JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 35, 1061-1074 (2019)
DOI: 10.6688/JISE.201909_35(5).0008

A QC-LDPC Code Based Digital Signature Algorithm

FANG REN[1], XUE-FEI YANG[1] AND DONG ZHENG[2]
[1]*School of Telecommunication and Information Engineering*
[2]*National Engineering Laboratory for Wireless Security*
Xi'an University of Posts and Telecommunications
Xi'an, 710121 P.R. China
E-mail: renfang_81@163.com; 451173933@qq.com;
zhengdong@xupt.edu.cn

Courtois-Finiasz-Sendrier (CFS) digital signature algorithm, which proposed in 2001, is the most important code based digital signature algorithm and can resist the known attack of quantum algorithms such as Shor algorithm and Grover algorithm. But the efficiency of CFS is very low because of the extremely low signing speed and the large public key size. In this paper, a variation of CFS algorithm is presented. Instead of the Goppa code and the Patterson decoding algorithm, the new algorithm selects the Quasi-Cyclic Low Density Parity Check (QC-LDPC) code and the Belief Propagation (BP) decoding algorithm in the signing process. Compared with CFS algorithm, the new algorithm greatly reduces the storage space of public key and improves the efficiency of signature without compromising the security.

Keywords: code based cryptography, digital signature, QC-LDPC code, decoding algorithm, CFS algorithm

1. INTRODUCTION

Digital signature technology, used to authenticate the identity of users and ensure the integrity of data, is an important aspect of public key cryptography which has become one of the important technologies in the fields of modern communications, computers and other information technologies. The security of widely used digital signature algorithms, such as Rivest-Shamir-Adleman (RSA) signature algorithm, Digital Signature Algorithm (DSA) signature algorithm and ElGamal signature algorithm, are based on mathematical difficult problems. However, the quantum attack algorithms proposed by Shor [1] and Grover [2] can effectively break the difficult problems these digital signature algorithms depend on, so these algorithms were seriously threatened. Four types of cryptography are considered resistant to the quantum attacks at the first post-quantum cryptography conference [3] in 2006, in which the code based cryptography received extensive attention from many researchers.

CFS digital signature algorithm, proposed by Courtois, Finiasz and Sendrier [4], is the first secure digital signature algorithm based on binary Goppa codes. CFS algorithm is constructed on the basis of the Niederreiter algorithm [5], an important code based public key encryption algorithm. The efficiency of CFS algorithm is very low because of its large public key overhead and the extremely low probability of successful signature. On average, the success probability of signing a random message is only $1/t!$, which t is the error correcting capability of the Goppa code used in CFS algorithm.

Low Density Parity-Check (LDPC) codes was proposed by Gallager [6] in 1962,

Received April 4, 2019; accepted March 20, 2019.
Communicated by Xiaohong Jiang.

which is a type of linear block code in which the parity-check matrix is sparse. Gallager proved that it is a good code with asymptotic behavior and the performance approaches the Shannon limit. However, due to the limitations of the computational capabilities at that time, the LDPC code was considered as an impractical code and was ignored for a very long time. Until 1996, MacKay and Neal proved that LDPC codes are a valuable code [7], and Gallager's probability iterative decoding algorithm was extended. At the same time, Belief Propagation (BP) algorithm used to decode LDPC code was discussed, which greatly promoting the development of LDPC codes. In 2007, Baldi [8] constructed a new code based public key cryptosystem using QC-LDPC codes instead of Goppa codes. The quasi-cyclic structure of the QC-LDPC codes compensates for the large key overhead of the code based cryptographic scheme based on the Goppa code, and effectively reduces the key storage space.

This paper presents a variation of CFS signature algorithm based on QC-LDPC for the first time. Compared with the Goppa code, the parity-check matrix of the LDPC code is sparse, which can greatly reduce the key storage space. The quasi-cyclic structure of the QC-LDPC code makes its decoding complexity lower than that of the Goppa code. In addition, the BP decoding algorithm can be implemented in parallel in hardware, which can increase the decoding speed, thereby improving the signature efficiency of the CFS algorithm. The blind signature [9], ring signature [10], and group signature [11] algorithms, which were constructed based on the CFS signature algorithms using Goppa codes, can also be improved by the new scheme proposed in this paper.

2. PRELIMINARIES

In this section, we give some basic knowledge and conclusions of error correcting code firstly. Then the hard problems in this field and the most important code based digital signature algorithm, CFS algorithm, are given.

