[CS304] Introduction to Cryptography and Network Security

Course Instructor: Dr. Dibyendu Roy — Winter 2022-2023
Scribed by: Akshay (202051018) — Lecture (Week 8, Week 9)

1 Second Preimage

Given x, we need to find x' ($\neq x$) s.t. h(x') = h(x).

$$h : A \to B$$

$|B| = M$

$X_0 \subseteq A \backslash \{x\}$ $\quad |X_0| = Q$ $\quad h(x_i)$ $\quad$ where $x_i \in X_0$
Probability in this case is same as probability of previous case (Pre-image finding algo).

2 Collision Finding Algo

Find x, x' $\in$ X s.t. x $\neq$ x' and $h(x) = h(x')$
$X_0 \subseteq X$, $|X_0| = Q$ $\quad X_0 = \{x_1, ..., x_Q\}$

for each $x \in X_0$:
- Compute $y_x = h(x)$
- If $y_x = y_{x'}$ for some $x \neq x'$
  then return (x,x')
- else
  return failure

(event) E_i: $h(x_i) \notin \{h(x_1), h(x_2), .., h(x_{i-1})\}$ means $h(x_i)$ does not collide (matches) with any of these previous images.
Each E_i is an event associated with one collision. If its a success then that means no collision.

$\Pr[E_1] = 1$ $\because$ it is sure that you won't find any collision in an empty set.

E_2: $h(x_2) \notin \{h(x_1)\}$
Size of co-domain is M. So, total possible values which $h(x_2)$ can take are M. Out of those, favourable values are all except the value taken by $h(x_1)$.

$\therefore \Pr[E_2 \mid E_1] = \frac{M-1}{M}$

E_3: $h(x_3) \notin \{h(x_1), h(x_2)\}$

$\Pr[E_3 \mid E_2, E_1] = \frac{M-2}{M}$

We have incorporated conditional probability is because if previous events got happened then

it means we have no collision till now and values in the set are distinct. This gives us write to say favourable values are M-2. On the contrary if E_2, E_1 would have not happened then there would be a possibility of collision (match) between $h(x_2)$ and $h(x_1)$ and in that case we would have not got the right to say M-2 favourable cases (M-1 distinct favourable values would have worked then).

$$Pr[E_3|E_1 \cap E_2] = \frac{M-2}{M}$$

$$\frac{Pr[E_1 \cap E_2 \cap E_3]}{Pr[E_1 \cap E_2]} = \frac{M-2}{M}$$

$$Pr[E_1 \cap E_2 \cap E_3] = \frac{M-2}{M} \times Pr[E_1 \cap E_2]$$

$$= \frac{M-2}{M} \times \frac{M-1}{M}$$

$$Pr[E_1 \cap E_2 \cap E_3 ... \cap E_Q] = \frac{M-1}{M} \times \frac{M-2}{M} \times \frac{M-3}{M} \times ... \times \frac{M-(Q-1)}{M}$$

Pr[Collision finding Algo return success] = 1 - $Pr[E_1 \cap E_2 \cap E_3 ... \cap E_Q]$

$$= 1 - \tfrac{M-1}{M} \times \tfrac{M-2}{M} \times \tfrac{M-3}{M} \times ... \times \tfrac{M-(Q-1)}{M}$$

$\frac{M-i}{M} = 1 - \frac{i}{M} \approx e^{-\frac{i}{M}}$ ($\because e^{-x} \approx 1 - x$, when x is small)

$\frac{M-1}{M} \times \frac{M-2}{M} \times \frac{M-3}{M} \times ... \times \frac{M-(Q-1)}{M} \approx \Pi_{i=1}^{Q-1} e^{-\frac{i}{M}} = e^{(-\sum_{i=1}^{Q-1} \frac{i}{M})}$

$$= e^{(-\frac{1}{M} \frac{(Q-1).Q}{2})}$$

Pr[Collision] $\approx 1 - e^{-\frac{1}{M} \frac{Q(Q-1)}{2}}$

$\epsilon \approx 1 - e^{-\frac{1}{M} \frac{Q(Q-1)}{2}}$

$e^{-\frac{1}{M} \frac{Q(Q-1)}{2}} \approx 1 - \epsilon$

$-\frac{Q(Q-1)}{2M} \approx ln(1-\epsilon)$

$\implies Q^2 - Q \approx -2Mln(1-\epsilon)$

$Q^2 - Q \approx 2M.ln(\frac{1}{1-\epsilon})$

$Q^2 \approx 2M.ln(\frac{1}{1-\epsilon})$ (on assuming Q very large)

$Q \approx \sqrt{2M.ln(\frac{1}{1-\epsilon})}$

$= \sqrt{2.ln(\frac{1}{1-\epsilon})}\sqrt{M}$

If we take $\epsilon = 0.9$ (high probability, close to 1)

$Q \approx \sqrt{2 \times ln(\frac{1}{1-0.9})}\sqrt{M}$

$= \sqrt{2 \times ln10}\sqrt{M}$

$\approx 2.14\sqrt{M}$

On taking $\epsilon = \frac{1}{2}$ i.e., probability of success = 1/2 then:
$Q \approx 1.177\sqrt{M}$

Conclusion: There is high chance of collision event to get succeeded when the Q is of order of $\sqrt{M}$.

Complexity of collision finding = O($\sqrt{M}$)

3 Summary of hash function problems

$h : X \to Y \qquad |Y| \geq 2|X|$

Preimage $\to$ O($|Y|$)

Collision $\to$ O($\sqrt{|Y|}$)

Second preimage $\to$ O($|Y|$)

4

$h : \{0,1\}^* \to \{0,1\}^m$