Zero-Correlation Linear Cryptanalysis

Mathias Hall-Andersen

Advanced Topics in Cryptology, 2018-04-25

Resources

Linear Hulls with Correlation Zero and Linear Cryptanalysis of Block Ciphers
Concerns: Introduction to zero correlation attacks
Authors: Andrey Bogdanov, Vincent Rijmen
DOI: 10.1007/s10623-012-9697-z

Zero Correlation Linear Cryptanalysis with Reduced Data Complexity
Concerns: Multidimensional zero correlation attacks
Authors: Andrey Bogdanov, Meiqin Wang
DOI: 10.1007/978-3-642-34047-5_3

Zero-Correlation Linear Cryptanalysis with FFT and Improved Attacks on ISO Standards Camellia and CLEFIA.
Concerns: Attack on Camellia and CLEFIA.
Authors: Andrey Bogdanov, Huizheng Geng et al.
DOI: 10.1007/978-3-662-43414-7_16

Setting

$$E : \mathbb{F}_2^m \times \mathbb{F}_2^n \to \mathbb{F}_2^n$$

$$\alpha, \beta \in \mathbb{F}_2^n$$

$$\alpha^T x : \mathbb{F}_2^n \to \mathbb{F}_2 \quad \beta^T x : \mathbb{F}_2^n \to \mathbb{F}_2$$

$$C_{E_k}(\alpha, \beta) = C(\alpha^T x, \beta^T E(k, x)) = 2 \cdot \Pr_x[\alpha^T x = \beta^T E(k, x)] - 1$$

Setting

Normally:

$$|C_{E_k}(\alpha, \beta)| \gg 0$$

For many keys k

Setting

From an attackers perspective is:

$$\forall k \in \mathbb{F}_2^m : C_{E_k}(\alpha, \beta) = 0$$

Useful?

Trails and Correlation Contribution

Suppose:

$$E_k(\cdot) = R_{k,r} \circ \ldots \circ R_{k,2} \circ R_{k,1}$$

Let:

$$T = (T_0, \ldots, T_r) \in (\mathbb{F}_2^n)^{r+1}$$

$$C_{E_k}(T) = \prod_{0 \leq i < r} C_{R_{k,r}}(T_i, T_{i+1})$$

Then:

$$C_{E_k}(\alpha, \beta) = \sum_{T \in \alpha \times (\mathbb{F}_2^n)^{r-1} \times \beta} C_{E_k}(T)$$

Trails and Correlation Contribution

Conclusion:

$$\forall\ T \in \alpha \times (\mathbb{F}_2^n)^{r-1} \times \beta : C_{E_k}(T) = 0 \quad \Longrightarrow \quad C_{E_k}(\alpha, \beta) = 0$$

$$C_{E_k}(T) = 0 \quad \Longleftrightarrow \quad \exists\ i : C_{R_{k,i}}(T_i, T_{i+1}) = 0$$

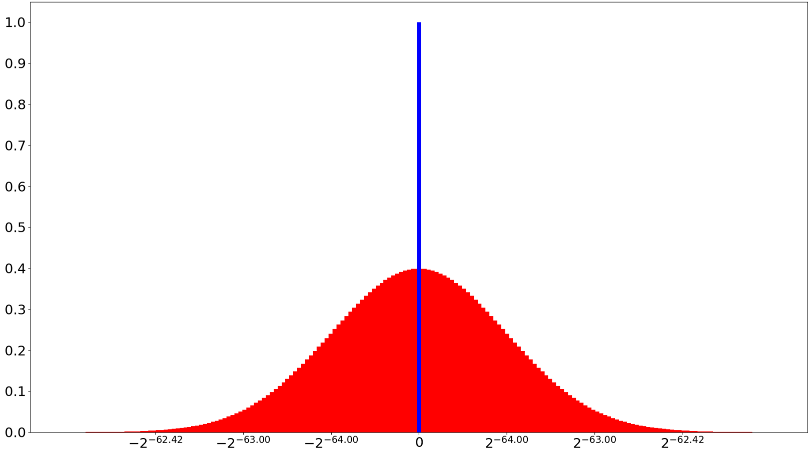
Can we find α, β st. $\forall k : C_{E_k}(\alpha, \beta) = 0$?

Wrong Key / Right Key distribution

$$C_P(\alpha, \beta) \sim \mathcal{N}(0, 2^{-n/2})$$

$$C_{E_k}(\alpha, \beta) = 0$$

Wrong Key / Right Key distribution



Finding Zero-Correlation Hulls

Can we construct concrete zero-correlation hulls? How?

