CS406 Course Project

Cryptanalysis of Block Ciphers

Ankit Kumar Misra (190050020)

Kartikey Gupta (190050044)

What is Cryptanalysis?

- The mathematical study of cryptographic security systems, with the objective of finding hidden aspects which can be exploited to obtain information about the plaintext from the ciphertext without the use of a key.
- Two major types of cryptanalysis:
  - Linear Cryptanalysis
  - Differential Cryptanalysis
- Other derived types:
  - Differential-Linear Cryptanalysis
  - Non-Linear Cryptanalysis
  - Chosen Plaintext Linear Cryptanalysis
  - Partial/Truncated Differential Cryptanalysis
  - Higher Order Differential Cryptanalysis

What is Cryptanalysis?

- Usually focus on analyzing the effects of the S-boxes (substitution boxes) involved in the encryption scheme.
- **S-boxes**: Non-linear functions, usually implemented as look-up tables, that are used primarily to obscure the relationship between the key and the ciphertext.
- Map n-bit inputs to m-bit outputs.
- Several encryption schemes are linear throughout the system, except at the S-boxes.
- S-boxes add confusion (each bit depends on several parts of the key).

Linear Cryptanalysis

Linear Cryptanalysis

- Known Plaintext Attack (KPA)
- Design a linear relation, between the bits of the input plaintext and the output ciphertext, that holds with a certain probability.
- **Component-Wise Analysis**: The first step involves analyzing non-linear components (usually only S-boxes) to obtain linear relationships between their input bits and output bits.
- Suppose an S-box maps n-bit inputs to m-bit outputs.
- S-box linear relationships are of the form:

$$X_{i_1} \oplus X_{i_2} \oplus \cdots \oplus X_{i_a} \oplus \cdots \oplus Y_{j_1} \oplus Y_{j_2} \oplus \cdots \oplus Y_{j_b} = 0$$

- All $(2^n-1)(2^m-1)$ possible linear relations checked, and those retained which have a large bias, i.e., have a probability significantly larger/smaller than ½ of being true.

Linear Cryptanalysis

- **Combining Component-Wise Approximations**: Linear relations from multiple S-boxes combined (by XORing) to get possible linear approximations for the entire encryption scheme.
- Such relations contain only plaintext bits, ciphertext bits, and key bits.
- Probabilities of combinations computed using the Piling-Up Lemma.
- **Piling-Up Lemma**: Given k independent Bernoulli random variables X_i such that $Pr[X_i=0] = p_i$ for all i, we have:

$$Pr[X_1 \oplus X_2 \oplus \cdots \oplus X_k = 0] = \frac{1}{2} + 2^{k-1} \prod_{i=1}^{k} (p_i - \frac{1}{2})$$

Linear Cryptanalysis

- **Attacking the Scheme:**
  - Suppose scheme has r rounds.
  - First, design a linear approximation for first r-1 rounds.
  - Suppose subkey involved in rth round is K_r.
  - Implement rth round for all possible keys K_r in the subkey space, and in each case, check the number of plaintexts satisfying the r-1 rounds linear approximation from above.
  - The key candidate for which the fraction of satisfying ciphertexts is most distant from ½ is chosen to be K_r as it has the most distinguishing power.
  - Next, a linear approximation for the first r-2 rounds is designed, the ciphertexts are decrypted through the rth round using K_r, and the same process as above is repeated to get an estimate of K_{r-1} and so on.

Linear Cryptanalysis

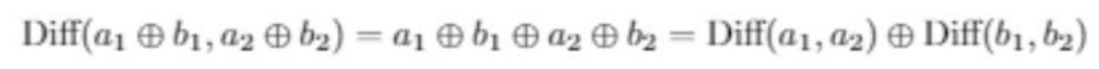
- **Attack Complexity**:
  - Number of plaintext-ciphertext pairs required for this attack is known to be of the order of 1 / (bias)^2.
  - Bias is the offset (from ½) of probability of the linear approximation used.
  - Probability of success increases with number of known pairs.

- **Prevention of Linear Cryptanalytic Attacks**:
  - Use as many S-boxes and non-linear components as possible.
  - S-box structures having lower bias of linear approximations should be used.

