

Automated Meet-in-the-Middle Attack Goes to Feistel

Qingliang Hou¹ Xiaoyang Dong^{2,4,5}(✉) Lingyue Qin^{2,4} (✉)
Guoyan Zhang^{1,3,5} (✉) Xiaoyun Wang^{1,2,3,4,5} (✉)

¹School of Cyber Science and Technology, Shandong University, Qingdao, China

²Institute for Advanced Study, BNRist, Tsinghua University, Beijing, China

³Key Laboratory of Cryptologic Technology and Information Security, Ministry of Education, Shandong University, Jinan, China

⁴Zhongguancun Laboratory, Beijing, China

⁵Shandong Institute of Blockchain, Jinan, China

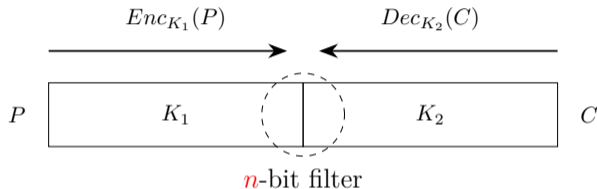
Asiacrypt 2023 / December 4 - 8, 2023

Outline

- 1 Background and preliminary
- 2 Improved matching strategy for MitM model
- 3 Preimage attack on 7-round Simpira-2
- 4 Conclusion

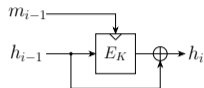
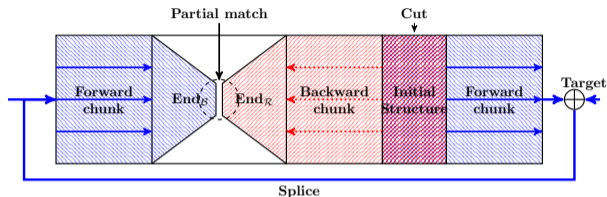
Meet-in-the-Middle(MitM) attack

Meet-in-the-Middle(MitM) attack was first introduced by Diffie and Hellman [DH77] in 1977.

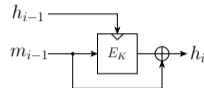


- Forward Encryption: store K_1 in table $L_1[Enc_{K_1}(P)]$.
- Backward Decryption: store K_2 in table $L_2[Dec_{K_2}(C)]$.
- Find match between L_1 and L_2 .
- $2^{|K_1|+|K_2|-n} (< 2^{|K_1|+|K_2|})$ keys survived.

MitM Preimage attack on compression function



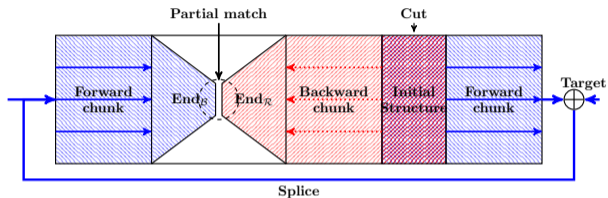
Davies-Meyer (DM) Mode



Matyas-Meyer-Oseas (MMO) Mode

- Splice-and-Cut[SAC:AS08]
 - Splice: the first and last steps are consecutive with feed-forward mechanism
 - Cut: chunk separation with *neutral sets* (■/■)
- Initial structure[EC:SA09]
 - starting steps before chunks where constraints are imposed on neutral sets
- Partial or indirect match[SAC:AS08]
 - match through $m(<n)$ bits
 - match through deterministic relation(i.e. MixColumns in AES-like hash).

MitM Preimage attack on compression function

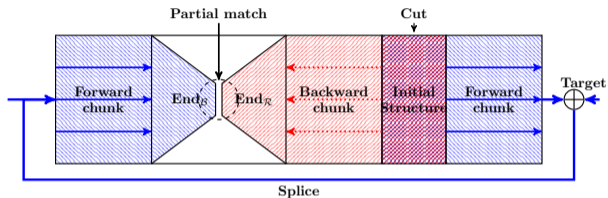


- Initial structure:

$$\begin{cases} \lambda_B \blacksquare, \lambda_R \blacksquare, \lambda_G \blacksquare \text{ in starting states.} \\ l_B \text{ constraints on } \blacksquare, l_R \text{ constraints on } \blacksquare. \end{cases}$$

- Forward: $\text{DoF}_B = \lambda_B - l_B$ values of \blacksquare .
- Backward: $\text{DoF}_R = \lambda_R - l_R$ values of \blacksquare .
- Match: DoM bytes from $\text{End}_B = \text{End}_R$.

