McEliece 공개키 암호체계의 암호해독을 위한 Polynomial-Time 알고리즘

박 창 섭
단국대학교 전자계산학과

A Polynomial-Time Algorithm for Breaking the McEliece's Public-Key Cryptosystem

Chang-Seop Park
Department of Computer Science
Dankook University

Abstract

McEliece 공개키 암호체계에 대한 새로운 암호해독적 공격이 제시되어진다. 기존의 암호해독 algorithm이 exponential-time의 complexity를 가지는 반면, 본고에서 제시되어지는 algorithm은 polynomial-time의 complexity를 가진다. 모든 linear codes에는 systematic generator matrix가 존재한다는 사실이 본 연구의 동기가 된다. Public generator matrix로 부터, 암호해독에 사용되어질 수 있는 새로운 trapdoor generator matrix가 Gauss-Jordan Elimination의 역할을 하는 일련의 transformation matrix multiplication을 통해 도출되어진다. 제시되어지는 algorithm의 계산상의 complexity는 주로 systematic trapdoor generator matrix를 도출하기 위해 사용되는 binary matrix multiplication에 기인한다. Systematic generator matrix로 부터 쉽게 도출되어지는 parity-check matrix를 통해서 인위적 오류의 수정을 위한 Decoding이 이루어진다.

본 연구는 한국 전자 통신 연구소의 출연금에 의해 이루어졌습니다.

I. Introduction.

Based on the idea that the general decoding problem for a linear code is NP-complete[1], McEliece[2] has introduced a public-key cryptosystem using the error-correcting codes, especially a binary irreducible Goppa code. His main idea is to transform the original generator matrix of a Goppa code into the one from which the corresponding parity-check matrix for decoding can't be obtained without knowing the matrices used for the transformation. Several cryptanalytic attacks against the McEliece's public-key cryptosystem have been proposed. Adams and Meijer[3] have calculated the optimum value for the number of errors to be introduced in the codeword for the purpose of maximizing the work factor, and Lee and Brickell[4] have lowered the work factor by a systematic method of checking the validity of the cryptanalyzed message.

Their ideas are based on repeatedly selecting the random bits from the ciphertext in the hope that none of the selected bits are affected by errors. The approach taken by them results in a high work factor due to the exhastive search for the random bits which are not in error.

In this paper, a polynomial-time algorithm for breaking the McEliece's system is presented, which is based on the existence of a new trapdoor generator matrix obtained from the public generator matrix. In chapter two, a description of the McEliece's cryptosystem is given, together with the conventional cryptanalytic attacks against it. A new cryptanalytic algorithm for breaking it, which runs in polynomial time, is presented in chapter three. Finally, conclusions are drawn.

II. The McEliece's Public-Key Cryptosystem.

II.1 System Description.

Given an irreducible polynomial of degree t over GF(2^m), the user generates a t-error-correcting Goppa code of length $n = 2^m$ and dimension $k \geq n - t \cdot m$, then produces the associated kxn generator matrix G. A kxk nonsingular matrix S and a nxn permutation matrix P are used to scramble the generator matrix G so that $G' = S \cdot G \cdot P$. The public key is G', while the private keys are S, G, and P.

Encryption : Given a k-bit plaintext message $\underline{m}$, the corresponding ciphertext $\underline{c}$ is calculated as follows:

$$\underline{c} = \underline{m}\cdot G' + \underline{e}, \tag{1}$$

where $\underline{e}$ is an artificial error vector of length n and weight t.

Decryption : Compute the syndrome vector

$$\underline{s} = \underline{c}\cdot P^{-1}\cdot H^T = \underline{m}\cdot S\cdot G\cdot H^T + \underline{e}\cdot P^{-1}\cdot H^T = \underline{e}\cdot P^{-1}\cdot H^T, \tag{2}$$

where H^T is a transpose of the parity-check matrix corresponding to G, and use a decoding algorithm such as Patterson algorithm to identify and remove $\underline{e}\cdot P^{-1}$. As a result, $\underline{m}\cdot S$ can be obtained. The sender's plaintext message $\underline{m}$ is then easily found by $\underline{m}\cdot S\cdot S^{-1}$.

II.2 Conventional Cryptanalytic Attacks.

