Improved cryptanalysis of Skein

Jean-Philippe Aumasson	(FHNW, Switzerland)
Çağdaş Çalık	(METU, Turkey)
Willi Meier	(FHNW, Switzerland)
Onur Özen	(EPFL, Switzerland)
Raphael C.-W. Phan	(Loughborough Uni, UK)
Kerem Varıcı	(KU Leuven, Belgium)

Design by Ferguson, Lucks, Schneier, Whiting, Bellare, Kohno, Callas, Walker

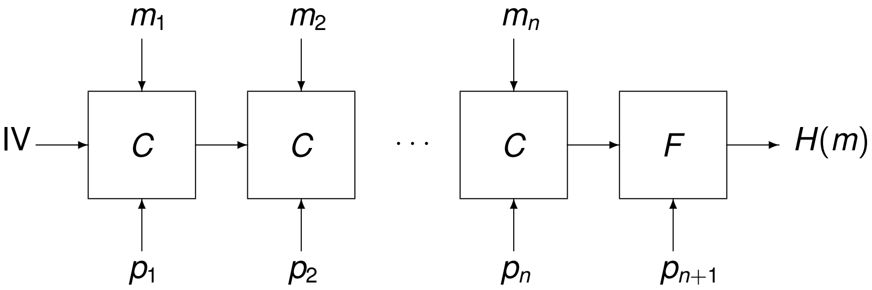
2nd round candidate in the SHA-3 competition

Iterated hash

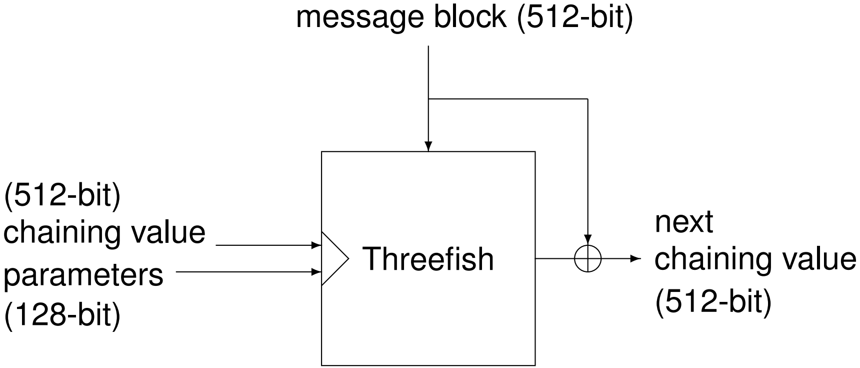
Compression function C, finalization function F

Process message $m = m_1 \| \ldots \| m_n$

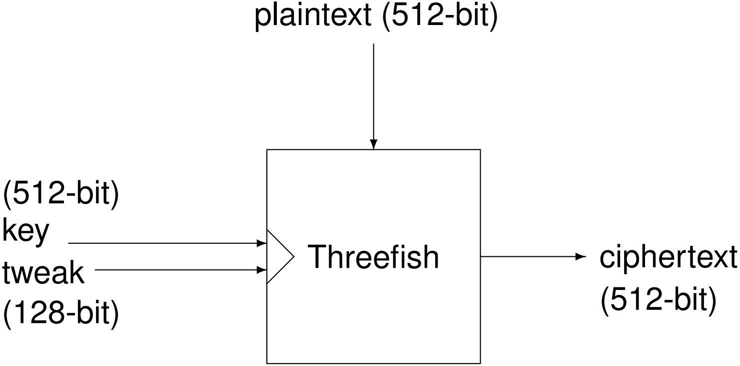
Parameters $p_1, \ldots, p_{n+1}$



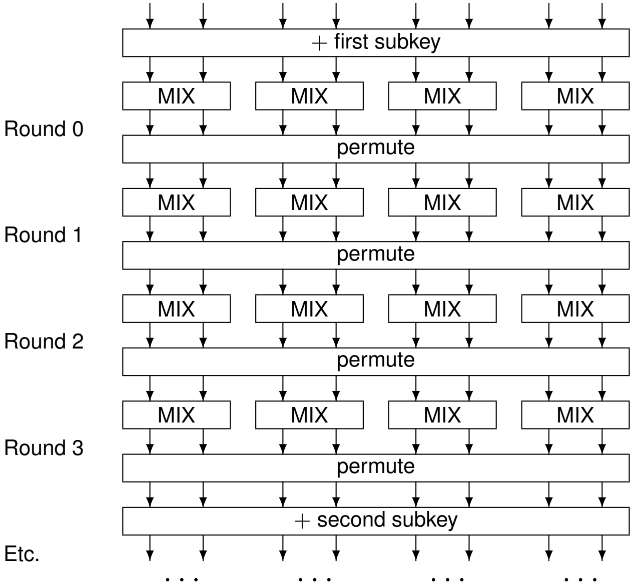
Block cipher-based compression function



The tweakable block cipher Threefish

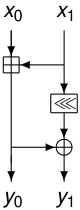


Substitution-permutation network with **72 rounds**
Subkeys words are XOR of key and tweak words