2.1 Linear Block Codes

Let F_2 represents the finite field of order 2. A (n, k) linear block code C on F_2 is a k-dimensional subspace of an n-dimensional linear space over F_2. The vectors in F_2^n are called word and the vectors in C are called codeword. n is the code length, and k is the dimension of the code.

If C is a (n, k) linear block code, the matrix G is named a *generator matrix* for C, having size $k \times n$. It should be noted that any $k \times n$ linearly independent codewords can be used to form a generator matrix, so G is not unique for a given code.

The matrix H is denoted as the *parity check matrix* of the (n, k) code C and the size of H is $(n-k) \times n$. $HG^T = O$ expresses the link between G and H, and it is not unique too. Every codeword c of C must satisfy $Hc^T = 0$ and Hc^T is called *syndrome* for any word c.

2.2 Hard Problems

Syndrome Decoding (SD)

For a binary $(n-k) \times n$ matrix $H \in F_q^n$, a vector $s \in F_q^{n-k}$ and an integer $\omega > 0$, is there a word $x \in F_q^n$ that satisfies $Hx^T = s$ and $w(x) \leq \omega$.

The SD problem is NP-complete.

Goppa Code Distinguishing (GCD)

For a binary $(n-k)\times n$ matrix $H\in F_q^n$, it is determined whether H is a check matrix of (n, k) Goppa codes or a random (n, k) codes.

Although there is no rigorous proof yet, most researchers believe that the GCD problem is difficult.

2.3 CFS Signature Scheme

The CFS signature algorithm is the first code based digital signature scheme that can be proved security, and it is constructed based on the Niederreiter public key cryptosystem. Security of the classical CFS signature scheme relies on SD problem and the GCD problem. The specific process of CFS algorithm can be expressed as follows.

(A) *Initialization*

Randomly select an irreducible Goppa code C with parameters (n, k, t), of which the error correction capability is t, described by its parity-check matrix $(n-k)\times n$ H_1, a non-singular matrix $S\in F_2$ of order $(n-k)\times(n-k)$, an $n\times n$ permutation matrix $P\in F_2$ and a hash algorithm h. Let γ represents a valid syndrome decoding algorithm for Goppa codes. The public key is $H^{pub}=S\times H_1\times P$, and the private key is $\langle S, H_1, P, \gamma\rangle$.

(B) *Signature*

Calculate $z=\gamma(S^{-1}h(h(m)\|i))$, where i is the smallest natural number that makes decoding successful. The signature of message m is $\sigma=[z|i]$.

(C) *Verification*

Compute $s_1=H^{pub}z^T$, $s_2=h[h(m)|i]$; If s_1 is equal to s_2, the signature σ is valid; otherwise, σ is invalid.

3. QC-LDPC CODES

The (n, k) LDPC codes $C\in F_2$ is a special kind of linear block codes defined by a $(n-k)\times n$ parity check matrix H. The number of non-zero elements in H is particularly small, and '1' is randomly arranged. That is to say, LDPC codes are a block code with a very low check matrix density. Where n denotes the length of code, k indicates the dimension, r is indicative of the check bit length, and satisfies $r=n-k$.

3.1 Tanner Graph

LDPC codes can be visually represented by the Tanner graph. The Tanner graph consists of n variable nodes and r check nodes. For example, the Tanner graph of a LDPC code with $n=7$, $r=3$ is illustrated in Fig. 1, where $v_j(j=0, 1, \ldots, 6)$ represent the variable nodes and $c_i(i=0, 1, 2)$ represent the check nodes.

The parity check matrix H corresponding to the Tanner graph in Fig. 1 is expressible as follows:

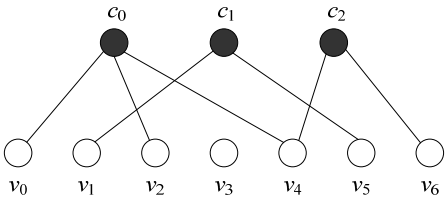

Fig. 1. Tanner graph of a (7,4) LDPC code.

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

There is an edge between c_i and v_j in the Tanner graph if and only If $h_{ij} = 1$ in the check matrix H, which also means that the component at position j of a codeword participates in the ith parity check equation.

