International Journal of Innovative Computing, Information and Control
Volume **18**, Number **2**, April **2022**

ICIC International ©2022 ISSN 1349-4198
pp. **633–648**

EFFICIENT AUTHENTICATION STEGANOGRAPHIC SYSTEMS BASED ON A CLIENT-SERVER MODEL WITH RANDOM-LIKE CODES

Hsi-Yuan Chang[1,2], Jyun-Jie Wang[3], Chin-Hsing Chen[1,2] and Chi-Yuan Lin[3,*]

[1]Institute of Computer and Communication Engineering
[2]Department of Electrical Engineering
National Cheng Kung University
No. 1, University Road, Tainan City 701, Taiwan
q38021149@ncku.edu.tw; chench@eembox.ncku.edu.tw

[3]Department of Computer Science and Information Engineering
National Chin-Yi University of Technology
No. 57, Sec. 2, Zhongshan Road, Taiping District, Taichung 41170, Taiwan
jjwang@ncut.edu.tw; *Corresponding author: chiyuan@ncut.edu.tw

Received October 2021; revised January 2022

Abstract. *In this study, to increase steganographic security, a matrix embedding code was developed as a commonly used public key system by exploiting matrix decomposition. However, in systems such as the public secret cryptography system, users cannot use the public key because of the high complexity required in the networking. To address this problem, a novel method of authentication in steganographic systems, which involves the generation of random-like codes, is presented in this paper. By using a Gaussian elimination technique, random-like codes can be represented as a product of a left submatrix, right submatrix, and systematic parity matrix. The proposed method involves using the aforementioned submatrices and matrix to generate public and private key. In the client, the nonshared selection channel is used to authenticate the user to reduce the risk of cheating. Experimental results confirmed that a reliable authentication performance was achieved when Hamming codes with a systematic parity matrix were used for authenticating steganographic systems.*

Keywords: Matrix embedding, Steganography, Public key, Network security

1. **Introduction.** In steganography, the cover must be modified, and the stego must be obtained. A high embedding efficiency is required for a steganographic scheme. Embedding efficiency, which is the average number of embedded bits per embedding change, is a critical factor in steganography and is involved in a steganographic technique called the matrix embedding (ME) code scheme. In coding theory, near-optimal ME codes, that is, ME codes close to the rate distortion bound, can be achieved using an excellent structured code, such as an LDGM code, and an efficient decoding algorithm that is sufficiently long [1]. Crandall [2] and Bierbrauer [3] have used excellent linear block codes for ME schemes, and linear block codes have been derived from covering codes [3-5] with a high embedding efficiency. Bierbrauer and Fridrich [5] described several families of covering codes constructed using the blockwise direct sum (BDS) of factorizations. In addition, they found that BDS(6) and BDS(8) led to the highest embedding efficiency, which indicates that a considerable increase in embedding efficiency was achieved using nonlinear covering codes. Fridrich and Soukal [6] proposed a high-embedding-efficiency scheme, namely an

DOI: 10.24507/ijicic.18.02.633

ME-based embedding technique with large payloads. They demonstrated the efficiency of this technique by using two types of linear block codes, namely simplex codes and random codes. Fridrich and Soukal [6] achieved superior steganographic security for a large payload. In addition, they used structured simple codes, that is, fast Hadamard decoding, to develop efficient ME codes and approach the efficiency bound for a large payload. Several approaches have been developed using structured codes [6-11]. Although the authors of [3-8] have proposed that structured linear block codes are suitable for ME, achieving a high embedding efficiency by promoting the subject to locate the coset leader in the case of sufficiently long linear block codes is difficult because of the complexity of the maximum likelihood (ML) method. Alternatively, wet paper codes (WPCs) [12] and ZZW (Zhang, Zhang, and Wang) construction [13-15] can be adopted to generate a family of codes with arbitrary small relative payloads from any code with a large payload. ZZW construction can provide a high embedding efficiency at a low embedding rate.

Parity matrix check delivery and user authentication are the two major ME-related problems exhibited by open networks [16,17]. In this study, the problem of protection against unauthorized access in the transmission of information between two parties in traditional ME was investigated by considering both parties to have the distinct parity check matrix H for the embedder and extractor. The emergence of open networks in which the majority of users are unacquainted with other users, that is, users who do not have a common parity matrix H, has enabled information security to be provided by employing traditional ME methods. In this study, an authentication scheme was designed for steganographic systems by using a public key system and nonshared selection channel.

In the master of networks for this study, matrix decomposition was used to construct a system similar to the public key system. Gaussian elimination plays a crucial role in the constructed system. This system has two keys: an encryption key E and a decryption key D. The encryption key is made public, whereas the decryption key is kept private. In the aforementioned system, when the encryption key is known, deriving the decryption key is computationally difficult; thus, this system provides the ability to encrypt messages without automatically providing the ability to decrypt them. On the basis of the method employed for generating the public and private keys, the original parity matrix, which embeds the logo symbols into the cover, can be divided into a left matrix, transformed parity matrix, and right permutation matrix. When the master wants to send an encrypted message to a slave, the transformed parity matrix and the left matrix are sent. The slave can use these matrices to embed the logo symbols. The cover embedded by the transformed parity matrix and left matrix is called the stego. The stego is secret when being retransmitted to the master as long as the transformed parity matrix and left matrix are unknown. In the master, the stego can be found using the right permutation matrix and original parity matrix.

For user verification, nonshared selection channels are used to not only transmit data from a slave to the master but also verify the authorized users in the slave. Nonshared selection channels in steganography are also called wet paper codes [12]. In the slave, the transmission is slowed down for marginally modifying the selection channel but not the forbidden cover. During transmission, the stego dries out; thus, the receiver has no information about which changeable covers are dry. Steganography schemes with nonshared selection channels can be realized using the STME algorithm [21], which is similar to ME in the sense that the message is transmitted as stego bits for some linear codes.

The remainder of this paper is organized as follows. In Section 2, coding theory and the bound of the embedding scheme are briefly discussed. In Section 3, the authentication steganographic system is described. In Section 4, the experimental results are provided.

Moreover, this section presents a constructive discussion on the performance of various embedding algorithms. Finally, the conclusions are presented in Section 5.