Differential Cryptanalysis

Differential Cryptanalysis

- Chosen Plaintext Attack (CPA)
- Analyzes the effects of differences in plaintext pairs on differences in the resulting ciphertext pairs.
- Difference of two bitstrings is simply their XOR.
- Linear components usually affect differences in an easily computable, bijective, and deterministic way. As in linear cryptanalysis, the problem is again at non-linear components, usually just S-boxes.
- E.g.: Differences can be easily propagated through an XOR operation.



$$\mathrm{Diff}(a_1 \oplus b_1, a_2 \oplus b_2) = a_1 \oplus b_1 \oplus a_2 \oplus b_2 = \mathrm{Diff}(a_1, a_2) \oplus \mathrm{Diff}(b_1, b_2)$$

Differential Cryptanalysis

- **Component-Wise Analysis**: At individual S-boxes, the distribution of output differences resulting from every possible input difference is analyzed.
- S-box differentials of the form (dX, dY, p) are thus obtained, interpreted as “an input difference of dX results in an output difference of dY with probability p”.
- Only those differentials having large probability are retained.
- **Differential Characteristics:** A differential characteristic for the encryption algorithm is of the form (dP, dC, p), interpreted as “a difference of dP in the plaintext results in a ciphertext difference of dC with probability p.
- Using the component-wise analysis from above, a differential characteristic having high probability is obtained for the encryption algorithm.
- Probability of a differential characteristic = Product of probabilities for each sub-differential used in it.

Differential Cryptanalysis

- **Attacking the Scheme:**
  - Suppose scheme has r rounds.
  - First, design a differential characteristic for the first r-1 rounds.
  - Suppose subkey involved in rth round is K_r.
  - Implement rth round for all possible keys K_r in the subkey space (for several plaintext pairs differing by the plaintext difference assumed in the characteristic), and in each case, check the number of ciphertext pairs satisfying the r-1 rounds differential characteristic from above.
  - The key candidate that satisfies the differential characteristic for the maximum number of plaintext pairs is chosen as the estimate for K_r.
  - Next, a differential characteristic for the first r-2 rounds is designed, ciphertext is decrypted through the rth round using K_r, and the same process as above is repeated to extract K_{r-1}.

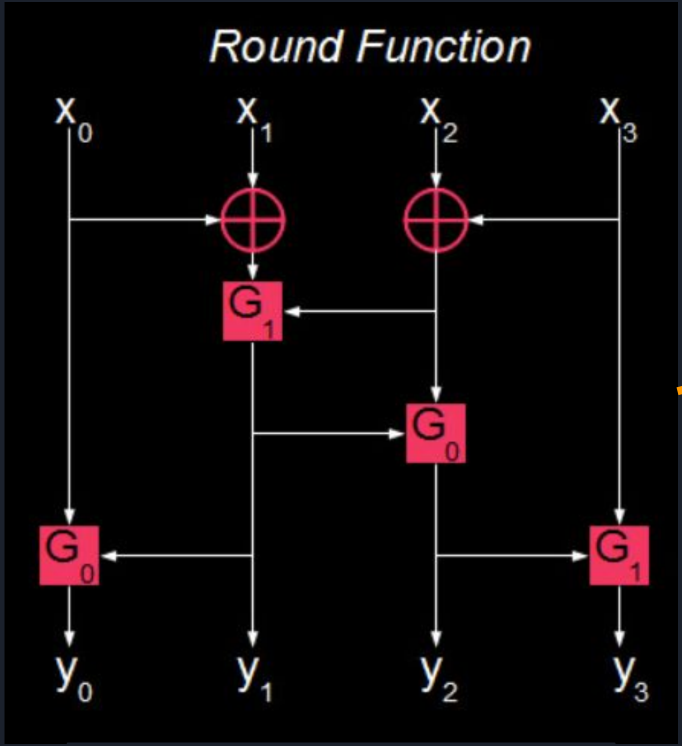
Differential Cryptanalysis

- **Attack Complexity**:
  - Number of chosen plaintext pairs required for this attack is known to be approximately c/p, where p is the probability of the differential characteristic used, and c is some small constant.

