Coefficient Grouping: Breaking Chaghri and More

Fukang Liu[1], Ravi Anand[2], Libo Wang[2,4], Willi Meier[5], Takanori Isobe[2,3]

[1]Tokyo Institute of Technology, Japan

[2]University of Hyogo, Japan

[3]NICT, Japan

[4]Jinan University, China

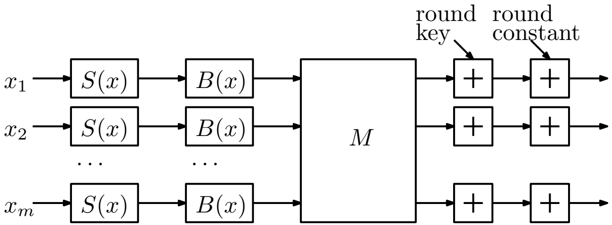
[5]FHNW, Switzerland

April 25, EUROCRYPT 2023

Outline

The Chaghri Primitive

- Proposed at ACM CCS 2022
- FHE-friendly block cipher
- Outperforms AES (in FHE setting) by 65%
- Over a large finite field $\mathbb{F}_{2^{63}}^3$



$$m = 3, \quad S(x) = x^{2^{32}+1}, \quad B(x) = c_0 x^{2^3} + c_1,$$

$$M : \text{MDS matrix}, \quad \#rounds = 16$$

Breaking and Rescuing Chaghri

- Broke Chaghri in less than 3 weeks after its publication
- Three different ways for the degree evaluation:
  - Method 1 (not tight but useful to break Chaghri)
  - Method 2 (tighter but only efficient for Chaghri)
  - Coefficient grouping (for a general construction)
- Identified countermeasures after breaking it

Impact of our attack

The designers of Chaghri have revised their designs with our proposed countermeasures:

$$B(x) = c_0 x^{2^8} + c_1 x^{2^2} + c_2 x + c_3$$

Basic Knowledge for $\mathbb{F}_{p^n}$

Polynomial basis

Let f be an irreducible polynomial over $\mathbb{F}_{p^n}$ and $f(\alpha) = 0$. Then, $\{1, \alpha, \ldots, \alpha^{n-1}\}$ is called a polynomial basis of $\mathbb{F}_{p^n}$. In this way, each element $x \in \mathbb{F}_{p^n}$ can be represented as

$$x = \sum_{i=0}^{n-1} \beta_i \alpha^i, \ \ \beta_i \in [0, p-1],$$

i.e. x is uniquely represented by $(\beta_0, \ldots, \beta_{n-1}) \in \mathbb{F}_p^n$.

Basic Knowledge for $\mathbb{F}_{p^n}$

Well-known properties

$$\begin{cases} (x+y)^{p^i} = x^{p^i} + y^{p^i},\ \forall x, y \in \mathbb{F}_{p^n}, \\ x^{p^n} = x,\ \forall x \in \mathbb{F}_{p^n}, \\ x^{p^n-1} = 1,\ \forall x \in \mathbb{F}_{p^n} \text{ and } x \neq 0. \end{cases}$$

Higher-order Differential Attack over $\mathbb{F}_{2^n}$

Algebraic degree of a univariate polynomial $\mathcal{F}(X)$ in $\mathbb{F}_{2^n}[X]$

Let

$$\mathcal{F}(X) = \sum_{i=0}^{2^n-1} u_i X^i.$$

Then, its algebraic degree $D_{\mathcal{F}}$ is defined as:

$$D_{\mathcal{F}} \quad = \quad max\{H(i) : i \in [0, 2^n - 1], u_i \neq 0\},$$

where $H(i)$ denotes the hamming weight of the integer i, i.e., the number of "1" in its binary representation.

Example

For $\mathcal{F} = X^{2^{30}+2^{31}} + X^{2^1+2^3+2^4}$, we have $D_{\mathcal{F}} = 3$.

