How to Launch A Birthday Attack Against DES

Zhengjun Cao

Computer Sciences Department, Université Libre de Bruxelles, Belgium. caoamss@gmail.com

Abstract We show how to launch a birthday attack against DES. It requires about 2^{16} ciphertexts of the same R_{16}, encrypted by the same key K. We conjecture it has a computational complexity of 2^{48}.

1 Introduction

The Data Encryption Standard (DES) is a cipher selected as an official Federal Information Processing Standard for the United States in 1976 and which has subsequently enjoyed widespread use internationally [1]. DES consequently came under intense academic scrutiny which motivated the modern understanding of block ciphers and their cryptanalysis [3]. There are some analytical results which demonstrate theoretical weaknesses in the cipher, although they are infeasible to mount in practice. In recent years, the cipher has been superseded by the Advanced Encryption Standard (AES).

Although more information has been published on the cryptanalysis of DES than any other block cipher, the most practical attack to date is still a brute force approach. Various minor cryptanalytic properties are known, and three theoretical attacks are possible which, while having a theoretical complexity less than a brute force attack, require an unrealistic amount of known or chosen plaintext to carry out, and are not a concern in practice. We refer to [4-13] for more details.

A birthday attack is a type of cryptographic attack [2]. Specifically, given a function f, the goal of the attack is to find two inputs x_1, x_2 such that $f(x_1) = f(x_2)$. Such a pair x_1, x_2 is called a collision. The method used to find a collision is to simply evaluate the function f for different input values that may be chosen randomly or pseudorandomly until the same result is found more than once. Because of the birthday paradox, this method can be rather efficient.

Up to the present, regretfully, nobody shows that how to apply a birthday attack to DES. In this paper, we present such an attack against DES. It requires about 2^{16} ciphertexts of the same R_{16}, encrypted by the same key K. Each ciphertext is of the same length 64-bit, namely the length of underlying block. We conjecture it has a computational complexity of 2^{48}.

2 Preliminary

DES processes plaintext blocks of $n = 64$ bits, producing 64-bit ciphertext blocks. The effective size of the secret key is $K = 56$ bits; more precisely, the input key K is specified as a 64-bit key, 8 bits of which (bits $8, 16, \cdots, 64$) may be used as parity bits. The 2^{56} keys implement (at most) 2^{56} of the $2^{64}!$ possible bijections on 64-bit blocks. We refer to the following figures for the DES inner function f, computation path and S-boxes.

Function f operates on two blocks of data: R_{n-1} and K_n. It produces 32-bit long block of data. Process of calculating f function consists of 4 steps:

1. E permutation
2. XOR with a subkey
3. S box transformation
4. P permutation

Each of the eight S-boxes replaces its six input bits with four output bits according to a non-linear transformation, provided in the form of a lookup table. The idea of transformation is straightforward: the first and the last bit of the first group of six bits form a binary number in the decimal range 0 to 3. This is the number of a row in the S1 table. The middle four bits of the group of six bits form a binary number in the decimal range 0 to 15. This is the number of a column in the S1 table. Those two coordinates indicates a decimal number, which as a 4-bit long binary number is the output. We repeat this operation for each of eight groups of six bits and as a result we obtain eight groups of 4 bits. The S-boxes provide the core of the security of DES. Without them, the cipher would be linear and trivially breakable.

S[1]-box	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7
1	0	15	7	4	14	2	13	1	10	6	12	11	9	5	3	8
2	4	1	14	8	13	6	2	11	15	12	9	7	3	10	5	0
3	15	12	8	2	4	9	1	7	5	11	3	14	10	0	6	13
⋮																
S[8]-box	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	13	2	8	4	6	15	11	1	10	9	3	14	5	0	12	7
1	1	15	13	8	10	3	7	4	12	5	6	11	0	14	9	2
2	7	11	4	1	9	12	14	2	0	6	10	13	15	3	5	8
3	2	1	14	7	4	10	8	13	15	12	9	0	3	5	6	11

3 Basic idea

We first point out that it's easy to compute the R_{16} and L_{16} for a known ciphertext c (of 64-bit length).
By the last round, we have

