



Hash functions

From CRC to SHA-2

Martin Šácha

2. 4. 2012



First look

Properties:

- Low cost
- Determinism
- Uniformity
- One-way - difficult to find data for given hash
- Collision resistance - difficult to find $x \neq y; \varphi(x) = \varphi(y)$
- Infeasible to modify data without changing hash



Merkle–Damgård construction

Collision-resistant compression \rightarrow hash function

Add true-bit at the end of message

Add 0 – 447 zero-bits to round msg. length to $k512 - 64b$

Add length of message in 64-bit number (round $(k + 1)512b$)



HMAC

Verify data integrity and authenticity

HMAC-<hash function> (ie HMAC-SHA1)

$$HMAC(K, m) = H((K \oplus opad) || H((K \oplus ipad) || m))$$



Birthday attack

$$n(p, H) \approx \sqrt{2H \ln \frac{1}{1-p}}$$

$$Q(H) \approx \sqrt{\frac{\pi}{2} H}$$

p - probability of collision

H - hash length in bits



CRC - Cyclic redundancy check

$$\begin{array}{r|l} M(x) = & 100101_2 = x^5 + x^2 + 1 \\ G(x) = & \text{XOR } 110011_2 = x^5 + x^4 + x + 1 \\ \hline R(x) = & 010110_2 = x^4 + x^2 + x \end{array}$$



Whirlpool and Tiger

	<i>Whirlpool</i>	<i>Tiger</i>
<i>Designed</i>	2000	1996
<i>Hash size</i>	512b	192b
– <i>string</i>	128B	48B
<i>Attacks</i>	<i>Coll. in 2^{120}</i>	<i>not known</i>



MD5 - Message digest version 5

Hash length = 128b (32B in human readable string)

- 1991, Donald Rivest, MIT
- 1993, pseudo-collision
- 1996, first particular collision
- 2004, birthday attack
- 2004, full collision
- 2005, two X.509 certificates with the same hash
- 2005, Vlastimil Klíma, alg. to find collision in one minute
- 2010, single-block collision, secret algorithm

Vlastimil Klíma - most efficient MD5 collision algorithm



SHA1 - Secure hash algorithm version 1

Hash length = 160b (40B in human readable string)

- 1995, NSA
- 2005, collision in 2^{69} operations (strong is 2^{80})
- 2007, collision in 2^{63} operations
- 2007, began distributed collisions (BOINC)
- 2008, Stéphane Manuel - theoretic collision in $2^{51} - 2^{57}$
- 2010, full attack in $2^{57.5}$ operations



SHA2 - Secure hash algorithm version 2

Family of algorithms (SHA-224, SHA-256, SHA-384, SHA-512)

No collisions found

	<i>SHA – 224/256</i>	<i>SHA – 384/512</i>
<i>Hash size</i>	224/256b	384/512b
<i>– string</i>	56/64B	96/128B
<i>Max input</i>	$2^{64} - 1b$	$2^{128} - 1b$



How can we find collision?

$$\varphi(x) = ed076287532e86365e841e92bfc50d8c$$

- Google it
- Rainbow tables
- Social engineering
- Bruteforce algorithms:
 - Dark room
 - Birthday attack
 - Classic bruteforce attack



And how to make hash stronger?

- Combine functions from different families
- Salt data
- Truncate hash to lower length (carefully)

MD5 ← *SHA1* ← *Tiger* ← *SHA2* ← *Whirlpool*

32 ← 40 ← 48 ← 56/64/96/128 ← 128

- Freely combine methods above



Further reading

`fjfi.lainspira.net/ukry/hash_functions.pdf`

- `cryptography.hyperlink.cz/`
- `en.wikipedia.org/wiki/MD5`
- `en.wikipedia.org/wiki/SHA-1`
- `en.wikipedia.org/wiki/SHA-2`
- `en.wikipedia.org/wiki/Whirlpool_%28cryptography%29`
- `en.wikipedia.org/wiki/Tiger_%28cryptography%29`
- `en.wikipedia.org/wiki/Crypt_%28Unix%29#Library_Function`