Cryptanalysis of Lightweight Block Ciphers: Theory Meets Dependencies

Orr Dunkelman

Computer Science Department
University of Haifa, Israel

December 14th, 2019

Outline

1-Round Differential Characteristics [BS91]

Definition A 1-**round differential characteristic** is a pair (Ω_P, Ω_T) where Ω_P and Ω_T are n-bit differences, such that the probability of a pair with input difference Ω_P to have an output difference Ω_T after one round is p.

r-Round Differential Characteristics [BS91]

Definition A r-**round differential characteristic** is a tuple $\Omega = (\Omega_P = \Omega_0, \Omega_1, \Omega_2, \ldots, \Omega_r = \Omega_T)$ where Ω_P, Ω_T, and all Ω_i are n-bit differences, where Ω_i are the differences predicted after each round of the scheme.

Probability of a Characteristic

- **Definition:** The **probability** of a characteristic is the probability that a random pair P, P^* which satisfies $P' = \Omega_P$ is a right pair with respect to a random independent key.
- The probability of an r-round characteristic is the product of all the probabilities of the 1-round characteristics which compose the n-round characteristic.
- There is an underlying assumption that all the transitions are independent.
- Usually, it is OK to assume that. **Usually**. **Usually**. **Usually**.

Underlying Assumptions for Differential Attacks

Formally, let

$$G_K\left(\Omega_P \xrightarrow{E} \Omega_T\right) = \left\{P \middle| E_K(P) \oplus E_K(P \oplus \Omega_P) = \Omega_T\right\}.$$

and

$$G_K^{-1}\left(\Omega_P \xrightarrow{E} \Omega_T\right) = \left\{C \middle| E_K^{-1}(C) \oplus E_K^{-1}(C \oplus \Omega_T) = \Omega_P\right\}.$$

These two sets contain all the right pairs (i.e., X is in the set if it is a part of a right pair).

Independence Assumptions for Differential Attacks

1. The probability of the differential characteristic in round i is independent of other rounds.

(formally: the event $X \in G_K^{-1}(\Omega_P \xrightarrow{E_0} \Omega_{r'})$ is independent of the event $X \in G_K(\Omega_{r'} \xrightarrow{E_1} \Omega_T)$ for all K and $\Omega_{r'}$)

2. Partial encryption/decryption under the wrong key makes the cipher closer to a random permutation.

Independent Subkeys

- A cipher whose subkeys are all chosen at random (independently of each other) can be modeled as a Markov chain.
- For such a cipher, the previous conditions are satisfied (under reasonable use of the keys) as the independent subkeys assure that the inputs to each round are truly random and independent.

Independent Subkeys — Where We Cheated

- The above assumes that the keys are chosen *during* the differential attack, and for each new pair of plaintexts, they are chosen again at random.
- This is of course wrong, as the key is fixed **a priori**, and the only source of "randomness" in the experiment is the plaintext pair.
- Hence, we need to assume *Stochastic Equivalence*, i.e.,

$$\Pr[\Delta C = \Omega_T | \Delta P = \Omega_P] =$$

$$\Pr[\Delta C = \Omega_T | \Delta P = \Omega_C \wedge K = (k_1, k_2, \ldots)]$$

  for almost all keys K.
- See more info at [LM93] where the Markov cipher is introduced.

Why the Stochastic Equivalence Assumption was Used?

- It works — most of the times it works.
- Even when it does not work for a large portion of the keys — it is mostly an issue of weak keys.
- Experiments showed it to hold many times.

However,

In theory there is no difference between theory and practice.

In practice, there is.

XOR Differences in Additive World [WangDK07]

A differential Characteristic used in [HKK+05] for SHACAL-1 from round 6 to round 12:

i	ΔA_i	ΔB_i	ΔC_i	ΔD_i	ΔE_i	ΔK_i	*Prob.*
6	e_3	0	0	$e_{13,31}$	0	0	2^{-3}
7	e_8	e_3	0	0	$e_{13,31}$	e_{31}	2^{-3}
8	0	e_8	e_1	0	0	0	2^{-2}
9	0	0	e_6	e_1	0	0	2^{-2}
10	0	0	0	e_6	e_1	0	2^{-2}
11	e_1	0	0	0	e_6	0	2^{-2}
12	0	e_1	0	0	0	0	2^{-1}

XOR Differences in Additive World [WangDK07]

