Lecture 23: Collision-Resistant Hash Function

Outline

- In this lecture, we will define various forms of collision resistance
- Construct Collision-resistant Hash Function (CRHF) families
- Finally, we will consider domain extension of CRHF families

Universal Hashing

- Consider a domain D and a range R, and a family of hash functions
$$\mathcal{H} := \{h_i : i \in I\},$$
where I is an index set and each h_i is a $D \to R$ function
- This family has the property that for all distinct inputs $x, x' \in D$, the probability of collision $h_i(x) = h_i(x')$, for randomly chosen index $i \in I$, is "small"
- Specifically,
$$\mathop{\mathbb{P}}_{i \in_R I}\left[h_i(x) = h_i(x')\right] \leqslant \frac{1}{|R|}.$$
- For example, consider the function $h_{a,b} \colon F^2 \to F$ (for any finite field F) defined by
$$h_{a,b}(x_1, x_2) := a \cdot x_1 + b \cdot x_2,$$
where $(a, b) \neq (0, 0)$
- Universal Hashing families exist information-theoretically

Universal One-way Hash Function (UOWHF)

- Consider a domain D and a range R, and a family of hash functions
$$\mathcal{H} := \{h_i : \ i \in I\},$$
where I is an index set and each h_i is a $D \to R$ function
- A UOWHF family has the following property. First, an adversary declares $x \in D$ where it chooses to be challenged. Next, a random index $i \in I$ is selected and sent to the adversary. Then, the adversary presents $x' \in D$ as a candidate collision. The adversary wins the game if $h_i(x) = h_i(x')$ and $x \neq x'$.
- Note that this primitive cannot exist against an adversary of unbounded computational power. Try to find an adversarial strategy with unbounded computational power that finds collisions with high probability when the hash function compresses a lot (i.e., $|R| \ll |D|$)
- Moni Naor and Moti Yung introduced UOWHFs, and Rompel constructed them from OWFs

Collision-Resistant Hash Function (CRHF)

- Consider a domain D and a range R, and a family of hash functions

$$\mathcal{H} := \{h_i : i \in I\},$$

  where I is an index set and each h_i is a $D \to R$ function
- A CRHF family has the following property. First, a random index $i \in I$ is selected and sent to the adversary. Then, the adversary presents two distinct $x.x' \in D$ as a candidate collision. The adversary wins the game if $h_i(x) = h_i(x')$ and $x \neq x'$.
- CRHF has strong security requirements as compared to UOWHF, so they also cannot exist information-theoretically

One-bit Compressing CRHF from Discrete Log

- Let Z_p^* be a (multiplicative) group with generator g such that the discrete log is hard with respect to this generator
- For $y \in Z_p^*$, define the hash function $h_y : Z_p^* \times \{0,1\} \to Z_p^*$ defined by

$$h_y(x, b) = y^b \cdot g^x.$$

- The hash function family

$$\mathcal{H} := \left\{ h_y : y \in Z_p^* \right\}$$

is a CRHF

Length Halving CRHF from Discrete Log

- Let Z_p^* be a (multiplicative) group with generator g such that the discrete log is hard with respect to this generator
- For $y \in Z_p^*$, define the hash function $h_y : Z_p^* \times Z_p^* \to Z_p^*$ defined by

$$h_y(x_1, x_2) = y^{x_2} \cdot g^{x_1}.$$

- The hash function family

$$\mathcal{H} := \left\{ h_y : y \in Z_p^* \right\}$$

is a length-halving CRHF

Domain Extension: Merkle-Damgard Transform

- Suppose $\mathcal{H}$ is a $\{0,1\}^{2B} \rightarrow \{0,1\}^{B}$ CRHF
- Our objective is to construct a $\{0,1\}^{*} \rightarrow \{0,1\}^{B}$ CRHF
- Given $h \in \mathcal{H}$ the Merkle-Damgard transform defines the following function $H(x)$ for an input $x \in \left(\{0,1\}^{B}\right)^{L}$