A Few Techniques for Block Cipher Cryptanalysis

Jiqiang Lu

Institute for Infocomm Research,
Agency for Science, Technology and Research,
1 Fusionopolis Way, Singapore 138632
lvjiqiang@hotmail.com

ASK 2011

Outline:

1.1 Block Cipher

An important primitive in symmetric-key cryptography.

- An algorithm that transforms a fixed-length data block into another data block of the same length under a secret user key.
  - Input: plaintext.
  - Output: ciphertext.
  - Three sub-algorithms: encryption, decryption, key schedule.
  - Main purpose: provide confidentiality.
- Constructed by repeating a simple function many times, known as the iterated method.
  - An iteration: a round.
  - The repeated function: the round function.
  - The key used in every round: a round subkey.
  - The number of iterations: the number of rounds.
  - The round subkeys are generated from the user key under a key schedule algorithm.

Block Cipher (continued)

Round structures:

- Feistel networks:
  * The plaintext is split into two halves.
  * The round function is applied to one half.
  * The output of the round function is xored with the other half.
  * Finally, the two halves are swapped.
- Substitution-Permutation Networks (SPNs):
  * The round function is applied to the whole block.
  * The output becomes the input of the next round.
- Other round structures.

Examples

- DES (Data Encryption Standard)
  - A 64-bit Feistel cipher, a 56-bit key, 16 rounds.
  - Designed by IBM, published in 1977.
  - A former USA and ISO standard.
- AES (Advanced Encryption Standard)
  - A 128-bit SPN cipher, a key of 128, 192 or 256 bits.
  - 10 rounds for AES-128, 12 rounds for AES-192, 14 rounds for AES-256.
  - Designed by Daemen and Rijndael, first published in 1998.
  - The next-generation data encryption standard in USA, to replace DES.
  - An European NESSIE selected cipher, an ISO standard.
- Camellia
  - A 128-bit Feistel cipher, a key of 128, 192 or 256 bits.
  - 18 rounds for Camellia-128, 24 rounds for Camellia-192/256.
  - Designed by Mitsubishi and NTT, first published in 2000.
  - An European NESSIE selected cipher, an ISO standard.

Examples (continued)

- MISTY1
  - A 64-bit Feistel cipher, a 128-bit key, 8 rounds.
  - Designed by Matsui, first published in 1997.
  - An European NESSIE selected cipher, an ISO standard.
- SHACAL-2
  - A 256-bit (generalised) Feistel cipher, a key of up to 512 bits, 64 rounds.
  - Based on the SHA-256 hash function.
  - Designed by Handschuh and Naccache, first published in 2001.
  - An European NESSIE selected cipher.
- Serpent
  - A 128-bit SPN cipher, a key of up to 256 bits, 32 rounds.
  - Designed by Anderson, Biham and Knudsen, first published in 1998.
  - One of the five AES finalists, second to the Rijndael cipher.
- CTC2
  - A toy SPN cipher, variable block size/key/number of rounds.
  - Designed by Courtois to show the strength of algebraic cryptanalysis on block ciphers, first presented in 2007.

1.2 A Cryptanalytic Attack

- An algorithm that distinguishes a cryptosystem from a random function.
- Usually measured using the following three metrics:
  * Data complexity
    - The numbers of plaintexts and/or ciphertexts required.
  * Memory (storage) complexity
    - The amount of memory required.
  * Time (computational) complexity
    - The amount of computation or time required, how many encryptions/decryptions or memory accesses.

1.3 Four Cryptanalysis Scenarios

- Ciphertext-only attack scenario
  - Have access to a number of ciphertexts.
- Known-plaintext attack scenario
  - Have access to a number of ciphertexts and the corresponding plaintexts.
- Chosen-plaintext/cipertext attack scenario
  - Can choose a number of plaintexts (or ciphertexts), and be given the corresponding ciphertexts (or plaintexts).
- Adaptive chosen plaintext and ciphertext attack scenario
  - Can choose plaintexts (or ciphertexts) and be given the corresponding ciphertexts (or plaintexts). Based on the information obtained, the attacker can then choose further plaintexts/ciphertexts, and be given the corresponding ciphertexts/plaintexts ...

1.4 Three Elementary Cryptanalysis Techniques

Assume an n-bit block cipher with a k-bit user key $E_K(\cdot)$.

- A dictionary attack
  - Build a table of all possible ciphertexts corresponding to one particular plaintext, with one entry for each possible key: $C_i = E_{K_i}(P)$.
  - Data: 2^k ciphertexts, Memory: 2^k n-bit, Time: negligible.
- A codebook attack:
  - Build a table of the ciphertexts for all the plaintexts encrypted using one unknown key: $C_i = E_K(P_i)$.
  - Data: 2^n plaintext-ciphertext pairs, Memory: 2^n n-bit, Time: negligible.
- An exhaustive key search (or brute force search) attack:
  - Try every possible key, given a known plaintext-ciphertext pair. The correct key will yield the correct correspondence: $E_{K_i}(P) \stackrel{?}{\rightarrow} C$.
  - Data: negligible, Memory: negligible, Time: 2^k encryptions.

1.5 Advanced Cryptanalysis Techniques

An attack is commonly regarded as effective if it is faster than an exhaustive key search.

A trade-off between data, time and/or memory.

- Meet-in-the-middle attack
  - The reflection-meet-in-the-middle attack
- Differential cryptanalysis
  - Truncated, Higher-order, Impossible differential cryptanalysis
  - Boomerang, Amplified boomerang, Rectangle attacks
- Linear cryptanalysis
- Differential-linear cryptanalysis
- Integral cryptanalysis
  - Square attack, Saturation attack
- Slide attack, Reflection attack
- Related-key attack
- Algebraic cryptanalysis

2.1 Differential Cryptanalysis

Take advantage of how a specific difference in a pair of plaintexts can affect a difference in the pair of ciphertexts.

- Introduced by Biham and Shair in 1990.
- A differential is the combination of the input difference and the output difference.
- The probability of the differential (α, β) for a block cipher $\mathbb{E}$, written $\Delta\alpha \rightarrow \Delta\beta$, is

$$\Pr_{\mathbb{E}}(\Delta\alpha \rightarrow \Delta\beta) = \Pr_{P \in \{0,1\}^n}(\mathbb{E}(P) \oplus \mathbb{E}(P \oplus \alpha) = \beta).$$

- For a random function, the expected probability of any differential is 2^{-n}.

Thus, if $\Pr_{\mathbb{E}}(\Delta\alpha \rightarrow \Delta\beta)$ is larger than 2^{-n}, we can use the differential to distinguish $\mathbb{E}$ from a random function.

2.2 Impossible Differential Cryptanalysis

A special case of differential cryptanalysis.

- Independently introduced by Knudsen in 1998 and Biham, Biryukov and Shamir in 1999.
- An impossible differential is a differential with a zero probability.

Thus:

- If there exists a plaintext-ciphertext pair for a key guess, then the key guess must be incorrect.
- Find the correct key by discarding all wrong keys, given a sufficient number of chosen plaintext pairs.

A Typical Key Recovery Attack

Treat a block cipher $\mathbb{E} : \{0,1\}^n \times \{0,1\}^k \rightarrow \{0,1\}^n$ as $\mathbb{E} = \mathbb{E}^a \circ \mathbb{E}^0 \circ \mathbb{E}^b$:

- $\mathbb{E}^0$ denotes the rounds for which an impossible differential $\Delta\alpha \nrightarrow \Delta\beta$ holds.
- $\mathbb{E}^a$ denotes the rounds before $\mathbb{E}^0$.
- $\mathbb{E}^b$ denotes the rounds after $\mathbb{E}^0$.

