Understanding Cryptography

Answers of Homework No.6

Chapter 11

1.

- Birth-date and the collision challenge!

a) What is the minimum number of students in a class needed to have at least two students with the same birth-date with probability more than 1/2?

b) If a year has N days and the number of students is K, find the probability of having at least two students with the same birth-date as a function of K and N.

c) If we want to observe collision in a hash function with outputs of size n bits with probability more than 1/2, how many random messages do we need?
(Hint: You can use the inequality $1 - x \leq e^{-x}$. $\quad x > 0$)

حل :

$a)$

$$P(no\ \ collision\ among\ t\ people) = \left(1 - {}^{1}\!/_{365}\right).\left(1 - {}^{2}\!/_{365}\right) \ldots\ldots \left(1 - {}^{n-1}\!/_{365}\right)$$

$$p(at\ least\ one\ collision) = 1 - p(\ no\ collision) = \ 1 - \left(1 - \frac{1}{365}\right) \ldots \ldots \left(1 - \frac{23-1}{365}\right) = 0.5$$

$$n = 23$$

حداقل 23 نفر در یک کلاس باید وجود داشته باشد که با احتمال بیشتر از پنجاه درصد حداقل دو نفر یک روز تولد یکسان در سال داشته باشند.

$$1 - \prod_{i=1}^{n}\left(1 - \frac{i-1}{365}\right) \geq \frac{1}{2} \Rightarrow \prod_{i=1}^{n}\left(1 - \frac{i-1}{365}\right) \leq \frac{1}{2} \Rightarrow n = 23$$

$$\prod_{i=1}^{23}\left(1 - \frac{i-1}{365}\right) = 0.49 < 0.51 \qquad \Rightarrow n \geq 23$$

$b)$

$$p(no\ collision) = \left(1-\frac{1}{N}\right).\left(1-\frac{2}{N}\right)\ldots\left(1-\frac{K-1}{N}\right) \Rightarrow$$

$$p(\ at\ least\ one\ collision) = 1-\prod_{i=1}^{K}\left(1-\frac{i-1}{N}\right)$$

$c)$

$$p\ (no\ collision) = \left(1-\frac{1}{2^n}\right).\left(1-\frac{2}{2^n}\right)\ldots\left(1-\frac{t-1}{2^n}\right) = \prod_{i=1}^{t-1}\left(1-\frac{i}{2^n}\right)$$

از طرفی طبق بسط تیلور داریم:

$$1-x \le e^{-x}.\quad x>0$$

$$p\ (no\ collision) = \prod_{i=1}^{t-1}\left(e^{-i/2^n}\right)$$

از طرفی با توجه به اینکه:

$$p(no\ collision) = \ e^{-\frac{1+2+3+\cdots+t-1}{2^n}}$$

$$1+2+3+\cdots\ldots.+t-1 = \frac{t(t-1)}{2}$$

$$p\ (no\ collision) = \prod_{i=1}^{t-1} e^{-t(t-1)/2.2^n}$$

$$p\ (\ at\ least\ one\ collision) = 1-\ p\ (no\ collision) = \alpha$$

$$\alpha = 1-\ e^{-t(t-1)/2^{n+1}}$$

$$\ln(1-\alpha) = -\frac{t(t-1)}{2^{n+1}}$$

$$t(t-1) = 2^{n+1}\ln(\frac{1}{1-\alpha})$$

$$for\ t \gg 1. \quad t^2 \approx t(t-1)$$

$$t \approx \sqrt{2^{n+1}\ \ln\left(\frac{1}{1-\alpha}\right)}$$

$$t \approx 2^{(n+1)/2}\ \sqrt{\ln(\frac{1}{1-\alpha})}$$

2.

We consider three different hash functions which produce outputs of lengths 64, 128 and 160 bit. After how many random inputs do we have a probability of $\varepsilon = 0.5$ for a collision? After how many random inputs do we have a probability of $\varepsilon = 0.1$ for a collision?

حل:

$$\boldsymbol{t \approx \sqrt{2^{n+1}.\ ln(\frac{1}{1-\varepsilon})}}$$

$Length$	$\varepsilon = 0.1$	$\varepsilon = 0.5$
$64\ bit$	$\approx \sqrt{2^{64+1}.\ ln\left(\frac{1}{1-0.1}\right)} =$ $2^{32}.\sqrt{2 \times ln\, ^{10}/_9} =$ $2^{32} \times 0.45$	$t \approx \sqrt{2^{64+1}.\ ln\left(\frac{1}{1-0.5}\right)} =$ $2^{32}.\sqrt{2.ln2} = 2^{32}.\sqrt{2.0.69} =$ $2^{32} \times 1.17$
$128\ bit$	$\approx \sqrt{2^{128+1}.\ ln\left(\frac{1}{1-0.1}\right)} =$ $2^{64}.\sqrt{2 \times ln\, ^{10}/_9} =$	$\approx \sqrt{2^{128+1}.\ ln\left(\frac{1}{1-0.5}\right)} =$ $2^{64}.\sqrt{2 \times ln2} =$ $2^{64} \times 1.17$

	$2^{64} \times 0.45$	
160 *bit*	$\approx \sqrt{2^{160+1} \cdot \ ln\left(\frac{1}{1-0.1}\right)} =$ $2^{80} \cdot \sqrt{2 \times ln \ {}^{10}/_{9}} = 2^{80} \times 0.45$	$\approx \sqrt{2^{160+1} \cdot \ ln\left(\frac{1}{1-0.5}\right)} =$ $2^{80} \cdot \sqrt{2 \times ln2} =$ $2^{80} \times 1.17$

Chapter 12

3.

