Patterns of Entropy Drop of the Key in an S-Box of the DES

(Extended Abtract)

K.C. Zeng, J.H.Yang, Z.T. Dai

Data and Communications Security Center

Graduate School of Academia Sinica

(I)

The S-boxes used in the DES have been looked at in various aspects like non-linearity,propagation characteristics and I/O correlation immunity. In the present work we try to make a cryptographic study of these boxes from a new viewpoint, namely,by investigating the way in which the uncertainty of the 6-bit key vector k which controls the work of an S-box diminishes,when a certain set of distinct plaintext vectors

$$Qr: x(1),x(2),\ldots,x(r)$$

put as queries to the box,together with the signals

$$Rr: y_1, y_2, \ldots, y_r,$$

appearing in response at a certain output position are assumed accessible to the cryptanalyst.

Such an approach to the problem of evaluating the S-boxes seems natural and necessary. As a matter of fact,the job of a cryptanalyst consists in nothing else than looking for observable query-response processes as inlets into the secrecy of a system,and in our case the query vectors are observable in reality,at least on the last round of the encryption algorithm.

C. Pomerance (Ed.): Advances in Cryptology - CRYPTO '87, LNCS 293, pp. 438-444, 1988.

Since every S-box is composed of four permutations

$$\pi_0, \pi_1, \pi_2, \pi_3$$

acting on the set

$$Z_{16}: 0, 1, 2, \ldots, 15,$$

it suffices to consider the problem for such "simplicial" S-boxes only. Here the vectors k and x belong to $V_4(F_2)$ and we have

$$y=f(x+k),$$

where the output function f(x) at any given position can be expressed in a unique way as a polynomial in $F_2[x_0,x_1,x_2,x_3]$, linear in each indeterminate separately. This is a polynomial of total degree at most 3, and balanced in the sense that it assumes the value 0 exactly eight times over $V_4(F_2)$.

Given any set Qr of r distinct queries to a permutation ,viewed as an S-box, we can introduce an equivalence relation in the set K of 16 possible key vectors by defining

$$k_1 \sim k_2 \iff f(x(i)+k_1) = f(x(i)+k_2),\ i=1,2,\ldots,r,$$

and decompose K accordingly into a disjoint union of equivalence classes

$$K = K_1 \cup k_2 \cup \ldots \cup K_5$$

If, by looking at the presumably accessible response signals y_i, $1 \le i \le r$, the cryptanalyst succeeds to decide that the key k in work falls in the class K_t, then the uncertainty of k diminishes from 4 bits to $\log_2|K_t|$ bits, resulting in $4-\log_2|K_t|$ bits of entropy drop. Thus the average entropy drop of the key which a given query set Qr may cause is

$$C(Q) = -\sum_{t=1}^{5} \frac{|K_t|}{16} \log_2 \frac{|K_t|}{16}$$

and a greater significance in evaluating the cryptographic strength of the permutation π is carried by the parameter

$$C_r = \max \{ C(Q_r) \}.$$

But one must have in mind,that the queries we are talking about here are of an in-active nature. Namely, one can observe which queries have been put to the box, but cannot organize them purposefully. So one has to take into consideration yet an other parameter, namely the probability

$$p_r = \text{Prob } (C(Q_r)=C_r),$$

assuming all the Q s equally possible.

We note in pass that the parameters C_r, p_r can as well be computed for an abitrary polynomial in $F_2[x_0,x_1,x_2,x_3]$.

DEFINITION. The sequence of parameter pairs

$$(C_1,p_1),(C_2,p_2),\ldots,(C_{16},p_{16})$$

computed at a given output position of a permutation is called the entropy drop pattern (EDP) of the working key at that position, or to meet the taste of an algebraist,the EDP of the corresponding output polynomial f(x).

THEOREM 1. The EDP of an arbitrary polynimial f(x) in $F_2[x_0,x_1,x_2,x_3]$ is invariant under the group G of 4-dimensional affine transformations acting on the indeterminates,extended by the involutional mapping $f(x) \longmapsto f(x)+1$.

