1.5 Traditional symmetric-key ciphers

This section will focus on traditional symmetric-key ciphers used in the past and explain the principles of such ciphers so as to prepare the you for the next few sections which will focuss on modern symmetric-key ciphers.

Objectives

- To define the terms and the concepts of symmetric-key ciphers
- To describe the categories of cryptanalysis used to break the symmetric ciphers
- To introduce the concepts of the stream ciphers and block ciphers

The terms and definitions presented here will be used in all later sections on symmetric-key ciphers.

Though traditional symmetric-key ciphers are not used today, we will study them here for several reasons:

- They are simpler than modern ciphers and easier to understand by the beginners.
- They give the basic foundation of cryptography and encipherment, the foundation which is used to better understand modern ciphers.
- A look at Cryptographic histories reveals stories of how cryptography, and in particular cryptanalysis (attacking cryptographic systems), has played a pivotal role in the development of the modern cryptography. Ciphers that were secure in earlier years are no longer secure in the present computer age.

Traditionally, cryptography involved two people trying to talk without being overheard.

Modern cryptography involves communication in the presence of powerful adversaries. It is about much more than plain communication between two people and now includes human activities such as finance and voting.

By the end of this chapter, you should be able to use modern cryptography in the real world to help defend things we care about, both abstract and concrete: secure messaging, Bitcoin and cryptographic voting schemes.

1.6 Symmetric encryption

Suppose I want to send messsages to you via some communications channel. Let us assume that an Intruder Eve has access to the channel, and she may eavesdrop on and tamper with anything sent over the channel. As the sender, I do not want Eve to be able to eavesdrop on my messages (also called plaintext). I want to communicate confidentially. When you receive a message seemingly from me, you want to be sure that I sent the message and that it was not tampered with. Both you and I want integrity. You and I share a secret (called a key).

Cryptography where the sender and the receiver (or the honest users) have the same key is called symmetric cryptography.

Symmetric encryption, also referred to as conventional encryption or single-key encryption is by far the most widely used of the two types of encryption.

Definition 2

Suppose I want you to learn a plaintext (original message). I will send you a ciphertext or encryption, from which you will learn the plaintext. Creating the ciphertext from the plaintext is called enciphering or encryption and it uses an encryption algorithm and a shared secret key. Restoring the plaintext from the ciphertext is called deciphering or decryption and it uses a decryption algorithm and the same secret key. In order to make this possible, both you and I have a shared secret called a secret key. The encryption and decryption algorithms are referred to as ciphers. A key is a set of values (numbers) that the cipher, as an algorithm, operates on.

Definition 3

A symmetric cryptosystem consists of a set $\mathcal{K}$ of keys, a set $\mathcal{P}$ of plaintexts, a set $\mathcal{C}$ of ciphertexts, and two algorithms:

- an encryption algorithm $\mathcal{E}$ that on input of a key and a plaintext outputs a ciphertext; and
- a decryption algorithm $\mathcal{D}$ that on input of a key and a ciphertext outputs either a plaintext or the special symbol $\perp$ (indicating an invalid ciphertext).

That is, for any key k and any plaintext m, we have that

$$\mathcal{D}(k, \mathcal{E}(k, k)) = m.$$

The plaintext set $\mathcal{P}$ will usually be a set of finite sequences (strings) of letters from an alphabet.

Note here that the symmetric-key encipherment uses same key for both encryption and decryption and that the encryption and decryption algorithms are inverses of each other.

Using symmetric-key encipherment, person A and person B can use the same key for communication from either direction, from person A to person B and vise versa. This is why the method is called symmetric.

Supposing person A wants to communicate to another person C, another different secret key is needed between them.

If there are m people in a group who need to communicate with each other then $(m \times (m-1))/2$ keys are needed since each person needs $m-1$ keys to communicate with the rest of the group members, but the key between A and B can be used in both directions. Later we see how this problem is handled.

Later sections will show that the asymmetric-key encipherment needs two keys, one for encryption and one for decryption.

As we discuss historic cryptosystems we should also get a gentle introduction to the basic concepts in cryptography that will assist us to see some insight into important attack strategies.

Confidentiality: Let us consider a situation where I send you messages while Eve eavesdrops. Eve wants to understand what I am saying to you.

