Journal of the Faculty of Environmental Science and Technology,
Okayama University
Vol.9, No.1, pp.19-26, February 2004

A Study on the Linear Cryptanalysis of AES Cipher

Kenichi Sakamura[1], Wang Xiao Dong[1] and Hirofumi Ishikawa[2]

(Received November 27, 2003)

We have investigated the linear cryptanalysis of AES cipher in this article. As the previous encryption standard DES could be broken by the linear cryptanalysis, NIST decided a new encryption standard AES in 2000. We try to analyze one and two rounds AES cipher by the method of the linear cryptanalysis and learn the limits of this method. AES cipher provides a conspicuous difficulty in breaking its keys because of small bias of its S-box. We report the experimental results of success rate and are led to conclusion that this method would not work well on more than 3 rounds to break keys.

Keywords: *AES, Chosen plaintext attack, Linear cryptanalysis*

1 Introduction

We have investigated the linear cryptanalysis of AES cipher in this article. The National Institute of Standard and Technology (NIST) made a formal call for Advanced Encryption Standard (AES) in 1997 and selected Rijndael block cipher as AES (FIPS 197) in 2000 (NIST, 2001), because Data Encryption Standard (DES), which was the previous encryption standard in US since 1977 (NIST, 1977), could be broken by the linear cryptanalysis, a kind of the chosen plain text attack (Matsui, 1994) and was anxious for a lowering of security. AES cipher uses 128-bit as a block length and allows a variable key length among 128, 192 and 256-bit. It iterates a round 10, 12 or 14 times depending key length that is composed four different transformations, ByteSub, ShiftRow, MixColumn and AddRoundkey. Lucks (2000) improved the block cipher square attack proposed by Daemen et al. (1997) on 7-rounds from 6-rounds and Ferguson et al. (2001), 8 rounds.

We try to analyze one and two rounds AES cipher by the method of the linear cryptanalysis proposed by Matsui (1994) and learn the limits of this method. The cryptanalysis is based on a bias of S-box. As the absolute value of maximum or minimum bias obtained from AES S-box is much smaller than that from DES S-boxes, AES cipher provides a conspicuous difficulty in breaking its keys. We report the experimental results of success rate. In due consideration of these results with bias, it seems that this method would not work well on more than 3 rounds to break keys.

2 Materials and Methods

2.1 Analysis of ByteSub

In this section, we describe a feature of ByteSub, a constituent element of a round in AES cipher, which is defined by S-box (S_A). S_A has 8 input and 8 output bits. For AES S-box (S_A), the number $NS_A(\alpha,\beta)$ of input (x) satisfying the linear relation for all XOR values of input (x) and output ($S_A(x)$) bits masked by α and β is defined as follows:

$$NS_A(\alpha,\beta) = \#\{x \mid 0 \le x < 255,\ (\bigoplus_{s=0}^{8}(x[s] \bullet \alpha[s])) = (\bigoplus_{t=0}^{8}(S(x)[t] \bullet \beta[t]))\} \tag{1}$$

Department of Environmental Synthesis and Analysis, Graduate School of Natural Science and Technology, Okayama University 700-8530, Japan[1], Department of Environmental and Mathematical Science, Faculty of Environmental Science and Technology, Okayama University, 700-8530, Japan[2]

