Multiple-Tweak differential Attack against SCARF

Christina Boura[1], Shahram Rasoolzadeh[2], Dhiman Saha[3], Yosuke Todo[4]
1 IRIF, 2 RUB, 3 Indian Institute of Technology, 4 NTT

Overview

NTT

- New types of block cipher are designed for each specific purpose.
  - SCARF (for cache attack) 10-bit block, 48-bit tweak
  - BipBip (for memory safety) 24-bit block, 40-bit tweak
- Unique feature
  - Block length is **extremely shorter** than the security level.
  - Tweak length is **enough higher** than the block length.
- Question?
  - Is the traditional statistical analysis enough?

    → More careful analysis is required !!
- Result
  - Precise differential probability evaluation for SCARF.
  - 7-round key recovery and full-round multi-key distinguisher on SCARF.

Differential cryptanalysis

NTT

$$\alpha \longrightarrow \boxed{E_K} \longrightarrow \beta$$

$$\mathrm{EDP}[\alpha \to \beta] = \frac{1}{|\mathcal{K}|} \sum_{K \in \mathcal{K}} \frac{\#\{x | E_K(x) \oplus E_K(x \oplus \alpha) = \beta\}}{2^n}$$

If $\mathrm{EDP}[\alpha \to \beta] > 2^{-n}$, we can distinguish the cipher with a query complexity of $\mathrm{EDP}[\alpha \to \beta]^{-1}$.
New types of block ciphers has a short n and wide tweak.

Is guaranteeing $\mathrm{EDP}[\alpha \to \beta] \approx 2^{-n}$ enough in this context?

→ Answer NO!!

Multiple tweak differential cryptanalysis

$$\alpha \longrightarrow \boxed{E_{K,T}} \longrightarrow \beta$$

$$\mathrm{EDP}[\alpha \to \beta] = \frac{1}{|\mathcal{K}|} \sum_{K \in \mathcal{K}} \frac{1}{|\mathcal{T}|} \sum_{T \in \mathcal{T}} \frac{\#\{x | E_{K,T}(x) \oplus E_{K,T}(x \oplus \alpha) = \beta\}}{2^n}$$

The ideal probability is $\frac{1}{2^n-1}$, and we observe the "bias" from this ideal probability.

$\mathbf{EDP}[\boldsymbol{\alpha} \to \boldsymbol{\beta}] = \frac{\mathbf{1}}{\mathbf{2^n-1}} + \boldsymbol{\epsilon}$.

Assuming the binomial distribution, we can distinguish with a query complexity of $\boldsymbol{\epsilon^2/(2^n-1)}$.

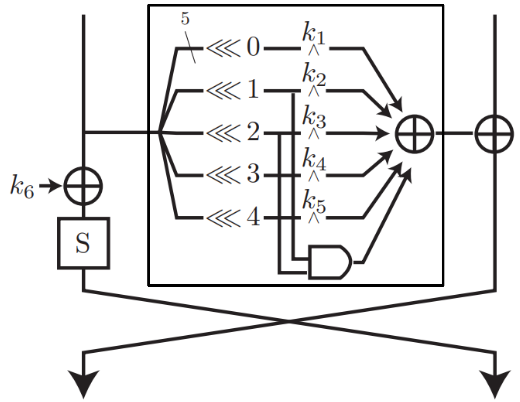
We can collect such pairs by **activating tweak**!!

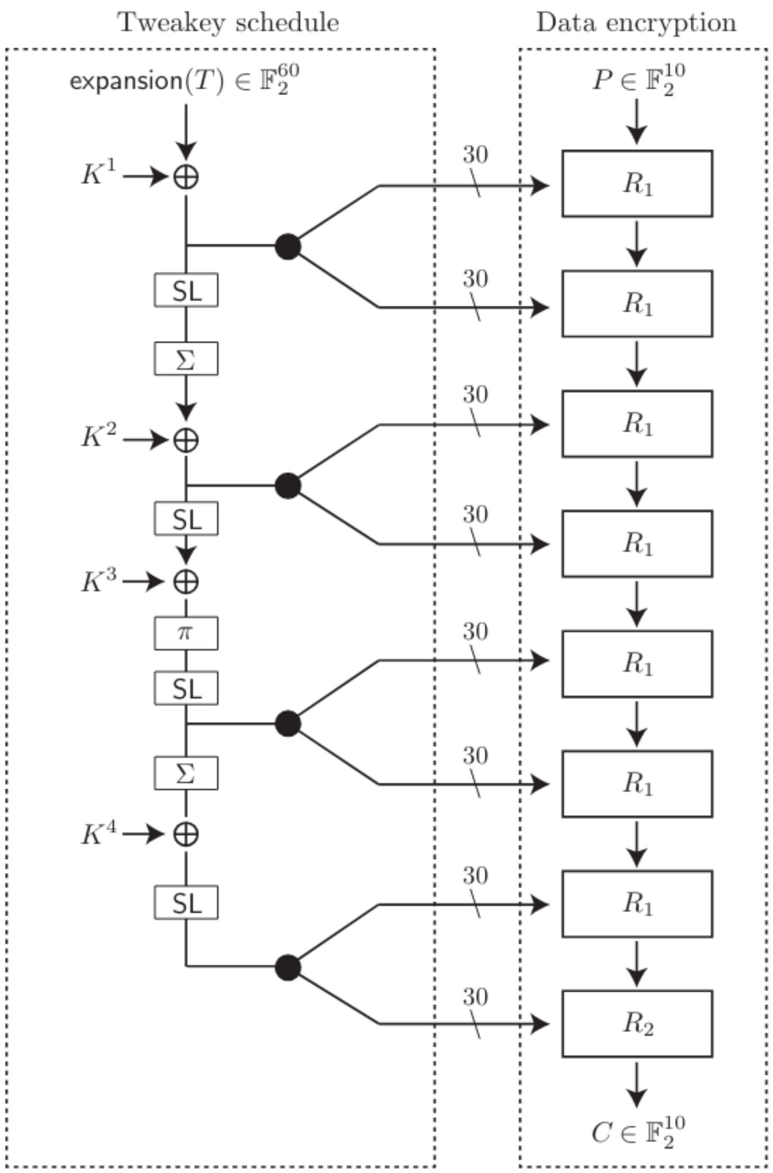
➡ We need to evaluate the differential probability accurately, but it's difficult in general.

