Lecture 8: Hash Functions

CS6903: Modern Cryptography
Spring 2010
Nitesh Saxena

Construction

- A hash function is typically based on internal compression function f() that works on fixed-size input blocks (Mi)

\[ \text{IV} \xrightarrow{f} h_1 \xrightarrow{f} h_2 \xrightarrow{f} \cdots \xrightarrow{f} h_{n-1} \xrightarrow{f} h \]

- Works sort of like a Chained Block Cipher
  - Produces a hash value for each fixed-size block based on its content and based on the hash value for the previous block
  - "Avalanche" effect (1-bit change in input produces "catastrophic" changes in output)

- In fact, can use symmetric encryption: \( f=E() \), and use \( M_i \) as the key (but it won’t be fast)
Secure Hash Algorithm (SHA)
- SHA was published by NIST as a standard in 1993
  - Revised in 1995 as SHA-1
    - Input: Up to 2^64 bits
    - Output: 160 bit digest
    - 80-bit collision resistance
  - Pad with at least 64 bits to resist padding attack
    - 1024...0\text{message length}>
  - Processes 512-bit block
    - Initialize 5x32-bit MD registers
    - Apply compression function
      - 4 rounds of 20 steps each
      - each round uses different non-linear fi
      - registers are shifted and switched

Digest Generation with SHA-1

3/17/2010 Lecture 1 - Introduction
SHA-1 of a 512-Bit Block

Figure 3.5 SHA-1 Processing of a Single 512-bit Block

Basic Steps

Step1: Padding
Step2:Appending length as 64 bit unsigned
Step3: Initialize MD buffer 5 32-bit words
   A|B|C|D|E
   A = 67452301
   B = efcdb89
   C = 98badcfe
   D = 10325476
   E = c3d2e1f0
Basic Steps...

Step 4: the 80-step processing of 512-bit blocks: 4 rounds, 20 steps each
Each step \( t \) (\( 0 \leq t \leq 79 \)):
- Input:
  - \( W_t \) - a 32-bit word from the message
  - \( K_t \) - a constant
  - ABCDE: current MD
- Output:
  - ABCDE: new MD

Basic Steps...

- Only 4 per-round distinctive additive constants
  - \( 0 \leq t \leq 19 \quad K_t = 5A827999 \)
  - \( 20 \leq t \leq 39 \quad K_t = 6ED9EBA1 \)
  - \( 40 \leq t \leq 59 \quad K_t = 8F1BCCDC \)
  - \( 60 \leq t \leq 79 \quad K_t = CA62C1D6 \)
Basic Steps - The Heart Of The Matter

A B C D E

Basic Logic Functions

- Only 3 different functions

Round    Function \( f_t(B, C, D) \)
0 ≤ t ≤ 19   \( (B \land C) \lor (\neg B \land D) \)
20 ≤ t ≤ 39  \( B \oplus C \oplus D \)
40 ≤ t ≤ 59  \( (B \land C) \lor (B \land D) \lor (C \land D) \)
60 ≤ t ≤ 79  \( B \oplus C \oplus D \)
Twist With $W_t$'s

- Additional mixing used with input message 512-bit block
  $W_0|W_1|...|W_{15} = m_0|m_1|m_2|...|m_{15}$
  For $15 \leq t < 80$:
  $W_t = W_{t-16} \oplus W_{t-14} \oplus W_{t-8} \oplus W_{t-3}$

- XOR is a very efficient operation, but with multilevel shifting, it should produce very extensive and random mixing!

Today’s Reading

- [http://www-cse.ucsd.edu/users/mihir/cse207/w-hash.pdf](http://www-cse.ucsd.edu/users/mihir/cse207/w-hash.pdf)
- HAC chapter on Hash Functions