2. **Preliminaries.** For source coding by using linear codes, an (n, k) linear code C can be specified as the null space of a parity check matrix $H \in \{0,1\}^{m\times n}$. The code C consists of 2^{n-m} codewords. Theoretically, the codewords of C can be regarded as a quantized message set $C = \{c\}$ for an arbitrary cover $y \in \{0,1\}^n$ over a binary symmetric source. The objective of a quantizer solves the quantization problem as described in the following. Given a linear code C and an arbitrary sequence y, the optimal quantizer is used to locate the codeword $c_{opt} \in C$ that is closest to the specified sequence y as follows:

$$c_{opt} = \arg\min_{c\in C} d(c, y) \tag{1}$$

According to Equation (1), the average quantization error per a bit is the Hamming distance $d = E[d_H(\hat{c}, y)]/n$. In rate distortion theory, an optimal trade-off between the compression rate $R = k/n$ and the average quantization error d is achieved using the function $R(d) = 1 - h(d)$, where $h(d) = \delta \log_2(1/d) + (1-d)\log_2(1/(1-d))$ denotes a binary entropy function. The geometric interpretation is displayed in Figure 1.

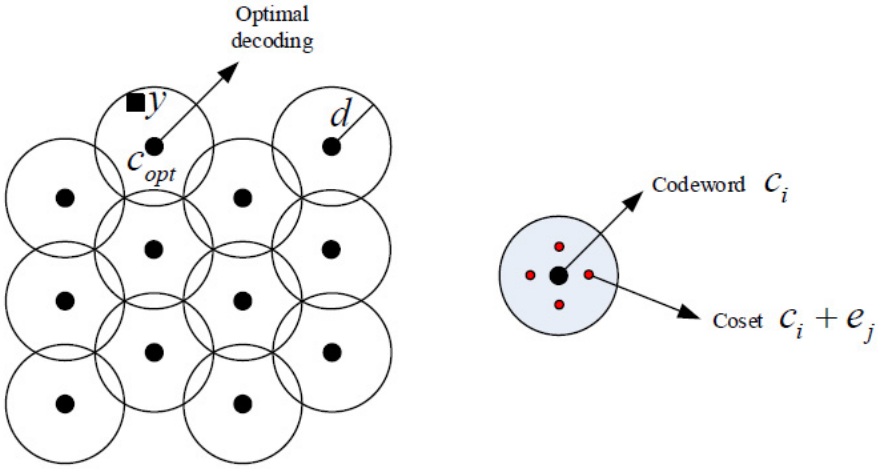


FIGURE 1. Coding for the theoretical bound $R(d) + h(d) = 1$ in geometric interpretation

In the embedding procedures, the y that is closest to the quantized codeword c_{opt} is identified using the embedding message s through quantization. The embedding algorithm with linear codes is described in the following text. For an embedding scheme with linear codes, an (n, k) linear code C is implemented at the embedding rate $R_e = (n-k)/n = m/n$. Assume that a message sequence $s \in \{0,1\}^m$ is embedded into a cover sequence $y \in \{0,1\}^n$ by using the linear code C and transmitted to the receiver. Under this assumption, the optimal stego sequence can be expressed as follows: $l' = y - e_{opt}$, where e_{opt} is provided by an embedder, that is, a stego l', modified from y and corresponding to the message s. These can be formulated as a rate distortion problem. Assuming that the linear codes C at the embedding rate R_e correspond to an embedding average distortion d, the theoretically achievable bound is derived to be $k/n \geq 1 - h(d)$.

Thus, the theoretical bound of the embedding rate is $h(d) \geq R_e$, and the minimal average distortion d is bounded as follows:

$$d \geq h^{-1}(R_e) = \delta \tag{2}$$

where the bound of d is $\delta \leq d \leq 0.5$ and $h^{-1}(*)$ is the inverse binary entropy function. The geometric interpretation of the embedding of the theoretical bound $R(d) + h(d) = 1$ is illustrated in Figure 2.

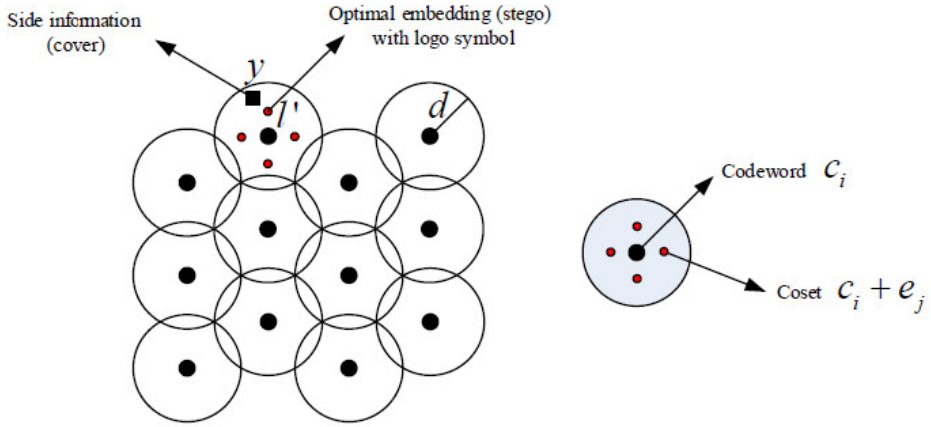


FIGURE 2. Geometric interpretation of the embedding of the theoretical bound $R(d) + h(d) = 1$

Without loss of generality, the embedding efficiency can be defined as follows:

$$\eta = \frac{R_e}{d} = \frac{m}{D} \tag{3}$$

where $D = nd$ is the average embedding distortion per each block. By using Equation (2), Equation (3) can be expressed as an asymptotic upper bound as follows:

$$\eta \leq \frac{R_e}{h^{-1}(R_e)} = \eta_\delta \tag{4}$$

where η_δ is the bound of the embedding efficiency.

The embedding efficiency formula deduced in this section serves as the primary basis for evaluating the embedding system. The next section introduces the focus of this study, that is, the authentication steganographic system. The embedding efficiency of this system is influenced by the number of users of the authentication steganographic system. The higher the number of users of this system, the lower is its embedding efficiency.

