

Differential Cryptanalysis

See:

Biham and Shamir,
Differential Cryptanalysis of the Data Encryption Standard, Springer Verlag, 1993.

Differential Cryptanalysis

The first method which reduced the complexity of attacking DES below (half of) exhaustive search.

Note: In all the following discussion we ignore the existence of the initial and the final permutations, since they do not affect the analysis.

Motivation:

1. All the operations except for the S boxes are linear.
2. Mixing the key in all the rounds prohibits the attacker from knowing which entries of the S boxes are actually used, and thus he cannot know their output.

Differential Cryptanalysis (cont.)

How can we inhibit the key from hiding the information?

The basic idea of differential cryptanalysis: Study the differences between two encryptions of two different plaintexts: P and P^* .

Notation: For any value X during the encryption of P , and the corresponding value X^* during encryption of P^* , denote the difference by $X' = X \oplus X^*$.

Differential Cryptanalysis (cont.)

Advantages: It is easy to predict the output difference of linear operations given the input difference:

- **Unary operations** (E, P, IP):

$$(P(X'))' = P(X) \oplus P(X^*) = P(X')$$

- **Binary operations** (XOR):

$$(X \oplus Y)' = (X \oplus Y) \oplus (X^* \oplus Y^*) = X' \oplus Y'$$

- **Mixing the key:**

$$(X \oplus K)' = (X \oplus K) \oplus (X^* \oplus K) = X'$$

We conclude that the differences are linear in linear operations, and in particular, **the result is key independent**.

Differences and the S Boxes

Assume we have two inputs X and X^* for the same S box, and that **we know only their difference X'** .

Denote $Y = S(X)$.

What do we know about Y' ?

The simple case: **when $X' = 0$** : $S(X) = S(X^*)$ for any X , and $Y' = 0$.

If $X' \neq 0$: we do not know the output difference.

Definition: Lets look on the distribution of the pairs (X', Y') of all the possible inputs X . We call the table containing this information **difference distribution table of the S box**.

The Difference Distribution Table of S1

Input XOR	Output XOR															
	0x	1x	2x	3x	4x	5x	6x	7x	8x	9x	Ax	Bx	Cx	Dx	Ex	Fx
0x	64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ax	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bx	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cx	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Dx	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ex	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Fx	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27x	10	4	2	0	2	4	2	0	4	8	0	4	8	8	4	4
28x	12	2	2	8	0	2	12	0	0	8	8	0	8	8	0	8
29x	4	2	2	10	0	6	8	2	14	10	10	10	10	10	10	10
30x	4	2	2	2	4	6	6	2	2	14	6	2	6	2	6	2
31x	12	2	2	2	4	6	6	2	2	14	6	2	6	2	6	2
32x	6	2	6	2	8	4	4	4	2	4	6	0	8	2	0	6
33x	6	2	6	2	2	6	8	8	2	4	4	6	8	2	4	2
34x	0	4	6	0	12	6	2	8	2	4	4	6	8	2	2	4
35x	4	8	2	10	2	2	2	6	0	0	2	2	2	4	10	8
36x	0	8	16	6	2	0	12	6	0	0	0	0	0	0	8	0
37x	2	2	2	4	0	8	0	0	14	4	6	8	0	2	14	0
38x	2	2	2	2	8	0	2	2	4	2	2	8	6	4	10	0
39x	0	2	12	4	2	4	4	10	4	4	2	6	0	2	10	4
40x	0	6	2	2	4	2	4	4	4	4	2	6	0	2	10	4
41x	6	2	2	4	12	6	4	8	4	0	2	4	2	4	4	0
42x	6	4	6	4	6	8	0	6	4	6	2	8	16	4	4	6
43x	0	10	4	0	12	0	2	6	0	4	12	4	4	2	0	0
44x	0	8	6	2	2	2	6	0	14	4	4	0	3	0	12	4
45x	4	8	2	2	2	4	4	14	4	2	4	8	8	6	2	2
46x	4	8	4	2	4	0	2	4	2	4	8	8	6	2	2	2

The Difference Distribution Table of S1 (cont.)

