Quadratic Relation of S-box and Its Application to the Linear Attack of Full Round DES

Takeshi Shimoyama*, Toshinobu Kaneko[†]

* TAO (Telecommunications Advancement Organization of Japan)
1-1-32 Shin'urashima-cho, Kanagawa-ku, Yokohama, 221 Japan
E-mail : shimo@yokohama.tao.go.jp

† Science University of Tokyo
2641 Yamazaki, Noda-shi, Chiba, 278 Japan
E-mail : kaneko@ee.noda.sut.ac.jp

Abstract. In this paper, we derive 7 quadratic relations over $GF(2)$ from the input and output bits of the S-boxes of DES. We apply one of those to an improved linear attack of full round DES. We describe an improved algorithm by combining the non-linear approximation method proposed by Knudsen and Robshaw, and the multiple approximation method proposed by Kaliski and Robshaw. This improvement can reduce the number of required plaintexts and ciphertexts pairs to 25/34 (73.5 %) of those number of pairs 2^{43} required in the linear attack by Matsui.

1 Introduction

It is well known that there is no linear relation between the input and output bits of each S-box of DES [Hel76, Bra77]. On the other hand, by representing S-boxes as Boolean polynomials [Sch82, Dav83, Way92, SAM97], it is easy to derive some algebraic relations of the input and output bits of S-boxes. We know that the degrees of these polynomials are less than or equal to 6, so there are algebraic relations of S-boxes with degree less than or equal to 6. Thus, the following problem may be natural to consider; what is the smallest degree of all algebraic relations of the S-boxes, and how are the algebraic relations which have the smallest degree represented? It can be shown that there is an algebraic relation over $GF(2)$ which has degree 3 in all S-boxes, so the above question is rewritten as follows; does there exist a quadratic relation? This paper shows that there are 7 quadratic relations of S-boxes S_1, S_4 and S_5 of DES; they can be derived by calculating the Gröbner bases of S-boxes with respect to the degree reverse lexicographic order in the Boolean polynomial ring. We apply one of these quadratic relations to improve the linear cryptanalysis offered by Matsui [Mat93].

In 1993, Matsui succeeded in recovering the secret key of the 16-round DES by using linear cryptanalysis in computational experiments [Mat94]. His main idea was the approximation of the S-boxes by linear relations. He recovered the key of 16-round DES by using 2^{43} pairs of plaintext and ciphertext, which took 50 days.

Since then, some theoretical and practical enhancements or extensions to linear cryptanalysis have been proposed [LH94, KR94, KR96, THHK98]. Kaliski and Robshaw proposed an algorithm using *multiple linear approximations* [KR94]. They applied it to small-round versions of DES to confirm its performance. As an example, they tried the 1-R attack and 2-R attack of 7-round DES, and succeeded to reduce the number of required texts. On the other hand, Knudsen and Robshaw proposed an algorithm using *non-linear approximation* [KR96]. They considered whether the linear approximations can be replaced with non-linear approximations. They constructed relatively simple non-linear approximations whose absolute bias are larger than that of the best linear approximation to S-box S_5, and adopted to 5-round DES. However, their techniques do not seem to offer any significant advantage over the existing attack to full round DES.

In this paper, we deal with derived quadratic relations of the round function of DES, like non-linear approximations, whose probabilities are 1. By using one of the quadratic relations, we construct an improved linear attack algorithm for full round DES. We combine the non-linear approximations method and the multiple approximations method. This improvement can reduce the number of plaintexts and ciphertexts to 25/34 (73.5 %) of the 2^{43} pairs required in Matsui's attack.

2   Deriving the algebraic relations of S-boxes

In [Sch82, Dav83, Way92, SAM97], the polynomial expressions of the S-boxes of DES in the Boolean polynomial ring over $GF(2)$ were constructed.

At first we summarize the notion of the Boolean polynomial ring. The Boolean polynomial ring over $GF(2)$ with n variables $t_1, ..., t_n$ is defined by the following quotient ring of the polynomial ring

$$GF(2)[t_1, ..., t_n]/Id(t_1^2 + t_1, ..., t_n^2 + t_n), \qquad (1)$$

where $Id(t_1^2 + t_1, ..., t_n^2 + t_n)$ is the ideal generated by the fundamental relations $t_1^2 + t_1 = 0, ..., t_n^2 + t_n = 0$ of Boolean variables $t_1, ..., t_n$.

