Cryptanalysis of Private-Key Encryption Schemes Based on Burst-Error-Correcting Codes

Hung-Min Sun

Department of Information Management
ChaoYang Institute of Technology
WuFeng, Taichung County, Taiwan

Shiuh-Pyng Shieh

Department of Computer Science and
Information Engineering
National Chiao Tung University
Hsinchu, Taiwan 30050

Abstract

Recently, Alencar et al. proposed a private-key encryption scheme based on the use of burst-error-correcting codes. After that, Campello de Souza et al. implemented Alencar et al.'s scheme by array codes which is a class of burst-error-correcting codes. In this paper, we will show that these two schemes are insecure against chosen plaintext attacks.

1. Introduction

In 1978, McEliece proposed a public-key cryptosystem based on algebraic coding theory [1]. The idea of the cryptosystem is based on the fact that the decoding problem of a general linear code is an NP-complete problem. Compared with other public-key cryptosystems, McEliece's scheme has the advantage of high-speed encryption and decryption. In 1989, Rao and Nam modified the McEliece's scheme to construct a private-key algebraic-code cryptosystem which allows the use of simpler codes [2]. However, the Rao-Nam system is subjected to some chosen plaintext attacks [2][3], and therefore is insecure. In 1993, Alencar et al. [4] proposed a private-key cryptosystem based on binary linear block burst-error-correcting codes, which has drawn much attention. The idea of the cryptosystem is based on the fact that the burst-correcting capacity of a binary linear block burst-error-correcting codes is, in general, larger thanits random error-correcting capacity. After that, Campello de Souza et al. analyzed the security of the Alencar et al.'s schemeand concluded Alencar et al.'s scheme is secure against chosen-plaintext attacks [5]. In addition, they implemented Alencar et al.'s technique by a class of array codes, which have a fixed random error-correcting capacity ($t = 1$) [5].

In this paper, we will show that Alencar et al.'s private-key cryptosystem based on the burst-correcting codes is insecure against chosen plaintext attacks. Therefore, Campello de Souza et al.'s scheme can be broken in the same way, too.

2. Alencar et al.'s Scheme

In this section, we will introduce the private-key burst-error-correcting code encryption proposed by Alencar et al [4]. First, we introduce the concept of burst-error-correcting codes. Let $B(n, k, d, b)$ denote a binary linear block burst-error-correcting code of length n, dimension k, minimum Hamming distance d, capable of correcting single bursts of lengths up to b. A burst of length b means that a binary vector of length n whose nonzero components are confined to b consecutive positions with ones in the first and the last positions. We also include the case of end-around burst whose errors confined to i high-order positions and b-i low-order positions [6]. Let t be the random error-correcting capacity of the code and $d = 2t+1$. We assume that $b > t$. Alencar et al.'s scheme works as follows.

Secret key: G is the generator matrix of a $B(n, k, d, b)$,
P is an $n \times n$ permutation matrix.

Encryption:
Let the plaintext M be a binary k-tuple.
The ciphertext C is calculated by the sender: $C =$

---

[1]This work was supported in part by the National Science Council, Taiwan, ROC under grant NSC 85-2213-E-009-032.

CCS '96, New Delhi, India

$(MG + E_{l,w})P$, where $E_{l,w}$ is a random burst of length l with Hamming weight w. It is assumed that $w_{\min} \leq w \leq l \leq b$, where $w_{\min}$ is a fixed number greater than t.

Decryption:

The receiver first calculates $C' = CP^{-1} = MG + E_{l,w}$, where P^{-1} is the inverse of P. Then the sender removes the errors embedded in C' to obtain M by using the decoding algorithm of the code $B(n, k, d, b)$.

Campello de Souza et al. [5] analyzed the security of Alencar et al.'s scheme as follows. The encryption algorithm can be rewritten as

$$C = (MG + E_{l,w})P = MG' + E'_{l,w}$$

where $G' = GP$ and $E'_{l,w} = E_{l,w}P$. The matrix G' can be found by a chosen plaintext attack suggested by Campello de Souza et al.'s [5]. The cryptanalyst chooses a plaintext of the form M_i with only one 1 in the ith position for i = 1, ..., k. He encrypts M_i a number of times and obtains an estimate of g'_i, the ith column of the matrix G', with a desired degree of certainty. Repeating this step for i = 1, ..., k gives G'. Campello de Souza et al. conclude that the Alencar et al.'s scheme is still secure against chosen plaintext attacks though the matrix G' is known. The security of the system relies on the difficulty of decoding a general linear code, as in the McEliece scheme [1], and on the difficulty of correcting a number of errors which is beyond the error-correcting capacity of a given code (the code with generator matrix G' can correct t random errors, but $E'_{l,w}$ is an error vector with Hamming weight w, where $w>t$).