Higher-order Differential Attack over $\mathbb{F}_{2^n}$

Higher-order differential attack over $\mathbb{F}_{2^n}$

Let

$$\mathcal{F}(X) = \sum_{i=0}^{2^n-1} u_i X^i.$$

With the polynomial basis, each X is uniquely represented by a vector $(\beta_0, \ldots, \beta_{n-1}) \in \mathbb{F}_2^n$.
In this way, we have

$$\sum_{(\beta_0,\beta_1,\ldots,\beta_{n-1})\in V} \mathcal{F}(\sum_{i=0}^{n-1} \beta_i \alpha^i) = 0 \text{ for } \mathrm{Dim}(V) \geq D_{\mathcal{F}} + 1$$

Previous Work on MiMC

- Round function: $R(x) = S(x) + K_i$ where
$$S(x) = x^3.$$
- Upper bound on the algebraic degree after r rounds:
  - $P_r(x)$: the polynomial representation after r rounds:
  - If $3^r < 2^n$, there must be
$$D_{P_r} \leq \lceil \log_2 3^r \rceil \approx r \log_2 3.$$
  - Good enough to break MiMC over $\mathbb{F}_{2^n}$.
- Follow-up work for the case $S(x) = x^{2^d+1}$: If $(2^d + 1)^r < 2^n$, there must be
$$D_{P_r} \leq \lceil \log_2 (2^d + 1)^r \rceil \approx rd.$$
  - Too loose for large d, e.g. the case of Chaghri.
  - How about larger r such that $(2^d + 1)^r \geq 2^n$, i.e., the set of exponents are within modulo $2^n - 1$.

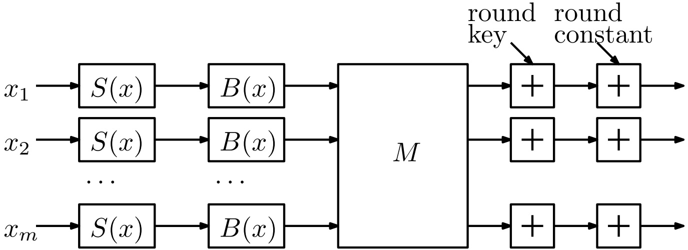
Degree Evaluation for Chaghri: Method 2

- The round function:

$$S(x) = x^{2^{32}+1}, \quad B(x) = c_0 x^{2^3} + c_1.$$

- State transitions:

$$(z_{0,1}, z_{0,2}, z_{0,3}) \to (z_{1,1}, z_{1,2}, z_{1,3}) \to \cdots \to (z_{r,1}, z_{r,2}, z_{r,3})$$



Degree Evaluation for Chaghri: Method 2

Our very naive idea:

- Step 1: set the input as a univariate polynomial in X:

$$\begin{aligned} z_{0,1} &= A_{0,1}X + B_{0,1}, \\ z_{0,2} &= A_{0,2}X + B_{0,2}, \\ z_{0,3} &= A_{0,3}X + B_{0,3}. \end{aligned}$$

  - $z_{r,i}$ is always a univariate polynomial $P_{r,i}(X) \in \mathbb{F}_{2^n}[X]$.

- Step 2: trace the evolution of $P_{r,i}$.
- Step 3: compute all possible exponents in $P_{r,i}$. (practical???)
- Step 4: find the exponent with the maximal hamming weight

Degree Evaluation for Chaghri: Method 2

Step 2: trace the evolution of polynomials

- New representation for $(z_{r,1}, z_{r,2}, z_{r,3})$

$$z_{r,1} = \sum_{i=1}^{|w_r|} A_{r,i} X^{w_{r,i}},\ \ z_{r,2} = \sum_{i=1}^{|w_r|} B_{r,i} X^{w_{r,i}}, z_{r,3} = \sum_{i=1}^{|w_r|} C_{r,i} X^{w_{r,i}}$$

- The set of all possible exponents after r rounds:

$$w_r = \{w_{r,1}, w_{r,2}, \ldots, w_{r,|w_r|}\} \subseteq \mathbb{N}, \quad w_0 = \{0, 1\}.$$

- Goal: find a relation between w_r and w_{r+1} to compute w_r iteratively.

Degree Evaluation for Chaghri: Method 2

Step 2: trace the evolution of polynomials

- Through $S(x) = x^{2^{32}+1}$:

$$
\begin{aligned}
S(z_{r,1}) &= \Big(\sum_{i=1}^{|w_r|} A_{r,i} X^{w_{r,i}}\Big)^{2^{32}+2^0} \\
&= \Big(\sum_{i=1}^{|w_r|} A_{r,i} X^{w_{r,i}}\Big)^{2^{32}} \times \Big(\sum_{i=1}^{|w_r|} A_{r,i} X^{w_{r,i}}\Big)^{2^0} \\
&= \sum_{i=1}^{|w_r|} \sum_{j=1}^{|w_r|} A_{r,i,j} X^{2^{32} w_{r,i} + 2^0 w_{r,j}}.
\end{aligned}
$$

where $A_{r,i,j} \in \mathbb{F}_{2^n}$ are key-dependent coefficients.