Now we review how to obtain representations of the input and output bits of S-boxes in Boolean polynomial. For example, since the output of S-box S_1 corresponding to input 4 (= (0,0,0,1,0,0)) is 13 (= (1,1,0,1)) (Figure 1), we have the following algebraic relation of input Boolean variables $x_1, ..., x_6$ and output Boolean variables $y_1, ..., y_4$.

$$(x_1 + 1)(x_2 + 1)(x_3 + 1)(x_4 + 0)(x_5 + 1)(x_6 + 1)$$
$$((y_1 + 0)(y_2 + 0)(y_3 + 1)(y_4 + 0) + 1) = 0 \qquad (2)$$

Since there is an algebraic relation corresponding to each input from 0 to 63, we have 64 algebraic relations for each S-box which are similar to equation (2).

In commutative algebra, the technique of *Gröbner basis* is well-known as a basic tool [Bec93]. By using this technique, we can obtain another representation of these algebraic relations. For example, we can obtain the representation of each output bit y_i by the polynomial of input bits $x_1, ..., x_6$ by computing the Gröbner
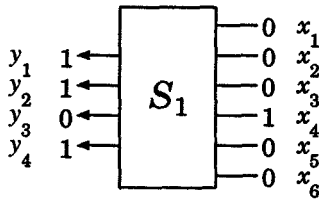
Fig. 1. S-box S_1 of DES

Table 1. The number of the quadratic and cubic relations of the S-boxes

S-box	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
quadratic	1	0	0	5	1	0	0	0
cubic	103	112	112	75	103	112	112	112

basis with respect to the lexicographic order of the sum set of polynomials in the above 64 algebraic relations and the fundamental relations $x_1^2 + x_1 = 0, ..., y_4^2 + y_4 = 0$ of all Boolean variables [SAM97]. We compute the Gröbner basis[1] in order to obtain algebraic relations which have much smaller *degrees* of S-boxes. With reference to the problem of the degree of the algebraic relations of S-boxes, we have the following lemma.

Lemma 1 *1. There is no linear relation for all S-boxes. [Hel76, Bra77].*
2. There is a cubic (that is, it has degree 3) algebraic relation for each S-box. (See Appendix A.)

Does there exists a quadratic (that is, it has degree 2) algebraic relation of each S-box? In order to see that, we can use the *reduced* Gröbner basis. By using the reduced Gröbner basis with respect to the degree reverse lexicographic order, we can obtain all algebraic relations of each S-box which are linearly independent over $GF(2)$. Table 2 shows the number of quadratic and cubic polynomials in the reduced Gröbner basis of the Boolean algebraic relations of S-boxes as derived above with respect to the degree reverse lexicographic order.

From Table 2, we know that there are 7 quadratic relations of S-boxes in total. All quadratic relations are given in Appendix B. Now we pay attention to the quadratic relation corresponding to the S-box S_5 : $(x_1, x_2, x_3, x_4, x_5, x_6) \rightarrow (y_1, y_2, y_3, y_4)$ as follows.

$$
\begin{aligned}
x_1 y_1 + x_1 y_2 + x_1 y_3 + x_1 y_4 + x_2 y_1 + x_2 y_2 + x_2 y_3 \\
+ x_2 y_4 + x_2 x_1 + x_5 y_1 + x_5 y_2 + x_5 y_3 + x_5 y_4 + x_5 x_2 \\
+ y_1 + y_2 + y_3 + y_4 + x_1 + x_2 + x_5 + 1 = 0
\end{aligned}
\tag{3}
$$

---

[1] In order to compute the Gröbner Basis over $GF(2)$, we used the computer algebra system Risa/Asir developed by Fujitsu LABORATORIES LTD. [Nor92].

It can be factorized into the polynomial as follows[2].

$$(y_1 + y_2 + y_3 + y_4 + x_2 + 1) \cdot (x_1 + x_2 + x_5 + 1) = 0 \qquad (4)$$

It is surprising that, in the first factor of the left side of the polynomial (4), there is the best linear approximation (5) with bias 5/16 corresponding to the input and output bits of S_5 discovered by Matsui [Mat93].

$$y_1 + y_2 + y_3 + y_4 + x_2 = 0 \qquad (5)$$

In the remaining part of this paper, we will try to apply the quadratic relation (4) for improving the linear attack of 16-round DES.

3   Application to non-linear cryptanalysis

We denote the sum of the coordinates from i_1 to i_j by $X[i_1, i_2, \ldots, i_j]$ for each vector $X \in GF(2)^n$. In particular, we denote the i-th coordinate of X by $X[i]$. We can easily extend the algebraic relation (4) to the algebraic relation of i-th round function $F_i : (X_i, K_i) \to F_i(X_i, K_i)$ as follows;

$$\text{A}^* : \quad \begin{aligned}(F_i[3, 8, 15, 24] + X_i[17] + K_i[26] + 1) \\ \cdot (X_i[16, 17, 20] + K_i[25, 26, 29] + 1) = 0,\end{aligned} \qquad (6)$$

where $X_i \in GF(2)^{32}$ is an input of i-th round and $K_i \in GF(2)^{48}$ is a i-th round key of the round function F_i.

