Lab Assignment-3
Due: 14th February, 2023
Total Mark: 30

The problems in this assignment explore the issue of insecurity when a "one-time pad" is reused. In this problem, all characters are represented as 8-bit bytes with the ASCII encoding. Let $M = (m_1, m_2, \ldots, m_n)$ be a message, consisting of a sequence of n message bytes, to be encrypted. Let $P = (p_1, p_2, \ldots, p_n)$ denote a pad, consisting of a corresponding sequence of (randomly chosen) "pad bytes" (key bytes). In the usual one-time pad, the sequence $C = (c_1, c_2, \ldots, c_n)$ of ciphertext bytes is obtained by xor-ing each message byte with the corresponding pad byte:

$$c_i = m_i \oplus p_i, \text{ for } i = 1 \ldots n .$$

Problem 1 (Mark =10)

Consider the following two 8-character English words encrypted with the same "one-time pad". What are the words?

e9 3a e9 c5 fc 73 55 d5

f4 3a fe c7 e1 68 4a df

Hint: Use ***/usr/share/dict/words*** in Ubuntu OS as a representation of dictionary for English words.

Problem 2 (Mark = 20)

Assume, one of your fellow students tries to fix this problem by making sure that you can't just cancel the pad bytes by xor-ing the ciphertext bytes. In her/his scheme the key is still as long as the ciphertext. Suppose $c_0 = 0$, then the ciphertext bytes $c_1, c_2, \ldots, c_n$ are obtained as follows:

$$c_i = m_i \oplus ((p_i + c_{i-1}) \bmod 256) .$$

That is, each ciphertext byte is added to the next key byte and the addition result (modulo 256) is used to encrypt to the next plaintext byte. The student is now confident s/he can reuse her/his pad, since $(p_i + c_{i-1}) \bmod 256$ will be different for different messages, so nobody would be able to cancel the p_i's out.

(1) Write a program that implements the new encryption and decryption scheme. Show the result with a 10 character message and key.

(2) You are provided with a text file containing ten ciphertexts C_1, C_2, . . . , C_{10} produced by the new modified encryption and decryption technique. All of the ciphertexts are generated using the same pad P. The messages contain valid English text. Find the messages and the pad.

Hint:

1. The valid character set (characters of the plaintext) contains lowercase a – z, uppercase A – Z, space , , . ? ! - ( ). Total number of characters is 60.

2. Try to figure out possible pads for the valid characters.

3. Try to find the words of the messages using the dictionary mentioned in problem set 1.

(3) Marks on Viva.