The typical attack is to repeatedly select k bits at random from the ciphertext $\underline{c}$ to form $\underline{c}_k(= \underline{m}\cdot G_k' + \underline{e}_k)$ in the hope that any of the selected k bits are not in error, namely $\underline{e}_k = 0$. If there is no error in them, $\underline{c}_k\cdot[G_k']^{-1}$ is equal to $\underline{m}$, where G_k' is the kxk submatrix obtained by choosing k columns of G' according to the same selection of $\underline{c}_k$. The probability that there is no error in randomly selected k bits among n bits with t errors is P =

$$\frac{\binom{n-t}{k}}{\binom{n}{k}} \tag{3}$$

The total work factor for this attack is $W = k^3\cdot P^{-1}$, assuming the matrix inversion for $[G_k']^{-1}$ requires k^3 steps.

Adams and Meijer showed that the optimum value of t that maximizes W for n = 1024 is 37, which results in $W \cong 2^{84.1}$. Lee and Brickell suggested the systematic method of checking whether the obtained $\underline{c}_k\cdot[G_k']^{-1}$ is really $\underline{m}$. If $\underline{c}_k\cdot[G_k']^{-1}$ is not the true $\underline{m}$, then $\underline{m}\cdot G' + \underline{c}_k\cdot[G_k']^{-1}\cdot G'$ must have a Hamming weight of at least 2t. Hence, if $\underline{c} + \underline{c}_k\cdot[G_k']^{-1}\cdot G'$ has a Hamming weight of less than or equal to t, it is claimed that $\underline{c}_k\cdot[G_k']^{-1} = \underline{m}$. Based on this idea, they described an efficient algorithm to cryptanalyze the McEliece's cryptosystem by allowing a very small number of errors in the selected $\underline{c}_k$. The algorithm requires a work factor of $W \cong 2^{73.4}$ for n = 1024.

The McEliece's public-key cryptosystem is still secure against the attacks described above. Now, we present a cryptanalytic algorithm to break it.

III. A New Cryptanalytic Attack.

III.1 Main Idea.

The attack to be discussed is motivated by the existence of a systematic generator matrix of a code. If $G'(= S\cdot G\cdot P)$ can be transformed into G'' in a systematic form, namely in a row-echelon form, the corresponding systematic parity-check matrix H'' can be easily constructed from a relationship between $G'' = [\ I_{kxk} \mid A\ ]$ and $H'' = [\ A^T \mid I_{(n-k)x(n-k)}\ ]$, where A is a kx(n-k) binary matrix and A^T is its transpose.

The transformation of G' into G'' can be carried out by the Gauss-Jordan elimination. That is, several elementary row operations are applied to G'. The key point of this cryptanalytic approach is to make a connection between G' and G'' through a series of transformation matrices. The effect of applying elementary row operations to G' can be also obtained by premultiplying G' by a series of transformation matrices as follows:

$$D_k\cdot D_{k-1}\cdot\cdot\cdot D_1\cdot G' = G'', \tag{4}$$

where D_i, $1 \le i \le k$, is binary kxk tranformation matrix whose function is to reduce G' to a systematic generator matrix G''. The transformation matrices can be represented as a single kxk matrix D $(= D_k\cdot D_{k-1}\cdot\cdot\cdot D_1)$. If D is invertible, then

$$G' = D^{-1}\cdot G''. \tag{5}$$

The equation (1) can be rewritten as

$$\underline{c} = \underline{m}\cdot D^{-1}\cdot G'' + \underline{e}. \tag{6}$$

From G'', the corresponding parity-check matrix H'' can be easily obtained. G'' generates a linear code with the same code rate and minimum distance as the code generated by G. The intercepted ciphertext $\underline{c}$ can be cryptanalyzed as follows:

$$\underline{c}\cdot(H'')^T = \underline{m}\cdot D^{-1}\cdot G''\cdot(H'')^T + \underline{e}\cdot(H'')^T = \underline{e}\cdot(H'')^T, \tag{7}$$

because $G''\cdot(H'')^T = 0$. Based on the syndrome vector $\underline{e}\cdot(H'')^T$, a decoding algorithm can be applied to eliminate the error vector. $\underline{m}\cdot D^{-1}$ is the first k componenets from $\underline{c} - \underline{e} = \underline{m}\cdot D^{-1}\cdot G''$ because G'' is in a systematic form. The plaintext message $\underline{m}$ is obtained as follows: $\underline{m} = \underline{m}\cdot D^{-1}\cdot D$.

In the next section, the method for obtaining the systematic generator matrix G'' is described, and it is shown that D is always invertible.

