r_i corresponding to c_j, judge the magnitude of r_i and c_j. If $r_i < c_j$, according to *Lemma 2*, change $2^{x-1} - 1 + i$ and $2^x - 1$. If $r_i > c_j$, the subscript of $g_1 = (r_i, c_j) > c_j$. According to *Lemma 3*, change the subscript of $g_2 = (r_i)$ and $g_2 = (c_j)$

A. EMBEDDED ALGORITHM FOR SECRET MESSAGES

Based on the main idea introduced above, this section presents the specific embedding algorithm. The flow of the embedding algorithm is shown in Fig. 2.

In row embedding, the F5 encoding algorithm is used to embed from row 1 to row N sequentially. The detailed embedding algorithm is described as follows:

1) Read the secret messages m bit, and use (4) to embed the secret messages m.
2) Repeat until all rows are embedded. Obviously, using the nature of F5 algorithm, the result of row embedding can be modified by 1 bit at most.

The proportion of 'dry' position in column embedding has been determined, Use SWPC algorithm to embed from column 1 to column $2^{x-1} - 1$, and determine the number of elements that can be modified k. The steps are follows:

1) Construct the $R \times N$ hamming check matrix A.
2) Eliminate the $2^x - 1 - k$ columns from matrix A to get the reduced matrix H.
3) Construct the $R \times (2^x - 1)$ zero matrix D, replace the subscripts of D with the subscripts of $\{l_j\}$ $(j = 1, 2, \ldots, k)$ for each column in H, and unreplaced

columns in D randomly generate binary sequences for replacement.

4) Using (8), solve vector v, and modify the position of carrier data L according to the position of non-zero elements in the vector to get L', so as to embed secret messages m.

B. THE EXTRACTION ALGORITHM OF SECRET MESSAGES

Voice-over IP is the typical application of streaming media, featuring strong real-time and dynamic features. The network environment of real-time secret communication based on VoIP needs to meet the requirements of small network delay and strong stability. Since the real-time voice call adopts IP packet encapsulation, after receiving the encrypted IP voice data packet, the receiving end obtains the G.729a encoded IP voice carrier bit stream by disassembling the data packet and pre-processing operations. The proposed extraction algorithm is extracted, and the extraction algorithm process is shown in Fig 3.

Identify N and R, the IP speech streams L encoded by the ITU-T G.729a encoder is divided into several independent matrix blocks with the length of $N \times (2^x - 1)$. The first and second layers of the carrier matrix are taken as a whole to extract the secret messages. The steps are as follows:

1) Extract the secret messages of each row of matrix, using (5) is used to extract secret messages m.
2) Construct the $R \times (2^x - 1)$ zero matrix D.
3) Construct the $R \times N$ hamming check matrix A, and eliminate the $2^x - 1 - k$ columns from matrix A to get the reduced matrix H.
4) Replace the subscripts of D with the subscripts of $\{l_j\}$ $((j = 1, 2, \ldots, k))$ for each column in H, and the other zero columns in D randomly generate binary sequences for replacement.
5) Identify L', using (6) to extract secret messages m.

It can be seen from the above process that the extraction process of the secret information is simpler than the embedded process of the secret information, and the main task is to solve the problem, so that the secret information can be easily extracted.

C. SYNCHRONIZATION MECHANISM

It is worth noting that it is very necessary to consider the synchronous processing of secret IP communication, which determines whether the extraction and recovery of secret information can be successful. The proposed method in this paper is to embed and extract secret information in the bitstream of IP speech carrier encoded by each voice packet on the basis of parameters agreed by both parties, and use the UDP protocol of connectionless transmission for real-time transmission over the IP network. The collected analog signals are converted into corresponding digital signals through PCM encoding, 8kHz sampling rate and 16-bit quantization. In order to ensure that the embedding and extraction of secret information in the communication process are not affected by network performance and other factors, to meet the real-time demand of VoIP secret synchronization mechanism [27], [28], we consider it from the following two aspects.