Observe that:

- In the first line $X' = 0$ and thus all the 64 pairs satisfy $Y' = 0$. $Y' \neq 0$ is impossible.
- In the rest of the lines: The average value is 4, the sum in each line is 64. The values are all even in the range 0-16.

The entries with value 16 mean that for a quarter of the pairs with input difference X' , the output difference is the particular Y' .

The entries with value 0 mean that there are no pairs with the corresponding input difference X' and the corresponding output difference Y' .

Differences and the S Boxes (cont.)

Definition: If the entry of the input difference X' and the output difference Y' is greater than zero, we say that **X' may cause Y' by the S box**, and denote $X' \rightarrow Y'$.

Definition: The probability of $X' \rightarrow Y'$ is the probability that for a pair with value X' , the output difference is Y' , among all the possible pairs. In DES, the probability is the corresponding value in the difference distribution table divided by 64.

Similarly we define $X' \rightarrow Y'$ by the **F-function**, and define the probability as the product of the probabilities by the eight S boxes.

Differences and the S Boxes (cont.)

Differential cryptanalysis uses the entries with large values, and in particular the $0 \rightarrow 0$ entry and the entries with value 16, and other large values.

Observation

Given an input and output differences of an S box, it is possible to list all the pairs with these differences.

Example: For the entry $09_x \rightarrow 1_x$ the 2 pairs are:

1. $33_x, 3A_x$
2. $3A_x, 33_x$

For the entry $01_x \rightarrow F_x$ the 4 pairs are:

1. $1E_x, 1F_x$
2. $1F_x, 1E_x$
3. $2A_x, 2B_x$
4. $2B_x, 2A_x$

The lists of pairs of all the differences can easily be computed in advance.

Example of a Simple Attack

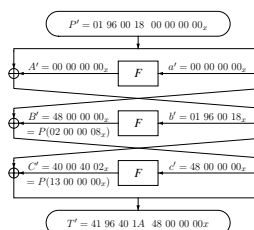
Assume a 3-round DES, in which for some pair of plaintexts $P' = 01\ 96\ 00\ 18\ 00\ 00\ 00\ 00_x$, and $T' = 41\ 96\ 40\ 1A\ 48\ 00\ 00\ 00_x$.

We also assume that $T = 00\ 00\ 00\ 00\ 08\ 00\ 00\ 00_x$ and $T^* = 41\ 96\ 40\ 1A\ 40\ 00\ 00\ 00_x$.

(We use the notation T for the ciphertexts, as we use C for the third round intermediate values.)

Example of a Simple Attack (cont.)

Then, the differences in the various rounds are



Example of a Simple Attack (cont.)

We identify that S1 in the third round accepts difference 09_x in the input and outputs difference 1_x in the output. Looking at the difference distribution table, we find only two possible pairs for this combination ($(33_x, 3A_x)$ and $(3A_x, 33_x)$).

Thus, we get the following equations:

$$S1_E \oplus S1_K = 33_x \text{ or } 3A_x$$

$$S1_E^* \oplus S1_K = 3A_x \text{ or } 33_x.$$

From the known ciphertexts we know that

$$S1_E = 01_x$$

$$S1_E^* = 08_x.$$

Therefore, we can find two possible values for $S1_K$

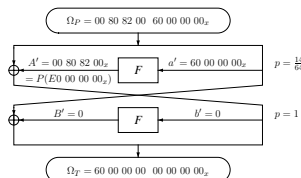
$$S1_K = 32_x \text{ or } 3B_x.$$

(Notice that the difference between these two values is always the input difference, 09_x in this case.)

Characteristics (תכונות)

In differential cryptanalysis we wish to know some statistical information on the differences in intermediate rounds during encryption, given only the plaintext difference.

