A KNOWN PLAINTEXT ATTACK OF FEAL-4 AND FEAL-6

Anne TARDY-CORFDIR and Henri GILBERT
Centre National d'Etudes des Télécommunications (CNET)
PAA/TSA/SRC
38-40 rue de la République
92131 Issy les Moulineaux
FRANCE

Abstract

We present new results on the cryptanalysis of the FEAL-N blockcipher. As a matter of fact, almost all the attacks of this cryptosystem published so far are chosen plaintext attacks [3,4,5,7], except the announcement in [7] of a non-differential known plaintext attack of FEAL-4 which requires about 100000 plaintext blocks. We describe known plaintext attacks of FEAL-4 and FEAL-6, which require about 1000 and 20000 plaintext blocks respectively and are based on correlations with linear functions. Using similar methods, we have also found more recently an improved attack on FEAL-4, which requires only 200 knwon plaintext blocks.

1 The FEAL-N cryptosystem

FEAL-N is an N-round blockcipher proposed by NTT [1,2]. The standard version FEAL-8 is well suited for a fast software execution. So far, chosen plaintext attacks of FEAL-4 [3,4] and FEAL-8 [5] and chosen plaintext attacks that break FEAL-N faster than an exhaustive search for any $N \leqslant 31$ [7] have been published. [7] contains also some bounds on the extension of differential attacks to known plaintext attacks, and the announcement of a non-differential known plaintext attack on FEAL-4, which requires about 100000 plaintext blocks.

In this paper we present known plaintext attacks of FEAL-4 and FEAL-6. These attacks are statistical in nature, and require a limited number of ciphertext blocks and the corresponding plaintext (about 1000 blocks for the attack of FEAL-4 described here, about 200 blocks for an improved attack of FEAL-4, and about 20000 blocks for the attack of FEAL-6). There are no particular constraints on the plaintext.

J. Feigenbaum (Ed.): Advances in Cryptology - CRYPTO '91, LNCS 576, pp. 172-182, 1992.

We are using the following notations :

- If X represents a 32-bit word $(x_{31}, x_{30}, \ldots, x_0)$, X_0 is the byte $(x_{31},x_{30},\ldots,x_{24})$; X_1 is the byte $(x_{23}, x_{22},\ldots,x_{16})$, etc; we also write : $X = (X_0, X_1, X_2, X_3)$;
- If X and Y are two binary strings of equal length, X⊕Y represents the bitwise xor between X and Y;
- If B represents the byte $(b_7,b_6,b_5,b_4,b_3,b_2,b_1,b_0)$, the byte $(b_5,b_4,b_3,b_2,b_1,b_0,b_7,b_6)$ is denoted by ROT2(B); the byte $(b_6,b_5,b_4,b_3,b_2,b_1,b_0,0)$ is denoted by SH1(B);
- If B_1 and B_2 are two bytes, the byte $B_1 + B_2$ represents the sum modulo 256 of the numbers represented by B_1 and B_2, using the usual binary convention (low weight bit right). We also define the ternary operator SBOX : $SBOX(B_1,B_2,\epsilon) = ROT2(B_1+B_2+\epsilon)$ where $\epsilon \in \{0,1\}$.

The FEAL-N algorithm can be divided in two components : the key schedule and the data randomizer.

We do not need here to consider the detail of the key schedule : let us only say that the key schedule transforms the 64-bit secret key into an expanded key composed of the 2N+16 bytes $K_0, K_1, \ldots, K_{2N+15}$.