- According to $A_{i+1} = K_i + ROTL_5(A_i) + F_i(B_i, C_i, D_i) + E_i + Con_i$, we get that $A_{7,8} = A_{6,3}$ and $A^*_{7,8} = A^*_{6,3}$.
- From the encryption algorithm, we get that $A_{11,1} = E_{10,1} = A_{6,3}$, $A^*_{11,1} = E^*_{10,1} = A^*_{6,3}$, $E_{11,6} = A_{7,8}$ and $E^*_{11,6} = A^*_{7,8}$.
- From the above two claims, we obtain that $A_{11,1} = E_{11,6}$ and $A^*_{11,1} = E^*_{11,6}$. By $A_{i+1} = K_i + ROTL_5(A_i) + F_i(B_i, C_i, D_i) + E_i + Con_i$, we obtain that $A_{12} \neq A^*_{12}$, i.e., $\Delta A_{12} \neq 0$, which is a contradiction with $\Delta A_{12} = 0$ in the differential characteristic.

The signs of the difference are not compatible.

Linear Cryptanalysis [M93]

- Linear cryptanalysis studies the relation between plaintext, ciphertext, and key bits.
- The key element is the linear approximation:

$$\lambda_P \cdot P \oplus \lambda_C \cdot C = \lambda_K \cdot K$$

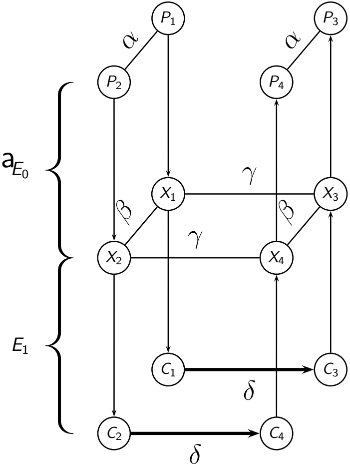
  that holds for non-trivial $\lambda_P, \lambda_C, \lambda_K$ with as large as possible bias*.
- Such approximations can be built by concatenating short 1-round approximations to form an r-round approximations.

Independence Assumptions in Linear Cryptanalysis

- Two 1-round approximations that are concatenated are independent,
- There are no other linear approximations (with the same input/output masks) that interfere with the approximation we use,
- Random wrong keys, produce a close to uniform distribution w.r.t. the probability of satisfying the approximation.

The Boomerang Attack

- Introduced by [W99].
- Targets ciphers with good short differentials, but bad long ones.
- The core idea: Treat the cipher as a cascade of two sub-ciphers. Where in the first sub-cipher a differential $\alpha \xrightarrow{E_0} \beta$ exists, and a differential $\gamma \xrightarrow{E_1} \delta$ exists for the second.
- The process starts with a pair of plaintexts: $P_1, P_2 = P_1 \oplus \alpha$.
- After the first sub-cipher, $X_1 \oplus X_2 = \beta$.



Underlying Assumptions for the Boomerang Attack

For $E = E_1 \circ E_0$, and any set of differences α, γ and δ,, we require that X is (part of) a right pair with respect to $\gamma \xrightarrow{E_1} \delta$ independently of the following three events:

1. X is (part of) a right pair with respect to $\alpha \xrightarrow{E_0} \beta$ for all β.
2. $X \oplus \beta$ is (part of) a right pair with respect to $\gamma \xrightarrow{E_1} \delta$ for all β, γ.
3. $X \oplus \gamma$ is (part of) a right pair with respect to $\alpha \xrightarrow{E_0} \beta$ for all β.

When Independence Fails — Part I

- The independence may fail if
  - There is one β whose most significant bit is 0 for which $\Pr\left[\alpha \xrightarrow{E_0} \beta\right] = 1/2$.
  - For all other β_1: $\Pr\left[\alpha \xrightarrow{E_0} \beta_1\right]$ is either 0 or 2^{-n+1}.
  - In all $X \in G_K^{-1}\left(\alpha \xrightarrow{E_0} \beta\right)$ and all $X \in G_K^{-1}\left(\alpha \xrightarrow{E_0} \beta\right)$ the most significant bit is 0.
  - There is one γ whose most significant bit is 1 for which $\Pr\left[\gamma \xrightarrow{E_1} \delta\right] = 1/2$.
  - For all other γ_1: $\Pr\left[\gamma_1 \xrightarrow{E_1} \delta\right]$ is either 0 or 2^{-n+1}.

When Independence Fails — Part II

- Consider the case where the last round of the first differential characteristic relies on the transformation $x \xrightarrow{S} y$ for some S-box S.
- If the difference distribution table of S satisfies that $DDT_S(x, y) = 2$, and if the difference in γ is such that the two pairs (X_a, X_c) and (X_b, X_d) have a non-zero difference in the bits of x, then the transition is impossible.

Is it Serious?

