Collision Attacks on Round-Reduced SHA-3 Using Conditional Internal Differentials

Zhongyi Zhang[1,2] Chengan Hou[1,2] Meicheng Liu[1,2]

[1]SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences, China

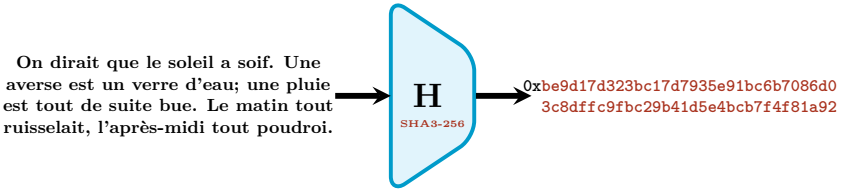
[2]School of Cyber Security, University of Chinese Academy of Sciences, China

Lyon, France
Eurocrypt 2023

Outline

Outline

Hash Function



- A cryptographic hash function is a mathematical algorithm that maps an **arbitrary length input** (the message M) to a **fixed length d-bit output**.
- Security goals
  - Pre-image resistance
  - Second pre-image resistance
  - Collision resistance
    - It should be difficult to find two message (m, m') such that $H(m) = H(m')$

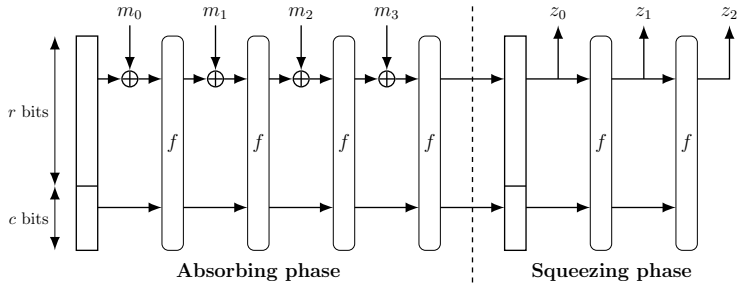
Keccak

- NIST SHA-3 hash function competition (2007-2012)
- Designers: Guido Bertoni, Joan Daemen, Michaël Peeters and Gilles Van Assche
- Submitted to SHA-3 competition in 2008
- The winner was announced to be Keccak in 2012
- In 2015, Keccak was standardized by NIST as SHA-3
  - SHA3-224/256/384/512
  - SHAKE128/256 (eXtendable Output Functions, XOFs)

TeamKeccak

Guido Bertoni[3], Joan Daemen[2], Seth Hoffert, Michaël Peeters[1], Gilles Van Assche[1] and Ronny Van Keer[1]
[1]STMicroelectronics - [2]Radboud University - [3]Security Pattern

Structure of Sponge construction



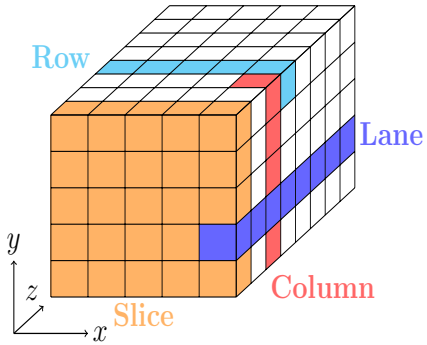
- b-bit permutation Keccak-f, f contains 24 rounds
- Two parameters: bitrate r and capacity c, $b = r + c$
- The message is padded and then split into r-bit blocks

The Internal State

- 1600 bits: seen as a 5×5 matrix, where each cell is a 64-bit lane in the direction of the z axis
  $A[x, y], 0 \leq x, y < 5$
- each round R consists of five steps:

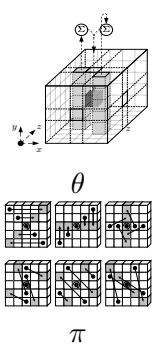
$$R = \iota \circ \chi \circ \pi \circ \rho \circ \theta, L \triangleq \pi \circ \rho \circ \theta$$

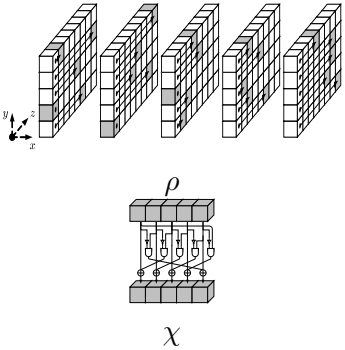
- χ: the only nonlinear operation, a 5-bit Sbox applies to each row



Round Function

- ι step: Adding one round-dependent constant to the first "lane", to destory the symmetry. It cannot be ignored in internal differential cryptanalysis.