On the one hand, according to the characteristics of SWPC coding, the compressed IP speech data bitstream is grouped, and the secret messages is grouped according to the carrier length, and the secret information is embedded by embedding algorithm, so the loss of any secret speech packet [29] will not affect other packets. For real-time transmission through IP network channels and the receiver to receive packets, only according to sending the original serial processing time sequence in turn, use extraction algorithm in time to carry secret hidden information extracted in VoIP packets, grouping and embedding secret messages almost at the same time, to achieve the purpose of reduced latency and packet loss.On the other hand, this paper designs a method to determine whether each IP voice packet can be embedded and extracted independently, which is more suitable to deal with the problem that VoIP communication synchronization can extract and restore secret information. Determine Current Operation (DCOP) is a method for determining a voice packet. It can synchronize the embedding and extraction of information by the sender and the receiver. The receiver will not miss the extraction of secret information in any voice packet; nor will the receiver conduct the extraction process for the voice packet without secret information. The calculation of DCOP is very small and easy to operate. Fig.4 shows the flow chart of the DCOP method.

To sum up, the whole communication process preprocesses the speech carrier from the perspective of coding, introduces steganography coding strategy, and proves through experiments that the algorithm proposed in this paper can guarantee the real-time broadcast of IP speech and simultaneously receive secret messages.

IV. PERFORMANCE EVALUATION AND ANALYSIS

A. STEGANOGRAPHIC PERFORMANCE

In order to evaluate the performance of the construction method of enhanced steganographic codes proposed in this paper, we use the embedding rate (ER) to evaluate the steganographic capacity, the bit change rate after embedding secret messages (BCR) to evaluate the transparency of steganographic and the embedding efficiency (EE) to comprehensively evaluate the steganographic performance. For the ease of description, we assume that an $N \times (2^x - 1)$ encoding matrix is employed. ER is the ratio between the number of secret messages and the total number of cover bits, which can be defines as follows:

$$ER = \frac{\sum_{t=1}^{2^{x-1}-1} Nx(2^x - t)/2^x(2^x - 1) + Nx}{N \times (2^x - 1)} \tag{9}$$

where N is the packet length encoded by SWPC, and x is the number of embedded secret messages bits.

