CS 435: Introduction to Cryptography **Fall 2016-17**

Homework 1 Solutions

Professor Somesh Jha

1. Consider a variant of the Vigenère cipher where instead of a word or short phrase, the key instead consists of a book or some other English-language text that is much longer than the message to be encrypted. Using this cipher, the key is never repeated, so our standard methods for retrieving the key length will fail. Now assume that Bob is a sleeper agent and Alice is his handler. Alice, using this cipher, has sent Bob a ciphertext that reads

   `zsfgsjvtpsnzycrvc`

   The plaintext is known to contain the day of the week that Bob is supposed to receive the dead drop, followed by the day of the week he is supposed to flee the country. Determine which plaintext Alice sent to Bob, and explain how you reached your answer.
   Note. Each ciphertext character c_i is equal to $m_i + k_i \pmod{26}$, where m_i is the i-th character of the plaintext message and k_i is the i-th character of the key. In particular, the alphabet is indexed from 0, so 'a' corresponds to 0, 'b' corresponds to 1, and so on.

   Solution: Note that the ciphertext has 17 characters. In the Vigenère cipher the number of characters in the plaintext is the same as the number of characters in the ciphertext, which tells us the plaintext must also have 17 characters. The plaintext is made up of two days of the week which in total have 17 letters. Monday, Friday and Sunday each have 6 letters, Tuesday has 7, Wednesday has 9, while Thursday and Saturday have 8 letters each. Therefore, one of the days in the plaintext must be Wednesday the other is either Thursday or Saturday.

   To find the correct plaintext, we use the fact that the key is an English language book. For each candidate plaintext we compute the corresponding key as follows: Given the i-th characters of the candidate plaintext (m_i) and the ciphertext (c_i), the i-th character of the key k_i is $c_i - m_i \pmod{26}$. Doing so for each candidate plaintext reveals that the only plaintext for which the corresponding key could appear in an English language book is `wednesdaysaturday`, for which the key is `doctorstrangelove`.

2. a. Assume an attacker knows that a user's password is either `abcd` or `bedg`. Say the user encrypts his password using the shift cipher, and the attacker sees the resulting ciphertext. Show how the attacker can determine the user's password, or explain why this is not possible.

   b. Repeat part (a) for the Vigenère cipher using period 2, using period 3, and using period 4.

Solution: The attacker can determine the user's password when the cipher used is either the shift cipher in part (a) or the Vigenère cipher with period 3. For both these ciphers the first and fourth letters of the user's password are shifted by the same amount. Let the ciphertext the attacker sees be $\mathsf{c}_1\mathsf{c}_2\mathsf{c}_3\mathsf{c}_4$. The attacker can identify the password used by computing $\mathsf{c}_4 - \mathsf{c}_1(\mod 26)$. If the password used is abcd, $\mathsf{c}_4 - \mathsf{c}_1(\mod 26) = 3$. On the other hand if the password used is bedg, $\mathsf{c}_4 - \mathsf{c}_1(\mod 26) = 5$.

To argue that the attacker can't determine the user's password for the Vigenère cipher with period 2 or 4, we show that in each case there are keys k_0 and k_1 such that $\mathsf{Enc}_{k_0}(\mathsf{abcd}) = \mathsf{Enc}_{k_1}(\mathsf{bedg})$. This shows that the attacker cannot identify the user's password from just the ciphertext. For the Vigenère cipher with period 2, take k_0 to be bd and k_1 to be aa. For the Vigenère cipher with period 4, take k_0 to be bdbd and k_1 to be aaaa.

3. Prove that if an encryption scheme $\Pi = (\mathsf{Gen},\mathsf{Enc},\mathsf{Dec})$ is perfectly secret then it is perfectly indistinguishable.

   Solution : In the adversarial indistinguishability experiment, the adversary succeeds if the algorithm $\mathcal{A}$ correctly outputs the bit b. Let $\mathcal{A}(c)$ denote the random variable that represents the output of the algorithm $\mathcal{A}$ on ciphertext c. The probability that the adversary succeeds is:

$$\Pr[\mathsf{PrivK}^{\mathrm{eav}}_{\mathcal{A},\Pi} = 1] = \Pr[\mathcal{A}(\mathsf{Enc}_K(m_b)) = 1 \wedge b = 1] + \Pr[\mathcal{A}(\mathsf{Enc}_K(m_b)) = 0 \wedge b = 0].$$