Degree Evaluation for Chaghri: Method 2

Step 2: trace the evolution of polynomials

- Through $B(x) = x^{2^3}$:

$$\begin{aligned} B \circ S(z_{r,1}) &= c_0 \left(\sum_{i=1}^{|w_r|} \sum_{j=1}^{|w_r|} A_{r,i,j} X^{(2^{32} w_{r,i} + 2^0 w_{r,j})} \right)^{2^3} + c_1 \\ &= \sum_{i=1}^{|w_r|} \sum_{j=1}^{|w_r|} A'_{r,i,j} X^{2^{35} w_{r,i} + 2^3 w_{r,j}}. \end{aligned}$$

- The matrix M does not affect this representation:

$$z_{r+1,1} = \sum_{i=1}^{|w_r|} \sum_{j=1}^{|w_r|} A_{r+1,i,j} X^{2^{35} w_{r,i} + 2^3 w_{r,j}}$$

Degree Evaluation for Chaghri: Method 2

Step 2: trace the evolution of polynomials

- The relation between w_r and w_{r+1} is obtained as

$$w_{r+1} = \{\mathcal{M}_{63}(e) | e = 2^{35} w_{r,i} + 2^3 w_{r,j}, 1 \le i, j \le |w_r|\},$$

where we define

$$\mathcal{M}_n(x) = \begin{cases} 2^n - 1 \text{ if } 2^n - 1 | x, x \ge 2^n - 1, \\ x \% (2^n - 1) \text{ otherwise.} \end{cases}$$

due to

$$\begin{cases} x^{2^n} = x \ \forall x \in \mathbb{F}_{2^n}, \\ x^{2^n - 1} = 1 \ \forall x \in \mathbb{F}_{2^n} \text{ and } x \neq 0. \end{cases}$$

- Why previous methods failed: they can not handle the modular addition!!!

Degree Evaluation for Chaghri: Method 2

Step 2: trace the evolution of polynomials

- The relation between w_r and w_{r+2} is obtained as

$$
\begin{aligned}
w_{r+1} &= \{\mathcal{M}_{63}(e) \mid e = 2^{35} w_{r,i} + 2^3 w_{r,j}, 1 \le i, j \le |w_r|\}, \\
w_{r+2} &= \{\mathcal{M}_{63}(e) \mid e = 2^{35}(2^{35} w_{r,i} + 2^3 w_{r,j}) + 2^3(2^{35} w_{r,s} + 2^3 w_{r,t}), 1 \le i, j, s, t \le |w_r|\}, \\
&= \{\mathcal{M}_{63}(e) \mid e = 2^{38}(w_{r,i} + w_{r,s}) + 2^7 w_{r,j} + 2^6 w_{r,t}, 1 \le i, j, s, t \le |w_r|\},
\end{aligned}
$$

- Why we consider w_{r+2}: 2 rounds are treated as 1 round in Chaghri.

Throughout this slide, we have

$$w_r = \{w_{r,1}, w_{r,2}, \ldots, w_{r,|w_r|}\}.$$

Degree Evaluation for Chaghri: Method 2

Step 3: Compute w_r

- Initial set:

$$w_0 = \{0, 1\}.$$

- Compute w_{r+2} with

$$\begin{aligned} w_{r+2} \quad = \quad & \{\mathcal{M}_{63}(e) \mid e = 2^{38}(w_{r,i} + w_{r,s}) + 2^7 w_{r,i} + 2^6 w_{r,t}, \\ & 1 \le i, j, s, t \le |w_r|\}. \end{aligned}$$

- Naive enumeration quickly becomes impractical as $|w_r|$ is too large even for small r.

Degree Evaluation for Chaghri: Method 2

Step 3: Compute w_r

- New observation:

$$w_r \subseteq \{e = e^H \vee e^L \mid e^H \wedge e^L = 0, e^H \in w_r^H, e^L \in w_r^L\}.$$

- w_r^H and w_r^L are much smaller (computed independently).
- Practically compute w_r^H and w_r^L for $r = 16$!!!