Internal Differential Cryptanalysis

Developed by Thomas Peyrin (Crypto 2010) and generalized by Itai Dinur, Orr Dunkelman and Adi Shamir (FSE 2013) in the analysis of Keccak.

Collision Attacks on Up to 5 Rounds of SHA-3 Using Generalized Internal Differentials

- Squeeze Attack
- Internal Differential Characteristic
- Target Internal Difference Algorithm
- Practical Results:
  - 3-round Keccak-384
  - 3-round Keccak-512
- Theoretical Results:
  - 5-round Keccak-256
  - 4-round Keccak-384

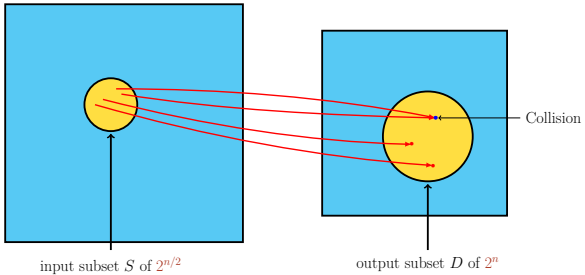
Collision Attacks on Round-Reduced SHA-3

Methods	SHA3-224	SHA3-256	SHA3-384	SHA3-512	SHAKE128	SHAKE256
Differential or SAT-based	2 (practical)	2 (practical)	-	-	-	-
Differential [DDS12]	4 (practical)	4 (practical)	-	-	-	-
Internal Diff. [DDS13]	-	5 (2^{115})	3 (practical) 4 (2^{147})	3 (practical)	-	-
Algebriac Diff. [GLLLQS20]	5 (practical)	5 (practical)	-	-	5 (practical)	-
Poly. Systems [Dinur21]	-	-	-	$4^{*}(2^{263})$	-	-
SAT-based Diff. [HBDM22]	-	-	4 ($2^{59.64}$)	-	-	-
SAT and Quant. [GLST22]	$6^{\dagger}(2^{96.15})$	$6^{\dagger}(2^{102.65})$	-	-	6 ($2^{123.5}$) $6^{\dagger}(2^{61.4})$	-
Conditional Internal Diff.	5 (2^{105})	5 (2^{105})	4 (2^{76})	4 (2^{237})	5 (2^{105})	4 (2^{76}) 5 (2^{185})

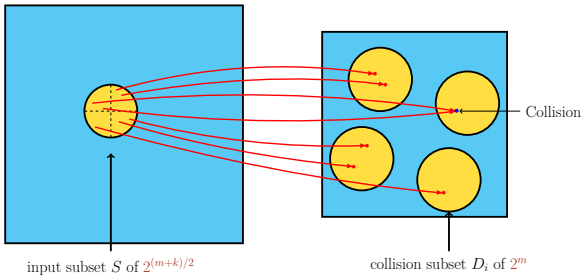
* Calculated by bit operation
† Quantum

Outline

- The general birthday attack
  - A hash function $H : \mathbb{F}_2^* \to \mathbb{F}_2^d$ maps one input subset S into one output subset D of size 2^n.
  - To find a collision, $2^{n/2}$ inputs from S are needed.



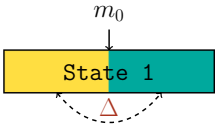
- A variant of birthday attack
  - A hash function $H : \mathbb{F}_2^* \to \mathbb{F}_2^n$ maps input subset S into 2^k output subsets $D_1, \ldots, D_{2^k}$ (collision subsets), D_j are both pairwise disjoint. $|D_j| = 2^m$.
  - To find a collision, $2^{(m+k)/2}$ inputs from S are needed.
- Parallel search can be performed on multiple subsets D_i
- Smaller hash table size



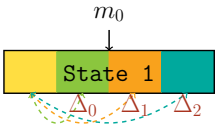
Internal Differential Cryptanalysis

- Standard differential cryptanalysis [BS91]:

- Internal differential cryptanalysis [Peyrin10]:

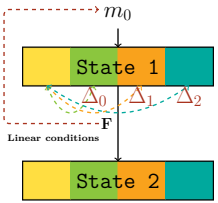


- Generalized internal differential cryptanalysis [DDS13]:



Internal Differential Cryptanalysis

- Conditional internal differential cryptanalysis (this work):



Find messages conforming 2-round internal differential characteristic by adding linear conditions to the initial state space.