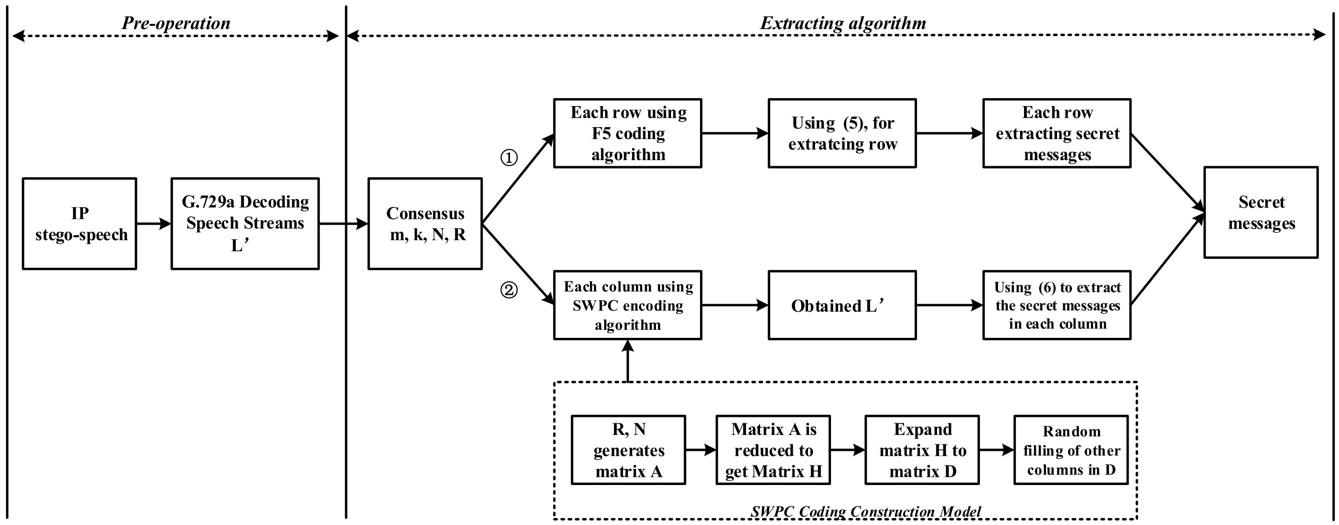

Pre-operation
Extracting algorithm
IP stego-speech
G.729a Decoding Speech Streams L'
Consensus m, k, N, R
①
②
Each row using F5 coding algorithm
Using (5), for extratcing row
Each row extracting secret messages
Each column using SWPC encoding algorithm
Obtained L'
Using (6) to extract the secret messages in each column
Secret messages
R, N generates matrix A
Matrix A is reduced to get Matrix H
Expand matrix H to matrix D
Random filling of other columns in D
SWPC Coding Construction Model

FIGURE 3. The process of extraction algorithm.

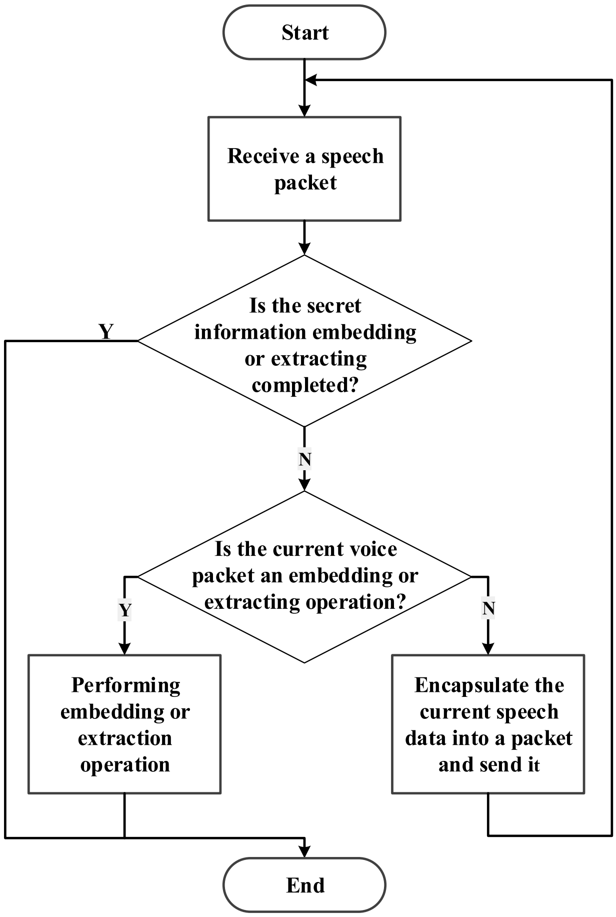

Start
Receive a speech packet
Is the secret information embedding or extracting completed?
Y
N
Is the current voice packet an embedding or extracting operation?
Y
N
Performing embedding or extraction operation
Encapsulate the current speech data into a packet and send it
End

FIGURE 4. Flow chart of DCOP method.