The MIX function

At round $r \in \{0, 1, \ldots, 71\}$ and position $s \in \{0, 1, 2, 3\}$:



$$
\begin{aligned}
\mathrm{MIX}_{r,s}(x_0, x_1) &= (y_0, y_1) \\
y_0 &= x_0 + x_1 \\
y_1 &= y_0 \oplus (x_1 \lll R_{r,s})
\end{aligned}
$$

$\Rightarrow$ Skein: Add-Xor-Rotate (AXR) algorithm

Basic properties

Full diffusion in 10 rounds

Simple and **linear key schedule**: subkeys $k_{s,0}, \ldots, k_{s,7}$ are derived from the key $k_0, \ldots, k_7$ and from the tweak t_0, t_1 as

$$
\begin{array}{lll|lll}
k_{s,0} & \leftarrow & k_{(s+0) \bmod 5} & k_{s,4} & \leftarrow & k_{(s+4) \bmod 5} \\
k_{s,1} & \leftarrow & k_{(s+1) \bmod 5} & k_{s,5} & \leftarrow & k_{(s+5) \bmod 5} + t_{s \bmod 3} \\
k_{s,2} & \leftarrow & k_{(s+2) \bmod 5} & k_{s,6} & \leftarrow & k_{(s+6) \bmod 5} + t_{(s+1) \bmod 3} \\
k_{s,3} & \leftarrow & k_{(s+3) \bmod 5} & k_{s,7} & \leftarrow & k_{(s+7) \bmod 5} + s
\end{array}
$$

Basic properties

Full diffusion in 10 rounds

Simple and **linear key schedule**: subkeys $k_{s,0}, \ldots, k_{s,7}$ are derived from the key $k_0, \ldots, k_7$ and from the tweak t_0, t_1 as

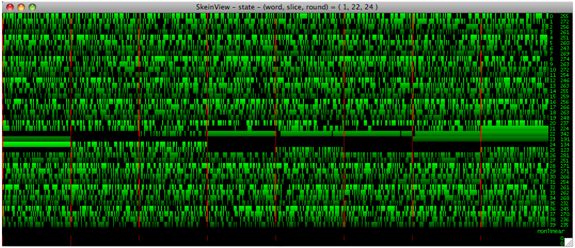
$$\begin{array}{lll|lll} k_{s,0} & \leftarrow & k_{(s+0) \bmod 5} & k_{s,4} & \leftarrow & k_{(s+4) \bmod 5} \\ k_{s,1} & \leftarrow & k_{(s+1) \bmod 5} & k_{s,5} & \leftarrow & k_{(s+5) \bmod 5} + t_{s \bmod 3} \\ k_{s,2} & \leftarrow & k_{(s+2) \bmod 5} & k_{s,6} & \leftarrow & k_{(s+6) \bmod 5} + t_{(s+1) \bmod 3} \\ k_{s,3} & \leftarrow & k_{(s+3) \bmod 5} & k_{s,7} & \leftarrow & k_{(s+7) \bmod 5} + s \end{array}$$

⇒ "**Subkey collisions**" easy to find, but. . .

- Impossible to find two consecutive collisions
- At least 7 subkeys between two collisions

Using differences in the plaintext, can delay full diffusion 8 rounds (then need 18 rounds for diffusion of differences)

SkeinView

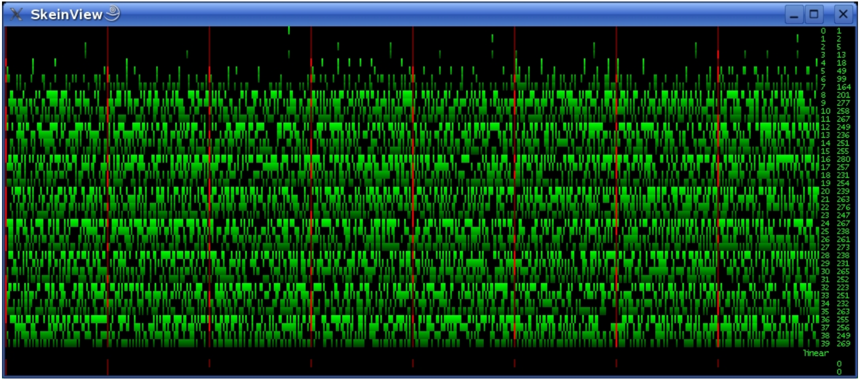


C++ program for Linux/Mac/Windows for studying Skein

- Visualization of differential trails
- Interactive choice of differences
- Differences in key, tweak, state
- Search for trails given conditions
- Normal and linearized modes
- LaTeX output of the trails