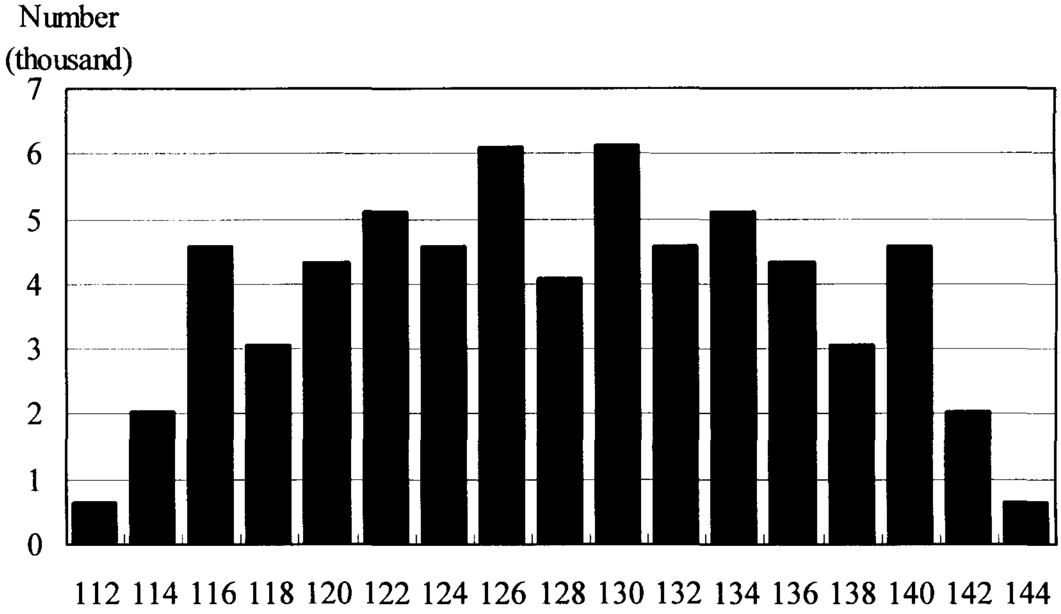

Fig. 1. The histogram of $NS_A(\alpha,\beta)$ for AES S-box. The axes of abscissas and ordinates indicate the value of $NS_A(\alpha,\beta)$, the number of pairs (α,β), respectively.

where α and β stand for mask of input bits and output bits($1\leq\alpha,\beta\leq255$), the symbols #, $\oplus$, $\bullet$, the number of x meeting the requirement, a bitwise XOR and AND. A bias is defined as $(NS_A(\alpha,\beta)-128)/256$. The eight different S-boxes are used for each 8 blocks in DES cipher, while the only one common S-box is used for all 16 blocks in AES cipher. The maximum absolute value of the biases for AES S-box is smaller than that for DES S-boxes. The maximal bias realized in the fifth S-box (S_{D5}) in DES, that is, 20/64=0.31 and $NS_{D5}(16,15)=12$ (Matsui,1994). On the other hand, we found that for AES S-box, the maximum absolute value of the bias was 16/256=0.0625 ($NS_A(\alpha,\beta)$ being 112 or 144). We showed the distribution of $NS_A(\alpha,\beta)$ in Fig. 1, and also the number of pairs (α,β) for the maximum and minimum biases in Table 1. AES S-box has remarkable features that the distribution of $NS_A(\alpha,\beta)$ is symmetrical about the center (128) (Fig.1), and that there exist just five linear approximations realizing the maximum or minimum bias (±0.0625) for any input mask $\alpha(1\leq\alpha\leq255)$, and for any output mask $\beta(1\leq\beta\leq255)$ too. Table 2 shows sets of above five linear approximations for 1 bit output and input cases.

2.2 Analysis of MixColumn

In this section, we describe a feature of MixColumn, a constituent element of a round in AES cipher. This MixColumn operation is carried out independently in each word (32bits), where a state (128bits) dividing into 4 words. On considering a word as 4 byte vector, this operation can be written by matrix multiplication on GF(2^8), where the field structure of GF(2^8) was defined by Daemen and Rijmen (2002).

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

a: input, b: output

The above expression is rewritten by the following formula:

$$b_i = '02'\bullet a_i \oplus '03'\bullet a_{(i+1)\bmod 4} \oplus '01'\bullet a_{(i+2)\bmod 4} \oplus '01'\bullet a_{(i+3)\bmod 4} \quad (i=0,1,2,3) \quad (2)$$

Table 1. The numbers of linear approximations for three maximum and minimum biases.