Example: A **two-round characteristic** with probability $\frac{14}{64}$ (In S1, $0C_x \rightarrow E_x$ with probability $\frac{14}{64}$):



Characteristics (תכונות) (cont.)

Informal Definition: Associated with any pair of encryptions are the XOR value of its two plaintexts, the XOR of its ciphertexts, the XORs of the inputs of each round in the two executions and the XORs of the outputs of each round in the two executions. These XOR values form an **n-round characteristic**. A characteristic has a probability, which is the probability that a random pair with the chosen plaintext XOR has the round and ciphertext XORs specified in the characteristic. We denote the plaintext XOR of a characteristic by Ω_P and its ciphertext XOR by Ω_T .

Characteristics (תכונות) (cont.)

Definition: An **n-round characteristic** is a tuple $\Omega = (\Omega_P, \Omega_A, \Omega_T)$ where Ω_P and Ω_T are m -bit numbers and Ω_A is a list of n elements $\Omega_A = (\Lambda_1, \Lambda_2, \dots, \Lambda_n)$, each is a pair of the form $\Lambda_i = (\lambda_i^1, \lambda_i^0)$ where λ_i^1 and λ_i^0 are $m/2$ bit numbers and m is the block size of the cryptosystem. A characteristic satisfies the following requirements:

- λ_i^1 = the right half of Ω_P
- λ_i^0 = the left half of $\Omega_P \oplus \lambda_0^0$
- λ_i^1 = the right half of Ω_T
- λ_i^0 = the left half of $\Omega_T \oplus \lambda_0^0$

and for every i such that $2 \leq i \leq n - 1$:

$$\lambda_0^0 = \lambda_i^1 \oplus \lambda_i^{i+1}.$$

Characteristics (תכונות) (cont.)

Definition: Characteristics can be concatenated if $\text{swap}(\Omega_T^1) = \Omega_T^2$. The resultant characteristic is

$$\Omega = (\Omega_P^1, \Omega_A^1 || \Omega_A^2, \Omega_T^2).$$

Definition: A **right pair** (זוג נכון) with respect to a characteristic Ω and a key K is a pair P, P^* , which satisfies $P' = \Omega_P$, and all whose differences in the rounds $1, \dots, n$ are as predicted by the characteristic.

Characteristics (תכונות) (cont.)

Definition: An **independent key** (מפתח בלתי תלוי) is a list of subkeys which is not necessarily derivable from some key via the key scheduling algorithm.

Probability of a Characteristic

Definition: The **probability** of a characteristic is the probability that a random pair P, P^* which satisfies $P' = \Omega_P$ is a right pair with respect to a random independent key.

Note: The probability of a characteristic is the product of all the probabilities of the S boxes in the characteristic.

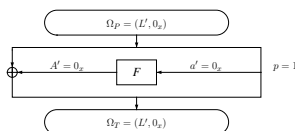
Probability of a Characteristic (cont.)

Note: The probability of characteristics of DES is the probability that any specific pair P, P^* ($P' = \Omega_P$) is a right pair among all random keys. We are more interested in the probability that for a specific (unknown) key, a random pair P, P^* ($P' = \Omega_P$) is a right pair. In practice, the first probability is a good approximation of the second probability.

Examples of One-Round Characteristics

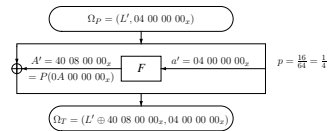
Choose the inputs of the S boxes by the best entries in the difference distribution tables.

Example: An one-round characteristic with probability 1 is (for any L'):



Examples of One-Round Characteristics (cont.)

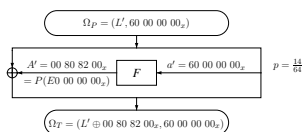
The second best one-round characteristic has probability $1/4$, using only one active S box (S2):



There is a similar characteristic using S6.

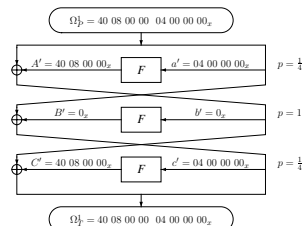
Examples of One-Round Characteristics (cont.)

