CS 639/839: Applied Cryptography **Spring 2018-19**

Homework 1

Professor Somesh Jha **Due:** February 15

1. Use definition 2 of perfect secrecy (Lemma 2.4 of the textbook) to prove that the mono-alphabetic substitution cipher is not perfectly secret.

   $$\Pr[\mathsf{Enc}_K(m) = c] = \Pr[\mathsf{Enc}_K(m') = c] \tag{2.1}$$

   Lemma 2.4. *An encryption scheme* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *with message space* $\mathcal{M}$ *is perfectly secret if and only if Equation (2.1) holds for every* $m, m' \in \mathcal{M}$ *and every* $c \in \mathcal{C}$.

   Solution:

   Since definition 2 of perfect secrecy should hold true for every $m, m' \in \mathcal{M}$ and every $c \in \mathcal{C}$, if we are able to come up with m, m' and c such that Equation 2.1 does not hold true, then we are done.

   Let $m = xy$, $m' = zz$ and $c = aa$. In the mono-alphabetic cipher each letter is mapped to a unique letter in the key. Encrypting m to c would require a K which maps $x \to a$ and $y \to a$. This is not possible. Thus, $\Pr[\mathsf{Enc}_K(m) = c] = 0$ (1).

   All keys which map $z \to a$ would encrypt m' as c. Thus, $\Pr[\mathsf{Enc}_K(m') = c] > 0$ (2).

   By (1) and (2) we can say that the mono-alphabetic substitution cipher is not perfectly secret.

2. 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

   ```
   YVBCXGJRYHHRCJIUL
   ```

   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:

Notice that the length of the ciphertext is 17. Comparing the length of the days (Monday → 6, Tuesday → 7, Wednesday → 9, Thursday → 8, Friday → 6, Saturday → 8 and Sunday → 6), one can figure out that there are two possibilities for the plain text, either "WednesdayThursday" or "WednesdaySaturday".

Both the plaintexts have a common prefix of "Wednesday". Using the plaintext and the ciphertext we can back calculate the first part of the key which is "Cryptogra". This is suggestive of "Cryptography". One can now use this key to figure out that the plaintext is "WednesdaySaturday".

3. a. Assume an attacker knows that a user's password is one of these: pqrs, mnop, svux, jmlo. 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.

   Solution:

   Notice that the letters pqrs and mnop are all consecutive. This is not the case with svux, jmlo (These two also have a same pattern). Since the shift cipher moves all letters in the password by an equal amount, the resulting cipher text for pqrs, mnop will also be consecutive, and for svux, jmlo, the resulting cipher text will have the same pattern as the plaintext passwords. If the cipher text is composed of consecutive letters, the password is either pqrs or mnop and if not, then the password is either svux, jmlo. So at best the attacker can narrow the possibilities down to two passwords, but will not be able to determine the password.

   b. The attacker has somehow narrowed down the passwords to pqrs or jmlo but now the user encrypts his password by Vigenère cipher. Show how the attacker can determine the user's password, or explain why this is not possible. Do this using period 2, using period 3, and using period 4.

   Solution:

   Period 2

   A period 2 key implies that the 1st and 3rd characters of the password will be shifted by the same amount. Similarly, the 2nd and 4th characters will be shifted by the same amount.

   In pqrs the difference between the 1st and 3rd character is 2.

   The difference between the 2nd and 4th character is 2.

   In jmlo the difference between the 1st and 3rd character is 2.

   The difference between the 2nd and 4th character is 2.

   Since there is no difference which can be exploited, the attacker will not be able to tell whether the password was pqrs or jmlo.

   Period 3

   A period 3 key implies that the 1st and 4th character of the plain text will be shifted by the same amount.

   In pqrs the difference between the 1st and 4th character is 3.

   In jmlo the difference between the 1st and 4th character is 5.

If the difference between the 1st and 4th character of the ciphertext is 3, the attacker knows that the password is pqrs, else it is jmlo.

Period 4

With period 4, for any ciphertext c, there exists keys K and K' such that $\mathsf{Enc}_K(\mathsf{pqrs}) = c$ and $\mathsf{Enc}_{K'}(\mathsf{jmlo}) = c$. Thus the attacker will not be able to tell whether the password was pqrs or jmlo.

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 $\{0, \ldots, 5\}$. $\mathsf{Enc}_k(m) = [k + m \mod 5]$.

   Solution:

   Let $m = 0$, $m' = 1$ and $c = 0$.

   $\Pr[\mathsf{Enc}_k(0) = 0] \equiv \Pr[(k + 0 \mod 5) = 0] \equiv \Pr[k \mod 5 = 0]$, since $k \in \{0, \ldots, 5\}$, there are only two possible values for k: 0 or 5. Thus, $\Pr[\mathsf{Enc}_k(0) = 0] = 2/6$.

   Similarly, $\Pr[\mathsf{Enc}_k(1) = 0] \equiv \Pr[(k + 1 \mod 5) = 0]$. This is only possible when $k = 4$. Thus, $\Pr[\mathsf{Enc}_k(1) = 0] = 1/6$.

   Since $\Pr[\mathsf{Enc}_k(m) = c] \neq \Pr[\mathsf{Enc}_k(m') = c]$, the above scheme is not perfectly secret (violation of definition 2 of perfect secrecy).

   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:

   $|K| = 2^n$ and $|\mathcal{M}| = 2^{2n}$. Notice that $|K| < |\mathcal{M}|$. This violates Theorem 2.10 of the text book which states that for a perfectly secret encryption scheme $|K| \geq |\mathcal{M}|$. Thus, the above scheme is not perfectly secret.