The average modification number of cover bits is as follows:

$$d = \sum_{t=1}^{2^{x-1}-1} \frac{N(2^x-5)(2^x-t)}{2^x} + N \times \frac{2^x-1}{2^x} \tag{10}$$

BCR refers to the ratio of the average modification number of cover bits to the number of cover bits, which can be defined as follows:

$$BCR = \frac{d}{N \times (2^x-1)} \tag{11}$$

Further, EE refers to the ratio of the number of bits embedded in secret messages to the average change number, which can be written as follows:

$$EE = \frac{\sum_{t=1}^{2^{x-1}-1} Nx(2^x-t)/2^x(2^x-1) + Nx}{d} \tag{12}$$

B. SPEECH QUALITY TESTING AND ANALYSIS

In order to test the effectiveness of the algorithm proposed in this paper, we selected the speech data in various audio libraries provided in ITU-T P.501 standard Appendix B of the International Telecommunication Union Standard as the speech samples, and selected the representative ITU-T G.729a encoder [30] as the IP voice encoder. The languages of the original speech samples include the following: Chinese male speech, Chinese female speech, English male speech and English female speech. During the test, in order to simulate the real VoIP communication, an external player is used to play the audio speech as the carrier speech input. First, all the speech samples were converted into PCM format, 8 kHz sampling frequency, and 16-bit quantized mono speech. Then, the secret messages are embedded into the IP speech data encoded by the ITU-T G.729a encoder. Finally, the secret information is extracted from the encrypted IP speech data decoded by the G. 729a decoder. Fig. 5 shows the original speech waveform before embedding in a test, and Fig. 6 is the corresponding speech waveform after embedding. As can be seen from the comparison figure, the waveform of the carrier hardly changed before and after hiding, indicating that the steganography algorithm is feasible.

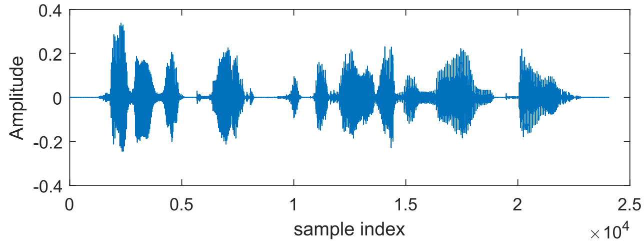


FIGURE 5. The original speech waveform before embedding.

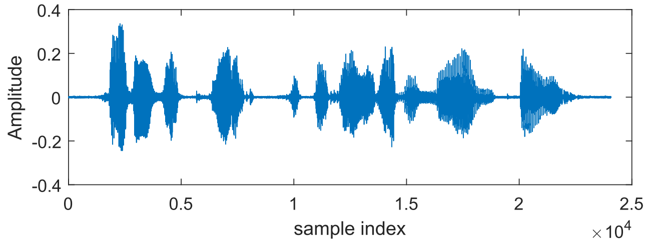


FIGURE 6. The speech wave graph of the embedded secret messages.

TABLE 1. Statistical results of PESQ.

Speech samples	***Algorithm used***	*[33]*	*[34]*
Chinese male speech	*3.625*	3.628	3.683
Chinese female speech	*3.607*	3.433	3.682
English male speech	*3.613*	3.565	3.702
English female speech	*3.592*	3.374	3.698

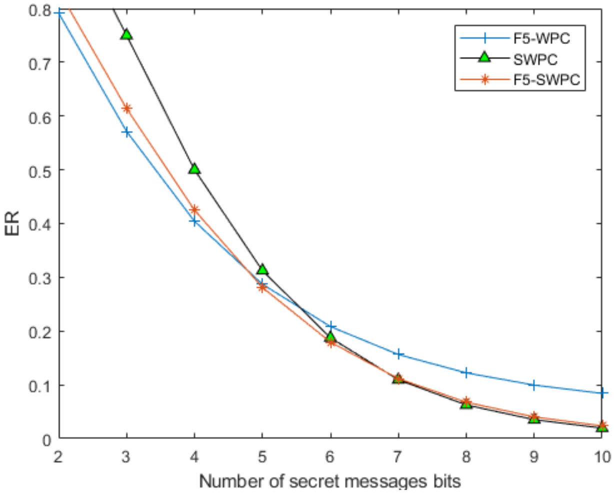


FIGURE 7. Statistical results of embedding rate (ER).

TABLE 2. Experiment results of BCR values of different method.