Degree Evaluation for Chaghri: Method 2

Step 4: Find the element with the maximal hamming weight in w_r

- The relation:

$$w_r \subseteq \{e = e^H \vee e^L \mid e^H \wedge e^L = 0, e^H \in w_r^H, e^L \leq w_r^L\},$$
$$w_0 = \{0,1\}, \quad w_0^H = \{0\}, \quad w_0^L = \{0,1\}$$

- The maximal hamming weight:

$$max\{H(i) \mid i \in w_r^H\} + max\{H(i) \mid i \in w_r^L\}.$$

- Get the upper bound (37): break Chaghri with time $O(2^{38})$.

Chaghri is broken. But not yet over. Not elegant enough!!!

Coefficient Grouping Technique

Motivation

- Do we really need to compute w_r round by round?
- The method is too dedicated for the parameters of Chaghri, i.e. $S(x), B(x)$.
- Can we have a more elegant and general method that can work for any

$$S(x) = x^{2^{k_0}+2^{k_1}}, B(x) = c_1 x^{2^{k_2}} + c_2$$

and a general finite field $\mathbb{F}_{2^n}$?

Coefficient Grouping Technique

Using $S(x) = x^{2^{k_0}+2^{k_1}} \in \mathbb{F}_{2^n}[x], \quad B(x) = c_1 x^{2^{k_2}} + c_2 \in \mathbb{F}_{2^n}[x]$

- Relation between w_r and w_{r+1}:

$$w_{r+1} = \{\mathcal{M}_n(e) \mid e = 2^{k_0+k_2} w_{r,i} + 2^{k_1+k_2} w_{r,j}, 1 \leq i,j \leq |w_r|\}$$

- Relation between w_r and w_{r+2}:

$$\begin{aligned} w_{r+2} & \\ = \quad & \{\mathcal{M}_n(e) \mid e = 2^{k_0+k_2}(2^{k_0+k_2} w_{r,i} + 2^{k_1+k_2} w_{r,j}) + 2^{k_1+k_2}(2^{k_0+k_2} w_{r,s} + 2^{k_1+k_2} w_{r,t}), \\ & 1 \leq i,j,s,t \leq |w_r|\} \\ = \quad & \{\mathcal{M}_n(e) \mid e = 2^{2k_0+2k_2} w_{r,i} + 2^{k_0+k_1+2k_2}(w_{r,j} + w_{r,s}) + 2^{2k_1+2k_2} w_{r,t}, \\ & 1 \leq i,j,s,t \leq |w_r|\}. \end{aligned}$$

Coefficient Grouping Technique

Using $S(x) = x^{2^{k_0}+2^{k_1}} \in \mathbb{F}_{2^n}[x], \quad B(x) = c_1 x^{2^{k_2}} + c_2 \in \mathbb{F}_{2^n}[x]$

- Relation between w_r and $w_{r+\ell}$:

$$w_{r+\ell} = \quad \{\mathcal{M}_n(e) \mid e = \sum_{i=1}^{N_{n-1}} 2^{n-1} w_{r,d_{i,n-1}} + \sum_{i=1}^{N_{n-2}} 2^{n-2} w_{r,d_{i,n-2}} + \ldots + \sum_{i=1}^{N_0} 2^0 w_{r,d_{i,0}},$$
$$\text{where } 1 \leq d_{i,j} \leq |w_r| \text{ for } 0 \leq j \leq n-1\}.$$

- Group all possible N_j coefficients sharing the same factor 2^j:

$$w_{r,d_{1,j}}, w_{r,d_{2,j}}, \ldots, w_{r,d_{N_j,j}} \in w_r \ (r = 0, \ w_0 = \{0,1\}),$$

  i.e., in the formula of e, $2^j w_{r,d_{i,j}}$ is possible to appear

- $w_{r+\ell}$ is fully described by a vector $(N_{n-1}, \ldots, N_0)$ and w_r.