3.2 QC-LDPC Codes

We use quasi-cyclic low-density parity-check codes to represent a special class of LDPC codes. Its check matrix has the characteristics of block cycling. Let C be a QC-LDPC code of length $n = n_0 \times p$ and dimension $k = k_0 \times p$, and the parity check matrix can be described as follows:

$$H_{QC} = \begin{bmatrix} A_{0,0} & A_{0,1} & \cdots & A_{0,n_0-1} \\ A_{1,0} & A_{1,1} & \cdots & A_{2,n_0-1} \\ \vdots & \vdots & & \vdots \\ A_{k_0-1,0} & A_{k_0-1,1} & \cdots & A_{k_0-1,n_0-1} \end{bmatrix}. \tag{1}$$

The parity check matrix H_{QC} is arranged from $k_0 \times n_0$ sub-matrices. $A_{ij}(1 \le i \le m, 1 \le j \le n)$ are zero matrixes or circulant matrixes of order $p \times p$. Due to the block-cyclic characteristic of the check matrix of the QC-LDPC code, only the position and cyclic shift bits of each non-zero sub-matrix need to be stored, which significantly reduces the key storage space required in the signing process.

3.3 BP Decoding Algorithm

The Belief Propagation (BP) decoding algorithm of a LDPC code, also called a sum-product algorithm (SPA), is the best-known iterative decoding algorithm for LDPC codes. The main idea is used by the information received during each iteration to continuously transfer information and iterative operations between variable nodes and check nodes for decoding. Compare with other decoding algorithms, the BP decoding algorithm has many advantages.

(1) The BP decoding algorithm can be implemented in parallel. Compared with the Patterson decoding algorithm used by Goppa code, this parallel feature can greatly improve the decoding speed.

(2) The computational complexity of the BP algorithm does not increase rapidly as the code length increases. The longer the code length of the LDPC code, the better the performance of the BP decoding algorithm. Because with the code length increases, the distance between the variable nodes increases, the check matrix becomes more and more sparse, the correlation between the nodes decreases and the decoding performance can be improved.

(3) The BP decoding algorithm is divided into a probability domain BP algorithm and a Log-Likelihood Ratios (LLR) BP algorithm [12]. The difference is that the message of the probability domain BP algorithm is expressed in the form of probability, and the LLR-BP algorithm is represented by the log likelihood ratio. Compared with the probability domain BP algorithm, LLR-BP algorithm converts a large number of multiplication operations into additional operations, which reduce the decoding complexity. LLR-BP algorithm is used in our algorithm.

The decoding processes of BP algorithm are shown in Fig. 2.

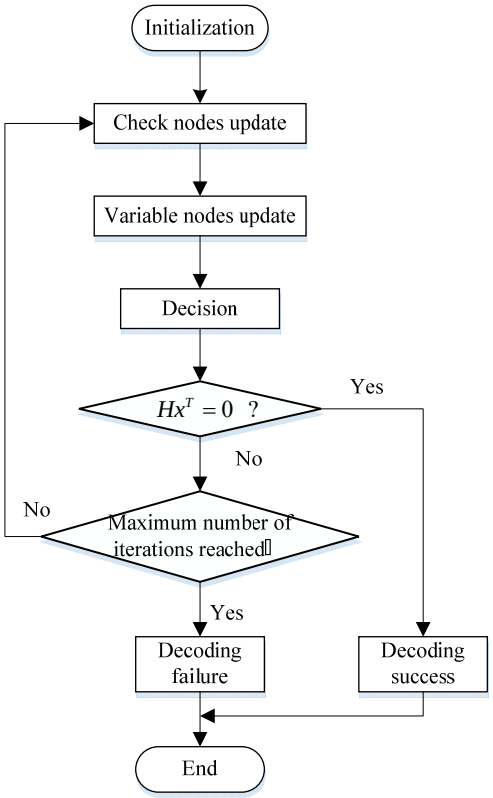

Fig. 2. Flow diagram of BP algorithm.

We first explain the meaning of the symbols used in the LLR-BP algorithm. Let us consider a LDPC code with codeword length n and dimension k and redundancy length r,

having a Tanner graph with n variable nodes $\{v_0, \ldots, v_{n-1}\}$ and r check nodes $\{c_0, \ldots, c_{r-1}\}$. $A(k)$ represents the set of all variable nodes connected with the check node c_k; $A(k)\backslash i$ is the set of all variable nodes connected with c_k, except the variable node v_i; $B(i)$ is the set of all check nodes connected with the variable node v_i; $B(i)\backslash k$ represents the set of all check nodes connected with v_i without the check node c_k.