$NS_A(\alpha,\beta)$	112	114	116	140	142	144
Bias	-0.0625	-0.05469	-0.04688	0.046875	0.054688	0.0625
Number	640	2040	4592	4588	2040	635

There is a perfect linear approximation for any output or input mask in the operation (2). The formulae (3) illustrate these perfect relations for 1-bit output.

$$\begin{array}{llll}
b_i[0]= a_i[7] & \oplus a_{(i+1)\mathrm{mod}4}[7]\oplus a_{(i+1)\mathrm{mod}4}[0] & \oplus a_{(i+2)\mathrm{mod}4}[0] & \oplus a_{(i+3)\mathrm{mod}4}[0] \\
b_i[1]= a_i[7]\oplus a_i[0] & \oplus a_{(i+1)\mathrm{mod}4}[7]\oplus a_{(i+1)\mathrm{mod}4}[1]\oplus a_{(i+1)\mathrm{mod}4}[0] & \oplus a_{(i+2)\mathrm{mod}4}[1] & \oplus a_{(i+3)\mathrm{mod}4}[1] \\
b_i[2]= a_i[1] & \oplus a_{(i+1)\mathrm{mod}4}[2]\oplus a_{(i+1)\mathrm{mod}4}[1] & \oplus a_{(i+2)\mathrm{mod}4}[2] & \oplus a_{(i+3)\mathrm{mod}4}[2] \\
b_i[3]= a_i[7]\oplus a_i[2] & \oplus a_{(i+1)\mathrm{mod}4}[7]\oplus a_{(i+1)\mathrm{mod}4}[3]\oplus a_{(i+1)\mathrm{mod}4}[2] & \oplus a_{(i+2)\mathrm{mod}4}[3] & \oplus a_{(i+3)\mathrm{mod}4}[3] \\
b_i[4]= a_i[7]\oplus a_i[3] & \oplus a_{(i+1)\mathrm{mod}4}[7]\oplus a_{(i+1)\mathrm{mod}4}[4]\oplus a_{(i+1)\mathrm{mod}4}[3] & \oplus a_{(i+2)\mathrm{mod}4}[4] & \oplus a_{(i+3)\mathrm{mod}4}[4] \\
b_i[5]= a_i[4] & \oplus a_{(i+1)\mathrm{mod}4}[5]\oplus a_{(i+1)\mathrm{mod}4}[4] & \oplus a_{(i+2)\mathrm{mod}4}[5] & \oplus a_{(i+3)\mathrm{mod}4}[5] \\
b_i[6]= a_i[5] & \oplus a_{(i+1)\mathrm{mod}4}[6]\oplus a_{(i+1)\mathrm{mod}4}[5] & \oplus a_{(i+2)\mathrm{mod}4}[6] & \oplus a_{(i+3)\mathrm{mod}4}[6] \\
b_i[7]= a_i[6] & \oplus a_{(i+1)\mathrm{mod}4}[7]\oplus a_{(i+1)\mathrm{mod}4}[6] & \oplus a_{(i+2)\mathrm{mod}4}[7] & \oplus a_{(i+3)\mathrm{mod}4}[7]
\end{array}$$
$$(i=0,1,2,3) \quad (3)$$

Table 2. The linear approximations of the maximal bias (±0.06) for AES S-box having 1 bit output or input only.

Output(β)	Bit	Input(α)									
1	0	45	-	103	-	142	-	163	-	196	-
2	1	77	-	106	-	151	-	176	-	253	-
4	2	106	+	128	+	176	+	218	+	234	+
8	3	53	+	64	+	117	+	216	+	237	+
16	4	34	+	70	+	100	+	170	+	206	+
32	5	4	-	185	-	189	-	224	-	228	-
64	6	2	-	112	-	114	-	220	-	222	-
128	7	57	+	110	+	129	+	184	+	239	+

Input(α)	Bit	Output(β)									
1	0	72	-	80	-	192	-	136	+	144	+
2	1	56	-	64	-	207	-	120	+	143	+
4	2	32	-	60	-	149	-	28	+	181	+
8	3	15	+	53	+	82	+	93	+	103	+
16	4	51	-	63	-	101	-	105	-	86	+
32	5	121	-	153	-	180	-	205	+	224	+
64	6	204	-	230	-	238	-	8	+	34	+
128	7	17	-	52	-	4	+	33	+	37	+

The symbol (+,-) after figure showing the signature of bias.

