HASHES AND MESSAGE DIGESTS

Method, application and standards

Features of a Hash Function (I)

- $f(\text{message})$ and normally < length of message
- $f(.)$ is a one way function
- secure if
  - knowing $f(m_1)$ infeasible to find $m_2$ such that $f(m_1) = f(m_2)$
  - infeasible to find $m_1$ and $m_2$ such that $f(m_1) = f(m_2)$

Features of a Hash Function (II)

- $f(m)$ may not be predicted from any part of $m$
- typical length of $f(m)$ is 128 bits but SHA-1 is 160 bits

Application of Message Digests

- Protection of stored data and programs
- Authentication or MAC generation
- Encryption
Length of secure message digest
(Birthday Paradox)

- if more than 23 people in a room then very likely that two will have the same birthday
- if $n$ people in a room and $k$ possible birthdays there are $n(n-1)/2$ possible pairs each with prob $1/k$ of a birthday match
- then prob of at least one match is $n(n-1)/2k$ which is $\approx \frac{n^2}{2k} \geq 0.5$ if $n \geq \sqrt{k}$

Birthday Paradox and MD length

- Let length of message digest be $L$ bits
- then there are $2^L$ possible message digests and from the Birthday Paradox $2^{L/2}$ messages should be tested before a match is found
- since testing $2^{64}$ would be infeasible, $L$ should be 128 bits

Authentication with a message digest

A \quad B

$\rightarrow$

$\rightarrow$

$\rightarrow$

$\rightarrow$

$\rightarrow$

MAC Generation

- compute $\text{MD}(K_{A\|l\|m})$
- HMAC uses two hashes
Encryption with a message digest

\[ b_1 = \text{MD}(K_{AD} \oplus IV) \]
\[ c_1 = m_1 \oplus b_1 \]
\[ b_2 = \text{MD}(K_{AD} \oplus c) \]
\[ c_2 = m_2 \oplus b_2 \]

\[ b_3 \quad b_2 \quad b_1 \]
\[ \quad m_3 \quad m_2 \quad m_1 \]

NB For decryption calculate \( b_1 \) and then \( m_1 = c_1 \oplus b_1 \)

Message Digest Standards

<table>
<thead>
<tr>
<th>MD2</th>
<th>RFC 1319</th>
<th>128 bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD4</td>
<td>RFC 1320</td>
<td>128 bit</td>
</tr>
<tr>
<td>MD5</td>
<td>RFC 1321</td>
<td>128 bit</td>
</tr>
<tr>
<td>SHA-1</td>
<td>NIST</td>
<td>160 bit</td>
</tr>
</tbody>
</table>

SHA-1 Overview

SHA-1 Padding

Original message 1000……000000 original message length in bits

Multiple of 512 bits
**SHA-1 Stage Operation (I)**

- 512 bit message block has 16 32bit words
- \( W_0, W_1, W_2, \ldots, W_{15} \)

For \( n \geq 16 \),

\[
W_n = W_{n-3} \oplus W_{n-8} \oplus W_{n-14} \oplus W_{n-16}
\]

where \( \oplus \) is bitwise exclusive or

**SHA-1 Stage Operation (II)**

- \( A' = E + (A \ll 5) + W_n + K_n + f(n,B,C,D) \)

where \( A \ll 5 \) is left rotate 5 bits, and the "constant" \( K_n \) takes different values for the ranges (0-19), (20-39), (40-59), (60-79), and the function \( f() \) also depends on the same four ranges

The 80 iterations \( (n = 0, 1, 2, \ldots, 79) \) is equivalent to 5 passes over a 16 x 32 = 512 bit message block

---

**HMAC**

- Provides a standard way to compute a MAC using a hash function
- Is a function of message and secret key
- Is secure if underlying hash function is secure
- May be used with SHA-1 to give a 160 bit MAC

**HMAC Overview**

- Key (variable length) padded to 512 bits
- SHA-1
- 512 bits
- 160 bits
- HMAC

160 bit MAC if SHA-1 is used