

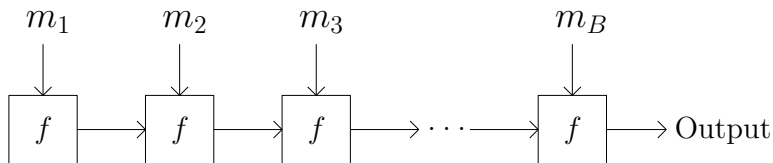
Time-Space Tradeoff for Collision Finding in Sponge Functions

Xiaoqi Duan, Akshima, Siyao Guo, Qipeng Liu

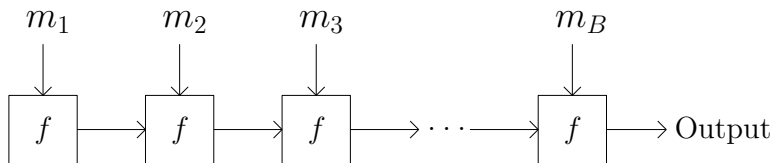
ETH Zürich; NYU Shanghai; UC San Diego

November 30, 2023

Domain Extension Hash Functions



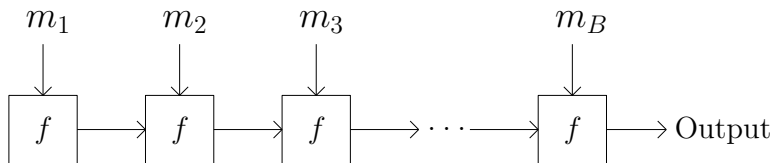
Domain Extension Hash Functions



SHA-2: Merkle-Damgård Hash Functions

$f : [N] \rightarrow [M] \quad [M] \neq [N]$ is a random function

Domain Extension Hash Functions



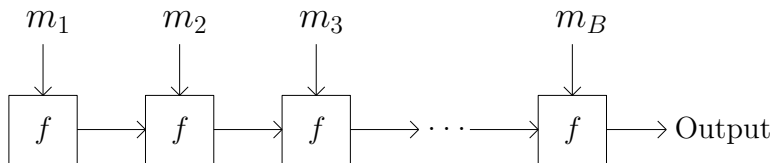
SHA-2: Merkle-Damgård Hash Functions

$f : [N] \rightarrow [M] ! [N]$ is a random function

SHA-3: Sponge Hash Functions

$f : [C] \rightarrow [R] ! [C] \rightarrow [R]$ is a random permutation

Domain Extension Hash Functions



SHA-2: Merkle-Damgård Hash Functions

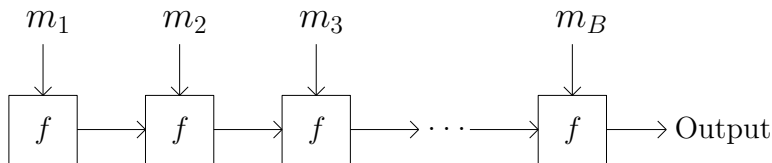
$f : [N] \rightarrow [M] ! [N]$ is a random function

SHA-3: Sponge Hash Functions

$f : [C] \rightarrow [R] ! [C]$ $[R]$ is a random permutation

Can access both f and f^{-1}

Domain Extension Hash Functions



SHA-2: Merkle-Damgård Hash Functions

$f : [N] \rightarrow [M] \mid [N]$ is a random function

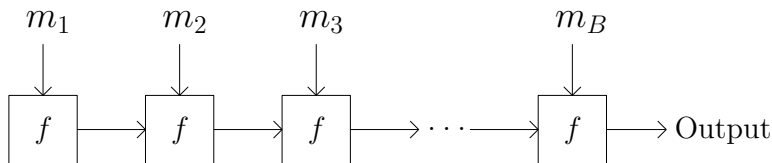
SHA-3: Sponge Hash Functions

$f : [C] \rightarrow [R] \mid [C]$ is a random permutation

Can access both f and f^{-1}

Target: Hard to find collisions

Domain Extension Hash Functions



SHA-2: Merkle-Damgård Hash Functions

$f : [N] \rightarrow [M] \mid [N]$ is a random function

SHA-3: Sponge Hash Functions

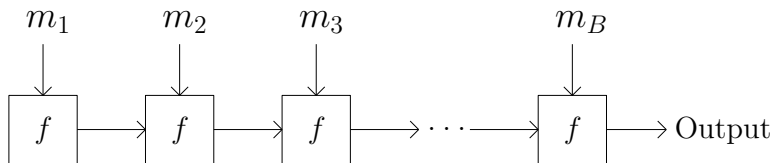
$f : [C] \rightarrow [R] \mid [R]$ is a random permutation

Can access both f and f^{-1}

Target: Hard to find collisions

SHA-3 is less secure?