In the case of SCARF

- SCARF
  - 10-bit block
  - 48-bit tweak
  - 80-bit security

G function



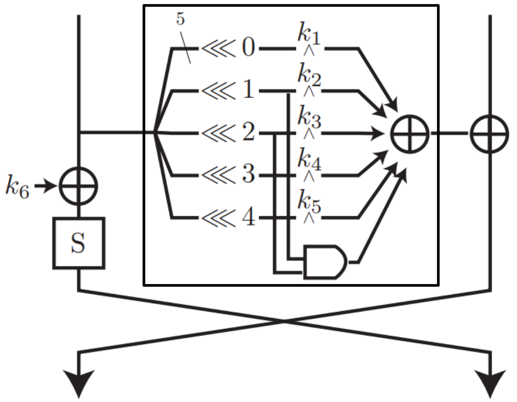


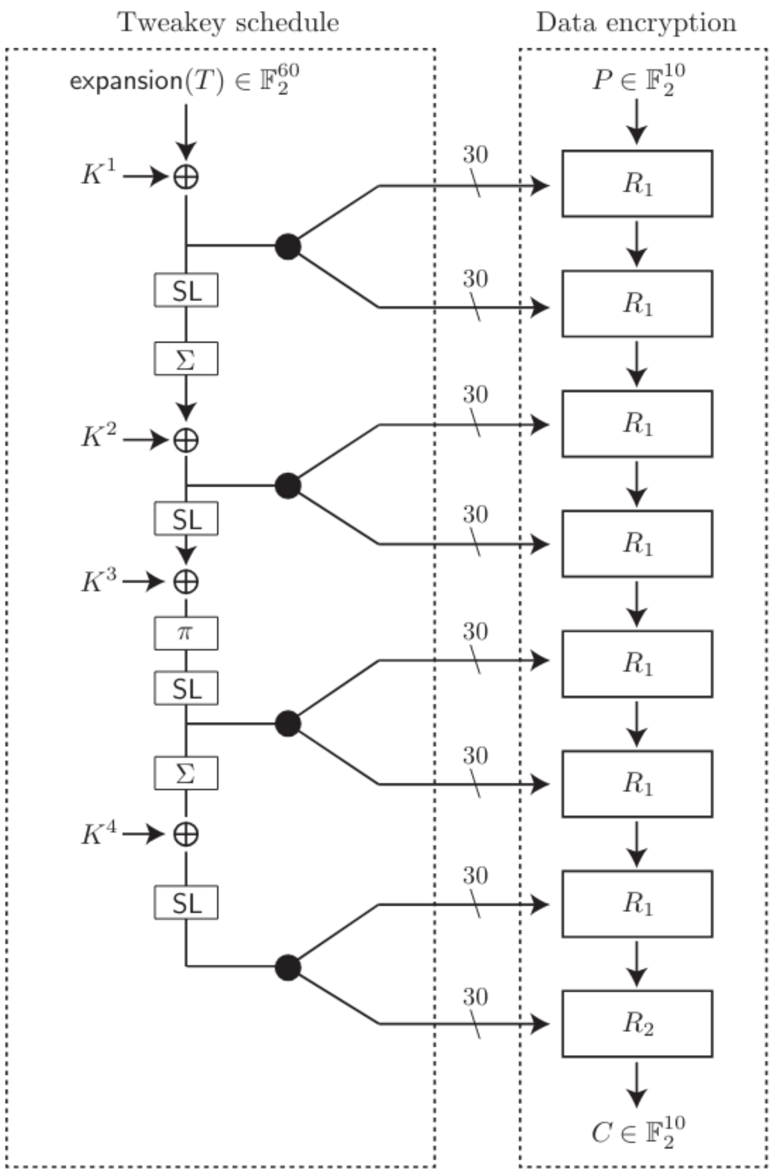
- Security requirement 1 (collision model)
  - Attacker can query (x,T) and (x′,T′).
  - If $E_T(x) = E_{T'}(x')$, the oracle return 1.
  - Attacker can't distinguish the cipher from ideal tweakable random permutation up to 2^{40} queries and 2^{80} time.