3. Authentication Steganographic Systems.

As per coding theory, the traditional embedding scheme with a random parity matrix is subject to the decoding problem. The general problems associated with coding theory and the embedding procedure that determines the security of code-based methods are as follows:

1) General decoding problem
2) Syndrome decoding problem

These problems were proved to be NP-complete problems for binary codes by Berlekamp et al. [18] and for codes over all finite fields by Barg [19]. In channel coding, the general decoding problem is a crucial problem. The general and syndrome decoding problems must be simultaneously considered in the embedding scheme. Syndrome decoding is the first step in an ME algorithm [6]; then, general decoding is performed for an arbitrary toggle sequence obtained using syndrome decoding. The syndrome decoding problem can be solved by a code with a systematic form, as described in Section 3.1. Random linear codes do not meet the condition that the parity check matrix must be in a systematic form;

however, random codes with a systematic form can be generated through Gaussian elimination. Gaussian elimination does not change the characteristics of random linear codes; therefore, the steganographic scheme based on random linear codes can be embedded in a systematic form.

3.1. **ME with a systematic parity matrix.** A code is called a binary linear embedding code when a parity check matrix is used. Such a code is constructed for two main purposes: 1) to find a well-defined coding structure or a well-behaved parity check matrix and 2) to perform decoding with an ML algorithm. Given a cover vector and logo vector intended for embedding, the syndrome of the cover vector must be found and then added to that of the logo vector to acquire a toggle syndrome. Ultimately, the coset leader corresponding to the toggle syndrome can be found through ML decoding. The coset leader is then added to the cover vector to obtain the closest vector, into which a secret logo vector is embedded. An (n, k) linear block code C can be characterized using a parity check matrix H of size $(n-k) \times n$, as described in the following text. A cover vector exists corresponding to an arbitrary sequence y with a length of n bits within the coset C^y of the standard array. The syndrome $s_y = Hy^T$ corresponding to C^y is called the cover vector syndrome. A known binary sequence s_l with a length of $n-k$ bits, which is also referred to as the logo vector, is employed for embedding. The coset leader e_{opt} must be located within a set C^x located ahead of a sequence that is closest to y and contains syndrome s_l. Then, the syndrome s_x is determined by adding the logo vector s_l with s_y. From the viewpoint of decoding, the coset leader e_{opt} can be determined through ML decoding as $e_{opt} = f_{opt}(s_y + s_l) = f_{opt}(s_x)$.

Suppose that a sequence $x \in C^x$ exists and represents a coset of the code. The parameter x represents the minimal weight; thus, $x = e_{opt}$, which can be expressed as follows:

$$e_{opt} = \arg \min_{x \in C^x} w(x) \tag{5}$$

After the coset leader e_{opt} is determined, this parameter is added to the cover vector y as follows: $l' = y + e_{opt}$. Essentially, l' is the sequence closest to the sequence y within an F_2^n dimensional space and contains the logo vector s_l. Consider a systematic parity check matrix $H = [I_{m \times m}\ P]$. Suppose that a logo vector l with a length of n bits exists within the coset C^l of code C. Regard s_l as the logo vector with a length of m bits intended for embedding and with k number of 0s added to its right. A vector l of length n bits is thus formed ($l = [s_l\ 0\ \cdots\ 0]$ and $Hl^T = s_l$). Consequently, the vector $l \in C^l$ corresponding to the sequence with a length of m bits can be found. For the systematic form, a vector $l \in F_2^n$ corresponding to the syndrome s_l can be determined with ease. The addition of l to the cover vector $y \in F_2^n$ yields $x \in C^x$, which is decoded for acquiring e_{opt} within C^x. The ML algorithm for embedding a binary linear code is described in the following text.

Algorithm 1: Optimal systematic embedding algorithm:

Given a systematic linear code C with a parity matrix H_s, a message symbol s_l with m bits, a cover sequence y with n bits, and a stego sequence with the n bits closest to the cover sequence y corresponding to the syndrome s_l can be located as described in the following text.

1. The systematic vector $l = [s_l\ 0\ \cdots\ 0]$ in C^l, derived from s_l, is subtracted from y to obtain x as follows: $x = y - l$.
2. The vector x is then decoded by the ML algorithm into a codeword c as follows: $\hat{c} = \arg \min_{c \in C} d(c, x)$.
3. The parameter x is subtracted from $\hat{c}$ to obtain e_{opt} as follows: $e_{opt} = x - \hat{c}$.
4. The parameter l' is obtained by subtracting e_{opt} to y as follows: $l' = y - e_{opt}$.
5. The embedded data are then extracted as per the following expression: $s_l = Hl'^T$.

When e_{opt} is known, the optimal embedding vector l' can be determined. However, finding e_{opt} in the case of a long (n, k) linear code C or a large value of k remains difficult because the complexity of ML decoding increases exponentially as 2^k. Thus, a suboptimal embedding algorithm is proposed in the following section to replace the ML algorithm for overcoming the aforementioned disadvantage.

3.2. **ME algorithm based on random codes in a systematic domain.** In code-based steganographic systems, generalized ME is employed, where the parity matrix H is divided into a left matrix L, right matrix P, and systematic matrix H_s. In this section, we present a detailed description of such systems. The embedding procedure in a code-based steganographic system consists of the following terms.

1) The parameter H denotes an $(n-k) \times n$ parity matrix of an (n, k) random code C defined over F_q. This matrix can embed m message symbols over F_q, and a simple embedding algorithm can be applied to it.
2) The matrix P denotes an $(n \times n)$ permutation matrix.
3) The matrix L denotes an $(n-k) \times (n-k)$ nonsingular matrix.

The embedding algorithm consists of the following components:

1) An embedding matrix H_s generated by the matrix $H = LH_sP$;
2) Embedded message symbols (s'_l) generated by the original message symbol s_l and left matrix L.