Domain Extension Hash Functions



SHA-2: Merkle-Damgård Hash Functions

$f : [N] \rightarrow [M] \mid [N]$ is a random function

SHA-3: Sponge Hash Functions

$f : [C] \rightarrow [R] \mid [C]$ is a random permutation

Can access both f and f^{-1}

Target: Hard to find collisions

SHA-3 is less secure?

Auxiliary-Input (AI) Model

Intuition: Extra knowledge about f , e.g. backdoors?

Auxiliary-Input (AI) Model

Intuition: Extra knowledge about f , e.g. backdoors?

Setting:

- S -bit information about f
- T queries to f (or f^{-1} for sponge)
- Output collision

Auxiliary-Input (AI) Model

Intuition: Extra knowledge about f , e.g. backdoors?

Setting:

- S -bit information about f
- T queries to f (or f^{-1} for sponge)
- Output collision

Trivial birthday attack advantage: $T^2=N$ (MD) or $T^2=R$ (Sponge)

Auxiliary-Input (AI) Model

Intuition: Extra knowledge about f , e.g. backdoors?

Setting:

- S -bit information about f
- T queries to f (or f^{-1} for sponge)
- Output collision

Trivial birthday attack advantage: $T^2=N$ (MD) or $T^2=R$ (Sponge)

Exist **non-trivial attacks!**

Short Collision Finding in Merkle Damgård [CDGS18,ACDW20,GK22,AGL22]

Message Length	Best Known Attack
$B = 1$	$S=N + T^2=N$
$B = 2$	$ST=N + T^2=N$
3 $B \quad T$	$STB=N + T^2=N$
$B > T$	$ST^2=N$

Short Collision Finding in Sponge [FGK22]

Message Length	Known Attack (MD)	Known Attack (Sponge)
$B = 1$	$S=N + T^2=N$	$\min((ST=C)^2; (S^2T=C^2)^{\frac{2}{3}}) + S=C + T^2=R$
$B = 2$	$ST=N + T^2=N$	$ST=C + T^2=\min(C; R)$
$3 \leq B \leq T$	$STB=N + T^2=N$	$STB=C + T^2=\min(C; R)$
$B > T$	$ST^2=N$	$ST^2=C + T^2=R$

Better attacks than MD even when $B = 1$

Short Collision Finding in Sponge [FGK22]

Message Length	Known Attack (MD)	Known Attack (Sponge)
$B = 1$	$S=N + T^2=N$	$\min((ST=C)^2; (S^2T=C^2)^{\frac{2}{3}}) + S=C + T^2=R$
$B = 2$	$ST=N + T^2=N$	$ST=C + T^2=\min(C; R)$
$3 \leq B \leq T$	$STB=N + T^2=N$	$STB=C + T^2=\min(C; R)$
$B > T$	$ST^2=N$	$ST^2=C + T^2=R$

Better attacks than MD even when $B = 1$
Utilizes the inverse oracle

What about security upper bounds?

What about security upper bounds?

Old Techniques [DGK17,CDGS18]: presampling, compression

What about security upper bounds?

Old Techniques [DGK17,CDGS18]: presampling, compression

Multi-Instance Games (MI): A recent technique for proving security bounds for preprocessing attacks [IK10,CGLQ20,ACDW20,AGL22,FGK22]

Upper Bounds in Merkle Damgård

[ACDW20,GK22,AGL22]

Message Length	Best Known Attack	Upper Bound Tight?
$B = 1$	$S=N + T^2=N$	×
$B = 2$	$ST=N + T^2=N$	×
$3 \quad B \quad T$	$STB=N + T^2=N$	
$B > T$	$ST^2=N$	×

MI works pretty well here

Upper Bounds in Sponge [FGK22]

Message Length	Best Known Attack	Upper Bound Tight?
$B = 1$	$\min((ST=C)^2; (S^2T=C^2)^{\frac{2}{3}})$ $+S=C + T^2=R$	
$B = 2$	$ST=C + T^2=\min(C;R)$	
3 B T	$STB=C + T^2=\min(C;R)$	
$B > T$	$ST^2=C + T^2=R$	X

Upper Bounds in Sponge [FGK22]

Message Length	Best Known Attack	Upper Bound Tight?
$B = 1$	$\min((ST=C)^2; (S^2T=C^2)^{\frac{2}{3}})$ $+S=C + T^2=R$	
$B = 2$	$ST=C + T^2=\min(C;R)$	
$3 \leq B \leq T$	$STB=C + T^2=\min(C;R)$	
$B > T$	$ST^2=C + T^2=R$	X

What happens at sponge?