III.2 How to obtain the Generator Matrix in a systematic form.

The operation of transforming G' into G'' consists of k stages. At each stage, a column vector of G' is reduced to an unit vector having a nonzero at the appropriate position through the following matrix multiplications:

$$G^{(j+1)} = E_j \cdot F_j \cdot G^{(j)} = D_j \cdot G^{(j)}, \ 1 \le j \le k, \qquad (8)$$

where $G^{(1)} = G'$. After the seventh stage, systematic generator matrix G'' is obtained. Two groups of binary kxk matrices, F_j and E_j, $1 \le j \le k$, are used to transform the public generator matrix G' into G''.

The role of the matrix E_j is to convert the j-th column of $G^{(j)}$ to the unit vector having a nonzero at the j-th position. Due to a peculiar structure of E_j which will be explained later, the j-th element of the j-th column of $G^{(j)}$, $a_{jj}{}^{(j)}$, should be nonzero before premultiplying $G^{(j)}$ by E_j. It is guaranteed through premultiplying $G^{(j)}$ by the matrix F_j whose role is to replace the j-th row of $G^{(j)}$ whose j-th element, $a_{jj}{}^{(j)}$, is not nonzero by another row of $G^{(j)}$ whose j-th element, $a_{ij}{}^{(j)}$, is nonzero. The replacing row should be selected among the rows below the j-th row of $G^{(j)}$, namely $j+1 \le i \le k$. If $a_{jj}{}^{(j)}$ is nonzero, $F_j = I_{kxk}$. Otherwise, F_j is obtained by swapping the two rows of the identity matrix I_{kxk} corresponding to the two rows of $G^{(j)}$ to be swapped. The matrix E_j is obtained by replacing the j-th column of the identity matrix I_{kxk} by the j-th column of $F_j \cdot G^{(j)}$.

At stage one, $G^{(2)} = E_1 \cdot F_1 \cdot G^{(1)}$ is obtained, where $G^{(1)} = G'$.

$$\begin{bmatrix} a_{11}{}^{(1)} & 0 & 0 & \cdot\cdot & 0 & 0 \\ a_{21}{}^{(1)} & 1 & 0 & \cdot\cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot\cdot & \cdot & \cdot \\ a_{k1}{}^{(1)} & 0 & 0 & \cdot\cdot & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_{11}{}^{(1)} & a_{12}{}^{(1)} & \cdots & a_{1k}{}^{(1)} & \vdots & \\ a_{21}{}^{(1)} & a_{22}{}^{(1)} & \cdots & a_{2k}{}^{(1)} & \vdots & \mathbf{A}^{(1)} \\ \cdot & \cdot & \cdots & \cdot & \vdots & \\ a_{k1}{}^{(1)} & a_{k2}{}^{(1)} & \cdots & a_{kk}{}^{(1)} & \vdots & \end{bmatrix}_{kxn} .$$

The kxk matrix on the left-hand side is E_1, and the kxn matrix on the right-hand side is $(F_1 \cdot G^{(1)})$, where $\mathbf{A}^{(1)}$ is a binary kx(n-k) submatrix of $(F_1 \cdot G^{(1)})$. The premultiplication of $G^{(1)}$ by F_1 results in $a_{11}{}^{(1)} = 1$. As a result of the above matrix multiplication, $G^{(2)} = E_1 \cdot (F_1 \cdot G^{(1)})$ is obtained. The first column of $G^{(2)}$ becomes an unit vector $[\ 1\ 0\ 0 \cdot\cdot\cdot 0\ ]^T$ by multiplying the rows of E_1 by the first column of $F_1 \cdot G^{(1)}$, where $a_{11}{}^{(1)} \cdot a_{11}{}^{(1)} = 1$ and $a_{i1}{}^{(1)} \cdot a_{11}{}^{(1)} + 1 \cdot a_{i1}{}^{(1)} = 0$ for $2 \le i \le k$, irregardless of the value of $a_{i1}{}^{(1)}$ because $a_{11}{}^{(1)}$ is nonzero.