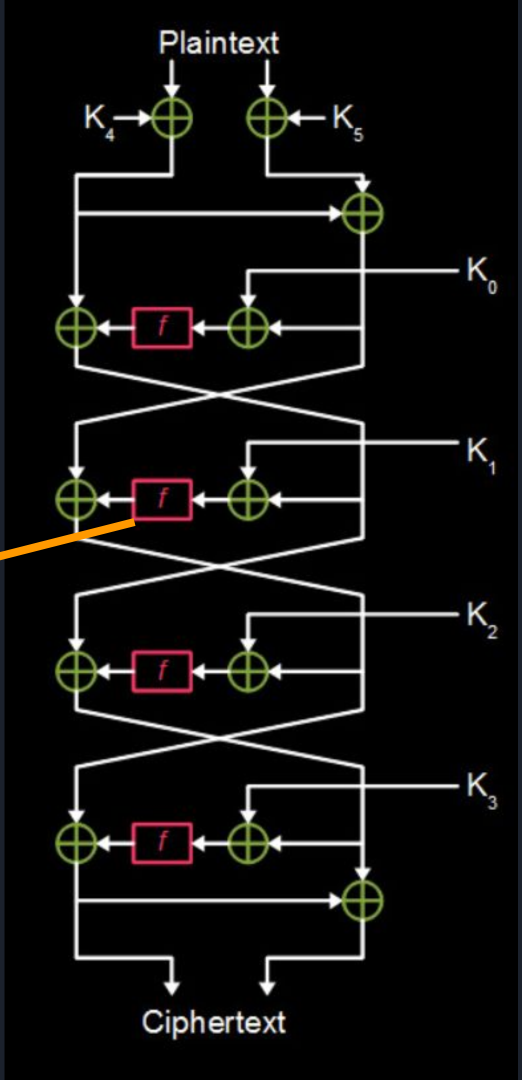
- **Prevention of Linear Cryptanalytic Attacks**:
  - Use as many S-boxes and non-linear components as possible.
  - Difference propagations within S-boxes should be kept as less probable as possible.