In the case of SCARF

- SCARF
  - 10-bit block
  - 48-bit tweak
  - 80-bit security

G function





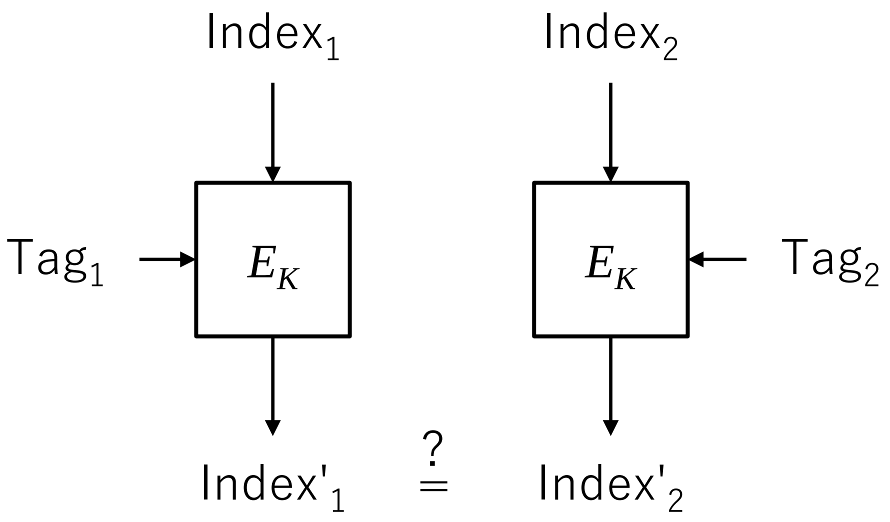
- Security requirement 2 (enc-then-dec)
  - Attacker can query (x,T,T′).
  - Attacker can learn $E_{T'}^{-1} \circ E_T(x)$ with 1 query.
  - Attacker can't distinguish the cipher from ideal tweakable random permutation up to 2^{40} queries and 2^{80} time.

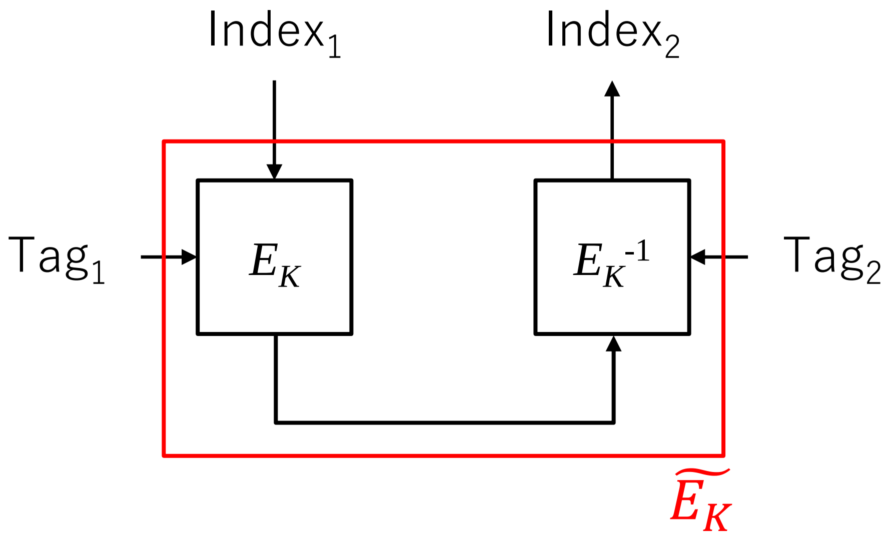
Relationship between two claims

Security requirement 1 (SR1, collision model)	Security requirement 2 (SR2, enc-then-dec model)

- Weaker security = more difficult to break.
- Respecting the real adversary scenario for cache attack.
- The definition is unfamiliar with cryptographers.

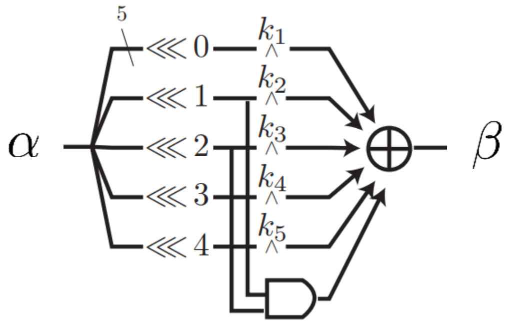
- Stronger security = easier to break.
- There is no such an oracle in real adversary scenario for cache attack.
- The definition is familiar with cryptographers.





Differential property of SCARF

- Differential property of the G function
  - The G function is key-dependent function.
  - If we ignore $(x \lll 1) \wedge (x \lll 2)$, it's a key-dependent linear function.
  - Unique feature of the G function.



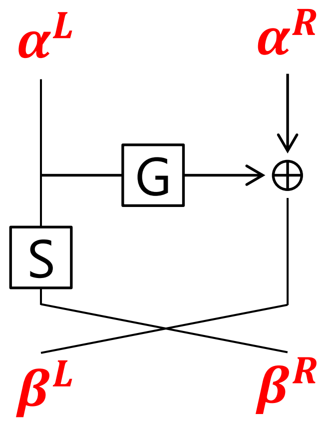
$$\mathrm{EDP}[\alpha \xrightarrow{G} \beta] = \begin{cases} 2^{-5}, & \text{if } \alpha \neq 0, \beta = *, \\ 1, & \text{if } \alpha = 0, \beta = 0, \\ 0, & \text{if } \alpha = 0, \beta \neq 0, \end{cases}$$

  - If the input difference is non-zero, the EDP is 2^{-5} independent of the output difference.