The next best characteristic has probability $\frac{14}{64}$:



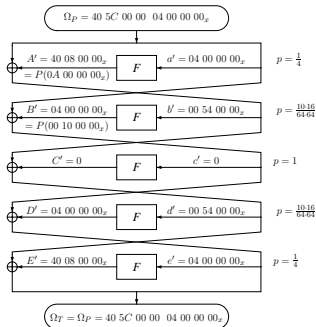
A Three-Round Characteristic

A three-round characteristic with probability $1/16$:



A Five-Round Characteristic

A five-round characteristic with probability about 1/10486:



Probabilities Versus Number of Rounds

The probabilities of the characteristics reduces very fast with the number of rounds:

Number of rounds	Probability
1	1
2	1/4
3	1/16
4	≈ 1/800
5	≈ 1/10000
6	≈ 1/1000000

Probabilities Versus Number of Rounds (cont.)

As the number of rounds is increased, the reduction rate grows. By the table, we may expect that at 9–10 rounds, the probabilities are smaller than 2^{-56} or 2^{-64} .

We are interested in longer characteristics with higher probabilities.

Differentials

Usually differential cryptanalysis use only the Ω_P and Ω_T of the characteristics, but not the intermediate values.

Definition: A **Differential** is a set of all the characteristics with the same Ω_P and Ω_T .

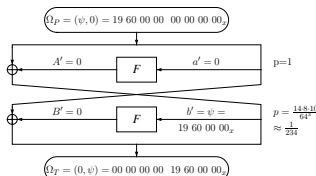
The probability of the differential is the sum of the probabilities of the various characteristics.

In most differential attacks we actually use differentials, rather than characteristics. The probabilities of the characteristics serve as **lower bounds** for the probabilities of the differentials.

Iterative Characteristics

Characteristics which can be concatenated to themselves are called **iterative characteristics**.

The best iterative characteristic of DES is:



where $\psi = 19 60 00 00_x$. Due to the importance of this iterative characteristic, we call it **the iterative characteristic**.

There is another value $\psi^\dagger = 1B 60 00 00_x$ for which the iterative characteristic has the same probability.

Iterative Characteristics (cont.)

These two characteristics are the best when iterated to seven or more rounds.

Note: In DES, in order to receive the same output of the F -function, two different inputs must differ in the input of at least three S boxes.

Probabilities Versus Number of Rounds

The probability of the iterative characteristic versus the number of rounds:

Number of rounds	Probability
3	$2^{-7.9} \approx 1/234$
5	$2^{-15.7} \approx 1/55000$
7	$2^{-23.6}$
9	$2^{-31.5}$
11	$2^{-39.4}$
13	$2^{-47.2}$
15	$2^{-55.1}$
16	2^{-62}
17	2^{-63}

Differential Attacks

The simplest differential attack (0R-attack) breaks ciphers with the same number of rounds as the characteristic. Using 3-round characteristics we can find key bits of 3-round DES, and using 5-round characteristics we can find key bits of 5-round DES.

Differential Attacks (cont.)

The basic algorithm:

1. Choose some $m = 2p^{-1}$ random pairs P, P^* such that $P' = \Omega_P$, and request the corresponding ciphertexts T and T^* under the unknown key K .
2. Choose only the pairs satisfying $T' = \Omega_T$, and discard the others. About $m(p + 2^{-64})$ pairs remain (from the m pairs): mp right pairs and $2^{-64}m$ wrong pairs. If $p \gg 2^{-64}$ we can assume that all the remaining pairs are right pairs.

Differential Attacks (cont.)

3. Each remaining right pair satisfies the difference predictions of the characteristics and its values of T and T^* are known. The differences of the inputs and the outputs of the S boxes of the last round are known from $T' = T \oplus T^*$ (and from the characteristic).