Cryptanalysis of the FEAL-4 Encryption Scheme

Structure of FEAL-4

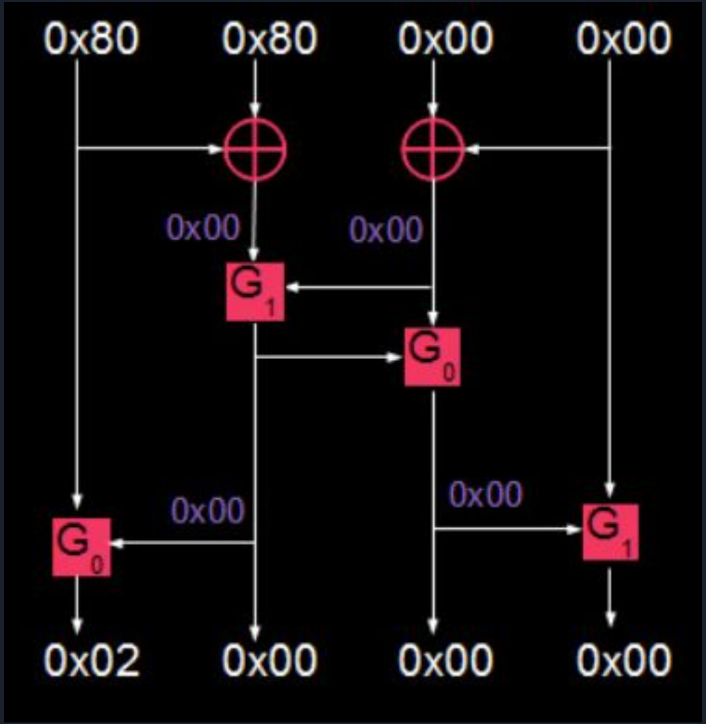


$$G_i(a,b) = ((a + b + i) \% 256) << 2$$



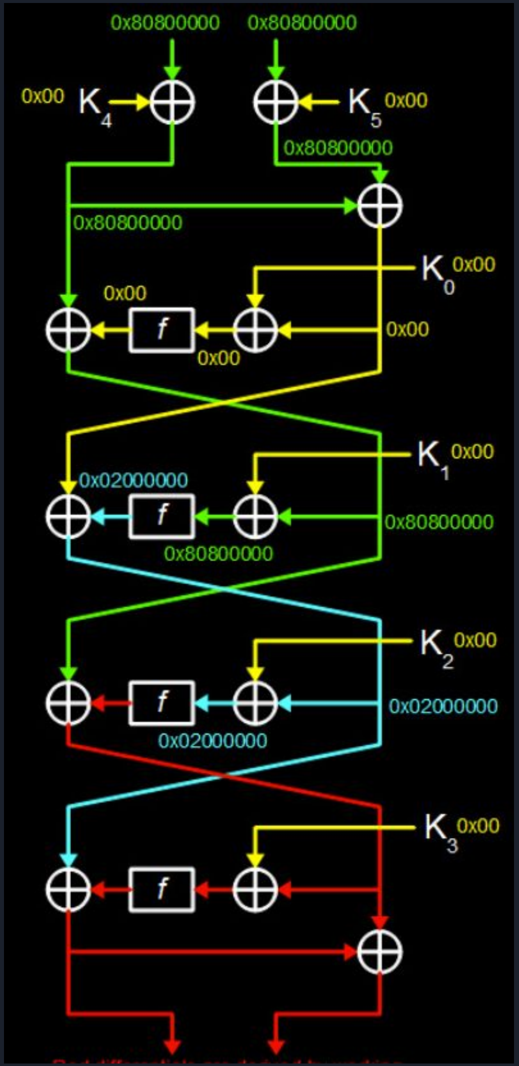
Differential Cryptanalysis of FEAL-4

- The figure on the right tracks an input difference of 0x80800000 through an f-Box of the FEAL-4 encryption scheme.
- For this input difference, an output difference of 0x02000000 is obtained with a probability of 1.
- Thus, (0x80800000, 0x02000000, 1) is a valid differential for the f-Box.
- It is easy to see that (0x00000000, 0x00000000, 1) is another valid differential for this f-Box.
- Since these hold with probability 1, they can be used to design very useful differential characteristics for the encryption scheme as a whole.



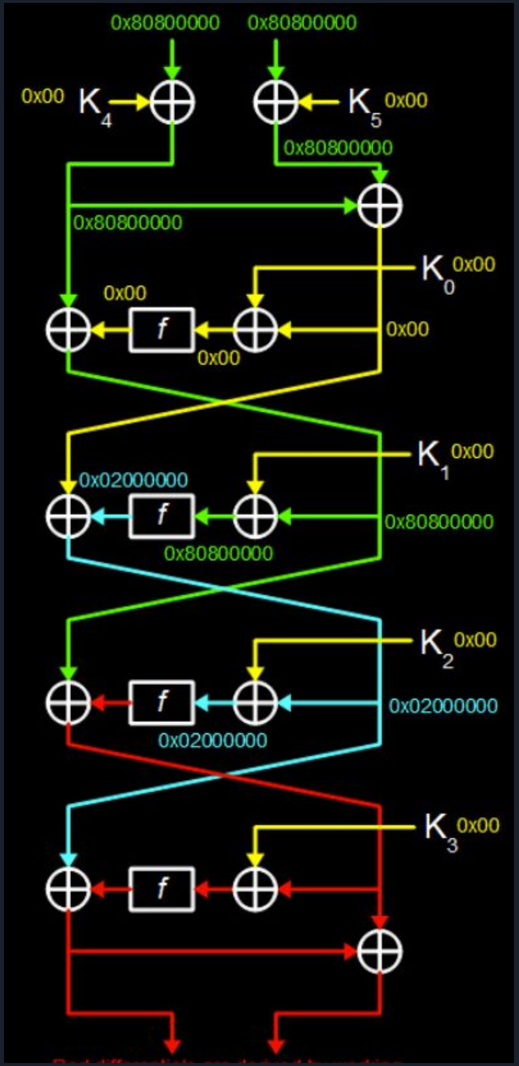
Differential Cryptanalysis of FEAL-4

- The figure on the right shows a differential characteristic for the first three rounds of the encryption scheme.
- All subkeys have been marked with 0x00 because the same keys are used for both plaintexts in the pair, so their difference becomes zero.
- Right-ciphertexts are propagated backwards from the bottom.
- A linear search over the key-space is performed for subkey K_3, computing the XOR with the bitstring at that stage, then passing it through the f-Box.
- The difference at this stage is XORed with the difference of left-ciphertexts, and then checked if it equals 0x02000000.