Given a candidate subkey (K_a, K_b) for $\mathbb{E}^a$ and $\mathbb{E}^b$, check whether there is a plaintext-ciphertext pair $((P, C), (P', C'))$ satisfying the following two conditions simultaneously:

$$\mathbb{E}^a_{K_a}(P) \oplus \mathbb{E}^a_{K_a}(P') = \alpha,$$
$$(\mathbb{E}^b_{K_b})^{-1}(C) \oplus (\mathbb{E}^b_{K_b})^{-1}(C') = \beta.$$

When checking whether a plaintext-ciphertext pair satisfies the two conditions, a general approach:

1. Guess all the required unknown subkey bits of a round necessary to partially encrypt/decrypt the pair.
2. Check whether the pair produces the expected difference just after/before the round.

2.3 Early Abort Technique

However, we find [LKKD:CT-RSA2008]:

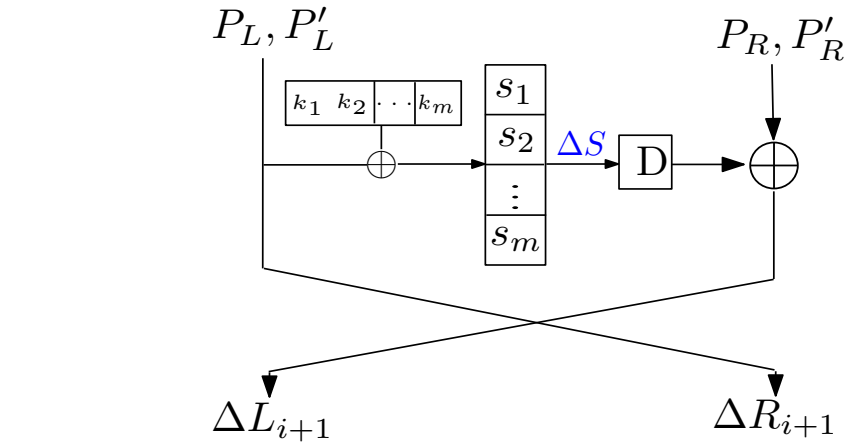
- Depending on the round structure, we can partially check whether the pair could produce the expected difference by guessing only a small fraction of the required round subkey bits at a time, and do a series of partial checks by guessing other fractions of the required round subkey bits, instead of all of them simultaneously.

Some invalid candidate pairs can be discarded after each guess, thus

- Reduce an attack's computational workload, and;
- May break more rounds.

2.4.1 Application to Feistel Cipher Camellia

- Known: $\Delta\alpha$, (L_i, R_i) for plaintext P, (L'_i, R'_i) for plaintext P'.
- Task: Check whether $(\Delta L_{i+1}, \Delta R_{i+1}) = \Delta\alpha$.



- General approach:
  1. Guess all the subkey bits corresponding to those active S-boxes;
  2. Decrypt the pair through the round;
  3. Check whether producing the expected difference $\Delta\alpha$.
- Early abort:
  1. Compute the expected difference just before the **D** function: $\Delta S = \mathbf{D}^{-1}(P_R \oplus P'_R \oplus \alpha_L)$.
  2. A partial check:
     I. Guess only the fraction of the required unknown subkey bits corresponding to one (or more) active S-box;
     II. Check whether producing the corresponding partial difference of ΔS.
  3. Another partial check, ...

A pair is a valid candidate only if it produces the corresponding partial difference of ΔS under each part of the required set of subkey bits.

Cryptanalytic Results

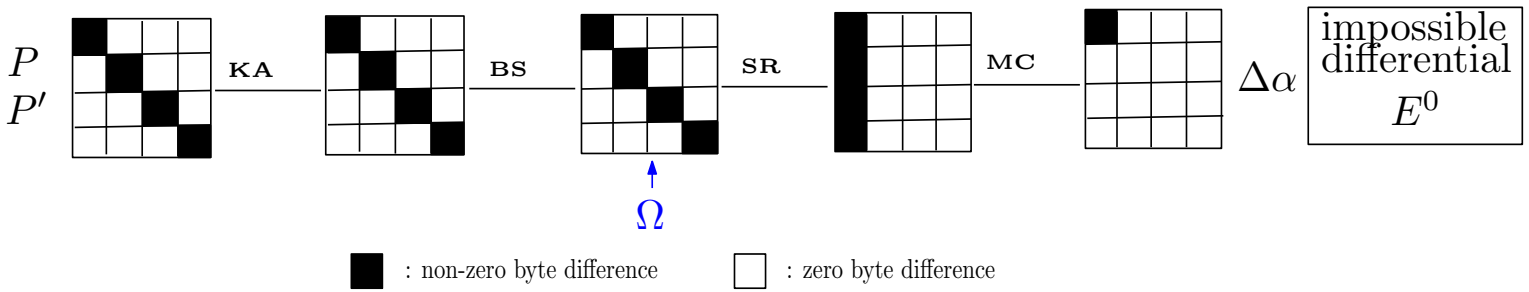
In 2007, Wu et al. presented impossible differential attacks on 12-round Camellia-192/256.

Applying the early abort technique, we can easily

- Improve Wu et al.'s 12-round Camellia-256 attack to break 13-round Camellia-256.
- Reduce the time complexity of Wu et al.'s 12-round Camellia-192 attack from 2^{181} to $2^{147.3}$.
- Attack 11-round Camellia-128.
- More results if considering the key schedule of Camellia.

2.4.2 Application to SPN Cipher AES

- Known: $\Delta\alpha$, plaintext P, plaintext P'.
- Task: Check whether the pair produces $\Delta\alpha$.



- General approach:
  1. Guess the four active subkey bytes;
  2. Decrypt the pair through the round;
  3. Check whether producing the expected difference $\Delta\alpha$.
- Early abort:
  1. Compute the expected differences just before the **SR** operation: Ω. (At most 255 values.)
  2. A partial check:
     I. Guess two of the four active subkey bytes;
     II. Check whether matching the corresponding partial difference of one in Ω.
  3. A partial check on one of the remaining two bytes.
  4. A partial check on the last one byte.

Cryptanalytic Results

In 2004, Phan presented impossible differential attacks on 7-round AES-192/256.

Applying the early abort technique, we can easily

- Improve Phan's 7-round AES-256 attack to break 8-round AES-256.
- Reduce the time complexity of Phan et al.'s 7-round AES-192 attack from 2 to 2.
- More results if considering the key schedule of AES.

2.4.3 Application to Other Structures

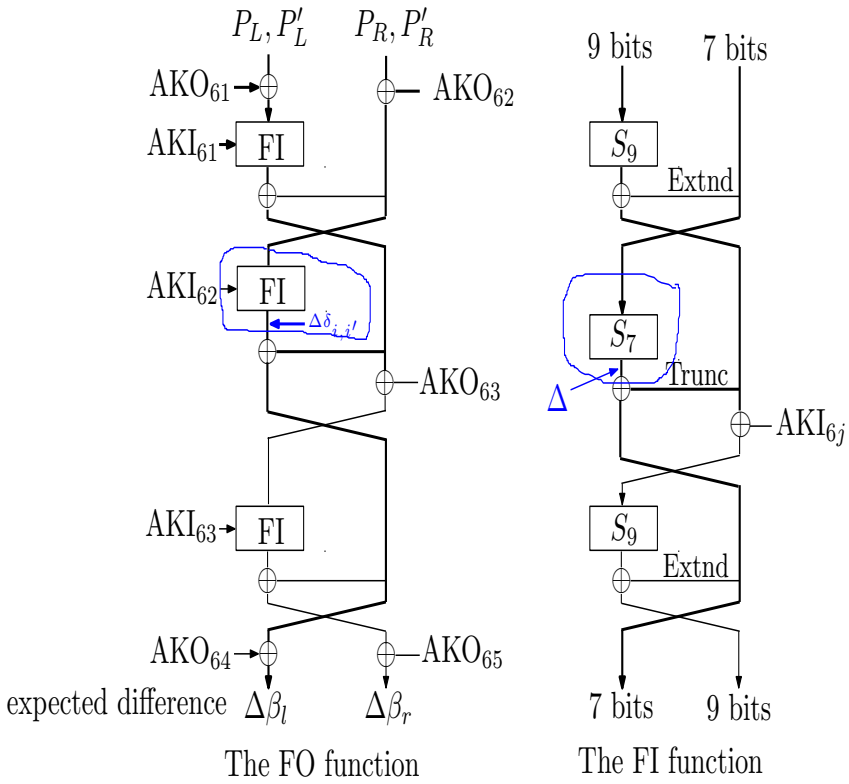
MISTY1:

In 2001, Kuhn presented an impossible differential attack on 6-round MISTY1 without FL functions.