MitM Preimage attack on compression function



- Initial structure:

$$\begin{cases} \lambda_B \blacksquare, \lambda_R \blacksquare, \lambda_G \blacksquare \text{ in starting states.} \\ l_B \text{ constraints on } \blacksquare, l_R \text{ constraints on } \blacksquare. \end{cases}$$

- Forward: $\text{DoF}_B = \lambda_B - l_B$ values of \blacksquare .
- Backward: $\text{DoF}_R = \lambda_R - l_R$ values of \blacksquare .
- Match: DoM bytes from $\text{End}_B = \text{End}_R$.

Attack Framework

- 1 Choose constants for λ_G and $l_B + l_R$ bytes.
- 2 For DoF_B values of \blacksquare , compute to End_B and store them in $L_1[\text{End}_B]$.
- 3 For DoF_R values of \blacksquare , compute to End_R and store them in $L_2[\text{End}_R]$.
- 4 Find match between L_1 and L_2 .

MitM Preimage attack on compression function

Attack Framework

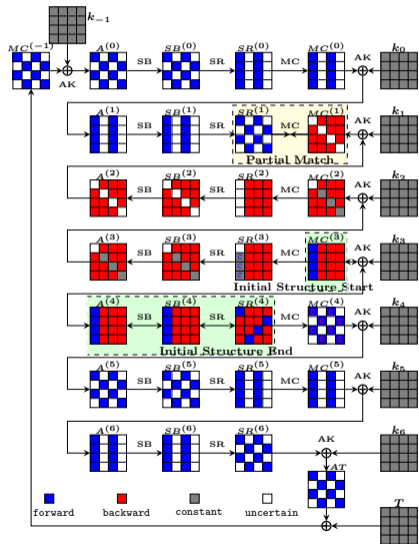
- 1 Choose constants for λ_G and $l_B + l_R$ bytes.
- 2 For DoF_B values of \blacksquare , compute to End_B and store them in $L_1[End_B]$.
- 3 For DoF_R values of \blacksquare , compute to End_R and store them in $L_2[End_R]$.
- 4 Find match between L_1 and L_2 .

Cplx.

According to $\lambda_G = n - \lambda_B - \lambda_R$, then $\lambda_G + l_B + l_R = n - \text{DoF}_B - \text{DoF}_R$,

$$\begin{aligned} T &= 2^{8 \times (n - \text{DoF}_B - \text{DoF}_R)} \times \left(2^{8 \times \text{DoF}_B} + 2^{8 \times \text{DoF}_R} + 2^{8 \times (\text{DoF}_B + \text{DoF}_R - \text{DoM})} \right) \\ &= 2^{8 \times (n - \min\{\text{DoF}_B, \text{DoF}_R, \text{DoM}\})} \\ M &= \min\{2^{8 \times \text{DoF}_B}, 2^{8 \times \text{DoF}_R}\} \end{aligned}$$

MitM (pseudo)Preimage attack on 7-round AES-like Hash [FSE:Sas11]



- Starting states(MC^3): $\lambda_B = 4$ ■, $\lambda_R = 12$ ■
- Initial structure: $l_B = 3$, $l_R = 8$

$$\begin{array}{c} MC^{(4)} \\ \text{█} \\ \text{█} \\ \text{█} \\ \text{█} \end{array} \xrightarrow{MC} \begin{array}{c} AK^{(4)} \\ \text{█} \\ \text{█} \\ \text{█} \\ \text{█} \end{array} \begin{bmatrix} 2 & 3 & 1 & 1 \\ 1 & 2 & 3 & 1 \\ 1 & 1 & 2 & 3 \\ 3 & 1 & 1 & 2 \end{bmatrix} \times \begin{bmatrix} MC^{(4)}[0] \\ MC^{(4)}[1] \\ MC^{(4)}[2] \\ MC^{(4)}[3] \end{bmatrix} = \begin{bmatrix} 2 \cdot MC^{(4)}[0] + c_0 \\ - \\ MC^{(4)}[0] + c_1 \\ - \end{bmatrix}$$

- Match through MixColumn (MC)

$$\begin{array}{c} MC^{(1)} \\ \text{█} \\ \text{█} \\ \text{█} \\ \text{█} \end{array} \xrightarrow{MC} \begin{array}{c} AK^{(1)} \\ \text{█} \\ \text{█} \\ \text{█} \\ \text{█} \end{array} \begin{bmatrix} MC^{(1)}[0] \\ - \\ MC^{(1)}[2] \\ - \end{bmatrix} = \begin{bmatrix} e & b & d & 9 \\ 9 & e & b & d \\ d & 9 & e & b \\ b & d & 9 & e \end{bmatrix} \times \begin{bmatrix} - \\ AK^{(1)}[1] \\ AK^{(1)}[2] \\ AK^{(1)}[3] \end{bmatrix}$$