The data randomization can be split into the three following steps :

The initial step
We start with a 64-bit word (I^0, I^1) as input. Then we compute a new 64-bit word (X^0,X^1) defined by :

$$X^0 = I^0 \oplus (K_{2N},K_{2N+1},K_{2N+2},K_{2N+3})$$
$$X^1 = X^0 \oplus I^1 \oplus (K_{2N+4},K_{2N+5},K_{2N+6},K_{2N+7})$$

The main step
The 64-bit word (X^0,X^1) is taken as the input to an N round Feistel scheme. Rounds are numbered from 0 to N-1. At round i, a new 32-bit word X^{i+2} is produced, given by the equation :

$$X^{i+2} = f_i(X^{i+1}) \oplus X^i$$

The function f_i is defined by :

$$\{0,1\}^{32} \rightarrow \{0,1\}^{32}$$
$$X = (X_0,X_1,X_2,X_3) \mapsto Y = (Y_0,Y_1,Y_2,Y_3)$$

where :

$$Y_1 = SBOX(X_0 \oplus X_1 \oplus K_{2i}, X_2 \oplus X_3 \oplus K_{2i+1}, 1),$$
$$Y_0 = SBOX(X_0, Y_1, 0),$$
$$Y_2 = SBOX(Y_1, X_2 \oplus X_3 \oplus K_{2i+1}, 0),$$
$$Y_3 = SBOX(Y_2, X_3, 1)$$

The function f_i is one to one and depends only on the two expanded key bytes K_{2i} and K_{2i+1}. In the (usual) 64-bit representation of the Feistel scheme, the output of round i is the 64 bit word (X^{i+1}, X^{i+2});

The final step

The 64-bit word (X^N, X^{N+1}) is taken as input to the final step. The 64-bit ciphertext block $(0^0,0^1)$ is defined by :

$$0^0 = X^{N+1} \oplus (K_{2N+8},K_{2N+9},K_{2N+10},K_{2N+11})$$
$$0^1 = X^N \oplus X^{N+1} \oplus (K_{2N+12},K_{2N+13},K_{2N+14},K_{2N+15})$$

2 Principle of the attack

Our attack is a statistical variant of the well known "meet in the middle" method. It is based on two kinds of relations :

(1) It uses some key-independent statistics which involve the plaintext and an intermediate block of the FEAL-N data randomizer (say the block X^{N-1}, which appears as an input to the last round of the Feistel scheme).

(2) In addition, the deciphering algorithm provides a key-dependent relation between this intermediate block and the ciphertext.

An exhaustive search for the value optimizing the agreement between the a priori expected statistics (1) and the statistics deduced from the ciphertext (2) provides the part of the expanded key involved in (2). The knowledge of this part of the expanded key can be generally used for the derivation of an additional part of the expanded key, based on further statistics of the form (1), etc... A full attack consists of the stepwise derivation of the entire expanded key.

The main difficulty of such an attack is (1), i.e. finding key-independent statistics. In order to obtain such statistics, our attack uses extensively the fact that in [0,255] considered as an 8-dimensional vector space over GF(2) the SBOX operator is nearly linear.

3 A linear approximation of the FEAL S-boxes.

We must find a good linear approximation of the S-box operator. Also we must find a good approximation of the two following operations in [0,255] :
- addition : $(B,B') \mapsto (B+B') \bmod 256$
- addition and successor : $(B,B') \mapsto (B+B'+1) \bmod 256$

We are led to study the addition in $\mathbb{N}$.

For $n \in \mathbb{N}^*$, f_n denotes the following boolean function :

$$\{0,1\}^{2n+2} \to \{0,1\}$$
$$(x_n,\ldots,x_0,y_n,\ldots,y_0) \mapsto z_n$$

where $z_{n+1}z_n z_{n-1}\ldots z_0$ is the binary representation of the sum in $\mathbb{N}$ of the two numbers x and y represented by $x_n x_{n-1}\ldots x_0$ and $y_n y_{n-1}\ldots y_0$ respectively.