Applying the early abort technique, we can

- Reduce the time complexity of Kuhn's attack from 2^{106} to 2^{85}.
- Improve Kuhn's attack to break 7-round MISTY1 (without FL functions) by also considering the key schedule of MISTY1.

An example of the early abort technique in MISTY1



2.5 Application in Other Cryptanalytic Techniques

Used in other cryptanalytic approaches to improve efficiency:

- Differential cryptanalysis and its extensions.
- Meet-in-the-middle attacks.
  - The first non-trivial attack on 8-round AES-192 (BDK:ASIACRYPT2010).

3.1 Boomerang Attack

An extension of differential cryptanalysis.

- Introduced by Wagner in 1999.
- Work in an adaptive chosen plaintext and ciphertext attack scenario.
- Based on the use of a boomerang distinguisher.

A Boomerang Distinguisher

- Treat a block cipher $\mathbb{E} : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$ as $\mathbb{E} = \mathbb{E}^0$ and $\mathbb{E}^1$.
- Defined to be a pair of differentials $(\Delta\alpha \to \Delta\beta, \Delta\gamma \to \Delta\delta)$.
  - $\Delta\alpha \to \Delta\beta$ for $\mathbb{E}^0$ with probability p;
  - $\Delta\gamma \to \Delta\delta$ for $\mathbb{E}^1$ with probability q.
- Concerned event: $\mathbb{E}^{-1}(\mathbb{E}(P) \oplus \delta) \oplus \mathbb{E}^{-1}(\mathbb{E}(P \oplus \alpha) \oplus \delta) = \alpha$.
- Probability: p^2q^2 approximately (under assumptions).
- For a random function, the expected probability of any boomerang distinguisher is 2^{-n}.

Thus, if $p^2q^2 > 2^{-n}$, we can distinguish between $\mathbb{E}$ and a random function.

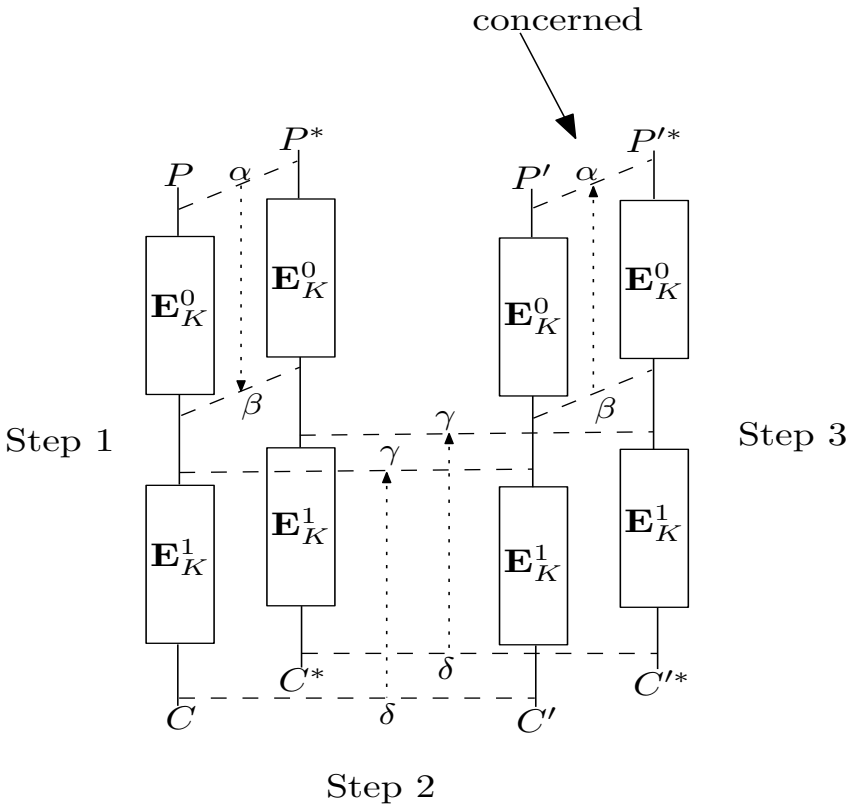


Figure: A boomerang distinguisher

3.2 Amplified Boomerang Attack

A variant of the boomerang attack.

- Introduced by Kelsey, Kohno and Schneier in 2000.
- Work in a chosen plaintext/ciphertext attack scenario.
- Based on the use of an amplified boomerang distinguisher.

An Amplified Boomerang Distinguisher

- Treat a block cipher $\mathbb{E} : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$ as $\mathbb{E} = \mathbb{E}^0 \circ \mathbb{E}^1$.
- Defined to be a pair of differentials $(\Delta\alpha \to \Delta\beta, \Delta\gamma \to \Delta\delta)$:
  - $\Delta\alpha \to \Delta\beta$ for $\mathbb{E}^0$ with probability p;
  - $\Delta\gamma \to \Delta\delta$ for $\mathbb{E}^1$ with probability q.
- Concerned event: $\mathbb{E}(P) \oplus \mathbb{E}(P') = \delta$ and $\mathbb{E}(P \oplus \alpha) \oplus \mathbb{E}(P' \oplus \alpha) = \delta$
- Probability: $p^2q^2 2^{-n}$ approximately (under assumptions).
- For a random function, the expected probability of any boomerang distinguisher is 2^{-2n}.

Thus, if $p^2q^2 > 2^{-n}$, we can distinguish between $\mathbb{E}$ and a random function.

