New Approaches for Estimating the Bias of Differential-Linear Distinguishers

Ting Peng, Wentao Zhang, Jingsui Weng, Tianyou Ding

Key Laboratory of Cyberspace Security Defense, Institute of Information Engineering, CAS

School of Cyber Security, University of Chinese Academy of Sciences

August 22, 2024

Overview

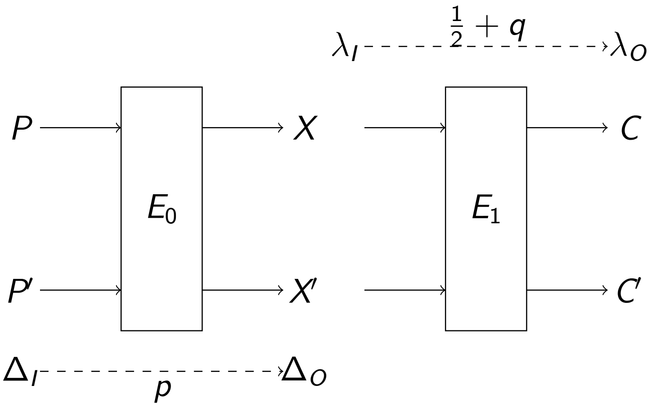
Symmetric Cryptanalysis

- Differential cryptanalysis
  - proposed by Biham and Shamir at CRYPTO 1990
  - broke DES at CRYPTO 1992

- Linear cryptanalysis
  - proposed by Matsui in 1993, broke DES again
  - the first experimental cryptanalysis of DES at CRYPTO 1994

Differential-Linear Cryptanalysis

- A combination of differential and linear cryptanalysis
  - proposed by Langford and Hellman at CRYPTO 1994
  - a chosen plaintext two-stage technique of cryptanalysis



Differential-Linear Approximation

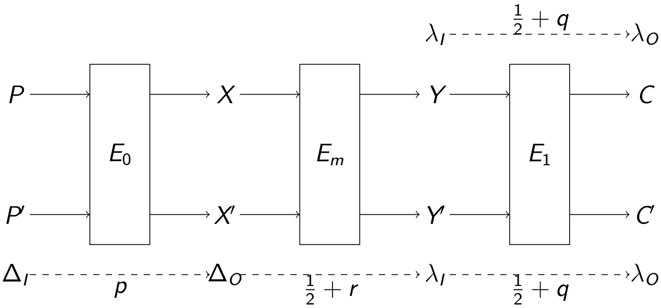
- Differential: $\Pr[\Delta_I \xrightarrow{p} \Delta_O] = p$

- Linear approximation: $\Pr[\lambda_I \xrightarrow{q} \lambda_O] = 1/2 + q$

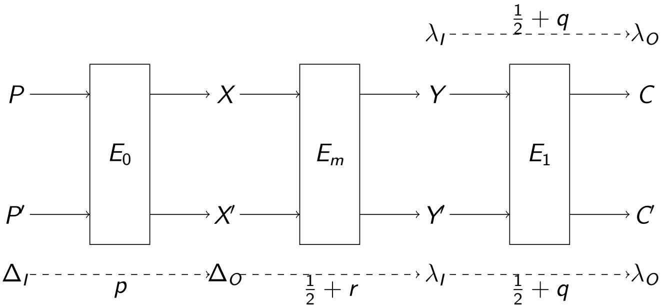
- Differential-linear approximation:
  $\Pr[C \cdot \lambda_O = C' \cdot \lambda_O | P \oplus P' = \Delta_I] = p(1/2 + 2q^2) + (1-p) \cdot 1/2 = 1/2 + 2pq^2$

Estimating the bias of a DL approximation in the middle

- Differential-Linear Connectivity Table (DLCT, EUROCRYPT 2019)
  - inspired by Boomerang Connectivity Table
  - more accurate than before
  - applications: ICEPOLE, DES, Serpent, Ascon



Estimating the bias of a DL approximation in the middle



The theoretical bias of a differential-linear approximation:

$$\mathcal{E}_{\Delta_I,\lambda_O} = 4p \cdot \overline{DLCT}_{E_m}(\Delta,\lambda) \cdot q^2 = 4prq^2$$

