T-79.5501 Cryptology Spring 2009

Homework 2

Tutor : Joo Y. Cho
joo.cho@tkk.fi

5th February 2009

Q1. Let us consider a cryptosystem where $\mathcal{P} = \{a, b, c\}$ and $\mathcal{C} = \{1, 2, 3, 4\}$, $\mathcal{K} = \{K_1, K_2, K_3\}$, and the encryption mappings e_K are defined as follows:

K	$e_K(a)$	$e_K(b)$	$e_K(c)$
K_1	1	2	3
K_2	2	3	4
K_3	3	4	1

The keys are chosen equiprobably, and the plaintext probability distribution is **Pr**[a] = 1/2, **Pr**[b] = 1/3, **Pr**[c] = 1/6. Compute H(**P**), H(**C**), H(**K**), H(**K**|**C**) and H(**P**|**C**).

$$H(\mathbf{K}|\mathbf{C}) = H(\mathbf{K}) + H(\mathbf{P}) - H(\mathbf{C})$$

$$H(\mathbf{P}|\mathbf{C}) = H(\mathbf{P}, \mathbf{C}) - H(\mathbf{C})$$

A1. Since $\mathbf{Pr}[a] = 1/2$, $\mathbf{Pr}[b] = 1/3$, $\mathbf{Pr}[c] = 1/6$.

$$H(\mathbf{P}) = \frac{1}{2}\log_2 2 + \frac{1}{3}\log_2 3 + \frac{1}{6}\log_2 6 = \frac{2}{3} + \frac{1}{2}\log_2 3 \approx 1.459.$$

In HW-1, the probability distribution of $\mathbf{C}$ was calculated as

$$\mathbf{Pr}[y = 1] = \frac{2}{9} \quad \mathbf{Pr}[y = 2] = \frac{5}{18} \quad \mathbf{Pr}[y = 3] = \frac{1}{3} \quad \mathbf{Pr}[y = 4] = \frac{1}{6}$$

thus, the entropy of the ciphertext is

$$H(\mathbf{C}) = -\frac{2}{9}\log_2\frac{2}{9} - \frac{5}{18}\log_2\frac{5}{18} - \frac{1}{3}\log_2\frac{1}{3} - \frac{1}{6}\log_2\frac{1}{6} \approx 1.955.$$

The keys are chosen equiprobably,

$$H(\mathbf{K}) = \frac{1}{3}\log_2 3 + \frac{1}{3}\log_2 3 + \frac{1}{3}\log_2 3 \approx 1.585.$$

By Theorem 2.10 in Stinson, we get

$$H(\mathbf{K}|\mathbf{C}) = H(\mathbf{K}) + H(\mathbf{P}) - H(\mathbf{C}) \approx 1.089.$$

In HW-1, we calculated

$$\mathbf{Pr}[x = a, y] = \frac{1}{6}, \text{ for } y = 1, 2, 3 \quad \mathbf{Pr}[x = b, y] = \frac{1}{9}, \text{ for } y = 2, 3, 4$$

$$\mathbf{Pr}[x = c, y] = \frac{1}{18}, \text{ for } y = 1, 3, 4$$

The three remaining probabilities were zero. Then,

$$H(\mathbf{P}, \mathbf{C}) = 3 \times [\frac{1}{6} \log_2 6 + \frac{1}{9} \log_2 9 + \frac{1}{18} \log_2 18] \approx 3.044,$$

Hence,

$$H(\mathbf{P}|\mathbf{C}) = H(\mathbf{P}, \mathbf{C}) - H(\mathbf{C}) \approx 1.089.$$

Remark. Since the plaintext and the ciphertext determine the key uniquely, $H(\mathbf{K}|\mathbf{P}, \mathbf{C}) = H(\mathbf{K}, \mathbf{P}, \mathbf{C}) - H(\mathbf{P}, \mathbf{C}) = 0$. Hence,

$$\begin{aligned} H(\mathbf{P}|\mathbf{C}) &= H(\mathbf{P}, \mathbf{C}) - H(\mathbf{C}) = H(\mathbf{K}, \mathbf{P}, \mathbf{C}) - H(\mathbf{C}) \\ &= H(\mathbf{K}) + H(\mathbf{P}) - H(\mathbf{C}) \approx 1.089. \end{aligned}$$

Q2. Prove that, in any cryptosystem

a) $H(\mathbf{C}|\mathbf{K}) = H(\mathbf{P})$, and

b) $H(\mathbf{P}|\mathbf{C}) \leq H(\mathbf{K}|\mathbf{C})$.

