NEOMA
Publishing House

On Novel Public-key Cryptosystem Using MDS Code

Karla Zayood

[2]Online Islamic University, Department Of Science and Information Technology, Doha, Qatar

zayyyoood134@gmail.com

ABSTRACT

In this paper we will build a public-key cryptosystem using the MDS code. The advantages of this cryptosystem are that the size of the plaintext is the same as the size of ciphertext. We started with *k*-bit message *m* and ended with a *k*-bit ciphertext C. This may be considered economical and time saving, if compared by other systems, the length of the ciphertext is about 100 times longer than the length of the message itself.

1. INTRODUCTION

There are various techniques for public-key encryption based on a hard mathematical problems [3, 4], one of these techniques is the McEliece public-key encryption which based on error correcting codes [5]. In this paper we will build public-key encryption using MDS codes and this public-key is preferable than McEliece public-key encryption that the size of the plaintext is the same that of ciphertext.

This paper is organized as follows. The first section surveys the literature for the use of MDS codes in cryptosystems. The second and the third sections expose the elementary definitions and results of both coding theory and cryptography respectively in order to set up the notation being used through out this paper. The fourth section explains the main result of how to build up the cryptosystem. The fifth section studies the security and complexity level of the cryptosystem. The sixth and last section is the conclusion.

2. *Basic Definitions in Coding Theory* [1]

A *word* is a sequence of digits. *The length* of a word is the number of digits in the word.

Let $\beta = \{\alpha_1, \dots, \alpha_q\}$ be a finite set called the alphabet, and let β^n be the set of all strings of length *n* over β. Any nonempty subset $\boldsymbol{C}$ of β^n is called *q-ary block code*. Each string in $\boldsymbol{C}$ is called a

Submitted: 13/1/2023, accepted: 12/2/2023

codeword. If $\boldsymbol{C} \subset \beta^n$ contains *M* codewords then $\boldsymbol{C}$ has length *n* and size *M* and is denoted by (*n*, *M*)-code. A code whose β = {0, 1} is called *binary code*. Commonly, alphabets have a group, a field or even a ring structure.

If v is a word in $\boldsymbol{C}$ that is sent and w in β^n is received, then $u = v + w$ is the *error pattern*, or the error.

Let a be a q-ary word of length *n* over an alphabet that has a group structure (that contains zero). The Hamming weight, or simply the weight of a is the number of non-zero components in a. We denote the weight of a by *w*(a). The minimum weight of a code $\boldsymbol{C}$ is the minimum weight of all nonzero code words in $\boldsymbol{C}$ and denoted by *w*($\boldsymbol{C}$).

Let x and y be two words of the same length. The Hamming distance or simply the distance between *x* and *y* is the number of positions in which *x* and *y* differ.

We denote this distance between *x* and *y* by *d*(*x*, *y*). A code $\boldsymbol{C}$ is said to have minimum distance *d* if $d = min\{d(x,y): x, y \epsilon\, \boldsymbol{C}, x \neq y\}$ and is denoted by *d*($\boldsymbol{C}$).

An (*n*, *M*, *d*)-code is a code of length *n* and size *M* and minimum distance *d*.

A code $\boldsymbol{C}$ is called a linear code if $\boldsymbol{C}$ is a vector subspace of a vector space V(*n*, *q*) of dimension *n* over the field *GF(q)*. If $\boldsymbol{C}$ has dimension *k* over GF(*q*), we say that $\boldsymbol{C}$ is an [*n*, *k*]-code, and if $\boldsymbol{C}$ has minimum distance *d* we say that $\boldsymbol{C}$ is an [*n*, *k*, *d*]-code.

We say that a code $\boldsymbol{C}$ corrects an error pattern *u* if, for all *v* in $\boldsymbol{C}$; *u* + *v* is closer to *v* than to any other word in $\boldsymbol{C}$.

A code $\boldsymbol{C}$ is called a *t*-error-correcting code if $\boldsymbol{C}$ corrects all error patterns of weights at most *t* and does not correct at least one error pattern of weight *t*+1.