Differential-Linear Probability

Definition

For a t-round differential-linear approximation $(\Delta \xrightarrow{t\ round} \lambda)$, where Δ is the input difference, and λ is the output difference, the differential-linear probability (DLP) is defined by

$$\mathrm{DLP}(\Delta, \lambda) = \mathrm{Pr}[\Delta \xrightarrow{t\ round} \lambda] = \frac{|\{X|\lambda \cdot (f(X) \oplus f(X \oplus \Delta)) = 0\}|}{2^n}$$

The differential-linear bias is $\varepsilon = \mathrm{DLP}(\Delta, \lambda) - \frac{1}{2}$.

An Important Observation on DLP

$$\begin{aligned}\mathrm{DLP}(\Delta,\lambda) &= \frac{|\{X|\lambda\cdot(f(X)\oplus f(X\oplus\Delta))=0\}|}{2^n}\\ &= \frac{\sum_{\Delta_i\in\mathbb{F}_2^n,\lambda\cdot\Delta_i=0}|\{X|f(X)\oplus f(X\oplus\Delta)=\Delta_i\}|}{2^n}\\ &= \sum_{\Delta_i\in\mathbb{F}_2^n,\lambda\cdot\Delta_i=0}\overline{\mathrm{DDT}}_f(\Delta,\Delta_i)\end{aligned} \quad (1)$$

An Important Observation on DLP

- $\lambda \cdot \Delta_i = \lambda^{\{n-1\}}\Delta_i^{\{n-1\}} \oplus \lambda^{\{n-2\}}\Delta_i^{\{n-2\}} \oplus \cdots \lambda^{\{1\}}\Delta_i^{\{1\}} \oplus \lambda^{\{0\}}\Delta_i^{\{0\}}$

- If $\lambda^{\{j\}} = 0$, then $\lambda^{\{j\}}\Delta_i^{\{j\}} = 0$ always holds, which means that the value of this bit of Δ_i does not affect the value of $\lambda \cdot \Delta_i$.

- $\mathrm{DLP}(\Delta, \lambda) = \sum_{\Delta_i \in \mathbb{F}_2^n, \lambda \cdot \Delta_i = 0} \mathrm{DP}(\Delta, \Delta_i) = \sum_{\substack{0 \le j < 2^{hw(\lambda)} \\ \lambda \cdot \mathcal{T}_{t,j} = 0}} \mathrm{Pr}[\Delta \xrightarrow{t \text{ rounds}} \mathcal{T}_{t,j}]$

The Truncated Difference Distribution Table

Definition

Let f: $\{0,1\}^n \rightarrow \{0,1\}^n$ be a bijective vectorial boolean function. the TDT of f is a three-dimensional table whose first parameter $\Delta_I \in \{0,1\}^n$ is an input difference of f, and whose second parameter $\mathcal{M} \in \{0,1\}^n$ is the TD mask of a truncated output difference $\mathcal{T} \in \{*,0,1\}^n$ of f and whose third parameter is $\mathcal{Z} \in \{0,1\}^n$. Define the TDT entry $(\Delta_I, \mathcal{M}, \mathcal{Z})$ of f as

$$\mathrm{TDT}_f(\Delta_I, \mathcal{M}, \mathcal{Z}) = |\{X | \mathcal{M} \& (f(X) \oplus f(X \oplus \Delta_I)) = \mathcal{Z}\}|$$

where the TDT entry is equal to zero when $\mathcal{M} \& \mathcal{Z} \neq \mathcal{M}$.

The Truncated Difference Distribution Table

Proposition 1

The TDT is an extension of the DDT. There is a connection between DDT and TDT:

$$\mathrm{TDT}_f(\Delta_I, \mathcal{M}, \mathcal{Z}) = \sum_{\Delta : \mathcal{M} \& \Delta = \mathcal{Z}} \mathrm{DDT}_f(\Delta_I, \Delta)$$

Proposition 2

Let f: $\{0,1\}^n \rightarrow \{0,1\}^n$ be a bijective vectorial boolean function, Δ and λ denote an input difference and an output mask of f respectively.