3. Cryptanalysis of Alencar et al.'s Scheme

In this section, we will show that the permutation matrix P in Alencar et al.'s scheme can be determined by a known plaintext attack if we have the matrix G'. Therefore, the matrix G can be computed by $G = G'P^{-1}$. Thus, the Alencar et al.'s scheme can be broken when the private key of the system, P and G , is known.

We assume that $E_{l,w} = \ < e_1, e_2, \ldots, e_i, \ldots, e_n >$ and $E'_{l,w} = < e_1', e_2', \ldots, e_i', \ldots, e_n' >$.
Because $E_{l,w}P = E'_{l,w}$ where P is a permutation matrix, we can write

$$\begin{aligned} E_{l,w}\mathrm{P} &= < e_1, e_2, \ldots, e_i, \ldots, e_n > \mathrm{P} \\ &= < e_{\tau(1)}, e_{\tau(2)}, \ldots, e_{\tau(i)}, \ldots, e_{\tau(n)} > \\ &= < e_1', e_2', \ldots, e_i', \ldots, e_n' >, \end{aligned}$$

where $\tau(\cdot)$ is an one-to-one and onto function from {1, 2, ..., n } to itself.

If we can find the mapping function $\tau(\cdot)$, then the permutation matrix P can be obtained.
In order to find the mapping function $\tau(\cdot)$, we give some definitions and propose some lemmas in the following.

Definition 1: The neighborhood of $\tau(i)$ with distance b-1 is
$N_b(\tau(i)) = \{ \tau(j) \mid |\tau(i) - \tau(j)| \leq b-1 \text{ or } |\tau(i) - \tau(j)| \geq n-b+1 \}$.

Note that: Each $N_b(\tau(i))$ has the size $2b$-1, i.e., $|N_b(\tau(i))| = 2b-1$.

In the following, we give an example describing the concept of $\tau(i)$ and $N_b(\tau(i))$.

Example: Let n=9, b=2, $E_{l,w} = < e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_9 >$ and

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

then we have

$$\begin{aligned} E_{l,w}P &= < e_{\tau(1)}, e_{\tau(2)}, e_{\tau(3)}, e_{\tau(4)}, e_{\tau(5)}, e_{\tau(6)}, e_{\tau(7)}, e_{\tau(8)}, e_{\tau(9)} > \\ &= < e_6, e_3, e_8, e_9, e_2, e_4, e_1, e_5, e_7 >. \end{aligned}$$

That is, $\tau(1)=6$, $\tau(2)=3$, $\tau(3)=8$, $\tau(4)=9$, $\tau(5)=2$, $\tau(6)=4$, $\tau(7)=1$, $\tau(8)=5$, and $\tau(9)=7$.
From Definition 1, we obtain $N_b(\tau(i))$ as follows.

$N_b(\tau(1)) = \{ \tau(1), \tau(8), \tau(9) \}$,
$N_b(\tau(2)) = \{ \tau(2), \tau(5), \tau(6) \}$,
$N_b(\tau(3)) = \{ \tau(3), \tau(4), \tau(9) \}$,
$N_b(\tau(4)) = \{ \tau(3), \tau(4), \tau(7) \}$,
$N_b(\tau(5)) = \{ \tau(2), \tau(5), \tau(7) \}$,
$N_b(\tau(6)) = \{ \tau(2), \tau(6), \tau(8) \}$,
$N_b(\tau(7)) = \{ \tau(4), \tau(5), \tau(7) \}$,
$N_b(\tau(8)) = \{ \tau(1), \tau(6), \tau(8) \}$,
$N_b(\tau(9)) = \{ \tau(1), \tau(3), \tau(9) \}$.

Lemma 1: If $|N_b(\tau(i)) \cap N_b(\tau(j))| = 2b-2$, then either $|\tau(i)-\tau(j)| = 1$ or $|\tau(i)-\tau(j)| = n-1$.

Definition 2: A sequence $x_1, x_2, \ldots, x_n$ is said to be cyclically sorted in increasing order if the smallest number in the sequence is x_i for some unknown i, and the sequence $x_i, x_{i+1}, \ldots, x_n, x_1, \ldots, x_{i-1}$ is sorted in increasing order.