In [KR96], Knudsen and Robshaw tried to apply the following non-linear approximations of S-boxes S_5 in order to raise the bias of best linear approximation of 5-round DES. Each non-linear approximation has bias 24/64, 18/64, respectively.

$$\begin{aligned}\text{A}' : y_1 + y_2 + y_3 + y_4 \\ = x_2 + x_1 x_2 + x_1 x_5 + x_2 x_6 + x_5 x_6 + x_1 x_2 x_6 + x_1 x_5 x_6 \\ \text{D}' : y_1 + y_2 + y_3 \\ = x_2 + x_4 + x_1 x_4 + x_1 x_6 + x_2 x_4 + x_2 x_6 + x_1 x_2 x_4 + x_1 x_2 x_6\end{aligned} \qquad (7)$$

By replacing linear approximation with these non-linear approximations, we can raise the bias 2.26 times more than those of linear approximation of 5-round DES, which reduces the number of plaintexts and ciphertexts pairs required for recovering one bit of key information of 5-round DES. They tried to recover more key bits using Matsui's 1-round and 2-round elimination method. They said, however, their techniques do not seem to offer any significant advantage over the basic attack.

In the above, we derived the quadratic representation (6) of the round function. In the first factor of equation (6), we can find the following best linear

---

[2] Because the Boolean polynomial ring is not a unique factored domain, each polynomial in the ring may have another factorized form. In fact, polynomial (3) can be factorized into another form $(y_1 + y_2 + y_3 + y_4 + x_1 + x_5 + 1) \cdot (x_1 + x_2 + x_5 + 1) = 0$.

approximation A for the i-th round function F_i with the absolute valued bias of $5/16$ that appeared in [Mat93].

$$A: \quad F_i[3, 8, 15, 24] + X_i[17] + K_i[26] = 0 \tag{8}$$

Matsui derived the following linear approximation (9) for 16-round DES by using best linear approximation of 14-round A-ACD-DCA-ACD- whose bias is $p_{14} = 1.19 \cdot 2^{-21}$ which is a concatenation of the three linear approximations A,C,D of the round function [Mat93],

$$\begin{aligned} P_r[3, 8, 14, 25] + P_l[17] + C_l[8, 14, 25] + F_1(P_r, K_1)[17] \\ + F_{16}(C_r, K_{16})[8, 14, 25] = K_2[26] + K_4[26] + K_5[4] + K_6[26] \\ + K_8[26] + K_9[4] + K_{10}[26] + K_{12}[26] + K_{13}[4] + K_{14}[26], \end{aligned} \tag{9}$$

where P_l, P_r are left and right halves of plaintext and C_l, C_r are left and right halves of ciphertext, respectively.

Since A* is a non-linear approximation with bias $1/2$, we obtain the following non-linear approximation A*-ACD-DCA-ACD- of 16-round DES by replacing the linear approximation A with quadratic relation A* which has higher bias than (9).

$$\begin{aligned} (P_r[3, 8, 14, 25] + P_l[17] + C_l[8, 14, 25] + F_1(P_r, K_1)[17] \\ + F_{16}(C_r, K_{16})[8, 14, 25] + K_2[26] + K_4[26] + K_5[4] + K_6[26] \\ + K_8[26] + K_9[4] + K_{10}[26] + K_{12}[26] + K_{13}[4] + K_{14}[26] + 1) \\ \cdot (P_l[16, 17, 20] + F_1(P_r, K_1)[16, 17, 20] + K_2[25, 26, 29] + 1) = 0 \end{aligned} \tag{10}$$

The bias of non-linear approximation (10) is higher than (9). We may not, however, be able to use (10) directly in order to reduce the number of required plaintexts and ciphertexts for recovering the *effective key* bits of 16-round DES, involved in (10), because the numbers of *effective text* bits and effective key bits involved in (10) become much larger than those in (9). In the next section, we will apply (10) to the *multiple approximations* to avoid this problem.

4 Application to multiple approximations method

In the previous section, we showed the non-linear approximation (10) of 16-round DES. The numbers of effective text bits and effective key bits corresponding to (10) are 24 and 26. We think it is not efficient to derive all 26 effective key bits at once, because the size of counter table corresponding to the effective keys is quite large. In order to avoid this problem, we deal with each factor in (10) independently. The following equation is the second factor of (10).

$$P_l[16, 17, 20] + F_1(P_r, K_1)[16, 17, 20] = K_2[25, 26, 29] \tag{11}$$

When (11) holds, the bias of (9) changes to $\epsilon_0 = (1/2)/(5/16)p_{14} = 8/5p_{14}$. When (11) does not hold, it changes to $\epsilon_1 = 2 \cdot (1 - (8/5)/2)p_{14} = 2/5p_{14}$. Thus, we deal with the linear approximation (9) as two linear approximations; one is when the equation (11) holds, and the other is when it does not hold.