$$\mathrm{DLP}(\Delta, \lambda) - \frac{1}{2} = \sum_{\substack{0 \leq \mathcal{Z} < 2^n \\ \lambda \cdot \mathcal{Z} = 0}} \mathrm{TDTP}(\Delta, \lambda, \mathcal{Z}) - \frac{1}{2}$$

Properties of the TDT

Property 1

$$\mathrm{TDT}_f(0, \mathcal{M}, \mathcal{Z}) = \begin{cases} 2^n, & \textit{if } \mathcal{Z} = 0 \\ 0, & \textit{if } \mathcal{Z} \neq 0 \end{cases}$$

Property 2

$$\mathrm{TDT}_f(\Delta_I, 0, \mathcal{Z}) = \begin{cases} 2^n, & \textit{if } \mathcal{Z} = 0 \\ 0, & \textit{if } \mathcal{Z} \neq 0 \end{cases}$$

Property 3

$$\mathrm{TDT}_f(\Delta_I, 2^n - 1, \mathcal{Z}) = \mathrm{DDT}_f(\Delta_I, \mathcal{Z})$$

Property 4

Given Δ_I and $\mathcal{M}$, there are at most $2^{hw(\mathcal{M})}$ non-zero entries in the TDT.

Estimation of the DLP based on TDT

The probability of a truncated differential characteristic

$$\Pr[\mathcal{T}_0 \xrightarrow{R} \mathcal{T}_1 \xrightarrow{R} \cdots \xrightarrow{R} \mathcal{T}_t] = \prod_{i=0}^{t-1} \prod_{|\mathcal{T}_i[j]|=1} \overline{\mathrm{TDT}}(\mathcal{T}_i[j], \mathcal{M}_i[j], \mathcal{Z}_i[j]) \quad (2)$$

The probability of a truncated differential

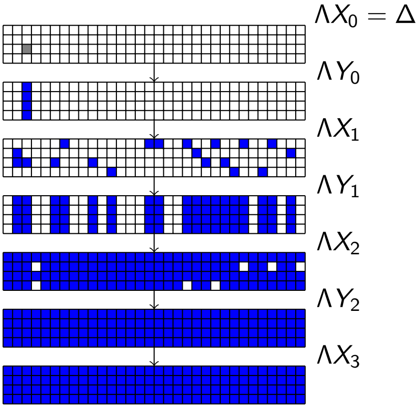
$$\Pr[\mathcal{T}_0 \xrightarrow{t\ rounds} \mathcal{T}_t] = \sum_{\mathcal{T}_1, \cdots, \mathcal{T}_{t-1}} \prod_{i=0}^{t-1} \prod_{|\mathcal{T}_i[j]|=1} \overline{\mathrm{TDT}}(\mathcal{T}_i[j], \mathcal{M}_i[j], \mathcal{Z}_i[j]) \quad (3)$$

Estimation of the DLP based on TDT

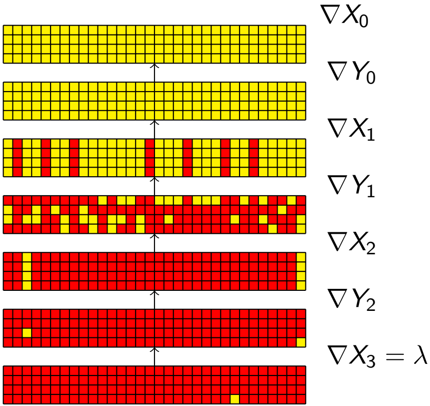
The relationship between DLP and TDT

$$\begin{aligned}\mathrm{DLP}(\Delta,\lambda) &= \sum_{\substack{0\le k<2^{hw(\lambda)}\\ \lambda\cdot\mathcal{T}_{t,k}=0}} \Pr[\Delta \xrightarrow{t\ rounds} \mathcal{T}_{t,k}] \\ &= \sum_{\substack{0\le k<2^{hw(\lambda)}\\ \lambda\cdot\mathcal{T}_{t,k}=0}} \sum_{\mathcal{T}_1,\cdots,\mathcal{T}_{t-1}} \prod_{i=0}^{t-1} \prod_{|\mathcal{T}_i[j]|=1} \overline{\mathrm{TDT}}(\mathcal{T}_i[j], \mathcal{M}_i[j], \mathcal{Z}_i[j])\end{aligned} \quad (4)$$

