The Best Differential Characteristics and Subtleties of the Biham-Shamir Attacks on DES

Nicolas T. Courtois

Axalto Cryptographic Research & Advanced Security,
36-38 rue de la Princesse, BP 45, F-78430 Louveciennes Cedex, France,
courtois@minrank.org,

Abstract. In about every book about cryptography, we learn that the plaintext complexity of differential cryptanalysis on DES is 2^{47}, as reported by Biham and Shamir in [2]. Yet few people realise that in a typical setting this estimation is not exact and too optimistic.
In this note we show that the two "best" differentials for DES used by Biham and Shamir [1, 2], are not the best differentials that exist in DES. For approximations over many rounds such as used in the Biham-Shamir attack from [2], the best characteristic is in fact a third, different differential already given by Knudsen in [17]).
A more detailed analysis shows that on average the best differential attack on DES remains the Biham-Shamir attack from [2], because it can exploit two differentials at the same time and their propagation probabilities are related. However for a typical fixed DES key, the attack requires on average about $2^{48.34}$ chosen plaintexts and not 2^{47} as initially claimed. In addition, if the key is changing frequently during the attack, then in fact Biham and Shamir initial figure of 2^{47} is correct.
We were surprised to find out that (with differential cryptanalysis) it is easier to break DES with a changing key, than for one fixed key.

Key Words: Block ciphers, Feistel schemes, DES, S-box design, differential cryptanalysis.

1 Introduction

In a typical cryptographic chosen-plaintext attack we assume that we have an access to an encryption oracle in which one fixed key is embedded. This key is assumed to be typical (behave as an average key).

Very few cryptographic attacks can tolerate a changing key. It happens to be true for the Biham-Shamir attack on DES, but we consider this case only in Section 4 and otherwise assume a fixed key.

The Biham-Shamir attacks on DES works as follows: we query the encryption oracle with known plaintext that are grouped in structures each containing $2 \cdot 2^{12}$ plaintexts (or twice as much with the "quartet" version).

Then, for each structure separately, by looking at the ciphertexts, we spot pairs having 20 bits that are identical, and each of them is a candidate for the so called "right" pair of plaintexts that satisfy a set of differential conditions throughout the whole encryption process. Then from such a pair we try to recover the key, and otherwise repeat the attack.

2 What is the Best Differential Characteristic for DES ?

On the contrary of most people think, the answer to this question is not well known.

In this paper, to indicate differences we use the usual hexadecimal notations of Shamir, Biham, and Knudsen. For completeness we will also indicate the numbers of active bits, with bit numbering compatible with the FIPS standard [14] (which differs from the bit numbering of Matsui [18]).

In [17] Knudsen's carries a very detailed study of differential characteristics. The most interesting characteristics are two-round invariants, if the difference of plaintexts is of type $(0, \psi)$ at some round, the difference is still $(0, \psi)$ two rounds earlier, with some probability that should be as high as possible. In general this probability is fixed for one fixed DES key, and may be different for different DES keys. In Table 2, page 505 in Knudsen paper [17] the first three lines give the following three differences:

1. Let $\psi = \texttt{19600000}_{\texttt{x}} = \{\texttt{4},\texttt{5},\texttt{8},\texttt{10},\texttt{11}\}$ be the main difference used in all Biham-Shamir attacks on DES [1, 2], that appears in the first line of this table.
2. Let $\psi' = \texttt{1b600000}_{\texttt{x}} = \{\texttt{4},\texttt{5},\texttt{8},\texttt{9},\texttt{10},\texttt{11}\}$ be another difference used in [1, 2], and the second line in the Table 2 of [17]. (It is called $\psi^\dagger$ in [2] and it is also called Γ in [17].) The two differentials $(0, \psi)$ and $(0, \psi')$ propagate over two rounds with probability being either 1/146 or 1/585, depending on the key. The average of 1/146 and 1/585 is about $1/234 \approx 2^{-7.87}$. Our simulations show that Biham-Shamir ψ (as well as ψ') do indeed propagate with probability $2^{-7.87}$ for a random (changing) key.
3. Finally, let $\psi'' = \texttt{00196000}_{\texttt{x}} = \{\texttt{12},\texttt{13},\texttt{16},\texttt{18},\texttt{19}\}$ be another difference, that appears in the third line in the Table 2 of [17]. With $(0, \psi'')$ we have the propagation probability of exactly $1/256 = 2^{-8}$ for any key.