At the j-th stage, $E_j \cdot (F_j \cdot G^{(j)})$ looks like:

$$\begin{bmatrix} 1 & 0 & 0 & \cdot & \cdot & 0 & a_{1j}^{(j)} & \cdot & \cdot & 0 \\ 0 & 1 & 0 & \cdot & \cdot & 0 & a_{2j}^{(j)} & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & 0 & a_{kj}^{(j)} & \cdot & \cdot & 1 \end{bmatrix} \cdot \left[\begin{array}{cccccccccc:c} 1 & 0 & 0 & \cdot & \cdot & 0 & a_{1j}^{(j)} & \cdot & \cdot & a_{1k}^{(j)} & \\ 0 & 1 & 0 & \cdot & \cdot & 0 & a_{2j}^{(j)} & \cdot & \cdot & a_{2k}^{(j)} & A^{(j)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ 0 & 0 & 0 & \cdot & \cdot & 0 & a_{kj}^{(j)} & \cdot & \cdot & a_{kk}^{(j)} & \end{array}\right] .$$

The operations explained above are repeated until we obtain $G^{(k+1)} = E_k \cdot F_k \cdot G^{(k)}$ which is the systematic generator matrix G'' derived from G'. As a result, $G''(= G^{(k+1)}) = E_k \cdot F_k \cdot \cdot \cdot E_1 \cdot F_1 \cdot G' = (D_k \cdot \cdot \cdot D_1) \cdot G'$, where $D_j = E_j \cdot F_j$ for $1 \le j \le k$. $(D_k \cdot \cdot \cdot D_1)$ can be rewritten as a single binary kxk matrix D. If D is invertible, a connection between G' and G'' is made through the matrix D^{-1} as follows: $G' = D^{-1} \cdot G''$.

Now, we prove that D is invertible by showing that E_j, $1 \le j \le k$, is invertible. F_j, $1 \le j \le k$, is obviously invertible because F_j is obtained by swapping two rows of the kxk identity matrix so that $|F_j| = |I|$.

Theorem 1 : E_j is invertible for $1 \le j \le k$.
<proof> We show that $|E_j| = 1$ for $1 \le j \le k$. A general form of $|E_j|$ looks like:

$$\begin{vmatrix} 1 & 0 & \cdot\cdot\cdot & 0 & a_{1j}^{(j)} & 0 & \cdot\cdot & 0 \\ 0 & 1 & \cdot\cdot\cdot & 0 & a_{2j}^{(j)} & 0 & \cdot\cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot\cdot\cdot & 1 & a_{(j-1)j}^{(j)} & 0 & \cdot\cdot & 0 \\ 0 & 0 & \cdot\cdot\cdot & 0 & a_{jj}^{(j)} & 0 & \cdot\cdot & 0 \\ 0 & 0 & \cdot\cdot\cdot & 0 & a_{(j+1)j}^{(j)} & 1 & \cdot\cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot\cdot\cdot & 0 & a_{kk}^{(j)} & 0 & \cdot\cdot & 1 \end{vmatrix}$$

When $|E_j|$ is expanded by the j-th column of E_j, $|E_j| = a_{1j}^{(j)} \cdot |L_1| + a_{2j}^{(j)} \cdot |L_2| + \cdot \cdot \cdot + a_{jj}^{(j)} \cdot |L_j| + \cdot \cdot \cdot + a_{kj}^{(j)} \cdot |L_k|$, where L_i, $1 \le i \le k$, is a (k-1)x(k-1) submatrix obtained by deleting the i-th row and the j-th column of E_j. Because the i-th column of L_i for $1 \le i \le j-1$ and $j+1 \le i \le k$ is always zero vector so that $|L_i| = 0$, $|E_j| = a_{jj}^{(j)} \cdot |L_j|$. $a_{jj}^{(j)}$ is nonzero and $|L_j| = |I_{(k-1)x(k-1)}| = 1$. Therefore, $|E_j| = 1$ for $1 \le j \le k$. Q.E.D.

III.3 Work Factor.

Most of the work is due to the matrix multiplications for deriving a systematic generator matrix G″ from G′. The following describes a simple algorithm to tranform G′ into G″ and to obtain $D = E_k \cdot F_k \cdot\cdot\cdot E_1 \cdot F_1$.