Differential property of SCARF

- Differential property of the round function

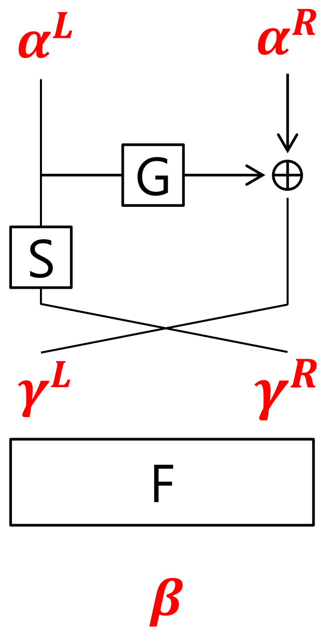
$$\mathrm{EDP}[\alpha \xrightarrow{R_1} \beta] = \begin{cases} P_S[\alpha^L, \beta^R] \times 2^{-5}, & \text{if } \alpha^L \neq 0, \\ 1, & \text{if } \alpha^L = 0 \text{ and } (\beta_L, \beta_R) = (\alpha_R, 0), \\ 0, & \text{if } \alpha^L = 0 \text{ and } (\beta_L, \beta_R) \neq (\alpha_R, 0), \end{cases}$$



  - If α^L is non-zero the EDP is $P_S[\alpha^L, \beta^R] \times 2^{-5}$ **independent of β^L**.
    - Any β^L appears with an equal probability.
    - $P_S[\alpha^L, \beta^R]$ is the differential probability of the S-box.

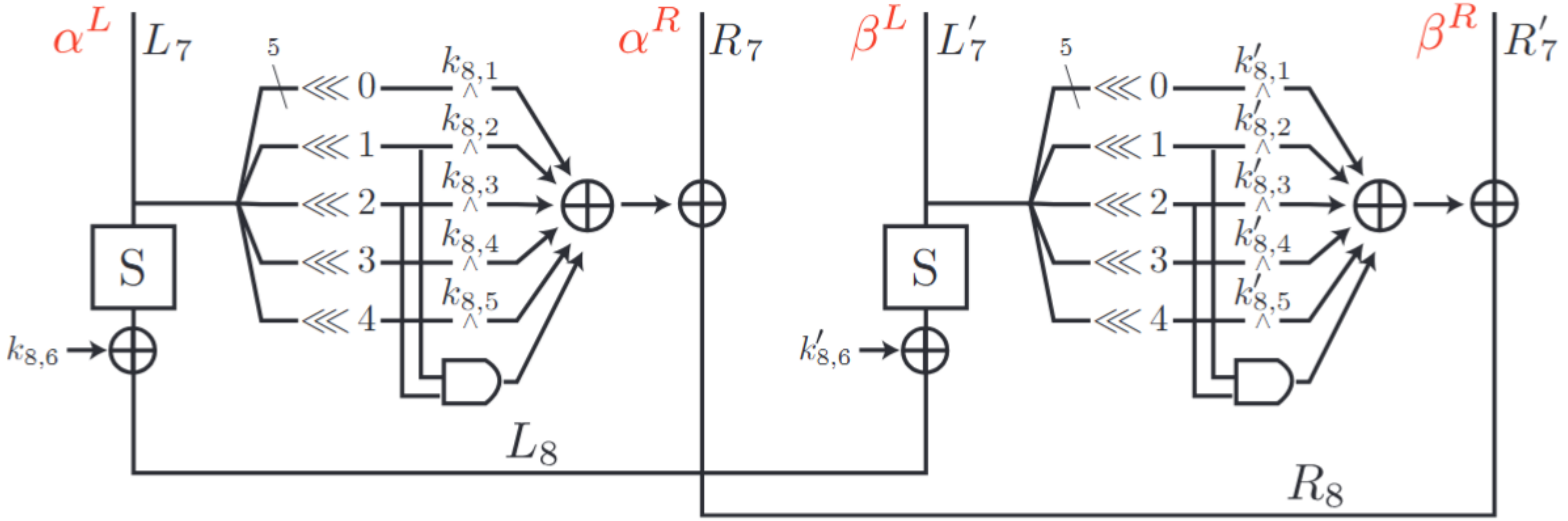
Differential property of SCARF

- Lemma
  - F is any permutation.
  - $$\mathrm{EDP}[\alpha \xrightarrow{F \circ R_1} \beta] = \sum_{\gamma} \mathrm{EDP}[\alpha \xrightarrow{R_1} \gamma] \times \mathrm{EDP}[\gamma \xrightarrow{F} \beta]$$
    $$= \sum_{\gamma} 2^{-5} \times P_S[\alpha^L, \gamma^R] \times \mathrm{EDP}[\gamma \xrightarrow{F} \beta].$$
  - Namely, $\mathrm{EDP}[\alpha \xrightarrow{F \circ R} \beta]$ does not depend on α^R if $\alpha^L \neq 0$.



Differential property of SCARF

- The property of enc-then-dec structure



$$\mathrm{EDP}[\alpha \xrightarrow{1+1} \beta] = \begin{cases} 1, & \alpha = (0, \alpha^R), \beta = (0, \beta^R), \alpha^R = \beta^R, \\ 0, & \alpha = (0, *), \beta = (\beta^L, *), \beta^L \neq 0, \\ 0, & \alpha = (\alpha^L, *), \beta = (0, *), \alpha^L \neq 0, \\ P_{S^{-1} \circ S}[\alpha^L, \beta^L] \times 2^{-5}, & \alpha = (\alpha^L, *), \beta = (\beta^L, *), \alpha^L \neq 0, \beta^L \neq 0, \end{cases}$$

Differential property of SCARF

- To compute the EDP accurately, we need to $O(2^{3n})$ time in general.
- For SCARF, thanks to the unique structure, the following 4 patterns are enough to compute the EDP.

$$\mathrm{EDP}[(0,\alpha^R) \xrightarrow{R+R'} (0,\beta^R)], \qquad \mathrm{EDP}[(0,\alpha^R) \xrightarrow{R+R'} (\beta^L,*)],$$
$$\mathrm{EDP}[(\alpha^L,*) \xrightarrow{R+R'} (0,\beta^R)], \qquad \mathrm{EDP}[(\alpha^L,*) \xrightarrow{R+R'} (\beta^L,*)].$$

- Then, the complexity is reduced to $O(2^{1.5n})$.