Symmetric States

- One state has **period** i in the z-axis is called a **symmetric state**

$$A[x][y][(z+i) \mod 64] = A[x][y][z], 0 \leq x, y < 5, 0 \leq z < 64$$

- The fundamental period corresponding to i is gcd (i, 64), i can attain non-trivial values $\{1, 2, 4, 8, 16, 32\}$

Examples

A symmetric state with i = 16

```
|2023202320232023|746a746a746a746a|b82eb82eb82eb82e|5642564256425642|6d586d586d586d58|
|0714071407140714|934a934a934a934a|858c858c858c858c|75cb75cb75cb75cb|9e8d9e8d9e8d9e8d|
|6d586d586d586d58|0255025502550255|dd9ddd9ddd9ddd9d|fce0fce0fce0fce0|4a064a064a064a06|
|8482848284828482|3e993e993e993e99|df29df29df29df29|7e547e547e547e54|2013201320132013|
|49ea49ea49ea49ea|f441f441f441f441|e371e371e371e371|c6d9c6d9c6d9c6d9|3541354135413541|
```

Symmetric states maintain symmetric after applying the θ, ρ, π, χ

a_1	a_1	a_1	a_1	b_1	b_1	b_1	b_1	c_1	c_1	c_1	c_1	d_1	d_1	d_1	d_1	e_1	e_1	e_1	e_1
f_1	f_1	f_1	f_1	g_1	g_1	g_1	g_1	h_1	h_1	h_1	h_1	i_1	i_1	i_1	i_1	j_1	j_1	j_1	j_1
k_1	k_1	k_1	k_1	l_1	l_1	l_1	l_1	m_1	m_1	m_1	m_1	n_1	n_1	n_1	n_1	o_1	o_1	o_1	o_1
p_1	p_1	p_1	p_1	q_1	q_1	q_1	q_1	r_1	r_1	r_1	r_1	s_1	s_1	s_1	s_1	t_1	t_1	t_1	t_1
u_1	u_1	u_1	u_1	v_1	v_1	v_1	v_1	w_1	w_1	w_1	w_1	x_1	x_1	x_1	x_1	y_1	y_1	y_1	y_1

$\downarrow \theta, \rho, \pi, \chi$

a_2	a_2	a_2	a_2	b_2	b_2	b_2	b_2	c_2	c_2	c_2	c_2	d_2	d_2	d_2	d_2	e_2	e_2	e_2	e_2
f_2	f_2	f_2	f_2	g_2	g_2	g_2	g_2	h_2	h_2	h_2	h_2	i_2	i_2	i_2	i_2	j_2	j_2	j_2	j_2
k_2	k_2	k_2	k_2	l_2	l_2	l_2	l_2	m_2	m_2	m_2	m_2	n_2	n_2	n_2	n_2	o_2	o_2	o_2	o_2
p_2	p_2	p_2	p_2	q_2	q_2	q_2	q_2	r_2	r_2	r_2	r_2	s_2	s_2	s_2	s_2	t_2	t_2	t_2	t_2
u_2	u_2	u_2	u_2	v_2	v_2	v_2	v_2	w_2	w_2	w_2	w_2	x_2	x_2	x_2	x_2	y_2	y_2	y_2	y_2

Internal Difference Sets

- Given a period i, the set by adding a single **representative state v** to all symmetric states is an **internal difference set** (internal difference)

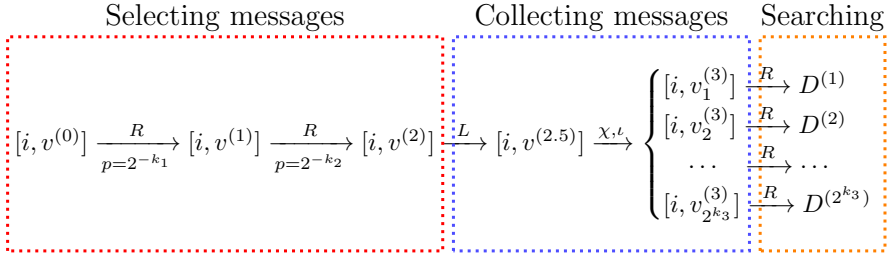
$$[i, v] \triangleq \{v + w \mid w \text{ is symmetric with period } i\}$$