Let N be the number of plaintexts and ciphertexts pairs. T_0 (, T_1) be the number of plaintexts and ciphertexts pairs such that the left side of equation (9) is equal to 0 and the equation (11) holds (, does not holds). We calculate the statistic $U = a_0 T_0 + a_1 T_1$ for some weights a_0 a_1 such that $a_0 + a_1 = 2$. For maximizing the distance between $N/2$ and the average $E[U]$ in terms of the standard deviation σ_U, we use Lemma 2.

Lemma 2 (Kaliski and Robshaw [KR94]) *The distance $|N/2 - E[U]|/\sigma_U$ is maximized for given N when the weights a_i are proportional to the biases of the linear approximations.*

From Lemma 2, we conclude that the best choices of the weights are a_0, a_1 such that $a_0 : a_1 = \epsilon_0 : \epsilon_1 = 4 : 1$.

Lemma 3 (Kaliski and Robshaw [KR94]) *The success rate of the algorithm with optimal weights a_i with respect to the biases ϵ_i is*

$$\Phi\left(2\sqrt{N}\sqrt{\frac{\sum \epsilon_i^2}{1 - 4\sum \epsilon_i^2}}\right). \tag{12}$$

Lemma 3 tells us that the success rate of original attack with N plain texts is the same as that of the improved attack with N' plaintexts as long as the following relation holds. On the assumption of random input, we can assume that the number of times of holding the equation (11) is $N'/2$.

$$2\sqrt{\frac{N'}{2}}\sqrt{\frac{(8/5 p_{14})^2 + (2/5 p_{14})^2}{1 - 4((8/5 p_{14})^2 + (2/5 p_{14})^2)}} = 2\sqrt{N}p_{14} \tag{13}$$

This is equivalent to

$$N' = \frac{25}{34}\left(1 - 4 \cdot \frac{68}{25}p_{14}\right) \cdot N \approx \frac{25}{34}N = 0.735 \cdot N. \tag{14}$$

Therefore, we can reduce the number of pairs to 73.5 % by using our attack.

5　Improved algorithm for attacking 16-round DES

In this section, we show the improved attack algorithm for 16-round DES. It still requires a large number of effective texts and effective keys in equations (9) and (11). In order to minimize the work spent in processing the data, we divide the algorithm into two parts. The first part is Matsui's original attack (part 1, 2, 3). The second is an improved part which replaces the exhaustive key search part in Matsui's attack with multiple approximations (part 4, 5, 6).

1　Compute plaintexts and ciphertexts pairs and count up the effective text bits of equation (9) and (11).

2 Count up the counters in the set $\mathcal{K}$ corresponding to effective key bits of (9) if the left side of (9) is zero.

3 Sort the effective keys of (9) using the counters $\mathcal{K}$ in order of reliability.

4 For the most reliable effective key of (9) when the right hand of (9) is zero, count up the counters in the set $\mathcal{H}_0$ corresponding to effective key bits of (11), with bias 4 or 1 by whether the left hand of (11) is zero or not, respectively, count up counters in the set $\mathcal{H}_1$ with bias 1 or 4, respectively, in the same way as $\mathcal{H}_0$.

5 Sort the effective keys of (11) using the counters $\mathcal{H}_0$ and $\mathcal{H}_1$ in order of reliability.

6 From the most reliable effective keys of (9) and (11), search for the remaining key bits.

In [Mat94], the effective text and key bits of (9) are shown. The 13 effective text bits of the left half of equation (9) are

$$P_r[32], P_r[1], ..., P_r[5], P_r[16], ..., P_r[21], P_r[3, 8, 14, 25] + P_l[17] + C_l[3, 8, 14], \quad (15)$$

and the 12 effective key bits of left half of equation (9) are

$$K_1[1], ..., K_1[6], K_1[25], ..., K_1[30]. \quad (16)$$

The 11 effective text bits of (11), if the key bits in (16) are fixed, are

$$P_l[16, 17, 20], P_r[8], ..., P_r[17], \quad (17)$$

and the 13 effective key bits of (11) if the key bits in (16) are fixed are

$$K_1[13], ..., K_1[24], K_2[25, 26, 29]. \quad (18)$$

Moreover, we can use another approximation replacing the plaintexts P and ciphertexts C in (9) and (11), similarly.