Multiple multi-tweak differential

- Multiple differential enhances the distinguishing advantage.

- When we use LLR statistics, the data complexity is about the inverse of the capacity.
  - Capacity using full-active differentials.

$$C_{all} = 1023 \times \sum_{\alpha \neq 0} \sum_{\beta \neq 0} \varepsilon^2_{\alpha,\beta}.$$

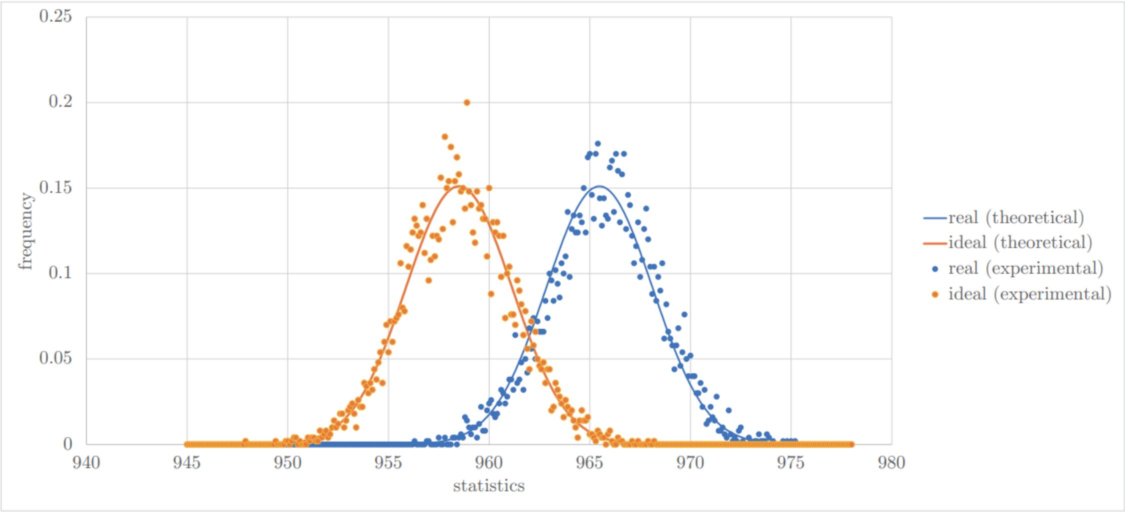
  - Capacity using right-hand active differentials.

$$C = 1023 \times \sum_{\alpha^R \neq 0} \sum_{\beta^R \neq 0} \varepsilon^2_{(0,\alpha^R),(0,\beta^R)},$$

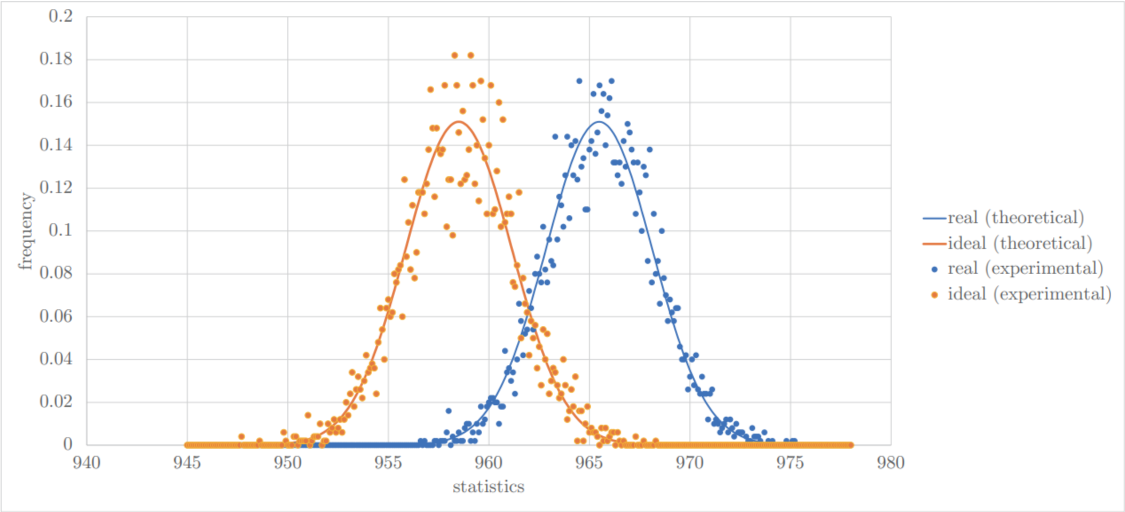
Multiple multi-tweak differential

- Summary of the capacity

Round	2+2	3+3	4+4	5+5	6+6	7+7	8+8
$-\log_2(C)$	3.37	13.39	32.20	42.40	60.89	70.89	88.95
$-\log_2(C_{all})$	2.38	13.38	31.20	42.19	59.89	70.88	87.95



Experiments for LLR test for 4+4.