We have $1/234 \approx 2^{-7.87} > 1/256 = 2^{-8}$ and therefore it appears that the third differential $(0, \psi'')$ is worse than any of the first two: $(0, \psi)$ or $(0, \psi')$. It is indeed so for two rounds. But **not when the number of rounds grows**.

2.1 The Tricky Point

Though the average of 1/146 and 1/585 is about $1/234 = 2^{-7.87} > 2^8 = 1/256$, their geometric mean is about $1/292.25 \approx 2^{-8.19} < 1/256 = 2^{-8}$. Since the probabilities of differential characteristics are multiplied, it is the geometric mean that is a good way of estimating the "average" effect of these probabilities: for $2k$ rounds and large k, we expect that each of these two characteristics will hold with probability $(1/292.25)^k \approx 2^{-8.19 \cdot k}$. this should be compared to $(0, \psi'')$ that propagates for $2k$ rounds with probability 2^{-8k}. Clearly, ψ'' is better and Knudsen, (and also Biham and Shamir in [1]) did **not** use the best differential characteristics of DES.

3 What is the Best Differential Attack on DES ?

More detailed analysis needs to be done if one wants to determine which will be the best differential attack on DES. The difference is quite small and the advantage of $(0, \psi'')$ could be compensated by the simultaneous usage of $(0, \psi)$ or $(0, \psi')$ as in Biham and Shamir [1, 2]. Unfortunately only ψ and ψ' can be combined in a single "structure" on plaintexts specified by Biham and Shamir [2], as have the same set of active S-boxes S1-S3, while for ψ'' the active S-boxes are S3-S5.

3.1 First Estimation - Typical Complexity

The simplified analysis is as follows: in Biham-Shamir attack the differential must hold for $2k$ rounds with $k = 6$. The expected average effect of a combination of an equal number of 1/146 and 1/585 is about $2^{-8.19k} = 2^{49.15}$. We have an additional factor of two: in one "structure" of Biham-Shamir [2] we have $2 \cdot 2^{12}$ plaintexts, while it allows to generate 2^{12} pairs with the prescribed difference after 2 rounds. Thus, an attack with ψ requires typically $2^{50.15}$ chosen plaintexts. The same holds for ψ', and if we use the two simultaneously (the quartet method from [1, 2]) we can divide the it by two and get back to $2^{49.15}$ chosen plaintexts. This attack should work with probability bigger than 0.5 (see below) over the keys. For some keys it will fail, because the less good case with 1/585 will be more frequent than 1/146.

In comparison, with ψ'' we get a simpler attack without "quartets", without the probabilities depending on the key and thus working with success probability very close to 1, that requires only exactly $2 \cdot 2^{8k} = 2^{49}$ chosen plaintexts.

3.2 More Detailed Analysis

If in the attack we use only one of the two differences ψ or ψ', we expect that fraction of the keys for which the attack will succeed will be close to the ratio of keys for which we have at least 3 cases with 1/146. This should be about:

$$\frac{\sum_{i=0}^{3} \binom{6}{i}}{2^6} \approx 0.66$$

This formula assumes that there are no dependencies due to the DES key schedule. Then with probability 0.66 our attack will require about $2^{49.15}$ plaintexts (or less).

Things are less simple when we use the two characteristics ψ and ψ' together. In Section 6.5. of [1] and in Section 5.1. of [17] the DES keys for which the approximation ψ is "worse" and gives the probability 1/585, are exactly the keys for which for ψ' we get the better case 1/146. Thus, following the analysis of Section 5.1. of Knudsen [17], we are always assured to get one characteristic with a success probability of at least $(\frac{1}{145 \cdot 585})^3 \approx 2^{49.15}$.

