HIT — Cryptography — Solutions 1

1603202-1150810613-Qiuhao Li

September 27, 2018

Problem 1. Show that the shift, Mono-Alphabetic sub., and Vigenère ciphers are all trivial to break using a known-plaintext attack. How much known plaintext (how many characters) is needed to completely recover the key for each of the ciphers? (show how to break the cipher)

Solution 1. :

As for the Shift Cipher, since $c = \mathsf{Enc}_k(m) = m + k \bmod 26 \Rightarrow k = c - m \bmod 26$, so we can completely recover the key using a single character.

For the other two problems, **I don't think there is a definitive answer.**

For Mono-Alphabetic Substitution Cipher, since the offset of each plaintext letter corresponding to the ciphertext may be different, we need a plaintext contains at least 25 different letters to completely recover the key (the remaining one plaintext letter mapped to the remaining one ciphertext letter), so the lengthof(plaintext) is an uncertain event...

For Vigenère Cipher, the n-th ciphertext letter is calculated by adding the n-th plaintext letter and the n-th key letter (where the key is repeated as many times as necessary to make it as long as the plaintext) modulo 26 (for the standard English alphabet), i.e.:

$$C_n \equiv M_n + K_n \bmod 26 \tag{1}$$

Conversely, if we know M_n and C_n, we can solve for K_n simply by subtracting pn from both sides of (1):

$$K_n \equiv C_n - M_n \bmod 26$$

So, to find out the key, given the ciphertext and plaintext, we simply need to decrypt the ciphertext using the plaintext as the key.

If the length of key is known, the encryption of $|k|$ (consecutive) characters of plaintext suces for recovering the entire key.

If the length of key is uncertain, we can assume that the key does not appear as "kk" (k is a sequence of letters) because it is meaningless for the encryption method, and since the length of the key is fixed, we need to obtain the minimum period length of the repeated sequence in the decrypted sequence. That is, we need a plaintext with a minimum length of 2 * lengthof(key), which is also an uncertain event(from the perspective of the attacker).

Problem 2. Show that the shift, Mono-Alphabetic sub., and Vigenère ciphers are all trivial to break using a chosen-plaintext attack. How much plaintext (how many characters) must be encrypted to completely recover the key? (show your chosen plaintext)

Solution 2. :

This problem is basically the same as the first problem, except that we can construct the plaintext ourselves at this time, making it easier to obtain the key.

As for the Shift Cipher, since $c = \mathsf{Enc}_k(m) = m + k \mod 26 \Rightarrow k = c - m \mod 26$, so we can completely recover the key using a single character.

As for the Mono-Alphabetic Substitution Cipher, since the offset of each plaintext letter corresponding to the ciphertext may be different, we need at least 25 different letters of plaintext to completely recover the key (the remaining one plaintext letter mapped to the remaining onr ciphertext letter). So we can simply construct the following plaintext/ciphertext pairs:

$$ABCDEFGHIJKLMNOPQRSTUVWXY \xrightarrow{c}$$

$$AZERTYUIOPQSDFGHJKLMWXCVB$$

So the first 25 letters of key is

$$AZERTYUIOPQSDFGHJKLMWXCVB$$

, and the last plaintext letter Z mapped to N, then the key is:

$$AZERTYUIOPQSDFGHJKLMWXCVBN$$

As for the Vigenère Cipher, if the length of key is known, the attack on the cipher remains the same as in the previous question.

If the length of key is uncertain, first we can assume that the key does not appear as "kk" (k is a sequence of letters) because it is meaningless for the encryption method. Next, we need a plaintext which contains many same letters and the corresponding ciphertext to recover the key. E.g:

$$AAAAAAAAAAAAAAAAAAAAAAAAA \xrightarrow{c}$$

$$ABCDEABCDEABCDEABCDEABCDE$$

We can observe that the minimum period of ciphertext repetition is 5 letters long, so the length of the key is also 5. Then we can solve the key according to the offset corresponding to each letter. Here the key is:

$$ABCDE$$

Therefore, in Virginia decryption, the required plaintext length is related to the length of the key k. Since we need to observe the minimum period length of the ciphertext, at least 2k length of plaintext is required.

