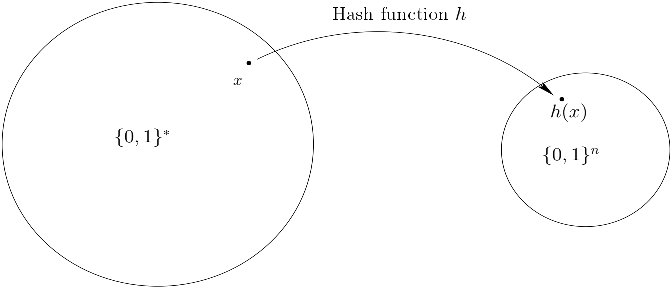
Deterministic Constructions of 21-Step Collisions for the SHA-2 Hash Family

Somitra Kr. Sanadhya and Palash Sarkar

Cryptology Research Group
Applied Statistics Unit
Indian Statistical Institute, Kolkata
India
{somitra_r, palash}@isical.ac.in

ISC, Taipei, 17th September 2008

Cryptographic Hash functions



- Fixed size "Fingerprint" of arbitrary length data.

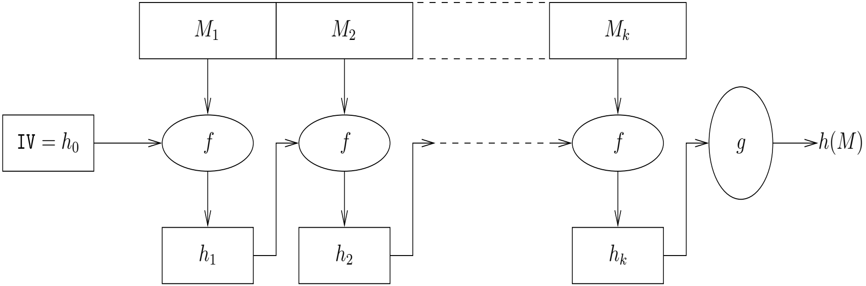
Cryptographic Hash functions

- Used in :
  - Verifying integrity of data.
  - Digital signatures.
  - Storing authentication information (Passwords).
  - . . .

Hash function Security Requirements

- Collision resistance.
  Difficult to find x_1 and x_2 s.t. $x_1 \neq x_2$ but $h(x_1) = h(x_2)$
- Preimage resistance.
  Given y, it is difficult to find an x s.t. $h(x) = y$
- Second Preimage resistance.
  Given x_1, it is difficult to find an x_2 s.t. $x_1 \neq x_2$ but $h(x_1) = h(x_2)$

Merkle-Damgard Hash Design



Hash Function Schema (for one block message)

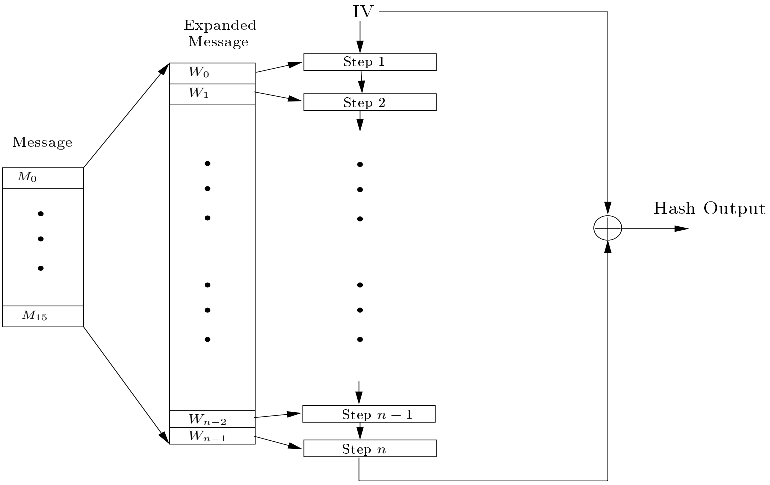
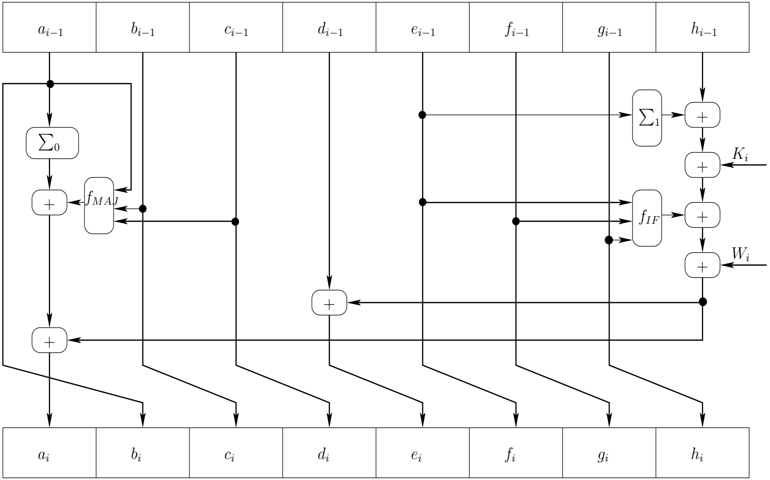