- $T = 2^{8 \times (3+8)} \times (2^8 + 2^{8 \times 4} + 2^{8 \times (1+4-4)}) = 2^{120}$

Table-based method to solve nonlinear constraints [C:DHSLWH21]

- Initial structure
 - Linear constraints: Solving linear systems of equations
 - Nonlinear constraints: Precomputing the concrete value of constraint constants c_B by traversing λ_B ■ and store DoF_B ■ in table $U[c_B]$. Similar to ■ and $V[c_R]$.

Table-based method to solve nonlinear constraints [C:DHSLWH21]

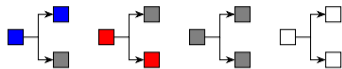
- Initial structure
 - Linear constraints: Solving linear systems of equations
 - Nonlinear constraints: Precomputing the concrete value of constraint constants c_B by traversing λ_B ■ and store DoF_B ■ in table $U[c_B]$. Similar to ■ and $V[c_R]$.

Improved attack framework

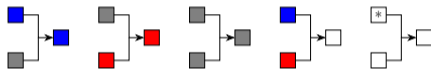
- 1 Choose constants for the starting states.
- 2 Build two table to solve nonlinear constraints c_B/c_R with additional time $\max\{2^{8 \times \lambda_B}, 2^{8 \times \lambda_R}\}$ and memory $\max\{2^{8 \times \lambda_B}, 2^{8 \times \lambda_R}\}$.
- 3 For each c_B , DoF_B values of ■ are computed to End_B and store in $L_1[\text{End}_B]$.
- 4 For each c_R , DoF_R values of ■ are computed to End_R and store in $L_2[\text{End}_R]$.
- 5 Find match between L_1 and L_2 .

Superposition states and Bi-direction attribute-propagation [C:BGST22]

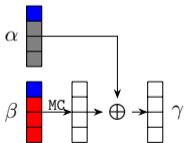
- Intermediate states are separated into two Superposition(SupP) states.
- SupP states are handled independently in linear operation and can be propagated equally in Bi-directions(BiDir).
- SupP states are combined before the next nonlinear operation.



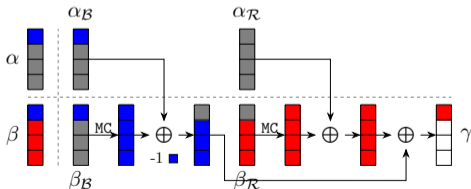
(a) Rules for separation



(b) Rules for combination

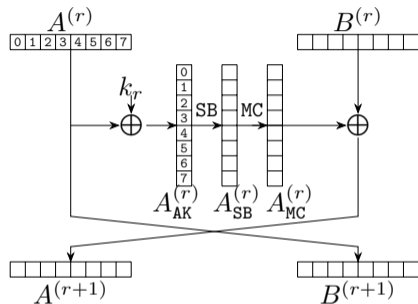


MC and Xor



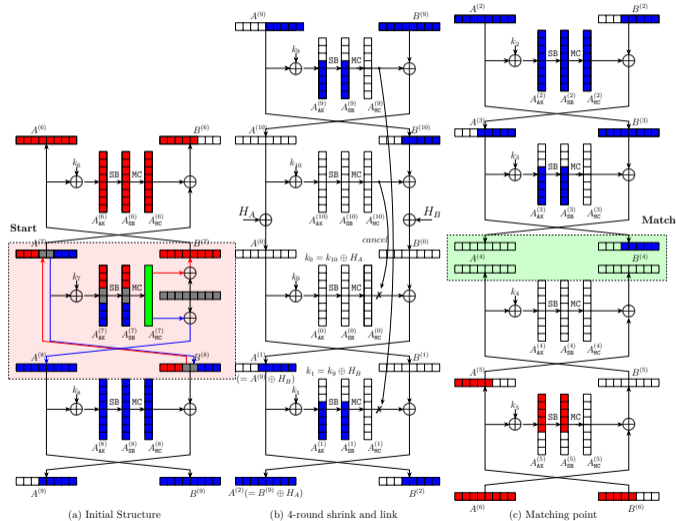
SupP MC and Xor

Description of Feistel-SP

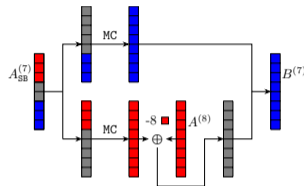


- **SubBytes(SB)**. Substitute each byte according to an S-box.
- **ShiftRows $_{\pi}$ (SR)**. Permute byte positions according to the permutation π .
- **MixColumns(MC)**. Left-multiply each column by one MDS matrix.
- **AddRoundKey(AK)**. Xor with a round key or a round constant.