Computing the Differential-Linear Pattern



(a) The forward propagation of Δ: a blank cell indicates a bit difference always inactive; a gray cell indicates an active bit difference; a blue cell indicates a bit difference undetermined



(b) The backward propagation of λ: a yellow cell indicates a bit of which the bit difference need to be determined; a red cell indicates a bit of which the bit difference is arbitrary

Computing the Differential-Linear Pattern

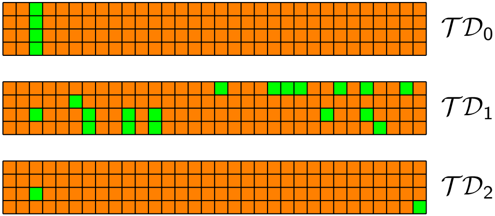


Figure: The DL pattern $(\mathcal{TD}_0, \mathcal{TD}_1, \mathcal{TD}_2)$ of 3-round Serpent: an orange cell indicates a bit of which the bit difference always be inactive or arbitrary, which is of no concern; a green cell indicates a bit difference need to be determined

- If $\wedge Y_i^{\{j\}} = 0$, $\mathcal{TD}_i^{\{j\}} = 0$.
- If $\nabla Y_i^{\{j\}} = 0$, $\mathcal{TD}_i^{\{j\}} = 0$.
- If $\wedge Y_i^{\{j\}} = 1$ and $\nabla Y_i^{\{j\}} = 1$, $\mathcal{TD}_i^{\{j\}} = 1$.

Estimate of the DL Bias when E_m Consists of Multiple Rounds

- Phase 1: Computing the differential-linear bias of E_m using the breadth-first method

- Phase 2: Extending a differential with high-probability in the differential part E_0 and a linear approximation with high-bias in the linear part E_1

- Phase 3: Computing the overall bias of the differential-linear bias of E

Computing the differential-linear bias of E_m

- For the $\mathcal{S}$ layer in the 0-th round and for each $0 \leq j < |A_0|$

$$\Pr[\Delta \xrightarrow{1\ round} A_{0,j}] = \prod_{k:\mathcal{TD}_0[k]\neq 0} \overline{\mathrm{TDT}}(\Delta[k], \mathcal{TD}_0[k], A_{0,j}[k])$$

- For the $\mathcal{S}$ layer in the i-th round ($1 \leq i < R_m$), and for each $0 \leq j < |A_i|$

$$\begin{aligned}\Pr[\Delta \xrightarrow{i\ rounds} A_{i,j}] &= \sum_{t=0}^{|B_{i-1}|-1} \mathrm{TDP}_{i-1,t} \cdot \Pr[B_{i-1,t} \xrightarrow{\mathcal{S}} A_{i,j}] \\ &= \sum_{t=0}^{|B_{i-1}|-1} \mathrm{TDP}_{i-1,t} \cdot \left(\prod_{k:\mathcal{TD}_i[k]\neq 0} \overline{\mathrm{TDT}}(B_{i-1,t}[k], \mathcal{TD}_i[k], A_{i,j}[k]) \right)\end{aligned}$$

- Finally, $\mathrm{DLP}(\Delta, \lambda) = \sum_{j=0}^{|B_{R_m-1}|-1} \mathrm{TDP}_{R_m-1,j} \cdot \pi(\lambda \cdot B_{R_m-1,j})$.
  where $\pi(x) = 1$ if $x = 0$ and $\pi(x) = 0$ otherwise. Finally, the bias of E_m is $\mathrm{DLP}(\Delta, \lambda) - \frac{1}{2}$.