THEOREM 2. The balanced polynomials of $F_2[x_0,x_1,x_2,x_3]$,linear in the indeterminates separately but non-linear in total,form three equivalent classes under the action of the extended affine group G defined above,with class representatives

$$f_1(x)=x_0x_1+x_2,\ f_2(x)=x_0x_1x_2+x_3,\ f_3(x)=x_0x_1x_2+x_0x_3+x_1$$

and corresponding EDPs as tabulated in the following

A			B			C		
$X_0X_1 + X_2$			$X_0X_1X_2 + X_3$			$X_0X_1X_2 + X_0X_3 + X_1$		
R	CR	PR	R	CR	PR	R	CR	PR
1	1	1	1	1	1	1	1	1
2	2	0.80	2	1.81	0.93	2	2	0.60
3	3	0.45	3	2.54	0.80	3	3	0.11
4	3	0.77	4	3	0.61	4	3.5	0.35
5	3	0.92	5	3.41	0.015	5	4	0.11
6	3	0.98	6	3.75	0.22	6	4	0.40
7	3	0.99	7	4	0.09	7	4	0.68
8	3	1	8	4	0.30	8	4	0.84
9	3	1	9	4	0.56	9	4	0.92
10	3	1	10	4	0.78	10	4	0.97
11	3	1	11	4	0.92	11	4	0.99
12	3	1	12	4	0.98	12	4	0.998
13	3	1	13	4	1	13	4	1
14	3	1	14	4	1	14	4	1
15	3	1	15	4	1	15	4	1
16	3	1	16	4	1	16	4	1

We see from the above table that output positions of entropy drop pattern C are more leaky for the key in work than those of pattern B, and positions of patern A are the most unleaky ones,though weaker in the aspect of non-linearity of the output function. This weakness, however, should not be taken too seriously in view of the iterative character of the encryption algorithm as a whole.

At the suggestion of this theorem we computed the 128 EDP's for the 32 permutations used in the DES with the following results

BOX0	BOX1	BOX2	BOX3
C C C C	C B A C	A C C B	C C C C
C C C C	A C C B	A B C C	C C C C
C C C C	B C B A	B C A B	C C C C
A C A C	C C C C	B C C A	C C C C

BOX4	BOX5	BOX6	BOX7
C C A B	C C C C	C B A C	B C C C
C C C C	C C C C	C B C A	A C B C
C B C C	C B B C	A C C B	C B B A
C B C B	A C B C	C A C B	C C A C

In addition to the uncanny appearance of Box 3, we have the following disapproving statistics with regard to the choice of the S-boxs in general. It turns out that more leaky patterns occur much more frequently in these boxes.

Pattern	Occurences	Frequency
A	18	14.06%
B	24	18.75%
C	86	67.19%

(II)

If the situation disclosed in the above is to some extent a weakness in the DES algorithm, then it is probably because of the fact that the non-linearity requirement has been unduly overstressed in choosing the boxes, while comparatively less attention has been paid to other equally important aspects such as I/O-correlation immunity and entropy drop of the key. In fact,there exist interesting mutual connections between these three aspects which we try to explain in a couple of words more in the following.

If we denote the input signal to and the output signal from a permutation π by x_i and y_j respectively, and define the I/O-correlation coefficient $c_{i,j}$ as

$$c_{i,j} = \text{Prob}\ (x_i = 0 \mid y_j = 0),$$

then for the 32 permutations used in the DES we have

$$c_{i,j} = 1/2 + h/8, \quad h = -1, 0, 1.$$

We call permutations satisfying this restriction selected ones,and the second author of the present paper has proved for

them the following result [1]

THEOREM 3. If for any permutation π define the degree of I/O-correlation immunity $c(\pi)$ to be the number of $c_{i,j}$ s which are equal to 1/2, and define the degree of non-linearity $d(\pi)$ to be the sum of the numbers of terms of degree 3 in the four output polynomials, then in the case of a selected permutation we have

$$c(\pi) + d(\pi) = 16.$$

This is a result much more refined than a similar one of Siegenthaler. In particular,it follows from this theorem and theorem2 above, that in the case of an I/O-correlation immune permutation, i.e., a permutation π with $c(\pi)=16$,all non-linear output polynomials are of degree 2, and hence of entropy drop pattern A.

I/O-correlation immune and, more generally,selected permutations satisfying the completenees and all requirements labelled in [2] as "Design Criteria", have been determined and classified by the same author. There are total 17433 equivalent classes of selected permutations under the transformation of subgroups of the symmetrical group S_{16} , including 46 equivalent classes of I/O-correlation immune permutations. His results show that there is enough space for constructing S-Boxes with correlation immune permutations exclusively. This will result in liquidating the leaky patterns B and C in the DES algorithm without bringing about substantial influence on the non-linearity and propagation requirements.

References

[1] J.H. Yang. Structure and Cryptographic Evaluation of S-Boxes of DES Type. Unpublished report.

[2] E.F. Brickell, J.H. Moore and M.E. Purtill, Structure in the S-boxes of the DES, Crypt-86.