In our algorithm, we prepare a counter corresponding to the effective keys of equation (9) in the first part whose size is 12 bits long, and those of equation (11) in the second part whose size is 13 bits long. Thus, we can reduce the total size of effective key counter from 2×2^{25} to $2 \times (2^{12} + 2^{13})$ by using the improved algorithm.

6 The computer experiments

In this section, we show the results of computer experiments. Detail of the algorithm is shown in Appendix C. In order to estimate the complexity of our improved attack on 16-round DES, we consider the attack of 8-round DES by using plaintexts and ciphertexts pairs whose number is

$$1.49 \cdot 25/34 \cdot 2^{17} = 1.09 \cdot 2^{17}, \quad (19)$$

which is equivalent to the attack of 16-round DES using $25/34 \cdot 2^{43}$ pairs. Our computer experiments recovered the all round keys 10,000 times. Table 2 shows

Table 2. Complexity and success rate of attacks on 16-round DES (%)

complexity	2^{37}	2^{38}	2^{39}	2^{40}	2^{41}	2^{42}	2^{43}	2^{44}
(1)	47.9	54.9	62.0	68.7	74.3	81.4	86.6	90.9
(2)	24.7	30.3	36.6	44.1	52.2	60.5	68.8	76.2
(3)	50.1	54.4	61.3	68.2	75.1	81.3	86.8	91.2

(1) : The original algorithm with 2^{43} pairs.
(2) : The original algorithm with $25/34 \cdot 2^{43}$ pairs.
(3) : Our improved algorithm with $25/34 \cdot 2^{43}$ pairs.

the comparison of the complexity and the success rate of the Matsui's original algorithm and our improved algorithm.

We also conducted a computer experiment of recovering all round keys of full round DES using $25/34 \times 2^{43}$ pairs by our improved algorithm. The computer environments we used are 16 Sun workstations (Ultra SPARC 167MHz × 14 and 200 MHz × 2)[3] and a DEC workstation (Alpha 21164A 500MHz). By using the above machines and by implementing the algorithms in a bitslice manner [Bih97, NM97, SAM97] with Kwan's instructions sets of S-boxes [Kwa98][4], we achieved 1.14 Gbps in total[5]. It took about 6 days to compute all pairs of plaintexts and ciphertexts, 44 seconds for arranging the order of the key bits and about 4 hours for exhaustive key search (= about 2^{37} times of encryption). In total, we could recover the all key bits in less than 7 days.

7 Concluding remarks

In this paper, we derived the 7 quadratic relations of S-boxes. We used one of those quadratic relations for improving the linear cryptanalysis with 2-round elimination method proposed by Matsui. We constructed an improved algorithm for attacking 16-round DES which is a combination of the non-linear apprximation method and multiple approximation method. Moreover, we showed an effective algorithm that consisted of two parts to reduce the size of counter table of effective keys and minimizing the effort in processing the data. Overall, we could reduce the number of required plaintexts and ciphertexts pairs to $25/34 = 73.5$ % of that demanded by Matsui's original attack for recovering the key of 16-round DES. From computer experiments, when we attack 16-round DES with $25/34 \cdot 2^{43}$ pairs, the probability of finding the secret key equals that of Matsui's original attack.

---

[3] These workstations construct the parallel computer AP3000 with 16 nodes developed by FUJITSU LTD.
[4] For calculating one S-box, 51 instructions are required on average.
[5] Ultra SPARC 167MHz : 51 Mbps, 200MHz : 62 Mbps, Alpha 500MHz : 336 Mbps

References

[Bec93] T. Becker, V. Weispfenning., *"Gröbner Bases."* Springer-Verlag, New York, (1993).

[Bih97] E. Biham., "A Fast New DES Implementation in Software," FSE'97, LNCS 1267, pp. 245–251, (1997).

[Bra77] D. Branstead, J. Gait, S. Katzke., "Report of the workshop on cryptography in support of computer security," National Bureau of Standards, Sept, 21-22 1976, NBSIR 77-1291, Sept. (1977).

[Dav83] M. Davio, Y. Desmedt, M. Fosseprez, R. Govaerts, J. Hulsbosch, P. Neutjens, P. Piret,J. Quisiquater, J. Vandewall, P. Wouters., "Analytical charactersitics of the DES," CRYPTO'83, (1984).

[Hel76] M. Hellman, R. Merkel, R. Schroeppel, L. Washington, W. Diffie, P. Schweiter., "Results of an initial attempt to cryptanalyze the NBS data encryptoion standard," SEL 76-042, Stanford Univ. (1976).

[KR94] B. Kaliski, and M. Robshaw., "Linear Cryptnalysis Using Multiple Approximations," CRYPTO'94, LNCS 839, pp. 26–38, (1994).