Proposition

For every $n \in \mathbb{N}^*$ a best linear approximation of f_n is the function $\tilde{f}_n$ defined by :

$$\{0,1\}^{2n+2} \to \{0,1\}$$
$$(x_n,\ldots,x_0,y_n,\ldots,y_0) \mapsto x_n \oplus x_{n-1} \oplus y_n$$

and $$d(f_n , \tilde{f}_n) = \frac{1}{4} \cdot 2^{2n+2} \qquad \text{(i)}$$

Proof

We first state (i). Our proof is basically the same as the one contained in [8], where (i) and similar relations are mentioned. On $\Omega = \{0,1\}^{2n+2}$ equipped with the uniform probability we define the boolean random variable c_k $(1 \le k \le n)$ of a sample $(x_n,\ldots,x_0;y_n,\ldots,y_0)$ as the left carry generated by adding the numbers $x_{k-1}\ldots x_0$ and $y_{k-1}\ldots y_0$. For instance $c_1(1;1)=1$ and $c_1(1;0)=0$. We also define the random variable $c_0=0$.

$$\begin{aligned}
Pr_{(x,y)\in\Omega}\{f_n(x,y)\neq\tilde{f}_n(x,y)\} &= Pr_\Omega\{x_n \oplus y_n \oplus c_n \neq x_n \oplus x_{n-1} \oplus y_n\} \\
&= Pr_\Omega\{x_{n-1}=1 \wedge c_n=0\} + Pr_\Omega\{x_{n-1}=0 \wedge c_n=1\} \\
&= Pr_\Omega\{x_{n-1}=1 \wedge y_{n-1}=0 \wedge c_{n-1}=0\} \\
&\quad + Pr_\Omega\{x_{n-1}=0 \wedge y_{n-1}=1 \wedge c_{n-1}=1\} \\
&= Pr_\Omega\{x_{n-1}=1\}.Pr_\Omega\{y_{n-1}=0\}.Pr_\Omega\{c_{n-1}=0\} \\
&\quad + Pr_\Omega\{x_{n-1}=0\}.Pr_\Omega\{y_{n-1}=1\}.Pr_\Omega\{c_{n-1}=1\}
\end{aligned}$$

(we are using the fact that for $\epsilon, \epsilon', \epsilon'' \in \{0,1\}$ the events $\{x_{n-1}=\epsilon\}$, $\{y_{n-1}=\epsilon'\}$ and $\{c_{n-1}=\epsilon''\}$ are independent)

$$= \frac{1}{4} \cdot (Pr_\Omega\{c_{n-1}=0\} + Pr_\Omega\{c_{n-1}=1\}) = \frac{1}{4}$$

Thus (i) is proved.

The fact that $\tilde{f}_n$ is a best linear approximation of f_n is a consequence of (i). Let L be any affine boolean function on $\{0,1\}^{2n+2}$ other than $\tilde{f}_n$. We have :

$$d(\tilde{f}_n, L) = \frac{1}{2} \cdot 2^{2n+2} \qquad \text{(ii)}$$

From (i) (ii) and the triangular inequality :

$d(\tilde{f}_n, L) \leq d(\tilde{f}_n, f_n) + d(f_n, L)$ we deduce :

$d(f,L) \geq \frac{1}{4} \cdot 2^{2n+2}$ Q.E.D.

Note : of course the function $\hat{f}_n$ defined in replacing x_{n-1} by y_{n-1} in the expression of $\tilde{f}_n$ is also a best linear approximation of f_n.

The above proposition suggests the following linear approximation for the addition :

$$B + B' \simeq B \oplus B' \oplus SH1(B') \quad (a)$$

(for each of the 8 bit positions in a byte, the equality between the bit at the left and the bit at the right of $\simeq$ holds with a probability of at least 0.75). Similarly we are led to the following approximation for the addition and successor operation :