Let $q_{i\to k}(x)$, $x\in\{0,1\}$ be the probability information that the node v_i sends to the node c_k according to $B(i)\backslash k$ and $r_{k\to i}(x)$, $x\in(0,1)$ be the probability information that the node c_k sends the node v_i according to $A(k)\backslash i$.

In the LLR-BP algorithm, the probability information of the messages sent from variable nodes to check nodes are expressed in logarithmic form, as follows:

$$\Gamma_{i\to k}(x_i) = \ln\left[\frac{q_{i\to k}(0)}{q_{i\to k}(1)}\right]. \tag{2}$$

The logarithmic form of the probability information sent from the check nodes to the variable nodes, as follows:

$$\Lambda_{k\to i}(x_i) = \ln\left[\frac{r_{k\to i}(0)}{r_{k\to i}(1)}\right]. \tag{3}$$

In the following, the specific steps of the LLR-SPA are reviewed.

- **Initialization**

  $\forall i$, $k\in R$, an edge exists in the Tanner graph connecting nodes c_k and v_i, in Binary Symmetric Channel (BSC),

$$\Gamma_{i\to k}(x_i) = LLR(x_i) \tag{4}$$

$$\Lambda_{k\to i}(x_i) = 0 \tag{5}$$

In Eq. (4), $LLR(x_i)$ indicates the initialization probability information of a variable node, given received signal, *i.e.*

$$LLR(x_i) = \ln\left[\frac{P(x_i = 0 \mid y_i = y)}{P(x_i = 1 \mid y_i = y)}\right]. \tag{6}$$

According to the BSC channel and the Bayes theorem, we find

$$LLR(x_i \mid y_i = 0) = \ln\left(\frac{1-p}{p}\right) = \ln\left(\frac{n-t}{t}\right), \tag{7}$$

$$LLR(x_i \mid y_i = 1) = \ln\left(\frac{p}{1-p}\right) = \ln\left(\frac{t}{n-t}\right). \tag{8}$$

- **Check nodes update**

In the second step, the messages sent from the check nodes to the variable nodes are computed by the following formula:

$$\Lambda_{k\to i}(x_i) = 2\cdot \tanh^{-1}\left\{\prod_{j\in A(k)\backslash i} \tanh\left[\frac{1}{2}\Gamma_{j\to k}(x_i)\right]\right\}. \tag{9}$$

• Variable nodes update

The messages sent from the variable nodes to check nodes are computed by the following formulas:

$$\Gamma_{i\to k}(x_i) = LLR(x_i) + \sum_{j\in B(i)\backslash k} \Lambda_{j\to i}(x_i) \tag{10}$$

$$\Gamma_{i}(x_i) = LLR(x_i) + \sum_{j\in B(i)} \Lambda_{j\to i}(x_i) \tag{11}$$

Eq. (11) will be used in the last step.

• Decision

In the last step, the value calculated by Eq. (11) is used to obtain an estimated value $\hat{x} = \{\hat{x}_0, \ldots, \hat{x}_n\}$ of the reliable codeword component x_i, according to the following decision rule:

$$\hat{x}_i = \begin{cases} 0, & \text{if } \Gamma_i(x_i) \geq 0 \\ 1, & \text{if } \Gamma_i(x_i) < 0 \end{cases} \tag{12}$$

Syndrome of the estimated codeword $\hat{x}$ by the matrix H is computed. If $H\hat{x}^T = 0$, the decoding is successful and $\hat{x}$ is given as the result; otherwise, we need go back to the updated check notes. If the maximum number of iterations is reached and $H\hat{x}^T = 0$ is not satisfied, the decoding has failed and the algorithm stops.

4. A QC-LDPC BASED CFS DIGITAL SIGNATURE

4.1 Analysis of CFS signature algorithm

Although the security of the CFS algorithm based on Goppa code is high, the signature rate is particularly low because multiple decoding attempts are required at the signing process.