Determining the toggle vector x in the equation $Hx^T = s_x$, where $s_x = s_u - s_l$, is a difficult problem in large-size random ME (RME). The problem of syndrome decoding is NP-complete [18,19]. When determining the toggle vector x by using syndrome decoding becomes difficult, the embedding algorithm becomes incapable. Fortunately, the problem can be transformed using the Gaussian elimination method into a systematic decoding problem so that the embedding algorithm with a random parity check matrix can be applied. However, this problem is simple as long as H_s is in a systematic form. The arbitrary vector l corresponding to the logo s_l can be obtained for parity with the systematic form, and the toggle vector can then be determined as follows: $x = y - l$. The embedding algorithms with systematic form parity have the advantage that the time complexity of the embedding is low. The embedding algorithm for a parity matrix with a systematic form is described in the following text. The coset leader x is subtracted to the cover vector y to obtain the closest stego vector v, into which a secret logo vector s_l is embedded as follows:

$$Hv^T = s_l \tag{6}$$

The parity matrix H can be divided as follows:

$$LH_sPv^T = s_l \tag{7}$$

where H_s is a parity matrix with a systematic form. In the aforementioned equation, the left inverse matrix L^{-1} is multiplied as follows:

$$H_sPv^T = L^{-1}s_l \tag{8}$$

The new transformed vector v', i.e., new stego vector, and transformed logo vector s'_l can be defined as follows:

$$v' = Pv^T \tag{9}$$

and

$$(s'_l)^T = L^{-1}s_l^T \tag{10}$$

Finally, the embedding equation with a systemic form can be expressed as follows:

$$H_sv'^T = s'_l \tag{11}$$

The embedding algorithm can be used to obtain the toggle vector x from the systematic form matrix H_s without performing syndrome decoding. The toggle vector x can be decoded using the decoding algorithm to determine the optimal toggle vector e_{opt} as follows: $v' = y - e_{opt}$. Therefore, the stego vector can be expressed as follows: $v' = y - e_{opt}$. The new stego vector v' is then retransmitted from a slave to the master, and the master inverses the stego v' as follows:

$$P^{-1}v' = v^T \tag{12}$$

The expression $H_s = L^{-1}HP^{-1}$ is substituted into the equation $H_s v' = s'_l$, and the permutation matrix P is multiplied as follows:

$$L^{-1}HP^{-1}\left(Pv^T\right) = L^{-1}s_l \tag{13}$$

Finally, the embedding logo vector can be obtained as follows:

$$Hv^T = s_l \tag{14}$$

Although the public keys H_s and L can be acquired by any user, the master can obtain the logo vector s_l only through the private keys H and P. When the master finds the secret logo vector, nonrepudiation is achieved for the master. For achieving nonrepudiation for the slave, nonshared selection channels are used.

In the slave, the logo vector is embedded into the cover vector as a stego vector; then, the stego is retransmitted to the master. Although H_s and L cannot be appropriated by a pretender, they can impersonate a slave to transmit fake data to the master. To avoid this situation, the nonshared selection channel is used. In the slave, the stego vector is found using the following equation:

$$H_s v' = s'_l \tag{15}$$

where

$$H_s = [I, P_\gamma] \tag{16}$$

Assume that the subchannel of the ith user is defined as follows:

$$S_i \subset \{1, 2, \ldots, n\} \tag{17}$$

and the size of subchannel is $|S_i| = \lambda$. Let $m \leq \lambda \leq n$ such that

$$H_i = \left[\begin{array}{cc} I & M_i \end{array}\right]_{m \times \lambda} \tag{18}$$

where M_i is a random matrix of size $m \times (\lambda - m)$ and some of the columns of M_i correspond to the position of the stego vector v'. Thus, M_i contains some columns from the parity matrix H_s and can be expressed as follows:

$$M_i \subseteq P_\gamma \tag{19}$$

The STME algorithm is an embedding algorithm with a subchannel [14]. The equation can be solved using the STME algorithm as described in the following text. Suppose that $n - k$ linearly independent row vectors are present within the parity check matrix H; that is, $\text{Rank}(H) = n - k$. For an (n, k) linear code over F_q, $S \subseteq \{1, 2, \ldots, n\}$, $|S| \geq m$, $S_\theta \subseteq S$, and $|S_\theta| = m$. A column vector m is selected randomly from H and verified as linearly independent. The expression $\theta = \{\theta_i | i \in S_\theta\}$, where $|\theta| = m$, is used for representing an arbitrary m toggle vector s_x. Assuming that an m syndrome vector $s_x = (s_{x,1}, \ldots, s_{x,m})$ corresponding to H is present, we can obtain an independent matrix $\theta = \{\theta_i | i \in S_\theta \subset n, |S_\theta| = m\}$ from m columns out of H such that the following equation is obtained:

$$s_x^{\mathrm{T}} = \begin{bmatrix} \theta_{1,i_1} & \theta_{1,i_2} & \cdots & \theta_{1,i_m} \\ \theta_{2,i_1} & \theta_{2,i_2} & \cdots & \theta_{2,i_m} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{m,i_1} & \theta_{m,i_2} & \ldots & \theta_{m,i_m} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix} \tag{20}$$

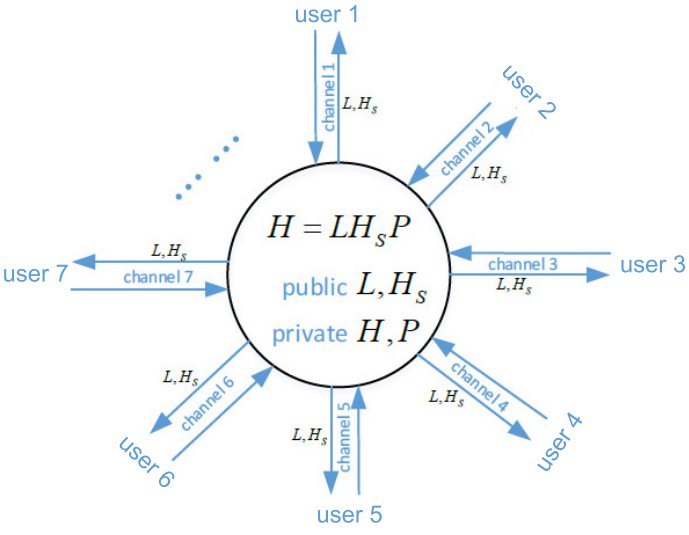
where $\lambda_i \in F_q$ and $\theta_i = (\theta_{1,i}, \theta_{2,i}, \ldots, \theta_{m,i})^{\mathrm{T}}$. Given θ and s_x, the coordinates $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ corresponding to θ can be evaluated as follows:

$$\lambda^{\mathrm{T}} = \begin{bmatrix} \theta_{1,i_1} & \theta_{1,i_2} & \cdots & \theta_{1,i_m} \\ \theta_{2,i_1} & \theta_{2,i_2} & \cdots & \theta_{2,i_m} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{m,i_1} & \theta_{m,i_2} & \ldots & \theta_{m,i_m} \end{bmatrix}^{-1} \begin{bmatrix} s_{x,1} \\ s_{x,2} \\ \vdots \\ s_{x,m} \end{bmatrix} \tag{21}$$