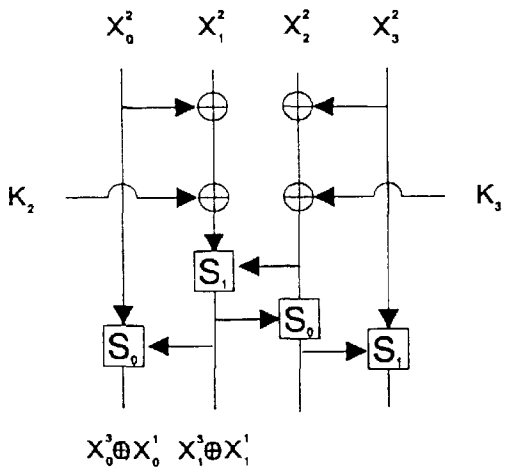
$$B + B' + 1 \simeq B \oplus B' \oplus SH1(B') \oplus 1 \quad (b)$$

Our attack uses a keyless linear approximation of the encryption scheme obtained by omitting the expanded key and by replacing the S-boxes with their linear approximation derived from (a) and (b). If X^n denotes one of the intermediate variables of the encryption scheme, the corresponding value obtained by replacing the encryption scheme with its keyless linear approximation will be denoted by $\tilde{X}^n$.

4 The attack of FEAL-4

4.1 Statistics

The attack of FEAL-4 uses key-independent statistics which involve the intermediate variable X^3 and the plaintext. The following diagram shows the relation between the bytes X_0^3 and X_1^3 and the bytes X_0^2, X_0^1 and X_1^1.



In using the explicit expression of the left S-box S_0 in the above diagram, we obtain the relation :

$$X_0^3 \oplus X_0^1 = ROT2\ (X_0^2 + (X_1^3 \oplus X_1^1\ )).$$

Now by using the linear approximation of the addition of Section 3 we obtain :

$$X_0^3 \oplus X_0^1 \simeq ROT2(X_0^2 \oplus X_1^3 \oplus X_1^1 \oplus SH1(X_1^3 \oplus X_1^1)).$$

We call φ the function defined by :

$$\varphi(X^n) = X_0^n \oplus ROT2(X_1^n \oplus SH1(X_1^n)).$$

We can restate the above relation :

$$\varphi(X^3) \simeq ROT2(X_0^2) \oplus \varphi(X^1) \quad (R)$$

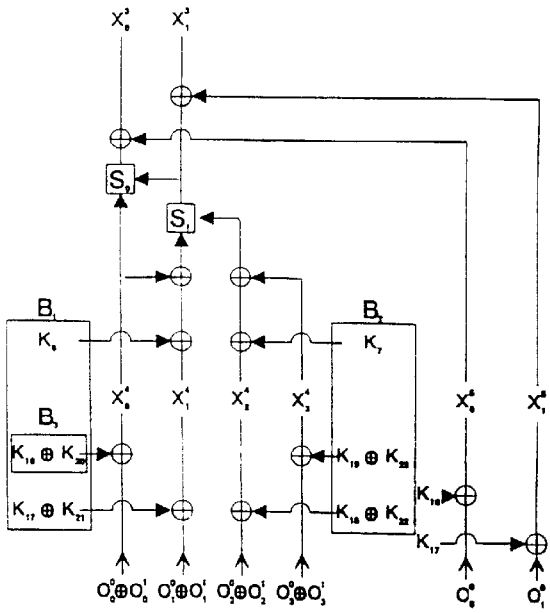
We studied the statistics of the byte $\varphi(X^3) \oplus ROT2(\underset{\sim}{X}_0^2) \oplus \varphi(\underset{\sim}{X}^1)$, where $\underset{\sim}{X}_0^2$ and $\underset{\sim}{X}^1$ are keyless linear approximations of X_0^2 and X^1 obtained as explained in Section 3. These statistics are summarized in Table 1. They are key-independent, i.e. for each bit position the absolute value of the deviation from 0.5 is independent of the key. They are non uniform and express a correlation between a function of X^3 (the term $\varphi(X^3)$) and a function of the plaintext (the two last terms), so they are of the desired form.

bit number	average value
0	0.578
1	0.580
2	0.667
3	0.413
4	0.373
5	0.434
6	0.377
7	0.572

Table 1 : Statistics obtained with 10000 blocks.
Only the absolute value of the deviation from 0.5 is significant.