BCR / *x*	*SWPC Algorithm*	*F5-WPC Algorithm*	***Algorithm used***
$x = 1$	0.4167	0.2600	*0.2550*
$x = 2$	0.2500	0.1407	*0.1329*
$x = 3$	0.1615	0.0820	*0.0722*
$x = 4$	0.945	0.0533	*0.0422*
$x = 5$	0.0622	0.0394	*0.0275*
$x = 6$	0.0378	0.0328	*0.0203*
$x = 7$	0.0260	0.0297	*0.0168*
$x = 8$	0.0164	0.0284	*0.0151*
$x = 9$	0.0117	0.0278	*0.0144*

In order to further test the hiding effect of the algorithm in this paper, we also adopted the method of the Perceptual Evaluation of Speech Quality(PESQ) proposed in ITU-T P.862 [31], [32]. The PESQ is used to predict perceived quality by computing a PESQ score by comparing the original speech signal with the degraded speech signal. The higher the PESQ score, the better the speech signal quality. In the experiment, the PCM speech data in the audio library were used as the original speech of the PESQ, and the steganographic of the PCM speech data was used as the degraded speech of the PESQ. The PESQ values of various steganographic speech samples in different languages were taken as the mean values of each group of test data, which are listed in Table 1.

According to the data analysis in Table 1, the average score of speech steganography PESQ is about 3.601, which meets the requirement of VoIP call quality. The PESQ value in [33] is low because the long delay in the process of communication affects the quality of speech. The PESQ value in [34] is slightly higher than the algorithm used in this paper, but the embedding capacity has limitations. The steganographic algorithm proposed in this paper has high embedding performance and meets the requirements of VoIP call quality, thus, it is very worthwhile.

C. STEGANOGRAPHIC PERFORMANCE COMPARISON ANALYSIS

According to the steganographic algorithm proposed in this paper, we also established the corresponding simulation test environment and compared it with F5-WPC and SWPC steganographic algorithm. Embedding rate can be depicted in Fig. 7.

In Fig. 7, we can observe that the embedding rate of the proposed algorithm is higher than that of F5-WPC algorithm when the embedding rate is higher than 0.1. However, when the embedding rate is lower than 0.1, the embedding performance is inferior to SWPC and F5-WPC steganalysis algorithm. This is because when the parameter x increases gradually, the embedding rate in rows of hamming code $[2^x - 1, x, 1]$ radually decreases, and the probability that one modification is needed to embed x bit secret messages in the data of $2^x - 1$ bit carrier gradually increases. When x increases to a certain extent, the probability to be modified in the column tends to be zero, thus reducing the performance of embedding.

According to the simulation results, the BCR values of algorithms, SWPC, F5-WPC, and the method proposed in this paper are shown in Table 2. In order to more vividly see the hidden transparency of the three algorithms, the experimental results of BCR in Fig. 8 are drawn according to the experimental results in Table 2.

In general, the smaller the BCR, the better the hidden transparency. As depicted in Fig.8, compareing with F5-WPC

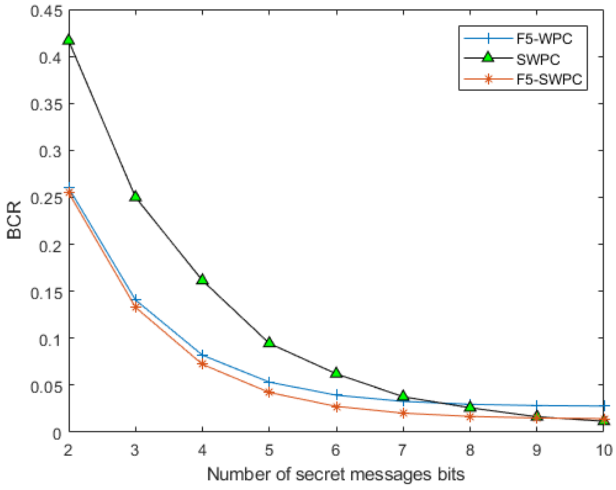


FIGURE 8. Statistical results of Bit-Change Rate (BCR).