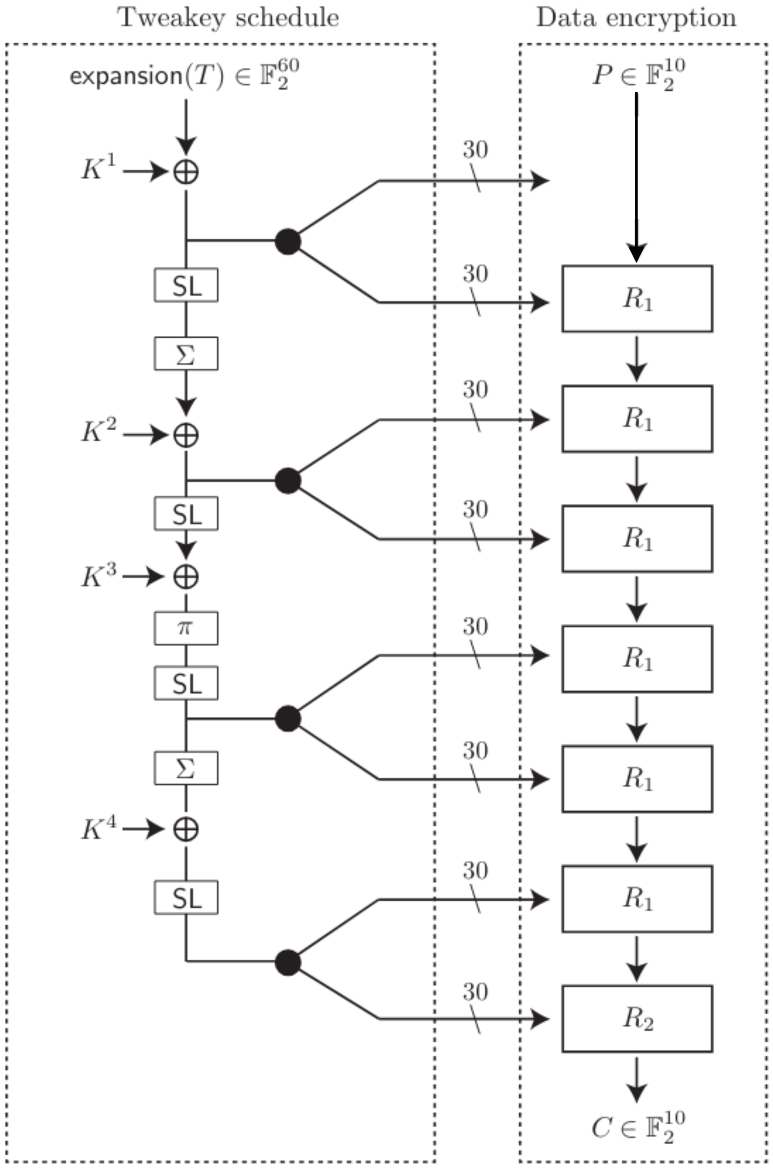
Experiments for LLR test for 5+5.

Our theoretical estimations of EDP and LLR statistics are experimentally verified.

Key recovery against 7 rounds

Define reduced-round SCARF

- We define a 7-round SCARF as the right one.
  - We use the same tweakey schedule as the original.
  - We remove the first round function.

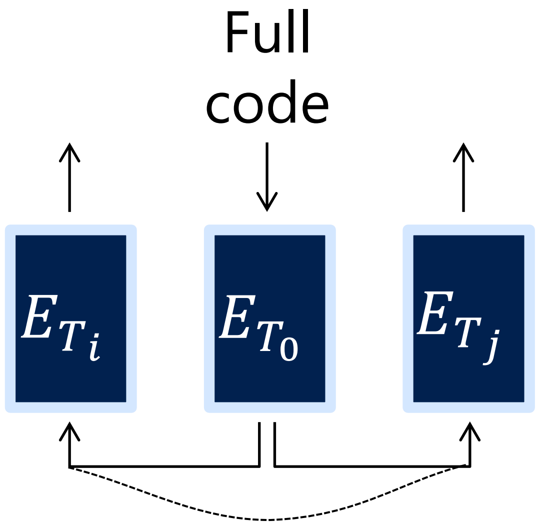


- Attack overview.
  1. Data collection.
  2. Partial-key recovery using (6+6)-round multiple differential distinguisher.
  3. Key recovery using (5S+5S)-round differential distinguisher.