4.2 Derivation of the expanded key

The relation between the bytes X_0^3 and X_1^3 (which are needed for computing the expression $\varphi(X^3) \oplus ROT2(\underset{\sim}{X}_0^2) \oplus \varphi(\underset{\sim}{X}^1)$) and the 64-bit ciphertext block $(0^0,0^1)$ is illustrated in the diagram hereafter :



This diagram shows that X_0^3 and X_1^3 can be calculated up to the unknown constants K_{16} and K_{17} using only the three unknown combinations B_1, B_2 and B_3. A more careful analysis shows that $\varphi(X^3)$ can be calculated up to an unknown constant byte using only the 7 lowest weight bits of B_1, B_2 and B_3 and the bit $B_1[7] \oplus B_2[7]$, i.e. 22 unknown keybits.

The procedure for testing a value of the 22 unknown bits is the following :

- for each plaintext block we calculate $ROT2(\tilde{X}_0^2) \oplus \varphi(\tilde{X}^1)$. This is done only once;
- for each ciphertext sample we calculate $\varphi(X^3)$ up to a constant, using the assumed value of the 22 keybits;
- we assign to that 22 keybits value a "criterion value" : the sum of the absolute values of the deviation from 0.5 of the average of each bit of the byte $\varphi(X^3) \oplus ROT2(\tilde{X}_0^2) \oplus \varphi(\tilde{X}^1)$ (this byte is calculated up to an unknown constant which has no effect on the criterion value).

We select the value of the 22 keybits for which that criterion value is maximal. In fact if the 22 keybits value is not the correct one, the calculated $\varphi(X^3)$ has no sense and each bit of the studied byte has an average value close to 0.5.

Experiments prove that this test leads to the correct value of the 22 keybits with only 1000 plaintext blocks and the corresponding ciphertext (we obtained some good results even with only 300 blocks). In order to improve

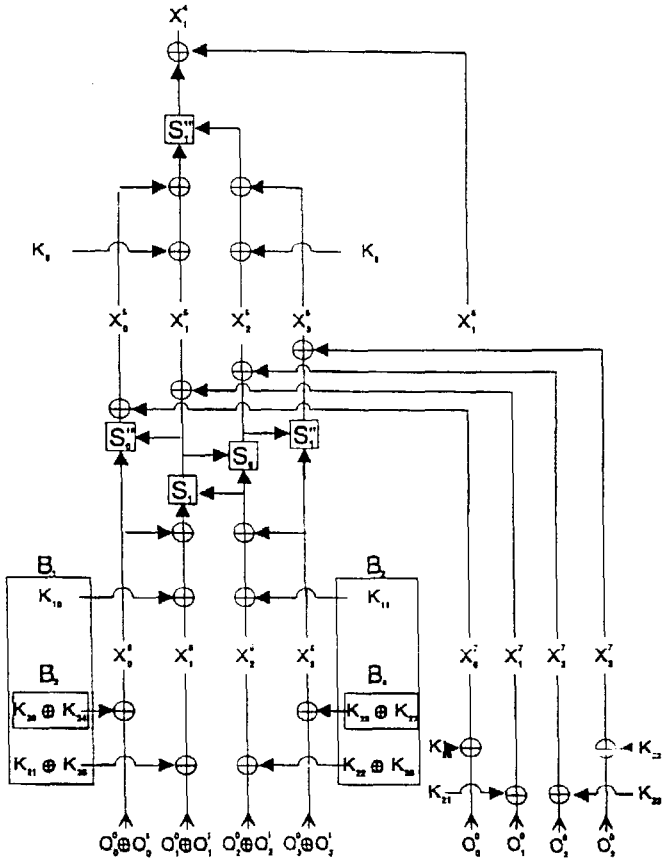
the computation time, we searched appropriately selected parts of the keybits first. With this improved method, only an half of hour computing time of a SUN4 workstationwas needed for deriving the correct 22 keybits from 1000 blocks.