In particular for a fraction of

$$\frac{\binom{6}{3}}{2^6} \approx 0.31$$

of all keys we will have $2^{49.15}$ for both ψ and ψ' and we can gain a factor of 2 with "quartet" method in the attack of [2]. Thus we need $2^{49.15}$ plaintexts in the attack.

For the other 0.69 keys, the plaintext complexity of the attack will be at most $(\frac{1}{145})^4(\frac{1}{585})^2 \approx 2^{47.15}$ but only one half of the differences that are used in the "quartet" method are useful. We do not gain a factor of 2 with the "quartet" method (but have to use it anyway to do the same attack for every key). Thus, here we will need $2^{48.15}$ plaintexts (or less) in the attack.

We neglected the case when 1/146 appears only once or not all, (see Table 2 in [17]) but it is clear that the weighted average number of plaintexts required in both attacks is lower than 2^{49}. The exact average plaintext complexity can be computed as follows:

$$\mathbf{2}\cdot\frac{\binom{6}{0}+\binom{6}{6}}{2^6}\cdot 145^6\cdot 585^0+\mathbf{2}\cdot\frac{\binom{6}{1}+\binom{6}{5}}{2^6}\cdot 145^5\cdot 585^1+$$
$$+\mathbf{2}\cdot\frac{\binom{6}{2}+\binom{6}{4}}{2^6}\cdot 145^4\cdot 585^2+\frac{\binom{6}{3}}{2^6}\cdot 145^3\cdot 585^3+\approx 2^{48.34}$$

4 Special Case of Attacks with Changing Key

There is another interesting and subtle point as suggested by Eli Biham [private communication]. The Biham-Shamir attack on DES from [2] is designed to work not only when the attacker has access to an encryption oracle that contains one single fixed DES key, which yields the results above. It is also possible to attack DES when the key changes with time, and we only assume that for each structure containing $2\cdot 2^{12}$ plaintexts the key is the same, and that it changes for the next structure[1]. In such an attack, it is possible to see that the arithmetic averaging such as originally used by Biham and Shamir in [2] is again the correct method.

The argument is as follows. For each structure after two rounds, the number of pairs that have the "good" difference (for example $(0,\psi)$) is 2^{12}. After 4 rounds, for one fixed key, the expected number of pairs with difference $(0,\psi)$ is

$$2^{12}\cdot\left(\frac{1/146+1/585}{2}\right)$$

After 14 rounds, the expected number of pairs with difference $(0,\psi)$, averaging over all keys, will be

$$2^{12}\cdot\left(\frac{1/146+1/585}{2}\right)^6\approx 2^{-35.2}.$$

We need on average $2^{35.2}$ structures with changing keys, to succeed. This will be about $2^{48.2}$ chosen plaintexts. Then with the "quartet" method we get $2^{47.2}$ chosen plaintexts.

[1] It is interesting to note that in such attack scenarios Linear Cryptanalysis does not work and Differential Cryptanalysis will be the best attack on DES

5 Conclusion

The best differential characteristics in DES is not based on one of the two differences $\psi = \texttt{19600000}_\texttt{x}$ and $\psi' = \texttt{1b600000}_\texttt{x}$ used by Shamir and Biham but a third differential using $\psi'' = \texttt{00196000}_\texttt{x}$ already found by Knudsen.

In a typical chosen plaintext attack on DES in which the key is fixed, we have the following. With probability about 0.31 over the keys, the basic version of Shamir-Biham attack from [2] using only ψ'' will require less plaintexts than the full version exploiting both ψ and ψ' described in [2]. However, the attack with two differences exactly as proposed by Biham and Shamir will be better. This is because the two characteristics ψ and ψ' can be combined in a single attack and their propagation probabilities for different keys are related in a useful way, while ψ'' can only be used alone.

However, if the DES key is changing for each structure that is used in the attack, then we need again $2^{47.2}$ chosen plaintexts, as initially expected. We discovered that, due to subtleties in averaging, for differential chosen plaintext attacks, it is easier to break DES with changing key than for one fixed key.