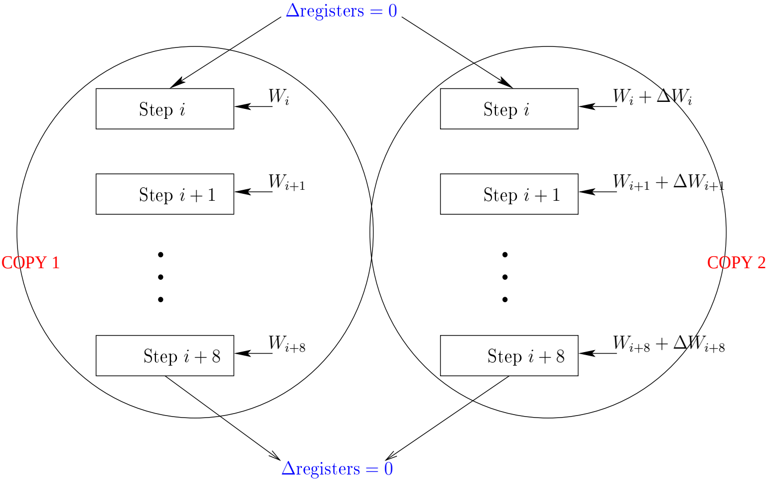
Figure: Round function of SHA-2 family



The Notation

- Message words :
  - I : $\{W_0, W_1, \ldots, W_{15}\}$,
  - II: $\{W'_0, W'_1, \ldots, W'_{15}\}$.
  - These message words are then expanded upto W_{20} and W'_{20} for this work. The word W_i is used in Step i, where the index i starts from zero.
- Differences $\{\delta W_0, \delta W_1, \ldots, \delta W_{15}\}$.
- $W'_i = W_i + \delta W_i$.

The Local Collision



9-step Non-Linear Local Collision

Step i	δW_i	δa_i	δb_i	δc_i	δd_i	δe_i	δf_i	δg_i	δh_i
$i-1$	0	0	0	0	0	0	0	0	0
i	x	x	0	0	0	x	0	0	0
$i+1$	δW_{i+1}	0	x	0	0	y	x	0	0
$i+2$	δW_{i+2}	0	0	x	0	z	y	x	0
$i+3$	δW_{i+3}	0	0	0	x	0	z	y	x
$i+4$	δW_{i+4}	0	0	0	0	x	0	z	y
$i+5$	δW_{i+5}	0	0	0	0	0	x	0	z
$i+6$	δW_{i+6}	0	0	0	0	0	0	x	0
$i+7$	δW_{i+7}	0	0	0	0	0	0	0	x
$i+8$	$-x$	0	0	0	0	0	0	0	0

1. Nikolić-Biryukov (NB) : FSE '08 = $\{x, y, z\}$=$\{1, -1, 0\}$;
2. Sanadhya-Sarkar (SS) : ACISP '08 = $\{x, y, z\}$=$\{1, -1, -1\}$

The Cross Dependence Equation

- We note a special and simple relation in the *a* and the *e* register.
- For example,

$$\begin{aligned} e_i &= d_{i-1} + \Sigma_1(e_{i-1}) + f_{IF}(e_{i-1}, f_{i-1}, g_{i-1}) + h_{i-1} + K_i + W_i \\ &= d_{i-1} + a_i - \Sigma_0(a_{i-1}) - f_{MAJ}(a_{i-1}, b_{i-1}, c_{i-1}) \\ &= a_{i-4} + a_i - \Sigma_0(a_{i-1}) - f_{MAJ}(a_{i-1}, a_{i-2}, a_{i-3}). \end{aligned} \quad (1)$$

  This relationship shows that the *e* register solely depends on the *a* register values of previous 5 steps.
- This relation also shows that the state update of the SHA-2 family can be written in terms of one variable only, as was also independently observed by Indesteege et al. (SAC '08).