Once the first 22 unknown keybits have been derived, new key-independent statistics must be used for deriving further unknown keybits. This process is quite similar to the derivation of the expanded key in the chosen plaintext attack of [5]. We restricted our experiments to the beginning of the derivation, but there is no substantial difficulty in continuing this derivation.

5 The attack of FEAL-6

5.1 Statistics

We must now find key-independent statistics which involve the intermediate variable X^5 .



In the previous Section we have used the approximate relation :

$$ROT2(X_0^{i+1}) \simeq \varphi(X^{i+2}) \oplus \varphi(X^i) \quad (R)$$

For the attack of FEAL-6 we are using another similar approximate relation. Instead of φ it involves a new function denoted by ψ, which is defined by :

$$\psi(X^i) = X_0^i \oplus X_1^i \oplus ROT2(SH1(X_1^i)) \oplus X_2^i$$

We have the following approximate relation :

$$ROT2(X_1^{i+1}) \simeq \psi(X^i) \oplus \psi(X^{i+2}) \quad (\mathfrak{R})$$

We can prove that the bits number 2 of the left and right bytes of $(\mathfrak{R})$ are strictly equal :

$$ROT2(X_1^{i+1})[2] = \psi(X^i)[2] \oplus \psi(X^{i+2})[2]$$

For the other bit positions, the correlations expressed by $(\mathfrak{R})$ are very low.

From the relations :

$$ROT2(X_1^2) \simeq \psi(X^3) \oplus \psi(X^1) \quad \text{and}$$

$$ROT2(X_1^4) \simeq \psi(X^3) \oplus \psi(X^5)$$

we deduce :

$$ROT2(X_1^2) \oplus \psi(X^1) \simeq ROT2(X_1^4) \oplus \psi(X^5)$$

We studied the statistics of the byte :

$$ROT2(\tilde{X}_1^2) \oplus \psi(\tilde{X}^1) \oplus ROT2(\tilde{\tilde{X}}_1^4) \oplus \psi(X^5)$$

where $\tilde{X}_1^2$ and $\tilde{X}^1$ are keyless approximations of X_1^2 and X^1 calculated from the plaintext as explained in Section 3 and $\tilde{\tilde{X}}_1^4$ is an approximation of X_1^4 which is derived from X^5 and the ciphertext by using a keyless approximation of the decryption scheme. We are using $\tilde{\tilde{X}}_1^4$ instead of X_1^4 (which would have given better statistics) in order to restrict the number of unknown keybytes involved in the calculation of the above expression to only four unknown combinations B_1, B_2, B_3 and B_4 (which are defined in the above diagram).

The obtained statistics are given in Table 2. They are key independent, i.e. for each bit position the absolute value of the deviation from 0.5 is independent of the key. The bit number 2 differs strongly from 0.5.

bit number	average value
0	0.498
1	0.499
2	0.624
3	0.493
4	0.497
5	0.501
6	0.499
7	0.496

Table 2 : Statistics obtained with 100000 blocks
Only the absolute values of the deviations from 0.5 are significant

5.2 Derivation of the expanded key

The attack method is similar to the one used for FEAL-4. We show the very beginning of the attack, i.e. the procedure for the test of a value of B_1 and B_2 (more precisely for the test of the 7 lowest weight bits of B_1 and B_2 and of the bit $B_1[7] \oplus B_2[7]$, i.e. 15 bits) :

- for each plaintext block we calculate $ROT2(X_0^{\sim 2})$ and $\psi(X^{\sim 1})$. This is done once only;
- for each ciphertext block we calculate X_1^5 and X_2^5 up to the unknown constants K_{21} and K_{22}, using B_1 and B_2. We calculate then approximate values of X_0^5 and X_3^5 from X_1^5, X_3^5 and the ciphertext , using a keyless linear approximation of the left and right S-boxes in the first round of the decryption scheme. We also calculate $X_0^{\approx 4}$ from the obtained approximate value of X^5 and the ciphertext, using a keyless linear approximation of the second round of the decryption scheme (all the calculations are represented in the diagram of Section 5.1);
- we assign to B_1 and B_2 a criterion value : the absolute value of the average deviation from 0.5 of the obtained approximate value of the bit $ROT2(X_1^{\sim 2})[2] \oplus \psi(X^{\sim 1})[2] \oplus ROT2(X_1^{\approx 4})[2] \oplus \psi(X^5)[2]$.

