Follow-up From Last Class

→ Review of Hash Functions

Hash Functions have the following properties:
1) Pre-image Resistance (PR)
2) Collision Resistance (CR)
3) Weak Collision Resistance (WCR)
   (also known as 2nd Pre-image Resistance)

NOTE CR \implies WCR

→ Hash Functions are Deterministic

For a hash \( y = H(x) \),
   given \( x \), \( H(x) = H(x) \) ALWAYS.

→ Examples of hash functions that don't have all the Properties

1) MDS
   \implies no longer CR
   \implies has PR, but no CR
2) MYTHIC HASH
   \implies has CR, but no PR
   \implies YES this is possible (See the following example)

Example

Given hash \( H \) which is secure and has output of \( d \)-bits
Construct a hash \( H' \) which has output \((d+1)\)-bits and is CR but not PR.

\[
y = H'(x) = \begin{cases} 
0 \| x, & \text{if } |x| = n \\
1 \| H(x), & \text{otherwise}
\end{cases}
\]
The fact that $H'$ isn't PR is trivial; when $|x| = n$, you are just concatenating 0 with $x$.

To show CR:

Case 1: $y = 0$...

$ightarrow$ collisions can't exist because same output $\Rightarrow$ same input

Case 2: $y = 1$...

$ightarrow$ for example take $y = 1 \| H(x_1)$

$y = 1 \| H(x_2)$

Since $H$ is a hash that is secure, by definition it is CR.

Hence, in case 2, $H'$ is also CR.

Taken together we see that $H'$ is in fact CR, but not PR.

\[ CR \not\Rightarrow PR \]

In Today's Lecture

- Other Properties
- Details of $H$

Two Other Properties (we sometimes assume)

4) Random Output (RO)
5) Non-Malleability
Internals

fixed length block

$M: \quad \downarrow \quad \downarrow \quad \downarrow$

compression function $f$

$\sim$ similar to

different: $\{0, 1\}^{2d} \rightarrow \{0, 1\}^{d}$

$\downarrow \quad \downarrow \quad \downarrow$

$f \quad f \quad f$

$\quad \downarrow \quad \downarrow \quad \downarrow$

$\text{ACCUMULATE}$

$\rightarrow \quad \rightarrow \quad \rightarrow \quad \rightarrow \quad \rightarrow$

$\text{Single Block}$

---

**Merkle-Damgård** (see Smart p 155 for more info)

$d = 256$ bits

$\begin{array}{c}
M \\
\downarrow \\
d \\
\downarrow \\
M_1 \\
\downarrow \\
f \\
\downarrow \\
f \\
\downarrow \\
f \\
\downarrow \\
\text{output}
\end{array}$

$d$-bits

$f$ is a compression function

$(CR, PR)$ and $\{0, 1\}^{2d} \rightarrow \{0, 1\}^{d}$

Question 1) What is the input to $f$ on the first instance?
(shown as $\square$ in the diagram)

Answer) A Random number, picked once by the designer

Answer) $\Rightarrow$ initialization vectors
Random Output (RO)

For a hash $y = H(x)$, it is infeasible to distinguish $y$ from a uniform random value of length $d$-bits.

\[ \downarrow \text{Desired} \]

\[ \downarrow \text{Sometimes you assume RO} \]

\[ \downarrow \text{this is CONTROVERSIAL} \]

\[ \downarrow \text{"RANDOM ORACLE MODEL" \rightarrow assumes RO} \]

\[ \rightarrow \text{has other nice properties that won't be discussed here} \]

\[ \downarrow \text{Fiat-Shamir Heuristic} \]

\[ \downarrow \text{OAEP \rightarrow padding scheme for RSA Public Key Encryption} \]

Non-Malleability

Given $y$ ($y = H(x)$), it is infeasible to produce $y'$ where $y' = H(f(x))$.

Examples: 

| $y' = h(x+1)$ | ✓ |
| $y' = h(2x)$ | ✓ |
| $y'' = h(x \| M)$ | × |

\[ \downarrow \text{length extension attacks} \]

AH!
Padding

Example that is NOT used:

Zero-pad: \( m = \overline{m_1} \overline{m_2} \overline{00} \ldots \overline{0} \)

\( \uparrow \) Breaks CR

For example, \( h(0101) = h(01010) = h(010100) \)

Another example:

\( m = \overline{m_1} \overline{m_2} \overline{100} \ldots \overline{0} \)

\( \uparrow \)

add as many zeros as needed to fill up the block

Length Extension Attacks

Merkle-Damgard

\( x \rightarrow H \rightarrow y \)

\( y = H(x) \)

\( y' = H(x || m) \)

Still PR

Break Malleability
Later we will see:
MACs which is a signed hash that's not used anymore

![Diagram of Alice to Bob communication]

"k" which is secret

How secure are well known hashes?

- **MD5**
- **SHA1**
- **SHA2 → SHA256, SHA512**
- **SHA3**

SHA 3 is NOT vulnerable to length extension attacks

- uses sponge construction instead of Merkle-Damgward

Merkle-Damgward based, so vulnerable to length extension problems
Moral of the Story: THIS STUFF IS COMPLICATED! KNOW AND APPRECIATE THE PROPERTIES!

Compression Function Internals (for more info see Smart p156)

\[
\begin{align*}
\text{Compression function} & \quad (\text{MD4}) \\
M & \xrightarrow{f} A B C D \\
f^{-1} & \\
\end{align*}
\]

function changes every round

\[
\begin{align*}
A & = A \oplus g(B, C, D) + M + \text{Const.} \\
(A, B, C, D) & = (D, +\ll\text{Const}, B, C) \\
\end{align*}
\]

changes every time

Repeat for 3 rounds. After each round change \( g \).

\[
\begin{align*}
g(B, C, D) \\
g_1 = (B \land C) \lor (\neg B \land D) \\
g_2 = (B \land C) \lor (B \land D) \lor (C \land D) \\
g_3 = B \oplus C \oplus D \\
\end{align*}
\]

Again, just appreciate this; won't be tested directly on the compression function internals.