The complexity

- The computational complexity:

$$|A_0| + |A_{R_m-1}| + \sum_{i=1}^{R_m-1} |A_{i-1}| \cdot |A_i| = 2^{hw(\mathcal{TD}_0)} + 2^{hw(\mathcal{TD}_{R_m-1})} + \sum_{i=1}^{R_m-1} 2^{hw(\mathcal{TD}_{i-1} \| \mathcal{TD}_i)}$$

- The memory complexity:

$$\max_{1 \leq i < R_m} (|A_{i-1}| + |A_i| + |\mathrm{TDP}_{i-1}| + |\mathrm{TDP}_i|) = \max_{1 \leq i < R_m} (2 \times (2^{hw(\mathcal{TD}_{i-1})} + 2^{hw(\mathcal{TD}_i)}))$$

Estimate of the DL Bias when E_m Consists of One Rounds

- Phase 1: Computing the probability of a truncated differential of E_0 using the depth-first method

- Phase 2: Searching a linear approximation with high-bias for the linear part E_1

- Phase 3: Using DLCT to connect the strong truncated differential and the strong biased linear approximation

Computing the probability of a truncated differential of E_0

Procedure Round-0

Begin the program.

Let $P_{TD} = 0$.

For each candidate for $\mathcal{Z}_0$ with fixed $\mathcal{TD}_0$, do the following:

Let $p_0 = \overline{\mathrm{TDT}}(\Delta X_0, \mathcal{TD}_0, \mathcal{Z}_0)$.

If $p_0 \geq \overline{TS}$, then call *Procedure Round-1*.

Exit the program.

Procedure Round-i $(1 \leq i < R_0 - 1)$

For each candidate for $\mathcal{Z}_i$ with fixed $\mathcal{TD}_i$, do the following:

Let $\Delta X_1 = \mathcal{L}(\mathcal{Z}_0)$ and $p_i = \overline{\mathrm{TDT}}(\Delta X_i, \mathcal{TD}_i, \mathcal{Z}_i)$.

If $\prod_{k=0}^{i} p_k \geq \overline{TS}$, then call *Procedure Round-*$(i+1)$.

Reture to the upper procedure.

Computing the probability of a truncated differential of E_0

Procedure Round-$(R_0 - 1)$

For each candidate for $\mathcal{Z}_{R_0-1}$ with fixed $\mathcal{TD}_{R_0-1}$, do the following:

Let $\Delta X_{R_0-1} = \mathcal{L}(\mathcal{Z}_{R_0-2})$.

Let $p_{R_0-1} = \overline{\mathrm{TDT}}(\Delta X_{R_0-1}, \mathcal{TD}_{R_0-1}, \mathcal{Z}_{R_0-1})$.

If $p = \prod_{k=0}^{R_0-1} p_k \geq \overline{TS}$, then a linear transformation is performed, i.e., $\Delta X_{R_0} = \mathcal{L}(\mathcal{Z}_{R_0-1})$.

Let $\mathcal{Z}_{R_0} = \lambda \& \mathcal{T}$. If $\Delta X_{R_0} = \mathcal{Z}_{R_0}$, then $P_{TD} = P_{TD} + p$.

Reture to the upper procedure.

Applications

- Authenticated encryption
  - Ascon
  - KNOT

- Bit-wise block cipher
  - Serpent

- Byte-wise block cipher
  - AES
  - CLEFIA

Conclusion: Ascon

Cipher	Rounds	Experimental value	Theoretical estimate				
			DLCT [1]	DATF [2]	HATF [3]	Method in Sect.4.2	Method in Sect.4.3
Ascon	4/12	2^{-2}	2^{-5}	$2^{-2.365}$	$2^{-2.09}$	$\mathbf{2^{-2}}$	
	5/12	2^{-10}					$\mathbf{2^{-10.1}}$
	6/12‡						$\mathbf{2^{-22.43}}$

- 4-round DL distinguisher: the same as the experimental result
- 5-round DL distinguisher: extrmely close to the experimental result
- 6-round DL distinguisher: first introduced

