Inside the hypercube

Jean-Philippe Aumasson   Eric Brier   Willi Meier
María Naya-Plasencia   Thomas Peyrin

Talk kindly given by Michael Gorski

CubeHash

D.J. Bernstein's SHA-3 candidate

"A simple hash function"

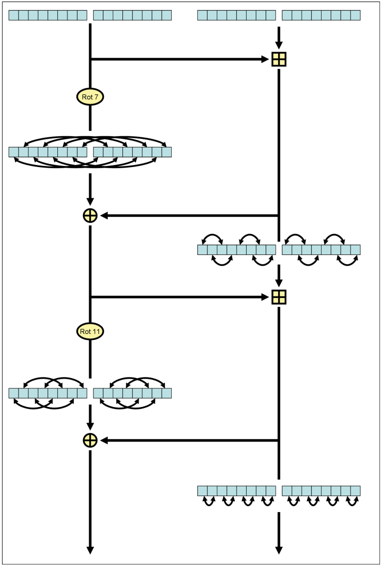
"ARX" algorithm ($+, \oplus, \ggg$)

Sponge-like construction:

- r-round permutation of a 1024-bit state, $r \in \{1, 2, \ldots\}$
- XOR b-byte message block, $b \in \{1, \ldots, 128\}$
- repeat for each block
- finalize the state: $10r$ rounds

Submission: $b = 1, r = 8$

CubeHash round



This talk

First third-party analysis of CubeHash

Improved generic attacks

Multicollisions strategy

State symmetries

Fixed point

Distinguisher

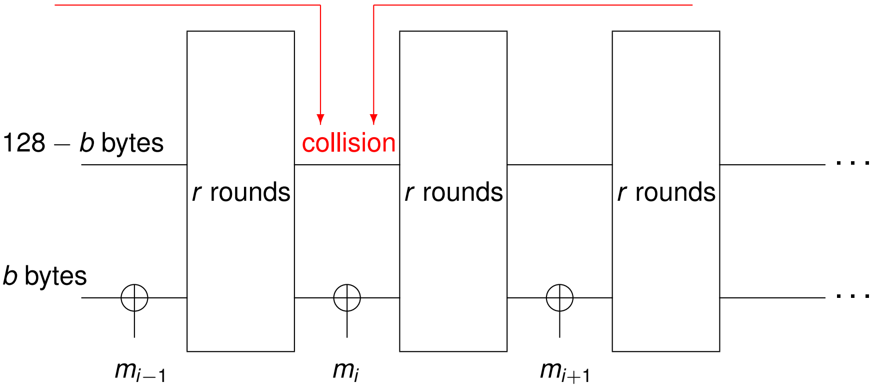
Previous (subsequent) works

Focus on collisions by linearization

- Aumasson $r = 2$, $b = 114$
- …
- Dai $r = 2$, $b = 3$
- Peyrin $r = 2$, $b = 12$
- Brier, Khazaei, Meier, Peyrin $r = 2$, $b = 2$
- Brier, Khazaei, Meier, Peyrin $r = 3$, $b = 64$
- …

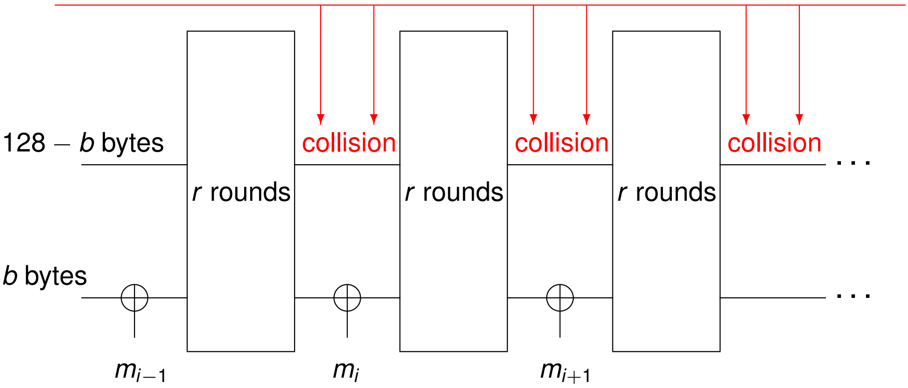
Generic preimage attacks

- meet-in-the-middle on $128 - b$ bytes
- fwrd/bwrd multiblock computation
- $\exp_2(522 - 4b - \log b)$ permutations



Improved generic preimage attacks

- multiple meet-in-the-middle on $128 - b$ bytes
- fwrd/bwrd multiblock computation
- $\exp_2(513 - 4b)$ permutations



Multicollisions

Key observations:

- the zero state is a fixed point for the permutation
- no counter and no padding of message

Technique for finding q-collisions:

1. meet-in-the-middle IV$\rightarrow X \leftarrow 0$
2. append zero message blocks

Costs 2^{513-4b} permutations (vs. Joux's $\log q \times 2^{512-4b}$)

State symmetries

Permutation T applied to 32 words $x[0], \ldots, x[31]$

If input of the form

```
AABBCCDD  EEFFGGHH  IIJJKKLL  MMNNOOPP
```

then output of the same form (with different values)

2^{512} states follow this pattern (out of 2^{1024} states)

Define a subgroup of 2^{512} elements

State symmetries

Can represent symmetry classes by a set of pairs (i, j), meaning $x[i] = x[j]$, for example

```
AABBCCDD  EEFFGGHH  IIJJKKLL  MMNNOOPP
```

represented by $(0, 1), (2, 3), \ldots, (28, 29), (30, 31)$

State symmetries

Can represent symmetry classes by a set of pairs (i, j), meaning $x[i] = x[j]$, for example

```
AABBCCDD  EEFFGGHH  IIJJKKLL  MMNNOOPP
```

represented by $(0, 1), (2, 3), \ldots, (28, 29), (30, 31)$

Use this representation for an exhaustive enumeration of all distinct (but non-disjoint) symmetry classes:

- 16 classes $C_0, \ldots, C_{15}$
- each class contains 2^{512} states
- 2^{516} symmetric states in total
- $T(C_i) = C_i$, $i = 0, \ldots, 15$

Exploiting symmetries

Key idea: classes sizes give upper bound on the size of the cycle of a symmetric state

$\Rightarrow$ search for near collisions within a permutation cycle

Preimage attack (for $b \geq 4$)

1. meet-in-a-same-class C_i
2. collision within C_i using symmetry-preserving blocks

For $b = 4$, meet in C_1: 2^{481} permutations
vs. 2^{493} with the improved generic attack

Works for any reasonable r

On fixed points

$1/k$ cycles of length k expected for a random permutation

$\Rightarrow$ one fixed point expected (length-1 cycle)

Same round permutation repeated r times, thus

- a r-cycle gives r fixed points
- cycles of length dividing r give more fixed points

Taking symmetry classes into account:

- 67 fixed points expected for one round
- 269 for 8 rounds of CubeHash

$\approx$ nonrandomness property

Distinguisher

Find a 3-round characteristic with weights $64 \rightarrow 1$

64 secret bits in $x[25]$ and $x[26]$

Weight-1 difference gives observable biases after 7 more rounds

$\Rightarrow$ truncated differential on 10 rounds

Not relevant to CubeHash hashing mode

Conclusion

CubeHash "broken", in the sense "less than 2^n permutations"...

Author considers "bit operations" (2^{11} per round)

Large parameters space, many safe choices

Which definition of nonrandomness is sufficient?

Inside the hypercube

Jean-Philippe Aumasson Eric Brier Willi Meier
María Naya-Plasencia Thomas Peyrin

Talk kindly given by Michael Gorski