A symmetric cryptosystem provides confidentiality if it is, without knowledge of the key, hard to learn anything at all about the decryption of a ciphertext from the ciphertext itself, except possibly the length of the decryption.

Cryptanalysis: Cryptanalysis is the process of decrypting a message without knowing the underlying key. We need to study cryptanalysis techniques not to break other people's codes but to learn how vulnerable our cryptosystem is. This will in turn help us create better secret codes.

1.7 Substitution ciphers

In substitution cipher each character (sometimes groups of characters) in a message is replaced with a different character according to fixed rules.

For instance, if the symbols in the plaintext are alphabetic characters we can replace letter C with letter M, and letter T with letter K and if the symbols are digits (0 to 9), we can replace 2 with 6, and 3 with 8.

Substitution ciphers are categorized as either monoalphabetic ciphers or polyalphabetic ciphers.

Monoalphabetic Ciphers

In monoalphabetic substitution, if the algorithm says that letter B in the plaintext is changed to letter D, every letter B is changed to letter D. That is, the relationship between letters in the plaintext and the ciphertext is one-to-one.

Example 6

Consider the following examples of plaintext with their corresponding ciphertext.

1. Plaintext: **hello** Ciphertext: **WTAAD**
   Note here that both l's are encrypted as A's and so the cipher is probably monoalphabetic.
2. Plaintext: **hello** Ciphertext: **ABNZF**
   Here, each l is encrypted by a different character and so the cipher is not monoalphabetic.

a) Shift Cipher

This is simplest monoalphabetic cipher. In case you have a message you want to transmit securely, you can encrypt it (translate it into a secret code). One of the simplest ways to do this is with a shift cipher. In the shift cipher, also known as the Caesar cipher or as additive ciphers, we give the alphabet G a group structure. Here, there is a bijection between the English alphabet $\{a, b, c, \cdots z\}$ and the group $\mathbb{Z}_{26}$ given by $0 \leftrightarrow a,\ 1 \leftrightarrow b,\ 2 \leftrightarrow c$, etc as shown it the table below.

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Thus, to add f and g we apply the bijection to get 5 and 6 and add them get 11 and then apply the bijection to get L.

The plaintext m is a sequence of letters $m_1 m_2 \cdots m_l$ from the alphabet. The key (also called encryption key) is an element k from G and is just the length of the shift we are using.

We encrypt the message by adding the key to each letter and using modular arithmetic, that is, the i^{th} ciphertext letter is

$$c_i = (m_i + k) \bmod n, \quad 1 \leq i \leq l. \tag{1}$$

where $k \in \{0, 1, 2, \cdots, 25\}$.
The ciphertext c is the sequence of letters $c_1 c_2 \cdots c_l$.

To decrypt a ciphertext $c = c_1 c_2 \cdots c_l$, we subtract the key from each ciphertext letter and reduce it modulo n, that is, the i^{th} plaintext letter is

$$m_i = c_i - k, \quad 1 \leq i \leq l.$$

Example 7

You agree with your friend to use the shift cipher with the key $k = d$ (d corresponds to the number 3).
Encrypt CYPRIAN SAKWA IS HERE .

Solution

We do the following computations:

$$\begin{array}{cccccccccccccccccccc}
 & \text{C} & \text{Y} & \text{P} & \text{R} & \text{I} & \text{A} & \text{N} & \text{S} & \text{A} & \text{K} & \text{W} & \text{A} & \text{I} & \text{S} & \text{H} & \text{E} & \text{R} & \text{E} & \\
 & 2 & 24 & 15 & 17 & 8 & 0 & 13 & 18 & 0 & 10 & 22 & 0 & 8 & 18 & 7 & 4 & 17 & 4 & \\
+ & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & \\
\hline
 & (5 & 1 & 18 & 20 & 11 & 3 & 16 & 21 & 3 & 13 & 25 & 3 & 11 & 21 & 10 & 7 & 20 & 7) & \bmod 26 \\
 & \text{F} & \text{B} & \text{S} & \text{U} & \text{L} & \text{D} & \text{Q} & \text{V} & \text{D} & \text{N} & \text{Z} & \text{D} & \text{L} & \text{V} & \text{K} & \text{H} & \text{U} & \text{H} &
\end{array}$$

The ciphertext is FBSULDQVDNZDLVKHUH

So you give the message FBSULDQVDNZDLVKHUH to our friend.