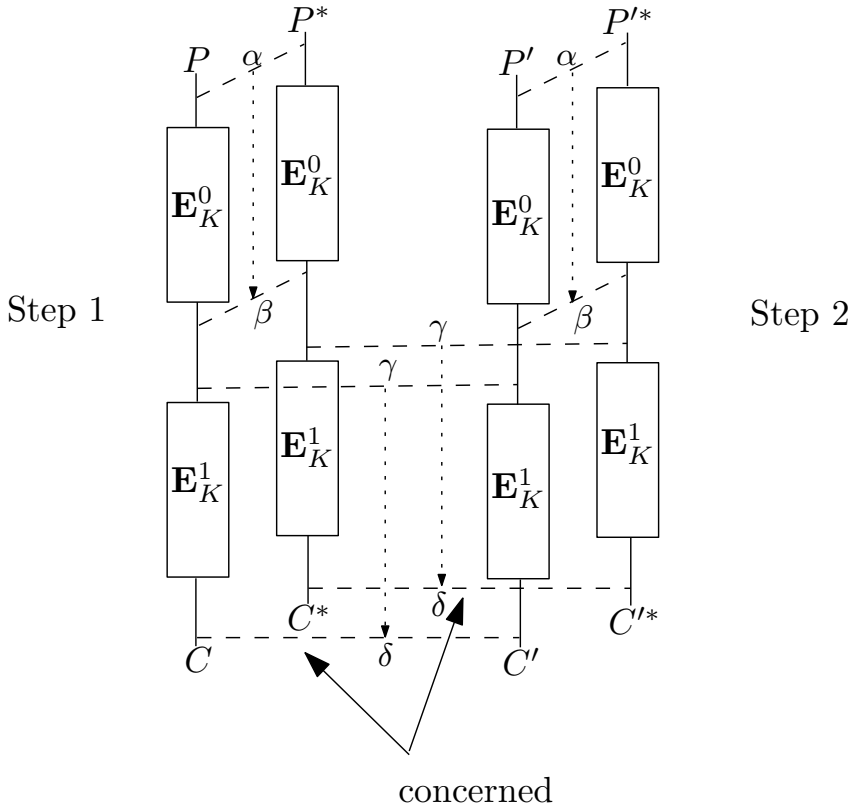


Figure: An amplified boomerang distinguisher

3.3 Rectangle Attack

The amplified boomerang attack using many possible differences for β and γ.

- Renamed as the rectangle attack by Biham, Dunkelman and Keller in 2001.
- Probability: $(\widehat{p} \cdot \widehat{q})^2 \cdot 2^{-n}$,
  - $\widehat{p} = \sqrt{\sum_{\beta'} \mathrm{Pr}^2_{\mathbb{E}^0_K}(\Delta\alpha \rightarrow \Delta\beta')}$;
  - $\widehat{q} = \sqrt{\sum_{\gamma'} \mathrm{Pr}^2_{\mathbb{E}^1_K}(\Delta\gamma' \rightarrow \Delta\delta)}$.

3.4 Related-Key Rectangle Attack

A combination of the rectangle attack and related-key cryptanalysis.

- Related-key cryptanalysis:
  - Independently introduced by Knudsen in 1992 and Biham in 1993.
  - Takes advantage of how a specific difference in a pair of plaintexts can affect a difference in the pair of ciphertexts, where the pair of ciphertexts are obtained by encrypting the pair of plaintexts using two different keys with a specific difference.
  - Probability of a related-key differential:

$$\Pr_{\mathbb{E}_K, \mathbb{E}_{K'}}(\Delta\alpha \rightarrow \Delta\beta) = \Pr_{P \in \{0,1\}^n}(\mathbb{E}_K(P) \oplus \mathbb{E}_{K'}(P \oplus \alpha) = \beta).$$

- Based on the use of a related-key rectangle distinguisher.

A Related-Key Rectangle Distinguisher

- Treat a block cipher $\mathbb{E}$ as $\mathbb{E} = \mathbb{E}^0 \circ \mathbb{E}^1$.
- Work typically in a related-key attack scenario with four related keys K_A, K_B, K_C, K_D:
  - $K_A \oplus K_B = K_C \oplus K_D$;
  - $K_A \oplus K_C = K_B \oplus K_D$.
- Consist of four related-key differentials.
- Concerned event: $\mathbb{E}_{K_A}(P) \oplus \mathbb{E}_{K_C}(P') = \delta$ and $\mathbb{E}_{K_B}(P \oplus \alpha) \oplus \mathbb{E}_{K_D}(P' \oplus \alpha) = \delta$.
- Probability: $\widehat{p}^2\widehat{q}^2 2^{-n}$ approximately (under assumptions).
- For a random function, the expected probability of any related-key rectangle distinguisher is 2^{-2n}.

Thus, if $\widehat{p}^2\widehat{q}^2 > 2^{-n}$, we can distinguish between $\mathbb{E}$ and a random function.

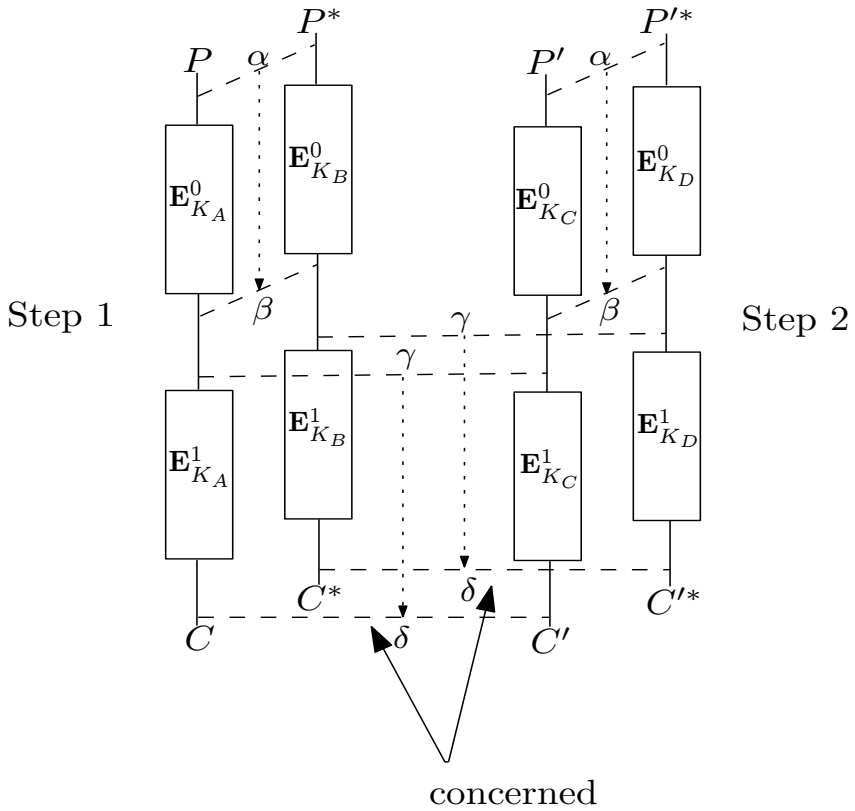


Figure: A related-key rectangle distinguisher

A Typical Key Recovery Attack

Treat a block cipher $\mathbb{E}$ as $\mathbb{E} = \mathbb{E}^a \circ \mathbb{E}^0 \circ \mathbb{E}^1 \circ \mathbb{E}^b$:

- $\mathbb{E}^0 \circ \mathbb{E}^1$ denotes the rounds for which a rectangle distinguisher holds.
- $\mathbb{E}^a$ denotes the rounds before $\mathbb{E}^0$.
- $\mathbb{E}^b$ denotes the rounds after $\mathbb{E}^0$.

Given a candidate subkey (K_a, K_b) for $\mathbb{E}^a$ and $\mathbb{E}^b$, check whether a plaintext-ciphertext quartet (consisting of two plaintext/ciphertext pairs) $((P, C), (P^*, C^*), (P', C'), (P'^*, C'^*))$ satisfies:

$$\mathbb{E}^a_{K_A^a}(P) \oplus \mathbb{E}^a_{K_B^a}(P^*) = \mathbb{E}^a_{K_C^a}(P') \oplus \mathbb{E}^a_{K_D^a}(P'^*) = \alpha,$$
$$(\mathbb{E}^b_{K_A^b})^{-1}(C) \oplus (\mathbb{E}^b_{K_C^b})^{-1}(C') = (\mathbb{E}^b_{K_B^b})^{-1}(C^*) \oplus (\mathbb{E}^b_{K_D^b})^{-1}(C'^*) = \delta.$$

When checking whether a quartet satisfies the two conditions, a general approach:

1. Guess all the required unknown subkey bits of one or more relevant rounds necessary to partially encrypt/decrypt the quartet.
2. Check whether the quartet produces the expected differences just after/or before the rounds.