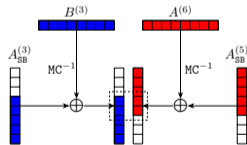
MitM Preimage attack on 11-round Feistel-SP [ACNS:Sas13]



- $\text{DoF}_B = 3, \text{DoF}_R = 11 - 8 = 3$



- $\text{DoM} = 2$

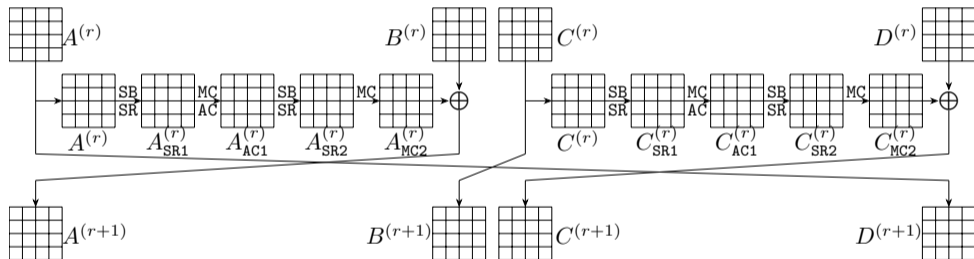


Cplx.

$$T = 2^{8 \times (16 - \min\{3, 3, 2\})} = 2^{112}$$

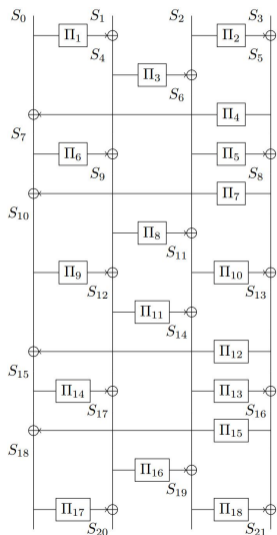
Guess and Determine Preimage attack on 9-round Simpira-4[C:SS22]

- Simpira-4: 4-brance Type-2 GFN with Double-SP round function



- **Output:** $\text{trunc}_{256}(\text{Simpira-4}(x) \oplus x) = H_A || H_B$

Guess and Determine Preimage attack on 9-round Simpira-4[C:SS22]



- **Simplified:**

$$\begin{array}{lll}
 S_0 \oplus S_7 = \Pi_4(S_5) & S_5 \oplus S_8 = \Pi_5(S_6) & S_7 \oplus S_{10} = \Pi_7(S_8) \\
 S_6 \oplus S_{11} = \Pi_8(S_9) & S_9 \oplus S_{12} = \Pi_9(S_{10}) & S_8 \oplus S_{13} = \Pi_{10}(S_{11}) \\
 S_{11} \oplus S_{14} = \Pi_{11}(S_{12}) & S_{10} \oplus S_{15} = \Pi_{12}(S_{13}) & S_{13} \oplus S_{16} = \Pi_{13}(S_{14}) \\
 S_{15} \oplus S_{18} = \Pi_{15}(S_{16}) & S_0 = S_{18} \oplus H_A &
 \end{array}$$

- **Expanded:**

$$\Pi_4(S_5) \oplus \Pi_7(S_8) \oplus \Pi_{12}(S_{13}) \oplus \Pi_{15}(S_{16}) = H_A$$

- **MitM language:**

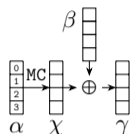
- ■ cells: S_9, S_{11} , ■ cells: S_8 , ■ cells: S_{10}
- match: $S_{10} \oplus \Pi_4(S_5) \oplus \Pi_7(S_8) = H_A \oplus \Pi_{12}(S_{13}) \oplus \Pi_{15}(S_{16}) \oplus S_{10}$
- no table is needed (only ■ and ■ appeared in direct matching)

Outline

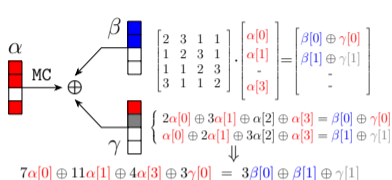
- 1 Background and preliminary
- 2 Improved matching strategy for MitM model
 - New view of Matching Strategy in MitM attack
 - Generalization of Sasaki's Matching Strategy for Feistel
- 3 Preimage attack on 7-round Simpira-2
- 4 Conclusion

