Differential-Linear Cryptanalysis

Susan K. Langford[1] and Martin E. Hellman

Department of Electrical Engineering
Stanford University
Stanford, CA 94035-4055

Abstract. This paper introduces a new chosen text attack on iterated cryptosystems, such as the Data Encryption Standard (DES). The attack is very efficient for 8-round DES,[2] recovering 10 bits of key with 80% probability of success using only 512 chosen plaintexts. The probability of success increases to 95% using 768 chosen plaintexts. More key can be recovered with reduced probability of success. The attack takes less than 10 seconds on a SUN-4 workstation. While comparable in speed to existing attacks, this 8-round attack represents an order of magnitude improvement in the amount of required text.

1 Summary

Iterated cryptosystems are encryption algorithms created by repeating a simple encryption function n times. Each iteration, or round, is a function of the previous round's output and the key. Probably the best known algorithm of this type is the Data Encryption Standard (DES) [6]. Because DES is widely used, it has been the focus of much of the research on the strength of iterated cryptosystems and is the system used as the sole example in this paper.

Three major attacks on DES are exhaustive search [2, 7], Biham-Shamir's differential cryptanalysis [1], and Matsui's linear cryptanalysis [3, 4, 5]. While exhaustive search is still the most practical attack for full 16 round DES, research interest is focused on the latter analytic attacks, in the hope or fear that improvements will render them practical as well. For example, linear cryptanalysis is much faster than exhaustive search, but requires an impractical 2^{43} known plaintexts. In contrast, exhaustive search requires only one known plaintext block or about 1000 bits in a ciphertext only attack. The goal of our work is therefore to reduce the amount of text required in the analytic attacks.

This paper builds on techniques from differential and linear cryptanalysis, creating an eight round attack which recovers 10 bits of key with only 512 chosen plaintexts. While the computation time is comparable to pre-existing attacks, the amount of required text is reduced by an order of magnitude. The best current

[1] This author was supported by NSF grant NCR-9205663

[2] Because FIPS PUB 46 specifies 16 rounds as part of the standard, strictly speaking, we should use the more cumbersome term "DES reduced to 8 rounds." While for ease of exposition we use the simpler "8-round DES" the reader should remember what is intended.

Y.G. Desmedt (Ed.): Advances in Cryptology - CRYPTO '94, LNCS 839, pp. 17-25, 1994.

Biham-Shamir 8-round attack requires over 5,000 chosen plaintexts and Matsui's 8-round attack requires approximately 500,000 known plaintexts. In comparing our attack with these others, it should be remembered that they recover more bits of key and that Matsui's is a more desirable known plaintext attack. They also extend more efficiently to 16 rounds than ours. Our attack should therefore be viewed as providing an interesting new possibility that supplements earlier attacks when the amount of required text is at a premium. Of course, it is our hope that the attack can be extended.

2 Notation

We use FIPSPUB-46's DES numbering so that plaintext, ciphertext, and the bits of the intermediate results (L_n, R_n) are numbered from 1 to 64 reading from left to right. This numbering differs from Matsui's paper, which numbers bits from 0 to 63, reading from right to left. Similarly, in this paper the input to an S-box is taken as $(x_1, x_2, x_3, x_4, x_5, x_6)$ while Matsui uses $(x_5, x_4, x_3, x_2, x_1, x_0)$. We will use Matsui's notation in which $A[i]$ represents the ith bit of A and $A[i, j, ..., k] = A[i] \oplus A[j] \oplus ... \oplus A[k]$.

We will ignore the initial and final permutations, IP and IP^{-1}, since they have no cryptographic significance in a chosen or known text attack. Thus, we refer to (L_0, R_0), the 64 bits after IP as the plaintext and (L_n, R_n), the 64 bits before IP^{-1}, as the ciphertext. This notation differs from both Biham-Shamir and Matsui in that they take (R_n, L_n) as the ciphertext. Our notation simplifies concatenation of k-round and l-round attacks into $(k + l)$ round attacks.

3 Review of Differential and Linear Cryptanalysis

This section is included for completeness. The reader familiar with differential and/or linear cryptanalysis can omit the corresponding subsections.