2.3 Linear Analysis (1-Step)

Now we investigate attacking the AES code of 1 round. We can ignore ShiftRow operation because it is easy to break through this operation in attacking the AES code. Thus, the attack process can deal with each word (32bit) separately which makes up a state (128bit).

Firstly, we introduce several notations using in this article. The symbols P, C and K are used as plaintext, ciphertext and key (32bit word). For a word A, A_i stands for a byte part (8bit) of A the position of which locates in $[8i, 8i+7]$ (i=0,1,2,3). For a byte B, $B[j]$ denotes j-th bit of B, and $B[j_1, j_2, \ldots, j_k]$, the bitwise XOR summation $B[j_1] \oplus B[j_2] \oplus \cdots \oplus B[j_k]$. The formula (4) illustrates 1 round of AES code.

$$\begin{aligned}&'02'\bullet S(P_i)\oplus'03'\bullet S(P_{(i+1)\bmod 4})\oplus'01'\bullet S(P_{(i+2)\bmod 4})\\&\oplus'01'\bullet S(P_{(i+3)\bmod 4})\oplus K_i=C_i\end{aligned}$$

$$(i=0,\ 1,\ 2,\ 3) \quad (4)$$

We can get the formula (5) when we pick up the linear relation of the 0-th bit key ($K_0[0]$) in (4).

$$\begin{aligned}&S(P_0)[7]\oplus S(P_1)[0,7]\oplus S(P_2)[0]\\&\oplus S(P_3)[0]\oplus C_0[0]=K_0[0]\end{aligned} \quad (5)$$

Combining the formula (5) with the maximum and minimum bias' linear approximations of 1-bit output modified by MixColumn operation (Table 2 and (3)), we find an effective linear approximation formula among the 0th-bit of key and ciphertext and several bits of plaintext:

$$\begin{aligned}&P_0[0,3,4,5]\oplus P_1[2,4,5]\oplus P_2[0,2,3,5]\oplus P_3[0,2,3,5]\\&\oplus 1\oplus C_0[0]=K_0[0]\end{aligned} \quad (6)$$

The reason why '1' appears in the left hand side is that the number of approximations used in (6) which have negative bias is odd. Using the above approximation, $K_0[0]$ is predictable by chosen plaintext attack. We illustrated the flow of this linear approximation in Fig. 2. By the same way as above, we can obtain linear approximation formulae for $K_0[1]$- $K_0[7]$, and estimate all bits of K_0. We can obtain 5 effective linear approximations for each P_i (i=0, 1, 2, 3) because there are just 5 linear approximations of the maximum or minimum bias for the demanded output mask. Therefore, we can use such 5^4=625 linear approximations to estimate a key, which produces good results; the more linear approximations one uses, the less chosen plaintexts one needs to estimate a cipher key.

2.4 Linear Analysis (2-Step)

We assigned a position number i (i=0, 1, 2, 3) to the operation multiplied by '02', '03', '01', '01' in MixColumn. To deal with the round operation which consists of ByteSub (S-box) and MixColumn, we introduce two functions, *Upstream* and *Downstream* as follows.

For an output mask (x), an input mask (y), a MixColumn position number i and h (h=0, 1, 2, 3, 4), *Upstream*(x,i,h), *Downstream*(y,i,h) denote an input mask, an output mask which derive from h-th maximum or minimum bias' linear approximation of output mask (x), input mask (y) modified by multiplicative operation of number i, respectively.