New view of Matching Strategy in MitM attack

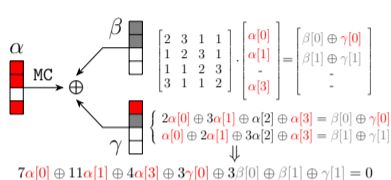
- Detect M_B bytes filter $\pi_B = 0$ and M_R bytes filter $\pi_R = 0$



(a) MC then XOR



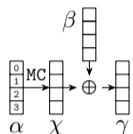
(b) DoM = 1 bytes matching



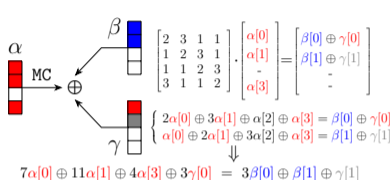
(c) $M_R = 1$ bytes matching

New view of Matching Strategy in MitM attack

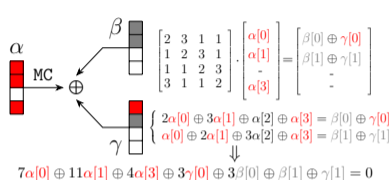
- Detect M_B bytes filter $\pi_B = 0$ and M_R bytes filter $\pi_R = 0$



(a) MC then XOR



(b) DoM = 1 bytes matching



(c) $M_R = 1$ bytes matching

New attack framework

- 1 Choose constants for the Initial structure.
- 2 For DoF_B values of ■, compute to End_B . If $\pi_B = 0$ holds, store them in $L_1[End_B]$ (Filter before store).
- 3 For DoF_R values of ■, compute to End_R . If $\pi_R = 0$ holds, store them in $L_2[End_R]$.
- 4 Find match between L_1 and L_2 .

Comparison with the traditional attack framework

- Traditional:

- ① Choose constants ..
- ② For $2^{\text{DoF}_{\mathcal{B}}}$ ■, .. store in $L_1[\text{End}_{\mathcal{B}}]$.
- ③ For $2^{\text{DoF}_{\mathcal{R}}}$ ■, .. store in $L_2[\text{End}_{\mathcal{R}}]$.
- ④ Find match between L_1 and L_2 .

- $M = 2^{8 \times \min\{\text{DoF}_{\mathcal{B}}, \text{DoF}_{\mathcal{R}}\}}$

- Ours:

- ① Choose constants ..
- ② For $2^{\text{DoF}_{\mathcal{B}}}$ ■, .. if $\pi_{\mathcal{B}} = 0$, store in $L_1[\text{End}_{\mathcal{B}}]$.
- ③ For $2^{\text{DoF}_{\mathcal{R}}}$ ■, .. if $\pi_{\mathcal{R}} = 0$, store in $L_2[\text{End}_{\mathcal{R}}]$.
- ④ Find match between L_1 and L_2 .

- $M = 2^{8 \times \min\{\text{DoF}_{\mathcal{B}} - M_{\mathcal{B}}, \text{DoF}_{\mathcal{R}} - M_{\mathcal{R}}\}}$

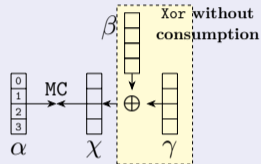
An extreme example

Let $\text{DoF}_{\mathcal{B}} = \text{DoF}_{\mathcal{R}} = 1$, and $M_{\mathcal{B}} = 1$. After 1-byte filter $\pi_{\mathcal{B}} = 0$, only one solution of ■ is derived. Therefore, no table is needed.

MILP Model to detect $\pi_{\mathcal{B}} = 0$ and $\pi_{\mathcal{R}} = 0$

Encoded with a pair 0-1 binary variables (x_i^α, y_i^α)

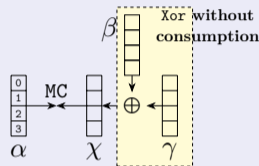
- Gray(\mathcal{G}) : $(x_i^\alpha, y_i^\alpha) = (1, 1)$, known in both chunks.
- Blue(\mathcal{B}) : $(x_i^\alpha, y_i^\alpha) = (1, 0)$, computed from ■ or ■.
- Red(\mathcal{R}) : $(x_i^\alpha, y_i^\alpha) = (0, 1)$, computed from ■ or ■.
- White(\mathcal{W}) : $(x_i^\alpha, y_i^\alpha) = (0, 0)$, unknown in both chunks.