By using Equation (21), we can obtain the solution $x' = (x_1, \ldots, x_n)$ for $Hx'^{\mathrm{T}} = s_x$. Subsequently, we construct the ith component of x' as follows: if $i \in S_\theta$, then $x_i = \lambda_i$, and if $i \notin S_\theta$, then $x_i = 0$. Thus, $Hx'^{\mathrm{T}} = \theta\lambda^{\mathrm{T}} = s_x$, where θ is the submatrix of H. The original equation $Hx^{\mathrm{T}} = s_x$ can also be resolved for x by using a linear combination of column vectors from H. Next, the least weight corresponding to the minimum embedding distortion is determined. Therefore, the least weight coordinate vector λ associated with a randomly selected basis θ can be obtained. Consequently, the adaptive toggle vector x corresponding to the least weight coordinate vector λ can be obtained. The output l' of the embedder, that is, the stego vector, is subsequently obtained as follows: $l' = y - x$. Finally, at the receiver, the secret message $s_{l'}$ is extracted as follows:

$$s_{\hat{l'}} = Hl'^{\mathrm{T}} = H(y - x)^{\mathrm{T}} = Hy^{\mathrm{T}} - \theta\lambda^{\mathrm{T}} = s_{l'} \tag{22}$$

In Figure 3, each user in the slave possesses only one secret subchannel to embed the logo vector. The user in the slave can be authenticated in the master as long as the selection subchannel cannot be detected.



$S_i \subset \{1, 2, \ldots, n\}$
$S_i = \lambda$ and $n \geq m$

FIGURE 3. Client transmission with private selection

Finally, the algorithm for authentication steganographic systems is described in the following text.

Algorithm 2: ME embedding with a random parity check matrix

(1) Master: send H_s and L
Decompose $H = LH_sP$ and send H_s and L
(2) Slave: receive H_s, L, and genetate S_i
1. Given s_l, determine $s'_l = L^{-1}s_l$
2. Use H_s to embed into subchannel S_i, s'_l into the cover y as the stego v' such that $v' = Emb(y, s'_l)$ over subchannel S_i, and send v', S_i to the master
(3) Master: receive S_i and v'
1. Use S_i, v', and P to determine $v = P^{-1}v'$
2. Use H to extract the logo sequence $s_l = Hv^{\mathrm{T}}$

3.3. **Embedding for small payloads.** In addition to the STME algorithm, some efficient embedding algorithms with systematic parity matrices have been developed for authentication steganographic systems. A simple method is presented in [17] for locating a low-weight toggle vector during the search of the coset leader vector e_{opt}. This method aims to locate the vectors e_{sub} and $w(e_{sub}) \geq w(e_{opt})$ according to the optimal coset leader $w(e_{opt})$. The vector e_{sub} is defined in C^x and must stay as close to $w(e_{opt})$ as possible. The target vector l', which is obtained through the addition of e_{sub} to the cover vector y, cannot be employed as the optimal vector. Although the linear block code is simple when used as an embedding algorithm, this article proposes an embedding algorithm based on the parity check matrix, which has low complexity and can be executed rapidly.

In [20], a method with low computational complexity was proposed for finding the suboptimal stego. This method is called weight approximation embedding (WAE), which can be used for RME with small payloads. In contrast to the ML embedding algorithm, in WAE, a minimum weighted toggle vector is searched in an iterative manner. In the ML algorithm, the minimum weight is obtained by performing a complete search of the codewords within the linear embedding code, whereas in WAE, a minimum weighted toggle vector is found in each iteration. In WAE, k numbers of hardware operations are performed in each iteration in an iterative manner until convergence is reached. WAE has a lower operation complexity but lower distortion efficiency than does the ML algorithm.

As discussed in Section 3, in the ML algorithm, $f(s_x)$ is decoded to obtain the codeword c and the coset leader vector e is obtained by adding c to $x' = x + c$. The parameter $f(s_x)$ is not directly decoded through ML decoding; instead, the syndrome s_x of the toggle x is maintained invariant. To achieve this in the simplest manner, the codeword c from the linear code C is added to the toggle x to obtain x' (i.e., $x' = x + c$). Therefore, the weight of x' is altered through the codeword c; however, x' still falls within the coset C^x. A total of 2^k c codewords are present in the linear code C, and testing all of them is unrealistic. Thus, for testing, only k number of codewords are selected from among the 2^k codewords. The selected codewords form the row vector g_i of a systematic generator matrix G_s.

4. **Simulation Results.** Simulations of authentication steganographic systems are performed using the client-server networking model, as displayed in Figure 4.

RME with a random parity matrix H of size $m \times n$ over F_q is described in the simulation. A total of $q^{m \times n}$ such matrices exist in the random matrix H. The following parameter settings are used: $q = 3$, $n = 13$, $m = 3$, $k = n - m = 10$. The authentication steganographic system introduces ternary random-like codes with parity matrix as follows:

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 0 & 0 & 1 & 1 & 2 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \end{pmatrix}$$