It can be easily verified that a linear [*n*, *k*, *d*] code $\boldsymbol{C}$ corrects all error patterns *e* of *t* where $t = \left\lfloor \frac{d-1}{2} \right\rfloor$. We denote such a code by (*n*, *k*)-*t-error correcting code.*

Let $\boldsymbol{C}$ be a linear [*n*, *k*]-code, a *k*×*n* matrix G whose rows form a basis for $\boldsymbol{C}$ is called a *generator matrix* for $\boldsymbol{C}$. If $\boldsymbol{C}$ is an [*n*, *k*]-code with generator matrix G, then the codewords in $\boldsymbol{C}$ are the linear combinations of the rows of G, i.e.,$\boldsymbol{C} = \{x\, G: x \in V(k,q)\}$.

3. *Basic Definitions and Notations in Cryptography* [2]

A plain message is called *a plaintext*. The process of disguising a message in such a way as to hide its substance is called *encryption*. An encrypted message is called *ciphertext*. The process of turning ciphertext back into plaintext is called *decryption*. Plaintext is denoted by P. Ciphertext is denoted by C. The encryption function E operates on P to produce C. This mean that E(P) = C. In the reverse process, the decryption function D operates on C to produce P; D(C) = P. Since the whole point of encryption and then decryption of a message is to recover the original plaintext, the following identity must hold: D(E(P)) = P.

Submitted: 13/1/2023, accepted: 12/2/2023

A key, denoted by *k*, is any thing that we keep secret such that without knowing it the decryption is infeasible process. *A cryptosystem* is the system consists of plaintext, ciphertext and the keys (encryption and decryption).

A symmetric key cryptosystem is a cryptosystem in which the decryption key can be calculated from the encryption key and vice versa.

In most symmetric cryptosystems the encryption and decryption keys are the same.

Public key cryptosystem is a system in which the encryption key (*called public key*) is different from the decryption key and it is made public. Decryption key can not be calculated from the encryption key.

The decryption key is often called the *private key*. A protocol is a set of instructions telling the sender and receiver what to do.

4. *The Main Results*

Let $\boldsymbol{C}$ be an MSD code of length n and dimension k whose generator matrix is G of orders $k \times n$. The idea is to create a public-key and corresponding private key as follow:

Each entity should perform the following steps:

1. Choose a $k \times n$ generator matrix G for (n, k)-linear MDS code.
2. Select a random $k \times k$ binary non-singular matrix S.
3. Select a random $n \times n$ permutation matrix P.
4. Compute the $k \times n$ matrix $\widehat{\mathrm{G}} = \mathrm{SGP}$.
5. *A*'s public-key is $(\widehat{\mathrm{G}}\ , r)$; *A*'s private key is (S, G, P), where *r* is the number of repetition of G.

Cipher design

A encrypts a massage *m* for B as follow:

1. Obtain B's public-key $(\widehat{\mathrm{G}}\ , \mathrm{r})$.
2. Represent a plaintext as a string m of length k.
3. Compute the following:

$$C_1 = m\widehat{\mathrm{G}}\,T,$$

$$C_i = C_{i-1}\widehat{\mathrm{G}}\,T,\ i=1,2,\dots,r,$$

where T is the matrix of the form

$$T = \begin{bmatrix} I_{k\times k} \\ 0_{n-k\times k} \end{bmatrix}$$

1. Take $C = C_r$

B decrypts the ciphertext C to *m* as follow:

1. Compute the missing components of $C_r = C$ by using MDS code to get $\hat{C}_r$.
2. Compute $C_r = \hat{C}_r\ \mathrm{p}^{-1}$, where $C_r = (\hat{C}_{r-1}\mathrm{S})\mathrm{G}$.
3. Solve the system of equation $C_r = (\hat{C}_{r-1}\mathrm{S})\mathrm{G}$ to get $C_{r-1} = \hat{C}_{r-1}\mathrm{S}$.

Submitted: 13/1/2023, accepted: 12/2/2023

4. Multiply $C_{r-1}S$ by $S^{-1}$to get C_{r-1}.
5. Repeat the last 4-steps (r –1) times until we get the system of equation $C_1 = (mS)G$, by solving it we get mS which we multiply by S^{-1} to get m (which the plain text).

Example.

Let $\boldsymbol{C}$ be a Reed-Solomon code (RS(2, 5)) whose generator matrix G is

$$G = \begin{pmatrix} 0 & 2 & 4 & 1 & 3 \\ 1 & 3 & 0 & 2 & 4 \end{pmatrix}$$

Let S be a non-singular matrix of order 2 × 2.