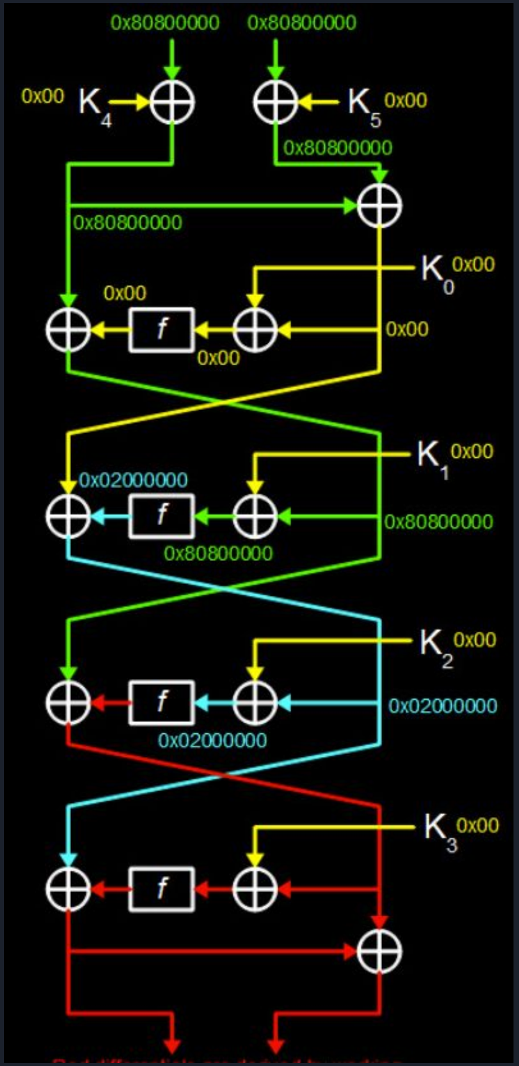
Differential Cryptanalysis of FEAL-4

- The key candidate which yields a difference of 0x02000000 for every plaintext pair tested is selected as K_3 (since the probability of the characteristic is 1; in general, the subkey at this round would be the candidate having maximum difference matches).
- Bitstrings are decrypted through K_3, and a differential characteristic for 2 rounds (plaintext difference = 0x00000000, 0x80800000) is now used to crack K_2 in the third round.
- Bitstrings are decrypted through K_2, and a differential characteristic for 2 rounds (plaintext difference = 0x00000000, 0x02000000) is now used to crack K_1 in the second round.