3.5 The Early Abort Technique

However, we find [LKKL:ISC2006, LK:IEICE2008]:

I. May partially determine whether a candidate quartet is valid one or more rounds earlier than usual.

II. Check the two plaintext/ciphertext pairs from a quartet one after the other, instead of checking them simultaneously.

III. Can use the early abort technique for impossible differential cryptanalysis when checking each pair.

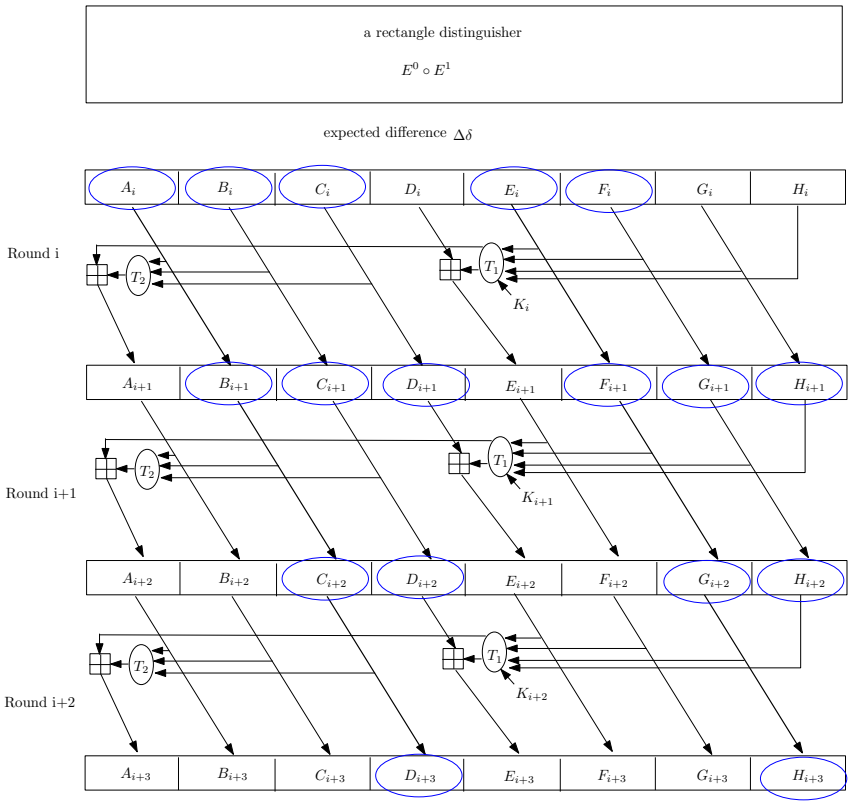
Some invalid candidate quartets can be discarded earlier, thus

- Reduce an attack's computational workload, and;
- May break more rounds.

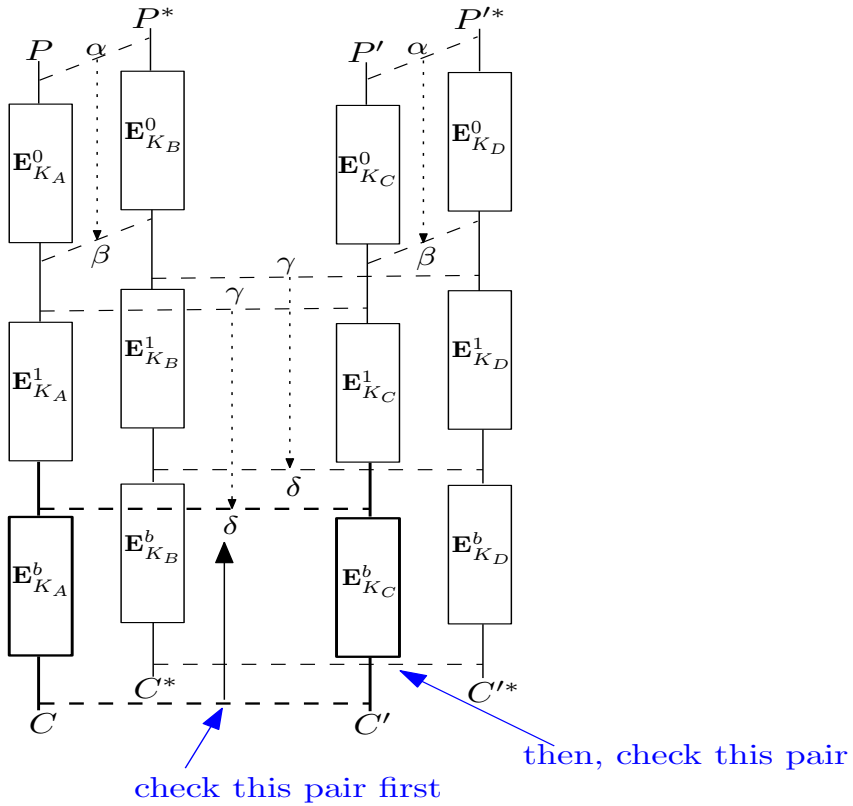
3.6 Application to SHACAL-2

- A 35-round related-key rectangle distinguisher with probability 2^{-460}.
- Observation I allows us to break 43 rounds.
- Observation II allows us to break one more round.
- In total, we can break 44 rounds.

An Example of Observation I



An Example of Observation II



4.1 Impossible Boomerang Attack

An extension of the boomerang attack.

- Work in a chosen plaintext/ciphertext attack scenario.
- Based on the use of an impossible boomerang distinguisher:
  - A combination of four n-bit blocks $\alpha, \alpha', \delta, \delta'$, such that any plaintext pair (X, X') cannot simultaneously meet

$$\mathbb{E}_K(X) \oplus \mathbb{E}_K(X') = \delta,$$
$$\mathbb{E}_K(X \oplus \alpha) \oplus \mathbb{E}_K(X' \oplus \alpha') = \delta'.$$

An Impossible Boomerang Distinguisher

- Treat a block cipher $\mathbb{E} : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$ as $\mathbb{E} = \mathbb{E}^0 \circ \mathbb{E}^1$.
- Consist of four differentials with probability 1:
  - $\Delta\alpha \to \Delta\beta$, $\Delta\alpha' \to \Delta\beta'$ for $\mathbb{E}^0$;
  - $\Delta\delta \to \Delta\gamma$, $\Delta\delta' \to \Delta\gamma'$ for $(\mathbb{E}^1)^{-1}$;
  - $\beta \oplus \beta' \oplus \gamma \oplus \gamma' \neq 0$.
- Probability:

$$\Pr_{P,P' \in \{0,1\}^n}(\mathbb{E}(P) \oplus \mathbb{E}(P') = \delta, \mathbb{E}(P \oplus \alpha) \oplus \mathbb{E}(P' \oplus \alpha') = \delta') = 0.$$

Thus, we can distinguish between $\mathbb{E}$ and a random function.