MILP Model to detect $\pi_{\mathcal{B}} = 0$ and $\pi_{\mathcal{R}} = 0$

Encoded with a pair 0-1 binary variables (x_i^α, y_i^α)

- $$\left\{ \begin{array}{l} \blacksquare \text{ Gray}(\mathcal{G}) : (x_i^\alpha, y_i^\alpha) = (1, 1), \text{ known in both chunks.} \\ \blacksquare \text{ Blue}(\mathcal{B}) : (x_i^\alpha, y_i^\alpha) = (1, 0), \text{ computed from } \blacksquare \text{ or } \blacksquare. \\ \blacksquare \text{ Red}(\mathcal{R}) : (x_i^\alpha, y_i^\alpha) = (0, 1), \text{ computed from } \blacksquare \text{ or } \blacksquare. \\ \square \text{ White}(\mathcal{W}) : (x_i^\alpha, y_i^\alpha) = (0, 0), \text{ unknown in both chunks.} \end{array} \right.$$

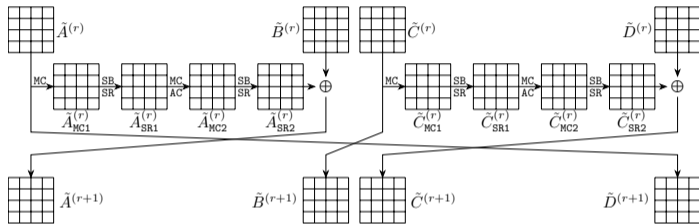


$$\left\{ \begin{array}{l} n_{\mathcal{B}}^\alpha = \sum_{i=0}^3 x_i^\alpha; \\ n_{\mathcal{R}}^\alpha = \sum_{i=0}^3 y_i^\alpha; \end{array} \right. \quad \left\{ \begin{array}{l} n_{\mathcal{B}}^\chi = \sum_{i=0}^3 x_i^\chi; \\ n_{\mathcal{R}}^\chi = \sum_{i=0}^3 y_i^\chi; \end{array} \right. \quad \left\{ \begin{array}{l} n_{\mathcal{G}} = \sum_{i=0}^3 \text{AND}(x_i^\alpha, y_i^\alpha) + \text{AND}(x_i^\chi, y_i^\chi); \\ M_{\mathcal{G}} = \max\{0, n_{\mathcal{G}} - 4\}. \end{array} \right.$$

$$M_{\mathcal{B}} = \max\{0, n_{\mathcal{B}}^\alpha + n_{\mathcal{B}}^\chi - 4\}, \quad M_{\mathcal{R}} = \max\{0, n_{\mathcal{R}}^\alpha + n_{\mathcal{R}}^\chi - M_{\mathcal{G}} - 4\}.$$

Generalization of Sasaki's Matching Strategy for Feistel

- Equivalent transformation of Simpira-4



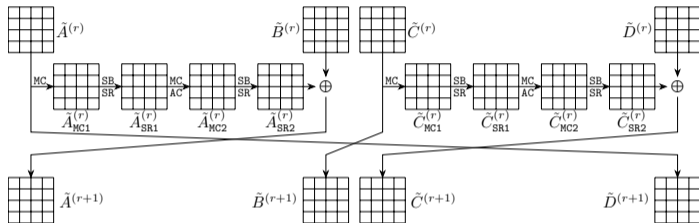
where the round function becomes

$$\mathcal{R}'_i = \text{SR} \circ \text{SB} \circ \text{AC} \circ \text{MC} \circ \text{SR} \circ \text{SB} \circ \text{MC},$$

and the input becomes $(\text{MC}^{-1}(A^{(r)}), \text{MC}^{-1}(B^{(r)}), \text{MC}^{-1}(C^{(r)}), \text{MC}^{-1}(D^{(r)}))$.

Generalization of Sasaki's Matching Strategy for Feistel

- Equivalent transformation of SImpira-4



where the round function becomes

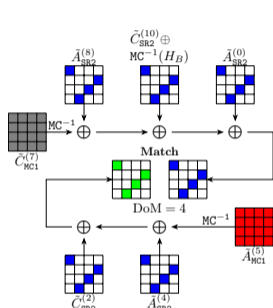
$$\mathcal{R}'_i = \text{SR} \circ \text{SB} \circ \text{AC} \circ \text{MC} \circ \text{SR} \circ \text{SB} \circ \text{MC},$$

and the input becomes $(\text{MC}^{-1}(A^{(r)}), \text{MC}^{-1}(B^{(r)}), \text{MC}^{-1}(C^{(r)}), \text{MC}^{-1}(D^{(r)}))$.