We finally select the value of the 15 unknown bits for which the criterion value is maximal.

The experiments of this attack made with 20000 plaintext blocks and the corresponding ciphertext led to the correct value of the 15 keybits. The derivation of these 15 bits took approximately 10 hours computation time on a SUN4 workstation (using a non optimised Pascal program). We did not experiment the whole continuation of the derivation of the expanded key, but there is no substantial difficulty in continuing the derivation.

Partial experiments of a similar attack based on the function φ of Section 4 showed that it should be possible to reduce the number of blocks required to a few thousands, at the expense of increasing the number of unknown keybits in the first step of the attack. This other attack requires an exhaustive search of the four bytes B_1, B_2, B_3 and B_4 with 3000 blocks; its full test was not within the reach of our computer.

6 Improved results on FEAL-4

The above attack on FEAL-6 suggested us an improvement to the FEAL-4 attack described in Section 4. The new attack is entirely based on the approximate relation ($\mathfrak{R}$), as for the FEAL-6 attack. It requires about 200 plaintext blocks. According to our experiments, the derivation of the first 12 unknown keybits takes 20 seconds on a SUN4 workstation. Our estimate of the time required for the entire key derivation is a few minutes.

7 Conclusion

The attacks presented here are an example of the use of correlations with linear functions for the cryptanalysis of blockciphers. They belong, together with the differential attacks [4,5,6,7], to a broader family of statistical attacks. They use approximate relations in the same way as differential cryptanalysis uses caracteristics [6]. The relations (R) and ($\mathfrak{R}$) are similar to eight one-round caracteristics each (one for each bit position), and the statistics used for the attack of FEAL-4 and FEAL-6 are similar to 2-rounds caracteristics and 4-rounds caracteristics respectively. A known plaintext attack of the standard version FEAL-8 would require an efficient 6-rounds approximate relation. We do not know whether such relations can be found.

8 Acknowledgements

The authors wish to thank Marc Girault and Janice Bell for their discerning remarks and valuable suggestions on the early version of this work, and François Allègre for generously cooperating to improve the presentation of the results.

REFERENCES

[1] A. Shimizu, S. Miyaguchi, "Fast Data Encipherment Algorithm FEAL", Advances in Cryptology - Eurocrypt 87, Lecture Notes in Computer Science, Vol 304, Springer Verlag, 1988.

[2] S. Miyaguchi, "News on Feal cipher", talk at the Rump session, Crypto'90, 1990.

[3] Bert Den Boer, "Cryptanalysis of FEAL", Advances in Cryptology - Eurocrypt 88, Lecture Notes in Computer Science, Vol 330, Springer Verlag, 1989.

[4] S. Murphy, "The Cryptanalysis of Feal-4 with 20 Chosen Plaintexts", Journal of Cryptology Vol2 N°3, 1990.

[5] H. Gilbert , G. Chassé, "A Statistical Attack of the FEAL-8 Cryptosystem", Proceedings of Crypto'90.

[6] E. Biham, A. Shamir, "Differential Cryptanalysis of DES-like Cryptosystems", Proceedings of Crypto'90.

[7] E. Biham, A. Shamir, "Differential Cryptanalysis of Feal and N-Hash", extended abstract, Proceedinds of Eurocrypt'91.

[8] W. Meier, O. Staffelbach, "Correlation Properties of Combiners with Memory in Stream Ciphers", to appear in the Journal of Cryptology.