Chapter 1

Introduction to Cryptography and Data Security

1.1 Literature Recommendations

Course Textbooks: [Sti95] or [Sch93].

Further Reading - the following books are excellent supplements to the course textbook:

1. [AM97] - great compilation of theoretical and practical aspects of many crypto schemes. Unique since it includes many theoretical topics that are hard to find otherwise. Highly recommended.
2. [Sta95] - Very readable treatment of algorithms and standards relevant to cryptography in networks.

1.2 Overview

Brief History of Cryptography

- Private-Key: all encryption and decryption schemes dating from BC to 1976.

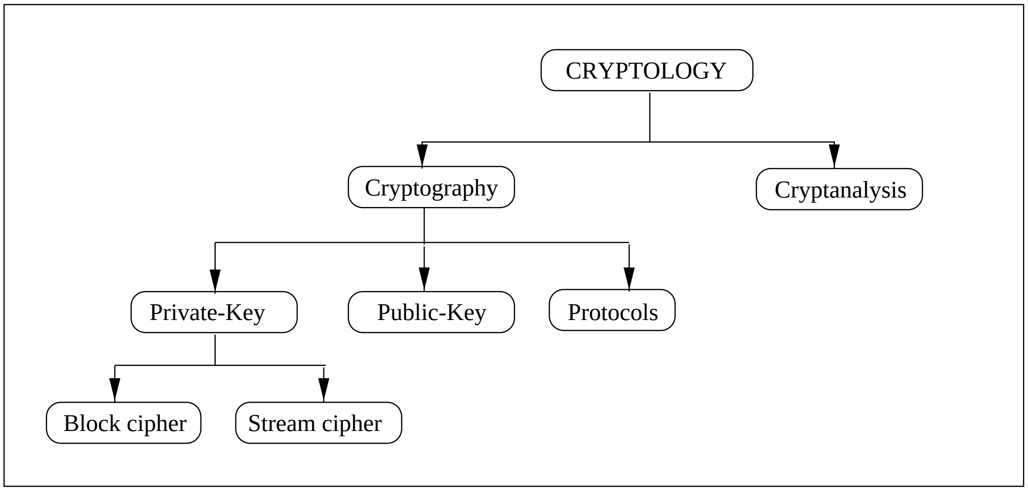

Figure 1.1: Overview on the field of cryptology

- Public-Key: in 1976 the first public-key scheme was introduced by Diffie-Hellman key exchange protocol.
- Hybrid Approach: in today's protocol, very often hybrid schemes are applied which use private and public-key algorithms.

1.3 Private-Key Cryptosystems

Sometimes these schemes are also referred to as *symmetric*, *single-key*, and *secret-key* approaches.

Problem Statement: Alice and Bob want to communication over an un-secure channel (e.g., computer network, satellite link). They want to prevent Oscar (the bad guy) from listening.

Solution: Use of private-key cryptosystems (these have been around since BC) such that if Oscar reads the encrypted version y of the message x over the un-secure channel, he will not be able to understand its content because x is what really was sent.

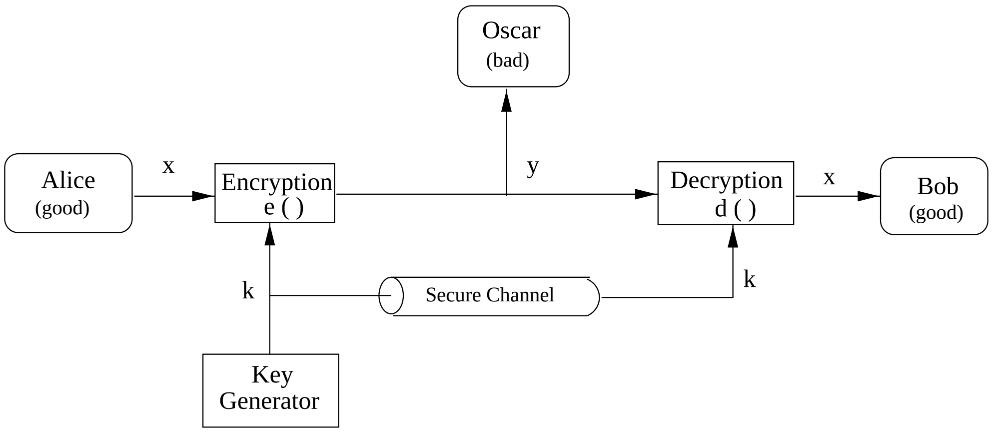

Figure 1.2: Private-key cryptosystem

Some important definitions:

1a)	x is called the "plaintext"
1b)	$\mathcal{P}= \{x_1, x_2, \ldots, x_p\}$ is the (finite) "plaintext space"
2a)	y is called the "ciphertext"
2b)	$\mathcal{C}= \{y_1, y_2, \ldots, y_c\}$ is the (finite) "ciphertext space"
3a)	k is called the "key"
3b)	$\mathcal{K}= \{k_1, k_2, \ldots, k_l\}$ is the finite "key space"
4a)	There are l encryption functions $e_{k_i} : \mathcal{P} \rightarrow \mathcal{C}$ (or: $e_{k_i}(x) = y$)
4b)	There are l decryption functions $d_{k_i} : \mathcal{C} \rightarrow \mathcal{P}$ (or: $d_{k_i}(y) = x$)
4c)	e_{k_1} and d_{k_2} are inverse functions if $k_1 = k_2 : d_{k_i}(y) = d_{k_i}(e_{k_i}(x)) = x$ for all $k_i \in \mathcal{K}$

Example: Data Encryption Standard (DES)

- $\mathcal{P} = \mathcal{C}= \{0, 1, 2, \ldots, 2^{64} - 1\}$ (each x_i has 64 bits: $x_i = 010 \ldots 0110$)
- $\mathcal{K}= \{0, 1, 2, \ldots, 2^{56} - 1\}$ (each k_i has 56 bits)
- encryption (e_k) and decryption (d_k) will be described in Chapter 4

1.4 Cryptanalysis

Definition: The science of recovering the plaintext x from the ciphertext y without the knowledge of the key (Oscar's job).

Rules of the game:

The cryptanalysis rules are known as *Kerckhoff's Principle*:

1. Oscar knows the cryptosystem (encryption and decryption algorithms).

2. Oscar does not know the key.

1.4.1 Attacks against Cryptoalgorithms

1. Ciphertext-only attack

Oscar's knowledge: some $y_1 = e_k(x_1)$, $y_2 = e_k(x_2)$, ...

Oscar's goal : obtain $x_1, x_2, \ldots$ or the key k.

2. Known plaintext attack

Oscar's knowledge: some pairs $(x_1, y_1 = e_k(x_1)), (x_2, y_2 = e_k(x_2)) \ldots$

Oscar's goal : obtain the key k.

3. Chosen plaintext attack

Oscar's knowledge: some pairs $(x_1, y_1 = e_k(x_1)), (x_2, y_2 = e_k(x_2)) \ldots$ of which he can choose $x_1, x_2, \ldots$

Oscar's goal : obtain the key k.

4. Chosen ciphertext attack

Oscar's knowledge: some pairs $(x_1, y_1 = e_k(x_1)), (x_2, y_2 = e_k(x_2)) \ldots$ of which he can choose

$y_1, y_2, \ldots$

Oscar's goal : obtain the key k.

1.5 Some Number Theory

Modulo operation:

Question: What is $12 \bmod 9$?

Answer: $12 \bmod 9 \equiv 3$

or $12 \equiv 3 \bmod 9$.

Definition 1.5.1 Modulo Operation

Let $a, r, m \in Z$ (where Z is a set of all integers) and $m > 0$. We write $a \equiv r \bmod m$ if m divides $r - a$.

"m" is called the modulus.

"r" is called the remainder.

Some remarks on the modulo operation:

- How is the remainder computed?

  It is always possible to write $a \in Z$, such that

  $a = q \cdot m + r;\ 0 \leq r < m$

  Now since $a - r = q \cdot m$ (m divides $a - r$) and $a \equiv r \bmod m$.

  Note that $r \in \{0, 1, 2, \ldots, m-1\}$.

  Example:

  $a = 42;\ m = 9$

  $42 = 4 \cdot 9 + 6$ therefore $42 \equiv 6 \bmod 9$.

- C programming command : "%" (C can return a negative value)

  `r = 42 % 9` returns r = 6

  but `r = -42 % 9` returns r = -6 $\rightarrow$ if remainder is negative, add modulus m:

  $-6 + 9 = 3 \equiv -42 \bmod 9$

Ring:

Definition 1.5.2 *The "ring Z_m" consists of:*

1. *The set* $Z_m = \{0, 1, 2, \ldots, m - 1\}$
2. *Two operations "+" and "×" for all $a, b \in Z_m$ such that:*
   - $a + b \equiv c \bmod m$ $(c \in Z_m)$
   - $a \times b \equiv d \bmod m$ $(d \in Z_m)$

Example: $m = 9$

$Z_9 = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$

$6 + 8 = 14 \equiv 5 \bmod 9$

$6 \times 8 = 48 \equiv 3 \bmod 9$

Definition 1.5.3 Some important properties of the ring $Z_m = \{0, 1, 2, \ldots, m-1\}$

1. *The* additive identity *is the element zero "0":* $a + 0 = a \bmod m$, *for any* $a \in Z_m$.

2. *The* additive inverse *"$-a$" of "a" is such that* $a + (-a) \equiv 0 \bmod m$: $-a = m - a$, *for any* $a \in Z_m$.