Forks

$$f(x) = x \| x$$

$$\alpha^T x = \beta^T (x \| x) = \beta_1^T x + \beta_2^T x = (\beta_1 + \beta_2)^T x \implies \alpha = \beta_1 + \beta_2$$

Exclusive Or

$$f(x\|y) = x + y$$

$$\alpha^T(x\|y) = \beta^T(x+y) = \beta^T x + \beta^T y = \alpha_1^T x + \alpha_2^T y \implies \alpha_1 = \alpha_2 = \beta$$

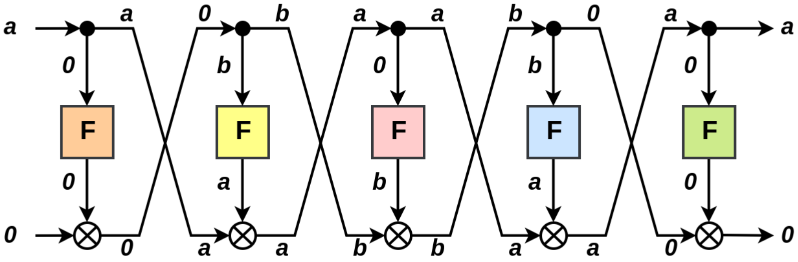
Permutations

$$\alpha \neq 0 \land C_P(\alpha, \beta) \neq 0 \implies \beta \neq 0$$

$$\beta \neq 0 \land C_P(\alpha, \beta) \neq 0 \implies \alpha \neq 0$$

Zero-Correlation Hulls for Feistel Cipher

Let $a \in \mathbb{F}_2^{n/2} \setminus \{0^{n/2}\}$, then $\alpha = \beta = 0^{n/2} \| a$ has zero correlation:



Note: How 'heavy' the F-function is does not affect the attack!

Evaluating the Correlation

Naively evaluating the correlation requires the full code-book:

$$\Pr_{x,y=E_k(x)}[\alpha^T x = \beta^T y] = \frac{|\{(x,y) \mid \alpha^T x = \beta^T y\}|}{2^n}$$

Evaluating the Correlation

$$T_{00} = \{(x,y) \mid \alpha^T x = 0 \wedge \beta^T y = 0\}$$
$$T_{01} = \{(x,y) \mid \alpha^T x = 0 \wedge \beta^T y = 1\}$$
$$T_{10} = \{(x,y) \mid \alpha^T x = 1 \wedge \beta^T y = 0\}$$
$$T_{11} = \{(x,y) \mid \alpha^T x = 1 \wedge \beta^T y = 1\}$$

$$\overbrace{|T_{00}| + |T_{01}|}^{(1)} = \overbrace{|T_{10}| + |T_{11}|}^{(2)} = 2^{n-1} = \overbrace{|T_{00}| + |T_{10}|}^{(3)} = \overbrace{|T_{01}| + |T_{11}|}^{(4)}$$

Then (4) - (2): $|T_{01}| - |T_{10}| = 0$
Then (1) - (2): $|T_{00}| + |T_{01}| - |T_{10}| - |T_{11}| =$
$|T_{00}| + |T_{10}| - |T_{10}| - |T_{11}| = |T_{00}| - |T_{11}| = 0$

Evaluating the Correlation

$$|T_{00}| = |T_{11}|$$

$$\Pr_{x,y=E_k(x)}[\alpha^T x = \beta^T y] = \frac{|\{(x,y) \mid \alpha^T x = \beta^T y\}|}{2^n}$$
$$= \frac{|T_{00}| + |T_{11}|}{2^n} = \frac{2 \cdot |T_{00}|}{2^n}$$

Note: Chosen plaintext attack.

Recap and attack example

Example of an attack:

1. Pick a 6 round balanced Feistel
2. Request the encryption of all plaintexts x st. $\alpha^T x = 0$
3. Guess the last round key k_6
   3.1 Partially decrypt the last round, and evaluate $|T_{00}|$
   3.2 If $|T_{00}| = 2^{n-2}$, add k_6 to key-candidates.

Implementation is super simple!

Recap and attack example

```
...

// collect ciphertexts (online phase)

uint32_t* ct = collect(alpha, data);

// trial decryption and correlation est.

printf("<attack>: begin key enumeration\n");

for (uint32_t key = 0; key < (1 << 16); key++) {

    size_t hits = 0;
    for (size_t i = 0; i < data; i++)
        if (parity(decrypt_round(key, ct[i]) & beta) == IN_PARITY)
            hits++;

    if (hits == (data / 2))
        printf("<attack>: possible key, %04x\n", key);
}
```

Took $\approx$ 12 hours on 96 cores for a 32-bit cipher.