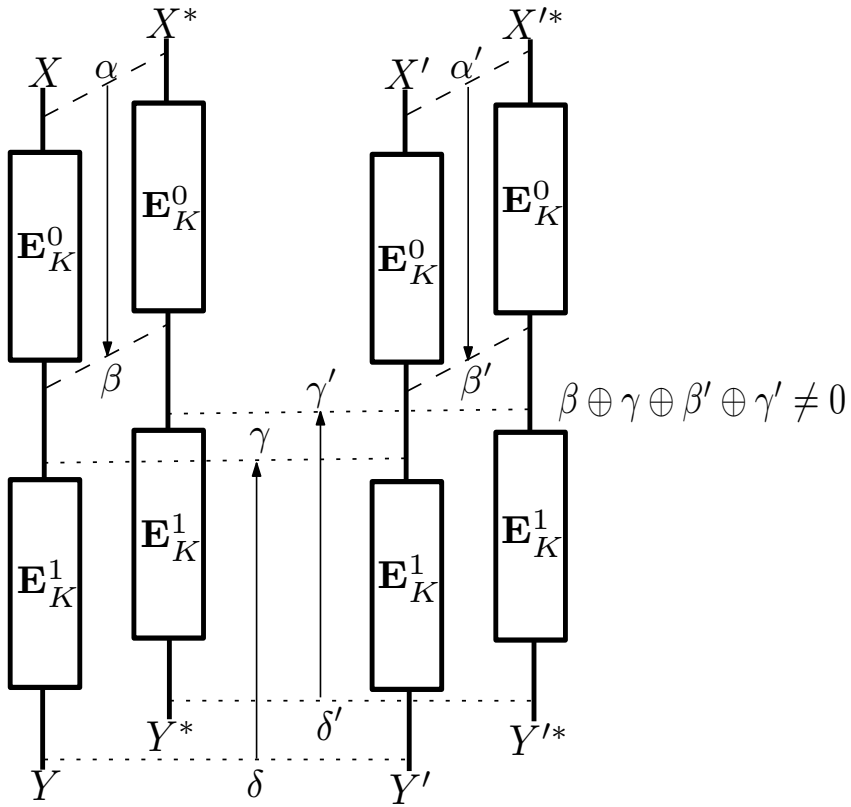


Figure: An impossible boomerang distinguisher

4.2 Related-Key Impossible Boomerang Attack

A combination of the impossible boomerang attack and related-key cryptanalysis.

- Work typically in a related-key attack scenario with four related keys K_A, K_B, K_C, K_D.
- Based on the use of a related-key impossible boomerang distinguisher:
  - A combination of four n-bit blocks $\alpha, \alpha', \delta, \delta'$, such that any plaintext pair (X, X') cannot simultaneously meet

$$\mathbb{E}_{K_A}(X) \oplus \mathbb{E}_{K_C}(X') = \delta,$$
$$\mathbb{E}_{K_B}(X \oplus \alpha) \oplus \mathbb{E}_{K_D}(X' \oplus \alpha') = \delta'.$$

A Related-Key Impossible Boomerang Distinguisher

- Treat a block cipher $\mathbb{E} : \{0,1\}^k \times \{0,1\}^n \rightarrow \{0,1\}^n$ as $\mathbb{E} = \mathbb{E}^0 \circ \mathbb{E}^1$.
- Consist of four related-key differentials with probability 1.
- Probability:

$$\Pr_{P,P' \in \{0,1\}^n}(\mathbb{E}_{K_A}(P) \oplus \mathbb{E}_{K_C}(P') = \delta, \mathbb{E}_{K_B}(P \oplus \alpha) \oplus \mathbb{E}_{K_D}(P' \oplus \alpha') = \delta') = 0.$$

Thus, we can distinguish between $\mathbb{E}$ and a random function.

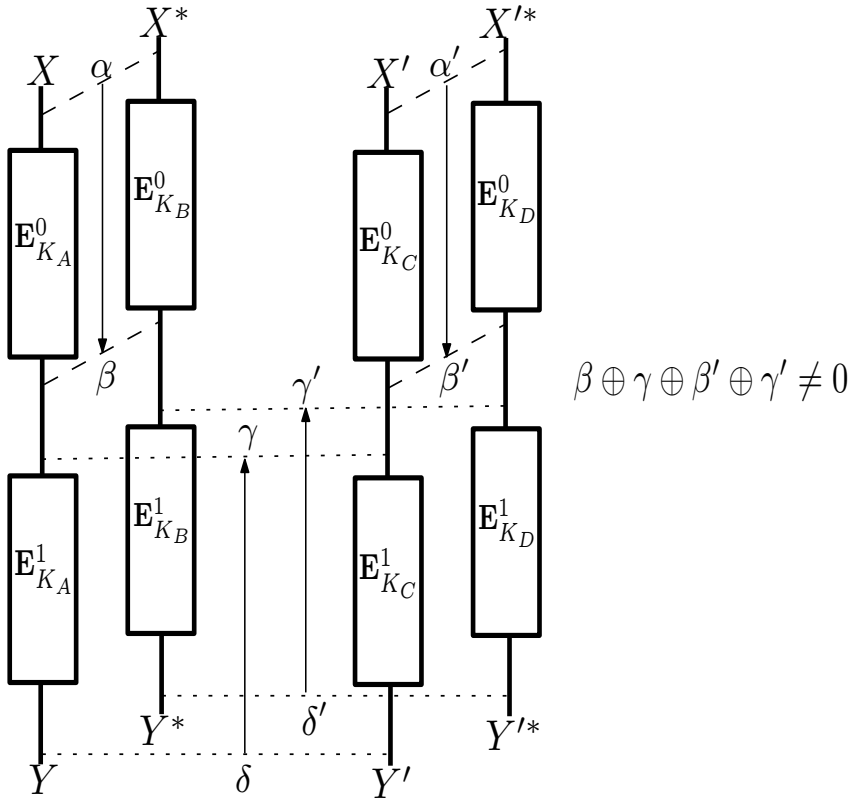


Figure: A related-key impossible boomerang distinguisher

4.3 A Comparison

1. From an impossible-boomerang distinguisher, an impossible differential for the same number of rounds can be obtained.

$$\beta \oplus \beta' \oplus \gamma \oplus \gamma' \neq 0 \Rightarrow \beta \oplus \gamma \neq \beta' \oplus \gamma'.$$

2. A block cipher resistant to related-key impossible differential cryptanalysis will not necessarily resist a related-key impossible boomerang attack.
3. A block cipher resistant to the boomerang-type attacks will not necessarily resist a (related-key) impossible boomerang attack.
4. A (related-key) impossible-boomerang distinguisher is more reasonable than the boomerang-type distinguishers.

4.4 Application to AES

- 4-round impossible-boomerang distinguishers of AES:
  - Break 6-round AES-128;
  - Break 7-round AES-192;
  - Break 7-round AES-256.
- 6-round related-key impossible-boomerang distinguishers of AES-192 under two related keys:
  - Break 8-round AES-192 in a related-key attack scenario with two keys.
- 6-round related-key impossible-boomerang distinguishers of AES-256 under two related keys:
  - Break 9-round AES-256 in a related-key attack scenario with two keys.

5.1 Linear Cryptanalysis

Exploit correlations between a particular linear function of the plaintexts and a second linear function of the ciphertexts.

- Introduced by Matsui and Yamagishi in 1992.
- A linear expression is the combination of the two linear functions.
- The probability of the linear expression (α, β) for a block cipher $\mathbb{E}$, written $\Gamma\alpha \rightarrow \Gamma\beta$, is defined to be

$$\Pr_{\mathbb{E}}(\Gamma\alpha \rightarrow \Gamma\beta) = \Pr_{P \in \{0,1\}^n}(P \odot \alpha = \mathbb{E}(P) \odot \beta).$$

- For a random function, the expected probability of a linear expression for any pair (α, β) is $\frac{1}{2}$.

Thus, if the bias $\epsilon = |\Pr_{\mathbb{E}}(\Gamma\alpha \rightarrow \Gamma\beta) - \frac{1}{2}|$ is sufficiently large, we can distinguish $\mathbb{E}$ from a random function.