Given a (n, k) Goppa code, in which $n = 2^m$, $k = n - mt$, the number of decodable syndromes is:

$$N_d = \sum_{i=1}^{t}\binom{n}{t} \approx \binom{n}{t} \approx \frac{n^t}{t!}. \tag{13}$$

The total number of syndromes is:

$$N_t = 2^{n-k} = 2^{mt} = n^t. \tag{14}$$

The probability that a random syndrome can be decoded is:

$$p = \frac{N_d}{N_t} = \frac{1}{t!}. \tag{15}$$

Therefore, the probability of successful signature of the CFS algorithm is $1/t!$. When $t = 9$, we need try $9! = 362880$ times to get a signature in average [13]. However, it has been proved in [14] that this parameter is no longer safe, and it is suggested that the parameter can take $m = 15$, $t = 12$ or $m = 16$, $t = 10$. With the continuous updating of attack methods, the value of t will continue to increase, which makes the number of attempts to decode increase exponentially, the signature speed becomes lower and lower, and the signature efficiency of the CFS algorithm also becomes lower and lower.

The public key amount of the classical CFS signature algorithm is $(n-k) \times n$ bit, which requires a large key storage space, making the algorithm rarely used in practice. The sparseness and blocking cycle characteristics of the check matrix H of the QC-LDPC code can greatly reduce the key storage space and improve the defects of the classic CFS algorithm. Therefore, a CFS signature algorithm based on QC-LDPC code is presented in next section.

4.2 QC-LDPC based CFS Digital Signature Algorithm

The classical CFS signature algorithm uses the Goppa codes. A new signature scheme based on LDPC codes is proposed in this section, where the Goppa codes is replaced by the QC-LDPC codes and the BP decoding algorithm replaces the decoding algorithm of the Goppa codes in the CFS algorithm. We replace the permutation matrix P with the reversible transformation matrix Q so that it can resist the density reduction attack [15]. Because the length of the message digest s is $r (r < n)$ in the classical CFS algorithm, but the input sequence length of the BP decoding algorithm of the LDPC code is n, therefore, the following improvements have been made in this paper. The message digest s is converted into a sequence of length n. For the specific process, see Algorithm 1, in which the matrix H is the check matrix of the LDPC code.

Given an $r \times n$ matrix H and r-dimensional vector v, we can calculate $Hv^T = s$ (v is not limited by weight). If the weight of v is limited to t, it is a SD problem.

Algorithm 1:

Input: Message digest s of length r and parity check matrix H.

(a) Transforming the matrix H into row simplest H' using row transformation, that is, there is an invertible matrix M so that $M \cdot H = H'$ and $H = M^{-1} \cdot H'$;

(b) According to $H = M^{-1} \cdot H'$ and $H \cdot v^T = s$, we can get $M^{-1} \cdot H' \cdot v^T = s$. Then, both sides of the equation are left-multiplied by an invertible matrix M simultaneously. That is, $M \cdot M^{-1} \cdot H' \cdot v^T = M \cdot s$, $H' \cdot v^T = M \cdot s$;

(c) Since H', M and s are known, v can be satisfied;

Output: Sequence v of length n.

Algorithm 2 is the specific process of a QC-LDPC based CFS signature algorithm.

Algorithm 2:

(a) Initialization: Randomly select a (n, k) QC-LDPC code C with error correction capability t, described by its parity-check matrix H_1, an invertible matrix $S \in F_2$ of order $(n-k) \times (n-k)$, an $n \times n$ reversible transformation matrix $Q \in F_2$. Q is a diagonal block matrix, and its weight of row and column $w > 1$. Computing $H^{pub} = S \times H_1 \times Q$. Let α

represent the BP decoding algorithm of LDPC codes and β represent Algorithm 1, which is an algorithm for converting a message of length r into a sequence of length n. Select the public security hash function h: $\{0,1\}^* \rightarrow F_2^{n-k}$. The public key is H^{pub}, and the private key is $\langle S, H_1, Q, \alpha \rangle$.

(b) Signature: The message to be signed is m.

- The signer hashes m using the hash function h to get the message digest s: $s=h(m)$;
- Signer uses Hash function h to calculate $s_i = h(s|i)$, $i = 0, 1, 2, \ldots$;
- Convert s_i to a sequence v_i of length n by the β algorithm.
- Try to decode v_i using the algorithm α and find the smallest i that makes $\alpha(v_i)$ exist, this i is denoted as i_0. And s_{i_0} denotes s_i corresponding to i_0, z is the translated word, satisfies $Hz^T = s_{i_0}$, $w(z) = t$.

The signature of message m is $\sigma = [z|i_0]$ and the message-signature pair is (m, σ).

(c) Verification: Suppose the verifier receives the message-signature pair (m, σ).