If the input difference is non-zero, not all the inputs are possible, and only a minority of the inputs satisfy the input and output differences: in each pair only about 0–16 possible values for the 6 input bits of the S box are possible. Each value suggests one value for the 6 corresponding key bits.

The right value of the 6 key bits must be suggested by all the right pairs, while other values are suggested arbitrarily by only a few of the pairs. By cutting the sets of keys suggested by all the pairs, we receive two possible values for each 6 key bits; in total we receive $2^8 = 256$ possible values for 48 key bits (if all the eight S boxes are active).

If a wrong pair still remains, still the keys suggested by the largest number of pairs are likely to include the right key.

Difficulty of Application to the Full DES

In order to attack the full DES (16-rounds) we need at least $2 \cdot 2^{62}$ pairs:

1. Their encryption costs more than exhaustive search.
2. **Include all the 2^{64} plaintext blocks** (who needs the key in this case?).
3. The identification of right pairs is not so good, since $p \gg 2^{-64}$

Enhancements: *R-Attacks

We observe that characteristics shorter than the cipher can be used. Attacks using characteristics shorter than the cipher by r rounds (in which the characteristic predicts the differences in the first $n - r$ rounds of the cipher) are called rR -attacks.

0R-attacks In 0R-attacks (as in the previous slides) we know that $T' = \Omega_T$, and thus it is easy to identify the right pairs. Then we use the information on the differences inside the characteristic. Still, we cannot identify between two possible values for each S box.

1R-Attacks

In these attacks, the characteristic predicts the differences except in the last round, and Ω_T is the predicted difference before the last round. The input difference of the F -function of the last round is known both from the characteristic and the ciphertexts $(T')_R = (\Omega_T)_L$, and it can be used to discard wrong pairs. On the other hand, the difference of the output of the F -function can be calculated as $(T')_L \oplus (\Omega_T)_R$.

Thus, we can use shorter characteristics with higher probabilities, although the identification of the right pairs is somewhat worse.

2R-Attacks

Allow to use a characteristic shorter than the cipher by two rounds.

In these attack, the attacker knows

1. The differences of the input to the last F -function, and the inputs themselves.
2. The predicted differences of the input to the F -function in the second-last round (from the characteristic).
3. The differences of the outputs of the last two F -functions can be calculated from Ω_T and T' .

2R-Attacks (cont.)

Identification and discarding of wrong pairs

For each S box in the last two rounds (total of 16 S boxes) we calculate the predicted input and output differences as above. If for some S box, the input difference may not cause the output difference (value 0 in the difference distribution table) the pair cannot be a right pair.

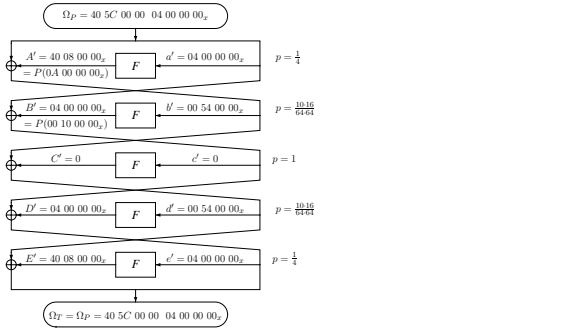
3R-Attacks / Attacking 8 Rounds

Allow to use a characteristic shorter than the cipher by three rounds.

Example: Breaking DES reduced to eight rounds using a 3R-attack:

Use the 5-round characteristic with probability about 1/10486:

3R-Attacks / Attacking 8 Rounds (cont.)



Attacking 8 Rounds: Brief Description

The attacker chooses pairs P, P^* satisfying $P' = \Omega_P$. With probability $p = 1/10486$ the difference after five rounds is Ω_T . In the sixth round $f' = (\Omega_T)_L = 40\ 5C\ 00\ 00_x$: $S1:08_x, S2:00_x, S3:0B_x, S4:38_x, S5:00_x, S6:00_x, S7:00_x, S8:00_x$. Thus, the output differences of S2, S5, S6, S7 and S8 are zero as well.

