现代密码学

第四十二讲 生日攻击

信息与软件工程学院

生日攻击-相关问题

- 问题1---第Ⅰ类生日攻击问题
  - 已知一杂凑函数H有n个可能的输出，H(x)是一个特定的输出，如果对H随机取k个输入，则至少有一个输入y使得H(y)=H(x)的概率为0.5时，k有多大？
  - 为叙述方便，称对杂凑函数H寻找上述y的攻击为第Ⅰ类生日攻击。

- 因为H有n个可能的输出，所以输入y产生的输出H(y)等于特定输出H(x)的概率是1/n，反过来说H(y)≠H(x)的概率是1-1/n。
- y取k个随机值而函数的k个输出中没有一个等于H(x)，其概率等于每个输出都不等于H(x)的概率之积，为$[1-1/n]^k$，所以y取k个随机值得到函数的k个输出中至少有一个等于H(x)的概率为$1-[1-1/n]^k$。
- 由当$|x|<<1$，$(1+x)^k \approx 1+kx$，可得

$$1-[1-1/n]^k \approx 1-[1-k/n]=k/n$$

- 若使上述概率等于0.5，则k=n/2。特别地，如果H的输出为m比特长，即可能的输出个数$n=2^m$，则

$$k=2^{m-1}。$$

生日攻击-相关问题

- 问题2---生日悖论
  - 在k个人中至少有两个人的生日相同的概率大于0.5时，k至少多大？

生日攻击-相关问题

- 问题2---生日悖论
  - 设有k个整数项，每一项都在1到n之间等可能地取值。
    P(n, k)：k个整数项中至少有两个取值相同的概率
  - 生日悖论就是求使得$P(365,k)\geq 0.5$的最小k。
    Q(365, k)：k个数据项中任意两个取值都不同的概率。

生日攻击-相关问题

- 问题2---生日悖论
  - 如果k>365，则不可能使得任意两个数据都不相同，因此假定k≤365。k个数据项中任意两个都不相同的所有取值方式数为

$$365 \times 364 \times \cdots \times (365-k+1) = \frac{365!}{(365-k)!}$$

  - 如果去掉任意两个都不相同这一限制条件，可得k个数据项中所有取值方式数为365^k。所以可得

$$Q(365,k) = \frac{365!}{(365-k)!\ 365^k}$$

$$P(365,k) = 1 - Q(365,k) = 1 - \frac{365!}{(365-k)!\ 365^k}$$

生日攻击-相关问题

- 问题2---生日悖论
  - 当k=23时，P(365,23)=0.5073，即上述问题只需23人，人数如此之少。
  - 若k=100，则P(365,100)=0.9999997，即获得如此大的概率。
  - 之所以称这一问题是悖论是因为当人数k给定时，得到的至少有两个人的生日相同的概率比想象的要大得多。

生日攻击-相关问题

- 问题---生日悖论推广问题
  - 已知一个在1到n之间均匀分布的整数型随机变量，若该变量的k个取值中至少有两个取值相同的概率大于0.5，则k至少多大？
  - $P(n,k)=1-\dfrac{n!}{(n-k)!n^k}$，令$P(n, k)>0.5$，可得

$$k=1.18\sqrt{n}\approx\sqrt{n}$$

  - 若取n=365，则

$$k=1.18\sqrt{365}=22.54$$

生日攻击

- 生日攻击
  - 设杂凑函数H有2^m个可能的输出（即输出长m比特），如果H的k个随机输入中至少有两个产生相同输出的概率大于0.5，则

$$k \approx \sqrt{2^m} = 2^{m/2}$$

  - 第II类生日攻击：寻找函数H的具有相同输出的两个任意输入的攻击方式。

安全应用

- 输出长度与碰撞
  - 这种生日攻击给出了消息摘要长度的下界，一个40比特的消息摘要是非常不安全的。
  - 因为在大约2^{20}个（大约100万）个随机值中就能以1/2的概率找到一个碰撞。
  - 通常建议消息摘要的最小可接受的长度为128比特，在DSS签名标准中使用160比特的消息摘要就是基于这个考虑。