- Using z and public key H^{pub} calculate $s_1 = H^{pub}z^T$;
- According to message digest $h(m)$ and i_0 calculate $s_2 = h[h(m)|i_0]$;
- If s_1 is equal to s_2, the signature σ is valid; otherwise, σ is invalid.

5. ANALYSIS OF SECURITY

5.1 Theoretical Analysis

In this scheme, the unidirectionality of the hash function h depends on the SD problem, and the hash value of the message m is the syndrome of the LDPC code. By attempting to decode, the relationship between the translatable code s_{i_0} and z can be interpreted as the syndrome and the error vector. From the public key H and s_{i_0}, directly solve the equation $Hz^T = s_{i_0}$, where $w(z) \le t$. According to theory of the error correction code, this is an NPC problem that can resist existing attacks of quantum algorithm such as the Shor algorithm and the Grover algorithm. Therefore, the proposed scheme can resist the existing attacks of quantum algorithms.

5.2 Stern's Attack

For the improved CFS algorithm, attacks can be performed against the low-density characteristics of the LDPC code. According to the Stern algorithm [16], the attacker obtains the message directly. According to the Stern algorithm of the (n, k) LDPC code proposed in [17], the probability that a codeword of weight w is found in one iteration by

$$\pi_w = \frac{\binom{w}{p}\binom{n-w}{k/2-p}\binom{w-p}{p}\binom{n-w-k/2+p}{k/2-p}}{\binom{n}{k/2}\binom{n-k/2}{k/2}} \times \frac{\binom{n-k-w+2p}{l}}{\binom{n-k}{l}}. \tag{16}$$

The p and l are two parameters for optimal performance, and their specific selection method is described in reference [16], so the average number of iterations required to find a low-weight codeword is π_w^{-1}.

To find the minimum distance efficiently, we can use Stern's algorithm instead of

the random codewords choice, and repeat the Stern's algorithm until the error probability is less than ε. The code words obtained for each iteration is independent. Therefore, after r iterations, the probability of failing to find the codeword of weight w is $(1-\pi_w)^r$.

The average number of bit operations is approximately the following form:

$$N \approx \frac{\frac{(n-k)^3}{2}+k(n-k)^2+2pl\binom{k/2}{p}+2p(n-k)\binom{k/2}{p}^2}{2^l}. \tag{17}$$

Generally, it is believed that the total work factor W satisfies $W \geq 2^{80}$, the scheme is safe. The total work factor W for finding a low weight code word is $W = \pi_w^{-1} \cdot N$. When we have the classical system parameters $n=4096$, $k=2048$, $d=82$, $p=3$ and $l=36$, the minimum work factor is $W = 2^{98.39}$, so the theme can resist Stern's attacks.

5.3 OTD Attacks

The attacker also exploits the sparseness of the check matrix. Both the matrix S and Q are composed of $p \times p$-order cyclic blocks and are sparse, which reduces decoding complexity. Let their generator polynomials be $s_{i,j}(x)$ and $q_{i,j}(x)$. Q be a diagonal form, satisfying $q_{i,j}(x) = 0 (i \neq j)$, can be expressed as

$$Q = \begin{bmatrix} Q_0 & & & \\ & Q_1 & & \\ & & \ddots & \\ & & & Q_{n_0-1} \end{bmatrix}. \tag{18}$$

According to the last $n-k$ columns of the parity check matrix H_1 and $H = S \times H_1 \times Q$, we can obtain

$$H_{n-k} = S^{-1} \times \begin{bmatrix} Q_0^{-1} & & & \\ & Q_1^{-1} & & \\ & & \ddots & \\ & & & Q_{n_0-1}^{-1} \end{bmatrix}. \tag{19}$$

The circulant block of H_{n-k}^1 at (i, j) is $Q_i S_{i,j}$, where $S_{i,j}$ is the circulant block at position (i, j) in S. Therefore, H_{n-k}^1 is still a circular matrix, its polynomial is expressed as $h_{i,j}(x) = q_i(x) \cdot s_{i,j}(x) \bmod (x^p - 1)$. Because of both S and Q are sparse, its row and column weight satisfying $w \ll n/n_0$. The highest power of $h_{i,j}(x)$ is w^2. The attacker can enumerate all the polynomial subsets of $h_{i,j}(x)$, and finally receive the secret key.

