Birthday Attacks

BIRTHDAY ATTACKS: Assume people are born uniformly at random on each day independently.

Birthday Paradox: How many people do we need in a room so that with prob. $\geq 50\%$, at least 2 people have the same birthday? (ignoring leap years)

A: $23$. (in comparison, we need $366$ people for this to happen $100\%$)

Why?

- $K = 2$: $\frac{1}{365}$
- $K = 3$: Analyze the prob. of no collision:

$$K = \Pr[\text{no collision}] = \frac{364 \cdot 363 \cdot \cdots \cdot 365}{365 \cdot 364 \cdot \cdots \cdot 366}$$

- $K = 23$:

$$\Pr[\text{no collision}] = \left(1 - \frac{1}{365}\right) \cdot \left(1 - \frac{2}{365}\right) \cdots \left(1 - \frac{22}{365}\right) \approx 0.493 \Rightarrow \Pr[\text{collision}] \approx 50.7\%$$

$N = 365$ $\Rightarrow$ How about general values of $N$?

When there are $k$ people,

$$\Pr[\text{collision}] = 1 - \left[1 \cdot \left(1 - \frac{1}{N}\right) \cdot \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{k-1}{N}\right)\right]$$

$$= 1 - \frac{(k-1)!}{N^k} \quad k \approx \sqrt{\ln N} \approx 1.77\sqrt{N}$$

Find collisions for hash functions

Suppose $h(x)$ can take $N$ values.

Goal: Find $x \neq y$ s.t. $h(x) = h(y)$.

Attack: Try approximately $\sqrt{N}$ random $x$ and evaluate $h(x)$.

Example: MD5 digest = 128 bits, $N = 2^{128}$.

$$\sqrt{N} = 2^{64} \approx 10^{19}$$

Runtime: $O(\sqrt{N} \cdot T)$ where $T$ is time we need to evaluate $h(\cdot)$ once.

Applications of hash functions:

1. Identity testing (quick test if two files are the same) objects
2. Data integrity
3. Digital signatures (in general to speed things up)
4. Hash tables as data structure