Lecture 7: Hash Functions

Viet Tung Hoang

The slides are loosely based on those of Prof. Mihir Bellare, UC San Diego.
1. Security Modeling for Hash Functions

2. Building Hash Function: MD Transform
Cryptographic Hash Functions

MD5  SHA-1  SHA-2  SHA-3

Obsolete, should not use

Crypto hash functions are often used for collision-resistant data compression
But there are also several other purposes
Collision-Resistance

\[ f : \text{Domain} \rightarrow \text{Range} \]

By Pigeonhole Principle, if \(|\text{Domain}| > |\text{Range}|\) then collision exists

**Want:** collisions are hard to find, although they exist
Defining Collision-Resistance

$\textbf{CR}_H$

procedure Initialize()  
\hspace{1cm} K \leftarrow \mathcal{K}; \text{ Return } K 

procedure Finalize(X_1, X_2)  
\hspace{1cm} \text{Return } (H_K(X_1) = H_K(X_2))

$K$  

$A$  

$(X_1, X_2)$

\[ \text{Adv}^{\text{cr}}_H(A) = \Pr[\text{CR}_H^A \Rightarrow 1] \]

\textbf{Question:} What’s the difference between this and almost-universal hash?

\textbf{Answer:} For almost-universal hash, the adversary is \textbf{not} given the key.
Keyless Hash Functions

A hash function is **keyless** if $\mathcal{K} = \{\varepsilon\}$

Write $H(X)$ instead of $H_\varepsilon(X)$

All practical cryptographic hash functions are keyless
**Application: File Comparison**

**Want:** Check the equality of the two big files using little bandwidth

\[ Y \leftarrow H(\text{File}_A) \]

Check if \( Y = H(\text{File}_B) \)
Application: Storage Auditing

Client

\[ K, H_K(x) \]

Info

Challenge \[ K \]

\[ x \]

Dropbox

Response \[ H_K(x) \]

Collision resistance is **not** enough here \(\rightarrow\) Need stronger properties
-Everybody, including the adversary, has access to RO
-RO returns a random answer for each query
-The answers are consistent
Birthday Attack

distinct $x_1, \ldots, x_q$

Domain

$H_K(x_1)$ \hspace{1cm} \ldots \hspace{1cm} H_K(x_q)$

Look for duplicate values

$\Pr[\text{collision}] \approx \frac{q^2}{2^n}$

If $H$ is modeled as a random oracle

Output length
Agenda

1. Security Modeling for Hash Functions

2. Building Hash Function: MD Transform
Compression Functions

For SHA-2, \( b = 512 \) and \( n = 256 \)
The Merkle-Damgard Transform

Fact: If $h$ is CR then $H = \text{MD}(h)$ is also CR

Can’t attack $H$ if $h$ has no weakness
Designing Compression Functions

Often designed from a blockcipher \( E : \{0, 1\}^b \times \{0, 1\}^n \rightarrow \{0, 1\}^n \)

An **incorrect** design \( h(x \| v) = E_x(v) \)

**Question:** Break the collision resistance of this compression function
Designing Compression Functions

Often designed from a blockcipher $E: \{0, 1\}^b \times \{0, 1\}^n \to \{0, 1\}^n$

A correct design $h(x||v) = E_x(v) \oplus v$