Step 1. Data collection

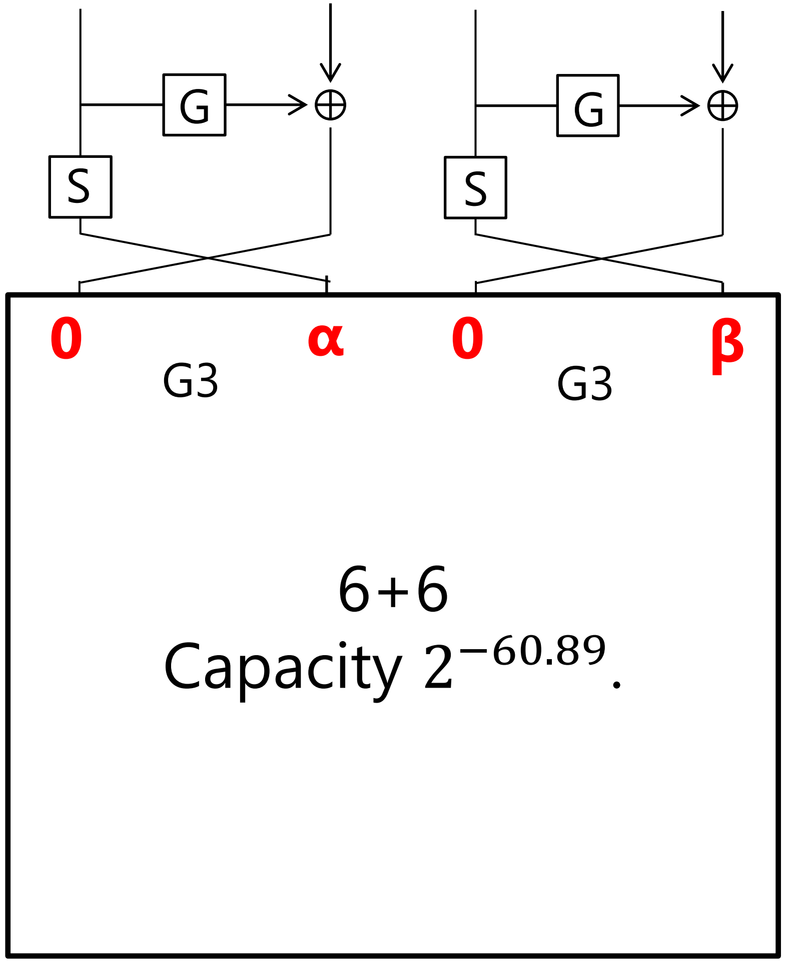
- We assume the SR2 and enc-then-dec oracle.

- Squared data collection.
  - Prepare 2^{30} tweaks T_i.
  - Query the full code book to $\tilde{E}_{T_i,T_0}$.
    - The data complexity is 2^{40}, which is the data limit.
  - Exploit the special property, $\tilde{E}_{T_i,T_j} = \tilde{E}^{-1}_{T_j,T_0} \circ \tilde{E}_{T_i,T_0}$.
    - It allows us to collect $2^{29+29+9} = 2^{67}$ (independent) pairs.



Step 2. multiple differential distinguisher

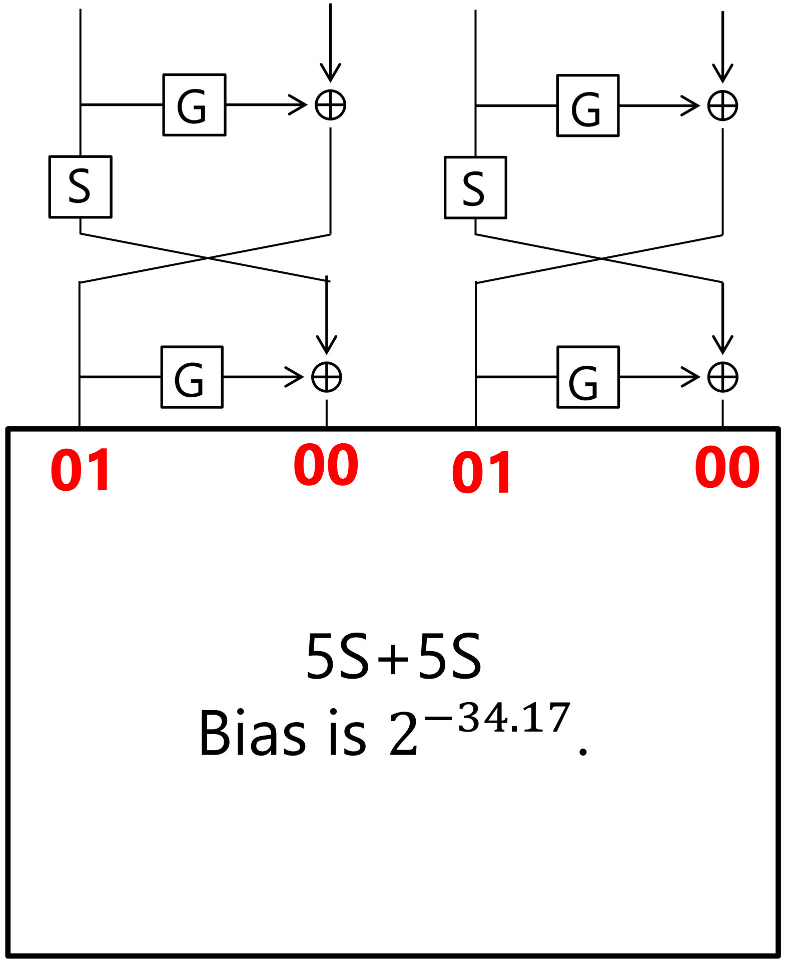
- Guess 30 bits of K^1.
- Use multiple differential distinguisher where $\alpha \in \mathbb{F}_2^5 \setminus 0$ and $\beta \in \mathbb{F}_2^5 \setminus 0$.
- We can recover the guessed 30 bits.
  - However, the remaining 30 bits of K^1 is not recovered.



Step 3. differential distinguisher

- Guess the remaining 30 bits of K^1 and 5 bits of K^2.
- Use differential distinguisher from (1,0) to (1,0).
- Recover the full K^1.

- We recover other keys with external procedure.



Multikey full-round distinguisher

Question?

- The definition of **bit security**.
  - Existing works [NW18,WY21,WY23] discuss how to define bit security properly.
- Bit security of Primitive [MW18]
  - Let T(A) be the time complexity of the algorithm A, that is linear under repetition. For any primitive, its bit security is defined as $\min_{A} \log \frac{\mathrm{T}(A)}{adv^{A}}$.