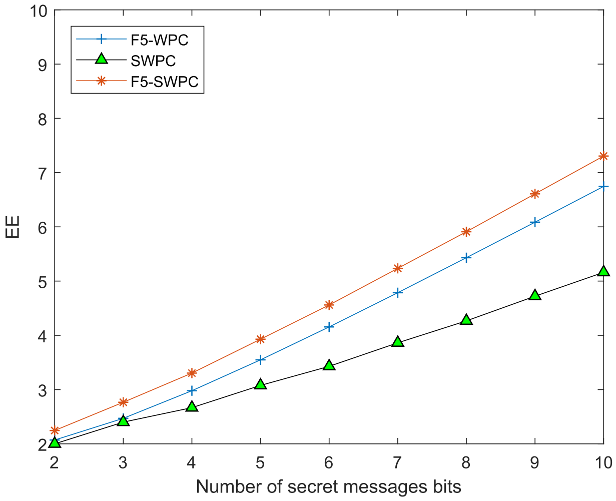


FIGURE 9. Statistical results of embedding efficiency (EE).

steganographic algorithm and SWPC steganographic algorithm, the carrier data of this algorithm are more transparent.

It can be seen from the combination of Fig.7 and Fig. 9 that the embedding rate decreases with the larger value of x, but the embedding efficiency decreases with the decrease of x. The embedding efficiency is related to the probability of the average change digit of the carrier. The higher the embedding rate is, the larger the average change digit of the carrier is, and the lower the embedding efficiency is. Therefore, you can compromise to choose the right x and modify the number of digits according to actual needs.

V. CONCLUSION

Through the analysis of F5 algorithm, WPC and the simplified hamming code check matrix, we combine F5 with SWPC algorithm, providing an enhanced steganography method for VoIP. The main idea is to divide the required encoding matrix, and use F5 and SWPC codes for the rows and columns of the matrix. On the premise of not affecting the embedding results of F5 algorithm, SWPC algorithm is used to improve the embedding efficiency. In order to evaluate the performance of the steganographic method proposed in this paper, we comprehensively evaluate the steganographic performance from the aspects of embedding rate, carrier bit change rate and embedding efficiency. Experiments are carried out on the IP speech streams encoded by the ITU-T G.729a encoder. The experimental results show that the steganography algorithm is feasible and guarantees the call quality of VoIP services. Compared with the other approaches, the proposed method not only can achieve steganography transparency and embedding efficiency while adequately satisfying the real-time requirement of VoIP. However, the tradeoff between the embedding rate and the embedding efficiency of the proposed method is what we will study next.

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their detailed reviews and constructive comments, which have helped improve the quality of this paper.

REFERENCES

[1] E. Zielińska, W. Mazurczyk, and K. Szczypiorski, "Trends in steganography," *Commun. ACM*, vol. 57, no. 3, pp. 86–95, 2014.

[2] R. J. Chen, J. L. Lai, and S. J. Horng, "Anti-forensic steganography using multi-bit minimum error replacement with flexible bit location," in *Proc. Int. Symp. Comput.*, Jun. 2012, pp. 175–178.

[3] W. Mazurczyk, "VoIP steganography and its detection—A survey," *ACM Comput. Surv.*, vol. 46, no. 2, p. 20, 2013.

[4] H. Tian, R. Guo, J. Lu, and Y. Chen, "Implementing covert communication over voice conversations with windows live messenger," *Adv. Inf. Sci. Service Sci.*, vol. 4, no. 4, pp. 18–26, 2012.

[5] H. Tian, J. Liu, and S. Li, "Improving security of quantization-index-modulation steganography in low bit-rate speech streams," *Multimedia Syst.*, vol. 20, no. 2, pp. 143–154, 2014.