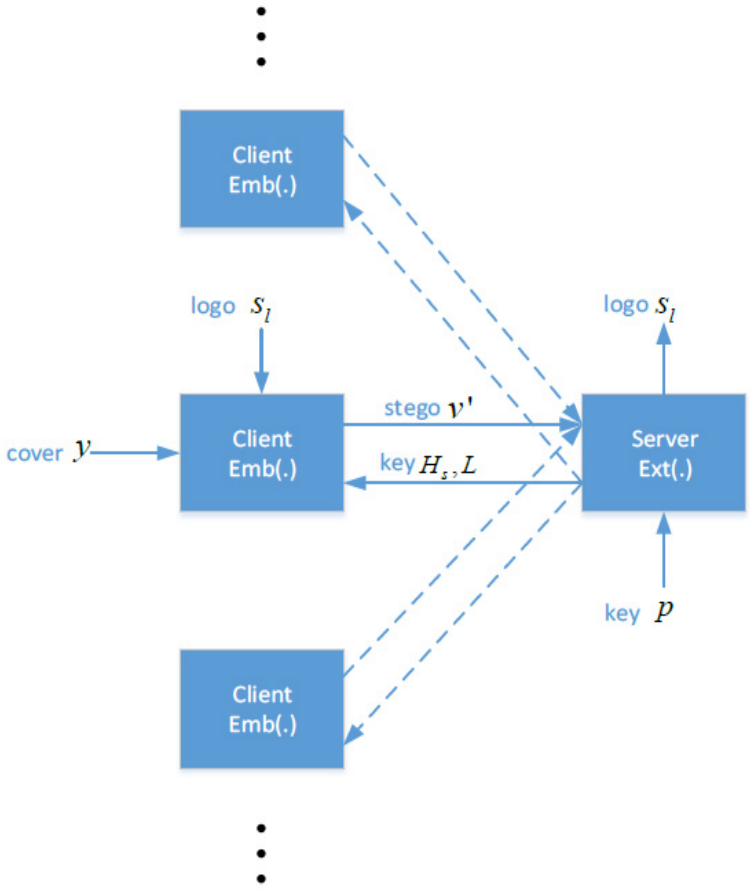


FIGURE 4. Block diagram of the authentication process in a steganographic system

Next, the following $n \times n$ permutation matrix P is selected ($n!$ possibilities exist):

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The following invertible $m \times m$ matrix (L) is selected:

$$L = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 2 \end{pmatrix}$$

In the server, the matrices L and H_s are public, whereas the matrices H and P are kept secret. The public keys L and H_s are transmitted to the client and used to embed the secret logos s_l in the client. In the simulation, the STME [14] embedding algorithm embeds

$m = 3$ symbol messages $s_l = (1, 2, 0)$ into the cover $y = (2, 0, 2, 1, 1, 2, 1, 0, 0, 2, 0, 0, 1)$ with a length of $n = 13$. For an embedding scheme with linear random codes, a Hamming code C over field F_3 is used to embed a secret message $s_l = (1, 2, 0)$. To generate a public key for each client, the parity matrix in the server can be divided into three matrices as follows:

$$H_s = LHP = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 2 & 2 & 0 & 0 & 1 & 2 & 2 & 2 \\ 0 & 1 & 0 & 1 & 2 & 0 & 1 & 1 & 2 & 1 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 & 1 & 2 & 0 & 2 & 2 & 1 & 1 & 0 & 1 \end{pmatrix} \tag{23}$$

such that

$$\begin{aligned} s'_l = H_s v' &= \begin{pmatrix} 1 & 0 & 0 & 1 & \cdots & 2 & 2 & 2 \\ 0 & 1 & 0 & 1 & \cdots & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 & \cdots & 1 & 0 & 1 \end{pmatrix} v' \\ &= L^{-1} s_l = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \end{aligned} \tag{24}$$

Assume that a secret message sequence $s'_l = (2, 1, 1)$ corresponding to a logo vector $l = (2, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ is embedded into a cover sequence $y = (1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ by using C and is transmitted to the receiver. The vector $x = y - l = (0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 2, 0)$ is obtained by modifying u corresponding to the embedded message sequence s'_l. Finally, the stego vector $v' = y - x = (1, 2, 0, 1, 1, 1, 2, 0, 1, 1, 0, 1, 2)$ is obtained using the optimal toggle vector e_x. The client assumes that the secret logos are embedded into the cover u as the stego v' by using the STME algorithm, and these logos are sent to the server through a secret subchannel. The client uses a secret subchannel to prevent cheating. The stego can be recovered from the private matrices P and H when the stego v' is transmitted to the server. The secret logos can be extracted as follows:

$$v = P^{-1} v' = (1, 2, 0, 1, 2, 0, 1, 1, 2, 1, 0, 1, 1)^{\mathrm{T}} \tag{25}$$

Finally, the following equation is obtained:

$$s_l = H v^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & \cdots & 2 & 2 & 2 \\ 0 & 0 & 1 & 1 & 2 & \cdots & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 1 \end{pmatrix} = (1, 2, 0)^{\mathrm{T}} \tag{26}$$

The authentication steganographic system is simulated in the application (app) window in the MATLAB 2021 program, as displayed in Figure 5.

To maintain a high embedding efficiency for authentication steganographic systems, q-ary Hamming codes are used as random-like codes. The embedding rate of the experimental result can be expressed as follows:

$$\alpha = \frac{m \log_2 q}{(q^m - 1)/(q - 1)} \tag{27}$$

The embedding efficiency can be defined as follows:

$$\eta = \frac{m \log_2 q}{1 - q^{-m}} \tag{28}$$

where m denotes the length of the logo. The embedding efficiency bound of ternary Hamming codes is displayed in Figure 6, in which the embedding efficiency η is a function of