   Recall that for a pair of events E_1, E_2, $\Pr[E_1 \wedge E_2] = \Pr[E_1|E_2]\Pr[E_2]$. It follows that

$$\begin{aligned}
&\Pr[\mathcal{A}(\mathsf{Enc}_K(m_b)) = 1 \wedge b = 1] + \Pr[\mathcal{A}(\mathsf{Enc}_K(m_b)) = 0 \wedge b = 0] \\
&= \Pr[\mathcal{A}(\mathsf{Enc}_K(m_b)) = 1|b = 1]\Pr[b = 1] + \Pr[\mathcal{A}(\mathsf{Enc}_K(m_b)) = 0|b = 1]\Pr[b = 0] \\
&= 1/2(\Pr[\mathcal{A}(\mathsf{Enc}_K(m_b)) = 1|b = 1] + \Pr[\mathcal{A}(\mathsf{Enc}_K(m_b)) = 0|b = 0]) \\
&= 1/2(\Pr[\mathcal{A}(\mathsf{Enc}_K(m_1)) = 1] + \Pr[\mathcal{A}(\mathsf{Enc}_K(m_0)) = 0]) \\
&= 1/2(\Pr[\mathcal{A}(\mathsf{Enc}_K(m_1)) = 1] + 1 - \Pr[\mathcal{A}(\mathsf{Enc}_K(m_0)) = 1]) \\
&= 1/2 + 1/2(\Pr[\mathcal{A}(\mathsf{Enc}_K(m_1)) = 1] - \Pr[\mathcal{A}(\mathsf{Enc}_K(m_0)) = 1])
\end{aligned}$$

   The second equality holds because b is a uniform bit, the third equality holds because we condition on the event b equals 1 in the first term and $b = 0$ in the second term. The fourth equality holds because $\Pr[\mathcal{A}(\mathsf{Enc}_K(m_0)) = 0]) + \Pr[\mathcal{A}(\mathsf{Enc}_K(m_0)) = 1] = 1$.

   Since $\Pi = (\mathsf{Gen},\mathsf{Enc},\mathsf{Dec})$ is perfectly secret, we have by Lemma 2.4 of the textbook that for every pair of messages m_0 and m_1 in M and for every ciphertext c

$$\Pr[\mathsf{Enc}_K(m_0) = c] = \Pr[\mathsf{Enc}_K(m_1) = c].$$

In other words, the random variables $\mathsf{Enc}_K(m_0)$ and $\mathsf{Enc}_K(m_1)$ have the same distribution. This immediately implies that the random variables $\mathcal{A}(\mathsf{Enc}_K(m_0))$ and $\mathcal{A}(\mathsf{Enc}_K(m_1))$ have the same distribution. Which tells us that

$$\Pr[\mathcal{A}(\mathsf{Enc}_K(m_1)) = 1] = \Pr[\mathcal{A}(\mathsf{Enc}_K(m_0)) = 1].$$

Therefore,

$$\begin{aligned}\Pr[\mathsf{PrivK}^{\mathrm{eav}}_{\mathcal{A},\Pi} = 1] & \\ &= 1/2 + 1/2(\Pr[\mathcal{A}(\mathsf{Enc}_K(m_1)) = 1] - \Pr[\mathcal{A}(\mathsf{Enc}_K(m_0)) = 1]) \\ &= 1/2\end{aligned}$$

4. For each of the following encryption schemes, state whether the scheme is perfectly secret. Justify your answer in each case.

   a. The message space $\mathcal{M}$ is $\{0, \ldots, 4\}$. Algorithm Gen chooses a uniform key k from the key space $\{0, \ldots, 5\}$. $\mathsf{Enc}_k(m) = [k + m \mod 5]$.

   b. The message space $\mathcal{M}$ is $\{0,1\}^{2n}$. Gen chooses an uniform key k from $\{0,1\}^n$. $\mathsf{Enc}_k(x) = \langle x_{1\ldots n} \oplus k, x_{n+1\ldots 2n} \oplus k\rangle$, where $\oplus$ denotes the bitwise XOR.

   Solution: Neither of the encryption schemes in this problem are perfectly secret.

   a A perfectly secret encryption scheme must satisfy the following: For every pair of messages m_0, m_1 and every ciphertext c

   $$\Pr[\mathsf{Enc}_K(m_0) = c] = \Pr[\mathsf{Enc}_K(m_1) = c].$$

   To see that the encryption scheme in part (a) is not perfectly secret, note that $\Pr[\mathsf{Enc}_K(0) = 0] = 2/6$, while $\Pr[\mathsf{Enc}_K(1) = 0] = 1/6$.

   b For part (b) one can invoke theorem 2.10 in the textbook which states that if an encryption scheme is perfectly secret then $|\mathcal{K}| \geq |\mathcal{M}|$. For encryption scheme in part (b), we have size of the key space $|\mathcal{K}| = 2^n$, while the size of the message space is $|\mathcal{M}| = 2^{2n}$.