3.1 Differential Cryptanalysis

The basic idea of differential cryptanalysis is that, while any single plaintext produces a ciphertext that appears random, the same is not true on a differential basis. Two chosen plaintexts, P and P^*, which XOR to a carefully chosen differential plaintext $P' = P \oplus P^*$ can encipher to two ciphertexts C and C^* such that $C' = C \oplus C^*$ takes on a specific value with non-negligible probability. As a trivial example, $P' = 0$ causes $C' = 0$ with probability 1 since $P = P^*$ implies $C = C^*$. More interestingly, Biham and Shamir found that, for 5-round DES, $P' = 405C000004000000_x$ causes $C' = 04000000405C0000_x$ with probability $\frac{1}{10485.76}$. They use this 5-round "characteristic" in an attack on 8-round DES by deciphering portions of the ciphertext (L_8, R_8) to determine when their characteristic has occurred, in which case they are able to derive a number of bits of key. Their attack is efficient because the partial deciphering of (L_8, R_8)

to tell when the characteristic has occurred depends on portions of the key small enough to allow a search. Making use of symmetries, they are able to break 8-round DES with 5,000 chosen plaintexts and 16-round DES with 2^{47} chosen plaintexts.

For the purposes of this paper, it is not necessary for the reader to understand further details of Biham and Shamir's attack. It is sufficient to be familiar with the concept of working differentially. The interested reader is referred to [1] for a complete description of Biham and Shamir's breakthrough in cryptanalysis.

3.2 Linear Cryptanalysis

A second breakthrough, linear cryptanalysis, was recently introduced by Matsui [1, 3, 4, 5]. This approach works with a known plaintext attack, as opposed to a chosen text attack. Linear cryptanalysis finds probabilistic parity relations between selected bits of the plaintext, the ciphertext, and the key. These parity relations derive from parity relations within the S-boxes that differ from the uniform 50-50 distribution and which can then be connected through multiple rounds.

Matsui was able to find useful parity relations for an arbitrary number of rounds of DES. For example, he found that for three round DES,

$$(L_0[3,8,14,25] \oplus R_0[17]) \oplus (R_3[3,8,14,25] \oplus L_3[17]) \oplus (K_1[26] \oplus K_3[26]) = 0, \quad (1)$$

with probability $p = 0.695$. In general, he uses either an $n-1$ round or $n-2$ round parity relation to attack n-round DES. He can use the 3-round relation (1) in a 4-round attack, by noting that although $K_1[26] \oplus K_3[26]$ is not known, its effect is to cause the reduced equation

$$(L_0[3,8,14,25] \oplus R_0[17]) \oplus (R_3[3,8,14,25] \oplus L_3[17]) = 0. \quad (2)$$

to be satisfied eigher with probability 0.695 or 0.305, both of which are different from 0.5, the value expected with random data. Matsui can decipher backwards through round 4, as shown in figure 1, to calculate the value of the necessary bits. This decipherment depends on six bits of key, requiring a search over only 64 values.

While we only describe Matsui's 4-round attack, the general idea follows in a straightforward manner from this example. Using the best 6-round parity relation, which holds with probability $0.5 - 1.95 * 2^{-9}$, Matsui was able to break 8-round DES with approximately 500,000 known plaintexts in less than a minute on a workstation. Similarly, using the best 14-round parity relation, which holds with probability $0.5 - 1.19 * 2^{-21}$, he was able to break 16-round DES with 2^{43} known plaintexts in 50 days using 12 HP9735 workstations. The interested reader is referred to [3, 4, 5] for more information.