Coefficient Grouping Technique

New representation of w_r

- $r = 0$:

$$\begin{aligned} w_0 &= \{0,1\} = \{\mathcal{M}_n(e) \mid e = 2^0 w_{0,i}, 1 \leq i \leq 2 = |w_0|\}, \\ &\rightarrow (N_{0,n-1}, \ldots, N_{0,1}) = (0, \ldots, 0), \quad N_{0,0} = 1. \end{aligned}$$

- Relation between w_r and w_{r+1}:

$$w_{r+1} = \{\mathcal{M}_n(e) \mid e = 2^{k_0+k_2} w_{r,i} + 2^{k_1+k_2} w_{r,j}, 1 \leq i,j \leq |w_r|\}$$

- Find $(N_{r,n-1}, \ldots, N_{r,0})$ to represent w_r:

$$N_{r+1,i} = N_{r,(i-(k_1+k_2))\%n} + N_{r,(i-(k_0+k_2))\%n} \text{ for } 0 \leq i \leq n-1.$$

- $(N_{r,n-1}, \ldots, N_{r,0})$ can be computed in time $O(n)$.

Coefficient Grouping Technique

Finding two representations of w_r

- Representation 1 of w_r:

$$w_r = \Big\{\mathcal{M}_n(e) \mid e = \sum_{i=1}^{N_{r,n-1}} 2^{n-1} w_{r,d_{i,n-1}} + \sum_{i=1}^{N_{r,n-2}} 2^{n-2} w_{r,d_{i,n-2}} + \ldots + \sum_{i=1}^{N_{r,0}} 2^0 w_{r,d_{i,0}},$$
$$\text{where } 1 \leq d_{i,j} \leq |w_0| \text{ for } 0 \leq j \leq n-1 \text{ and } w_0 = \{0,1\}\Big\}.$$

- For each term 2^j, there are N_j possible coefficients

$$w_{r,d_{1,j}}, w_{r,d_{2,j}}, \ldots, w_{r,d_{N_j,j}} \in w_0 = \{0,1\},$$

which implies $\sum_{i=1}^{N_{r,j}} 2^j w_{r,d_{i,j}} \in \{2^j \gamma_j \mid 0 \leq \gamma_j \leq N_{r,j}\}$.

Coefficient Grouping Technique

Finding $e \in w_r$ with $H(e)$ maximal

- Representation 2 of w_r:

$$w_r = \{\mathcal{M}_n(e) \mid e = \sum_{i=0}^{n-1} 2^i \gamma_i,\ 0 \le \gamma_i \le N_{r,i}\}.$$

- Problem reduction (optimization problem):

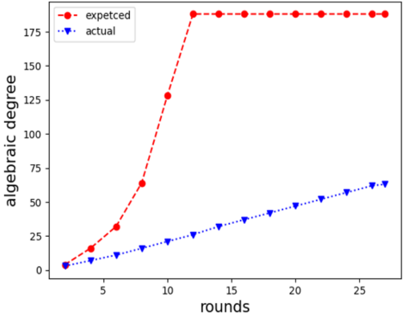
$$\begin{aligned} \text{maximize} \quad & H\Big(\mathcal{M}_n(\sum_{i=0}^{n-1} 2^i \gamma_i)\Big), \\ \text{subject to} \quad & 0 \le \gamma_i \le N_{r,i} \text{ for } i \in [0, n-1]. \end{aligned}$$

- Solved in time $O(n)$!!! or by blackbox solvers.
  - finding and proving the $O(n)$ algorithm require significant additional work

Breaking Chaghri and even More rounds

Table: The upper bounds of the algebraic degree for Chaghri

r	0	2	4	6	8	10	12	14	16	18	20	22	24	25	26
deg	1	3	7	12	17	22	27	32	37	42	47	52	58	60	63



Rescuing Chaghri

Achieving an (almost) exponential degree growth

- The slow growth is mainly caused by a sparse polynomial of $B(x)$, i.e. $B(x) = c_0 x^{2^3} + c_1$
- Reason: the growth of the number of possible monomials is highly related to the density of $B(x)$
  - requires significant additional work
- Intuition: more possible monomials, higher probability that a monomial with $\deg = 2^r$ appears
- Use $B(x) = c_0 x^{2^8} + c_1 x^{2^2} + c_2 x + c_3$ instead

Conclusion

- An efficient degree evaluation technique in time $O(n)$ for a special cipher over $\mathbb{F}_{2^n}$
- Be careful of the symmetric-key primitive design over a large finite field! (less understood)
- Open problem: other novel cryptanalytic techniques for ciphers over a large finite field