Symmetric Key Cryptography Optional Worksheet (for self-study)

1. Alice and Bob have decided to use the Vigenère cipher to send private emails to each other. At their last meeting in person, they exchanged a 6 character word to be used as the secret key for the Vigenère cipher. Marvin, their adversary, intercepted the following encrypted email content sent by Alice to Bob:

   `VXUWPFSTRWINHHBFCPZDVSOYWRX`

   Marvin suspects that Alice usually begins her emails to Bob with the greeting `HIBOB` and has overheard that the key is 6 letters. He heard that the Vigenère cipher is vulnerable to known plaintext attacks, and would like to exploit this weakness to decrypt Alice's message.

   (a) Explain how Marvin can exploit a known plaintext attack to determine some information (or all) on Alice and Bob's secret key.

   (b) Find the secret key using the above data available to Marvin.

   (c) Use the information obtained in previous step to decrypt Alice's email to Bob.

   (a) Assuming the first 5 characters of the plaintext are `HIBOB`, Marvin can subtract the numerical value of those characters from the numerical value of the ciphertext characters to recover each letter of the secret key that is used to shift the letters of the plaintext (and hence the letters in the first 5 positions), e.g. first ciphertext letter V = 21, first plaintext letter H = 7 would reveal the first letter of the secret key as shift = 21 - 7 = 14 = O.

   (b) Similarly, Marvin gets `OPTIO` as first 5 of 6 characters of the secret key using the first 5 ciphertext letters `VXUWP`. He could also guess `N` as the last character of the secret key, knowing it is an English word.

   (c) Knowing the keyword is 6 letters long, Marvin can now use the first 5 keyword characters `OPTIO` to decrypt all except every 6th character of the message by subtracting from the ciphertext, recovering `HIBOB*EEYOU*TSIXO*LOCKA*ICE` where the missing letters `*` depend on the 6th letter of the secret key.

2. Does a substitution need to be a permutation of the plaintext symbols? Why or why not?

   No. A substitution can be to an entirely different alphabet. One plaintext symbol can convert to several ciphertext symbols, or vice versa. For example, Morse code is a form of substitution of alphabetic letters to dots and dashes. Two plaintext characters could map the same ciphertext character as long as the recipient could distinguish between the two.

3. **Attack Models:** In the security community, the security of any security scheme is typically analyzed with respect to attack models, which formally define two aspects:

   - security goal (the attacker's goal will be the opposite of this)
   - adversarial capability

   (a) **Security goal:** Differentiate between the different variants of security goals that could be considered for CONFidentiality: infeasibility of key recovery (KR), infeasibility of plaintext recovery (PR), indistinguishability (IND)

      i. Discuss which security goal is the strongest (hardest to achieve by the security defender, easiest to achieve by the attacker).

      ii. Hence, explain why (give brief arguments) the goal you identified above is the strongest security goal.

(b) **Adversarial capability:** Discuss the different types of adversarial capabilities that have so far been considered in the security literature, e.g. ciphertext-only attack (COA), known-plaintext attack (KPA), chosen plaintext/ciphertext attack (CPA/CCA), including which one is the strongest capability and why.

(c) **Attack:** Choose one of the encryption schemes you have learnt or are aware of, or you can google for more.

i. Explain a simple attack that can be mounted on this encryption scheme.

ii. Explain which attack model (what security goal, what adversarial capabilities) is required and relevant for this attack.

(d) **Defend:** Now switch sides and wear the security defender's hat.

i. Discuss why this attack was possible, what weakness it exploited, what adversarial capability was required by the attacker.

ii. Could we change the attack model to cause this attack to no longer be valid? e.g. we only consider an attack model with weaker adversarial capabilities, or we only consider an attack model with a weaker security goal.

iii. Then discuss how this attack can be prevented, e.g. could the encryption scheme be tweaked to resist attacks in this model?

(a) **Security goal:** Differentiate between the different variants of security goals that could be considered for CONfidentiality: infeasibility of key recovery (KR), infeasibility of plaintext recovery (PR), indistinguishability (IND)

i. The strongest (hardest to achieve by the security defender, easiest to achieve by the attacker) security goal is indistinguishability (IND).

ii. Intuitively, the reason IND is a stronger goal than PR and KR is that IND is the easiest for an attacker to break: the attacker only needs to be able to decrypt 1 bit of information on the plaintext to be able to break IND, while in PR the attacker has to recover the whole (potentially multi-bit) message, and in KR the attacker has to recover the key (which allows the attacker to recover plaintexts by decryption and of course break IND). More precisely, breaking KR implies breaking PR (by decrypting with the key) and breaking PR implies breaking IND (because if you can recover the plaintext you can check which of two possible messages it matches). But the reverse of those two implications may not be true. Hence IND is the easiest to break (or at least no harder than the other goals), and hence IND security is the strongest goal one can aim for. In particular, the above two implications show that IND security implies PR security (since if PR security was broken then so would IND security be broken) and PR security implies KR security (since if KR security was broken, so would PR security).

