T-79.5501 Cryptology Spring 2009

Homework 1

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

29th January 2009

Before we start

- Prepare solutions before tutorials and be on time.
- Tick the problems (Q1,Q2,...,Q6) that you wish to present in the list.
- Even though you could not solve the problems completely, you can still present your idea for the solutions (and tick the problems, too).
- Tutor will pick your name from the list and ask him/her to present the solutions, if there is no volunteer.
- Solutions will be uploaded around Monday (after tutorial).
- Tutor will not mark your presentation.

Q1. Kasiski's method and the method of Index of Coincidence are efficient methods for breaking the Vigenère cipher. The purpose of this exercise is to prove that any such method can be used to break the autokey cipher. Show how the ciphertext produced by an autokey cipher can be transformed to a ciphertext produced by a Vigenère cipher. What about the converse?

Autokey Cipher

Let $\mathcal{P} = \mathcal{C} = \mathcal{K} = \mathbf{Z}_{26}$. Let $z_1 = K$ and define $z_i = x_{i-1}$ for all $i \geq 2$. For $0 \leq z \leq 25$, define

$$e_z(x) = (x + z) \bmod 26$$

and

$$d_z(y) = (y - z) \bmod 26$$

$(x, z \in \mathbf{Z}_{26})$.

A1. Autokey cipher is a stream cipher with $\mathcal{P} = \mathcal{C} = \mathcal{K} = \mathbf{Z}_{26}$ and

$$\begin{aligned} z_1 &= K \\ z_i &= x_{i-1}, \text{ for } i \geq 2 \\ e_z(x) &= x + z \bmod 26 \end{aligned}$$

Given a plaintext sequence $x_1, x_2, \ldots, x_n$ the ciphertext sequence is

$$\begin{aligned} y_1 &= (x_1 + K) \bmod 26 \\ y_2 &= (x_2 + x_1) \bmod 26 = (x_2 + y_1 - K) \bmod 26 \\ &\Rightarrow \underbrace{y_2 - y_1}_{y_2'} \equiv x_2 - K \bmod 26 \\ y_3 &= (x_3 + x_2) \bmod 26 = (x_3 + y_2 - y_1 + K) \bmod 26 \\ &\Rightarrow \underbrace{y_3 - (y_2 - y_1)}_{y_3 - y_2'} \equiv x_3 + K \bmod 26 \end{aligned}$$

and so on.

We define a transformed sequence $y'_1, y'_2, \ldots, y'_n$ as follows:

$$\begin{aligned} y'_1 &= y_1 \\ y'_i &= (y_i - y'_{i-1}) \bmod 26, \text{ for } i \geq 2 \end{aligned}$$

Then

$$y'_i = (x_i + (-1)^{i-1}K) \bmod 26, \text{ for all } i \geq 1, \tag{1}$$

which is the ciphertext obtained by encrypting the plaintext $\{x_1, x_2, \ldots x_n\}$ using Vigenère cipher with $\mathcal{P} = \mathcal{C} = \mathbf{Z}_{26}$ and $\mathcal{K} = \{(K, -K) \,|\, K \in \mathbf{Z}_{26}\}$.

For $i \geq 2$, a ciphertext produced by a Vigenère cipher is

$$\begin{aligned} y_i &= x_i + k \bmod 26 \\ y_{i+1} &= x_{i+1} + k \bmod 26 \\ \Rightarrow y_{i+1} - y_i &= x_{i+1} - x_i \bmod 26. \end{aligned} \tag{2}$$

However, for a ciphertext produced by an autokey cipher, we have

$$\begin{aligned} y_i &= x_i + x_{i-1} \bmod 26 \\ y_{i+1} &= x_{i+1} + x_i \bmod 26 \\ \Rightarrow y_{i+1} - y_i &= x_{i+1} - x_{i-1} \bmod 26, \end{aligned} \tag{3}$$

Since (3) is dependent on x_{i-1} and (2) is independent of x_{i-1}, Hence, transforming a ciphertext produced by a Vigenère cipher to a ciphertext produced by an autokey cipher is not generally possible.