Problem 3. Prove or refute: For every encryption scheme that is perfectly secret it holds that for every distribution over the message space $\mathcal{M}$, every $m, m' \in \mathcal{M}$, and every $c \in \mathcal{C}$:

$$\Pr[M = m | C = c] = \Pr[M = m' | C = c].$$

Solution 3. :

This statement is false.

Proof. According to definition of perfectly secret,

$$\Pr[M = m|C = c] = Pr[M = m].$$

$$\Pr[M = m'|C = c] = Pr[M = m'].$$

So,

$$\Pr[M = m|C = c] = \Pr[M = m'|C = c] \Rightarrow Pr[M = m] = Pr[M = m']$$

But, since distribution over the message space $\mathcal{M}$ is uncertain, this is flase. For example, in written English,

$$Pr[M = e] > Pr[M = q]$$

□

Problem 4. Study conditions under which the shift, mono-alphabetic sub., and Vigenère cipher ciphers are perfectly secret:

- (a) Prove that if only a single character is encrypted, then the shift cipher is perfectly secret.
- (b) What is the largest plaintext space M you can find for which the mono-alphabetic sub. cipher provides perfect secrecy?
- (c) Show how to use the Vigenère cipher to encrypt any word of length t so that perfect secrecy is obtained (note: you can choose the length of the key). Prove your answer.

Solution 4. :

- (a)

  Proof. Obviously, every $k \in \mathcal{K}$ is chosen with probability $1/|\mathcal{K}|$ by Gen. And for $\forall m \in \mathcal{M}$ and $\forall c \in \mathcal{C}$, $\exists$ unique $k \in \mathcal{K}$: $c = \mathsf{Enc}_k(m) = m + k \mod 26$. According to the Shannons Theorem, proved. □

- (b) 26!

  Proof. Since it should be a perfect secrecy, then $|\mathcal{K}| \geq |\mathcal{M}| \wedge max(|\mathcal{K}|) = 26!$. When the length of M equals 26(without duplicate letters), $|\mathcal{M}| = max(|\mathcal{K}|)$, as with the (a) problem, according to Shannons Theorem, this is a perfect secrecy. So the largest plaintext space $|\mathcal{M}|$ is 26!. □

- (c) I would choose the length of the key to be t.

Proof. If the length of the key is t, then every $k \in \mathcal{K}$ is chosen with probability $1/|\mathcal{K}|$ by Gen. And for $\forall m \in \mathcal{M}$ and $\forall c \in \mathcal{C}$, $\exists$ unique $k \in \mathcal{K}$: $c = \mathsf{Enc}_k(m)$ (To obtain the key, we can simply decrypt the ciphertext using the plaintext as the key). According to the Shannons Theorem, proved. □

Problem 5. In the one-time pad encryption scheme, it can sometimes happen that the key is the all-zero string. In this case, the encryption of a message m is given by $m \oplus 0^l = m$ and therefore the ciphertext is identical to the message!

- (a) Do you think the one-time pad scheme should be modified so that the all-zero key is not used? Explain.
- (b) Explain how it is possible that the one-time pad is perfectly secure even though the above situation can occur with non-zero probability.

Solution 5. :

- (a) I don't think all-zero key should be avoided in the one-time pad cipher, which will give the attack imformation about our plaintext. Give a simple counterexample, if we intentionally avoid using all-zero key, when attack sniffed a ciphertext "Yes", then he can be sure that the plaintext wouldn't be "Yes"(Which may be "Non").
- (b) "Perfectly Secure" means that attacker can not get any imformation about out plaintext from the ciphertext he has sniffed. As long as we followed the definition of perfectly secrecy above, for one ciphertext, it can correspond to any plaintext. For instance, we use the all-zero key with plaintext "attack", so attacker also get the "attack" message. But, can he be sure that plaintext is also "attack" instead of "defend"? — If our cipher is correct, "attack" and "defend" have the same probability to be encrypted as "attack".

Happy  Hacking!