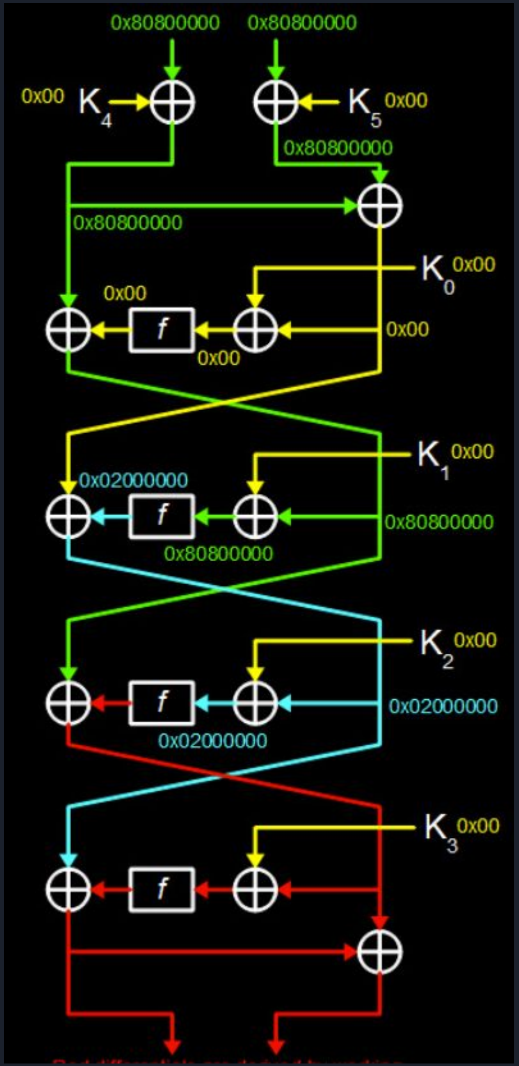
Differential Cryptanalysis of FEAL-4

- For the first round, all three of K_0, K_4, and K_5 need to be extracted. For this, we do a linear search over the 32-bit space for K_0, and for every candidate, the bitstrings are decrypted through the candidate K_0, and used to compute K_4 and K_5 by using the input plaintext.
- The candidate key which yields the same K_4 and K_5 for all plaintext pairs it is tested with is selected as K_0, and the corresponding K_4 and K_5 are evaluated.



Differential Cryptanalysis of FEAL-4

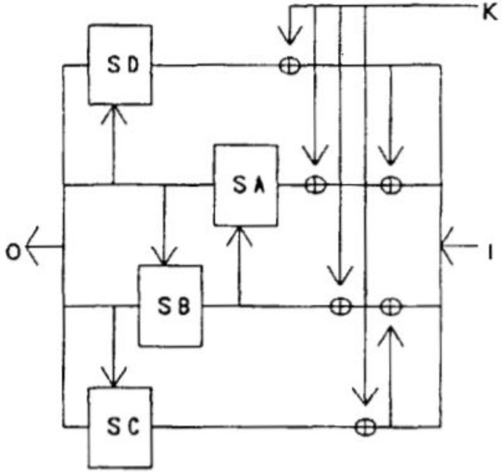
- **Efficiency**:
  - This technique is extremely better than a brute force attack on the FEAL-4 scheme.
  - A brute force attack must do a linear search over the complete key space of 6 x 32 = 192 bits, i.e., over 2^{192} possible keys, which is close to 10^{57}.
  - The technique described here must only do a linear search over four 32-bit key spaces, and individually one-by-one. This adds up to 4 x 2^{32} possible keys to be searched, which is close to 10^{10}. This is a drastic improvement!
  - Furthermore, the number of ciphertexts required for good accuracy here is only constant, as the characteristics used all hold with probability equal to 1.



Linear Cryptanalysis of FEAL-4

- Matsui and Yagamishi (1992) devised a linear cryptanalytic attack on the FEAL-4 block cipher which requires only 5 known plaintexts to extract the encryption key.
- They worked with a slightly differently structured FEAL-4 cipher, which is shown on the next slide.
- This paper invented linear cryptanalysis, applying it first to FEAL cipher family.