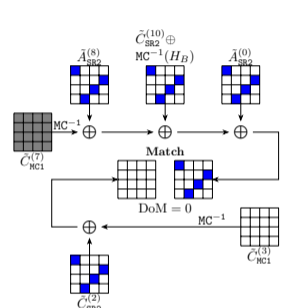
Remark. This is equivalent to searching for the MitM characteristic of the linear transformation of Preimage (or Collision).

Generalization of Sasaki's Matching Strategy for Feistel

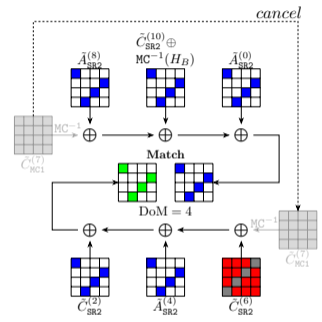
- Full Round match of Simpira-4



(a) (6,4)-round matching



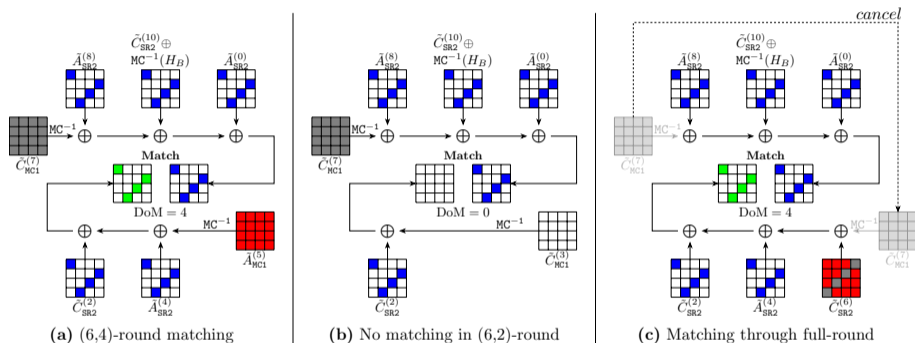
(b) No matching in (6,2)-round



(c) Matching through full-round

Generalization of Sasaki's Matching Strategy for Feistel

- Full Round match of Simpira-4



Advantage

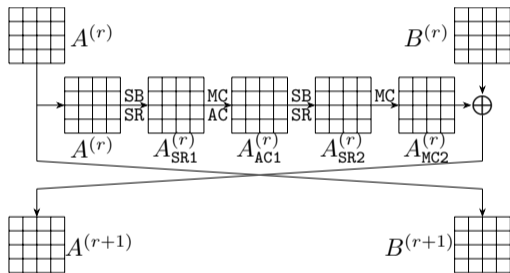
- 1 Initial structure preserves more useful information or even can be canceled.
- 2 Only Xor operation remained is friendly to program.

Outline

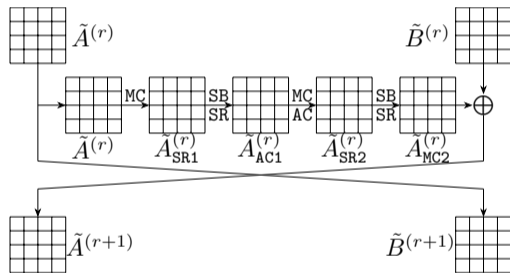
- 1 Background and preliminary
- 2 Improved matching strategy for MitM model
- 3 Preimage attack on 7-round Simpira-2**
- 4 Conclusion

Preimage attack on 7-round Simpira-2

- Description of Simpira-2

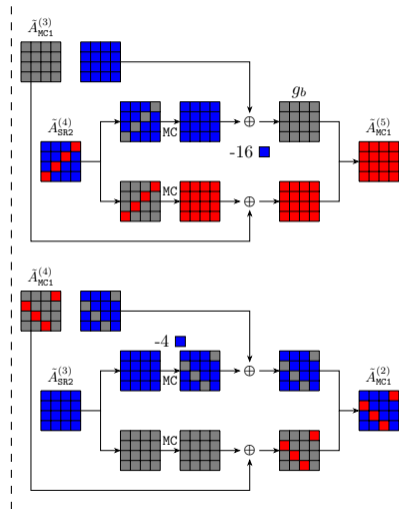
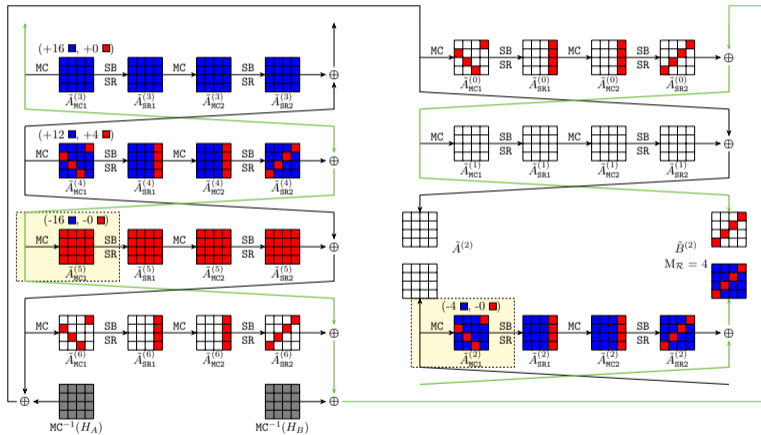