<u>Algorithm 1</u> :

```
D ← $I_{kxk}$
DO  j = 1   TO   k
    Obtain $F_j$ to compute $F_j \cdot G^{(j)}$ and $F_j \cdot D$
          { i ← j ;
            while ( $a_{ij}{}^{(j)} = 0$ ) i ←i + 1 ;
            if ( i = j ) then $F_j$ ← $I_{kxk}$
                              $F_j \cdot G^{(j)} = G^{(j)}$
                              $F_j \cdot D = D$
                         else $F_j$ ← swap(i-th row, j-th row) of $I_{kxk}$
                              $F_j \cdot G^{(j)}$ ← swap(i-th row, j-th row) of $G^{(j)}$
                              $F_j \cdot D$ ← swap(i-th row, j-th row) of D; }
    Obtain $E_j$ { replace the j-th column of $I_{kxk}$
                by the j-th column of $F_j \cdot G^{(j)}$ }
    Compute $G^{(j+1)}$ ← $E_j \cdot (F_j \cdot G^{(j)})$ ;
    Compute D ← $E_j \cdot (F_j \cdot D)$ ;
END DO
```

The effect of premultiplying $G^{(j)}$ and D by F_j is to swap the two rows of $G^{(j)}$ and two rows of D, respectively, which actually needs no matrix multiplication. At most $2 \cdot k \cdot n$ bit operations are needed in order to compute $E_j \cdot (F_j \cdot G^{(j)})$ because each row of E_j has at most two nonzero elements. Computing $E_j \cdot (F_j \cdot D)$ also needs at most $2 \cdot k^2$ bit operations. The total work factor to derive G″ and D is at most $W = k \cdot (2 \cdot k \cdot n + 2 \cdot k^2)$ bit operations. Therefore, the complexity of the algorithm 1 is $O(k^3)$, which is polynomial time. For the value of k = 524 which is suggested by McEliece, $W \cong 8.5 \cdot 10^8 \cong 2^{30}$ bit operations. Using a powerful microcomputer, it would take about a few hours to obtain both D and G″.

Once both G″ and D are derived, the remaining job of the cryptanalyst is the application of a decoding algorithm which is also performed by the legitimate receiver. Hence, W is the additional burden on the cryptanalyst over the work performed by the receiver.

III.4 Trapdoors.

In his cryptanalytic attack against iterated knapsack cryptosystems, Brickell[5] has transformed the public knapsack into one of several easy knapsacks which could have been used to break it without knowing the receiver's original easy knapsack.

Adams and Meijer[3] claimed that a Brickell-like attack against the McEliece's public-key cryptosystem would not be possible . They argued that the only transformation which converts the public generator matrix G' into another generator matrix G'' whose algebraic structure allows the cryptanalyst to use a decoding algorithm is the original transformation $G = S^{-1} \cdot G' \cdot P^{-1}$.

The attack suggested in the previous sections is based on a kind of trapdoor G'' which can be utilized to obtain the syndrome vector. The trapdoor G'' is obtained from the transformation $G'' = D^{-1} \cdot G'$, which contradicts the arguement of Adams and Meijer.

IV. Concluding Remarks.

A new cryptanalytic attack against the McEliece's public-key cryptosystem was given. Contrary to the other cryptanalytic algorithms which run in exponential time, the new algorithm suggested in this paper runs in polynomial time. It is based on the existence of a new trapdoor generator matrix which can be easily obtained from the public generator matrix. The algebraic structure of the trapdoor allows the cryptanalyst to easily decrypt the intercepted ciphertext.

Diffie[6] predicted that the McEliece's public-key cryptosystem would be fated to fall because it bears a structural similarity to Merkle-Hellman's knapsack system. The similarity between the McEliece's system and the knapsack system might be the likelihood of there being transformations from the public key into the trapdoor which can play the same role as the private key.

In conclusion, the McEliece's system is not secure any more.

References

[1] E.R. Berlekamp, R.J. McEliece, and H.C.A. van Tilborg, "On the Inherent Intractability of Certain Coding Problems," IEEE Trans. Inform. Theory, vol. IT-24, pp. 384-386, 1978.

[2] R.J. McEliece, "A Public-Key Cryptosystem based on Algebraic Coding Theory," DSN Progress Report, Jet Propulsion Lab., Pasadena, Ca., Jan-Feb., 1978.

[3] C.M. Adams and H. Meijer, "Security-Related Comments Regarding McEliece's Public-Key Cryptosystem," Advances in Cryptology - Crypto'87.

[4] P.J. Lee and E.F. Brickell, "An Observation on the Security of McEliece's Public-Key Cryptosystem," Advances in Cryptology - Eurocrypto'88.

[5] E.F. Brickell, "Breaking Iterated Knapsacks," Advances in Cryptology - Crypto'84.

[6] W. Diffie, "The First Ten Years of Public-Key Cryptography," Proc. of The IEEE, 1988, vol. 76, no. 5, pp. 560-577.