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

Let P be permutation matrix of order 5 × 5.

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

The public-key is ($\hat{G}$= SGP, r) where r is the number of repetition of a matrix $\hat{G}$, the private key is (S, G, P), take r = 2.

For instance, to encrypt the message m = (12). Let T be the fixed matrix of the form

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

1. Multiply $m\widehat{\mathrm{G}}$ to get $\hat{\mathrm{C}}_1 = (2\ 2\ 2\ 2\ 2)$.

2. Multiply $\hat{\mathrm{C}}_1\mathrm{T}$ to get $\mathrm{C}_1 = (2\ 2)$.

3. Multiply $\mathrm{C}_1\widehat{\mathrm{G}}$ to get $\hat{\mathrm{C}}_2 = (4\ 2\ 1\ 3\ 0)$.

4. Multiply $\hat{\mathrm{C}}_2\mathrm{T}$ to get $\mathrm{C}_2 = (4\ 2)$.

For decryption we reverse the process of encryption

1. Compute the missing components of C_2. Using MDS code to get $\hat{\mathrm{C}}_2 = (4\ 2\ 1\ 3\ 0)$.
2. Compute $(4\ 2\ 1\ 3\ 0)\mathrm{P}^{-1} = (2\ 4\ 1\ 3\ 0)$.
3. Solve the system of equations $(2\ 4\ 1\ 3\ 0) = (\mathrm{C}_1\mathrm{S})\mathrm{G}$ we get $\mathrm{C}_1\mathrm{S} = (4\ 2)$.
4. $\mathrm{C}_1 = (4\ 2)\mathrm{S}^{-1} = (4\ 2)\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = (2\ 2)$.
5. Complete the missing component of C_1 using MDS code to get $\hat{\mathrm{C}}_1 = (2\ 2\ 2\ 2\ 2)$
6. Complete $(2\ 2\ 2\ 2\ 2)\mathrm{P}^{-1} = (2\ 2\ 2\ 2\ 2)$
7. Solve the system of equation $(2\ 2\ 2\ 2\ 2) = (m\mathrm{S})\mathrm{G}$ to get $m\mathrm{S} = (3\ 2)$.
8. $m = (3\ 2)S^{-1} = (1\ 2)$ which is the plaintext.

Submitted: 13/1/2023, accepted: 12/2/2023

5. Security and Complexity of the Cryptosystem

Because any k-components of a codeword in an (n, k) MDS code form an information set, i.e., the value of these k-components can be arbitrary selected and they determine exactly one codeword. The number of selections is the number of choices $\binom{n}{k}$, so in this cryptosystem we must use an MDS code of large length n.

The security of this cryptosystem is the secrecy of G and the number of repetitions n. We know that the number of generating matrices G of a given MDS code is $\frac{1}{K!}\prod_{i=0}^{k-1}(q^k - q^i))$, where q is the order of the field.

As far as attacks on this system, we should say that for a plaintext ciphertext attack, we need to solve $\binom{n}{k} \times k$ nonlinear equations of degree n in $k \times n$ unknowns to get solutions of the system, which is considered as hard problem (See RSA cryptosystem).

6. *Conclusion*

We have built up a cryptosystem based on MDS codes. The public-key of this system is the pair $(\widehat{\mathrm{G}} = \mathrm{SGP}, r)$, where G is the generator matrix of MDS code and r is the number of repetition of G, and the private key of this system is (S, G, P). The advantages of this cryptosystem are:

1. The size of the plaintext is the same that of ciphertext.
2. It has high security level.
3. This cryptosystem withstands by the plaintext- ciphertext attacks.

References

[1] Roman, S., Coding and Information Theory, Springer-Verlag, (1992).

[2] Menezes, A. and Vanstone, S., Handbook of applied Cryptography, CRC Press (1996).

[3] Chabaud, F., "*On the Security of some Cryptosystems based on errorcorrecting codes*", Advances in Cryptography, pp.131-139, (1995).

[4] Massey, J.L. "*Some applications of coding theory in Cryptography Codes and Cyphers*", Cryptography and Coding IV. Essex. England:
Formare Ltd. pp. 33-47 (1995).

[5] McEliece, R.J. "*A public key Cryptosystem based on algebraic coding theory"* DSN Progress report 42-44 pp. 114-116. (1978).

Submitted: 13/1/2023, accepted: 12/2/2023