The output differences of S2, S5, S6, S7 and S8 in the last round can be calculated from Ω_T, T' and these zeroes. The inputs to the last round are known, and thus the inputs to the S boxes are known up to XOR with the last subkey K8.

Attacking 8 Rounds: Brief Description (cont.)

We can find several possible values for the key bits entering each of the five S boxes in the last round, total of 30 key bits. The right value of these 30 key bits is expected to appear as the most frequent value: it is suggested by all the right pairs (by about $1/10486$ of the pairs). Any other value is suggested by about $\frac{4^5}{2^{30}} = 2^{-20} = \frac{1}{1048576}$ of the pairs.

The right value will be suggested 100 times more frequently than any other value, and thus is easily identified by counting the frequency of the suggested values.

About 100000 pairs (and even less) suffice for this attack.

Attacking 8 Rounds: Detailed Description

1. Choose 100000 pairs P, P^* satisfying $P' = \Omega_P$, and request their ciphertexts T, T^* under the unknown key K .
2. Initialize an array of 2^{30} entries with zeroes.

Attacking 8 Rounds: Detailed Description (cont.)

3. Compute the inputs and the input difference of the last F -function:

$$\begin{aligned} h &= T_R \\ h^* &= T_R^* \\ h' &= h \oplus h^* \end{aligned}$$

and 20 bits of the output difference

$$H' = (\Omega_T)_R \oplus F' \oplus T'_L$$

where 20 bits of F' are known to be zero, and the same 20 bits are calculated for H' : the output of five S boxes.

Attacking 8 Rounds: Detailed Description (cont.)

4. For each of the five S boxes in the last round for which the inputs X, X^* as well as the output differences Y' are known, calculate all the possible values of their 6 key bits, which satisfy $S(X \oplus k) \oplus S(X^* \oplus k) = Y'$, and create a list of all the possible 30 bits of the key. For each 30-bit value, increment (by one) the corresponding entry in the array.
5. After all the pairs are processed, the highest entry should correspond to the right value of the 30 key bits.
6. Complete the remaining 26 key bits (by exhaustive search or by a differential attack).

A variant of this algorithm requires an array of only 2^{18} bytes, and it finds the key within a few seconds on a PC.

Conversion to a Known Plaintext Attack

Differential chosen plaintext attacks can be converted to known plaintext attacks with higher complexities:

1. Assume a chosen plaintext attack requires m pairs P, P^* with difference $P' = \Omega_P$.
2. Request $2^{32}\sqrt{2m}$ random known plaintexts.
3. There are $(2^{32}\sqrt{2m})^2/2$ pairs in these plaintexts, which are $2^{64}m$ pairs.
4. Each value of P' appears for about 2^{-64} of the pairs, i.e., for about m pairs.
5. In particular, there are about m pairs with the plaintext difference $P' = \Omega_P$. (These pairs can be identified efficiently using hash tables).
6. The original chosen plaintext attack is executed on these m pairs.

Conversion to a Known Plaintext Attack (cont.)

The number of required chosen plaintexts vs. the number of required known plaintexts:

m	$2^{32}\sqrt{2m}$
2	2^{33}
8	2^{34}
2^7	2^{36}
2^{15}	2^{40}
2^{31}	2^{48}
2^{55}	2^{60}

The Attack on the Full 16-Round DES

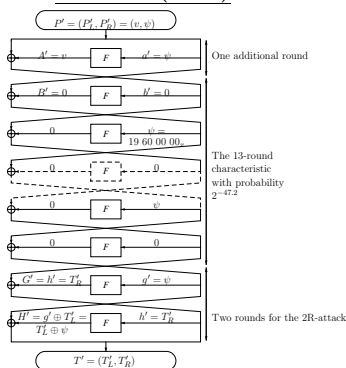
Motivation:

1. The 15-round characteristic has probability $2^{-55.1}$, and clearly cannot be used to reduce the complexity of attack below 2^{55} .
2. The 14-round characteristic has probability $2^{-54.1}$.
3. In order to attack DES, we must then use characteristics of at most 13 rounds.
4. However, 3R-attacks are infeasible, since due to lack of data the right key cannot be identified.