[6] W. Mazurczyk, P. Szaga, and K. Szczypiorski, "Using transcoding for hidden communication in IP telephony," *Multimedia Tools Appl.*, vol. 70, no. 3, pp. 2139–2165, 2014.

[7] Y. F. Huang, S. Tang, and J. Yuan, "Steganography in inactive frames of VoIP streams encoded by source codec," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 2, pp. 296–306, Jun. 2011.

[8] Y. Huang, C. Liu, S. Tang, and S. Bai, "Steganography integration into a low-bit rate speech codec," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 6, pp. 1865–1875, Dec. 2012.

[9] B. Banai, L. Laustsen, I. P. Banai, and K. Bovan, "Presidential, but not prime minister, candidates with lower pitched voices stand a better chance of winning the election in conservative countries," *Evol. Psychol.*, vol. 16, no. 2, 2018, Art. no. 1474704918758736.

[10] L. Jin, Z. Ke, and T. Hui, "Least-significant-digit steganography in low bitrate speech," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2012, pp. 1133–1137.

[11] T. Xu and Z. Yang, "Simple and effective speech steganography in G.723.1 low-rate codes," in *Proc. Int. Conf. Wireless Commun. Signal Process.*, Nov. 2009, pp. 1–4.

[12] R. Miao and Y. Huang, "An approach of covert communication based on the adaptive steganography scheme on voice over IP," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2011, pp. 1–5.

[13] W. Mazurczyk, P. Szaga, and K. Szczypiorski, "Using transcoding for hidden communication in ip telephony," *Multimedia Tools Appl.*, vol. 70, no. 3, pp. 2139–2165, 2014.

[14] B. Xiao, Y. Huang, and S. Tang, "An approach to information hiding in low bit-rate speech stream," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2008, pp. 1–5.

[15] H. Tian, J. Qin, C.-C. Chang, Y. Huang, and Y. Chen, "Improved wet paper code using simplified Hamming parity-check matrix and its application in voice-over-IP steganography," *Internet Technol. J.*, vol. 18, no. 3, pp. 551–559, 2017.

[16] X.-X. Zhu, W.-M. Zhang, and J.-F. Liu, "A steganographic algorithm based on Hamming code and wet paper code," *J. Electron. Inf. Technol.*, vol. 1, no. 32, pp. 162–165, 2010.

[17] C. Kim, D. Shin, C.-N. Yang, and Y.-S. Chou, "Generalizing Hamming+k data hiding by overlapped pixels," *Multimedia Tools Appl.*, to be published.

[18] R. Crandall, "Some notes on steganography," in *Proc. Posted Steganography Mailing List*, 1998, pp. 1–6.

[19] Y. Cao, T. Wang, O. Kaiwartya, G. Min, N. Ahmad, and A. H. Abdullah, "An EV charging management system concerning drivers' trip duration and mobility uncertainty," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 48, no. 4, pp. 596–607, Nov. 2018.

[20] W. Tong, W. Jiyi, X. He, Z. Jinghua, and C. Munyabugingo, "A cross unequal clustering routing algorithm for sensor network," *Meas. Sci. Rev.*, vol. 13, no. 4, pp. 200–205, 2013.

[21] J. Fridrich, M. Goljan, P. Lisonek, and D. Soukal, "Writing on wet paper," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3923–3935, Oct. 2005.

[22] X. Hu, L. Zhang, T. Huang, and X. Lei, "A security evaluation method for voice-over-IP streaming media information hiding," in *Proc. 14th Int. Conf. Comput. Intell. Secur. (CIS)*, Nov. 2018, pp. 228–232.

[23] A. Westfeld, "F5—A steganographic algorithm," in *Proc. Int. Workshop Inf. Hiding*. Berlin, Germany: Springer, 2001, pp. 289–302.

[24] H. Rifà-Pous and J. Rifà, "Product perfect codes and steganography," *Digit. Signal Process.*, vol. 19, no. 4, pp. 764–769, 2009.