5.2 Langford and Hellman's Methodology

- Introduced in 1994.
- Treat $\mathbb{E}$ as a cascade of two sub-ciphers $\mathbb{E} = \mathbb{E}_0 \circ \mathbb{E}_1$.
- Use a linear expression $\Gamma\gamma \rightarrow \Gamma\delta$ with bias ϵ for $\mathbb{E}_1$.
- Use a differential $\Delta\alpha \rightarrow \Delta\beta$ with probability 1 for $\mathbb{E}_0$, which has a zero output difference in the bits concerned by $\Gamma\gamma$.
- Concerned event: $\delta \odot \mathbb{E}(P) = \delta \odot \mathbb{E}(P \oplus \alpha)$.
- Probability: $\frac{1}{2} + 2\epsilon^2$ under Assumptions (I) and (II).
  - I. The round keys are independent and uniformly distributed.
  - II. The two inputs of the linear expression(s) for $\mathbb{E}_1$ have an independent propagation.

Thus, if the bias $2\epsilon^2$ is sufficiently large, we can distinguish $\mathbb{E}$ from a random function.

5.3 Biham, Dunkelman and Keller's Methodology

- Aim to make a distinguisher cover more rounds of the cipher.
- Introduced in 2001.
- Use a differential $\Delta\alpha \rightarrow \Delta\beta$ with probability p for $\mathbb{E}_0$, with $\beta \odot \gamma = 0$. $(p \leq 1)$
- Probability: $\frac{1}{2} + 2p\epsilon^2$ under Assumptions (I), (II) and (III).

  III. For the possible output differences of $\mathbb{E}_0$ under the input difference α excluding β, the output parities $\delta \odot \mathbb{E}(P)$ and $\delta \odot \mathbb{E}(P \oplus \alpha)$ have a random distribution in $\{0, 1\}$.

Thus, if the bias $2p\epsilon^2$ is sufficiently large, we can distinguish $\mathbb{E}$ from a random function.

A Counterexample to Biham et al.'s Methodology

Caution: Work only when Assumption (III) holds; otherwise it may give probability values that are highly inaccurate.

- An intuitive situation: $\Delta\alpha \to \Delta\beta$ has probability $\frac{1}{2}$ and meets $\beta \odot \gamma = 0$, and all other possible differentials $\{\Delta\alpha \to \Delta\widehat{\beta}\}$ meet $\widehat{\beta} \odot \gamma = 1$.
  - By Biham et al.'s methodology, the bias is $2 \times \frac{1}{2} \times \epsilon^2 = \epsilon^2$, and the distinguisher is useful (if ϵ^2 is large enough).
  - In fact, the bias is zero, and the distinguisher is useless.

Be careful with using an assumption, and be preferable to use as few assumptions as possible.

5.4 Our Methodology

Use only Assumptions (I) and (II).

- Treat $\mathbb{E}$ as a cascade of two sub-ciphers $\mathbb{E} = \mathbb{E}_0 \circ \mathbb{E}_1$.
- An input difference $\Delta\alpha$ for $\mathbb{E}_0$.
- A linear expression $\Gamma\gamma \rightarrow \Gamma\delta$ with bias ϵ for $\mathbb{E}_1$.
- Compute $\sum_{\beta \in \{0,1\}^n, \gamma \odot \beta = 0} \mathsf{Pr}_{\mathbb{E}_0}(\Delta\alpha \rightarrow \Delta\beta) = \widehat{p}$.
- Concerned event: $\mathbb{E}(P) \odot \delta = \mathbb{E}(P \oplus \alpha) \odot \delta$.
- Probability: $\frac{1}{2} + 2(2\widehat{p} - 1)\epsilon^2$ under Assumptions (I) and (II).

5.5 Implications

- Our result is more reasonable and general than Biham et al.'s result:
  1. Biham et al.'s result uses three assumptions; ours uses only two of them.
  2. Ours holds in some situations where Biham et al.'s result does not hold.
- Suggest a different methodology.
  * Biham et al.'s result computes p.
  * Ours suggests to compute $\sum_{\beta \in \{0,1\}^n, \gamma \odot \beta = 0} \Pr_{\mathbb{E}_0}(\Delta\alpha \rightarrow \Delta\beta)$.
- Can lead to some better differential-linear cryptanalytic results.

5.6.1 Application to DES

A 11-round differential-linear distinguisher with bias $2^{-24.05}$:

* $\mathbb{E}_0$: Rounds 1–5.
* $\mathbb{E}_1$. Rounds 6–11.
* $\Gamma\gamma \rightarrow \Gamma\delta$: $0x0000000001040080 \rightarrow 0x2104008000008000$ with bias $2^{-8.04}$.
* $\alpha = 0x4000000000000000$.
* $\widehat{p} = 0.5009935476482946 25$.

A differential-linear attack on 13-round DES.

– Langford and Hellman's methodology broke 8-round DES.
– Biham, Dunkelman and Keller's methodology broke 9-round DES.

5.6.2 Application to Serpent

A 9-round differential-linear distinguisher with bias $2^{-59.41}$:

* $\mathbb{E}_0$: Rounds 2 to 4.
* $\mathbb{E}_1$: Rounds 5 to 10.
* $\Gamma\gamma \longrightarrow \Gamma\delta$: $0x00400000000000000000000000000002 \longrightarrow 0x000B0000B000030000B0200E00000010$ with bias 2^{-27}.
* $\alpha = 0x000000A00000000000000000000000$.
* $\widehat{p} = 0.4944110107421875$.

A differential-linear attack on 12-round Serpent-256:

– Biham, Dunkelman and Keller's methodology broke the same number of rounds.

5.6.3 Application to CTC2

Consider the version when used with a 255-bit block size and a 255-bit key.

A 8.5-round differential-linear distinguisher with bias 2^{-68}:

* $\mathbb{E}_0$: Rounds 1 – 3.
* $\mathbb{E}_1$: Rounds 4 to immediately before the permutation operation of Round 9.
* $\Gamma\gamma \longrightarrow \Gamma\delta$: $e_{5,33,49,54,101,112,131,138,155,168,188,193,217,247,251} \longrightarrow e_{32,151}$ with bias 2^{-33}.
* $\alpha = e_0$.
* $\widehat{p} = 0.5625$.

A differential-linear attack on 10-round CTC2.

– Biham, Dunkelman and Keller's methodology broke 8 rounds. (Incorrect.)

6.1 Meet-in-the-Middle Attack

- Introduced by Diffie and Hellman in 1977.
- Work in a known-plaintext attack scenario.
- Treats a block cipher $\mathbb{E}$ as $\mathbb{E} = \mathbb{E}^a \circ \mathbb{E}^b$.
- A candidate subkey (K_a, K_b) for $\mathbb{E}^a$ and $\mathbb{E}^b$ is likely to be correct if

$$\mathbb{E}^a_{K_a}(P) = (\mathbb{E}^b_{K_b})^{-1}(C),$$

given a known plaintext-ciphertext pair (P, C).

6.2 Higher-Order Meet-in-the-Middle Attack

- Work in a chosen-plaintext attack scenario.
- Treats a block cipher $\mathbb{E}$ as $\mathbb{E} = \mathbb{E}^a \circ \mathbb{E}^b$.
- A candidate subkey (K_a, K_b) for $\mathbb{E}^a$ and $\mathbb{E}^b$ is likely to be correct if

$$\begin{aligned}&\mathbf{f}(E^a_{K_a}(P_1), E^a_{K_a}(P_2), \cdots, E^a_{K_a}(P_l))\\ &= \mathbf{f}((E^b_{K_b})^{-1}(C_1), (E^b_{K_b})^{-1}(C_2), \cdots, (E^b_{K_b})^{-1}(C_l)),\end{aligned}$$

given a set of chosen plaintext-ciphertext pairs (P_i, C_i).

