Hash functions

CRYPTOGRAPHY AND COMPUTER SECURITY

Ana I. González-Tablas Ferreres

José María de Fuentes García-Romero de Tejada

Lorena González Manzano

Sergio Pastrana Portillo

uc3m | Universidad Carlos III de Madrid

COSEC

OUTLINE

- 9. Hash functions
  - Hash functions
  - Cryptographic hash functions
  - Examples

OUTLINE

- 9. Hash functions
  - Hash functions
  - Cryptographic hash functions
  - Examples

HASH FUNCTIONS

- Takes as input a variable-length block of data (M) and produces a fixed-size hash value

hash =H(M)

- Main goal: data integrity

HASH FUNCTIONS

- There are infinite possible input messages (variable size)
- Collision
- Each hash function has a Hash Space of size |h|

$$|h| = 2^n$$

- with n being the hash function output length (in bits)
- It is possible to find two messages M and M' such that:

$$H(M) = H(M') \rightarrow \text{Collision}$$

OUTLINE

- 9. Hash functions
  - Hash functions
  - Cryptographic hash functions
  - Examples

CRYPTOGRAPHIC HASH FUNCTIONS

- Hash function with extra requirements:
  - Must work with messages of any size
  - Must compute fixed-size hash values
  - The output must satisfy pseudo-randomness requirements

Compression

CRYPTOGRAPHIC HASH FUNCTIONS

- **Diffusion**: if a single bit of the message M is changed, then H(M) must change approximately half of its bits
- **Determinism**: for a given input, multiple runs of the function must always generate the same hash value
- **Efficiency**: fast calculation of the hash value in both software and hardware implementations

CRYPTOGRAPHIC HASH FUNCTIONS

- **One-way property**: for any given value h, it is computationally unfeasible to find an M’ such that:

$$H(M') = h$$

- **Weak collision resistant**: for any given message M, it is computationally unfeasible to find a message *M’□M* such that

$$H(M) = H(M')$$

- **Strong collision resistant**: It is computationally unfeasible to find two messages M y M’ such that:

$$H(M) = H(M')$$

CRYPTOGRAPHIC HASH FUNCTIONS

- "Computationally unfeasible"
  - There is no algorithm more efficient than brute force for producing collisions
  - If hash space is large enough, the probability of finding a collision is null in a reasonable time (in HW or SW)
- Hash function strength depends on:
  - Design: only brute force attack available
    - not cryptanalyzable
  - Hash length (n) should be large enough

CRYPTOGRAPHIC HASH FUNCTIONS

- Probabilidades de encontrar una colisión (fuerza bruta)
  - One-way property attack: $\frac{1}{2^n}$
  - Weak collision attack: $\frac{1}{2^n}$
  - Strong collision attack: $\frac{1}{2^{n/2}}$ (p≥50%) (birthday paradox)

CRYPTOGRAPHIC HASH FUNCTIONS

- A hash function is said to be “broken” if there is no technique for producing collisions in less than brute force time

- 2^{80} is the minimum accepted barrier for algorithmic complexity

- MD5 is broken (produces hashes of 128 bits)

CRYPTOGRAPHIC HASH FUNCTIONS

- POSSIBLE ATTACK:
- One-way attacks
  - Impersonation at the password-hashes storage systems
  - Forcing false positives in hashing tables
- Weak collision attacks
  - Faking public key certificates,digitally signed documents,source code,etc.
- Strong collision attacks
  - Birthday attack to fake digitally signed documents

CRYPTOGRAPHIC HASH FUNCTIONS

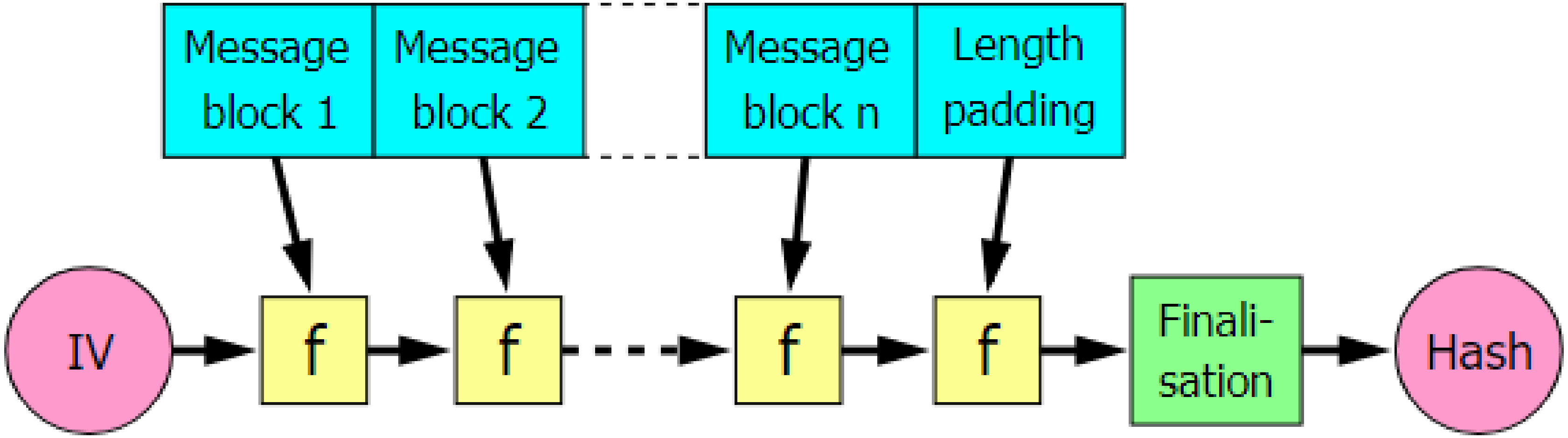
- APPLICATIONS
  - Integrity verification
  - Digital signatures
  - Use in MAC (Message Authentication Code) functions
  - Database index
  - Passwords storage
  - Intrusion detection
  - Pseudorandom number generators
  - Etc.