[25] Z. Lin, Y. Huang, and J. Wang, "RNN-SM: Fast steganalysis of VoIP streams using recurrent neural network," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 7, pp. 1854–1868, Jul. 2018.

[26] J. Fridrich, M. Goljan, and D. Soukal, "Wet paper codes with improved embedding efficiency," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 1, pp. 102–110, Mar. 2006.

[27] F. Shamieh and X. Wang, "Dynamic cross-layer signalling exchange for real-time and on-demand multimedia streams," *IEEE Trans. Multimedia*, to be published.

[28] G. Kaur, J. Kaur, S. Aggarwal, C. Singla, N. Mahajan, S. Kaushal, and A. K. Sangaiah, "An optimized hardware calibration technique for transmission of real-time applications in VoIP network," *Multimedia Tools Appl.*, vol. 78, no. 5, pp. 5537–5570, 2019.

[29] A. Bakri, A. Amrouche, M. Abbas, and L. Bouchakour, "Automatic speech recognition for VoIP with packet loss concealment," *Procedia Comput. Sci.*, vol. 128, pp. 72–78, Jan. 2018.

[30] *Coding of Speech at 8k bit/s Using Conjugate-Structure Algebraic-Code-Excited Linear Prediction (CS-ACELP)*, document G.729, 2007.

[31] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra, "Perceptual evaluation of speech quality (PESQ)—A new method for speech quality assessment of telephone networks and codecs," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 2, May 2001, pp. 749–752.

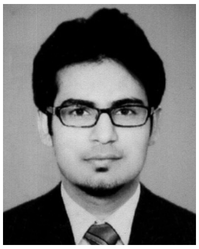
[32] J. Rozhon and M. Voznak, "Development of a speech quality monitoring tool based on ITU-T P.862," in *Proc. 34th Int. Conf. Telecommun. Signal Process. (TSP)*, Aug. 2011, pp. 62–66.

[33] W.-X. Yang, D.-H. Sun, and Y.-F. Huang, "Steganographic method in self-adaptive codebooks of speech codec," *Comput. Eng. Des.*, vol. 8, no. 34, pp. 2656–2661, 2013.

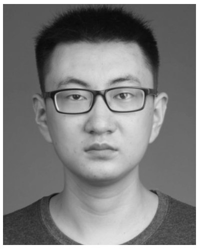
[34] W. U. Qiu-Ling and W. U. Meng, "Steganographic method of VOIP streaming media based on G.729b algorithm," *Comput. Eng. Des.*, to be published.

LEJUN ZHANG received the M.S. degree in computer science and technology from the Harbin Institute of Technology and the Ph.D. degree in computer science and technology from Harbin Engineering University. He was a Professor with Yangzhou University. His research interests include computer networks, social network analysis, dynamic network analysis, and information security.

XIAOYAN HU received the B.Eng. degree in computer science and technology engineering from the Huaiyin Institute of Technology. She is currently pursuing the master's degree in software engineering with Yangzhou University. Her research interest includes network security.



WAQAS RASHEED received the B.Eng. degree in software engineering from the University of Sindh, Pakistan. He is currently pursuing the master's degree in software engineering with Yangzhou University. His research interest includes network security.



TIANWEN HUANG received the B.Eng. degree in the Internet of Things engineering from the Huaiyin Institute of Technology. He is currently pursuing the master's degree in computer technology engineering with Yangzhou University. His research interest includes network security.

CHUNHUI ZHAO received the Ph.D. degree from the Department of Automatic Measure and Control, Harbin Institute of Technology, in 1998. He is currently with the College of Information Engineering, Yangzhou University, as a Professor and a Ph.D. Supervisor. He has published four works and more than 500 papers. His research interests include digital signal and image processing, mathematical morphology, and nonlinear filters. He is a Senior Member of the Chinese Electronics Academy.