1, 2-round keys stand for $K1$, $K2$. AES cipher of 2-round is expressed by:

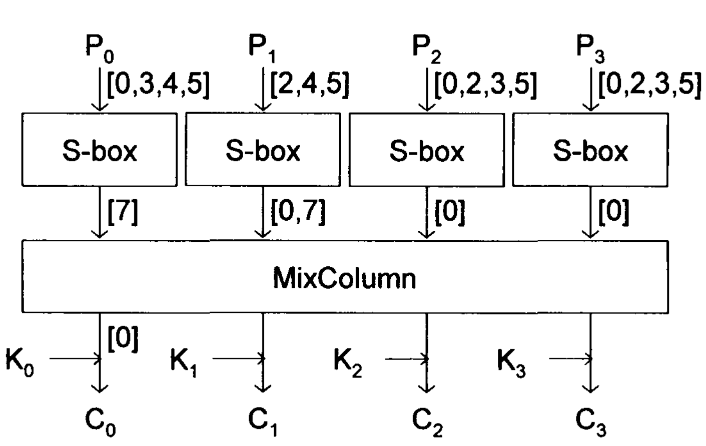

Fig. 2. 1-round AES cryptanalysis (see text).

$$MixColumn\ (S(MixColumn\ (S(P)) \oplus K1)) \oplus K2 = C \tag{7}$$

We can construct a linear approximation formula using *Upstream* or *Downstream* in the way of $K1$-pivot or $K2$-pivot, respectively.

Case 1. $K1$ -pivot.

The linear approximation involving $K1[0]$ was given by

$$P_0[Upstream(0,0,h_0)]) \oplus P_1[Upstream(0,1,h_1)] \oplus C_0[Downstream(0,0,h_2)] = K1_0[0] \oplus K2_0[Downstream(0,0,h_2)]$$

$$(0 \leq h_0,\ h_1,\ h_2 \leq 4) \tag{8}$$

When we choose opportune h_0, h_1, h_2 of *Upstream* and *Downstream* in (8), we get approximation formula:

$$P_0[0,3,4,5] \oplus P_1[2,4,5] \oplus C_0[1,2,3,7] = K1_0[0] \oplus K2_0[1,2,3,7] \tag{9}$$

We illustrated the flow of the above linear approximation in Fig. 3.

Case 2. $K2$-pivot.

The linear approximation involving $K2[0]$ was given by

$$P_0[Upstream(Upstream(0,0,h_2),0,h_0)]) \oplus P_1[Upstream(Upstream(0,0,h_2),1,h_1)] \oplus C_0[0] = K2_0[0] \oplus K1_0[Upstream(0,0,h_2)] \tag{10}$$

When we also choose opportune h_0, h_1, h_2 of *Upstream* in (10), we get approximation formula:

$$P_0[2,3,6] \oplus P_1[0,1,2,4,5] \oplus C_0[0] = K2_0[0] \oplus K1_0[0,3,4,5] \tag{11}$$

By the same way as (9), (11), the linear approximation that involves only one bit of $K1$ or $K2$ can be constructed. Since there are each five combinations of masks of P_0 and P_1 of (9) in §2.3, the key can be estimated accurately by using many combinations.

By the chosen plaintext attack against the linear approximations, a system of linear equations whose unknown quantities are all bits of keys $K1_0$ and $K2_0$ is obtained. The keys can be forecasted by solving a system of linear equations on GF(2). Their coefficients are derived from *Upstream* and *Downstream* in Case 1 and 2. As there are 16 unknown quantities in the equations, the system is in need of more than 16 linearly independent