OUTLINE

- 9. Hash functions
  - Hash functions
  - Cryptographic hash functions
  - Examples

EXAMPLES: MERKLE-DAMGÄRD CONSTRUCTION

- The most commonly used in modern hash functions



CV_0 = IV = initial value of the hash

$CV_i = f(CV_{i-1}, B_{i-1}) \quad 1 \leq i \leq L$

$H(M) = CV_L$

EXAMPLES: MERKLE-DAMGÄRD CONSTRUCTION

- Algorithm with chained iterations (stages)
- Initial phase:
  - The message is divided in L blocks (B) of length b
  - The total message length is appended to the last block
  - If necessary, a padding is also appended. It makes harder to find collisions:
    - 2 equal length messages that collide
    - 2 different length messages that collide when appending their own length

EXAMPLES: MERKLE-DAMGÄRD CONSTRUCTION

- COMPRESSION FUNCTION
  - 2 inputs: previous output (or IV for the first stage) + corresponding block
  - Each stage produces an n bit hash value
  - The final hash value is n bits length
- If the compression function is collision resistant, so is the hash function (not necessary the reverse)
- Compression function design à security core
  - The hash function cryptanalysis is focused on the compression function

EXAMPLES: MD5

- Designed by Ronald L.Rivest in 1991
- Mode of operation
  - Hash value of 128 bits length
  - Input message is divided into blocks of 512 bits length
  - Padding addition to the last block
  - Each block is again divided in16 sub-blocks of 32 bits length
  - 4 rounds are performed,having 16 operations each of them:
    - Non-lineal functions
    - Addition modulo 2^32
    - Bit rotation

EXAMPLES: MD5

- ATTACKS
  - First weaknesses discovered (1996)
  - First algorithms to find collisions (2004) http://eprint.iacr.org/2004/199
  - Lenstra, Wang and Weger, they were able to generate two different public key certificates with the same digital signature (MD5-RSA)
  - (2005)
  - http://eprint.iacr.org/2005/067
  - Algorithm that finds collisions in a single minute (2006) http://eprint.iacr.org/2006/105

EXAMPLES: SHA-0, SHA-1

- SHA-0
  - Hash value of 160 bits length
  - Broken in 2005
  - Published an algorithm for finding collisions with just 2^{39} operations
- SHA-1
  - Designed by the NSA
  - Hash value of 160 bits length
  - Similar structure as MD5
  - In 2005, an algorithm to find collisions using 2^{69} operations was published (with brute force,it would be 2^{80})
  - In 2005,the algorithm complexity was reduced to 2^{63}operations

EXAMPLES : SHA-2 FAMILY

- SHA-224, SHA-256, SHA-384 y SHA-512
- Designed by NSA
- New common structure
- SHA-224 and SHA-384 are reduced versions of SHA-256 and SHA-512 (64 rounds instead of 80 and with different initial values)
- No vulnerabilities found yet
- Good solution by now

EXAMPLES

Algorithm	Output size	Internal state size	Block size	Collision
HAVAL	256/224/192/160/128	256	1024	Yes
MD2	128	384	128	Almost
MD4	128	128	512	Yes
MD5	128	128	512	Yes
RIPEMD	128	128	512	Yes
RIPEMD-128/256	128/256	128/256	512	No
RIPEMD-160/320	160/320	160/320	512	No
SHA-0	160	160	512	Yes
SHA-1	160	160	512	With flaws
SHA-256/224	256/224	256	512	No
SHA-512/384	512/384	512	1024	No
WHIRLPOOL	512	512	512	No

EXAMPLES : FAMILIA SHA-3

- Currently, there is a competition to select a new hash function family:SHA-3
  - 2007:Requisites establishment
  - 2008:Call for proposals
  - 2009 (February): First SHA-3 Candidate Conference. Public revision of the candidates
  - 2010 (2Q): Second SHA-3 Candidate Conference. Result analysis and proposed improvements
  - 2010 (3Q):Selection of the finalists
  - 2010 (4Q):Final author touches
  - 2011: Global scientific community analysis.
  - 2012: Last conference. Winner selection http://csrc.nist.gov/groups/ST/hash/sha-3/

CRYPTOGRAPHY AND COMPUTER SECURITY

COSEC

uc3m | Universidad Carlos III de Madrid