[KR96] L. Knudsen, M. Robshaw., "Non-Linear Approximations in Linear Cryptanalysis," Eurocrypt'96, LNCS 1070, pp. 224–236, (1996).

[Kwa98] M. Kwan., "Bitslice DES," http://www.cs.mu.oz.au/~mkwan/bitslice, 12 May, (1998).

[LH94] S. Langford, M. Hellman., "Differential-Linear Cryptanalysis," CRYPTO'94, LNCS 839, pp. 17–25, (1994).

[Mat93] M. Matsui., "Linear Cryptanalysis Method for DES Cipher," Eurocrypt'93, LNCS 765, pp. 386–397, (1993).

[Mat94] M. Matsui., "The First Experimental Cryptanalysis of the Data Encryption Standard," CRYPTO'94, LNCS 839, pp. 1–11, (1994).

[NM97] J. Nakajima, M. Matsui., "Fast Software Implementation of MISTY on Alpha Processors," Proceedings of Korea-Japan joint workshop on information security and cryptology, pp. 55–64, (1997).

[Nor92] M. Noro, T. Takeshima., "Risa/Asir – a computer algebra system," Proceedings of ISSAC '92, ACM Press, pp. 387-396, (1992). (anonymous ftp from ftp://endeavor.fujitsu.co.jp/pub/isis/asir)

[Sch82] I. Schaumuller-Bichl., "Cryptanalysis of the Data Encryption Standard by method of formal coding," Proceedings of the workshop on cryptography, LNCS 149, pp. 235–255, (1982).

[SAM97] T. Shimoyama, S. Amada, S. Moriai., "Improved Fast Software Implementation of Block Ciphers," Proceedings of ICICS'97, LNCS 1334, pp. 269–273, (1997).

[Sti95] D. Stinson., "Cryptography : Theory and Practice," CRC Press Inc., (1995).

[THHK98] M. Takeda, T. Hamade, K. Hisamatsu, T. Kaneko., "Linear Cryptanalysis by Linear Sieve Method," IEICE Trans., Vol.E81-A,No.1, pp. 82–87, (1998).

[Way92] P. C. Wayner., "Content-Addressable Search Engines and DES-like Systems," CRYPTO'92, LNCS 740, pp. 575–586, (1992).

Appendix A. Proof of Lemma 1 (2)

For each S-box $S_i : (x_1, x_2, x_3, x_4, x_5, x_6) \to (y_1, y_2, y_3, y_4)$, there are Boolean polynomial representations in input bits $x_1, x_2, x_3, x_4, x_5, x_6$ of output bits y_1, ..., y_4 as follows.

$$\begin{cases} y_1 = f_1(x_1, ..., x_6) \\ y_2 = f_2(x_1, ..., x_6) \\ y_3 = f_3(x_1, ..., x_6) \\ y_4 = f_4(x_1, ..., x_6) \end{cases} \tag{20}$$

From these Boolean polynomials, we obtain the following algebraic relations.

$$x_i y_j = x_i f_j(x_1, ..., x_6) \quad (i \in \{1, ..., 6\}, j \in \{1, ..., 4\}) \tag{21}$$

The number of these polynomials in (20) and (21) is 28 ($= 4 + 4 \times 6$). It is easy to see that the right halves of these Boolean polynomials in (20) and (21) have degrees at most 6, and these polynomials are linearly independent over $GF(2)$. Since the number of terms with degree more than 3 in a Boolean polynomial with 6 variables are 22 ($= 15 + 6 + 1$), we can eliminate the terms with degree more than 3 from these algebraic relations.

Appendix B. All quadratic relations of S-boxes

We label the input and output bits to S-box as follows.

$$S_i : (x_1, x_2, x_3, x_4, x_5, x_6) \to (y_1, y_2, y_3, y_4)$$

$S_1 :$ • $x_2 x_1 + x_3 x_2 + x_4 x_2 + x_5 x_2 + x_6 x_2 + y_1 x_1 + y_1 x_2 + y_1 x_3 + y_1 x_4 + y_1 x_5 + y_1 x_6 + y_2 x_1 + y_2 x_2 + y_2 x_3 + y_2 x_4 + y_2 x_5 + y_2 x_6 + y_3 x_1 + y_3 x_2 + y_3 x_3 + y_3 x_4 + y_3 x_5 + y_3 x_6 + y_4 x_1 + y_4 x_3 + y_4 x_4 + y_4 x_5 + y_4 x_6 + y_4 y_1 + y_4 y_2 + y_4 y_3 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + y_1 + y_2 + y_3 + y_4 + 1 = 0,$

