Birthday Attack:

(a) From N identical objects, if one is randomly picked & marked at a time (with replacement), after q trials, what is the probability that a marked object was picked twice or more?

Let

$P[Coll_q]$: Prob of ≥ 1 collision after q trials

$= 1 - P[NoColl_q]$

$P[Coll_1] = 0$, $P[NoColl_1] = 1$

$P[NoColl_2 | NoColl_1] = 1 - 1/N$

$P[NoColl_3 | NoColl_2] = 1 - 2/N$

$P[NoColl_q | NoColl_{q-1}] = 1 - \frac{q-1}{N}$

$$P[NoColl_q] = P[NoColl_q | NoColl_{q-1}] \cdot P[NoColl_{q-1}]$$
$$= P[NoColl_q | NoColl_{q-1}] P[NoColl_{q-1} | NoColl_{q-2}] P[NoColl_{q-2}]$$
$$\vdots$$
$$= P[NoColl_q | NoColl_{q-1}] P[NoColl_{q-1} | NoColl_{q-2}] \ldots P[NoColl_1]$$
$$= \left(1 - \frac{q-1}{N}\right)\left(1 - \frac{q-2}{N}\right)\left(1 - \frac{q-3}{N}\right) \ldots \left(1 - \frac{1}{N}\right) \cdot 1$$
$$= \prod_{i=1}^{q-1} \left(1 - \frac{i}{N}\right)$$

$$P[Coll_q] = 1 - \prod_{i=1}^{q-1}\left(1 - \frac{i}{N}\right)$$

$\forall\, 0 \leq x \leq 1$, $x \in \mathbb{R}$,

$$\underbrace{1 - x \leq e^{-x}}_{①} \underbrace{\leq 1 - x/2}_{②}$$

$$P[NoColl_q] \leq \prod_{i=1}^{q-1} e^{-i/N} \qquad \text{From ①}$$
$$\leq e^{-\sum_{i=1}^{q-1} i/N}$$
$$\leq e^{-\frac{1}{N}\sum_{i=1}^{q-1} i}$$
$$\leq e^{-\frac{q(q-1)}{2N}}$$
$$\leq 1 - \frac{q(q-1)}{2N \times 2} \qquad \text{from ②}$$

$$\boxed{P[Coll_q] = 1 - P[NoColl_q] \geq \frac{q(q-1)}{4N}}$$

Condition: $0 \leq \frac{q(q-1)}{2n} \leq 1$

For high probability of collision, $q \in O(\sqrt{N})$

SHA 0 started with 80 bits. To get a collision with reasonable probability, it takes $O(2^{40})$ trials only.

$$\frac{q(q-1)}{4N} \leq P[Coll_q] \leq \frac{q(q-1)}{2N}$$

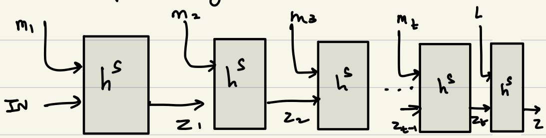
Merkle-Damgard Transform :

Given $h^s : \{0,1\}^{2n} \to \{0,1\}^n$ is collision resistant

Goal : $H^s : \{0,1\}^* \to \{0,1\}^n$

$H^s(m)$
$\hookrightarrow m_1 m_2 \dots m_t$, $|m_i| = n$

with 10^* padding



Collision in H^s : $x \neq y$ ∋ $H^s(x) = H^s(y)$

Case 1 : $|x| \neq |y|$

$H^s(x) = H^s(y)$

$\Rightarrow$ the last h^s gives the same z for two different L since $|x| \neq |y|$

$\Rightarrow h^s(x') = h^s(y')$, where $x' \neq y'$

where $x' = x_{last} \,\|\, |x|$, $y' = y_{last} \,\|\, |y|$

Which contradicts the fact that h^s is collision resistant.

Case 2 : $|x| = |y|$

There must be some block where h^s will have same outputs for diff. inputs since $H^s(x) = H^s(y)$ but $x \neq y$.

This works for any compressive collision resistant hash fn. ex : $n+1 \to n$, by changing the block size of m_i to 1.

Consider DLP in Group G with generator g on which DLP is hard.

$\mathbb{Z}_p^*$; $g^x \bmod p$

Public $r \in \mathbb{Z}_p^*$,

$h^s(x, y) = g^x \cdot r^y$

is a collision resistant $2n \to n$ hash fn.

Collision : $h^s(x_1, y_1) = h^s(x_2, y_2)$ where $(x_1, y_1) \neq (x_2, y_2)$

$$g^{x_1} r^{y_1} = g^{x_2} r^{y_2}$$

$$\Rightarrow g^{x_1 - x_2} = r^{y_2 - y_1}$$

$$\Rightarrow r = g^{\frac{x_1 - x_2}{y_2 - y_1}}$$ where $\frac{1}{y_2 - y_1}$ = inverse of $y_2 - y_1$ modulo p

It is easy to compute $\frac{x_1 - x_2}{y_2 - y_1}$ using extended euclidian algorithm.

But, $\frac{x_1 - x_2}{y_2 - y_1} = \log_g r$, i.e. DLP of r.

but DLP is hard, so the hash fn. is collision resistant.