Extended Abstract

PRIVATE-KEY ALGEBRAIC-CODE CRYPTOSYSTEMS WITH HIGH INFORMATION RATES

Tzonelih Hwang

National Cheng Kung University
Institute of Information Engineering
Tainan, Taiwan, R.O.C.

T.R.N. Rao

University of Southwestern Louisiana
The Center for Advanced Computer Studies
Lafayette, Louisiana 70504

1. Introduction

Algebraic codes have been proven to be extremely powerful to combat errors in communications and computer systems. They can *reveal* information reliably in the presence of sustained interference due to channel noise or hardware failures [Lin 83, Rao 89]. Algebraic codes can also be used to *conceal* information from any unauthorized user. For examples, McEliece public-key cryptosystem applies the error-correcting capability of Goppa codes to provide data secrecy [McEliece 78]; Rao-Nam (RN) scheme applies simple algebraic codes to construct private-key cryptosystems [Rao 87]. These systems are called here *algebraic-code cryptosystems* (ACC).

Algebraic-code cryptosystems may encipher one plaintext in several different ways under one key, while retaining the independence of individual ciphertext. This property is also called *splitting* [Stinson 88], and these systems are referred here as *ACC with splitting*. ACC with splitting are important to prevent *ciphertext search attacks* and *pattern matching attacks* [Denning 82].

ACC with splitting require data expansion, a disadvantage of low information rate. To improve the information rate, we propose a new algebraic-code cryptosystem with splitting. The security of the system is investigated and appears to be as secure as the RN scheme modified by Struik and van Tilburg [Struik 87] which is called here the *ST scheme*. Although ACC could be constructed with nearly 100% information rates, but such systems do not provide splitting [Hwang 88].

A cipher is generally called *computationally secure* if it cannot be broken by systematic analysis with available resources [Denning 82]. By assuming that the ST scheme is computationally secure, the proposed scheme appears to be as secure as the ST scheme. Some comparisons will be given to show that the proposed scheme provides better information rate than the RN or ST schemes of the same block length.

1.1 RN and ST Scheme

Rao and Nam have proposed a private-key ACC using only simple codes (e.g., $d_{min} \approx 10$) which is called RN scheme here. RN scheme performs encryption by the following equation [Rao 87].

$$C = (M\mathbf{SG}+Z)\mathbf{P},$$

where M is the plaintext of length k; C is the ciphertext of length n; $\mathbf{S}$ is a random nonsingular matrix of rank k; $\mathbf{G}$ is the generator matrix of an (n, k) block code $\mathbf{C}$ which can correct t errors; $\mathbf{P}$ is a random permutation matrix of rank n; Z is an error vector randomly selected from the *syndrome-error table* that is constructed from the standard-array-decoding table of the code $\mathbf{C}$.

Based on the linear structure of the system, Struik and van Tilburg proposed a chosen-plaintext attack to crack the RN scheme. They further modified the RN scheme to withstand similar chosen-plaintext attacks [Struik 87]. The ST scheme can be described as the following.

$$C = f(M, Z)\mathbf{G}+Z,$$

J.J. Quisquater and J. Vandewalle (Eds.): Advances in Cryptology - EUROCRYPT '89, LNCS 434, pp. 657-661, 1990.

where $f^{-1}(f(M, Z), Z) = M$. The details of enciphering and deciphering are simple and can be found in [Struik 87].

2. The Proposed Scheme

2.1 Encryption and Decryption

Let M be a plaintext block of $(n+k)$ bits to be enciphered into a ciphertext C of $2n$ bits (n and k are the parameters of the linear code C to be used here). M can be divided into two sub-blocks M_1 and M_2 where M_1 is an n-bit sequence and M_2 is a k-bit sequence. P is a random permutation matrix of rank $2n$. The syndrome-error table in the RN scheme is replaced by an arbitrary nonlinear function h for the purpose of saving storage space and also increasing the security level. The function h takes two parameters X and Y, where X is a k-bit sequence and Y is an n-bit sequence, and produces a k-bit block $h(X, Y)$.

Encryption. Let Ψ be an invertible, nonlinear (n-bit to n-bit) function. G is the generator matrix of an (n, k) linear code C that can correct t random errors. The encryption is performed by simple steps as follows.