$S_4 :$ • $x_3 x_1 + x_5 x_1 + x_5 x_3 + y_1 x_3 + y_1 x_5 + y_2 x_3 + y_2 x_5 + y_3 x_3 + y_3 x_5 + y_4 x_3 + y_4 x_5 + x_1 + x_3 + x_5 + y_1 + y_2 + y_3 + y_4 + 1 = 0,$

• $x_2 x_1 + x_3 x_2 + x_4 x_1 + x_5 x_1 + x_5 x_4 + y_1 x_5 + y_1 x_6 + y_2 x_5 + y_2 x_6 + y_3 x_5 + y_3 x_6 + y_4 x_5 + y_4 x_6 + x_1 + y_2 + y_3 = 0,$

• $x_3 x_2 + x_4 x_1 + x_4 x_3 + x_5 x_1 + x_5 x_2 + x_5 x_3 + y_1 x_1 + y_1 x_2 + y_1 x_4 + y_1 x_6 + y_2 x_1 + y_2 x_2 + y_2 x_4 + y_2 x_6 + y_3 x_1 + y_3 x_2 + y_3 x_4 + y_3 x_6 + y_4 x_1 + y_4 x_2 + y_4 x_4 + y_4 x_6 + x_1 + x_3 + x_6 + y_1 + y_3 + 1 = 0,$

• $x_3 x_1 + x_3 x_2 + x_4 x_3 + x_5 x_1 + x_5 x_2 + x_5 x_4 + y_3 y_1 + y_3 y_2 + y_4 y_1 + y_4 y_2 + x_2 + x_3 + x_4 + y_1 + y_3 + 1 = 0,$

• $x_4 x_1 + x_5 x_1 + x_5 x_4 + y_1 x_2 + y_1 x_4 + y_2 x_2 + y_2 x_4 + y_2 y_1 + y_3 x_2 + y_3 x_4 + y_3 y_2 + y_4 x_2 + y_4 x_4 + y_4 y_1 + y_4 y_3 + x_2 + x_3 + x_4 + x_5 + x_6 + y_1 + y_4 + 1 = 0,$

$S_5 :$ • $x_2 x_1 + x_5 x_2 + y_1 x_1 + y_1 x_2 + y_1 x_5 + y_2 x_1 + y_2 x_2 + y_2 x_5 + y_3 x_1 + y_3 x_2 + y_3 x_5 + y_4 x_1 + y_4 x_2 + y_4 x_5 + x_1 + x_2 + x_5 + y_1 + y_2 + y_3 + y_4 + 1 = 0.$

Appendix C. Detail of our algorithm

In this section, we show detail of our improved algorithm. The following two non-linear equations are obtained from best linear expression of 14-round DES with 2-R elimination method and the quadratic relation of the input and output bits of S-box S_5;

$$
\begin{aligned}
&(P_r[3,8,14,25] + P_l[17] + C_l[8,14,25] + F_1(P_r,K_1)[17] \\
&\quad + F_{16}(C_r,K_{16})[8,14,25] + K_2[26] + K_4[26] + K_5[4] + K_6[26] \\
&\quad + K_8[26] + K_9[4] + K_{10}[26] + K_{12}[26] + K_{13}[4] + K_{14}[26] + 1) \\
&\qquad \cdot (P_l[16,17,20] + F_1(P_r,K_1)[16,17,20] + K_2[25,26,29] + 1) = 0,
\end{aligned}
\tag{22}
$$

$$
\begin{aligned}
&(C_r[3,8,14,25] + C_l[17] + P_l[8,14,25] + F_{16}(C_r,K_{16})[17] \\
&\quad + F_1(P_r,K_1)[8,14,25] + K_3[26] + K_4[4] + K_5[26] + K_7[26] \\
&\quad + K_8[4] + K_9[26] + K_{11}[26] + K_{12}[4] + K_{13}[26] + K_{15}[26] + 1) \\
&\qquad \cdot (C_l[16,17,20] + F_{16}(C_r,K_{16})[16,17,20] + K_{15}[25,26,29] + 1) = 0,
\end{aligned}
\tag{23}
$$

where P_l, P_r are left and right halves of plaintext and C_l, C_r are left and right halves of ciphertext, respectively, and K_i is i-th round key with 48 bit long. We define the notations of the vectors of effective text and key bits corresponding to equation (22) as follows.