- The **zero internal difference** is the set of all symmetric states with period i

$$[i, \mathbf{0}] = \{w \mid w \text{ is symmetric with period } i\}$$

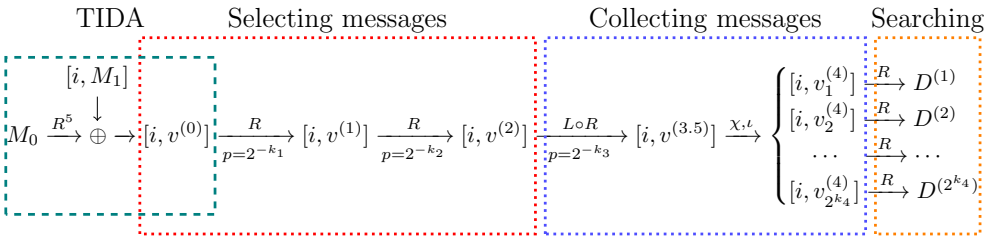
- The action of linear mapping L on any internal difference is equivalent to acting on the representative state

$$L([i, v]) = [i, L(v)]$$



4-round collision attack on SHA-3

- Select M in the initial internal difference such that the state enters the internal difference of the second round in a given characteristic
- Calculate the internal difference of M after 3 rounds of round function and store the state into the subset $[i, v_j^{(3)}]$
- Calculate the collision subset of each subset $[i, v_j^{(3)}]$ in turn until one collision is found in a certain collision subset $D^{(j)}$



5-round collision attack on SHA-3

- Select M_0 and M_1 by TIDA such that the state enters the internal difference of the second round in a given characteristic
- Calculate the internal difference of M after 4 rounds of round function and store the state into the subset $[i, v_j^{(4)}]$
- Calculate the collision subset of each subset $[i, v_j^{(4)}]$ in turn until one collision is found in a certain collision subset $D^{(j)}$

Outline

Estimate Transition Probability

Given the input difference $\delta_{in} = (\delta_0, \delta_1, \delta_2, \delta_3, \delta_4)$ of the Keccak Sbox, the output difference δ_{out} is determined by q $(2 \leq q \leq 4)$ linear conditions $\{l_t(x)\}_{t=0}^{q-1}$ (denoted as differential transition conditions) with respect to the actual input x.

$$\delta_{out} = S(x) \oplus S(x \oplus \delta_{in}) = C \cdot x \oplus S(\delta_{in})$$

$$C = \begin{pmatrix} & \delta_2 & \delta_1 & & \\ & & \delta_3 & \delta_2 & \\ & & & \delta_4 & \delta_3 \\ \delta_4 & & & & \delta_0 \\ \delta_1 & \delta_0 & & & \end{pmatrix}$$

- Difference Conditions Table (DCT)

$$\delta_{in} = 0x3,\ C \cdot x = \begin{pmatrix} & 1 & & \\ & & & 1 \\ 1 & 1 & & \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \simeq l_t(x) = \begin{pmatrix} x_4 \\ x_2 \\ x_1 + x_0 \end{pmatrix}$$

- Values of Difference Conditions Table (VDCT)

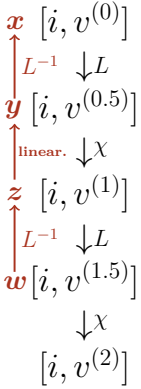
δ_{out}	0x0b	0x1b	0x0a	0x1a	0x03	0x13	0x02	0x12
$l_t(x)$	$(0,0,0)$	$(0,0,1)$	$(0,1,0)$	$(0,1,1)$	$(1,0,0)$	$(1,0,1)$	$(1,1,0)$	$(1,1,1)$

- The rank of linear equation system E composed of all differential transition conditions of an internal difference is the transition condition number. The transition probability is lower bounded by 2^{-k} ($rank(E) = k$).

