ISSN No:-2456-2165

McEliece's Crypto System based on the Hamming Cyclic Codes

KABEYA TSHISEBA Cedric
Chef of practical Works, Department of Mathématics and Computer Science
Faculty of Sciences, Pédagogic National University (UPN)
KINSHASA, DRC

Abstract:- The rapid development of global networks and the tremendous possibilities offered by electronic transactions in continuous communication today pose a critical challenge to the protection of information against transmission errors in a network. On the other hand, this data must be unintelligible except to the intended audience. In order to overcome these two constraints, the coding of information is used to combat transmission errors and data encryption is often used to combat any spy system. In this paper, we try to introduce a crypto public key system using error correcting codes (it's a two-in-one system). The system studied is the McEliece crypto system using the Hamming cycle code correcting a simple error.

Keywords:- *Mceliece Crypto System, Hamming Code, Cryptography, Encryption, Public Key.*

I. INTRODUCTION

The major problem of telecommunication devices in addition to the increase of the transmission rate which can be regulated by the compression methods, is on the one hand, the errors introduced by the transmission mediums ie. the information received is wrong. It is therefore necessary to protect this information. And on the other hand, the growing need to secure data in the IT and telecommunications fields. Especially now with the advent of networks, the use of satellite links and the use of the Internet the situation has changed dramatically, as the same message passes through several machines before reaching the recipient. At each stage, it can be copied, lost or altered. Encryption is therefore necessary for the data to be unintelligible except to the intended audience.

In 1976, with the invention of the first public key cryptosystem by Diffie and Hellman [8]. The new idea was to base the security of a system not on the knowledge of a key (shared secretly by users), but on the difficulty of reversing a one-way function with trapdoor.

A one-way function is simply a computationally difficult function to reverse.

A trapdoor is a secret algorithm making this reversal easy. Thus the hatch is only known to a person, only to be able to decipher the messages created by using the one-way function which is public.

As early as 1978, McEliece [6] devised the first and most famous public-key cryptosystem using error-correcting codes. As we will see in this paper, code theory also contains many well-structured and difficult-to-solve problems, more or less well adapted for use in cryptography.

In this work we will try to introduce a crypto public key system using error correcting codes. The studied system is the McEliece crypto system using the Hamming cycle code.

❖ *Generator / Control Matrices And Random Codes*

The construction of a code word comprising n bits is performed from k bits of the source message k-uplet binary $U=(u_1,u_2,u_3,.....,u_k)$, called generally information message, and r bits of redundancy. The simplest coding method is to leave the k information bits unchanged and to transfer them as they are in the code word by adding the r(= n-k) redundancy bits $\{a_1,a_2,...,a_r\}$, which are usually called control bits, the V^T line vector called codeword :

$V^T=[v_1v_2...v_n]=[u_1,u_2,u_3,.....u_ka_1,a_2,...,a_r]$.

- When the control bits are calculated only from the information bits of the block to which they belong, the code C (n, k) is called the block code.
- When the control bits are calculated from information bits belonging to several blocks, the code is called convolutional or recurrent.

The hamming code that is studied in this communication is of cyclic linear type:

1. The linear codes have the property that the set of code words form a vector space.
2. Cyclic codes have the property that any circular permutation of a code word is a code word.

The Hamming code is very used in practice to protect short information (16, 32 or 64 bits), it is very used for operations on computer memories but little used in transmission, it also has the following advantages:

- Easier to implement in hardwired logic.
- Lends itself well to a length extension of the information to be coded.

A. *Generator and Control Matrices*

As soon as it is possible to determine two matrices M_1 and M_2 having columns such as:

$[M_1][M_2]^T = 0$ (1)

We have defined a linear code with n positions [3]. Indeed, if the matrix M_1, has for dimension (mxn) and the matrix M_2 (kxn), one obtains all the words of the code by premutipliant M_1 by all the m-tuples X_i:

$<X_i>[M_1]=<C_i>$ i = 0 to 2^m-1

But, as we can always put the matrix M_1 under its canonical form in step:

[G] = [Im A] We have

$<X_i>$ [G] = <m information position, n-m control position>

But, for equality(1) to always be satisfied, it will be noted that M_2, put in its canonical form in step, must be equal to $[-A^T I_k]$.

We will designate this matrix by [H] and check that:

$[G][H]^T=0$ with k+m=n

The matrix G is the generator matrix of the code (n, m), the matrix H is its matrix of control. Now, since $[G][H]^T=0$ implies $[H][G]^T=0$, any code (n,m) corresponds to a code (n,n-m) called dual first, the role of the matrices being reversed.

The control matrix H plays a crucial role in detecting or correcting errors. Indeed:

If $<X_i>[G]=<C_i>$, then obligatorily [H](Ci)=0

This matrix equation represents a system of k equations with k unknowns which are the values to be assigned to the k control positions when the values information positions are known. All the n-tuples Ci checking this system are code words. The k equations represented by $[H]<C_i>=0$ define the control relationships, they make it possible to detect and correct the errors.