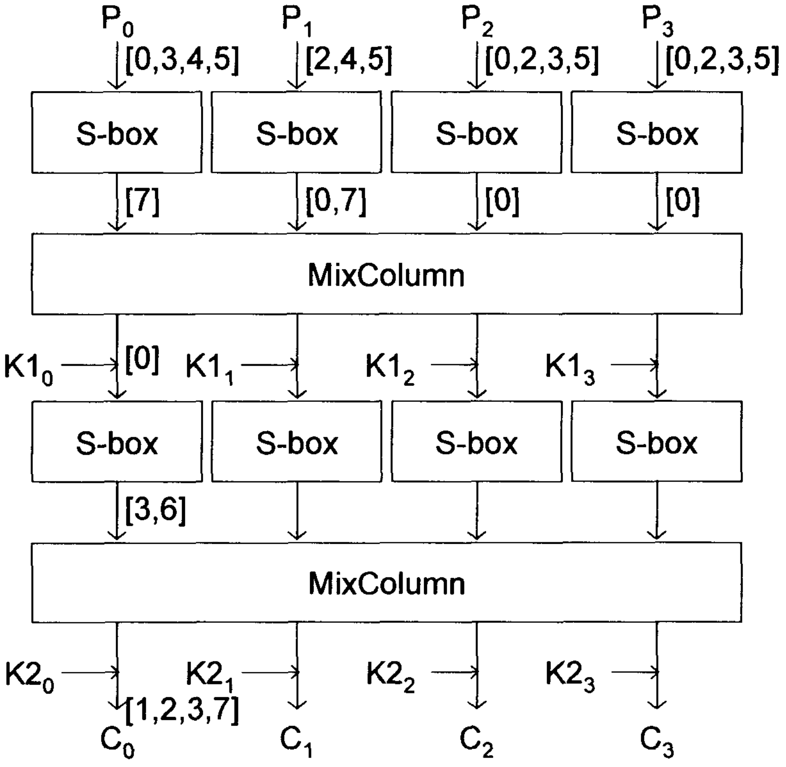

Fig. 3. 2-rounds AES cipher analysis (see text)

equations. We prepare 32 linear equations for an assurance of linear independence. These 32 equations are arranged in order: $K1_0[0]$, $\cdots$,$K1_0[7]$- pivot for *Downstream*(, ,*h*), $K2_0[0]$, $\cdots$, $K2_0[7]$ - pivot for *Upstream*(, ,*h*'), $K1_0[0]$, $\cdots$,$K1_0[7]$- pivot for *Downstream*(, ,*h*''), $K2_0[0]$, $\cdots$, $K2_0[7]$- pivot for *Upstream*(, ,*h*'''). The system of linear equations is formulated by a coefficient matrix $B = (b_{ij})$, a vector of unknown quantities $x = (x_j)$ and a constant vector $c = (c_i)$, that is, $Bx=c$. The bits of keys are arranged as follows:

$x = (K1_0[0], \ldots, K1_0[7], K2_0[0], \ldots, K2_0[7])$

The element of vector is determined by:

$$c_j = \begin{cases} 0 \cdots \text{if } j\text{ - th linear approximation holds} \\ \qquad \text{in the majority of chosen plain texts} \\ 1 \cdots \text{otherwise} \end{cases}$$

The elements of B are determined as follows.

(1) $b_{i,i} = b_{i+8,i+8} = b_{i+16,i} = b_{i+24,i+8} = 1\left(0 \le i \le 7\right)$

(2) If $\mathit{Downstream}(2^i,0,h_2)[j\,\mathrm{mod}8] = 1$,
then $b_{ij} = 1(0 \le i \le 7)$

(3) If $\mathit{Downstream}(2^{i\,\mathrm{mod}8},0,h'2)[j\,\mathrm{mod}8] = 1$,
then $b_{ij} = 1(16 \le i \le 23)$

(4) If $\mathit{Upstream}(2^{i\,\mathrm{mod}8},0,h''2)[j] = 1$,
then $b_{ij} = 1(8 \le i \le 15)$

(5) If $\mathit{Upwnstream}(2^{i\,\mathrm{mod}8},0,h'''2)[j] = 1$,
then $b_{ij} = 1(24 \le i \le 31)$