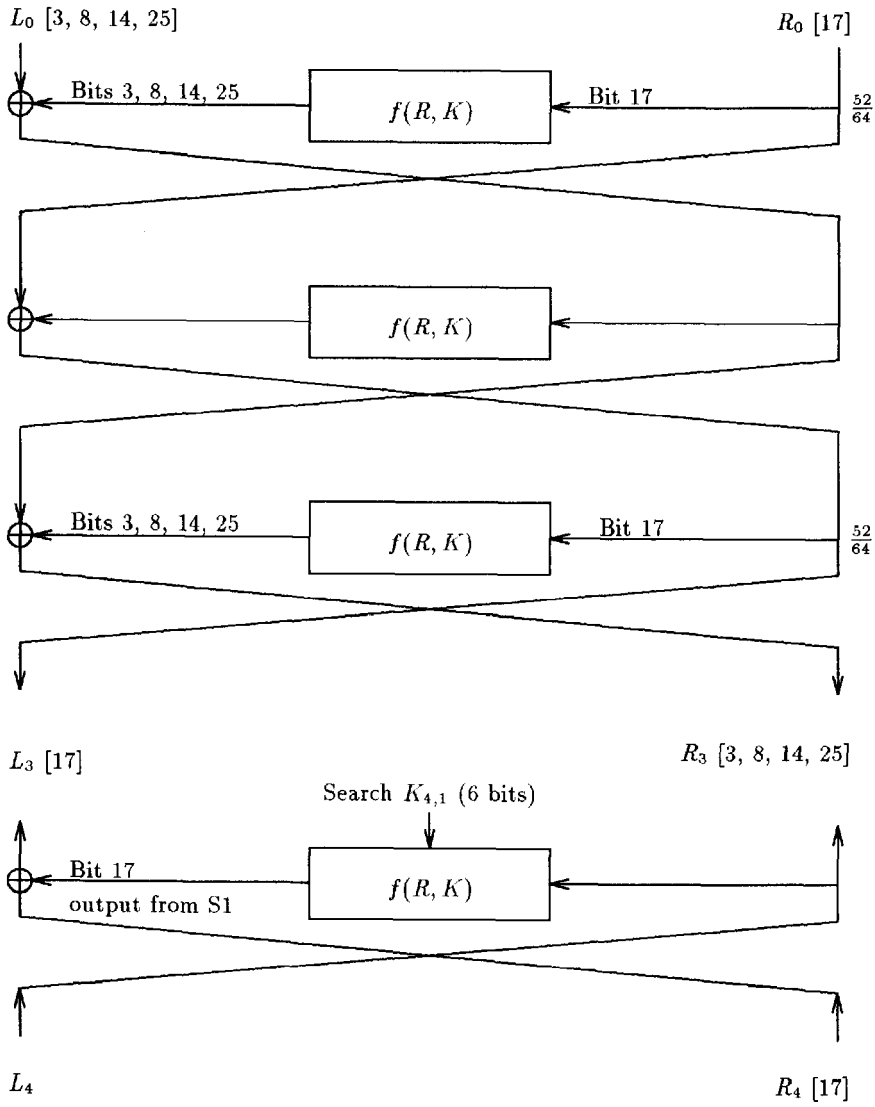

Fig. 1. Matsui's 4-round attack on DES

4 Differential-Linear Cryptanalysis

By the technique to be described in section 5, we can complement bits 2 and/or 3 of L_1 and keep the other 62 bits of (L_1, R_1) unchanged. The key observation in our attack is that this behavior in (L_1, R_1) leaves many bits of (L_4, R_4) unchanged. In particular, the input bits to Matsui's best 3-round

