Algebraic Attack on FHE-Friendly Cipher HERA Using Multiple Collisions

Fukang Liu[1], Abul Kalam[2], Santanu Sarkar[2], Willi Meier[3]

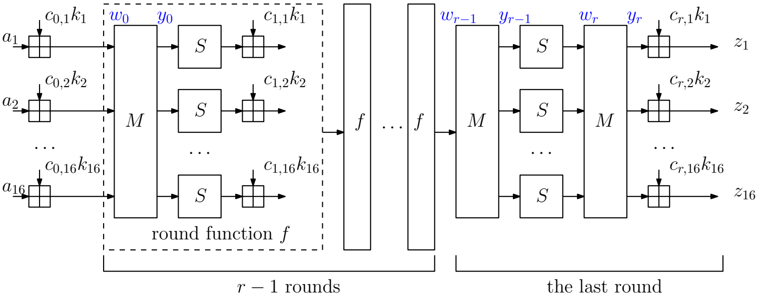
[1]Tokyo Institute of Technology, Tokyo, Japan
[2]Indian Institute of Technology Madras, Chennai, India
[3]FHNW, Windisch, Switzerland

FSE 2024

Overview

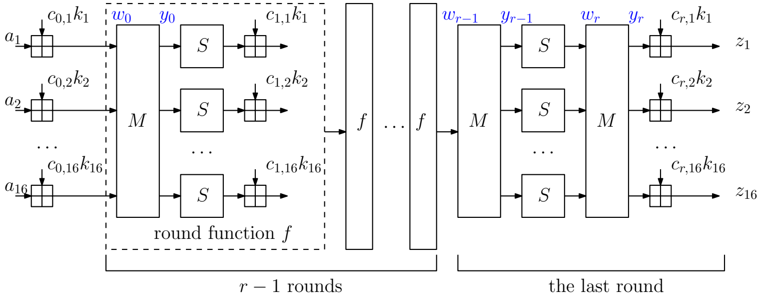
Background of HERA

- A stream cipher friendly to the CKKS FHE scheme (Asiacrypt 2021).
- SPN-based cipher with a (simple) randomized key schedule.
- No third party cryptanalysis so far.



Description of HERA

- $(a_1, \ldots, a_{16}) = (1, \ldots, 16)$.
- Cubic S-box $S(x) = x^3$ over prime fields $\mathbb{F}_p$ with $p > 2^{16}$.
- $M \in \mathbb{F}_p^{16\times 16}$ is a fixed invertible matrix.
- $k = (k_1, \ldots, k_{16}) \in \mathbb{F}_p^{16}$ is the secret key.
- $(c_{0,1}, \ldots, c_{r,16}) \in \mathbb{F}_p^{16\times(r+1)}$ are randomly generated constants.
- $z = (z_1, \ldots, z_{16}) \in \mathbb{F}_p^{16}$ is the keystream.

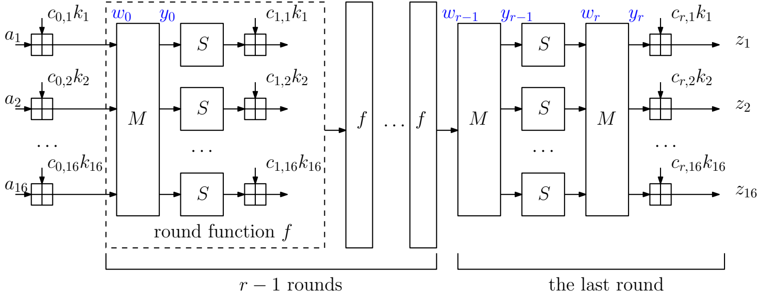


Description of HERA

- Providing $\lambda \in \{80, 128, 192, 256\}$ bits of security
- The length of nonce and cnt is related to λ:

$$\mathsf{IV} = \mathsf{nonce}||\mathsf{cnt} \in \mathbb{F}_2^{\lambda+\frac{\lambda}{2}}.$$

- Procedure to generate the keystream:
  1. Generate $(c_{0,1}, \ldots, c_{r,16}) \in \mathbb{F}_p^{16\times(r+1)}$ seeded with IV.
  2. Generate the keystream by running the encryption algorithm.