(6) Otherwise $b_{ij} = 0$

3 Results

3.1 Cryptanalysis for 1-step

We conducted a computer experiment in the linear cryptanalysis for two cases: (A) 2 -byte plaintexts (P_0, P_1) attack, (B) 4-byte plaintexts (P_0, P_1, P_2, P_3) attack. A set that is composed of two or four linear approximations for (P_0, P_1) (A) or (P_0, P_1, P_2, P_3) (B) which are chosen from 5 candidates (§2.3) is utilized for attack of 1-round AES. The experiments of (A) were made under the condition of 1, 5 and 25 sets served with attack, while the experiments of (B) were made under the condition of 1, 5 and 625 sets served. 25-sets (A) and 625-sets (B) consisted of all 5^2, 5^4 combinations, while 5-sets (A , B) were selected to have no common candidates. We summarized these results in Tables 3 and 4.

3.2 Cryptanalysis for 2-step

We carried out a computer experiment in the linear cryptanalysis attack against the incomplete 2-rounds AES that MixColumn of second round was omitted from, due to a technical reason. The above limitation made 2 bytes plaintext attack effective, so the whole 256^2 plaintexts were used in the attack. We could not increase the number of chosen plaintexts because the whole texts were already used, therefore we adopted the majority rule among the combination of systems of the linear equations that were derived from the selections of (*h*', *h*''') (§2.4) to estimate the keys of 2 rounds correctly, which was named as multi-system in contrast with the former as single-system. The experimental results were shown in Table 5.

Table 3. Success rates of the linear cryptanalysis by 2-byte chosen texts (A).

Number of chosen plain texts	Number of linear approximations sets served with attack					
	1		5		25	
	NT[a]	SR[b]	NT[a]	SR[b]	NT[a]	SR[b]
1,000	800	5.10	800	32.63	800	79.13
3,000	800	17.75	800	81.00	800	98.50
5,000	800	33.50	800	94.38	800	99.75
10,000	800	60.50	800	100.00	800	100.00
15,000	800	80.25	800	100.00	800	100.00
20,000	800	89.88	800	100.00	800	100.00

[a] Number of trials, [b] Success Rate (%)

Table 4. Success rates of the linear cryptanalysis by 4-byte chosen texts (B).

Number of chosen plain texts	Number of linear approximations sets served with attack					
	1		5		625	
	NT[a]	SR[b]	NT[a]	SR[b]	NT[a]	SR[b]
1,000					800	1.63
10,000	800	0.38	800	0.50	800	5.50
50,000	800	0.52	800	0.73	800	48.00
100,000	800	0.75	800	1.12	800	78.00
500,000	800	1.00	800	3.25	800	85.38
1,000,000	800	1.88	800	7.13		
5,000,000	800	6.00	800	37.50		
10,000,000	800	14.25	800	70.13		
20,000,000	800	29.13	800	95.13		

[a] Number of trials, [b] Success Rate (%)

Table 5. Success rates in the incomplete 2 rounds AES. The number of trials being 1000.

Method	Success rate (%)
Single-system	91.8
Multi-system	97.7

4 Discussion

Firstly, for the attack against 1-round AES we compared the success rates obtained from the computer experiment with the theoretical ones based on Lemma 2 in (Matsui, 1994). The bias that was derived from XOR sum of two linear approximations the biases of which were equally at 2^{-4} was computed at $2^{1}(2^{-4})^{2}=2^{-7}$ on the basis of Piling-up Lemma (Matsui, 1994). While the success rate on the attack of 2^{12} (1/4 $(2^{-7})^{-2}$) chosen plaintexts would be theoretically expected to be at 84.1%, the attainment of this rate was in need of about 16,000 chosen plaintexts in the experiment (Table3). In this case, the substantial bias would be estimated at 3.6×10^{-3}. In Table 6, we listed the substantial biases corresponding to the success rates in Table 3. Actually, the substantial bias for 5 sets served with attack was nearly equal to the theoretical bias, and such bias for 25sets was about double as much as the theoretical bias. As the theoretical bias for XOR sum of four linear approximations was computed at $2^{3}(2^{-4})^{4}=2^{-13}$, 2^{24} chosen plaintexts attack

Table 6. The substantial biases of 2-byte chosen texts attack (A).

