MD4-family, SHA-1, SHA-2, SHA-3

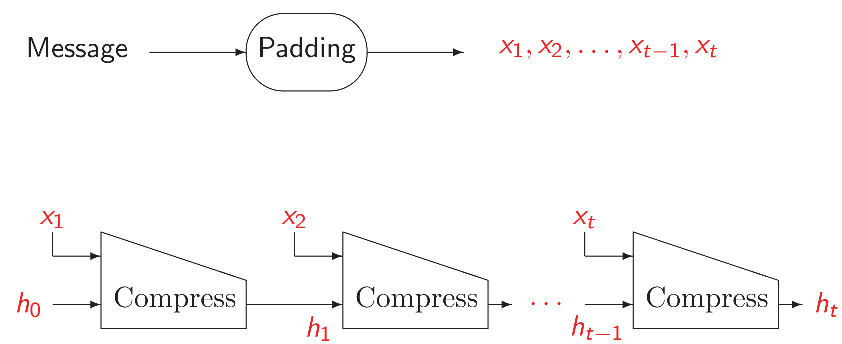
Lars R. Knudsen

April 21, 2008

Definition - hash function

$H : \{0,1\}^* \to \{0,1\}^n$, for fixed value of n

Iterated hash functions



Generic attacks

For $H : \{0,1\}^* \rightarrow \{0,1\}^n$ and $h : \{0,1\}^m \rightarrow \{0,1\}^n$, $m > n$

attack	rough complexities
collisions	$\sqrt{2^n} = 2^{n/2}$
2nd preimages	$k2^{n/2} + 2^{n-k}$ with 2^k blocks
preimage	2^n

Goal: generic attacks are best (known) attacks

MD4-family

- MD4, Rivest 1990
- MD5, Rivest 1991
- SHA-0, 1993
- SHA-1, 1995
- SHA-256, 2002
- SHA-512, 2002
- all hash functions of Davies-Meyer form
- "block ciphers" with feed-forward

MD4-family - RIPEMDs

- RIPEMDs, MD4-variants by Bosselaers, Dobbertin, Preneel
- 1992: RIPEMD (128 bits)
- 1995: RIPEMD-128, RIPEMD-160

Hash functions in real-life

Scheme	Bits in hash code	Compression fct. bits in $\rightarrow$ bits out	Designer	Year
MD4	128	$512 + 128 \rightarrow 128$	Rivest	1990
MD5	128	$512 + 128 \rightarrow 128$	Rivest	1991
RIPEMD-128	128	$512 + 128 \rightarrow 128$	BDP	1992
SHA-1	160	$512 + 160 \rightarrow 160$	US Gov.	1995
RIPEMD-160	160	$512 + 160 \rightarrow 160$	BDP	1995
SHA-256	256	$512 + 256 \rightarrow 256$	US Gov.	2002
SHA-512	512	$1024 + 512 \rightarrow 512$	US Gov.	2002

MD: Message Digest
SHA: Secure Hash Algorithm

From MD4 over MD5 to SHA

- Iterated hash functions
- Compression functions process message blocks of 512 bits
- Message blocks processed in words of 32 bits
- Message expanded from 512 to $32 \times r$ bits, where r is number of steps of algorithm
-

Algorithm	Steps	# registers of 32 bits
MD4	48	4
MD5	64	4
SHA-1	80	5

Secure Hash Standard (SHS)

also known as SHA-1 (Secure Hash Algorithm 1)

notation:

$+$	addition modulo 2^{32} of 32 bit words
$\mathrm{rot}_i(x)$	rotate word x (32 bits) left by i positions
$\oplus$	bitwise exclusive-or
$\&$	bitwise and
$\vert$	bitwise or

initialize variables A, B, C, D, E, each of 32 bits, by
$(A, B, C, D, E) =$
$(\mathtt{67452301}_x, \mathtt{EFCDAB89}_x, \mathtt{98BADCFE}_x, \mathtt{10325476}_x, \mathtt{C3D2E1F0}_x)$

Hashing with SHA-1

1. pad message, s.t. last block is 512-64 bits
2. append 64-bit block containing length of original message
3. for each message block $\tilde{M}$ of 512 bits:
   1. compute $t = h(\tilde{M}, (A, B, C, D, E)) = t_0, \ldots, t_4$
   2. set $A = t_0, B = t_1, C = t_2, D = t_3$, and $E = t_4$
4. output the hash value $[t_0 \parallel t_1 \parallel t_2 \parallel t_3 \parallel t_4]$.