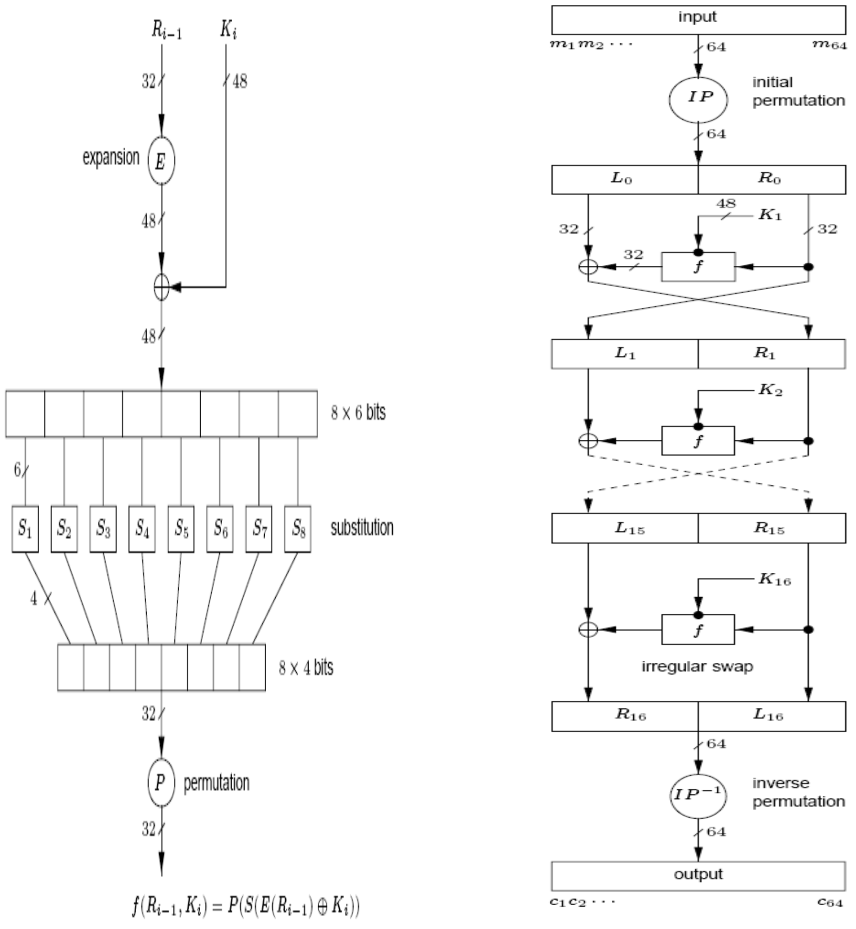
$$R_{16} = L_{15} \oplus f(R_{15}, K_{16}), \quad L_{16} = R_{15}$$

Hence

$$f(L_{16}, K_{16}) = R_{16} \oplus L_{15}$$

Note that both L_{15} and K_{16} are not accessible for an adversary.

DES inner function *f* and computation path



Assumption-1: Suppose that there is a pair of ciphertexts (c, c') generated by the same key K_{16} and satisfying

$$R'_{16} = R_{16},\ L'_{16} \neq L_{16},\ L'_{15} = L_{15}$$

Hence

$$f(L'_{16}, K_{16}) = f(L_{16}, K_{16}) \tag{1}$$

Denote $E(L_{16})$ by EL_{16} where E is the expansion transformation in function f. Express EL_{16}, K_{16} as

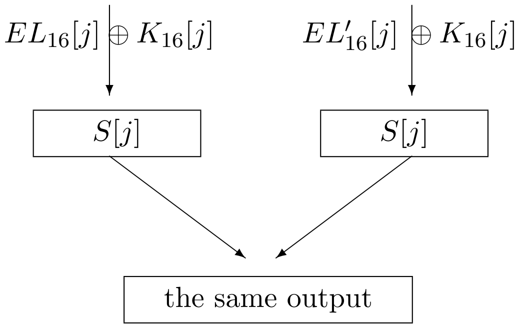
$$\begin{aligned} EL_{16} &= EL_{16}[1]\,||\,EL_{16}[2]\,||\,EL_{16}[3]\,||\,EL_{16}[4]\,||\,EL_{16}[5]\,||\,EL_{16}[6]\,||\,EL_{16}[7]\,||\,EL_{16}[8] \\ K_{16} &= K_{16}[1]\,||\,K_{16}[2]\,||\,K_{16}[3]\,||\,K_{16}[4]\,||\,K_{16}[5]\,||\,K_{16}[6]\,||\,K_{16}[7]\,||\,K_{16}[8] \end{aligned}$$

where each $EL_{16}[j], K_{16}[j], j = 1, \cdots, 8$, is of length 6-bit, $\alpha||\beta$ denotes the concatenation of the two strings α, β. Thus for each S-box $S[j], j = 1, \cdots, 8$, the input of $S[j]$ is

$$EL_{16}[j] \oplus K_{16}[j]$$

By the structure of f and Eq.(1), we have

$$S[j](EL_{16}[j] \oplus K_{16}[j]) = S[j](EL'_{16}[j] \oplus K_{16}[j]) \tag{2}$$