To decrypt, your friend does the following computations:

$$\begin{array}{cccccccccccccccccccc}
 & \text{F} & \text{B} & \text{S} & \text{U} & \text{L} & \text{D} & \text{Q} & \text{V} & \text{D} & \text{N} & \text{Z} & \text{D} & \text{L} & \text{V} & \text{K} & \text{H} & \text{U} & \text{H} & \\
 & 5 & 1 & 18 & 20 & 11 & 3 & 16 & 21 & 3 & 13 & 25 & 3 & 11 & 21 & 10 & 7 & 20 & 7 & \\
- & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & \\
\hline
 & (2 & 24 & 15 & 17 & 8 & 0 & 13 & 18 & 0 & 10 & 22 & 0 & 8 & 18 & 7 & 4 & 17 & 4) & \bmod 26 \\
 & \text{C} & \text{Y} & \text{P} & \text{R} & \text{I} & \text{A} & \text{N} & \text{S} & \text{A} & \text{K} & \text{W} & \text{A} & \text{I} & \text{S} & \text{H} & \text{E} & \text{R} & \text{E} &
\end{array}$$

The friend then breaks it into words and uses the appropriate punctuation to be able to read the message as CYPRIAN SAKWA IS HERE.

Exercise 2

a) Use the shift cipher with key $= 5$ to encrypt the message “hellobuilders”.

b) Use the shift cipher with key $= 5$ to decrypt the message “mjqqtgznqijwx”.

c) How many different keys are there for the shift cipher given that the alphabet has 26 elements?

d) Find all the possible decryptions of KTKSE. How many are English words?

e) Choose a random key for the shift cipher and use it to encrypt a message. Let someone else decrypt the ciphertext without telling them the key.

1.8 Cryptanalysis of shift ciphers

Let us suppose Eve intercepted a transmission of an encrypted message, and she knows that the sender had used a shift cipher on the English alphabet, but she doesn't know the encryption key. How difficult would it be for her to break the code?

If we exclude the encryption key 0, there are only 25 distinct shifts that might have been used. It probably wouldn't take very long (especially with computer help) to test each of these shifts in turn. This type of attack on the shift cipher where someone trys all possible choices of keys is called an exhaustive search for the key, or a brute force attack. If there are few keys, we can decrypt with all possible keys in reasonable time. The correct key will be the one that gives a reasonable decryption.

Example 8

You intercepted the ciphertext “mnfgjcbfxuwjxjsy”. Show how you can use a brute-force attack to break the cipher.

Solution

Trying keys from 1 to 5 shows that with a key of 5, the plaintext is not very secure since it makes sense.

$k_1 \rightarrow$ plaintext: lmepibaewtviwirx

$k_2 \rightarrow$ plaintext: kldohazdvsuhvhqw

$k_3 \rightarrow$ plaintext: jkcngzycurtgugpv

$k_4 \rightarrow$ plaintext: ijbmfyxbtqsftfou

$k_5 \rightarrow$ plaintext: Hialexwaspresent

Another kind of attack is the known plaintext attack, where you know part or all of the solution. For example, let's say you know that I start all my messages with "HI THERE", you can easily determine the key for the message LMEPIB.

The other type of attack is the frequency analysis attack. This is especially true where long ciphertexts are involved. The attacker can use the frequency of occurrence of characters for a particular language. A frequency analysis attack involves looking at how many times each letter appears in the encrypted message, and using this information to crack the code. For instance T is more likely to appear many times in a message than Z, while a, e, i, o, u appear more than the other letters of the English alphabet and the letter q in English is virtually always followed by the letter u.

Advantages of Shift Ciphers:

a) Is easy to implement and use thus and therefore suitable for beginners to learn about encryption.

b) It requires only a small set of pre-shared information.

c) It can be modified easily to create a more secure variant, such as use of a multiple shift values or keywords.

Disadvantages of Shift Ciphers:

a) Not secure against modern decryption methods.

b) It is vulnerable to a brute force attack due to a small number of possible keys since an attacker can easily try all possible keys until the correct one is found.

c) It is not suitable for long text encryption as it would be easy to crack.