Remark: The above implications (also known as *security reductions*) illustrate a general method to simplify security analysis of systems. Suppose we have two security goals A and B we want to achieve for a system. Suppose we can show the following reduction relation between two security goals A and B: any attack that breaks goal A can also be used to break goal B. Then achieving goal B implies that goal A is also satisfied. It allows us to focus on just designing a system that satisfies goal B. In above example, if we design an encryption system that satisfies IND security goal, then it will also satisfy the PR and KR security goals.

(b) **Adversarial capability:** A ciphertext-only attack (COA) is the weakest attacker capability; the attacker only has access to the target ciphertext C^* that the attacker wants to decrypt, but no other information to help the attack. A stronger attacker capability exists in a known-plaintext attack (KPA), where in addition to the target ciphertext C^*, the attacker also observes other ciphertexts $C_1, C_2, \ldots, C_N$ which are the encryption of plaintexts $P_1, P_2, \ldots, P_N$ respectively under the key K, i.e.

$C_i = Enc(P_i, K)$ for $i = 1, \ldots, N$, and $P_1, \ldots, P_N$ are also known to the attacker (e.g. they may have been disclosed by the sender). Here, N is the number of known plaintext/ciphertext pairs available to the attacker (the larger N, the stronger the attacker capability). For example, the known plaintext, ciphertext pairs (P_i, C_i) may provide useful information on the secret key K to the attacker, which may help the attacker to decrypt C^*. An even stronger attacker capability exists in a chosen plaintext attack (CPA), which is similar to KPA, except that in CPA the attacker not only *knows* P_i but can even *choose* the P_i to be very special plaintexts that may reveal more information on the key, for instance. A stronger still capability exists in a chosen ciphertext attack (CCA), where in addition to the chosen plaintext/ciphertext pairs, the attacker can also choose ciphertexts $C'_1, \ldots, C'_M$ (different from the target C^*) and ask the decryptor to give the attacker the correponding decrypted plaintexts $P'_i = Dec(C'_i, K)$ for $i = 1, \ldots, M$. The P'_i may give the attacker some information on the desired plaintext P^*. For instance, C'_i may be related to target ciphertext C^* and the corresponding $P'_i = Dec(C'_i, K)$ may be related to the target plaintext $P^* = Dec(C^*, K)$.

(c) **Attack:** As an example, consider the one-time pad encryption algorithm with a re-used key K: $C = P + K \mod 26$, where K is random key in $\{0, \ldots, 25\}$.

i. Suppose we observe a known plaintext $P_1 = 3$ is encrypted with K to a known ciphertext $C_1 = 10$. To decrypt another target ciphertext $C^* = 5$ encrypted with K, the attacker can use the known plaintext/ciphertext pair (P_1, C_1) to compute the key $K = C_1 - P_1 \mod 26 = 10 - 3 \mod 26 = 7$. Now attacker can decrypt C^* to plaintext $P^* = C^* - K \mod 26 = 5 - 7 \mod 26 = 24$.

ii. The attack model used in the above attack is: goal = KR (key recovery) and attacker capabilities: KPA (known plaintext attack with $N = 1$ known plaintext/ciphertext pair).

(d) **Defend:** Now switch sides and wear the security defender's hat.

i. The weakness exploited was using the one-time pad method to encrypt more than one message with the same key K.

ii. Could we change the attack model to cause this attack to no longer be valid? If we change the attack model to a *one* ciphertext only attack (COA) so that the attacker can only observe one target ciphertext $C*$, then as we know from security of one-time pad, IND security is achieved unconditionally.

iii. To achieve (computationally secure) IND security even against KPA attacks with many plaintext/ciphertext pairs and same K, we can change scheme to one of the streaming modes of operation - similar to one-time pad, except that plaintext is added with pseudorandom keys output by a secure block-cipher (with some fixed key K) in say OFB mode with a new random IV chosen for encrypting each plaintext. Hence key added to plaintext is different for each ciphertext and the above attack is prevented.