The Idea

Add an additional round as a **first** round, not included in the characteristic, and **without cost**.

The Idea (cont.)



The Data

1. Let $\{v_j\}$ be the set of 2^{12} possible output values of S1, S2 and S3, after the P permutation, where all the other 20 bits are zero (assume $v_0 = 0$).
2. Choose the plaintexts in structures of 2^{14} , using the two best iterative characteristics:
 - (a) Choose (random) P_0 .
 - (b) $P_1 = P_0 \oplus (0, \psi^1)$, where $\Omega_P^1 = (\psi^1, 0)$.
 - (c) $P_2 = P_0 \oplus (0, \psi^2)$, where $\Omega_P^2 = (\psi^2, 0)$.
 - (d) $P_3 = P_0 \oplus (0, \psi^1 \oplus \psi^2)$.
 - (e) For $0 \leq i \leq 3, 0 < j < 2^{12}: P_{i+4j} = P_i \oplus (v_j, 0)$.
3. In this structure, for every P_i there is some **unknown** P_j whose difference (before round 2) is Ω_P^1 . Similarly for Ω_P^2 .
4. Therefore, for each characteristic, there are 2^{13} pairs in the structure, and in total 2^{14} for both characteristics.

The Data (cont.)

5. Right pairs: the 13-round characteristic probability is $2^{-47.2}$. In a structure there are on average $2^{14} \cdot 2^{-47.2} = 2^{-33.2}$ right pairs.
6. One right pair is expected to exist in $2^{33.2}$ structures on average, i.e., in about $2^{47.2}$ chosen plaintexts.

Identification of Wrong Pairs

$\Omega_T = (\psi, 0)$, thus the input of the F -function in the second-last round differs by ψ in the right pairs. ψ is non-zero only in the input to S1, S2 and S3. Thus, the 20-bit output difference of S4,S5,S6,S7,S8 is zero.

The input difference of the last round must be zero in these 20 bits.

This difference can be easily calculated for any pair, and can be used to discard most of the wrong pairs: A wrong pair passes the test with probability 2^{-20} , in total there are 2^{26} pairs in each structure, and thus only about 2^6 wrong pairs pass the test.

These remaining pairs can be found efficiently: Hash the 2^{14} plaintexts by the 20 bits of T_R , and process only those hashed to the same entry. It requires only about 2^{14} steps, instead of 2^{26} .

Identification of Wrong Pairs (cont.)

We now discard additional wrong pairs by examining the other S boxes in the first, 15th and the 16th, and verifying that their computed input difference may cause their computed output difference. This test discards about $1 - \left(\frac{14}{16} \cdot \frac{13}{16} \cdot \frac{15}{16}\right)^2 \cdot 0.8^8 = 1 - 0.0745 = 92.55\%$ of the remaining wrong pairs. Only about $2^6 \cdot 0.0745 = 4.768$ wrong pairs from each structure remain after this test.

(Consult the book for the exact calculation).

Finding the Key Using One Right Pair

In previous differential attacks we counted the frequency of the keys, and thus needed several right pairs.

We observe that when we count by a large number of bits, it is more efficient to compute a trial encryption to verify key directly.

We first find 52-bit values corresponding to the 48 bits of the last subkey plus 4 bits accessible in rounds 1 and 15. For this, we now take into consideration that the subkeys are not independent.

Instead of counting on the 52 key bits, we complete the 52 bits to 56 bits (with all the possible values of the additional 4 bits), and compute a trial encryption on each of the 56-bit keys:

Finding the Key Using One Right Pair (cont.)