Multidimensional Zero-Correlation Linear Cryptanalysis

Half the code-book is still quite a lot...

- Do we need to evaluate the correlation exactly to distinguish the distributions?
- Is there a way to use multiple approximations simultaneously to distinguish the ciphers.

Right key distribution

ℓ zero-correlation approximations.
N ct/pt pairs.

$$\text{Sample correlation: } \hat{c}_i = 2\frac{T_i}{N} - 1 \sim \mathcal{N}(0, 1/\sqrt{N})$$

Notice, no longer chosen plaintext.

Right key distribution

How do we distinguish based on ℓ dimensions?
How about mapping to a single dimension?

Right key distribution

$$\sum_{i=1}^{\ell} \hat{c}_i^2 = \sum_{i=1}^{\ell} \left(2\frac{T_i}{N} - 1\right)^2$$

Why is:

$$\sum_{i=1}^{\ell} \hat{c}_i = \sum_{i=1}^{\ell} \left(2\frac{T_i}{N} - 1\right)$$

A bad idea?

Right key distribution

Assuming iid. (big assumption)

$$\sum_{i=1}^{\ell} \hat{c}_i^2 \sim \sum_{i=1}^{\ell} \mathcal{N}^2(0, 1/\sqrt{N}) = \frac{1}{N} \sum_{i=1}^{\ell} \mathcal{N}^2(0,1) = \frac{1}{N}\chi_\ell^2$$

For sufficiently large ℓ

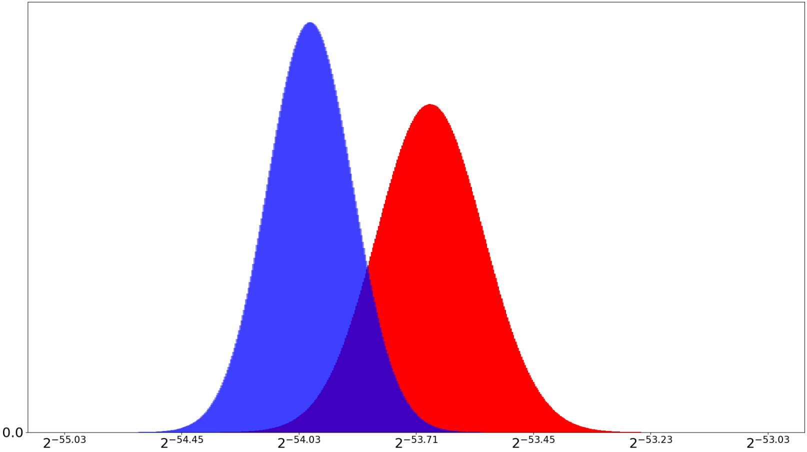
$$\frac{1}{N}\chi_\ell^2 \approx \frac{1}{N}\mathcal{N}\left(\ell, \sqrt{2\ell}\right) = \mathcal{N}\left(\frac{\ell}{N}, \frac{\sqrt{2\ell}}{N}\right)$$

Right key distribution:

$$\mathcal{N}\left(\mu_0 = \frac{\ell}{N}, \sigma_0 = \frac{\sqrt{2\ell}}{N}\right)$$

Wrong key distribution:

$$\mathcal{N}\left(\mu_1 = \frac{\ell}{N} + \frac{\ell}{2^n}, \sigma_2 = \frac{\sqrt{2\ell}}{N} + \frac{\sqrt{2\ell}}{2^n}\right)$$


0.0
$2^{-55.03}$
$2^{-54.45}$
$2^{-54.03}$
$2^{-53.71}$
$2^{-53.45}$
$2^{-53.23}$
$2^{-53.03}$

Results

Cipher	Rounds	Data	Time	Memory
AES-192	6	$2^{128}KP$	$2^{188.4}$	-
AES-192	5	$2^{127}CP$	$2^{156.3}$	-
TEA	21	$2^{62.62}KP$	$2^{121.51}$	-
XTEA	25	$2^{62.62}KP$	$2^{124.53}$	2^{32}
CLEFIA-192	14	$2^{127.5}KP$	$2^{180.2}$	2^{115}
CLEFIA-256	15	$2^{127.5}KP$	$2^{244.2}$	2^{115}
Camellia-128	11	$2^{125.3}KP$	$2^{125.8}$	2^{112}
Camellia-192	12	$2^{125.7}KP$	$2^{188.8}$	2^{112}

The zero-correlation attack on TEA, CLEFIA and Camellia are the best known attacks!