Revisiting Designers' Analysis

- Straightforward linearization attack:
  - $z_i = f_{\mathrm{IV}}(k_1, \ldots, k_{16})$ where f_{IV} is of degree 3^r.
  - Use sufficiently many IV to generate about $\binom{16+3^r}{3^r}$ equations $z_i = f_{\mathrm{IV}}(k_1, \ldots, k_{16})$.
  - Solve the equations in $(k_1, \ldots, k_{16})$ by simple Gaussian elimination (each monomial is renamed as a new variable).
- Reason:
  - At most $\binom{16+3^r}{3^r}$ monomials for a polynomial in 16 variables of degree 3^r.

Revisiting Designers' Analysis

- Complexity analysis:
  - Time complexity of the linearization attack on r-round HERA:

$$\mathcal{T}(r,\omega) = \binom{16+3^r}{3^r}^{\omega},$$

  where $2 \leq \omega \leq 3$ is the algebra constant.
- Secure parameters:
  - Select the minimal r such that

$$\mathcal{T}(r,2) = \binom{16+3^r}{3^r}^{2} > 2^{\lambda}.$$

Revisiting Designers' Analysis

- Parameters for HERA:

λ	80	128	192	256
r	4	5	6	7

- Cost to break r rounds of HERA with different (r, ω):

λ	80	128	192	256
r	4	5	6	7
brute force	p^{16}	p^{16}	p^{16}	p^{16}
$T_0(r, 2)$	2^{119}	2^{167}	2^{217}	2^{267}
$T_0(r-1, 2)$	2^{76}	2^{119}	2^{167}	2^{217}
$T_0(r, 2.8)$	2^{167}	2^{234}	2^{303}	2^{374}
$T_0(r-1, 2.8)$	2^{107}	2^{167}	2^{234}	2^{303}
$T_0(r-2, 2.8)$	2^{59}	2^{107}	2^{167}	2^{234}
$T_0(r, 3)$	2^{179}	2^{251}	2^{325}	2^{401}
$T_0(r-1, 3)$	2^{114}	2^{179}	2^{251}	2^{325}
$T_0(r-2, 3)$	2^{63}	2^{114}	2^{179}	2^{251}

Observations

If we can can set up equations of degree 3^{r-1} for r-round HERA:

1. HERA can be broken under $\omega = 2$.
2. Security margin will be reduced to 1 round under $\omega \in \{2.8, 3\}$.

λ	80	128	192	256
r	4	5	6	7
brute force	p^{16}	p^{16}	p^{16}	p^{16}
$T_0(r, 2)$	2^{119}	2^{167}	2^{217}	2^{267}
$T_0(r-1, 2)$	2^{76}	2^{119}	2^{167}	2^{217}
$T_0(r, 2.8)$	2^{167}	2^{234}	2^{303}	2^{374}
$T_0(r-1, 2.8)$	2^{107}	2^{167}	2^{234}	2^{303}
$T_0(r-2, 2.8)$	2^{59}	2^{107}	2^{167}	2^{234}
$T_0(r, 3)$	2^{179}	2^{251}	2^{325}	2^{401}
$T_0(r-1, 3)$	2^{114}	2^{179}	2^{251}	2^{325}
$T_0(r-2, 3)$	2^{63}	2^{114}	2^{179}	2^{251}

A New Attack Framework for HERA

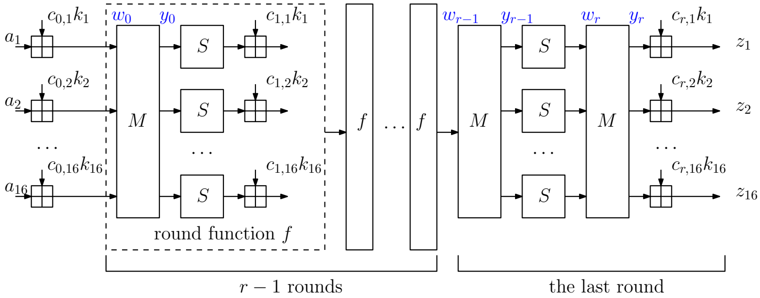
- Main idea:
  - Set up a low-degree ($< 3^r$) equation in $(k_1, \ldots, k_{16})$ from a keystream pair (z, z') rather than a single z.

- Overall procedure:
  1. **Offline phase**: Find sufficiently many good input pairs $(\mathrm{IV}, \mathrm{IV}')$ by the offline computation.
  2. **Online phase**: For each input pair $(\mathrm{IV}, \mathrm{IV}')$, compute the corresponding output pair (z, z'). If (z, z') satisfy certain conditions, we can set up some low-degree equations in k.
  3. **Solving equations**: After collecting many low-degree equations, we solve them with the linearization technique.