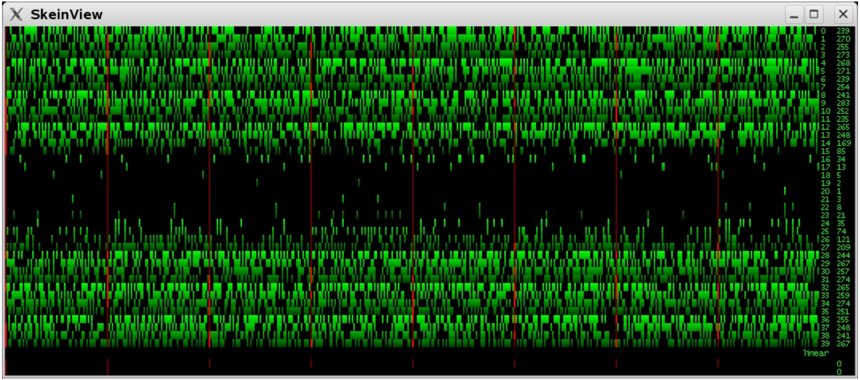
Threefish's diffusion of differences (1/2)

1-bit difference in the plaintext



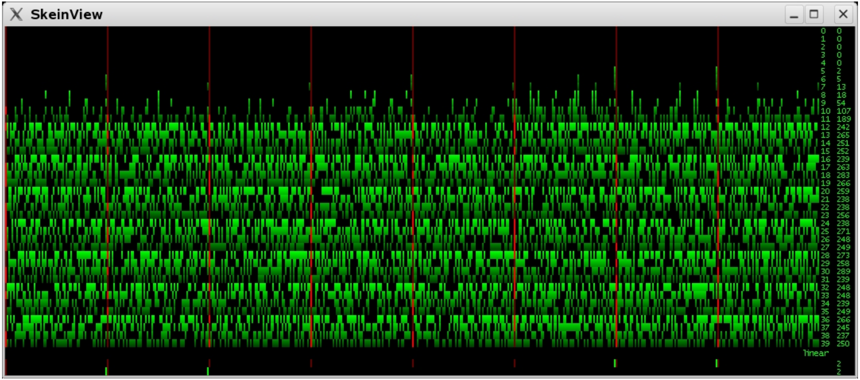
Threefish's diffusion of differences (2/2)

1-bit difference in the internal state



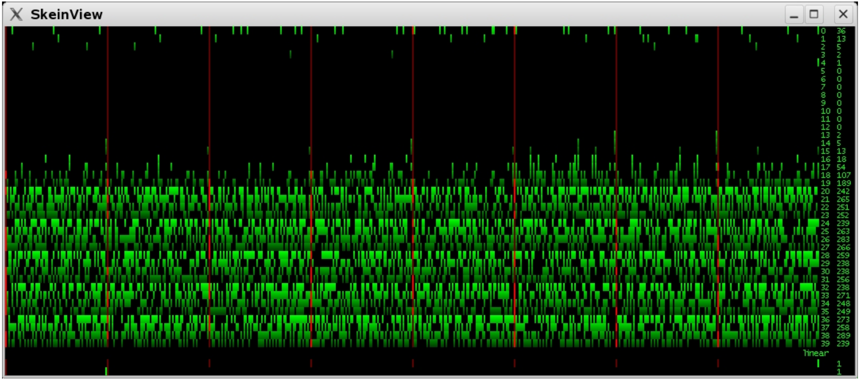
Subkey collision (initial subkey)

- No difference in the plaintext
- Difference introduced in the state at round 4
- Bias observable until round ≈ 13



Subkey collision (third subkey)

- Difference in the plaintext
- Difference introduced in the state at round 12
- Bias observable until round ≈ 21



Exploiting subkey collisions

- Distinguisher on 21 rounds with < 16 samples
- Near collisions on 17 rounds for the compression function in 2^{24}
- Impossible differentials…
- Boomerang attacks…

Finding impossible differentials

Miss in the middle

Proof by contradiction that $(\alpha \to \gamma)$ cannot occur

$$\alpha \xrightarrow{\textit{prob.}1} \beta \neq \delta \xleftarrow{\textit{prob.}1} \gamma$$

In practice, β and δ are differences over a subset of the internal state (that is, truncated differentials)

Impossible differentials were previously found for

- 8 rounds of AES-192 (of 12)
- 5 rounds of Twofish (of 16)

Miss in the middle of Threefish

We found probability-1 differentials:

- Forwards: on rounds **0 to 12**, over 92 output bits

  XXXXXXXXXXXXXX40 XXXXXXXXX2000000 XXXXXXXXXXXXX100 XXXXXXXXXXXXXX10
  XXXXXXXXXXXXX800 XXXXXXXXXXXXXXXX XXXXXXXXX2000000 XXXXXXXXXXXXXX40

- Backwards: on rounds **20 to 13**, over 134 output bits

  XXXXXXXXXXXX8000 XXXXXXXXXXXX8000 XXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXX
  XXXXXXXXXXXXXXXX XXXXXXXXXX400000 XXXXXXXXXX800000 XX50000000800000

$\Rightarrow$ Impossible differential for 21 rounds of Threefish

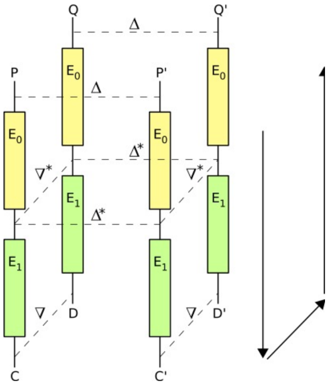
ブーメラン attack (outside the box)

1. Pick 2 plaintexts with difference α
2. **Encrypt** the 2 plaintexts
3. Set a difference β to each ciphertext
4. **Decrypt** the 2 new ciphertexts
5. Check that the new plaintexts have difference α

For well-chosen (α, β), step 5 succeeds with prob. $\gg 2^{-n}$

ブーメラン attack (inside view)

3-dimensional structure



Use of **2 differential characteristics**

1. For the first half of the cipher
2. For the inverse second half

Differential trails used for Threefish

Differences in the key and the tweak

Use locally optimal differentials

First half (rounds 1,...,16)

- Difference in the plaintext
- Probability 2^{-86}

Second half (rounds 34,...,17)

- Difference in the ciphertext
- Probability 2^{-113} from **round 34**

Each trail needs to be followed twice

$\Rightarrow$ Distinguisher with complexity $\approx 2^{2\times 86+2\times 113} = 2^{398}$

Variants

Key-recovery on 32 rounds

- Find inputs conforming to the boomerang relation
- Use them to determine half the whitening key
- 2^{312} decryptions, memory 2^{71} bytes

Variants

Key-recovery on 32 rounds

- Find inputs conforming to the boomerang relation
- Use them to determine half the whitening key
- 2^{312} decryptions, memory 2^{71} bytes

Known-key distinguisher on 35 rounds

- Key known but not chosen $\Rightarrow$ “white-box” attack
- Distinguisher: exhibition of inputs conforming to the boomerang relation
- Start with decryption instead of encryption
- complexity: 2^{478} trials

Summary

Rounds	Time	Memory	Type
16	2^{6}	–	459-bit near-collision
17	2^{24}	–	434-bit near-collision
21	$2^{3.4}$	–	related-key distinguisher
21	–	–	related-key impossible differential
25	$2^{416.6}$	–	related-key key recovery
26	$2^{507.8}$	–	related-key key recovery
32	2^{312}	2^{71}	related-key boomerang key recovery
34	2^{398}	–	related-key boomerang distinguisher
35	2^{478}	–	known-related-key boomerang distinguisher

Conclusion

At least 36 rounds needed for optimal security guarantees

The full Skein is not attacked (72 rounds)

Recent work by Chen and Jia: improved key-recovery using $+$-differences instead of $\oplus$-differences
See `http://eprint.iacr.org/2009/526`

Open issues:

- How to better exploit key collisions?
- Distinguishers using $+$-differences?
- Tweak…

Conclusion

NIST authorized “tweaks” for the second SHA-3 round

We have submitted a Tweak to the Skein algorithm. Specifically, we have changed – improved –the rotation constants. (Schneier)

Do the known attacks work on the new version of Skein?

Conclusion

NIST authorized "tweaks" for the second SHA-3 round

We have submitted a Tweak to the Skein algorithm. Specifically, we have changed – improved –the rotation constants. (Schneier)

Do the known attacks work on the new version of Skein?

From the revised Skein documentation:

We are confident that one can easily adopt (sic) *the attacks to Threefish-512 and its new rotation constants, mainly by finding new differential trails and performing new frequency tests.*

The 32-round attack by Chen and Jia is even **faster** with the new constants

Improved cryptanalysis of Skein

Jean-Philippe Aumasson	(FHNW, Switzerland)
Çağdaş Çalık	(METU, Turkey)
Willi Meier	(FHNW, Switzerland)
Onur Özen	(EPFL, Switzerland)
Raphael C.-W. Phan	(Loughborough Uni, UK)
Kerem Varıcı	(KU Leuven, Belgium)