- Given the $2^{47.2}$ ciphertexts, there is a right pair with a high probability.
- Discard wrong pairs by the algorithm in the previous slides.
- For each remaining pair do:
- Compute all the possible values of the 52 key bits: a total of 4^8 values on average for the last subkey for each pair, complete additional 4 bits using rounds 1 and 15, and discard contradicting values. Each analyzed pair proposes about $2^{52} \cdot \frac{2^{-32}}{0.88} \cdot \frac{2^{-12}}{\frac{16}{16} \cdot \frac{16}{16}} \cdot \frac{2^{-12}}{\frac{16}{16} \cdot \frac{16}{16}} = 0.84$ values for the 52 bits. Thus, each structure proposes $4.768 \cdot 0.84 = 4$ values on average.
- Complete the 52 bits to 56 bits by adding all the possible 4-bit values.

Finding the Key Using One Right Pair (cont.)

- Compute a trial encryption on each of the $4 \cdot 16 = 64$ 56-bit keys proposed by each structure.
- A total of $2^{47.2}/2^{14} \cdot 64 = 2^{39.2}$ trial encryptions are applied (and it can be reduced further to 2^{37}).
- During processing of the **first right pair**, the key **must** be found. Then, it can easily be verified with additional tests.

Results

Summary of the cryptanalysis of DES: The number of operations and plaintexts required to break the specified number of rounds.

No. of Rounds	Dependent Key		Independent Key	
	Chosen Plaintexts	Known Plaintexts	Chosen Plaintexts	Known Plaintexts
4	2^3	2^{33}	2^4	2^{33}
6	2^8	2^{36}	2^8	2^{36}
8	2^{14}	2^{38}	2^{16}	2^{40}
9	2^{24}	2^{44}	2^{26}	2^{45}
10	2^{24}	2^{43}	2^{35}	2^{49}
11	2^{31}	2^{47}	2^{36}	2^{50}
12	2^{31}	2^{47}	2^{43}	2^{53}
13	2^{39}	2^{52}	2^{44}	2^{54}
14	2^{39}	2^{51}	2^{51}	2^{57}
15	2^{47}	2^{56}	2^{52}	2^{58}
16	2^{47}	2^{55}	2^{60}	2^{61}

Additional Results

The effect of modifying the operations on the strength of DES:

The P permutation: Cannot strengthen DES, since the iterative characteristic is not affected by P. However, bad choices can crucially reduce the strength (for example the identity permutation).

Eliminating E, 4x4 S boxes: 2^{26} .

Order of E and the subkeys: 2^{44} (32-bit subkeys).

Additional Results (cont.)

The order of the S boxes: Can weaken much (the order S1, S7, S4, ... weakens to 2^{38}). Can strengthen only up to 2^{48} .

Modifying the S boxes:

- Random: 2^{18} - 2^{20} .
- Random permutations: 2^{33} - 2^{41} .
- Modifying one entry: 2^{33} .
- Uniform difference distribution tables: 2^{26} .

S³DES S boxes: This set of replacement S boxes was proposed by Kim et al. The 2-round iterative characteristics become impractical since they require the whole eight S boxes to be active to get a zero output difference in the F function. These S boxes (when S1 and S2 are exchanged) are immune against differential and linear cryptanalysis with complexities over 2^{60} . The (Improved) Davies' attack is not applicable at all.

Additional Results (cont.)

Independent keys:

- Eight rounds: finds the 384 key bits with the same complexity and data as in the case of the dependent keys (56 bits).
- 16 rounds: finds the 768 key bits with 2^{60} chosen plaintexts and 2^{60} complexity, or 2^{61} known plaintexts.

Extensions of Differential Cryptanalysis

- Conditional characteristics (Ben-Aroya, Biham)
- Higher-order differential cryptanalysis (Lai ; Biham)
- Markov Ciphers (Lai, Massey)
- Truncated Differentials (Knudsen)
- Provable Security against Differential Attacks (Knudsen, Nyberg)
- Impossible Differentials (1998, Biham, Biryukov, and Shamir).
- Boomerang, amplified boomerang, and rectangle attacks.