Number of chosen plaintexts	Number of linear approximations set served with attack		
	1	5	25
	Substantial bias		
1,000			1.28E-02
3,000		8.02E-03	1.98E-02
5,000		1.12E-02	1.99E-02
10,000	1.34E-03		
15,000	3.48E-03		
20,000	4.51E-03		

Table 7. The substantial biases of 4-byte chosen texts attack (B).

Number of chosen plaintexts	Number of linear approximations set served with attack		
	1	5	625
	Substantial bias		
100,000			1.22E-03
500,000			1.66E-03
10,000,000		8.35E-05	
20,000,000		1.86E-04	

would expect the success rate at 84.1%. However, we needed about 3 billion in number to attain this rate in the experiment. In Table 7, we also listed the substantial biases corresponding to the success rates in Table 4. Similarly, the substantial bias for 5, and 625 sets served with attack were about one and ten times as much as the theoretical bias in this situation (B).

Secondly, we analyzed the experimental results in the incomplete 2 rounds AES attack. It is impossible for one set of linear approximations the bias of which would be estimated at 2^{-10} to attack 2 round keys with success rate at 84.1%, because it would require 2^{18} plaintexts, which exceed the whole number 2^{16} of plaintexts. Therefore, we carried out 25 sets attack (§3.2). In this situation, the substantial bias was computed at $2{\bullet}1.98{\times}10^{-2} {\bullet}2^{-4}=2.48{\times} 10^{-3}$ that can realize the success rate at 89.8 % with the whole plaintext (2^{16}) attack, which agrees with the rate at 91.8% in the experiment. The adoption of the majority rule, the multi-system attack, also contributed to an improvement of the success rate (Table 5). For the complete 2-rounds AES attack involving MixColumn of the second round, the substantial bias is estimated at $2{\bullet}1.66{\times}10^{-3} {\bullet}2^{-4}=2.08{\times} 10^{-4}$, when 625 sets attack (§3.1) are used. About 5.8 million plaintext attack can attain the success rate at 84.1%.

Finally, we study the limitation of this method. For 3 or 4 rounds AES attack, the substantial bias would be reduced to about $2^{-15.5}$ or $2^{-18.5}$, even if 625-sets attack would be used. Therefore, the necessary number of plain texts would increase to 2^{29} or 2^{37}, respectively. In due consideration of the whole number of plaintexts (2^{32}), we come to the conclusion that this method would not work well on more than 3 rounds.

References

Daemen, J. and Rijmen, V.: The Design of Rijndael, Springer-Verlag, Berlin, 2002.

Daemen, J., Knudsen, L. R. and Rijmen, V.: The block cipher square, In: Biham, E. (Ed.), Fast software encryption 97, Lecture Notes in Computer Science 1267, Springer-Verlag , 1997, pp. 149-165.

Feguson, N., Kelsey, J., Lucks, S., Schneier, B., Stay, M., Wagnes, D. and Whiting, D: Improved Cryptanalysis of Rijndael. In: Schneier, B. (Ed.) Fast Software Encryption, Lecture Notes in Computer Science 1998, Springer-Verlag, Berlin, 2001, pp. 213-230.

Lucks, S: Attacking 7 rounds of Rijndael under 192-bit and 256-bit keys, In Proceedings of the 3rd AES conference, April 13-14, 2000, New York, 2000, pp. 215-229

Matsui, M.: Linear Cryptanalysis Method for DES Cipher, In: Helleseth, T. (Ed.) Advances in Cryptology, Proceedings Eurocrypt '93, Springer-Verlag, Berlin, 1994, pp. 386-397.

NIST: Data Encryption Standard, FIPS 46, Washington DC, 1977.

NIST: Announcing the Advanced Encryption Standard (AES), FIPS 197, 2001.