for compression function $h : \{0,1\}^{512} \times \{0,1\}^{160} \to \{0,1\}^{160}$

Compression fct: $\{0,1\}^{512} \times \{0,1\}^{160} \to \{0,1\}^{160}$

- 80 basic steps in compression function
- Message $W = [W_0 \parallel W_1 \parallel \ldots \parallel W_{15}]$, where W_i are 32-bit words.
- Expansion:
  $W_i = \mathrm{rot}_1(W_{i-3} \oplus W_{i-8} \oplus W_{i-14} \oplus W_{i-16}), \quad 16 \le i \le 79$
- constants K^i are defined

$$\begin{aligned} K^i &= \mathtt{5A827999}_x, & 0 \le i \le 19 \\ K^i &= \mathtt{6ED9EBA1}_x, & 20 \le i \le 39 \\ K^i &= \mathtt{8F1BBCDC}_x, & 40 \le i \le 59 \\ K^i &= \mathtt{CA62C1D6}_x, & 60 \le i \le 79 \end{aligned}$$

Compression function (continued)

$$A^0 = A, B^0 = B, C^0 = C, D^0 = D, E^0 = E$$

$$\begin{aligned} A^{i+1} &= W_i + \mathrm{rot}_5(A^i) + f^i(B^i, C^i, D^i) + E^i + K^i \\ B^{i+1} &= A^i \\ C^{i+1} &= \mathrm{rot}_{30}(B^i) \\ D^{i+1} &= C^i \\ E^{i+1} &= D^i \qquad \text{for } i = 0\dots,79, \text{ where} \end{aligned}$$

$$\begin{aligned} f^i &= f_{if} &= (X\&Y)|(\neg X\&Z), \qquad & 0 \le i \le 19 \\ f^i &= f_{xor} &= X \oplus Y \oplus Z, \qquad & 20 \le i \le 39, 60 \le i \le 79 \\ f^i &= f_{maj} &= (X\&Y)|(X\&Z)|(Y\&Z), \qquad & 40 \le i \le 59. \end{aligned}$$

$$A = A + A^{80}, B = B + B^{80}, C = C + C^{80}, D = D + D^{80}, E = E + E^{80}$$

Cryptanalysis - highlights

- 1996: MD4 broken, Dobbertin
- 2004: MD5 broken, Wang
- 2004: SHA-0 broken, Joux et al
- 2004: RIPEMD broken, Wang
- 2005, claim: collisions for SHA-1 in time $\approx 2^{69}$ (Wang)
- 2006, claim: collisions for SHA-1 in time $\approx 2^{63}$ (Wang)
- 2007, claim: collisions for SHA-1 in time $\approx 2^{60}$ (Mendel, Rechberger, Rijmen)
- 2007: `http://boinc.iaik.tugraz.at/`

SHA-3

NIST

SHA-3 - Call for candidates

- announcement: October 29, 2007
- must provide digests of 224, 256, 384, and 512 bits, not 160.
- available worldwide royalty-free, no IPR
- capable of protecting sensitive information for decades
- should be suitable for
  - digital signatures, FIPS 186-2
  - HMAC, FIPS 198
  - key establishment, SP 800-56A
  - random number generation, SP 800-90
- security strength at least that of the SHA-2s with **significantly** improved efficiency

SHA-3 - Desirable properties

- efficient integral options, e.g., randomized hashing, that "fundamentally improve security"
- parallelizable
- avoid "generic properties" of Damgård/Merkle constructions
- attack on SHA-2 should not lead to attack on SHA-3
- flexible for a wide variety of implementations
- a single family, except if good arguments for more families
- tunable security parameter, e.g., number of rounds, with recommendations

SHA-3 - Timeline

- hard submission deadline: 31/10-2008
- submissions by 31/8-2008 checked by NIST for inconsistencies
- Round 1: 12 months. Workshop 1. Workshop 2. No modifications during Round 1.
- Round 2: ≈ 5 candidates selected. 12-15 months. Tweaks allowed. Workshop 3.
- AHS(s).
- documentation and testing like AES
- review is public

SHA-3 - Security

Message digest of n bits

- Collisions should take $2^{n/2}$
- Preimages should take 2^n
- 2nd preimages should take 2^{n-k} for messages shorter than 2^k bits

Higher levels of security against 2nd preimage will be viewed positively

- NIST open to other designs than Damgård/Merkle