- We estimate the bit security of SCARF respecting this definition.
  - Based on the LLR-based multiple differential distinguisher

$$\log \frac{\mathrm{T}(A)}{adv^{A}} = 44.95 + 22.60 = 67.55 \ll 80\,!!$$

  - Namely, SCARF doesn't provide the 80-bit security…??

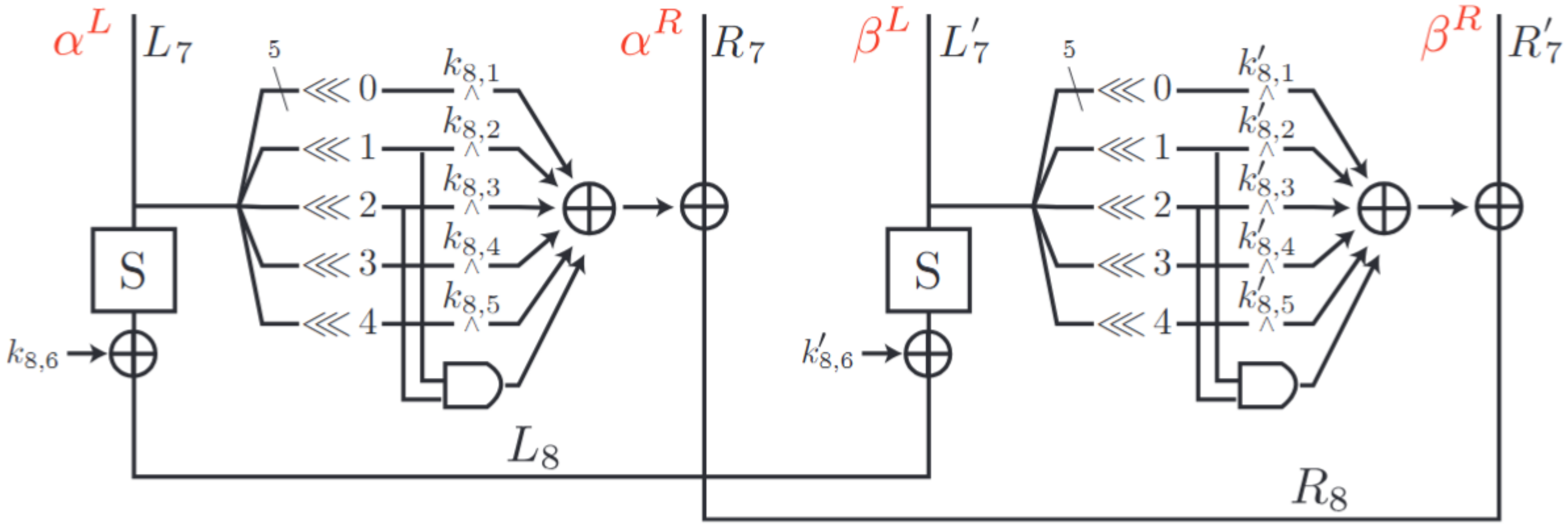
Multikey distinguisher

- WY21 and WY23 discuss a concrete attack procedure matching to the bit security.
  - Their procedure repeats the algorithm multiple times and store each bias.
  - To distinguisher in practice, these biases are combined.
- The attack can be valid assuming the multikey distinguisher.
  - Run the attack algorithm and compute the LLR statistics.
  - Repeat this procedure multiple times and distinguish by combining all statistics.
- For both SR1 and SR2, the complexity is lower than 2^{80}.
  - The attack complexity is $2^{67.55}$ for SR2.
  - The attack complexity is $2^{78.6}$ for SR1.

Observation

Connection to the BCT

- Connection to the Boomerang Connectivity Table (BCT)
  - The DDT of $S^{-1} \circ S$ is highly related to the BCT of S.
  - Up to 3+3 rounds, the BCT strongly affects the differential probability.
  - From 4 rounds, the differential probability is more influential.



Better S-box

- Are there better S-box?
  - Using the APN as the S-box is better than SCARF S-box.
  - How about low-latency S-box?
    - We explore **1016 × 5! S-boxes** satisfying the same design criteria of SCARF.
    - The following is the best alternation of the SCARF S-box.

$$
\begin{aligned}
S_{alt} = [&\mathtt{00}, \mathtt{01}, \mathtt{03}, \mathtt{0D}, \mathtt{06}, \mathtt{13}, \mathtt{16}, \mathtt{0F}, \mathtt{19}, \mathtt{10}, \mathtt{0B}, \mathtt{17}, \mathtt{09}, \mathtt{1D}, \mathtt{1A}, \mathtt{1C}, \\
&\mathtt{1E}, \mathtt{0C}, \mathtt{15}, \mathtt{04}, \mathtt{08}, \mathtt{1B}, \mathtt{11}, \mathtt{0A}, \mathtt{1F}, \mathtt{14}, \mathtt{12}, \mathtt{02}, \mathtt{05}, \mathtt{07}, \mathtt{18}, \mathtt{0E}]
\end{aligned}
$$

Conclusion

- More careful analysis is required for short-block cipher.
  - Guaranteeing 2^{-n} probability is not enough.
  - Estimating the EDP is required.
    - In general, it's difficult.
    - Thanks to the property of the G function, it's very efficient for SCARF.
- Attack on SCARF.
  - The 7-round key recovery on SR2.
  - The full-round multi-key distinguisher on both SR1 and SR2.
- Observation
  - The original SCARF S-box is not optimal in the context of the differential cryptanalysis.