(a) The round function of Simpira-2

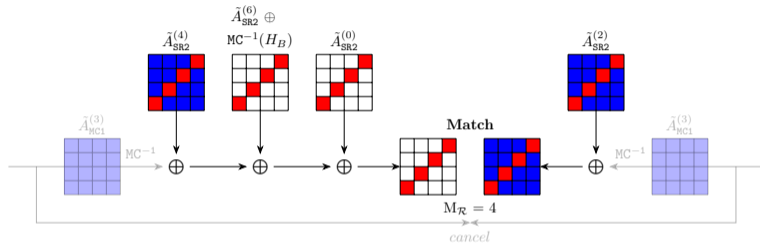


(b) Equivalent transform of Simpira-2

Preimage attack on 7-round Simpira-2



Preimage attack on 7-round Simpira-2



Cplx.

- $\lambda_B = 28$ ■, $\lambda_{\mathcal{R}} = 4$ ■, $l_B = 20$, $l_{\mathcal{R}} = 0$
- $\text{DoF}_B = \lambda_B - l_B = 8$, $\text{DoF}_{\mathcal{R}} = \lambda_{\mathcal{R}} - l_{\mathcal{R}} = 4$, $\text{DoM} = 0$, $M_B = 0$, $M_{\mathcal{R}} = 4$

$$T = 2^{8 \times 28} + 2^{8 \times (32 - \min\{8, 4, 4\})} \approx 2^{225}$$

Summary of applications to preimage and collision attacks

Target	Attacks	Settings	Rounds	Time	Memory	Generic	Ref.
Feistel-SP-128	Preimage	Classical	11	2^{112}	2^{24}	2^{128}	[ACNS:Sas13]
		Classical	12	2^{113}	2^{48}	2^{128}	[This]
Simpira-2	Preimage	Classical	5/15	2^{128}	-	2^{256}	[C:SS22]
		Quantum	5/15	2^{64}	-	2^{128}	[C:SS22]
		Classical	7/15	2^{225}	2^{96}	2^{256}	[This]
Simpira-4	Preimage	Classical	9/15	2^{128}	-	2^{256}	[C:SS22]
		Quantum	9/15	2^{64}	-	2^{128}	[C:SS22]
		Classical	11/15	2^{225}	2^{160}	2^{256}	[This]
Simpira-6	Preimage	Classical	11/15	$2^{193.6}$	2^{193}	2^{256}	[This]
Lesamnta-LW	Collision	Classical	11/64	2^{97}	2^{96}	2^{128}	[ICISC:HIK10]
		Classical	17/64	$2^{113.58}$	2^{112}	2^{128}	[This]
		Classical	20/64	2^{124}	2^{124}	2^{128}	[This]
Areion256-DM	Preimage	Classical	5/10	2^{248}	2^8	2^{256}	[CHES:IIL23]
		Classical	5/10	2^{193}	2^{88}	2^{256}	[This]
		Classical	7/10	2^{240}	2^{64}	2^{256}	[This]
Areion512-DM	Preimage	Classical	10/15	2^{248}	2^8	2^{256}	[CHES:IIL23]
		Classical	11/15	2^{241}	2^{48}	2^{256}	[This]

Outline

- 1 Background and preliminary
- 2 Improved matching strategy for MitM model
- 3 Preimage attack on 7-round Simpira-2
- 4 Conclusion

Conclusion

- Automated MitM Preimage and Collision attack on the linear transformation of Feistel-based hash functions with full-round match.
- Detect matching bytes filter with same color to reduce memory.
- Leading to improved or first MitM attack on Simpira, Areion, and Lesamnta-LW.
- Limitation:
 - In Double-SP round function, the output can be easily unknown with only one input unknown byte, so how to avoid it in SupP states even with additional cost?

Full Version: <https://eprint.iacr.org/2023/1359.pdf>

Source Code: <https://github.com/Hql-code/MitM-Feistel.git>

Thank you!