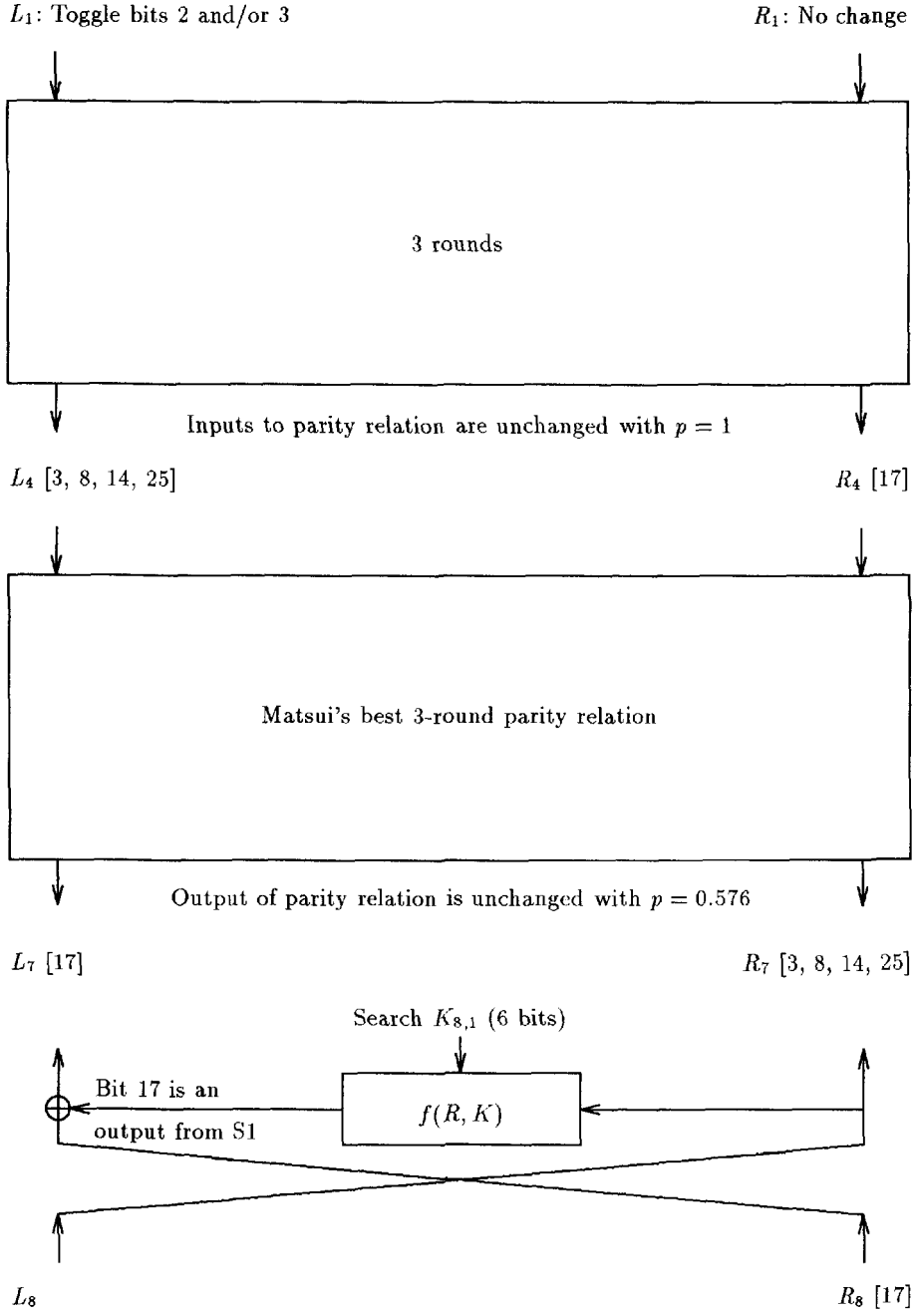

Fig. 2. Differential-Linear attack on 8-round DES

parity relation (bits 3, 8, 14, and 25 of L_4 and bit 17 of R_4) *never* change, [3] so that the parity of the output bits (bits 3, 8, 14, and 25 of R_7 and bit 17 of L_7) is unchanged with probability $r = p^2 + q^2 = 0.576$ where $p = 0.695$ is the probability of Matsui's parity relation holding once, and $q = 1 - p$. The probability is $p^2 + q^2$ because Matsui's parity relation must hold, or fail, twice – once for the reference plaintext and once for the plaintext which toggles only bits 2 and/or 3 in round 1. Unlike in ethics, two wrongs do make a right in mod-2 arithmetic. This behavior is depicted in the upper two blocks of figure 2, with the upper block being differential in nature and the second block being primarily linear.

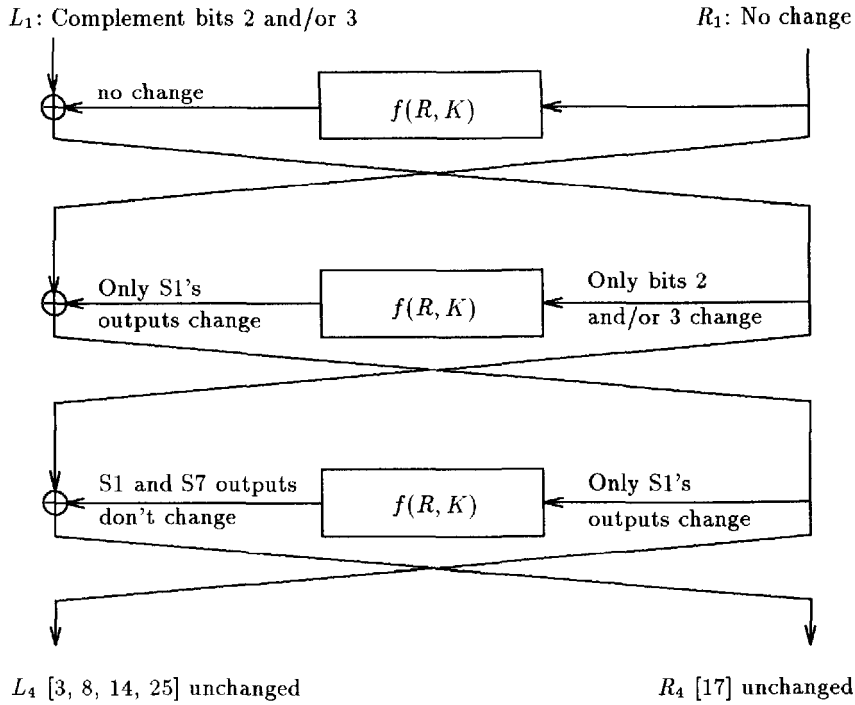

Fig. 3. Differential characteristic

Figure 3 shows why the differential characteristic holds going from (L_1, R_1) to (L_4, R_4). Because $R'_1 = 0$ differentially, the output of $f(R, K)$ in round 2 must also be 0 differentially. Thus, $R'_2 = L'_1$ and R'_2 has only bits 2 and/or 3 toggled. These two bits only affect the input of S1, so only the outputs of S1