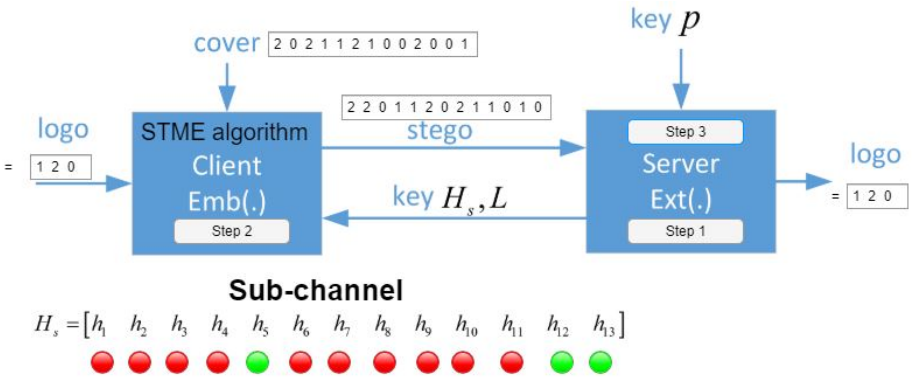


FIGURE 5. App simulation of the authentication steganographic system

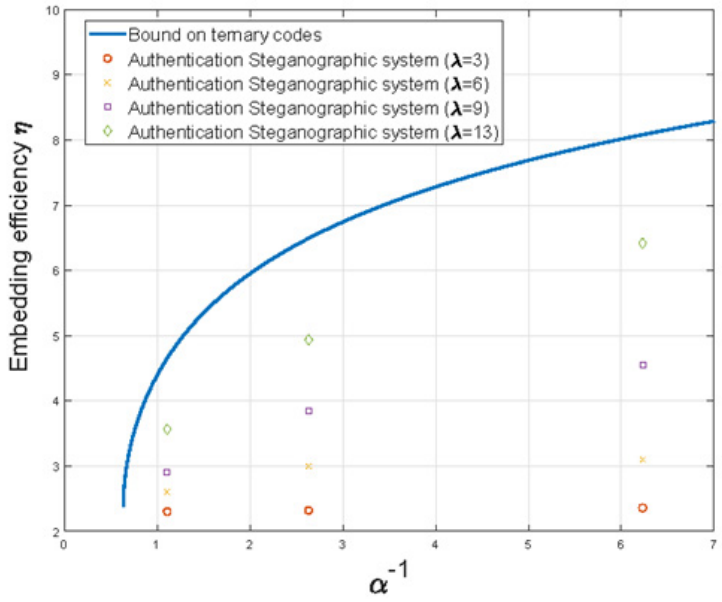


FIGURE 6. Plot of the embedding efficiency versus the embedding rate

α^{-1}, where α is the relative message length. The graph in Figure 6 displays the performance of embedding for authentication steganographic systems. In the simulation, the parameter λ is changed to estimate the embedding performance. The embedding efficiency η is presented as a function of λ in Figure 6 and decreases with λ; however, the proposed algorithm requires a high number of iterations for searching a linear independent column vector.

The security of the logos is dependent on the selection of λ by the subchannel from the client to the server. The range of λ is presented in Table 1. The higher the value of λ, the higher is the number of search iterations required to reach a high level of security and the higher is degradation of q. Therefore, the value of λ can be varied to change the operation complexity when the algorithms are executed.

In this study, the q-ary Hamming code was used as a substitute for the Hamming code in the experiment on the authentication steganographic system. The authors of [21] proposed a probability experiment for determining the linear independent basis of a random code. When the authentication steganographic system is used by multiple users, one must ensure that each user receives a set of public keys. If random codes are used in the STME algorithm, an extremely long random code or a small number of embedded

TABLE 1. λ values of $\left(\frac{q^m-1}{q-1}, \frac{q^m-1}{q-1} - m\right)$ q-ary Hamming codes for various numbers of secret subchannels

R_e	q	m	n
5.97×10^{-4}	13	5	30941
0.0062	13	4	2380
0.0607	13	3	183
0.0011	11	5	16105
0.0095	11	4	1464
0.078	11	3	133
8.59×10^{-4}	7	6	19608
0.005	7	5	2801
0.0281	7	4	400
0.1478	7	3	57
8.32×10^{-4}	5	7	19531
0.0036	5	6	3906
0.0149	5	5	781
0.0595	5	4	156
0.2247	5	3	31
5.36×10^{-4}	3	10	29524
0.0014	3	9	9841
0.0039	3	8	3280
0.0102	3	7	1093
0.0261	3	6	364
0.0655	3	5	121
0.1585	3	4	40
0.3658	3	3	13
8.54×10^{-4}	2	12	16383

symbols results in a high probability of a linear independent basis [21]. If the q-ary Hamming code is used instead in the STME algorithm, the aforementioned shortcoming can be overcome because the parity check matrix is composed of all linear independent row vectors in $q^m - 1$; thus, each set of linear conversion matrices contains an inverse matrix. Consequently, each steganographic symbol can be mapped to the location that must be concealed. The solutions of the STME algorithm yield a probability level because of the coordinate transform of the submatrix selected from the parity check matrix of devising the adaptive STME algorithm by using random codes. The parameter H can be deleted from numerous columns to form $(n', k) = (n - i, k)$ random codes at a fixed embedding rate of $\alpha = \frac{m \log_2 q}{(q^m-1)/(q-1)}$, where i is the number of deleted columns. In the case of the same q-ary, the solution probability increases as the cover size n increases, as shown in Figure 7. If the Hamming code is used in the STME algorithm, the Hamming code can achieve a 100% embedding possibility. In the authentication steganographic system, the use of the Hamming code is more reliable than that of random codes.

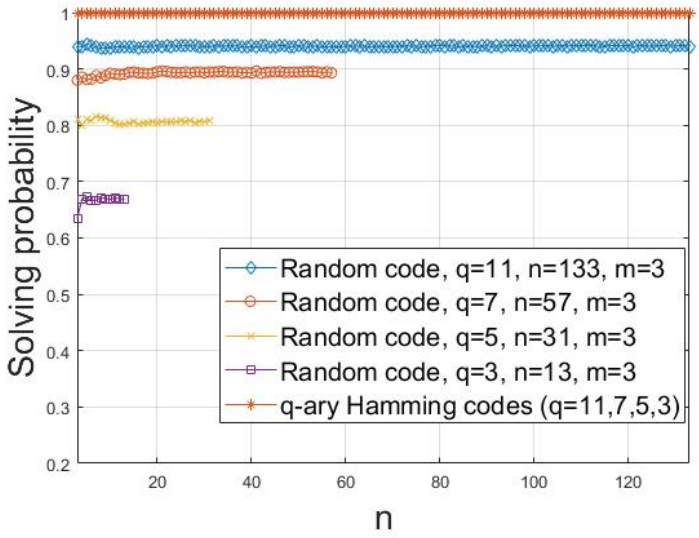


FIGURE 7. Solving probability of the adaptive STME algorithm

5. **Conclusions.** In this study, an alternative to secret steganographic schemes, namely authentication steganographic systems, is proposed. Designing a secret steganographic system is essential for achieving a suitable public key system. Therefore, a reliable authentication steganographic system is presented for implementation in security communication. Compared with traditional ME systems, authentication steganographic systems offer the following advantages: (i) the use of a public system at the server end and (ii) the use of a nonshared selection channel for protecting from cheating. This study found that authentication steganographic systems can use the STME algorithm to generate a nonshared selection channel for users in the client. The simulation results indicate that the proposed scheme not only embeds the secret logo message but also enables authentication in the server. Furthermore, the proposed scheme involves a realistic method for designing a client-server model for networking.