Q2. Define a stream cipher as follows:

$$
\begin{aligned}
\mathcal{P} &= \mathcal{C} = \mathbf{Z}_7,\ \mathcal{K} = \{(a,b) \mid \gcd(a,7) = 1\} \\
\text{plaintext} &= x_1, x_2, x_3, \ldots \\
z_i &= (a \times i + b) \bmod 7,\ \ i = 1, 2, \ldots, \text{ where } (a,b) \text{ is the key.} \\
e_{z_i}(x_i) &= (x_i + z_i) \bmod 7 \\
\text{ciphertext} &= y_1, y_2, y_3, \ldots \\
d_{z_i}(y_i) &= (y_i - z_i) \bmod 7
\end{aligned}
$$

a) Using (5,3) as the key, compute the decryption of the message 25542531.

b) If you know that some part of the plaintext is 110503, and this encrypts to give the ciphertext 501153, then derive as much as you can about the unknown key (a, b). What additional information you need to derive the entire key?

A2-a). Let x_i and y_i be the plaintext and ciphertext sequences. Then

$$\begin{aligned} y_i &\equiv x_i + a \cdot i + b \mod 7, \\ x_i &\equiv y_i - a \cdot i - b \mod 7 \end{aligned}$$

Since $(a, b) = (5, 3)$, we know $(-a, -b) = (-5, -3) = (2, 4) \mod 7$. Hence, the decryption of the ciphertext message 25542531 is

$$\begin{aligned} x_1 &= 2 + 2 \cdot 1 + 4 = 1 \\ x_2 &= 5 + 2 \cdot 2 + 4 = 6 \\ x_3 &= 5 + 2 \cdot 3 + 4 = 1 \\ x_4 &= 4 + 2 \cdot 4 + 4 = 2 \\ x_5 &= 2 + 2 \cdot 5 + 4 = 2 \\ x_6 &= 5 + 2 \cdot 6 + 4 = 0 \\ x_7 &= 3 + 2 \cdot 7 + 4 = 0 \\ x_8 &= 1 + 2 \cdot 8 + 4 = 0 \end{aligned}$$

A2-b). Using the known plaintext ciphertext pair, we get

$$
\begin{aligned}
5 &= 1 + a \cdot i + b && (4)\\
0 &= 1 + a \cdot (i+1) + b && (5)\\
1 &= 0 + a \cdot (i+2) + b && (6)\\
1 &= 5 + a \cdot (i+3) + b && (7)\\
5 &= 0 + a \cdot (i+4) + b && (8)\\
3 &= 3 + a \cdot (i+5) + b. && (9)
\end{aligned}
$$

The initial contents i of the index counter is not known. From $(5) - (4)$, one gets $a = -5 = 2 \bmod 7$. Then, from (6), one gets $2i + b \equiv 4 \bmod 7$. Hence b can be found if and only if i is known.

Q3. The plaintext and ciphertext alphabet consists of the 26 letters A-Z and the space between words. Each plaintext letter x is encrypted separately using a randomised substitution as follows. The key $K = (k_0, k_1, \ldots, k_9)$ is a permutation of the ten digits $\{0, 1, \ldots, 9\}$. The encryption process has the following steps.

1. Pick a character y from the plaintext alphabet at random. Interpret the pair (y, x) as the representation of an integer I to the base 27, that is, $I = 27 \cdot y + x$. Let $a_2,\ a_1,\ a_0$ be the digits of I in the decimal system, where a_2 is the most significant digit.
2. Use the key K to substitute a_i by k_{a_i}, $i = 0,\ 1,\ 2$.
3. The ciphertext $(c_2,\ c_1,\ c_0)$ is obtained as the 27-base representation of the integer $100 \cdot k_{a_2} + 10 \cdot k_{a_1} + k_{a_0}$.

An attacker is observing plaintext-ciphertext pairs produced by this encryption system with the same fixed key. An encryption of the character ‘space’ is ‘ABX’ and an encryption for character ‘B’ is ‘ACB’. Based on this information, derive a and b such that $k_a = 0$ and $k_b = 5$.

A3.