Acknowledgements: We thank Eli Biham for his kind help and valuable comments.

References

1. Eli Biham, Adi Shamir, *Differential Cryptanalysis of DES-like Cryptosystems,* Journal of Cryptology, vol. 4, pp. 3-72, IACR, 1991. (An extended abstract appears in Crypto'90).
2. Eli Biham, Adi Shamir, *Differential cryptanalysis of the full 16-round DES,* In Crypto'92, pp. 487-496, LNCS 740, Springer-Verlag, 1992.
3. Anne Tardy-Corfdir, Henri Gilbert: *A Known Plaintext Attack of FEAL-4 and FEAL-6,* Crypto'91, LNCS 576, Springer, pp. 172-181, 1992.
4. Nicolas Courtois: *Feistel Schemes and Bi-Linear Cryptanalysis,* in Crypto 2004, LNCS 3152, pp. 23-40, Springer, 2004.
5. Don Coppersmith, *The Data Encryption Standard (DES) and its strength against attacks,* Technical Report RC 18613, IBM T.J. Watson Center, December 1992.
6. Don Coppersmith, *The Data Encryption Standard (DES) and its strength against attacks,* IBM Journal of Research and Development, Vol. 38, n. 3, pp. 243-250, May 1994.
7. Don Coppersmith, *The development of DES, Invited Talk, Crypto'2000, August 2000.*
8. *D.W. Davies, Some Regular Properties of the Data Encryption Standard,* Crypto'82, pp. 89-96, Plenum Press, New-York, 1982.
9. D. Davies and S. Murphy, *Pairs and Triplets of DES S-Boxes,* Journal of Cryptology, vol. 8, Nb. 1, pp. 1-25, 1995.
10. H. Feistel, W.A. Notz, J.L. Smith, *Cryptographic Techniques for Machine to Machine Data Communications,* RC 3663 (#16560), Dec. 27, 1971, Communications, IBM T.J.Watson Research. Popular version published in [11].
11. Horst Feistel: *Cryptography and computer privacy;* Scientific American, vol. 228, No. 5, pp. 15-23, May 1973.
12. H. Feistel, W.A. Notz, and J.L. Smith: *Some cryptographic techniques for machine-to-machine data communications;* Proc. IEEE, Vol. 63, No. 11, 1975, pp. 1545-1554.
13. *Data Encryption Standard*, Federal Information Processing Standards Publication (FIPS PUB) 46, National Bureau of Standards, Washington, DC (1977).
14. *Data Encryption Standard (DES)*, Federal Information Processing Standards Publication, FIPS PUB 46-3, National Bureau of Standards, Gaithersburg, MD (1999).
15. M.E. Hellman, R. Merkle, R. Schroppel, L. Washington, W. Diffie, S. Pohlig, and P. Schweitzer: *Results of an initial attempt to cryptanalyze the NBS Data Encryption Standard,* Technical report, Stanford University, U.S.A., September 1976.
   Known also as "Lexar Report", Lexar Corporation, Unpublished Report, 11611 San Vicente Blvd., Los Angeles, 1976.
16. Lars R. Knudsen, John Erik Mathiassen: *A Chosen-Plaintext Linear Attack on DES.* FSE 2000, LNCS 1978, Springer, pp. 262-272, 2001.
17. Lars R. Knudsen, *Iterative characteristics of DES and s^2-DES.* In Crypto'92. Springer Verlag, LNCS 746, pp. 497-511, Berlin Heidelberg 1993.
18. M. Matsui: *Linear Cryptanalysis Method for DES Cipher,* Eurocrypt'93, LNCS 765, Springer, pp. 386-397, 1993.
19. M.Matsui: *The First Experimental Cryptanalysis of the Data Encryption Standard,* In Crypto 94, LNCS 839, Springer, pp. 1-11, 1994.
20. Alfred J. Menezes, Paul C. van Oorschot, Scott A. Vanstone: *Handbook of Applied Cryptography*; CRC Press, 1996.