Conclusion: Serpent

Cipher	Rounds	Experimental value	Theoretical estimate			
			DLCT [1]	DATF [2]	HATF [3]	Method in Sect.4.2
Serpent	**3/32**†	$\mathbf{2^{-1.415}}$				$\mathbf{2^{-1.415}}$
	4/32	$2^{-13.75}$	$2^{-13.68}$	$2^{-13.736}$		$\mathbf{2^{-13.696}}$
	4/32†	$\mathbf{2^{-5.30}}$				$\mathbf{2^{-5.415}}$
	5/32	$2^{-17.75}$		$2^{-17.736}$		$\mathbf{2^{-17.696}}$
	5/32†	$\mathbf{2^{-11.44}}$				$\mathbf{2^{-11.415}}$
	6/32†					$\mathbf{2^{-19.61}}$
	7/32†					$\mathbf{2^{-29.45}}$
	8/32†					$\mathbf{2^{-39.45}}$
	9/32		$2^{-57.68}$	$2^{-57.736}$		$\mathbf{2^{-57.696}}$
	9/32†					$\mathbf{2^{-52}}$
	9/32†					$\mathbf{2^{-55.33}}$

Conclusion: Serpent

- revisiting 4-round and 5-round DL distinguisher
- searching for the DL distinguisher up to 9 rounds
- ignoring the key recovery, a 9-round DL distinguisher with bias of 2^{-52}
- considering the key recovery, two better 9-round DL distinguishers with bias of $2^{-55.33}$

Conclusion: KNOT256

Cipher	Rounds	Experimental value	Theoretical estimate			
			DLCT [1]	DATF [2]	HATF [3]	Method in Sect.4.2
KNOT-256	9/52	$\mathbf{2^{-1.20}}$				$\mathbf{2^{-1.20}}$
	10/52	$\mathbf{2^{-3.27}}$				$\mathbf{2^{-3.66}}$
	11/52	$\mathbf{2^{-4.31}}$				$\mathbf{2^{-6.38}}$
	12/52	$\mathbf{2^{-9.91}}$				$\mathbf{2^{-9.27}}$
	13/52	$\mathbf{2^{-14.04}}$				$\mathbf{2^{-12.27}}$
	14/52					$\mathbf{2^{-16.23}}$
	15/52					$\mathbf{2^{-23.31}}$
	16/52					$\mathbf{2^{-30.52}}$

Conclusion: KNOT256

- focusing on the initialization phase

- searching for the DL distinguishers up to 16 rounds

- 16-round DL distinguisher: $2^{-30.52}$

Conclusion: AES, CLEFIA

Cipher	Rounds	Experimental value	Theoretical estimate			
			DLCT [1]	DATF [2]	HATF [3]	Method in Sect.4.2
AES	2/10	$\mathbf{2^{-1}}$				$\mathbf{2^{-1}}$
	3/10	$\mathbf{2^{-8.66}}$				$\mathbf{2^{-8.66}}$
	4/10					$\mathbf{2^{-27.85}}$
	5/10					$\mathbf{2^{-51.85}}$
CLEFIA	4/18	$\mathbf{2^{-1}}$				$\mathbf{2^{-1}}$
	5/18	$\mathbf{2^{-2.415}}$				$\mathbf{2^{-2.415}}$
	6/18	$\mathbf{2^{-6.81}}$				$\mathbf{2^{-6.80}}$
	7/18	$\mathbf{2^{-11.81}}$				$\mathbf{2^{-11.80}}$
	8/18					$\mathbf{2^{-32.70}}$
	9/18					$\mathbf{2^{-54.37}}$

Conclusion: AES, CLEFIA

- 3/4/5-round AES's DL distinguishers: $2^{-8.66}/2^{-27.85}/2^{-51.85}$

- searching for CLEFIA's DL distinguishers up to 9 round

- 9-round CLEFIA DL distinguisher: $2^{-54.37}$

References

Bar-On, A., Dunkelman, O., Keller, N., Weizman, A.: DLCT: a new tool for differential-linear cryptanalysis. In: Ishai, Y., Rijmen, V. (eds.) EUROCRYPT 2019. LNCS, vol. 11476, pp. 313–342. Springer, Cham (2019).

Liu, M., Lu, X., Lin, D.: Differential-linear cryptanalysis from an algebraic perspective. In: Malkin, T., Peikert, C. (eds.) CRYPTO 2021, Part III. LNCS, vol. 12827, pp. 247–277. Springer Cham (2021).

Hu, K., Peyrin, T., Tan, Q.Q., Yap, T.: Revisiting Higher-Order Differential-Linear Attacks from an Algebraic Perspective. ASIACRYPT 2023. LNCS, vol 14440. Springer, Singapore.

The End