1.9 Affine Cipher

The encryption key for an affine cipher is an ordered pair (k_1, k_2) of integers, both of which come from the set $\{0, 1, \cdots, n-1\}$, where n is the size of the set being used (for us, the set is the English alphabet, so we have $n = 26$).

Encryption: It uses modular arithmetic to transform an integer that each plaintext letter corresponds to into another letter corresponding to a ciphertext letter. The encryption function gives the ciphertext as

$$c_i = (k_1 x + k_2) \bmod n$$

with k_1 chosen so that it is relatively prime to n.

Decryption: Here we perform the inverse functions of the ciphertext to retrieve the plaintext. First, convert each of the ciphertext letters into integer values then use the decryption function

$$D_i = a^{-1}(x_i - b) \bmod n$$

Example 9

You agree with your friend to use the Affine cipher with the key $(3,5)$. Ecrypt the plaintext message CYPRIAN SAKWA IS HERE.

Solution

1. First turn CYPRIAN SAKWA IS HERE into the sequence of numbers 2 24 15 17 8 0 13 18 0 10 22 0 8 18 7 4 17 4
2. Feed these numbers into function $f(x) = (3x + 5) \mod 26$ one at a time to produce the sequence 11 25 24 4 3 5 18 7 5 9 19 5 3 7 0 17 4 17
3. Interpret the resultant numbers again as letters to produce the ciphertext L Z Y E D F S H F J T F D H A R E R

So you give to your friend the message LZYEDFSHFJTFDHARER

To decrypt, your friend does the following computations:

1. First turns LZYEDFSHFJTFDHARER into the sequence of numbers 11 25 24 4 3 5 18 7 5 9 19 5 3 7 0 17 4 17
2. Finds $3^{-1} \bmod 26 = 9$. This could be achieved by either extended Euclidean Algorithm or Euler's Theorem.
3. Feeds these numbers into function $D(x) = 9(x - 5) \bmod 26$ one at a time to produce the sequence 2 24 15 17 8 0 13 18 0 10 22 0 8 18 7 4 17 4
4. Interprets the resultant numbers again as letters to produce the plaintext CYPRIANSAKWAISHERE

   The friend then breaks it into words and uses the appropriate punctuations to be able to read the message as CYPRIAN SAKWA IS HERE.

Breaking an Affine Cipher

Example 10

Now, suppose EVE intercepted this ciphertext message “LZYEDFSHFJTFDHARER”, and didn’t know what the original message was, but wished to discover it. Suppose she knew that the affine cipher was used.

Let us suppose she guesses $f(C) = L$ and $f(E) = R$. Such a guess can come from a variety of sources, such as analyzing the frequency of certain letters; or pairs of letter; or from a known or guessed part of the original message.

She would then translate the guesses involving ‘C’,‘E’,‘L’ and ‘R’ into integers using $A = 0,\ B = 1, \cdots Z = 25$, then we have $f(2) = 11$ and $f(4) = 17$ which yield

$$4x + y \equiv 17 \bmod 26$$

$$2x + y \equiv 11 \bmod 26$$

Example (conti...)

Subtracting the second congruence from the first one yields

$$2x \equiv 6 \bmod 26 \quad \Rightarrow \quad x \equiv 3 \bmod 26$$

Substituting the value of x in $4x + y \equiv 17 \bmod 26$ gives

$$y \equiv 5 \bmod 26.$$

Knowing,$x = 3$ and $y = 5$ she gets the encrypting function

$$f(x) = 3x + 5 \bmod 26.$$

With the encrypting function (i.e., the "secret key" for the affine cipher), she proceeds to decrypt the message as explained above.

Exercise 3

a) Suppose you have intercepted this ciphertext message “CTOOX”, and you didn’t know what the original message was, but wished to discover it. Suppose you know that the affine cipher was used. Suppose further that you guess $f(\mathsf{L}) = \mathsf{O}$ and $f(\mathsf{E}) = \mathsf{T}$. Find the encrypting function and then decrypt the message.

b) Suppose you have intercepted this ciphertext message “UBLJYMBXKTXPVX”, from your classmate and you didn’t know what the original message was, but wished to discover it. Suppose you know that the affine cipher was used and that your classmate always begins messages with HI. Find the key that was used and decrypt the message.

c) Choose a random key for the affine cipher and encrypt a message. Let someone else decrypt the ciphertext without knowing the key, but give them some known plaintext.