Upper Bound in Sponge [FGK22]

Message Length	Best Known Attack	Upper Bound
$B = 1$	$\min((ST=C)^2; (S^2T=C^2)^{\frac{2}{3}})$ $+S=C + T^2=R$	$ST=C + T^2=R$
$B = 2$	$ST=C + T^2=\min(C; R)$	$ST=C + S^2T^4=C^2$ $+T^2=\min(C; R)$
$3 \quad B \quad T$	$STB=C + T^2=\min(C; R)$	$ST^2=C + T^2=R$
$B > T$	$ST^2=C + T^2=R$	$ST^2=C + T^2=R$

What happens at sponge?

Upper Bound in Sponge [FGK22]

Message Length	Best Known Attack	Upper Bound
$B = 1$	$\min((ST=C)^2; (S^2T=C^2)^{\frac{2}{3}})$ $+S=C + T^2=R$	$ST=C + T^2=R$
$B = 2$	$ST=C + T^2=\min(C;R)$	$ST=C + S^2T^4=C^2$ $+T^2=\min(C;R)$
$3 \quad B \quad T$	$STB=C + T^2=\min(C;R)$	$ST^2=C + T^2=R$
$B > T$	$ST^2=C + T^2=R$	$ST^2=C + T^2=R$

What happens at sponge?

Can we prove better bounds (via MI)?

Our Results

Message Length	Best Known Attack	Upper Bound Tight?
$B = 1$	$\min((ST=C)^2; (S^2T=C^2)^{\frac{2}{3}})$ $+S=C + T^2=R$	Almost
$B = 2$	$ST=C + T^2=\min(C;R)$	
3 $B \quad T$	$STB=C + T^2=\min(C;R)$	
$B > T$	$ST^2=C + T^2=R$	X

Better bounds for $B = 1$, Simpler proofs for $B = 2$

Our Results

Message Length	Upper Bound Tight?	Better bounds for MI?
$B = 1$	Almost	
$B = 2$		
$3 \leq B \leq T$		
$B > T$	X	-

Limit of MI games

Advantages between MI and AI adversaries [AGL22]

We can reduce an AI adversary with success probability 2^{-s} to an MI adversary with probability $\tilde{O}(2^{-s})$.

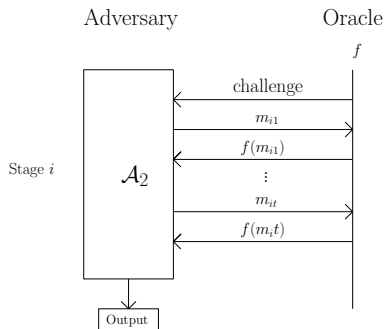
Advantages between MI and AI adversaries [AGL22]

We can reduce an AI adversary with success probability 2^{-S} to an MI adversary with probability $\tilde{O}(2^{-S})$.

Proof Idea. Guess the S -bit advice and run AI with that advice each round. The 2^S guessing probability will be amortized into 2^{-S} .

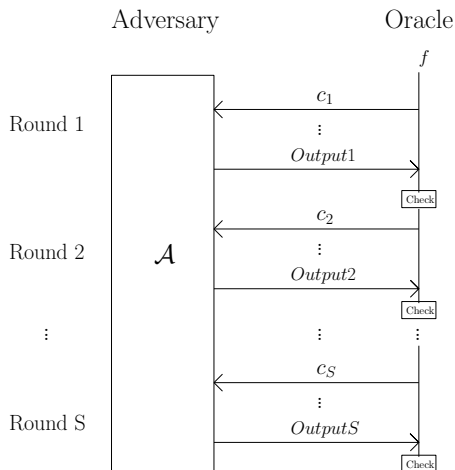
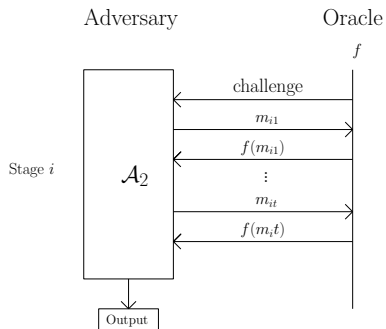
Multi-Instance Games