(a) Generate an n-bit random vector E of weight $\leq t$.
(b) Compute an error vector Z by
$$Z = \tau + E,$$
where $\tau = (h(M_2, E))\mathbf{G}$.
(c) Obtain the ciphertext $C = (C_1, C_2)\cdot\mathbf{P}$, where
$$C_1 = \Psi(M_1 + \tau) \tag{1}$$
$$C_2 = M_1 + M_2\mathbf{G} + Z. \tag{2}$$

Since Z is a function of both M_2 and E, the total number (N) of Z's is 2^k if h is chosen carefully. Because the value N is not determined by the number of cosets in the standard array decoding table of the code, simple codes can be used in the system and still achieve a high level of security as can be seen from the discussion in Sec. 2.2.

Decryption. The decryption can be carried out easily by the following steps.

(a) Compute $(C_1, C_2) = C\cdot\mathbf{P}^{-1}$.
(b) Compute $\Psi^{-1}(C_1)$.
(c) Obtain $M_2\mathbf{G} + E = \Psi^{-1}(C_1) + C_2$.
(d) Decode the result of (c) by applying the decoding of the introduced code C: Recover M_2 and obtain E.
(e) Compute $\tau = h(M_2, E)$.
(f) Recover $M_1 = \Psi^{-1}(C_1) + \tau$.

Note that any invertible, nonlinear function that can withstand ciphertext-only attacks can be used as the function Ψ. Therefore, Ψ can be very easy to implement. This will be shown in the next section.

2.2 Security of the Newly Developed Scheme

The encryption and decryption steps given above are fairly simple and are easy to implement. Clearly, due to Step (a) of encryption, it is indeed an ACC with splitting. What remains to be studied is its security. The following discussion on the security of the new scheme is based on the assumption that the ST scheme is *computationally secure*. The lemmas and theorem are given here without proofs but these proofs will be given in the full paper.

First we show that the partial encryption specified by Equation (2) is at least as secure as the ST scheme that uses the same code. The method used to prove the lemma is to show that the ST scheme can be reduced to the partial encryption given in Equation (2). On the other hand, Equation (2) cannot be reduced to the ST scheme.

Lemma 1

The partial encryption given in Equation (2) is at least as secure as the ST scheme.

Next, we show how the proposed scheme can be secure by investigating the structure of the scheme. First, we investigate a simplified scheme obtained by removing both functions Ψ and $\mathbf{P}$ from the original scheme and show that the simplified scheme can be broken by a chosen-plaintext attack as follows.

Lemma 2

The encryption scheme
$C = (M_1+r,\ M_1+M_2\mathrm{G}+Z)$
can be broken by a chosen-plaintext attack in $O(kn^2)$ *bit operations.*

If a linear function $\Psi_{(L)}$ is introduced to scramble the first part (M_1+r) of ciphertext C in Lemma 2, then the scheme is still insecure as shown by the following.

Lemma 3

The encryption scheme
$C = (\Psi_{(L)}(M_1+r),\ M_1+M_2\mathrm{G}+Z)$
can be partially broken by ST chosen-plaintext attacks in $O(n^2N^2\log N)$.

In what follows, we will show that if a nonlinear function $\Psi_{(N)}$, that is secure under ciphertext-only attacks while can be broken by a known-plaintext attack in polynomial time, is introduced, then the partial encryption given in Equation (1) is computationally secure.

Lemma 4

The partial encryption specified by Equation (1) is computationally secure if $\Psi_{(N)}$ *is secure under ciphertext-only attacks.*

So far we have shown that both Equations (1) and (2), i.e.,
$C_1 = \Psi_{(N)}(M_1+r)$, and
$C_2 = M_1+M_2\mathrm{G}+Z$
of the original scheme are computationally secure. However, it doesn't mean that the overall scheme is also computationally secure. For example, if $\Psi_{(N)}$ is specified by a key KEY_1 of length k_1 bits and the encryption of M_2 is specified by the key KEY_2 of k_2 bits, then an attack can crack the scheme in W operations, where $2^{k_2} \le W \le 2^{k_1}+2^{k_2}$, as follows.

(a) The cryptanalyst searches the key space of KEY_2 exhaustively to obtain M_2 and r. It requires 2^{k_2} operations in the worst case.
(b) $C_1 = \Psi_{(N)}(M_1+r)$ can be broken by a known-plaintext attack in less than 2^{k_1} operations because we have assumed that $\Psi_{(N)}$ can withstand ciphertext-only attacks while can be broken by a known-plaintext attack in polynomial time.

Note that an exhaustive search on the key space of $\Psi_{(N)}$ will not be productive because it involves the search for key spaces of both KEY_1 and KEY_2. Based on this analysis, we also have the following.