[3] The same applies to toggling bits 10 and/or 11, but for simplicity we will not deal with that here.

can change in round 3. Since $R'_3 = L'_4$ and bits 3, 8, 14, and 25 are not outputs of S1, these bits will be unchanged in L'_4. Further, because of the E-expansion, the 4 outputs of S1 affect the inputs of 6 S-boxes in round 4. Two S-boxes will therefore be unchanged, namely S1 and S7. Bit 17 of R'_4 is the output of S1, so it will remain unchanged.

As figure 2 shows, the parity invariance to be observed occurs in round 7 with probability 0.576. Following Matsui, we decipher the two ciphertexts (L_8, R_8) and (L^*_8, R^*_8) backward through one round to get the output bits of the parity relation: bits 3, 8, 14, and 25 of R_7, and bit 17 of L_7. Bits 3, 8, 14, and 25 of R_7 are known because $R_7 = L_8$, the left half of the ciphertext. Only bit 17 of L_7 must be computed. This computation involves only S1 in round 8, so we can test the 6-bit subkey $K_{8,1}$. When the correct value of $K_{8,1}$ is used, we expect to observe parity invariance 57.6% of the time; when an incorrect value is used, the produced data is more random and we expect to observe parity invariance closer to 50% of the time.

Based on Matsui's rule of thumb that approximately $8/(r-0.5)^2$ observations are needed when r is the probability of observing a parity relation, one would expect our attack to require about 1,400 pairs of chosen plaintexts. While one would expect this number must be increased to create the desired toggling in round 1 as opposed to round 0 (the plaintext), the next section develops an approach that obtains the desired behavior while reducing the required amount of text.

5 Structures

Our attack requires plaintext pairs which toggle bits 2 and/or 3 in round 1. We produce this behavior with structures similar to those used by Biham and Shamir. Choose any reference plaintext and let $P(0)$ through $P(64)$ be the 64 plaintexts obtained by varying bits 9, 17, 23, and 31 of L_0 and bits 2 and 3 of R_0. These bits of L_0 correspond to the four outputs of S1, and bits 2 and 3 of R_0 become bits 2 and 3 of L_1, the bits to be toggled.

Bits 2 and 3 are the middle input bits to S1. Since these bits are the only bits that can change in R_0, only the outputs of S1 can change in round 1, as shown in figure 4. Because we included all 16 possibilities for these bits in the structure, if we knew the 6-bit subkey $K_{1,1}$, for each of the 64 $P(i)$'s we could choose three other $P(j)$'s which had the desired toggling in round 1. One $P(j)$ would toggle bit 2, one would toggle bit 3, and one would toggle bits 2 and 3. The 64 chosen plaintexts might therefore seem to produce $64 * 3 = 192$ differential pairs (observations), but only half of these, or 96 pairs, are distinct since (i, j) and (j, i) are the same pair.

All 96 pairs in a structure could be used to help determine $K_{8,1}$ if $K_{1,1}$ were known. Since we do not know $K_{1,1}$, we search over all values of $(K_{1,1}, K_{8,1})$. Two of the bits of $K_{1,1}$ are also part of $K_{8,1}$, so there are only 10 bits and 1024 possible subkeys to search over. Since each structure of 64 plaintexts produces 96 differential pairs and 1400 pairs are required by Matsui's rule of thumb, our

