В інших випадках знайти кількість точок еліптичної кривої є дуже складною задачею і потребує дуже багато обчислювальних ресурсів.

Література: 1. Dan Beaureal. Efficient algorithms for implementing elliptic curve public-key schemes. A Thesis submitted to the Faculty of the Worcester Polytechnic Institute. 1996. 2. R.Mullin, I.Onyszchuk, S.Vanstone and R.Wilson. Optimal normal bases in GF(p^n). Discrete Applied Mathematics, 22(1988/89). 149-161.

УДК 638.253.231

Differential Cryptanalysis of Feistel's Iterated Block Ciphers

Alexander Telizhenko, Sergey Limar

Kiev Military Institute for Command and Communication

*Анотація:***В статті обговорюються базові принципи Диференційного криптоаналіза, концепції, алгоритми, ідеї і методи, які забезпечують цей тип атаки, а також математичне обгрунтування.**

Summary: **Here are described the basic principles of Differential Cryptanalysis, concepts, algorithms, ideas and methods which provide this kind of attack and also its mathematical background.**

*Ключові слова:***Differential attack, round differentials, conditional characteristic, probabilistic influence, chosen plaintext.**

I Introduction

This paper will attempt to introduce some concepts of cryptography, and especially some ideas pertaining to cryptanalysis, the breaking of encryption. 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.

In this research announcement, we describe a related attack (which we call Differential Cryptanalysis), and show that it is applicable to almost any secret key cryptosystem proposed so far in the open literature. In particular, we have actually implemented it in the case of DES, and demonstrated that under the same software differential model, we can extract the full DES key from a sealed tamperproof DES encryptor by analysing fewer than 200 ciphertexts generated from unknown plaintexts. The power of Differential Cryptanalysis is demonstrated by the fact that even if DES is replaced by triple DES (whose 168 bits of key were assumed to make it practically invulnerable), essentially the same attack can break it with essentially the same number of given ciphertexts.

II Motivation

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

How can we inhibit the key from hiding the information?

3. Ideas, methods and principles of Differential Attack.

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 $X' = X \oplus X^*$.

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

1. Unary operations (E, P, IP):

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

2. Boolean operations (XOR):

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

3. Mixing the key:

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

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

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.

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 $X' \to Y'$.

Definition: The probability of $X' \to Y'$ is the probability that for a pair with the input difference 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 $X' \to Y'$ by the F-function, and define the probability as the product of the probabilities by the eight S boxes.

Informal Definition: Associated with any pair of encryptions arc 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.

Definition: An n-round characteristic is a tuple $\Omega = (\Omega_p, \Omega_\Lambda, \Omega_T)$ where Ω_p and Ω_T are m-bit numbers and Ω_Λ is a list of n elements $\Omega_\Lambda = (\Lambda_1, \Lambda_2, ..., \Lambda_n)$ each is a pair of the form $\Lambda_i = (\lambda_I^i, \lambda_O^i)$ where λ_I^i and λ_O^i are $m/2$ bit numbers and m is the block size of the cryptosystem. A characteristic satisfies the following requirements:

$$\lambda_I^1 = \text{the right half of } \Omega_P$$

$$\lambda_I^2 = \text{the left half of } \Omega_P \oplus \lambda_O^1$$

$$\lambda_I^n = \text{the right half of } \Omega_T$$

$$\lambda_I^{n-1} = \text{the left half of } \Omega_T \oplus \lambda_O^n$$

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

$$\lambda_0^i = \lambda_I^{i-1} \oplus \lambda_I^{i+1}$$

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

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

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

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

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

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.

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.

Characteristics, which can be concatenated to themselves arc called iterative characteristics.

The simplest differential attacks (later called $0R$-attacks) break 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.

The algorithm:

1. Choose some $m = 2p^{-1}$ random pairs P, P^* such that $P' = \Omega_P$, and request the corresponding ciphertexts C and C^* under the unknown key K.
2. Choose only the pairs satisfying $C' = \Omega_T$, 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 $p >> 2^{-64}$ we can assume that all the remaining pairs are right pairs.
3. Each remaining right pair satisfies the difference predictions of the characteristics and its values C and C^* are known. The differences of the inputs and the outputs of the S-boxes of the last round are known from $C' = C \oplus C^*$ (and from the characteristic).

If the input difference is non-zero, not all the inputs arc 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.

In order to attack the full DES (16-rounds) we need at least $2 * 2^{62}$ 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 $p \lhd 2^{64}$

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 (which were studied in the previous slides) we know that $C' = \Omega_T$, 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 $(C')_L = (\Omega_T)_L$, 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 $(C')_L \oplus (\Omega_T)_R$.

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 C'.

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: 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 $1/10486$:

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

$$f' = (\Omega_T)_L = 405C0000_x :$$
$$S1:08_x, S2:00_x, S3:0B_x, S4:38_x,$$
$$S5:00_x, S6:00_x, S7:00_x, S8:00_x.$$

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, C' 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 $K8$.

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.

1. Choose 100000 pairs P, P^* satisfying $P = \Omega_P$, and request their cipher-texts C, C^* under the unknown key K.
2. Initialize an array of' 2^{30} entries with zeroes.
3. Compute the inputs and the input difference of the last F -function:
$$h = C_R$$
$$h^* = C_R^*$$
$$h' = h \oplus h^*$$
and 20 bits of the output difference
$$H' = (\Omega_T)_R \oplus F' \oplus C_L'$$
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.
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 2^{18} bytes.