Definition 3: A sequence $x_1, x_2, ..., x_n$ is said to be cyclically sorted in decreasing order if the largest number in the sequence is x_i for some unknown i, and the sequence $x_i, x_{i+1}, ..., x_n, x_1, ..., x_{i-1}$ is sorted in decreasing order.

Definition 4: Two sequences $x_1, x_2, ..., x_n$ and $y_1, y_2, ..., y_n$ is cyclically equivalent if there exists an integer i such that the sequence $x_1, x_2, ..., x_n$ is the same as the sequence $y_i, y_{i+1}, ..., y_n, y_1, ..., y_{i-1}$.

If we collect all the sets of $N_b(\tau(i))$ for $1 \le i \le n$, then we obtain the cyclically sorting of these $\tau(i)$'s in increasing order or decreasing order according to Lemma 1. We assume that $\tau(k_1), \tau(k_2), ..., \tau(k_n)$ is the cyclically sorting of $\tau(i)$'s for $1 \le i \le n$, where k_j and $k_i \in \{1, 2, ..., n\}$, $k_j \ne k_i$ if $i \ne j$. Then the sequence $\tau(k_1), \tau(k_2), ..., \tau(k_n)$ is cyclically equivalent to either the sequence $1, 2, ..., n-1, n$, or the sequence $n, n-1, ..., 2, 1$. Therefore, we can guess the sequence $\tau(k_1), \tau(k_2), ..., \tau(k_n)$ only from $2n$ possible sequences, i.e.,

sequence $1, 2, ..., n-1, n$,
sequence $2, 3, ..., n, 1$,
$\vdots$
sequence $n, 1, ..., n-2, n-1$,
sequence $n, n-1, ..., 2, 1$,
sequence $n-1, n-2, ..., 1, n$,
$\vdots$
sequence $1, n, ..., 3, 2$.

We can verify the correctness of each guess by testing whether the resulted P (obtained from $\tau(i)$'s) and G ($= G'P^{-1}$) can correctly decrypt the ciphertext into plaintext. Therefore, in order to break the Alencar et al.'s scheme, all we have to do is collecting all the sets of $N_b(\tau(i))$ for $1 \le i \le n$.

4. Collection of $N_b(\tau(i))$

In the following, we will discuss the working factor to obtain all the sets of $N_b(\tau(i))$ for $1 \le i \le n$. Because $C = MG' + E'_{l,w}$ and G' can be known from the analysis in section 2, we can collect error patterns of $E'_{l,w}$ as follows.

Given a pair of plaintext and ciphertext, (M, C), an error pattern of $E'_{l,w}$ can be computed by $E'_{l,w} = C - MG'$. Depending on the error pattern, it is clear that if $e_i' = 1$ and $e_j' = 1$, then either $|\tau(i) - \tau(j)| \le b-1$ or $|\tau(i) - \tau(j)| \ge n-b+1$, i.e., $\tau(i) \in N_b(\tau(j))$ and $\tau(j) \in N_b(\tau(i))$. It is clear that given $E_{l,w}$ with weight w in the encryption phase, $E'_{l,w}$ has the same weight w. From $E'_{l,w}$, we obtain $\binom{w}{2} = \frac{w(w-1)}{2}$ pairs of relations between $\tau(i)$ and $\tau(j)$.

Therefore, if the error patterns of $E_{l,w} =$
$< e_1, e_2, ..., e_j = 1, e_{j+1} = 1, ..., e_n >$,
$< e_1, e_2, ..., e_j = 1, e_{j+1}, e_{j+2} = 1, ..., e_n >$,
, ...,
$< e_1, e_2, ..., e_j = 1, ..., e_{j+b-1} = 1, ..., e_n >$, for $j=1, ..., n$, are randomly selected in the encryption phase, then we can collect all the sets of $N_b(\tau(i))$ for $1 \le i \le n$. In the following, we will estimate the probabilities of occurrence of these error patterns.

Lemma 2: If $\frac{a_i}{b_i} \ge k$, for $a_i, b_i > 0$, $1 \le i \le n$, then $\frac{a_1 + a_2 + ... + a_n}{b_1 + b_2 + ... + b_n} \ge k$.

Proof: Note that we have $\frac{a_i}{b_i} \ge k$, so $a_i \ge kb_i$.

Therefore, $\frac{a_1 + a_2 + ... + a_n}{b_1 + b_2 + ... + b_n} \ge \frac{kb_1 + kb_2 + ... + kb_n}{b_1 + b_2 + ... + b_n} = k$. (Q.E.D.)