Linear Cryptanalysis of FEAL-4



$SA(x, y) = SC(x, y) = ROL2(x+y+1 (mod256))$

$SB(x, y) = SD(x, y) = ROL2(x+y (mod256))$

Figure 1 : Modified F function

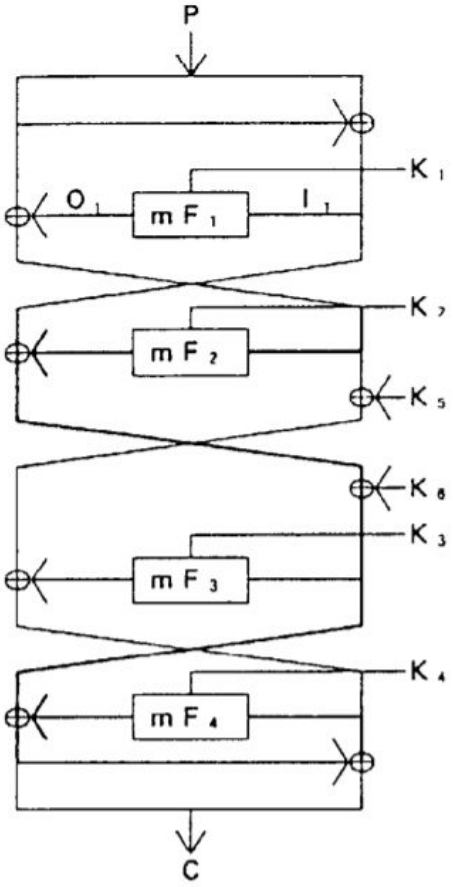
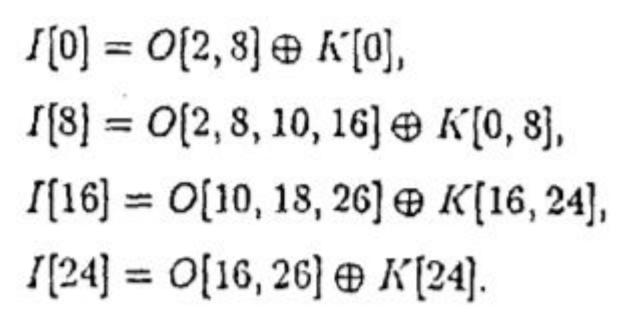


Figure 3: FEAL-4 cipher

Linear Cryptanalysis of FEAL-4

- Using the structure of the modified f-Box, the following linear relations were obtained:

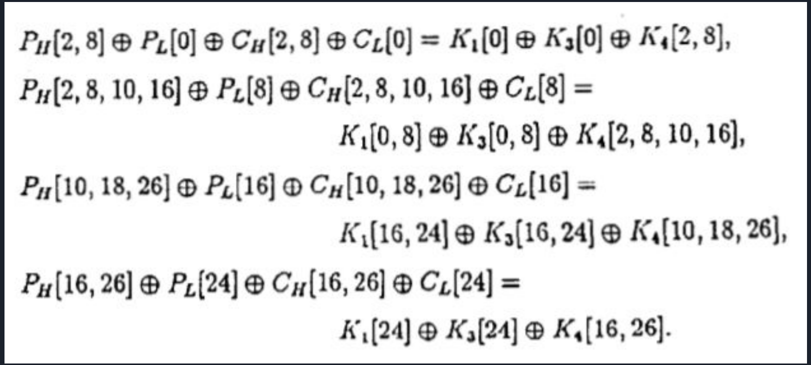


$$
\begin{aligned}
I[0] &= O[2,8] \oplus K[0], \\
I[8] &= O[2,8,10,16] \oplus K[0,8], \\
I[16] &= O[10,18,26] \oplus K[16,24], \\
I[24] &= O[16,26] \oplus K[24].
\end{aligned}
$$

- Notation: z[a,b,c] := z[a] XOR z[b] XOR z[c]
- I, O, K are the input, output, and subkey respectively.

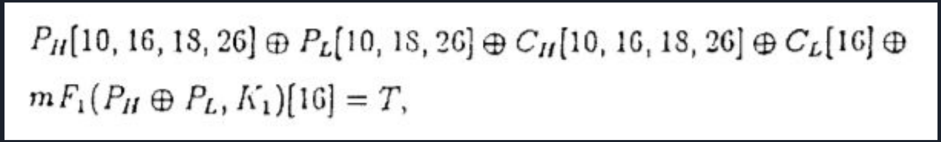
Linear Cryptanalysis of FEAL-4

- The linear relations on the previous slide were then used to produce the following relations for a 3-layer FEAL cipher:



$$\begin{aligned}
&P_H[2,8] \oplus P_L[0] \oplus C_H[2,8] \oplus C_L[0] = K_1[0] \oplus K_3[0] \oplus K_4[2,8],\\
&P_H[2,8,10,16] \oplus P_L[8] \oplus C_H[2,8,10,16] \oplus C_L[8] =\\
&\qquad\qquad K_1[0,8] \oplus K_3[0,8] \oplus K_4[2,8,10,16],\\
&P_H[10,18,26] \oplus P_L[16] \oplus C_H[10,18,26] \oplus C_L[16] =\\
&\qquad\qquad K_1[16,24] \oplus K_3[16,24] \oplus K_4[10,18,26],\\
&P_H[16,26] \oplus P_L[24] \oplus C_H[16,26] \oplus C_L[24] =\\
&\qquad\qquad K_1[24] \oplus K_3[24] \oplus K_4[16,26].
\end{aligned}$$