Intuitively, the second result means that given the ciphertext, uncertainty of the plaintext is at most the same as uncertainty about the key.

a) Following the idea of Proof of Thm 2.10 (Stinson), we get

$$\begin{aligned} H(\mathbf{C}|\mathbf{K}) &= H(\mathbf{K},\mathbf{C}) - H(\mathbf{K}) \\ &= H(\mathbf{K},\mathbf{P},\mathbf{C}) - H(\mathbf{K}) = H(\mathbf{K},\mathbf{P}) - H(\mathbf{K}) = H(\mathbf{P}). \end{aligned}$$

b) By Proof of Theorem 2.10, we know that $H(\mathbf{K},\mathbf{P},\mathbf{C}) = \mathbf{H}(\mathbf{K},\mathbf{P}) = \mathbf{H}(\mathbf{K}) + \mathbf{H}(\mathbf{P})$. By Theorem 2.8, $H(\mathbf{C}) + \mathbf{H}(\mathbf{P}|\mathbf{C}) = \mathbf{H}(\mathbf{P},\mathbf{C})$ and $H(\mathbf{K}|\mathbf{P},\mathbf{C}) = \mathbf{H}(\mathbf{K},\mathbf{P},\mathbf{C}) - \mathbf{H}(\mathbf{P},\mathbf{C})$. Since the entropy is always nonnegative, we get, by Theorem 2.10,

$$\begin{aligned} H(\mathbf{K}|\mathbf{C}) - \mathbf{H}(\mathbf{P}|\mathbf{C}) &= H(\mathbf{K}) + \mathbf{H}(\mathbf{P}) - \mathbf{H}(\mathbf{C}) - \mathbf{H}(\mathbf{P}|\mathbf{C}) \\ &= H(\mathbf{K},\mathbf{P},\mathbf{C}) - \mathbf{H}(\mathbf{P},\mathbf{C}) = \mathbf{H}(\mathbf{K}|\mathbf{P},\mathbf{C}) \geq \mathbf{0} \end{aligned}$$

Q3. A PIN code for a smart card is a number of four decimal digits (p_1, p_2, p_3, p_4), where each p_i, $i = 1, 2, 3, 4$, is derived from a uniformly distributed random string of 16 bits $(r_1, r_2, ..., r_{16})$ by computing

$$p_i = (r_{4i-3} + r_{4i-2} \cdot 2 + r_{4i-1} \cdot 2^2 + r_{4i} \cdot 2^3) \bmod 10.$$

Determine the entropy of the PIN code. Compare it with the maximum entropy of a string of four decimal digits.

A3. Each p_i, $i = 1, 2, 3, 4$ is chosen independently. Thus, $H(\mathbf{P}) = 4H(\mathbf{P}_i)$. Since p_i is computed by mapping a 4-bit random string into a decimal digit, we know that
$Pr[p_i = 0] = \cdots = Pr[p_i = 5] = \frac{2}{16}$ and
$Pr[p_i = 6] = \cdots = Pr[p_i = 9] = \frac{1}{16}$.
Hence, we get

$$H(\mathbf{P}) = 4H(\mathbf{P}_i) = 4\left(-6\left(\frac{2}{16}\log_2\frac{2}{16}\right) - 4\left(\frac{1}{16}\log_2\frac{1}{16}\right)\right) = 13.$$

The maximum entropy of a four decimal digit number equals $\log_2 10^4 \approx 13.29$.

Q4. (Stinson 2.4) Prove that if a cryptosystem has perfect secrecy and $|\mathcal{K}| = |\mathcal{C}| = |\mathcal{P}|$, then every ciphertext has equal probability.

Theorem 2.4 (Stinson)

Suppose $|\mathcal{P}| = |\mathcal{C}| = |\mathcal{K}|$. The cryptosystem provides perfect secrecy iff every key is used with $1/|\mathcal{K}|$ and for every $x \in \mathcal{P}$ and every $y \in \mathcal{C}$, there is a unique key K such that $e_K(x) = y$.

A4. In Theorem 2.4 (Stinson) it is proved that under the given assumptions every key has equal probability, and for each x and y there is a unique key $K_{x,y}$ such that $e_{K_{x,y}}(x) = y$. It follows that

$$\begin{aligned} p_{\mathcal{C}}(y) &= \sum_{K:y\in C(K)} p_{\mathcal{K}}(K)p_{\mathcal{P}}(d_K(y)) \\ &= \sum_{x\in\mathcal{P}} p_{\mathcal{K}}(K_{x,y})p_{\mathcal{P}}(x) \\ &= \frac{1}{|\mathcal{K}|}\sum_{x\in\mathcal{P}} p_{\mathcal{P}}(x) = \frac{1}{|\mathcal{K}|}. \end{aligned}$$

This proves the claim.