The probability of occurrence of the error pattern $< e_1, e_2, ..., e_j = 1, e_{j+1} = 1, ..., e_n >$ is denoted by $p(e_j=1, e_{j+1}=1)$. Therefore,

$$p(e_j=1,\ e_{j+1}=1) = \frac{\sum_{w=w_{\min}}^{b} \sum_{i=w}^{b} (i-1)\times\binom{i-3}{w-3}}{\sum_{w=w_{\min}}^{b} \sum_{i=w}^{b} n\times\binom{i-2}{w-2}}$$

$$= \frac{1}{n}\left(\frac{\sum_{i=w_{\min}}^{b} (i-1)\times\binom{i-3}{w_{\min}-3} + .. + \sum_{i=b-1}^{b} (i-1)\times\binom{i-3}{b-4} + \sum_{i=b}^{b} (i-1)\times\binom{i-3}{b-3}}{\sum_{i=w_{\min}}^{b} \binom{i-2}{w_{\min}-2} + .. + \sum_{i=b-1}^{b} \binom{i-2}{b-3} + \sum_{i=b}^{b} \binom{i-2}{b-2}}\right)$$

$(\because t < w_{\min} \le i,\ \therefore i-1 \ge t)$

$$\ge \frac{t}{n}\left(\frac{\sum_{i=w_{\min}}^{b} \binom{i-3}{w_{\min}-3} + ... + \sum_{i=b-1}^{b} \binom{i-3}{b-4} + \sum_{i=b}^{b} \binom{i-3}{b-3}}{\sum_{i=w_{\min}}^{b} \binom{i-2}{w_{\min}-2} + ... + \sum_{i=b-1}^{b} \binom{i-2}{b-3} + \sum_{i=b}^{b} \binom{i-2}{b-2}}\right) . \cdots\cdots\cdots\cdots\cdots\cdots (1)$$

Note that

$$\frac{\binom{i-3}{w-3}}{\binom{i-2}{w-2}} = \frac{w-2}{i-2} \ge \frac{w_{\min}-2}{b-2} \text{ for } w_{\min} \le w \le i \le b.$$

By Lemma 2, we get

$$\frac{\sum_{i=w}^{b} \binom{i-3}{w-3}}{\sum_{i=w}^{b} \binom{i-2}{w-2}} = \frac{\binom{w-3}{w-3} + \binom{w-2}{w-3} + ... + \binom{b-3}{w-3}}{\binom{w-2}{w-2} + \binom{w-1}{w-2} + ... + \binom{b-2}{w-2}} \ge \frac{w_{\min}-2}{b-2},$$

for $w_{\min} \le w \le b$.

Therefore, from (1) and Lemma 2, we obtain

$$p(e_j=1,\ e_{j+1}=1) \geq \frac{t}{n}\frac{w_{\min}-2}{b-2}.$$

$$p(e_j=1,\ e_{j+2}=1) = \frac{\sum\limits_{w=w_{\min}}^{b}\sum\limits_{i=w}^{b}(i-2)\times\binom{i-3}{w-3}}{\sum\limits_{w=w_{\min}}^{b}\sum\limits_{i=w}^{b}n\times\binom{i-2}{w-2}} < p(e_j=1,\ e_{j+1}=1),$$

Similarly, $p(e_j=1,\ e_{j+3}=1) < p(e_j=1,\ e_{j+2}=1)$,
$\ldots, p(e_j=1,\ e_{j+b-1}=1) < p(e_j=1,\ e_{j+b-2}=1)$.

Therefore, we need only consider the probability of occurrence of error pattern

$$E_{l,w} = < e_1, e_2, \ldots, e_j = 1, \ldots, e_{j+b-1} = 1, \ldots, e_n >.$$

$$p(e_j=1,\ e_{j+b-1}=1)$$

$$= \frac{\sum\limits_{w=w_{\min}}^{b}\binom{b-2}{w-2}}{\sum\limits_{w=w_{\min}}^{b}\sum\limits_{i=w}^{b}n\times\binom{i-2}{w-2}}$$

$$= \frac{1}{n}\left(\frac{\binom{b-2}{w_{\min}-2}+\ldots+\binom{b-2}{b-3}+\binom{b-2}{b-2}}{\sum\limits_{i=w_{\min}}^{b}\binom{i-2}{w_{\min}-2}+\ldots+\sum\limits_{i=b-1}^{b}\binom{i-2}{b-3}+\sum\limits_{i=b}^{b}\binom{i-2}{b-2}}\right). \cdots\cdots\cdots\cdots(2)$$