REFERENCES

[1] J. Fridrich and T. Filler, Practical methods for minimizing embedding impact in steganography, *Proc. of SPIE, Secur., Steganogr., Watermarking of Multimed. Contents IX*, San Jose, CA, vol.6050, pp.2-3, 2007.
[2] R. Crandall, Some notes on steganography, *Steganography Mailing List*, 1998.
[3] J. Bierbrauer, *Crandall's Problem*, http://www.ws.binghamton.edu/fridrich/ covcodes.pdf, 1998.
[4] F. Galand and G. Kabatiansky, Information hiding by coverings, *Proc. of ITW2003*, Paris, France, pp.151-154, 2003.
[5] J. Bierbrauer and J. Fridrich, Constructing good covering codes for applications in steganography, *LNCS Transactions on Data Hiding and Multimedia Security*, vol.4920, pp.1-22, 2008.
[6] J. Fridrich and D. Soukal, Matrix embedding for large payloads, *IEEE Trans. Inf. Forensics Secur.*, vol.1, no.3, pp.390-394, 2006.
[7] J. Chen, Y. Zhu, Y. Shen and W. Zhang, Efficient matrix embedding based on random linear codes, *Proc. of MINES2010*, Nanjing, China, pp.879-883, 2010.
[8] Y. Gao, X. Li and B. Yang, Employing optimal matrix for efficient matrix embedding, *Proc. of IIH-MSP2009*, Kyoto, Japan, pp.161-165, 2009.
[9] T. Hiraoka, A method for embedding another photographic image in a photographic image, *ICIC Express Letters*, vol.14, no.4, pp.311-317, 2020.
[10] S. Dash, M. Das and D. K. Behera, High capacity information hiding using enhanced difference expansion technique, *ICIC Express Letters*, vol.15, no.8, pp.819-827, 2021.

[11] D. SchÄonfeld and A. Winkler, Embedding with syndrome coding based on BCH codes, *Proc. of ACM 8th Workshop on Multimed. Secur.*, pp.214-223, 2006.
[12] J. Fridrich, M. Goljan and D. Soukal, Efficient wet paper codes, *The 7th Int. Workshop on Inf. Hiding, Lect. Notes Comput. Sci.*, Barcelona, pp.204-218, 2005.
[13] J. Fridrich, Asymptotic behavior of the ZZW embedding construction, *IEEE Trans. Inf. Forensics Secur.*, vol.4, no.1, pp.151-154, 2009.
[14] W. Zhang and X. Wang, Generalization of the ZZW embedding construction for steganography, *IEEE Trans. Inf. Forensics Secur.*, vol.4, pp.564-569, 2009.
[15] T. Filler and J. Fridrich, Wet ZZW construction for steganography, *IEEE Workshop Inf. Forensic Secur. (WIFS)*, London, UK, 2009.
[16] Y. Song, H. Ni and X. Zhu, Analytical modeling of optimal chunk size for efficient transmission in information-centric networking, *International Journal of Innovative Computing, Information and Control*, vol.16, no.5, pp.1511-1525, 2020.
[17] M. Ntahobari and T. Ahmad, Securing computer network by increasing the performance of intrusion detection system, *ICIC Express Letters*, vol.14, no.12, pp.1217-1223, 2020.
[18] E. Berlekamp, R. J. McEliece and H. Van Tilborg, On the inherent intractability of certain coding problems, *IEEE Trans. Inf. Theory*, vol.24, no.3, pp.384-386, 1978.
[19] A. Barg, Complexity issues in coding theory, *Electronic Colloquium on Computational Complexity (ECCC)*, vol.4, no.46, 1997.
[20] H.-Y. Chang, J.-J. Wang, C.-Y. Lin and C.-H. Chen, Development of low-complexity matrix embedding with an efficient iterative strategy, *ICIC Express Letters, Part B: Applications*, vol.9, no.7, pp.707-713, 2018.
[21] H.-Y. Chang, J.-J. Wang, C.-Y. Lin and C.-H. Chen, An efficient matrix embedding technique by using submatrix transform for grayscale images, *International Journal of Innovative Computing, Information and Control*, vol.15, no.4, pp.1565-1580, 2019.

Author Biography

Hsi-Yuan Chang is a general manager of the Company of the Jet-Dar at Taichung, Taiwan. He received the M.S. degree in Innovation Technology and Information Management from National Chin-Yi University Technology in 2013. He is currently a Ph.D. Candidate of Institute of Computer and Communication Engineering, Department of Electrical Engineering, National Cheng Kung University. His research interests include multimedia coding, data hiding, and image processing.

Jyun-Jie Wang received the B.S. degree in Electronic Engineering from National Chi-Yi University of Science and Technology, Taiwan, in 2003, the M.S. degree and the Ph.D. degree in Electrical Engineering from National Chung Hsing University, Taiwan, in 2005 and 2012, respectively. He is also a member of IEEE. His research interests include multimedia, image processing, watermarking, information theory and coding theory.

Chin-Hsing Chen received the B.S. degree in Electrical Engineering from National Taiwan University, Taiwan, in 1980, and the M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of California at Santa Barbara, in 1983 and 1987, respectively. Since 1988, he has been with the Department of Electrical Engineering at National Cheng Kung University in Taiwan, where he is now a professor. His current research interests include pattern recognition, image processing and VLSI array design. He has published over 240 papers and given more than 80 technical presentations in public in more than 15 countries.

Chi-Yuan Lin received the B.S. degree in Electronic Engineering from National Taiwan University of Science and Technology, Taiwan, in 1988, the M.S. degree in Electronic and Information Engineering from National Yunlin University of Science and Technology, Taiwan, in 1998, and the Ph.D. degree in Electrical Engineering from National Cheng Kung University, Taiwan, in 2004. Since 1983, he has been with the Department of Electronic Engineering at National Chin-Yi University of Technology in Taiwan where he is now a professor. His research interests include neural networks, multimedia coding, data hiding, and image processing.