B. *Random Codes*

To obtain a random code it is enough to draw a random generator matrix and to look for its image. Of course, once chosen, the code is no longer random, but in general, a code built in this way will have good properties on average: it usually has a good minimum distance. Unfortunately for such a code there is no polynomial decoding algorithm. This last category is not really a family of codes since it is in fact all the codes built without particular structures.

❖ *Hamming Code*

A Hamming code is a cyclic code C(n,k) generated by a primitive polynomial g(x) of degree m ≥ 3 [1,2,4]. With the characteristics:

1. Length of the code word $n=2^m-1$.
2. Number of control bits m=n-k.
3. Number of information bits $k=2^m-m-1$.
4. $d_{min}=3$, the code corrects t errors (a simple error) $t=[(d_{min}-1)/2]=1$.

A. *Coding principle*

Let $U=(U_0,U_1,...,U_{k-1})$ the k-tuplet to be encoded, to which the polynomial U(x) is associated, The coding consists of:

a) Pre multiply the polynomial U(x) associated with the k-tuplet to be encoded by x^{n-k}.

b) Obtain the remainder D(x) of the division of $x^{n-k} *U(x)$ by the generating polynomial g(x) .

c) Add D(x) and $x^{n-k} * U(x)$ to obtain the word of code V (x) = D (x) + $x^{n-k} * U$ (x).

B. *Decoding Principle*

The decoding process is divided into two steps:

- Detection of errors in the received word.
- Correction of these errors in case of need.

The received word S(x) syndrome R(x) is defined by:

$S(x)=rest(x^{n-k} * R(x)/g(x))$ (2)

With $R^{(i)}(x)$: the polynomial obtained after the ith right shift R(x). The syndrome corresponding to ith cyclic shift of R(x) can be calculated by:

$S^{(i)}(x)\ rest(X^{n-k}*R^{(i)}(x)/g(x))$ (3)

$S^{(i)}(x)$ can be calculated by [7]:

$S^{(i)}(x)=rest(xS^{(i-1)}(x)/g(x))$ if $i>0$ (4)

$S^{(i)}(x)= S(x)$ if $i = 1$

In (3) we show that the $S^{(i)}(x)$ can be calculated from S(x) using the recurrence equation and therefore we are not obliged to make offsets on R(x) for the computed. The formation of the syndrome is:

$S(x)=S^{(0)}+S^{(1)}x+S^{(2)}x^2+S^{(n-k-1)}x^{n-k-1}$ (5)

The detection is done according to the value of S(x) , we can say if there are errors or not, for that we have two cases:

- S(x) = 0, R(x) is a multiple of g(x) so the word received is considered as the word
- S(x) <> 0, R(x) is not a code word

V(x)=R(x)+E(x), we will correct the errors.

S(x) depends only on the error pattern E introduced and not V. Indeed we have:

$S(x) = rest(x^{n-k} *R(x)/g(x)) = rest(x^{n-k} *E(x)/g(x))$ (6)

❖ *Crypto McEliece System*

This is the oldest crypto public key system using error correcting codes. It was imagined by McEliece [6] in 1978, roughly at the same time as RSA [7]. Like all public key crypto systems, this system consists of 3 algorithms:

1. key generation,
2. encryption (using the public key) and
3. decryption (using the secret key).

McEliece suggested using Goppa codes, which are linear codes with a fast decoding algorithm. We propose to do this with the Hamming cyclic code.

A. Generation of key

We start by generating a Hamming code and its parity matrix G of size k×n. We will mix this matrix to make it indistinguishable from a random matrix, for that we need:

1. A random permutation matrix P of size n×n having a 1 in each row and column and 0 everywhere and,
2. A random invertible matrix S of size k×k (S is a scrambler matrix).

The public key will be the matrix G'=S×G×P which is indistinguishable from a random matrix. The security of this system is based on the problem of distinguishing the permuted Hamming code from a random code. The secret key is composed of the three matrices S, P and G which make it possible to find the structure of the Hamming code and give access to the decoding algorithm.

B. Encryption

Let m be a message of k bits that we want to encrypt. For this purpose we only have the public key G'. We start by calculating the code word C of length n associated with m:

$$C=m\times G' \qquad (7)$$

Then we generate a random error e of length n. The cipher will be simply the noisy code word:

$$C'=c+e \ (8)$$

C. Decryption

To decipher by knowing P, S and G it suffices to calculate:

$$C'\times P^{-1} = mG'P^{-1}+eP^{-1}=mS\times G+eP^{-1} \qquad (9)$$

MS × G is a word of the Hamming code and eP-1 is a weight error t (because P is a permutation and therefore keeps the weight of the words), so we can decode this error and find the initial message m S. all that remains is to multiply by S^{-1} to find the message m and to have finished deciphering.