Repeat S times:

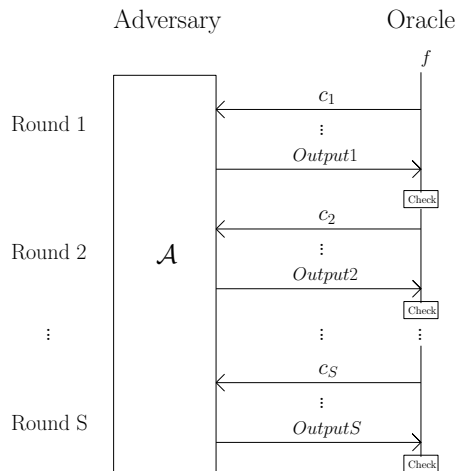


Multi-Instance Games

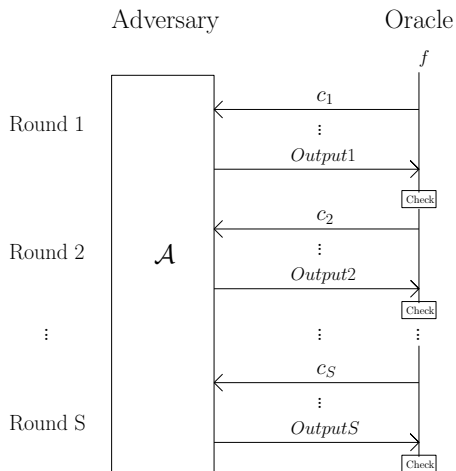
Repeat S times:



Multi-Instance Games



Multi-Instance Games



Something to mention:

- No advice string
- f doesn't change within rounds
- Has “memory” of previous rounds
- Need to win all S rounds

Multi-Instance Techniques

Main idea: By bounding the success probability of the MI game, we directly have upper bound for the original AI adversary.

Multi-Instance Techniques

Main idea: By bounding the success probability of the MI game, we directly have upper bound for the original AI adversary.

Advantages of MI game:

- No advice bits

Multi-Instance Techniques

Main idea: By bounding the success probability of the MI game, we directly have upper bound for the original AI adversary.

Advantages of MI game:

- No advice bits
- Ability to use lazy sampling and other techniques

Multi-Instance Techniques

Main idea: By bounding the success probability of the MI game, we directly have upper bound for the original AI adversary.

Advantages of MI game:

- No advice bits
- Ability to use lazy sampling and other techniques

Often easier to find upper bounds

Our Results

	Upper Bound	Known Attack
$B = 1$	$S^2 T^2 = C^2 + T^2 = R$ $+ S = C + T = C$	$\min((ST=C)^2; (S^2 T=C^2)^{\frac{2}{3}})$ $+ S = C + T^2 = R$
$B = 2$	$ST = C + S^2 T^4 = C^2$ $+ T^2 = \min(C; R)$	$ST = C + T^2 = \min(C; R)$
$B = 3$	$ST^2 = C + T^2 = R$	$STB = C + T^2 = \min(C; R)$

Our proof uses Multi-Instance Games technique
and highly non-trivial compression argument (please refer to original paper)

Showed limitations of MI Techniques:

	Upper Bound Given by MI	Best Attack in MI
$B = 1$	$(\tilde{O}(S^2 T^2 = C^2 + T^2 = R + S = C + T = C))^S$	$(\tilde{\Omega}(S^2 T^2 = C^2))^S$
$B = 2$	$(\tilde{O}(ST = C + S^2 T^4 = C^2 + T^2 = \min(C; R)))^S$	$(\tilde{\Omega}(S^2 T^4 = C^2))^S$
$B = 3$	$(\tilde{O}(ST^2 = C + T^2 = R))^S$	$(\tilde{\Omega}(ST^2 = C))^S$

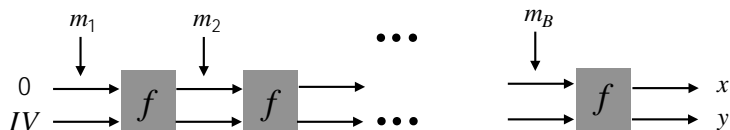
Showed limitations of MI Techniques:

	Upper Bound Given by MI	Best Attack in MI
$B = 1$	$(\tilde{O}(S^2 T^2 = C^2 + T^2 = R + S = C + T = C))^S$	$(\tilde{\Omega}(S^2 T^2 = C^2))^S$
$B = 2$	$(\tilde{O}(ST = C + S^2 T^4 = C^2 + T^2 = \min(C; R)))^S$	$(\tilde{\Omega}(S^2 T^4 = C^2))^S$
$B = 3$	$(\tilde{O}(ST^2 = C + T^2 = R))^S$	$(\tilde{\Omega}(ST^2 = C))^S$

It means we can't use MI to further bridge the gaps.

Sponge Hash Functions

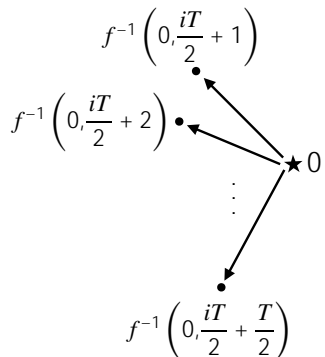
Input $m = m_1 || \dots || m_B, m_i \in [R]$



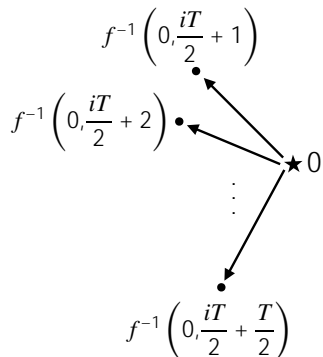
Sponge ^{f} (IV, m) := x

where $f : [R] \times [C] \rightarrow [R] \times [C]$ is a permutation

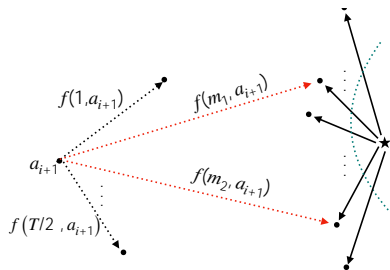
MI Attack, $B=2$



(1) Query $f^{-1}(0; i)$ for different i

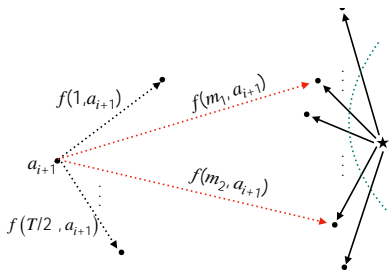


MI Attack, $B=2$



- (1) Query $f^{-1}(0; i)$ for different i
- (2) For challenge salt a , query $f(j; a)$ for different j

MI Attack, B=2



- (1) Query $f^{-1}(0; i)$ for different i
 - (2) For challenge salt a , query $f(j; a)$ for different j
- If two queries in (2) hits two salts visited in (1),

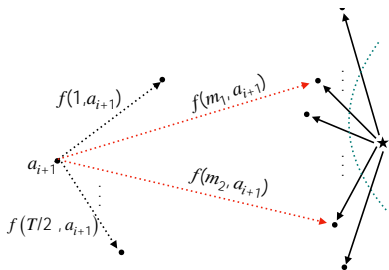
$$f^{-1}(0; i_1) = (m_1; a_1)$$

$$f(m_3; a) = (m_5; a_1)$$

$$f^{-1}(0; i_2) = (m_2; a_2)$$

$$f(m_4; a) = (m_6; a_2)$$

MI Attack, B=2



- (1) Query $f^{-1}(0; i)$ for different i
 - (2) For challenge salt a , query $f(j; a)$ for different j
- If two queries in (2) hits two salts visited in (1),

$$f^{-1}(0; i_1) = (m_1; a_1)$$

$$f(m_3; a) = (m_5; a_1)$$

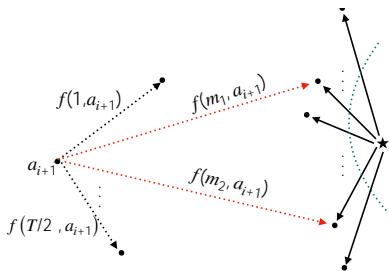
$$f^{-1}(0; i_2) = (m_2; a_2)$$

$$f(m_4; a) = (m_6; a_2)$$

then we have found valid collisions
on challenge salt a

$$(m_3/jm_5 \quad m_1); (m_4/jm_6 \quad m_2)$$

MI Attack, $B=2$



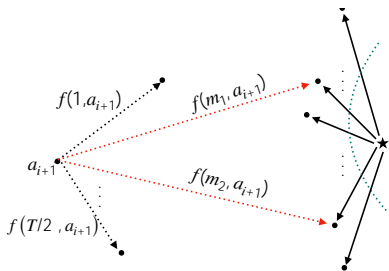
Each round: wins if we hit two old salts within $T=2$ queries
of different salts in (1):