Finding Messages Conforming 2-Round Internal Differential Characteristic

- Symmetric states x, y, z, w
- Find x such that y, w satisfy the respective differential transition conditions (denoted as $E_0(y)$ and $E_1(w)$)

$$L(x) = y, E_0(y) = E_0(L(x))$$



Finding Messages Conforming 2-Round Internal Differential Characteristic

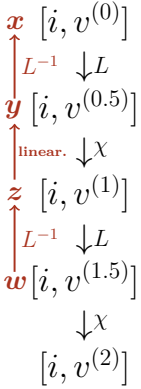
- Symmetric states x, y, z, w
- Find x such that y, w satisfy the respective differential transition conditions (denoted as $E_0(y)$ and $E_1(w)$)

$$L(x) = y, E_0(y) = E_0(L(x))$$

- Linearization of $z_0 = y_0 \oplus (y_1 \oplus 1) \cdot y_2$ by fixing y_1

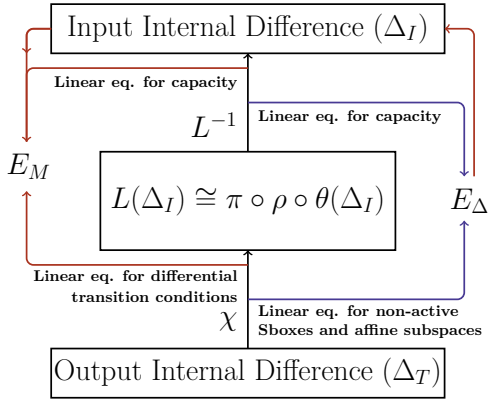
$$z_0 = \begin{cases} y_0 \oplus y_2, & y_1 = 0, \\ y_0 \quad\quad , & y_1 = 1. \end{cases}$$

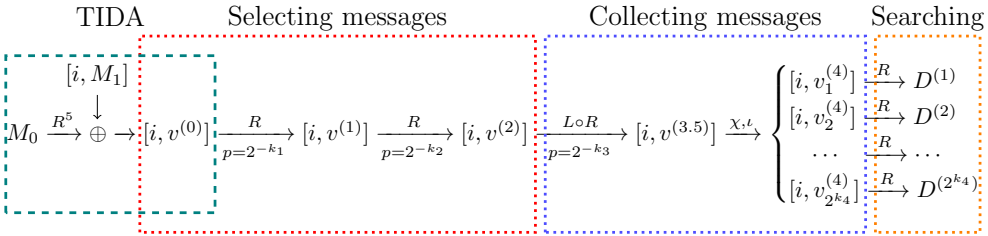
- Linearize each z_i in $E_1(w) = E_1(L(z))$ such that $E_1(w)$ is a linear equation system in x



Target Internal Difference Algorithm (TIDA)

- TDA was introduced by Dinur, Dunkelman and Shamir (FSE 2012) and developed into TIDA (FSE 2013) in the analysis of Keccak
- We separate the difference phase and value phase
- We reduce the number of interations for the two phases





5-round collision attack

- Calculate the transition probability $p = 2^{-k_i}$
- Select M_1 that can pass the first two rounds internal differential characteristic
- Select M_0 and $[i, v^{(0)}]$ by TIDA

Outline

Results of Attacks on Reduced SHA-3

Target	n_r	i	DF†	k_1	k_2	k_3	k_4	Complexity ($\log_2$)
SHA3-384	4	8	100	11	8	78	-	$79-3=76$
SHA3-512	4	32	284	16	16	170	-	$238-1=237$
SHAKE256	4	8	130	16	8	78	-	$79-3=76$
SHA3-224/256/SHAKE128	5	32	≥ 540	-	21	18	16	$106-1=105$
SHAKE256	5	32	538	-	21	18	16	≤ 185

† Degree of freedom of the initial message space.

Table: The parameters of characteristics and complexities

Summary and Future Work

- Summary
  - Utilize conditional internal differential to find collisions for up to 5 rounds of all the six SHA-3 functions
  - Present the **first collision attacks** on 4- and 5-round **SHAKE256**
  - and the best collision attack on 4-round **SHA3-512**

Summary and Future Work

- Summary
  - Utilize conditional internal differential to find collisions for up to 5 rounds of all the six SHA-3 functions
  - Present the **first collision attacks** on 4- and 5-round **SHAKE256**
  - and the best collision attack on 4-round **SHA3-512**
- Future work
  - Find **better** internal differential characteristics
  - Apply internal differential analysis to other ciphers

Thank you for your attention!