Analysis

- How to define good IV pairs?
  - We aim to find a good pair (c_r, c_r') generated from $(\mathrm{IV}, \mathrm{IV}')$, respectively, such that the corresponding (c_r, c_r') can satisfy certain conditions, where

$$c_r = (c_{r,1}, \ldots, c_{r,16}),\ \ c_r' = (c'_{r,1}, \ldots, c'_{r,16}).$$



Observations

For the last-round S-box, we have

$$w_{r,i} \;=\; S(y_{r-1,i}) = y_{r-1,i}^3,$$

- **Case 1**: if

$$w_{r,i} = w'_{r,i},$$

  we have

$$y_{r-1,i} = y'_{r-1,i}.$$

- **Case 2**: if $\beta \neq 0 \in \mathbb{F}_p$ is known and

$$w_{r,i} = \beta w'_{r,i}, \tag{1}$$

  we have

$$y_{r-1,i} = \beta^{\frac{1}{3}} y'_{r-1,i}.$$

- As $y_{r-1,i}$ and $y'_{r-1,i}$ are polynomials in k of degree 3^{r-1}, in both cases, we can set up an equation in k of degree 3^{r-1} for r-round HERA.

Observations

- **Goal**: check from (z, z') whether the following equation holds:

$$w_{r,i} = \beta w'_{r,i}.$$

- **Relation**:

$$\begin{aligned} w_r &= M^{-1}(z - c_r \cdot k) = M^{-1}(z) - M^{-1}(c_r \cdot k), \\ \rightarrow w_{r,i} &= M^{-1}(z)[i] - M^{-1}(c_r \cdot k)[i] \end{aligned} \tag{2}$$

where $c_r \cdot k$ denotes the element-wise multiplication.

- **Question**: $w_{r,i}$ cannot be known without guessing k, how is it even possible to check

$$w_{r,i} = \beta w'_{r,i}$$

and compute β?

Observations

$$w_{r,i} = M^{-1}(z)[i] - M^{-1}(c_r \cdot k)[i].$$

- **Our solution**: turn to checking conditions:

$$\begin{aligned} (c_{r,1}, \ldots, c_{r,16}) &= (\beta c'_{r,1}, \ldots, \beta c'_{r,16}), && (3) \\ M^{-1}(z)[i] &= \beta \times M^{-1}(z')[i], && (4) \end{aligned}$$

which requires no knowledge of k.

Observations

- **Offline phase**: (c_r, c'_r) are generated from an XOF seeded with $(\mathrm{IV}, \mathrm{IV}')$, respectively, which dose not depend on k, and hence

$$(c_{r,1}, \ldots, c_{r,16}) = (\beta c'_{r,1}, \ldots, \beta c'_{r,16})$$

can be checked at the offline phase.

- **Online phase**: computing (z, z') requires to call the encryption algorithm and hence

$$M^{-1}(z)[i] = \beta \times M^{-1}(z')[i],$$

can only be checked at the online phase.

Observations

$$w_{r,i} \quad = \quad M^{-1}(z)[i] - M^{-1}(c_r \cdot k)[i].$$

- **Goal**: compute β such that
$$w_{r,i} = \beta w'_{r,i}.$$
- **Relaxed conditions**: by guessing n_1 words, e.g., guessing $(k_1, \ldots, k_{n_1})$, we only need conditions
$$\begin{aligned} (c_{r,n_1+1}, \ldots, c_{r,16}) &= (\beta c'_{r,n_1+1}, \ldots, \beta c'_{r,16}), \quad (5) \\ M^{-1}(z)[i] - \delta &= \beta \times (M^{-1}(z')[i] - \delta'), \quad (6) \end{aligned}$$
where
$$\delta = \sum_{j=1}^{n_1} M^{-1}[i][j] c_{r,j} k_j, \quad \delta' = \sum_{j=1}^{n_1} M^{-1}[i][j] c'_{r,j} k_j.$$

Observations

- **Drawback**: Overhead caused by guessing n_1 key variables:

$$p^{n_1} \times \binom{16 - n_1 + 3^{r-1}}{3^{r-1}}^{\omega}.$$

Offline Phase

- **Goal**: find $(\mathsf{IV}, \mathsf{IV}')$ such that

$$(c_{r,n_1+1}, \ldots, c_{r,16}) \;=\; (\beta c'_{r,n_1+1}, \ldots, \beta c'_{r,16}), \tag{7}$$

  which is equivalent to finding the following collision

$$(1, \frac{c_{r,n_1+2}}{c_{r,n_1+1}}, \ldots, \frac{c_{r,16}}{c_{r,n_1+1}}) = (1, \frac{c'_{r,n_1+2}}{c'_{r,n_1+1}}, \ldots, \frac{c'_{r,16}}{c'_{r,n_1+1}}). \tag{8}$$