We study two methods for integrity protection with encryption.

3.1. Assume we apply a technique for combined encryption and integrity protection in which a cipher text c is computed as

$$c = ek(x||h(x))$$

Where $h()$ is a hash function. This technique is not suited for encryption with stream ciphers if the attacker knows the whole plaintext x. Explain *exactly* how an active attacker can now replace x by an arbitrary x' of his/her choosing and compute c' such that the receiver will verify the message correctly. Assume that x and x' are of equal length. Will this attack work too if the encryption is done with a one-time pad?

3.2. Is the attack still applicable if the checksum is computed using a keyed hash function such as a MAC:

$$c = e_{k_1}(x||MAC_{k_2}(x))$$

Assume that $e()$ is a stream cipher as above.

حل:

3.1.

$$c_i = z_i \oplus \{ x_1 \, x_2 \ldots . x_n \parallel H_1(x)H_2(x) \ldots . H_m(x)\}; \quad i = 1,2, \ldots . , n+m$$

با فرض اینکه x تعداد n بیت دارد. اسکار ابتدا مقدار زیر را محاسبه می کند.

$$z_i = x_i \oplus c_i \qquad i = 1.2. \ldots . n$$

اسکار به دلیل اینکه مقدار x را می داند مقدار $H(x)$ را محاسبه می کند. با فرض اینکه $H(x)$ دارای m بیت خروجی است اسکار مقدار زیر را محاسبه می کند.

$$z_{j+n} = H_j(x) \oplus c_{j+n} \qquad j = 1,2,\dots,m$$

اسکار مقدارهای زیر را نیز محاسبه می کند.

$$H(x')$$
$$c'_i = z_i \oplus x'_i \qquad i = 1,2,\dots,n$$
$$c'_{j+n} = z_{j+n} \oplus H_j(x') \qquad j = 1,2,\dots,m$$

به دلیل اینکه مهاجم کلید را بااستفاده از $plaintext$ و $ciphertext$ به دست می آورد این حمله در صورت استفاده از OTP نیز قابل انجام است. البته مهاجم باید برای هر پیام ردو بدل شده همه مراحل بالا را انجام دهد، به دلیل اینکه کلید هر سری متفاوت می باشد.

3. 2.

خیر.

با وجود اینکه اسکار مقادیر $z_1, z_2, \dots . z_n$ را می تواند بازیابی کند ولی او قادر به بازیابی رشته بیت $z_{n+1}, z_{n+2}, \dots, z_{n+m}$ که برای رمزنگاری $MAK_{k_2}(x)$ استفاده می شود نیست. حتی اگر کل $bit-stream$ را بداند باز هم به دلیل اینکه مقدار k_2 را نمی داند قادر به محاسبه $MAK_{k_2}(x')$ نیست.

Chapter 13

4.

People at your new job are deeply impressed that you worked through this book. As the first job assignment you are asked to design a digital $pay - TV$ system which uses encryption to prevent service theft through wiretapping. As key exchange protocol, a strong $Diffie\text{–}Hellman$ with, e.g., 2048 –bit modulus is being used.
However, since your company wants to use cheap legacy hardware, only DES is available for data encryption algorithm. You decide to use the following key derivation approach:

$$K^i = f\,(K_{AB} \parallel i)$$

Where f is an irreversible function.

4.1. First we have to determine whether the attacker can store an entire movie with reasonable effort (in particular, cost). Assume the data rate for the TV link is $1\ Mbit/s$, and that the longest movies we want to protect are 2 hours long. How many Gbytes (where $1M = 106$ and $1G = 109$) of data must be stored for a 2 hour film (don't mix up bit and byte here)? Is this realistic?

4.2. We assume that an attacker will be able to find a DES key in 10 minutes using a $brute-force$ attack. Note that this is a somewhat optimistic assumption from an attacker's point of view, but we want to provide some medium-term security by assuming increasingly faster key searches in the future.

How frequently must a key be derived if the goal is to prevent an offline decryption of a $2-hour$ movie in less than 30 days?

حل:

4.1.

$$t = rate = 1Mbit/s = 10^6\ bit/sec$$
$$r = hours = 2h = 2 \times 60 \times 60\ sec$$

داریم:

$$storage = t \times r = 10^6 \times 2h\ \frac{bits}{sec} = 2 \times 3600\ (sec) \times 10^6 \left(\frac{bit}{sec}\right) = 7.2\ Gbit$$
$$= \frac{7.2}{8}\ G\ Byte = 0.9\ GBytes$$

بنابراین میزان داده کمتر از $1\ GByte$ می تواند با هزینه های متوسط بر روی دیسک ذخیره شود.

4.2.

محاسبه تعداد کلیدهایی که یک مهاجم می تواند در مدت 30 روز بازیابی کند به شکل زیر است:

$$\text{تعداد کلیدها} = \#Keys = \frac{30\ days}{10\ min} = \frac{30.24.60\ \ min}{10\ min} = 4320$$

مدت زمان استخراج کلید به صورت زیر است:

با توجه به اینکه تعداد کلیدهایی که در 30 روز به دست می آید برابر 4320 است و فیلم $2h$ است. پس داریم:

$$T_{Kder} = \frac{2h}{4320} = 1.67\ sec$$

به دلیل اینکه $hash\ function$ ها سریع هستند، فرایند استخراج کلید می تواند به راحتی در چنین ریتی صورت پذیرد.