Q5. The keystream z_i, $i = 1, 2, \ldots$ of a binary stream cipher is generated by repeating a finite random sequence $K = (k_1, \ldots, k_m)$ of m bits, which is the key. Hence $z_i = k_i$, for $i = 1, 2, \ldots, m$, and $z_{i+m} = z_i$, for all $i = 1, 2, \ldots$.

a) This stream cipher is used to encrypt plaintext with redundancy R_L. Give an estimate for the unicity distance.

b) Suppose that $m = 5$ and the plaintext bit string is formed by repeating the following procedure (a finite number of times): two bits are generated at random, and a third bit is computed as an xor sum of these two bits. The first fifteen bits of the ciphertext are: 0 1 0 1 0 1 1 1 1 1 0 0 0 0 1. Attempt to find the key $K = (k_1, k_2, k_3, k_4, k_5)$.

A5-a).
Since the key is a randomly generated string of m bits, and the plaintext character is one bit, we know $|\mathcal{K}| = 2^m$ and $|\mathcal{P}| = 2$. Then, the unicity distance n_0 can be estimated by the formula

$$n_0 \approx \frac{\log_2 |\mathcal{K}|}{R_L \log_2 |\mathcal{P}|} = \frac{\log_2 |2^m|}{R_L \log_2 2} = \frac{m}{R_L}.$$

A5-b).

P	x_1	x_2	$x_3 = x_1 \oplus x_2$	x_4	x_5	$x_6 = x_3 \oplus x_4$	x_7	x_8	$x_9 = x_7 \oplus x_8$	$\cdots$
K	k_1	k_2	k_3	k_4	k_5	k_1	k_2	k_3	k_4	$\cdots$
C	0	1	0	1	0	1	1	1	1	$\cdots$

For each third plaintext bit:

$$\begin{cases} x_3 &= x_1 + x_2 = k_1 + k_2 + 1 \\ x_3 &= k_3 \end{cases}$$

$$\begin{cases} x_6 &= x_4 + x_5 = k_4 + k_5 + 1 \\ x_6 &= k_1 + 1 \end{cases}$$

$$\begin{cases} x_9 &= x_7 + x_8 = k_2 + k_3 \\ x_9 &= k_4 + 1 \end{cases}$$

$$\begin{cases} x_{12} &= x_{10} + x_{11} = k_5 + k_1 + 1 \\ x_{12} &= k_2 \end{cases}$$

$$\begin{cases} x_{15} &= x_{13} + x_{14} = k_3 + k_4 \\ x_{15} &= k_5 + 1 \end{cases}$$

The resulting system of five equations and five unknown key bits

$$
\begin{array}{ccccccccccc}
k_1 & + & k_2 & + & k_3 & & & & & = & 1 \\
k_1 & & & & & + & k_4 & + & k_5 & = & 0 \\
 & & k_2 & + & k_3 & + & k_4 & & & = & 1 \\
k_1 & + & k_2 & & & & & + & k_5 & = & 1 \\
 & & & & k_3 & + & k_4 & + & k_5 & = & 1
\end{array}
$$

By solving above equations, we find that the key is $K = (1, 0, 0, 1, 0)$.

Note that this result is in accordance with the theoretical estimate given in Q5-a.

Q6. The DES keys are 64 bits long, where each eighth bit is a parity bit computed as a modulo 2 sum of the preceding seven bits. A key management center uses AES encryption algorithm and a 128-bit "master" key to encrypt DES keys to end users. Each ciphertext block consists of two encrypted DES keys. Using the concept of unicity distance give an estimate of the number of encrypted DES keys that an attacker needs to have to be able to determine the master key uniquely given enough computing time.

A6. In the AES encryption algorithm

$$\mathcal{P} = \{0,1\}^{128}, \mathcal{C} = \{0,1\}^{128}, \mathcal{K} = \{0,1\}^{128}$$

Each 128-bit ciphertext consists of two encrypted 64-bit DES keys. Data(L) to be encrypted; DES keys, 64 bits each, each eighth bit is a parity check bit.

$$H_L = 64 - 8 = 56; R_L = 1 - \frac{H_L}{\log_2 |L|} = 1 - \frac{56}{64} = \frac{1}{8}.$$

Then, unicity distance can now be estimated as:

$$n_0 \approx \frac{\log_2 |\mathcal{K}|}{R_L \cdot \log_2 |\mathcal{P}|} = \frac{128}{\frac{1}{8} \cdot 128} = 8.$$

Since each ciphertext block of AES contains two DES keys, an attacker needs 16 encrypted DES keys to determine the master key uniquely.