Note that

$$\frac{\binom{b-2}{w-2}}{\sum\limits_{i=w}^{b}\binom{i-2}{w-2}}$$

$$= \frac{\binom{b-2}{w-2}}{\binom{w-2}{w-2}+\binom{w-1}{w-2}+\ldots+\binom{b-2}{w-2}}$$

$$= \frac{1}{\frac{\binom{w-2}{w-2}}{\binom{b-2}{w-2}}+\frac{\binom{w-2}{w-1}}{\binom{b-2}{w-2}}+\ldots+1}$$

$$\geq \frac{1}{b-w+1}$$

$$\geq \frac{1}{b-w_{\min}+1}, \text{ for } w_{\min} \leq w \leq b.$$

From (2) and Lemma 2, we obtain

$$p(e_i=1,\ e_{i+b-1}=1)$$

$$= \frac{1}{n}\left(\frac{\binom{b-2}{w_{\min}-2}+\ldots+\binom{b-2}{b-3}+\binom{b-2}{b-2}}{\sum\limits_{i=w_{\min}}^{b}\binom{i-2}{w_{\min}-2}+\ldots+\sum\limits_{i=b-1}^{b}\binom{i-2}{b-3}+\sum\limits_{i=b}^{b}\binom{i-2}{b-2}}\right)$$

$$\geq \frac{1}{n}\frac{1}{b-w_{\min}+1}.$$

Hence, the expected number of encryption using the error pattern $<e_1, e_2, \ldots, e_j = 1, \ldots, e_{j+b-1} = 1, \ldots, e_n>$ is less than or equal to $n(b-w_{\min}+1)$.

The expected number of pairs (*M*, *C*) needed to collect all the sets of $N_b(\tau(i))$ is equal to $\underset{1\leq j\leq n, 1\leq k\leq b-1}{Max}$ {the expected number of encryption using the error patterns $<e_1, e_2, \ldots, e_j = 1, \ldots, e_{j+k} = 1, \ldots, e_n>$}. Because $p(e_j=1,\ e_{j+1}=1) > p(e_j=1,\ e_{j+2}=1) > \ldots > p(e_j=1,\ e_{j+b-2}=1) > p(e_j=1,\ e_{j+b-1}=1)$, $\underset{1\leq j\leq n, 1\leq k\leq b-1}{Max}$ { the expected number of encryption using the error patterns $<e_1, e_2, \ldots, e_j = 1, \ldots, e_{j+k} = 1, \ldots, e_n>$ } is equal to the expected number of encryption using the error pattern $<e_1, e_2, \ldots, e_j = 1, \ldots, e_{j+b-1} = 1, \ldots, e_n>$. So, the expected number of pairs (*M*, *C*) needed to collect all the sets of $N_b(\tau(i))$ is equal to $n(b-w_{\min}+1)$. It is obvious that the system can be broken by chosen-plaintext attacks.

5. Conclusions

In this paper, we analyze the security of Alencar et al.`s private-key cryptosystem based on burst-correcting codes. We show that the system is insecure against the chosen-plaintext attacks. Similarly, the Campello de Souza et al's private-key cryptosystem based on the array codes is also insecure.

References

[1] McEliece, R.J., "A Public-Key Cryptosystem Based on Algebraic Coding Theory," *DSN Progress Report* 42-44, pp. 114-116, Jet Propulsion Laboratory, CA, January and February, 1978.

[2] Rao, T.R.N. and Nam, K., "Private-Key Algebraic-Code Encryption," IEEE Transactions, Vol. IT-35, No. 4, pp. 829-833, July, 1989.

[3] Brickell, E.F, and Odlyzko, A., "Cryptanalysis: a survey of recent results," Proc. IEEE, 76, (5), pp. 153-165, 1988.

[4] Alencar, F.M.R., Léo, A.M.P., and Campello de Souza, R.M., "Private-Key Burst Correcting Code Encryption," Proc. IEEE Int. Symp. Information Theory, January, pp. 227, 1993.

[5] Campello de Souza, R.M. and Campello de Souza, J., "Array Codes for Private-Key Encryption," Electronics Letters, Vol. 30, No. 17, pp. 1394-1396, August, 1994.

[6] Lin, S., Error Control Coding, Reading, Prentice-Hall, 1983.