Lemma 5

The security level of the encryption scheme
$C = (\Psi_{(N)}(M_1+r),\ M_1+M_2\mathrm{G}+Z)$
is $O(2^{k_1}+2^{k_2})$.

Obviously, the cryptanalysis mentioned above highly depends on the correct partition of the ciphertext into C_1 and C_2. If the ciphertext is scrambled by a random permutation matrix P of rank $2n$, then the security level of the scheme will be highly increased. Based on the above discussion, the new scheme appears to be very secure. However, proving the security level of the scheme remains open and requires further research.

3. Comparison with the ST Scheme and Conclusion

First we note that the encryption block length of the proposed scheme is $2n$ where n is the block length of the error correcting code C. (The parameters of C are $(n, k, 2t+1)$.) The information rate of the code C is k/n, but for the encryption scheme, the rate is $(n+k)/2n$. To make a fair comparison, we choose an ST scheme (or RN scheme) of code of length $2n$ and t error correcting capability. If we choose BCH codes (or shortened codes) as examples for this comparison we arrive at Table 1.

Table 1. Comparison of information rates under the same ciphertext length $2n$.

New Scheme				ST Scheme			
n	k	t	I.R. (%)	I.R. (%)	2n+1	k'	t
15	11	1	86.6	83.9	31	26	1
	7	2	73.3	67.7		21	2
	5	3	66.7	51.6		16	3
31	26	1	91.9	90.5	63	57	1
	21	2	83.9	81.0		51	2
	16	3	75.8	71.4		45	3
	11	5	67.7	57.1		36	5

If we compare the information rate of the two schemes under the condition that they are using the same $(n, k, 2t+1)$ base code, then Table 2 shows that our scheme also provides better information rates i.e., $R' = \frac{n+k}{2n} > \frac{k}{n}$. Note that in this case, the ciphertext length of the new scheme is $2n$, while that of the ST scheme is n.

Table 2. Comparison of information rates under the same base code.

New Scheme				ST Scheme			
n	k	t	I.R. (%)	I.R. (%)	2n+1	k'	t
15	11	1	86.6	73.3	15	11	1
	7	2	73.3	46.7		7	2
	5	3	66.7	33.3		5	3
31	26	1	91.9	83.9	31	26	1
	21	2	83.9	67.7		21	2
	16	3	75.8	51.6		16	3
	11	5	67.7	35.5		11	5

We have constructed a private-key algebraic-code encryption scheme with splitting property whose encryption/decryption can be carried out efficiently. We have investigated its security and show how the scheme can be secure. The new scheme provides higher information rate

than the RN or ST schemes that use the same code or have the same ciphertext length and hence appears to be more practical.

References

[Denning 82] Dorothy E. Denning, *Cryptography and Data Security*, Addison Wesley, 1982.

[Denny 88] W.F. Denny and T.R.N. Rao, "Encryptions Using Linear and Nonlinear Codes: Implementation and Security Considerations", Ph.D Dissertation, Univ. of SW Louisiana, Spring 1988.

[Hwang 88] Tzonelih Hwang, *Secret Error-Correcting Codes and Algebraic-Code Cryptosystems*, Ph.D. Dissertation, Univ. of SW Louisiana, Summer, 1988.

[McEliece 78] R.J. McEliece, "A Public-Key Cryptosystem Based on Algebraic Coding Theory", *DSN Progress Report*, Jet Propulsion Laboratory, CA., Jan. & Feb. 1978, pp 42-44.

[NBS 77] "Data Encryption Standard", FIPS PUB 46, National Bureau of Standard, Washington, D.C., Jan. 1977.

[Rao 87] T.R.N. Rao and K.H. Nam "A Private-Key Algebraic-Code Cryptosystem," Advances in CRYPTO 86, editor A.M. Odlyzko, New York, Springler Verlag, pp. 35-48, 1987.

[Rao 89] T.R.N. Rao and E. Fujiwara, *Error Control Coding for Computer Systems*, Prentice Hall, 1989.

[Stinson 88] D.R. Stinson, "Some Construction and Bounds for Authentication Codes," *Journal of Cryptology*, Vol. 1, No. 1, 1988.

[Struik 87] R. Struik and van Tilburg J., "The Rao-Nam Scheme is Insecure Against a Chosen-plaintext Attack," CRYPTO '87.

[Tilburg 88] J. van Tilburg, "On the McEliece Public-key Cryptosystem," a paper presented in CRYPTO '88, to appear.