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 $(2^{32}\sqrt{2m})/2$ 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 Ω_P. (These pairs can be identified efficiently using hash tables).
6. The original chosen plaintext attack is executed on these m pairs.

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 arc infeasible, since due to lack of data the right key cannot be identified.

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 $0 \le i \le 3, 0 < j < 2^{12} : P_{i+4j} = P_i \oplus (v_j, 0)$.
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.
5. Right pairs: the 13-round characteristic probability is $2^{-47,2}$. In a structure there are in average $2^{14} * 2^{-47,2} = 2^{-33,2}$ right pairs.
6. One right pair is expected to exist in $2^{33,2}$ structures in average, i.e., in about $2^{33,2}$ 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 C_R, and process only those hashed to the same entry. It requires only about 2^{14} steps, instead of 2^{26}.

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.

Instead of counting on the 48 key bits, we complete the 48 bits to 56 bits (with all the possible values of the additional 8 bits), and compute a trial encryption on each of the 56-bit keys:

1. Given the $2^{47,2}$ ciphertexts, there is a right pair with a high probability.
2. Discard wrong pairs by the algorithm in the previous slides.
3. For each remaining pair do:
4. Compute all the possible values of the last 48-bit subkey: a total of 4^8 values in average for each pair.
5. Complete the 48 bits to 56 bits by adding all the possible 8-bit values.
6. Compute a trial encryption on each of the $4^8 * 2^8 = 2^{24}$ keys.
7. During processing of the first right pair, the key must be found. With a high probability, only the right pair passes the trial encryptions.

III Conclusion

Differential Cryptanalysis can break many additional secret key cryptosystems, including IDEA, RC5 and Feal. Some ciphers, such as Khufu, Khafre and Blowfish compute their S boxes from the key material. In such ciphers, it may be even possible to extract the S boxes themselves, and the keys, using the techniques of Differential Cryptanalysis.

Differential Cryptanalysis can also be applied against stream ciphers, but the implementation might differ by some technical details from the implementation described above.

Література: E. Biham, A. Shamir. Differential Cryptanalysis of DES – like Cryptosystems, Journal of Cryptography. Vol. 1. № 1. pp 3 – 72. 1991. E. Biham, A. Shamir. Differential Cryptanalysis of Data Encryption Standard. Springer – Verlag - 1993. NBS. Data Encryption Standard. U.S. Department of Commitie. FIPS pub.16. January, 1997. E. Biham. New Types of Cryptanalytic Attacks Using Related Keys. Eurocrypt'93. LNCS 765. Springer - 1994.

УДК 621.391

ПРО ВІДНОШЕННЯ ЕКВІВАЛЕНТНОСТІ НА МНОЖИНІ БУЛЬОВИХ ФУНКЦІЙ

Сергій Мельник
Київський військовий інститут управління і зв'язку

Анотація: **Розглядається питання класифікації бульових функцій по відношенню еквівалентності, що визначене на множині F$_n$ всіх бульових функцій від n аргументів, а також деякі криптографічні властивості бульових функцій на одному з визначених класів.**

Summary: **In this article was examined classification of boolean functions on the attitude ecvivalation, which determined on the ensemble F$_n$ of all boolean functions from n arguments, as well as some cryptographs particularity of boolean functions on one of determined classes.**

Ключові слова: **бульова функція, клас, криптографічні властивості.**

Нехай G – деяка група взаємно однозначних перетворень векторного простору Z$_n$ (Z={0,1}). Функції $f_1(x_1,...,x_n)$ та $f_2(x_1,...,x_n)$ називаються еквівалентними відносно групи G, якщо для деякого елемента g∈G

$$f_2(x_1,...,x_n) = f_1(g(x_1,...,x_n)).$$

Представлене відношення є відношенням еквівалентності ([1,2]), за сяким множина всіх бульових функцій розбивається на класи. У випадку, коли $G = \langle S_n, \sum_n \rangle$, класи еквівалентності називають типами, а також якщо $G = S_n$ - розрядами (де S_n - симетрична група перестановок координат векторів із Z$_n$, $\sum_n$ - група зсувів простору Z$_n$).

Однотипні функції представляють собою одну логічну форму, записану в різних системах координат, тому значна частина властивостей бульових функцій одного типу співпадає.

З точки зору криптографічних застосувань, важливим є питання про співпадання деяких криптографічних властивостей бульових функцій ([3]), що належать одному типу, а саме: нелінійність, кореляційний імунітет та виконання строгого потокового критерію (“strict avalanche criterion”). З [3] відомо, що перелічені властивості можуть бути сформульовані в термінах перетворення Уолша: $\hat{F}(\overline{w}) = \sum_{\bar{x} \in Z_n} (-1)^{f(\bar{x}) \oplus \overline{wx}}$, де $\overline{xw}$ - скалярний добуток векторів $\bar{x}, \overline{w} \in$ Z$_n$, яке для функцій одного типу відрізняється несуттєво (щодо вказаних властивостей). Отже, для бульових функцій одного типу ці криптографічні властивості співпадають.

Слід відзначити, що в літературі найбільш відомою класифікацією є Гарвардський каталог, що перераховує за типами всі бульові функції від 4-х аргументів і містить інформацію про потужність типів та оптимальну лампову реалізацію функцій – представників типів. Інший, більш точний каталог, був складений групою японських вчених під керівництвом професора Нінномія. Представлений Ніконовим в [2] каталог побудований на використанні принципів класифікації бульових функцій за допомогою графів зв'язності вершин, що декілька полегшує роботу з ним. Каталог Ніконова також містить інформацію про потужність типів та мінімальну структуру реалізації бульових функцій.