$$\tilde{\Omega}(iT)$$

Winning Probability this round:

$$\tilde{O}((iT^2=C)^2)$$

MI Attack, B=2



Each round: wins if we hit two old salts within $T=2$ queries
of different salts in (1):

$$\tilde{\Omega}(iT)$$

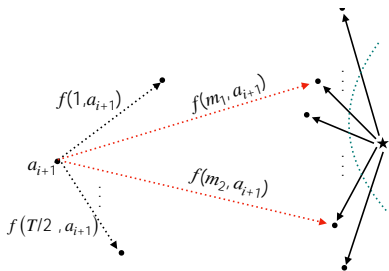
Winning Probability this round:

$$\tilde{O}((iT^2=C)^2)$$

Winning Probability for MI-game:

$$(\tilde{O}(S^2 T^4 = C^2))^S$$

MI Attack, B=2



Each round: wins if we hit two old salts within $T=2$ queries
of different salts in (1):

$$\tilde{\Omega}(iT)$$

Winning Probability this round:

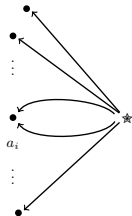
$$\tilde{O}((iT^2=C)^2)$$

Winning Probability for MI-game:

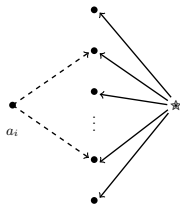
$$(\tilde{O}(S^2 T^4 = C^2))^S$$

Matches current upper bound
(proved by MI)

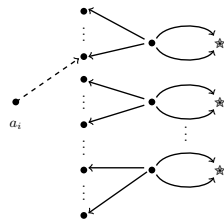
MI Attacks



(a) 1-Block collision attack

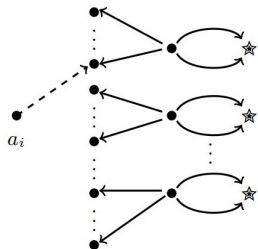


(b) 2-Block collision attack



(c) 3-Block collision attack

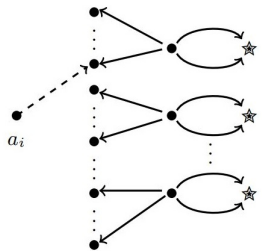
MI Attack, $B=3$



(c) 3-Block collision attack

- (1) Query out collisions at some salts a (via birthday attack)
- (2) Query $f^{-1}(\cdot; a)$ on these salts
- (3) Query $f(\cdot; a_i)$ for challenge salt a_i

MI Attack, $B=3$



(c) 3-Block collision attack

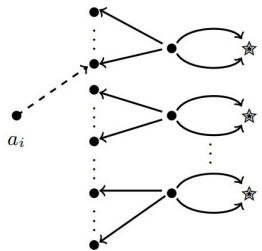
- (1) Query out collisions at some salts a (via birthday attack)
- (2) Query $f^{-1}(; a)$ on these salts
- (3) Query $f(; a_i)$ for challenge salt a_i

Wins if one query in (3) hits one salt in step (2)

Winning Probability for MI-game:

$$(\tilde{O}(ST^2=C))^S$$

MI Attack, $B=3$



(c) 3-Block collision attack

- (1) Query out collisions at some salts a (via birthday attack)
- (2) Query $f^{-1}(; a)$ on these salts
- (3) Query $f(; a_i)$ for challenge salt a_i

Wins if one query in (3) hits one salt in step (2)

Winning Probability for MI-game:

$$(\tilde{O}(ST^2=C))^S$$

Matches current upper bound (proved by MI)

Better bounds for $B = 1$, Showed limitation of MI

Message Length	Upper Bound Tight?	Better bounds for MI?
$B = 1$	Almost	
$B = 2$		
$3 \leq B \leq T$		
$B > T$	X	-

Open problems:

- Tight bounds (even for $B = 2$)?
- Better methods than MI?
- Better attacks?