Claim-1: There are about 2^2 possible values for $K_{16}[j]$ if $EL_{16}[j] \neq EL'_{16}[j]$, and 2^6 values for $K_{16}[j]$ if $EL_{16}[j] = EL'_{16}[j]$.

In fact, the pair $(EL_{16}[j] \oplus K_{16}[j], EL'_{16}[j] \oplus K_{16}[j])$ is just a collision of the nonlinear function $S[j]$. Roughly speaking, $S[j]$ can be treated as a random or pseudorandom function. To find a collision of it for the given $EL_{16}[j], EL'_{16}[j]$, about 2^4 different arguments should be evaluated. Thus, there are 2^2 possible values for $K_{16}[j]$.

For each box $S[j], j = 1, \cdots, 8$, integrate each string a of 6-bit with $EL_{16}[j], EL'_{16}[j]$. Check Eq.(2) to determine all candidates for $K_{16}[j]$. Thus the corresponding candidates for K_{16} are achieved.

4 Description of the birthday attack against DES

𝟙 *Collecting proper ciphertexts.* Choose ciphertexts (64-bit) generated by the same key K. For each ciphertext c, compute its corresponding L_{16}, R_{16}. Collect the ciphertexts with the same R_{16} and denote the set by $\mathbb{C}_{R_{16},K}$. Denote $E(L_{16})$ by EL_{16}, where E is the expansion transformation in function f. Express EL_{16} as

$$EL_{16} = EL_{16}[1] \,||\, EL_{16}[2] \,||\, EL_{16}[3] \,||\, EL_{16}[4] \,||\, EL_{16}[5] \,||\, EL_{16}[6] \,||\, EL_{16}[7] \,||\, EL_{16}[8]$$

𝟚 *Computing the candidates for each $K_{16}[j]$.* Randomly pick two ciphertexts $c, c' \in \mathbb{C}_{R_{16},K}$. Integrate each string a of 6-bit with $EL_{16}[j], EL'_{16}[j]$. Determine the candidates for $K_{16}[j]$ by checking

$$S[j](EL_{16}[j] \oplus a) \overset{?}{=} S[j](EL'_{16}[j] \oplus a)$$

By the way, the 8×2^6 integrations can be run in parallel.

𝟛 If there does not exist any candidate for some $K_{16}[i]$, goto step 𝟚.

𝟜 *Determining the candidates for K_{16}.* Derive the candidates for K_{16} from the candidates for $K_{16}[1], \cdots, K_{16}[8]$.

𝟝 *Determining the candidates for K.* Derive the candidates for K from K_{16} by the key schedule of DES.

𝟞 *Distinguishing K from the candidates.* Given a plaintext and its corresponding ciphertext, the key (or its equivalent) is distinguished from its candidates by evaluations.

𝟟 *Outputting K.* If the key cannot be derived from the pair (c, c'), goto step 𝟚. Otherwise, output the key.

Remark 1 In the above attack, we aim at finding a collision (L_{15}, L'_{15}), which is achieved by evaluating possible values for $K_{16}[j], j = 1, \cdots, 8$. This is the reason for calling it a *birthday attack.*

5 Complexity

5.1 On the amount of ciphertexts

By $L_{15} = R_{16} \oplus f(L_{16}, K_{16})$ and the definition of $\mathbb{C}_{R_{16},K}$, we define

$$\mathcal{P}_{R_{16},K_{16}} : L_{16} \mapsto L_{15}$$

It's reasonable to assume that $\mathcal{P}_{R_{16},K_{16}}$ is random or pseudorandom. To find a collision for it, i.e.,

$$\mathcal{P}_{R_{16},K_{16}}(L_{16}) = L_{15} = L'_{15} = \mathcal{P}_{R_{16},K_{16}}(L'_{16})$$

about 2^{16} arguments should be evaluated. Practically speaking, it is not difficult to construct such a set $\mathbb{C}_{R_{16},K}$ satisfying $D \geq 2^{16}$ where D is the cardinal number of $\mathbb{C}_{R_{16},K}$, because each ciphertext is of only 64-bit.