$$
\begin{aligned}
A(P,C,K) &= P_r[3,8,14,25] + P_l[17] + C_l[8,14,25] \\
&\quad + F_1(P_r,K_1)[17] + F_{16}(C_r,K_{16})[8,14,25] & \in GF(2) \\
B(P,C) &= (P_r[3,8,14,25] + P_l[17] + C_l[8,14,25], \\
&\quad P_r[32], P_r[1], ..., P_r[5], C_r[16], ..., C_r[21]) & \in GF(2)^{13} \\
D(K) &= K_2[26] + K_4[26] + K_5[4] + K_6[26] + K_8[26] \\
&\quad + K_9[4] + K_{10}[26] + K_{12}[26] + K_{13}[4] + K_{14}[26] & \in GF(2) \\
E(K) &= (K_1[1], ..., K_1[6], K_{15}[25], ..., K_{15}[30]) & \in GF(2)^{12} \\
G(P) &= P_l[16,17,20] + F_1(P_r,K_1)[16,17,20] & \in GF(2) \\
H(P) &= (P_l[16,17,20], P_r[8], ..., P_r[17]) & \in GF(2)^{11} \\
I(K) &= K_2[25,26,29] & \in GF(2) \\
J(K) &= (K_1[13], ..., K_1[24]) & \in GF(2)^{12}
\end{aligned}
$$

Similarly, we define the following notations of effective text and key bit vectors related with equation (23).

$$
A'(P,C,K), B'(P,C), D'(K), E'(K), G'(C), H'(C), I'(K), J'(K)
$$

Algorithm 1 (Improved Attack Algorithm)

1 (Data Counting Phase) For N pairs $\{(P_1,C_1), ..., (P_N,C_N)\}$ of input and output, count up the following counters.

$$
\begin{aligned}
V(b,d) &= \#\{\ n \mid b = B(P_n,C_n), d = H(P_n)\} \\
V'(b',d') &= \#\{\ n \mid b' = B'(P_n,C_n), d' = H'(C_n)\}
\end{aligned}
$$

2 (Original Linear Attack Phase)

$$W(b) = \sum_d V(b,d), \ (0 \le b < 2^{13})$$
$$W'(b') = \sum_{d'} V'(b',d'), \ (0 \le b' < 2^{13})$$

By using the above counters W, W', sort the effective key vectors corresponding to the following in order of reliability by using the original linear attack.

$$((D(K), E(K)), \ (D'(K), E'(K)))$$

3 (Data Counting Phase II) Let $(k, k') \in GF(2)^{26}$ be the most reliable key vector obtained in the step (2).

$$T(d) = \sum_{cond_1} V(b,d),$$
$$T'(d') = \sum_{cond_2} V'(b',d'),$$
$$cond_1 : A(b,k) = D(k), \ cond_2 : A'(b',k') = D'(k').$$

4 (Key Counting Phase) For m, m' $(0 \le m, m' < 2^{12})$, calculate the following counters.

$$U(e,a) \ = 4 \sum_{cond_3} T(d) + \sum_{cond_4} T(d)$$
$$cond_3 : G(d,a) = e, \ cond_4 : G(d,a) = e + 1,$$
$$U'(e',a') = 4 \sum_{cond_5} T'(d') + \sum_{cond_6} T'(d')$$
$$cond_5 : G'(d',a') = e', \ cond_6 : G'(d',a') = e' + 1.$$

5 (Key Sort Phase) Sort the set of key vectors $\{h_j(= (e,a))\}$, $\{h'_{j'}(= (e',a'))\}$ which are belong to effective key vectors of

$$((I(K), J(K)), \ (I'(K), J'(K))),$$

in order of $|U(h_j) - 5/4N|$, $|U'(h'_{j'}) - 5/4N|$, respectively, and sort the set of pairs of key vector $(h_j, h'_{j'}) \in GF(2)^{26}$ in order of reliability.

6 (Exhaustive Search Phase) For each key vectors $(k_i, k'_{i'}, h_j, h'_{j'})$, search for the remaining 14 secret key bits in order of reliability until the correct value is found.

In exhaustive search phase (Algorithm 1,(6)), the reliability of a vector $(k_i, k'_{i'}, h_j, h'_{j'})$ has been determined in order of the magnitude of $((i + 1) \times (i' + 1))^{128} \times (j + 1) \times (j' + 1)$ which is the formula derived experimentally from the case of 8-round DES.

In the improved attack algorithm, all of effective key bits are 52 bits, that is, $D(K)$, $E(K)$, $D'(K)$, $E'(K)$, $I(K)$, $J(K)$, $I'(K)$, $J'(K)$ in total. There are, however, 10 of 52 bits are duplicated as below. Therefore, the number of the remaining key bits which should be executed exhaustive search is $56 - 52 + 10 = 14$.

$$K_1[3] = K_{16}[15], \ K_1[5] = K_{16}[24], \ K_1[13] = K_{16}[4], \ K_1[14] = K_{16}[22],$$
$$K_1[15] = K_{16}[23], \ K_1[16] = K_{16}[6], \ K_1[17] = K_{16}[21], \ K_1[19] = K_{16}[2],$$
$$K_1[20] = K_{16}[18], \ K_1[23] = K_{16}[1].$$