Note that $\{A, B, \cdots, Z, space\} \mapsto \{0, 1, \cdots, 25, 26\}$.

$$\begin{aligned} \mathsf{ABX} &= 0 \cdot 27^2 + 1 \cdot 27 + 23 = 050 = k_a k_b k_a, \\ \mathsf{ACB} &= 0 \cdot 27^2 + 2 \cdot 27 + 1 = 055 = k_a k_b k_b \end{aligned}$$

Then, the plaintext and ciphertext of the "space" and B have the following relations:

$$\begin{aligned} k_a k_b k_a \mapsto 100a + 10b + a &= 27y_1 + 26 \leftarrow \text{“space”} \\ k_a k_b k_b \mapsto 100a + 10b + b &= 27y_2 + 1 \leftarrow \mathsf{B} \end{aligned}$$

Hence, we get

$$\begin{aligned} 101a + 10b &\equiv 26 \bmod 27 \\ 100a + 11b &\equiv 1 \bmod 27, \end{aligned}$$

which has the following unique solution $a = 2$ and $b = 4$.

Q4. 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

Given that keys are chosen equiprobably, and the plaintext probability distribution is $\mathbf{Pr}[a] = 1/2$, $\mathbf{Pr}[b] = 1/3$, $\mathbf{Pr}[c] = 1/6$, compute the following probabilities

a) $\mathbf{Pr}[\mathbf{y} = i]$, $i = 1, 2, 3, 4$.

b) $\mathbf{Pr}[\mathbf{x} = j, \mathbf{y} = i]$, $j = a, b, c$, and $i = 1, 2, 3, 4$..

A4-a). Note that
$\mathbf{Pr}(\mathbf{x} = a) = 1/2,\ \mathbf{Pr}(\mathbf{x} = b) = 1/3,\ \mathbf{Pr}(\mathbf{x} = c) = 1/6$ and
$\mathbf{Pr}(\mathbf{K} = K_i) = 1/3,\ i = 1, 2, 3.$
Let us calculate $\mathbf{Pr}(\mathbf{y} = 1)$ first.

$$\begin{aligned}\mathbf{Pr}(\mathbf{y} = 1) &= \mathbf{Pr}(\mathbf{K} = K_1)\,\mathbf{Pr}(\mathbf{x} = a) + \mathbf{Pr}(\mathbf{K} = K_2) \cdot 0 + \\ &\quad \mathbf{Pr}(\mathbf{K} = K_3)\,\mathbf{Pr}(\mathbf{x} = c) \\ &= \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{6} = \frac{2}{9}\end{aligned}$$

The other probabilities are calculated similarly to be

$$\mathbf{Pr}(\mathbf{y} = 2) = \frac{5}{18} \quad \mathbf{Pr}(\mathbf{y} = 3) = \frac{1}{3} \quad \mathbf{Pr}(\mathbf{y} = 4) = \frac{1}{6}.$$

A4-b). Note that
$\mathbf{Pr}(\mathbf{x} = a) = 1/2,\ \mathbf{Pr}(\mathbf{x} = b) = 1/3,\ \mathbf{Pr}(\mathbf{x} = c) = 1/6$ and
$\mathbf{Pr}(\mathbf{K} = K_i) = 1/3,\ i = 1, 2, 3.$
Let us calculate $\mathbf{Pr}(\mathbf{x} = a, \mathbf{y} = 1)$ as an example:

$$\mathbf{Pr}(\mathbf{x} = a, \mathbf{y} = 1) = \mathbf{Pr}(\mathbf{x} = a)\,\mathbf{Pr}(\mathbf{y} = 1|\mathbf{x} = a) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$$

The other nonzero probabilities are

$$\begin{aligned}
\mathbf{Pr}(\mathbf{x} = a, \mathbf{y} = i) &= \frac{1}{6}, i = 2, 3, \\
\mathbf{Pr}(\mathbf{x} = b, \mathbf{y} = i) &= \frac{1}{9}, i = 2, 3, 4 \\
\mathbf{Pr}(\mathbf{x} = c, \mathbf{y} = i) &= \frac{1}{18}, i = 1, 3, 4.
\end{aligned}$$

The other probabilities are zero.

Q5. Does the cryptosystem of the preceding problem achieve perfect secrecy?