- **#collisions**: suppose 2^b different such collisions are required. Let

$$\ell = (15 - n_1) \times \lceil \log_2 p \rceil,$$

  we need to test $2^{\frac{b+\ell+1}{2}}$ different IV.
- **Cost**:

$$T_{\text{offline}} = 2^{\frac{b+\ell+1}{2}}. \tag{9}$$

Online Phase

Procedure:

- Generate the corresponding (z, z') under each good $(\mathrm{IV}, \mathrm{IV}')$.
- For each guess of $(k_1, \ldots, k_{n_1})$, compute

$$\delta = \sum_{j=1}^{n_1} M^{-1}[i][j] c_{r,j} k_j, \quad \delta' = \sum_{j=1}^{n_1} M^{-1}[i][j] c'_{r,j} k_j.$$

  and check the following condition

$$\exists i: \ M^{-1}(z)[i] - \delta = \beta \times (M^{-1}(z')[i] - \delta'), \qquad (10)$$

- If Eq.(10) holds (with probability of about $\frac{16}{p}$), we set up an equation of degree 3^{r-1} until in total

$$\binom{16 - n_1 + 3^{r-1}}{3^{r-1}}$$

  equations are collected.

Solving Equations

- Solve the system of

$$\binom{16 - n_1 + 3^{r-1}}{3^{r-1}}$$

equations in $(k_{n_1+1}, \ldots, k_{16})$ of degree 3^{r-1} with Gaussian elimination.

- **Cost of online phase + solving equations**:

$$T_{\text{online}} = p^{n_1} \times 2^{b+1} + p^{n_1} \times \binom{16 - n_1 + 3^{r-1}}{3^{r-1}}^{\omega}. \tag{11}$$

Additional Constraints

- **Additional constraints**: due to the length of nonce and cnt, the following constraints should be satisfied:

$$\begin{cases} 3\lambda \geq b + \ell + 1 \\ \dfrac{16}{p} \times 2^b \geq \dbinom{16 - n_1 + 3^{r-1}}{3^{r-1}} \\ b + 1 \leq \dfrac{\lambda}{2} \end{cases} \tag{12}$$

Results

Table: Summary of the time complexity of our successful attacks on various parameters of under $\omega \in \{2, 2.8, 3\}$.

λ	Rounds	ω	$\lceil \log_2 p \rceil$											
			17	18	19	20	21	22	23	24	25	26	27	28
192	6 (full)	2	2^{185}	2^{187}	–	–	–	–	–	–	–	–	–	–
	5	2.8	2^{167}	2^{175}	2^{179}	2^{180}	2^{187}	–	–	–	–	–	–	–
	5	3	2^{179}	2^{179}	2^{183}	2^{191}	–	–	–	–	–	–	–	–
256	7 (full)	2	2^{217}	2^{224}	2^{225}	2^{226}	2^{227}	2^{228}	2^{229}	2^{243}	2^{245}	2^{247}	2^{249}	2^{251}
	6	2.8	2^{234}	2^{234}	2^{234}	2^{234}	2^{234}	2^{234}	2^{234}	2^{235}	2^{243}	2^{249}	2^{250}	2^{251}
	6	3	2^{251}	2^{251}	2^{251}	2^{251}	2^{251}	2^{251}	2^{251}	2^{251}	2^{251}	2^{251}	–	–

1. HERA with $\lambda \in \{80, 128\}$ is not affected by the attacks.
2. For $\lambda \in \{192, 256\}$, we can break some parameters under $\omega = 2$.
3. For $\lambda \in \{192, 256\}$, the security of some variants of HERA are reduced to only 1 round under $\omega \in \{2.8, 3\}$.

Future Research

- Can we apply the new insight into HERA to the cryptanalysis of FHE-friendly cipher Rubato, which also takes a randomized key schedule, but has an extremely small number of rounds, e.g. 2 rounds?

- Several obstacles:
  1. larger prime fields ($p \approx 2^{26}$).
  2. larger state (16, 36, 64 state words).
  3. noise in the keystream.