D. Example of Encryption

Let be the Hamming code (7, 4) with G generator matrix that corrects all the simple errors. The scrambler matrix S and a permutation matrix P are chosen. The public key corresponding to the matrix G' is calculated.

$$G=\begin{vmatrix} 1&0&0&0&1&1&0\\ 0&1&0&0&1&0&1\\ 0&0&1&0&0&1&1\\ 0&0&0&1&1&1&1 \end{vmatrix} \qquad P=\begin{vmatrix} 0&1&0&0&0&0&0\\ 0&0&0&1&0&0&0\\ 0&0&0&0&0&0&1\\ 1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0\\ 0&0&0&0&0&1&0\\ 0&0&0&0&1&0&0 \end{vmatrix}$$

$$S=\begin{vmatrix} 1&1&0&1\\ 1&0&0&1\\ 0&1&1&1\\ 1&1&1&0 \end{vmatrix} \qquad G'=S\,G\,P=\begin{vmatrix} 1&1&1&1&0&0&0\\ 1&1&0&0&1&0&0\\ 1&0&0&1&1&0&1\\ 0&1&0&1&1&1&0 \end{vmatrix}$$

Let X=(1101) be the message to send. If we suppose that the transmission channel introduces a simple error of weight 1 of value e=(00000100). Instead of sending the message X it is another message Y that is sent.

$$y=xG'+e=(0110010)+(0000100)=(0110110) \ (10)$$

On receipt of Y’ we calculate first: y’= yP^{-1}

$$P^{-1}=\begin{vmatrix} 0&0&0&1&0&0&0\\ 1&0&0&0&0&0&0\\ 0&0&0&0&1&0&0\\ 0&1&0&0&0&0&0\\ 0&0&0&0&0&0&1\\ 0&0&0&0&0&1&0\\ 0&0&1&0&0&0&0 \end{vmatrix} \qquad S^{-1}=\begin{vmatrix} 1&1&0&1\\ 1&1&0&0\\ 0&1&1&1\\ 1&0&0&1 \end{vmatrix}$$

$$yP\text{-}1=(xG'+e)P^{-1}=xSG+eP^{-1}=xSG+e' \qquad (11)$$

e’ is a vector of the weight t (since P^{-1} is also a permutation matrix). We obtain: y’=(1000111) We apply the Meggit decoding to determine the error vector e’ and consequently the code word (xS)G. The syndrome found is $(1110)^T$, so the error occurs in position 7 (details omitted). The receptor now has the code word y=(1000110).

The vector xS=(1000) can now be obtained by multiplying by G^{-1} on the right side (however, we can write G in the standard format $[I_k\ A]$, and then the xS would be just the first positions of k of the xSG and this multiplication would not be necessary). In conclusion, we obtain x=(1000) S^{-1}=(1101) by multiplying xS on the right side by S^{-1}.

II. CONCLUSION

We have introduced in this paper a new application of the cryptography error-correcting code theory. This application is very concrete, but it certainly requires some small adjustments before being really usable in practice. Because we have identified three disadvantages for this McEliece crypto system.

1. The size of the public key (G') is large. This will certainly pose problems of execution.
2. The encrypted message is longer than the clear message. This increase in the width of the encrypted message makes the system more susceptible to transmission errors.
3. The crypto system is not used for signing or authentication because the encryption algorithm is not linear and the whole algorithm is really asymmetric.

The security of a system is evaluated by the cost of the best attacks, but the effort to try and find better ones is also an important part of that security. From this point of view, an old system well known to all is often preferable to a younger system, even if it offers good properties.

REFERENCES

[1]. Clark GC, JB Cain, " Error Correcting Coding for Digital Communication ", Plenum Press 1981.

[2]. DJ Costello, S.Lin, " Error Control Coding: Fundamentals and Applications ", Prentice Hall 1983.

[3]. AL. Spataru, 'Transmission of Information II: Codes and Static Decisions', MASSON and CIE, 1973.

[4]. A. Poli, Li Huguet, " Correcting Codes: Theory and Applications ", Masson Paris 1989.

[5]. S .Foughali, S.Khelifa, " Concatenation of Cyclic Codes (Reed Solomon - Hamming) Applied to still images ", Institute of Computer Science, USTO 1998.

[6]. RJ McEliece. A public key cryptosystem based on algebraic coding theory. DSN Prog. Rep., Jet Prop. Lab., California Inst. Technol., Pasadena, CA, pp. 114-116, January 1978.

[7]. R. L. Rivest, A. Shamir, and L. M. Adleman. A method for obtaining digital signatures and public-keycryptosystems. Communications of the ACM, 21 (2): 120-126, February 1978.

[8]. W. Diffie and M. Hellman. New directions in cryptography. IEEE Transactions on Information Theory, 22 (6): 644-654, November 1976.