Constructing the 21-Step SHA-2 Attack

- Have a single local collision spanning from Step 6 to Step 14.
- We take other message words to have no differences. That is $\delta W_i = 0$ for $i \in \{0, 1, 2, 3, 4, 5, 15\}$.
- For the SS local collision, we have $\delta W_i = 0$ for $i \in \{10, 11, 12, 13\}$.
- First 5 steps of message expansion of SHA-2 are shown next.

$$\left.\begin{array}{rcl}
W_{16} & = & \underline{\sigma_1(W_{14}) + \ W_9} + \sigma_0(W_1) + W_0, \\
W_{17} & = & \sigma_1(W_{15}) + W_{10} + \sigma_0(W_2) + W_1, \\
W_{18} & = & \underline{\sigma_1(W_{16})} + W_{11} + \sigma_0(W_3) + W_2, \\
W_{19} & = & \sigma_1(W_{17}) + W_{12} + \sigma_0(W_4) + W_3, \\
W_{20} & = & \underline{\sigma_1(W_{18})} + W_{13} + \sigma_0(W_5) + W_4.
\end{array}\right\}$$

Underlined terms may have non-zero differences.

If $W_{16} = W'_{16}$ i.e. $\delta W_{16} = 0$ then we have a 21-step collision.

Constructing the 21-Step SHA-2 Attack

- $W_{16} = \sigma_1(W_{14}) + W_9 + \sigma_0(W_1) + W_0$.
- $\delta W_{14} = -1$.
- I.e. $\delta W_{16} = 0, \implies \sigma_1(W_{14}) + W_9 = \sigma_1(W'_{14}) + W'_9$.
- I.e. $\delta W_9 = \sigma_1(W_{14}) - \sigma_1(W_{14} - 1)$.
- In ACISP '08, we developed an improved probabilistic attack using the fact that the term $\sigma_1(X) - \sigma_1(X - 1)$ is highly skewed.
- We created a list of pairs $(X, \sigma_1(X) - \sigma_1(X - 1))$.
- Now we can make the attack deterministic.

Constructing the 21-Step SHA-2 Attack

- The SS local collision allows δW_9 to be set to any value.
- Even though $\sigma_1(W_{14}) - \sigma_1(W_{14} - 1)$ is highly skewed, we can suitably choose δW_9 so as to satisfy the equality of these two terms.
- This allows the deterministic 21-step SHA-2 attack.
- Prior work had not been able to show 21-step SHA-512 collisions. We provide the first colliding message pair for 21-step SHA-512.

Constructing the 21-Step SHA-2 Attack

- There are two different 21-step SHA-2 attacks in this work.
- Both the attacks are deterministic.
- There are 6 free words in the first attack.
- There are 5 free words in the second attack.
- In the first attack, the SS local collision has 4 consecutive $\delta W_i = 0$.
- In the first attack, the SS local collision has 3 consecutive $\delta W_i = 0$.

The Case of the NB Local Collision

- δW_9 depends on e_6, e_7 and e_8.
- Assuming that these three register values are random, we have that

$$Pr[\delta W_9 \geq 2^j] < \frac{1}{2^{j-1}}.$$

- I.e. δW_9 rarely takes very large values.
- On the other hand, for SHA-512, we have that

$$\sigma_1(X) - \sigma_1(X-1) \geq (2^{42} + 2^{39} + 2^{38} + 2^{36} - 2^3).$$

- The two lemmas above show that, using the NB local collision, obtaining equality of the two terms is very unlikely for SHA-512.

The Case of the NB Local Collision

- More generally, if the NB local collision is started at Step i, then
- δW_{i+3} depends on e_i, e_{i+1} and e_{i+2}.
- Assuming that these three register values are random, we have that

$$Pr[\delta W_{i+3} \geq 2^j] < \frac{1}{2^{j-1}}.$$

- But we still need to satisfy the equality

$$\delta W_{i+3} = \sigma_1(W_{i+8}) - \sigma_1(W_{i+8} - 1).$$

Bridging the Gap for the NB Local Collision

- First of all, note that

$$\delta W_{i+3} = -f_{IF}(e_{i+2}, e_{i+1} - 1, e_i) + f_{IF}(e_{i+2}, e_{i+1}, e_i).$$

- It is possible to ensure that the values of the registers e_{i+2}, e_{i+1} and e_i are such that the term δW_{i+3} is large. This allows attaining the equality possible.
- But, to attain such values of e_{i+2}, e_{i+1} and e_i, one needs to iterate over many initial words.
- $\implies$ Equality is achieved at the cost of more work in the beginning of the search for message words.

Recent work on SHA-2

- 22-step Deterministic SHA-512 Attack. (IACR eprint)
- 23-step SHA-512 attack with effort $2^{16.5}$ calls. (CoRR archive)
- 24-step SHA-512 attack with effort $2^{32.5}$ calls. (CoRR archive)

Thank You

The Organizers & The Audience.