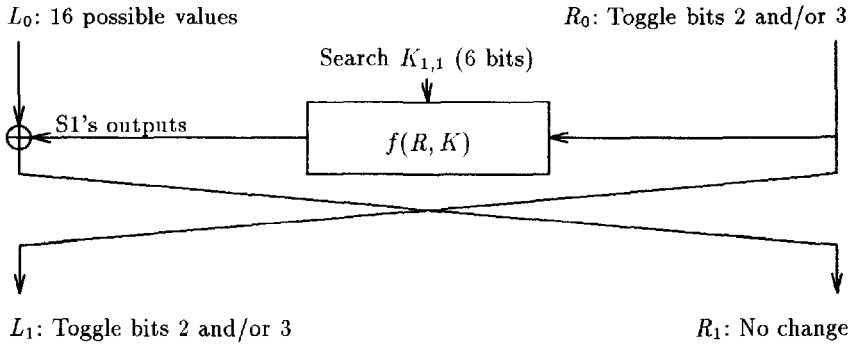

Fig. 4. First round of 8-round differential-linear attack

attack should work with approximately $64/96 * 1400 = 900$ chosen plaintexts. Our experiments find good agreement: 512 chosen plaintexts produce an 80% success rate and 768 chosen plaintexts have a 95% success rate in determining the 10 bits of key in $(K_{1,1}, K_{8,1})$. These two attacks use 8 and 12 structures of 64 chosen plaintexts respectively.

An even higher success rate can be obtained, with no increase in the number of required plaintexts, if we use ideas related to "list decoding" of error correcting codes. The most likely $(K_{1,1}, K_{8,1})$ is tried first in the semi-exhaustive search over the 46 remaining key bits and, if it does not work (which happens one time in five for the 512 chosen plaintexts attack), the next most likely value is tried, etc. With a list of two, this method increases the average computation by only 20%, while increasing the probability of success to 87%. A list of size 8 increases the probability of success to 95%.

6 Additional Bits of Key

While a conservative definition of security would regard a cryptosystem as broken when even one bit of key can be recovered, the 10 bits recovered in our attack leaves a large semi-exhaustive search over 2^{46} keys. However, we can recover additional bits of key using other, lower probability 3-round parity relations in rounds 5 to 7, in place of Matsui's optimal 3-round parity relation. For example, bits 5, 15, 21, 27, and 63 of round 4 also remain unchanged when bits 2 and/or 3 of round 1 are toggled. Therefore, we can use the relation

$$(L_4[5, 15, 21, 27] \oplus R_4[31]) \oplus (R_7[5, 15, 21, 27] \oplus L_7[29]) = 0. \qquad (3)$$

Differentially, equation (3) holds with probability 0.527, instead of 0.576 for Matsui's optimal relation. Use of this parity relation requires searching over $(K_{1,1}, K_{8,6})$ instead of $(K_{1,1}, K_{8,1})$, thereby recovering the six additional bits of key $K_{8,6}$.

The lower probability of the second relation increases the probability of error. The 768 chosen plaintexts which had a 95% success rate on the first ten bits of key, have an 85% success rate for all sixteen bits of key. The idea of list decoding, mentioned above, can be applied here with even greater success.

7 Work in Progress

The above analysis treats each differential pair as if it were independent of all other pairs. Since each structure of 64 chosen plaintexts yields 96 differential pairs, this assumption is clearly not true. However, the close agreement between the predicted and experimental results shows that the effect of dependence is small for the eight round attack.

We have begun a more precise analysis based on the fact that in each structure of 64 chosen plaintexts, there are 16 sets of four plaintexts. These four plaintexts differ only in bits 2 and/or 3 of (L_1, R_1), and therefore have the same input to Matsui's parity relation. Starting from the probability of the parity relation, we can calculate the probability that each set of four texts will have a particular set of four output parity bits. When we ran our eight round attack using the results from this more complex analysis, we obtained only a one percent improvement in the probability of success. Although the more complex analysis makes only a small difference in the eight round attack, it might be more useful in attacks on a larger number of rounds. Such work is in progress.

References

1. E. Biham and A. Shamir, *Differential Cryptanalysis of the Data Encryption Standard*, Berlin: Springer-Verlag, 1993.
2. W. Diffie and M. Hellman, "Exhaustive Cryptanalysis of the NBS Data Encryption Standard," *Computer*, vol.10, no. 6, June 1977, pp. 74–84.
3. M. Matsui, "Linear Cryptanalysis Method for DES Cipher," *Advances in Cryptology-EUROCRYPT '93 Proceedings*, Berlin: Springer-Verlag, 1994, to appear.
4. M. Matsui, "Linear Cryptanalysis of DES Cipher (I)," *Journal of Cryptology*, to appear.
5. M. Matsui, "The first experimental cryptanalysis of the Data Encryption Standard," *Advances in Cryptology–Crypto '94 Proceedings*, Springer-Verlag, to appear.
6. National Bureau of Standards, *Data Encryption Standard*, U.S. Department of Commerce, FIPS pub. 46, January 1977.
7. M. Wiener, "Efficient DES Key Search," *Advances in Cryptology Crypto '93 Proceedings*, Springer-Verlag, to appear.