The core: Use multiple plaintexts to cancel some key-dependent component(s)/parameter(s) when constructing a basic unit of "value-in-the-middle".

6.3 Two Construction Methods

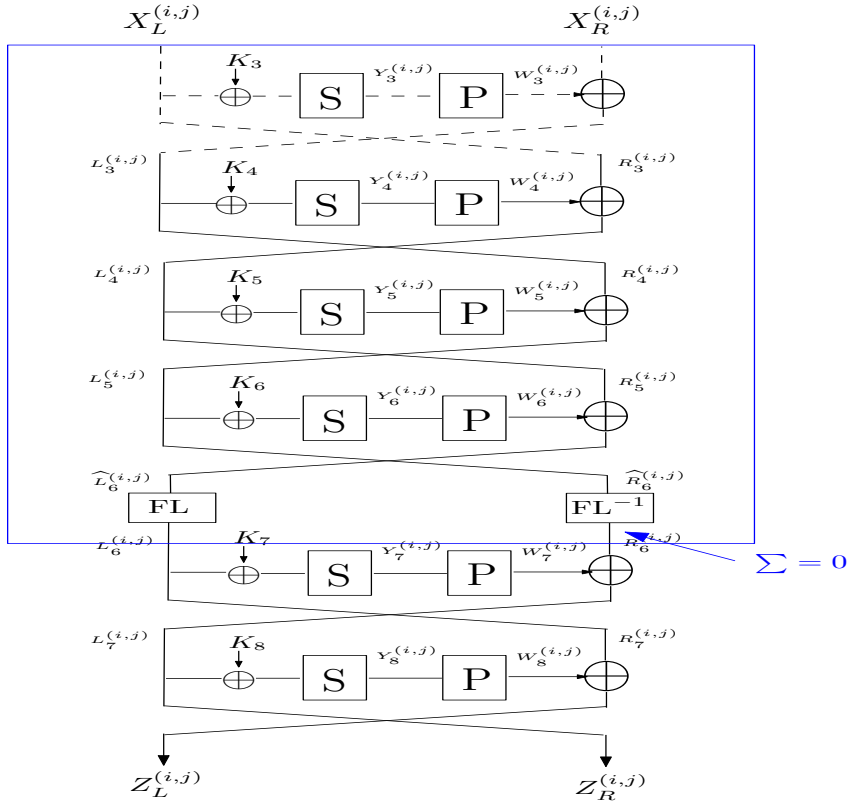
1. Combine the meet-in-the-middle attack with integral cryptanalysis.
   - The integral-meet-in-the-middle attack.
2. Combine the meet-in-the-middle attack with a general differential property.
   - XOR two value-in-the-middle's to cancel some secret parameters.
   - Often used before.

6.4 Applying Method I to Camellia

A known integral property for Rounds 3/4 to 6 with $\mathbf{FL}/\mathbf{FL}^{-1}$.

- [Input] A set of 256 sixteen-byte values $X^{(j)} = (m_1, \cdots, m_8, m_9, x^{(i)}, m_{10}, \cdots, m_{15})$.
  - $x^{(i)}$ takes all the possible values;
  - $m_1, m_2, \cdots, m_{15}$ constant.
- [Integral property]

$$\bigoplus_{j=1}^{256} \mathbf{FL}^{-1}(\widehat{R}_6^{(j)}, KI_2) = 0.$$


$X_L^{(i,j)}$
$X_R^{(i,j)}$
K_3
S
P
$Y_3^{(i,j)}$
$W_3^{(i,j)}$
$L_3^{(i,j)}$
$R_3^{(i,j)}$
K_4
$Y_4^{(i,j)}$
$W_4^{(i,j)}$
$L_4^{(i,j)}$
$R_4^{(i,j)}$
K_5
$Y_5^{(i,j)}$
$W_5^{(i,j)}$
$L_5^{(i,j)}$
$R_5^{(i,j)}$
K_6
$Y_6^{(i,j)}$
$W_6^{(i,j)}$
$\widehat{L}_6^{(i,j)}$
$\widehat{R}_6^{(i,j)}$
FL
FL^{-1}
$L_6^{(i,j)}$
K_7
$Y_7^{(i,j)}$
$W_7^{(i,j)}$
$R_6^{(i,j)}$
$\Sigma = 0$
$L_7^{(i,j)}$
$R_7^{(i,j)}$
K_8
$Y_8^{(i,j)}$
$W_8^{(i,j)}$
$Z_L^{(i,j)}$
$Z_R^{(i,j)}$

Cancel key-dependent components $\mathbf{FL}^{-1}$.

A set of 2^{16} sixteen-byte values
$X^{(i,j)} = (X_L^{(i,j)} || X_R^{(i,j)}) = (m_1, \cdots, m_8, x^{(i)}, y^{(j)}, m_9, \cdots, m_{14})$.

- $x^{(i)}$ and $y^{(j)}$ take all the possible values;
- $m_1, m_2, \cdots, m_{14}$ fixed.

For Rounds 4 to 8:

$$\begin{aligned}
& \bigoplus_{j=1}^{256} \mathbf{P}^{-1}(Z_R^{(i,j)}) \\
= \; & (\bigoplus_{j=1}^{256} \mathbf{P}^{-1}(\mathbf{FL}^{-1}(X_L^{(i,j)} \oplus W_5^{(i,j)}, KI_2))) \oplus (\bigoplus_{j=1}^{256} Y_7^{(i,j)}) \\
= \; & (\bigoplus_{j=1}^{256} \mathbf{P}^{-1}(\mathbf{FL}^{-1}(\widehat{R}_6^{(i,j)}, KI_2))) \oplus (\bigoplus_{j=1}^{256} Y_7^{(i,j)}) \\
= \; & \bigoplus_{j=1}^{256} Y_7^{(i,j)}.
\end{aligned}$$

Then, $\mathbf{P}^{-1}(\bigoplus_{j=1}^{256} Z_R^{(i,j)})[49 \sim 56]$ can be expressed as a function of $x^{(i)}$ and 13 constant 8-bit parameters.

For Rounds 3 to 8:

$$\begin{aligned}
& \bigoplus_{j=1}^{256} \mathbf{P}^{-1}(Z_R^{(i,j)}) \\
= \; & (\bigoplus_{j=1}^{256} \mathbf{P}^{-1}(\mathbf{FL}^{-1}(X_R^{(i,j)} \oplus W_3^{(i,j)} \oplus W_5^{(i,j)}, KI_2))) \oplus (\bigoplus_{j=1}^{256} Y_7^{(i,j)}) \\
= \; & (\bigoplus_{j=1}^{256} \mathbf{P}^{-1}(\mathbf{FL}^{-1}(\widehat{R}_6^{(i,j)}, KI_2))) \oplus (\bigoplus_{j=1}^{256} Y_7^{(i,j)}) \\
= \; & \bigoplus_{j=1}^{256} Y_7^{(i,j)}.
\end{aligned}$$

Then, $\mathbf{P}^{-1}(\bigoplus_{j=1}^{256} Z_R^{(i,j)})[41 \sim 48]$ can be expressed as a function of $x^{(i)}$ and 21 constant 8-bit parameters.

Cryptanalytic Results

- Break 10-round Camellia-128 with FL/FL^{-1} functions.
- Break 11-round Camellia-192 with FL/FL^{-1} functions.
- Break 12-round Camellia-256 with FL/FL^{-1} functions.

7 Conclusions

Have described five cryptanalytic techniques:

- The early abort technique for impossible differential cryptanalysis
- The early abort technique for the (related-key) rectangle attack
- The (related-key) impossible boomerang attack
- A methodology for differential-linear cryptanalysis
- The higher-order meet-in-the-middle attack

Thank you!