Fall 2023 2023-09-12

Solutions for Assignment 1

Humesh Reddy Venkatapuram *CS 380/525: Intro to Crypto*

1) Show that the Caesar (shift by 3) cipher and Vigen‘ere cipher defined over the 128-character ASCII character set, are easy to break by doing a chosen-plaintext attack. How much plaintext is needed to recover the key for each of the ciphers?

Answer: In a chosen-plaintext attack is when you pick some specific text and see how it looks when it is encrypted. For a caesar cipher, which can shift the each character by the 3 positions with in the 128-character ASCII set and this can help you crack the code that was used to encrypt it(only need one letter to crack the code). Some codes are easier to crack than others.

-One code is a Caesar cipher. It moves each letter in your text by a fixed number of places in the alphabet. To crack this code, you just need to encrypt one letter, like any alphabet, and see what it turns into. Then, you can find the number that was used to move the letters. That number is the code's key.

(A Caesar cipher with a fixed shift of 3, you dont need any of the plaintext to figure it out the key because everyone knows that. So the key isn't really a secret in this case.)

- Another code is a Vigenère cipher. It uses a word as the key, and moves each letter in your text by a different number based on the word. To crack this code, you need to encrypt more text than before. You need to guess how long the word is, and use different letters for each spot in the word. Then, you can compare how each letter moved in the encrypted text and find out the word that was used as the key.

-Example: I chosen a plaintext attack with the keyword length of 3: HELLOHELLO

-If you encrypt the chosen plaintext into suspected keyword will be: OEDRVIDVR
-Here I used the key - "KEY"

- Both of these codes are bad at hiding the data, because they can be easily cracked with a little bit of chosen plaintext. They also don't work well with many different characters, like numbers or symbols. We should use a good/better code if you want to keep the data safe.

P = C - K(mod 128)

2) Consider the case where the Caesar (shift by 3) cipher and Vigen‘ere cipher are defined over the 26-character English alphabet. How much plaintext is needed to recover the key for each of the ciphers?

Answer: A Caesar cipher is a code that can change the each character in the text by moving it by a certain number of places in the list of all possible characters. For example, if the number is

3, then A becomes D, B becomes E, and so on. To break this code, you need only one character. Encrypt any character (for example, A), and see what it turns into. The difference between the characters tells you the number can be used to move them. This number is the key to the code.

(A Caesar cipher with a fixed shift of 3, you dont need any of the plaintext to figure it out the key because everyone knows that. So the key isn't really a secret in this case.)

A Vigenère cipher is a code that uses a word as the key, and changes each character in your text by moving it by a different number based on the word. The amount of the plaintext is needed to break the cipher that depends on the length of the keyword used by the encryption. For example, if the word is CAT, then the first character in your text moves by 3 places (because C is the third character in the list), the second character moves by 1 place (because A is the first character), and the third character moves by 20 places (because T is the twentieth character). Then, it repeats for the next characters in your text. To break this code, you need more characters than before. You need to guess how long the word is, and use different characters for each spot in the word.

In simple terms, if you're dealing with a 128-character ASCII character set: To break a Caesar cipher, you only need to look at the encryption of a single character. For a Vigenère cipher, you'd need to examine the encryption of several characters, specifically, at least as many characters as there are in the keyword you suspect they used.

3) Consider the MAC presented in the "Overview" slides set. What happens when Charlie tries to modify the Tag, as well as the ciphertext in transit? Explain why this attack will fail, i.e, Bob will always be able to detect a modified Tag/ciphertext.

Answer: Alice and Bob have a secret code that only they know. They use this code to make a special sticker that they put on their messages. The sticker shows that the message is really from them and no one has changed it.

Charlie is someone who wants to mess with their messages. He tries to change the message or the sticker, but he doesn't know the secret code. So he can't make a new sticker that matches the old one.

When Bob gets the message, he checks the sticker using the secret code. If the sticker is different from what he expects, he knows that Charlie has messed with it. This way, he can tell if the message is real or not.

When a sender sends a message, they use the secret code to make a sticker for it. This sticker is called a MAC tag. The receiver gets the message and the sticker, and uses the same secret code to check if they match. If they don't match, the receiver knows that someone has changed the message or the sticker.

Even if Charlie attempts to tamper with both the message M and the ciphertext C and tries to create a Tag, denoted as Tag', it will not match the original Tag due to his lack of knowledge about the secret key K. In simple terms, Tag' will not equal Tag, which means that Bob's security measures will thwart Charlie's attack.

4) Consider the CBC mode of encryption:

(a) (20 points) Assume Alice is encrypting all the blocks, and Bob is decrypting them. Draw the corresponding diagram for Bob's decryption.

(b) (20 points) Assume Alice has encrypted 5 message blocks, M1, M2, M3, M4, M5, and the corre- sponding ciphertext blocks are C1, C2, C3, C4, C5. Now assume Charlie tampers with block C1 in transit, and changes it to C1.

How will this affect Bob's decryption? Will he be able to able to decrypt all blocks correctly? Or none? Or some?

Answer: Alice and Bob use a special way to send messages that are hard to break. They use a secret code and a random number to mix up the message. They also split the message into smaller parts, called blocks, and mix them up with each other.

Charlie wants to mess with one of the blocks, called C1. He changes it to C1'. But this will make two parts of the message look wrong when Bob gets it.

The first part, M1, will look wrong because C1' will not work well with the random number that Alice and Bob use. The second part, M2, will also look wrong because it depends on C1' for its mixing. The other parts of the message will be fine because they don't need C1 for anything.

So here M1 will be grabled because C1' will not decrypt correctly.
M2 will also be grabled because it depends on the C1' for its XOR operation after decryption of C2.
The blocks M3,M4,M5,.......... will be decrypted correctly because they dont depend on C1'

(c) (20 points) What happens if the Initialization Vector (IV), i.e, the initial randomness fed in to the first block arrives mangled at Bob's end? How will this affect Bob's decryption? Will he be able to able to decrypt all blocks correctly? Or none? Or some?

-"Cipher Block Chaining" (CBC) mode is a way of encrypting the data that makes it more harder for someone to break the code. It works by adding some random bits (It is called an "initialization vector" or "IV") to the first piece of data (called a block) before encrypting it. Then, it adds the encrypted block to the next block before encrypting that one, and so on.So, this way, each block depends on the previous one, and changing the one bit will affect all the following blocks.

-To decrypt the data, you need to do the opposite of what you did to encrypt it. You also need to know the initialization vector (IV) that was used in the beginning. The IV is not a secret, so you can send easily along with the encrypted data to any person who wants to decrypt it.

-Is the initial randomness fed in to the first block arrives mangled at Bob's end? No, that won't happen because CBC mode is designed to be very secure and prevent anyone from tampering with the data. Even if someone changes the IV, they won't be able to figure out what the original data was. The only thing that will happen is that the first block of decrypted data will be wrong, because it depends on the IV. But all the other blocks will be correct, because they depend on the encrypted blocks, not on the IV.

-Will he be able to able to decrypt all blocks correctly? Or none? Or some? Yes, that's right. You can still recover most of the data, even if the IV is wrong. The first block might be scrambled, but the rest will be intact.

-This shows that CBC mode is very good at protecting your data from attacks. It makes sure that no one can change or read your data without knowing the key. It also makes sure that each block is different from the others, so that no patterns or repetitions can be found.

REFERENCES:

1) https://www.geeksforgeeks.org/transforming-a-plain-text-message-to-cipher-text/

2) https://www.includehelp.com/cryptography/caesar-cipher.aspx