A5. A cryptosystem provides perfect secrecy if and only if $\mathbf{Pr}(\mathbf{x} = j, \mathbf{y} = i) = \mathbf{Pr}(\mathbf{x} = j)\,\mathbf{Pr}(\mathbf{y} = i)$ for all plaintext-ciphertext pairs $(\mathbf{i}, \mathbf{j})$. Since

$$\mathbf{Pr}(\mathbf{x} = a)\,\mathbf{Pr}(\mathbf{y} = 2) = \frac{1}{2} \cdot \frac{5}{18} = \frac{5}{36} \neq \frac{1}{6} = \mathbf{Pr}(\mathbf{x} = a, \mathbf{y} = 2),$$

the cryptosystem in Problem 4 does not provide perfect secrecy.

Q6. Plaintext is composed of independently generated bits that are arranged in blocks of four bits. The probability that a plaintext bit equals 0 is p. Each block x_1, x_2, x_3, x_4 is encrypted using one key bit z by adding it modulo 2 to each plaintext bit. Hence the ciphertext block is y_1, y_2, y_3, y_4 where $y_i = x_i \oplus z$, $i = 1, 2, 3, 4$. It is assumed that every key bit is generated uniformly at random. Let us assume that a ciphertext block has k zeroes and $4 - k$ ones, $k = 0, 1, 2, 3, 4$.

a) Compute the probability (as a function of k) that the encryption key was $z = 0$.

b) What value of k maximizes this probability?

c) For which value of k the probability that $z = 0$ is equal to $\frac{1}{2}$, that is, the ciphertext does not give any information at all about the used key bit?

A6-a). Let A_k be the event that a plaintext block has exactly k zeroes. Let B_k be the event that the ciphertext has k zeroes, $k = 0, 1, 2, 3, 4$. Then using the definition of conditional probability

$$Pr(z = 0 \,|\, B_k) \quad = \quad \frac{Pr(z = 0,\, B_k)}{Pr(B_k)}$$

Since $Pr(z = 0) = Pr(z = 1) = \frac{1}{2}$ and the key is independent of the plain text,

$$\begin{aligned} Pr(z = 0,\, B_k) &= Pr(z = 0)Pr(A_k) = \frac{1}{2}\binom{4}{k}p^k(1-p)^{4-k} \\ Pr(B_k) &= Pr(z = 0, A_k) + Pr(z = 1, A_{4-k}) \\ &= Pr(z = 0)Pr(A_k) + Pr(z = 1)Pr(A_{4-k})) \\ &= \frac{1}{2}\binom{4}{k}p^k(1-p)^{4-k} + \frac{1}{2}\binom{4}{4-k}p^{4-k}(1-p)^k \end{aligned}$$

Note $\binom{4}{k} = \binom{4}{4-k}$.

Hence, we get

$$\begin{aligned} Pr(z=0 \mid B_k) &= \frac{\frac{1}{2}\binom{4}{k}p^k(1-p)^{4-k}}{\frac{1}{2}\binom{4}{k}p^k(1-p)^{4-k} + \frac{1}{2}\binom{4}{4-k}p^{4-k}(1-p)^k} \\ &= \frac{1}{1+\left(\frac{p}{1-p}\right)^{4-2k}}, \end{aligned}$$

where $p \neq 1$.

A6-b).

- If $p < \frac{1}{2}$, that is $p < 1 - p$, then $Pr(z = 0 \mid B_k)$ decreases with k and is maximized with $k = 0$.
- If $p > \frac{1}{2}$, that is $p > 1 - p$, then $Pr(z = 0 \mid B_k)$ increases with k and is maximized with $k = 4$.
- If $p = \frac{1}{2}$, that is $p = 1 - p$, then $Pr(z = 0 \mid B_k)$ equals $\frac{1}{2}$, for all $k = 0, 1, 2, 3, 4$. In this case we do not get any information about the key by counting the number of zeroes and ones in the cipher text.

A6-c).
If $k = 2$, then $Pr(z = 0 \mid B_2) = \frac{1}{2}$ if $p \neq 0, 1$. Hence, the cipher text does not give any information about the key if $k = 2$.