- These can be applied to FEAL-4 to obtain a linear relation over r-1 = 3 rounds. For example, the following relation for FEAL-4 is derived on applying the third relation above on rounds 2-4 of FEAL-4:



$$\begin{aligned}
&P_H[10,16,18,26] \oplus P_L[10,18,26] \oplus C_H[10,16,18,26] \oplus C_L[16] \oplus\\
&mF_1(P_H \oplus P_L, K_1)[16] = T,
\end{aligned}$$

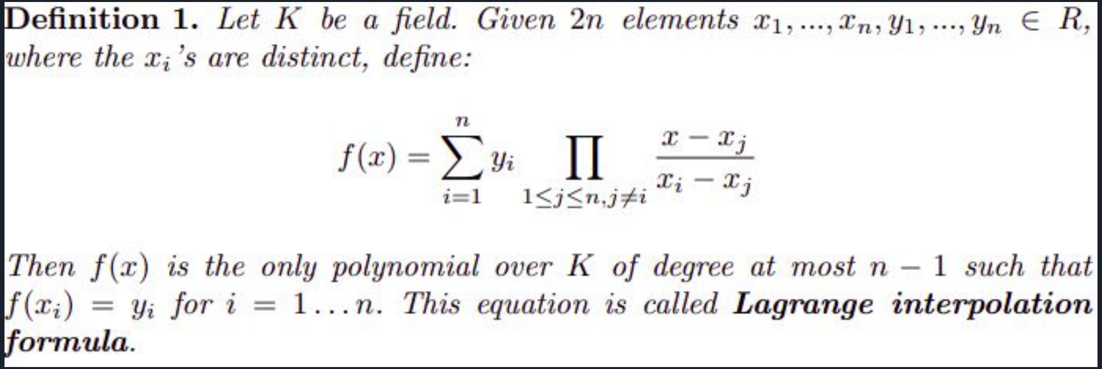
Linear Cryptanalysis of FEAL-4

- For each 3-round linear relation obtained for FEAL-4, we analyze which bits of the subkey (for the round being cracked) actually affect that relation.
- An exhaustive linear search is performed over all possible values of these particular bits, and the bits which give a constant value of the linear expression for all input plaintexts are chosen as the actual values in the subkey.
- In particular for K_1 of FEAL-4, the linear relations can be used to obtain bits 0 to 5 and bits 8 to 29, with high probability of success.
- This adds up to 28 known bits. For the remaining 4 bits, a simple brute force search suffices as there are only a few choices.
- The same procedure is repeated to extract the other subkeys.

Interpolation Attacks

Interpolation Attacks

- To attack ciphers whose round functions can be written as simple algebraic expressions
- If round function is polynomial with reasonable number of terms, then the whole cipher can be written as polynomial too with coefficients depending on keys
- Using Lagrange interpolation formula to compute coefficients



Definition 1. *Let K be a field. Given $2n$ elements $x_1, ..., x_n, y_1, ..., y_n \in R$, where the x_i's are distinct, define:*

$$f(x) = \sum_{i=1}^{n} y_i \prod_{1 \leq j \leq n, j \neq i} \frac{x - x_j}{x_i - x_j}$$

Then $f(x)$ is the only polynomial over K of degree at most $n-1$ such that $f(x_i) = y_i$ for $i = 1 \ldots n$. This equation is called ***Lagrange interpolation formula****.*

Interpolation Attacks

- It is a global deduction attack i.e., equivalent expression is constructed but key is not recovered
- Using last round key guess technique, can be converted to key recovery attack
- Attack is easily adaptable to ciphers that can be expressed as a rational expression

References

- Cryptanalysis of Block Ciphers: A Survey
- A New Method for Known Plaintext Attack of FEAL Cipher
- Differential Cryptanalysis of FEAL by Jon King
- Wikipedia, the free encyclopedia