5.2 On the amount of candidates for K in each iteration

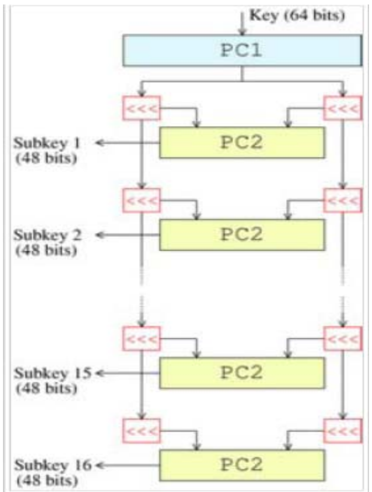
Define the *block-distance* between $c, c' \in \mathbb{C}_{R_{16},K}$ as

$$d = \#\{\, \lambda \,:\, EL_{16}[\lambda] \neq EL'_{16}[\lambda] \,\}$$

By the claim-1 and the definition of block-distance, we know the amount of candidates for K_{16} mainly depends on the block-distance of the pair (EL_{16}, EL'_{16}). In the best case, i.e., the block-distance is the maximum, 8, the amount of candidates for K_{16} is about 2^{16}. In the worst case, i.e., the block-distance is 1, the amount is 2^{44}.

Obviously, we are concerned about the average amount of candidates for K in each iteration. On average, a K_{16} leads to $\frac{7}{6}$ candidates for K. We conjecture the amount of candidates for K in each iteration is 2^{18}.

We refer to the following figure of the key schedule in DES.



5.3 On the amount of iterations

In the worst case, the amount of iterations is $\frac{D(D-1)}{2}$, namely we should try all ciphertext pairs of $\mathbb{C}_{R_{16},K}$. We conjecture the average amount of iterations is 2^{30}. Hence, the birthday attack should evaluate 2^{48} candidates for K. Thus, the attack has a computational complexity of 2^{48}.

6 Conclusion

We believe that the simple derivation of candidates for K from K_{16} can be a serious problem in DES. Possibly, it is due to historical considerations instead of a contrived process.

References

[1] http://en.wikipedia.org/wiki/Data_Encryption_Standard

[2] http://en.wikipedia.org/wiki/Birthday_attack

[3] http://dhost.info/pasjagor/des/start.php?id=0

[4] E.Biham, A.Biryukov. An Improvement of Davies' Attack on DES, Journal of Cryptology. 1997, 10(3), 195-206

[5] E.Biham, O.Dunkelman, N.Keller. Enhancing Differential-Linear Cryptanalysis. Advances in Cryptology-ASIACRYPT'2002. LNCS 2501, Springer-Verlag, 1990, 254-266

[6] S.Burton, J.Kaliski, R.Matthew. Linear Cryptanalysis Using Multiple Approximations, Advances in Cryptology-CRYPTO'1994. LNCS 839, Springer-Verlag, 1994, 26-39

[7] E.Biham, A.Shamir. Differential Cryptanalysis of DES-like Cryptosystems, Advances in Cryptology-CRYPTO'1990. LNCS 537, Springer-Verlag, 1990. 2-21

[8] D.Coppersmith. The data encryption standard (DES) and its strength against attacks. IBM Journal of Research and Development. 1994, 38 (3), 243-250

[9] K.Campbell, M.Wiener. DES is not a Group. Advances in Cryptology-CRYPTO'1992. LNCS 740, Springer-Verlag, 1992, 512-520

[10] J.Gilmore. Cracking DES: Secrets of Encryption Research, Wiretap Politics and Chip Design. O'Reilly, 1998.

[11] L.Knudsen, J.Mathiassen. A Chosen-Plaintext Linear Attack on DES, Fast Software Encryption-FSE'2000. LNCS 1978, Springer-Verlag, 2000, 262-272

[12] M.Matsui. Linear Cryptanalysis Method for DES Cipher, Advances in Cryptology-EUROCRYPT'1993. LNCS 765, Springer-Verlag, 1993, 386-397

[13] M.Matsui. The First Experimental Cryptanalysis of the Data Encryption Standard, Advances in Cryptology-CRYPTO'1994. LNCS 839, Springer-Verlag, 1994, 1-11