- It is possible to construct not-so-artificial examples of boomerangs that fail one of the above two examples [M09].
- On the other hand, the failure is with respect to a pair of intermediate differences β', γ'.
- When truly taking all possible differences (in the boomerang attack or in the rectangle attack), this problem tends to "shrink".

Differential-Linear Cryptanalysis

- Introduced first by [LH93] combines a differential with a linear approximation.
- Later extended to deal with probabilistic differentials [L94,BDK02,...]
- Very subtle dependency issues.

Dependency in DL Cryptanalysis

- Local issues — the differential and the linear approximation must not have internal dependency issues,
- Transition issues — wrong pairs (w.r.t. the differential) behave randomly w.r.t. the linear approximation,
- Transition issues 2 — right pairs (w.r.t. the differential) behave randomly w.r.t. the linear approximation,

Dependency Can Also Help!

- We can utilize dependency for improving attacks.
- Differential/linear cryptanalysis — conditional variants [BB93,BP18], multidimensional linear attacks [JV03,KR94,BDQ04,...], yoyo [BBD+99], mixture differentials [G18]
- Boomerang — boomerang switch [W99,BK09], middle-round trick [BCD03], Sandwich [DKS10], Boomerang Connectivity Table [CHP+18]
- Differential-Linear — Differential-Linear Connection Table [BDK+19]

Conditional Differential Cryptanalysis [BB93]

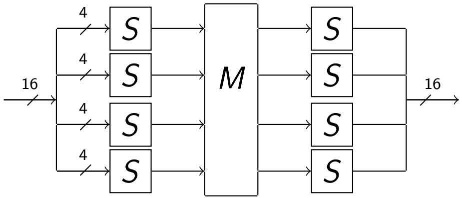
- Condition the differential transition on "events".
- Key conditions can be viewed as "weak-key" classes (very large ones).
- For hash functions — very related to collision finding techniques.
- Can be conditioned on actual plaintext/ciphertext values.

Conditional Linear Cryptanalysis [BP18]

- Condition the linear approximation on externally observable events.
- For example, fix a bit to some value.
- Or condition on a second linear approximation.

Piccolo (Linear Cryptanalysis & S-boxes)

- Piccolo is a generalized Feistel construction [SIH+11] for lightweight environments.
- Its round function has the following structure:



Finding a Linear Approximation through F

- The matrix M is an MDS.
- Just look for 5 active S-boxes approximations.
- Or treat the entire function as a 16-bit function:

Linear approximation of F	Bias
$0029_x \to 8808_x$	2^{-5}
$2229_x \to 0008_x$	2^{-5}
$2922_x \to 0800_x$	2^{-5}
$1022_x \to 0088_x$	2^{-5}
$9022_x \to 0088_x$	2^{-5}
$4046_x \to 8900_x$	2^{-5}
$C046_x \to 8900_x$	2^{-5}
$2222_x \to 8888_x 2222_x \to 8888_x$	-2^{-5}
$2430_x \to 0608_x$	-2^{-5}
$8862_x \to 000D_x$	$2^{-5.2}$

Finding Conditional Approximations of F

Linear approximation of F	Toatal Bias	MSB=0	MSB=1
$5B01_x \to 0029_x$	$2^{-5.83}$	$2^{-5.01}$	$2^{-8.38}$
$9022_x \to 0088_x$	$2^{-5.01}$	$2^{-6.05}$	$2^{-4.44}$
$1022_x \to 0088_x$	$2^{-5.01}$	$2^{-6.05}$	$-2^{-4.44}$
$4046_x \to 8900_x$	$2^{-5.01}$	$2^{-5.44}$	$2^{-4.71}$
$C046_x \to 8900_x$	$2^{-5.01}$	$2^{-5.44}$	$-2^{-4.71}$
$62A6_x \to 0D00_x$	$2^{-5.21}$	$2^{-4.87}$	$2^{-5.71}$
$E2A6_x \to 0D00_x$	$2^{-5.21}$	$2^{-4.87}$	$-2^{-5.71}$
$662A_x \to 00D0_x$	$2^{-5.21}$	$2^{-4.87}$	$2^{-5.71}$

Experiments

- Can be used to verify the different assumptions.
- Important tool in truly assessing the complexity of an attack.
- Guarantee the "science" in cryptanalysis (reproducibility).
- Sometimes can help in producing better results. . .

Open Problems

- Maybe it is time to test the differential attack on the full DES?
- Efficient detection of conditional differential characteristics/linear approximations?
- More work with values instead of differences?
- MILP modeling of “long” relations and consistency checks?
- Improved analysis techniques for dependency checks?

Questions?

Thank you for your attention!