The descriptions of the OTD attacks are as follows:

(a) Let R_i represents the ith row of H_{n-k}^1, *i.e.*

$$R_i = [Q_i S_{i,j} | Q_i S_{i,j} | \ldots | Q_i S_{i,n-1}];$$

(b) The parity check of the linear code is expressed as

$$H_{OTD3} = (Q_i S_{i,k})^{-1} \cdot R_i = \left[I \mid S_{i,k}^{-1} S_{i,k+1} \mid \ldots \mid S_{i,k}^{-1} S_{i,n-1} \right];$$

(c) The parity check matrix H_{OTD3} converts to the form

$$H'_{OTD3} = S_{i,k} \cdot H_{OTD3} = \left[S_{i,k} \mid S_{i,k+1} \mid \ldots \mid S_{i,n-1} \right];$$

Due to S is sparse, the code contains low weight codewords. The row weight of H'_{OTD3} equals $w \cdot (n-k)$, that is very small compared to the code length. We use the Stern's algorithm to find the low-weight codeword, and then recover the $[S_{i,k}|S_{i,k+1}|\ldots|S_{i,n-1}]$. However, if the matrix S is dense, the workload of the attack will become very large and the attacker will have difficulty obtaining information about the secret key, thus the solution is safe.

6. ANALYSIS OF KEY COST

The public key cryptosystem has the disadvantages of large key storage and low information transmission rate, making it rarely used in practice. The key cost of the classical CFS algorithm based on Goppa code is $(n-k) \times n$ bit. For example, the typical parameter of the Goppa code [15] is code length $n = 1024$ and the error correction capability $t = 50$. When the parity matrix is in system form, the key cost approximately is 262×10^3 bits. For $n = 2048$ and $t = 92$, the key cost approximately is 1048×10^3 bits. Although the security of classical CFS algorithm is higher, the key storage space is larger, and the efficiency of the algorithm is also reduced.

The CFS signature scheme based on QC-LDPC codes proposed in this paper exploits the sparseness of the parity check matrix. It can greatly reduce the key storage space, only need to store the number of non-zero elements in the H matrix and the location of each non-zero element. The cyclic characteristic of the QC-LDPC code parity check matrix not only reduces the amount of key storage, but also makes the information bits of the codeword larger, error correction capability enhanced, and the information transmission rate increased. The public key is a quasi-cyclic matrix in this scheme. It only needs to store the first row of each cyclic block.

For (n, k) LDPC codes with the parity matrix in system form, the size of public key is $(n-k) \times (np-1)\log_2 k$, in which p represent the average probability of nonzero entries in each row, while the size of public key is $k_0 \times n_0 \times q$ when using QC-LDPC codes.

For the typical parameters of QC-LDPC codes: $n_0 = 4$, $k_0 = 3$, $q = 4096$, $n = n_0 \times q = 16384$ and $k = k_0 \times q = 12288$, the size of public key is $k_0 \times n_0 \times q = 49152$ bit = 6144 byte, while this number is 1133794 byte in LDPC code with $p = 0.01$.

Obviously, the QC-LDPC code reduces the key storage space of the classical CFS signature algorithm very well. The performance of CFS signatures algorithm based on different error correction codes is shown in Table 1.

From Table 1, we observe that using the QC-LDPC code greatly reduce the key size of the CFS algorithm. Due to the structural characteristics of the QC-LDPC code, the information bits of the codeword are larger, which effectively improves the transmission rate.

Table 1. Data analysis.

Codes	Parameters	Key size/Bytes	Information rate
Goppa	(1024,524)	32750	0.51
Goppa	(2048,1036)	131054	0.51
Goppa	(4096,2056)	524280	0.50
LDPC	(4096, 2048)	112640	0.50
QC-LDPC	(4096, 2048)	1024	0.50
Goppa	(8192,4097)	2097151	0.50
LDPC	(8192, 4096)	497664	0.50
QC-LDPC	(8192, 4096)	2048	0.50
Goppa	(16384,12296)	6283256	0.75
LDPC	(16384,12288)	1133794	0.75
QC-LDPC	(16384,12288)	6144	0.75

7. CONCLUSIONS

The classical CFS algorithm is one of the important digital signature algorithms based on error correction codes. Due to its high security, it has been widely concerned and studied since it was proposed. However, this algorithm has the disadvantages of large key size and the rapid drop in the signature with the increase of Goppa code error correction capability *t*. The QC-LDPC code-based CFS signature algorithm is proposed in this paper. This algorithm takes advantage of the sparseness and circularity of the QC-LDPC code parity check matrix, which effectively improves the shortcomings of the classic CFS signature algorithm. The LLR-BP decoding algorithm improves the signature efficiency without reducing the security. This algorithm, which still falls into the code based digital signature algorithm, can also resist the existing quantum algorithm attacks.

ACKNOWLEDGMENT

This paper was supported by the National Nature Science Foundation of China (Program Nos. 61472472, 41504115, 61772418) and Natural Science Basic Research Plan in Shaanxi Province of China (Program Nos. 2015JQ6262, 2017JQ6010, 2018JZ6001).

REFERENCES

1. P. W. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in *Proceedings of IEEE 35th Annual Symposium on Foundations of Computer Science*, 1994, pp. 124-134.
2. L. K. Grover, "A fast quantum mechanical algorithm for database search," in *Proceedings of the 28th Annual ACM Symposium on Theory of Computing*, 1996, pp. 212-219.
3. D. J. Bernstein, "Introduction to post-quantum cryptography," *Post-Quantum Cryptography*, Springer, Berlin, Heidelberg, 2009, pp. 1-14.

4. N. Courtois, M. Finiasz, and N. Sendrier, "How to achieve a McEliece-based digital signature scheme," in *Proceedings of International Conference on Theory and Application of Cryptology and Information Security*, 2001, pp. 157-174.
5. H. Niederreiter, "Knapsack-type cryptosystems and algebraic coding theory," *Problem Control and Information Theory*, Vol. 15, 1986, pp. 159-166.
6. R. G. Gallager. "Low-density parity-check codes," *IRE Transactions on Information Theory*, Vol. 8, 1960, pp. 21-28.
7. D. J. C. Mackay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *Electronics Letters*, Vol. 32, 1996, pp. 1645-1646.
8. M. Baldi, F. Chiaraluce, R. Garello, *et al.*, "Quasi-cyclic low-density parity-check codes in the McEliece cryptosystem," in *Proceedings of IEEE International Conference on Communications*, 2007, pp. 951-956.
9. O. Blazy, P. Gaborit, J. Schrek, *et al.*, "A code-based blind signature," in *Proceedings of IEEE International Symposium on Information Theory*, 2017, pp. 2718-2722.
10. S. Chen, P. Zeng, K. K. R. Choo, *et al.*, "Efficient ring signature and group signature schemes based on *q*-ary identification protocols," *Computer Journal*, Vol. 61, 2018, pp. 545-560.
11. Q. Alamélou, O. Blazy, S. Cauchie, *et al.*, "A code-based group signature scheme," *Designs*, *Codes and Cryptography*, Vol. 82, 2017, pp. 469-493.
12. M. Baldi, *QC-LDPC Code-Based Cryptosystems*, Springer, Berlin, Heidelberg, 2014.
13. F. Ren, D. Zheng, and W. J. Wang, "An efficient code based digital signature algorithm," *International Journal of Network Security*, Vol. 19, 2017, pp. 1072-1079.
14. M. Finiasz and N. Sendrier, "Security bounds for the design of code-based cryptosystems," in *Proceedings of International Conference on the Theory and Application of Cryptology and Information Security*, 2009. pp. 88-105.
15. L. X. Yang, "Research on Niederreiter public key cryptosystem based on error correcting codes," School of Computer and Communication Engineering, Changsha University of Science and Technology, 2011.
16. J. Stern, "A method for finding codewords of small weight," in *Proceedings of International Colloquium on Coding Theory and Applications*, 1988. pp. 106-113.
17. M. Hirotomo, M. Mohri, and M. Morii, "A probabilistic computation method for the weight distribution of low-density parity-check codes," in *Proceedings of IEEE International Symposium on Information Theory*, 2005, pp. 2166-2170.

Fang Ren (任方) received his Ph.D. degree in Cryptography from Xidian University in 2012. Now he is an Associate Professor of Xi'an University of Posts and Telecommunications. His research interests include information security, code based cryptography, and digital image watermark.

Xue-Fei Yang (杨雪菲) is currently pursuing the Master's degree in Information Security from Xi'an University of Posts and Communications, Shannxi. Her research interests include error correcting code, digital signature and algorithm analysis.

Dong Zheng (郑东) received his Ph.D. degree from Xidian University in 1999. Now he is a Professor of National Engineering Laboratory for Wireless Security, Xi'an University of Posts and Telecommunications. His research interests include cloud security, code based systems and other new cryptographic technology.