3. *Addition is* closed*: i.e., for any* $a, b \in Z_m$, $a + b \in Z_m$.

4. *Addition is* commutative*: i.e., for any* $a, b \in Z_m$, $a + b = b + a$.

5. *Addition is* associative*: i.e., for any* $a, b \in Z_m$, $(a + b) + c = a + (b + c)$.

6. *The* multiplicative identity *is the element one "1":* $a \times 1 \equiv a \bmod m$, *for any* $a \in Z_m$.

7. *The* multiplicative inverse *"a^{-1}" of "a" is such that* $a \times a^{-1} = 1 \bmod m$*: An element* a *has a multiplicative inverse "a^{-1}" if and only if* $\gcd(a, m) = 1$.

8. *Multiplication is* closed*: i.e., for any* $a, b \in Z_m$, $ab \in Z_m$.

9. *Multiplication is* commutative*: i.e., for any* $a, b \in Z_m$, $ab = ba$.

10. *Multiplication is* associative*: i.e., for any* $a, b \in Z_m$, $(ab)c = a(bc)$.

Some remarks on the ring Z_m:

- Roughly speaking, a ring is a structure in which we can add, subtract, multiply, and sometimes divide.

- **Definition 1.5.4** *If* $\gcd(a, m) = 1$, *then* a *and* m *are* "relatively prime" *and the multiplicative inverse of* a *exists.*

  Example:

  i) **Question**: does multiplicative inverse exist with 15 mod 26?

  Answer: yes — $gcd(15, 26) = 1$

  ii) **Question**: does multiplicative inverse exist with 14 mod 26?

  Answer: no — $gcd(14, 26) \neq 1$

- The modulo operation can be applied whenever we want:

  $(a + b) \bmod m = [(a \bmod m) + (b \bmod m)] \bmod m.$

  $(a \times b) \bmod m = [(a \bmod m) \times (b \bmod m)] \bmod m.$

  Example: $3^8 \bmod 7 = ?$

  i) $3^8 = 3^4 \cdot 3^4 = (81 \bmod 7) \cdot (81 \bmod 7) \equiv 4 \cdot 4 = 16 \equiv 2 \bmod 7.$

  ii) $3^8 = 6561 \equiv 2 \bmod 7$, since $6561 = 937 \cdot 7 + 2$.

  As we see, it is almost always of computational advantage to apply the modulo reduction as soon as we can.

- The ring Z_m, and thus the integer arithmetic with the modulo operation, is of central importance to modern public-key cryptography. In practice, the integers are represented with 150–2048 bits.

1.6 Simple Blockciphers

Recall:

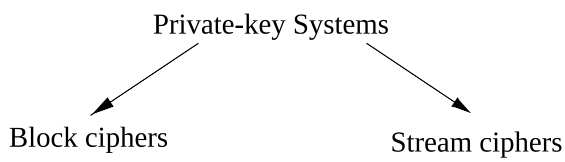

Figure 1.3: Classification of private-key systems

Idea: The message string is divided into blocks (or cells) of equal length that are then encrypted and decrypted.

Input: message string $\bar{X} \to \bar{X} = x_1, x_2, x_3, \ldots, x_n$, where each x_i is one block.

Cipher: $\bar{Y} = y_1, y_2, y_3, \ldots, y_n$; with $y_i = e_k(x_i)$ where the key k is fixed.

1.6.1 Shift Cipher

One of the most simple ciphers where the letters of the alphabet are assigned a number as depicted in Table 1.1.

A 0	B 1	C 2	D 3	E 4	F 5	G 6	H 7	I 8	J 9	K 10	L 11	M 12
N 13	O 14	P 15	Q 16	R 17	S 18	T 19	U 20	V 21	W 22	X 23	Y 24	Z 25

Table 1.1: Shift cipher table

Definition 1.6.1 Shift Cipher

Let $\mathcal{P} = \mathcal{C} = \mathcal{K} = Z_{26}$. $x \in \mathcal{P}$, $y \in \mathcal{C}$, $k \in \mathcal{K}$.

Encryption: $e_k(x) = x + k \bmod 26$.

Decryption: $d_k(y) = y - k \bmod 26$.

Remark:

If $k = 3$ the the shift cipher is given a special name — "Caesar Cipher".

Example:

$k = 17$,

plaintext:

$X = x_1, x_2, \ldots, x_6 = ATTACK.$

$X = x_1, x_2, \ldots, x_6 = 0, 19, 19, 0, 2, 10.$

encryption:

$y_1 = x_1 + k \bmod 26 = 0 + 17 = 17 \bmod 26 = R$

$y_2 = y_3 = 19 + 17 = 36 \equiv 10 \bmod 26 = K$

$y_4 = 17 = R$

$y_5 = 2 + 17 = 19 \bmod 26 = T$

$y_6 = 10 + 17 = 27 \equiv 1 \bmod 26 = B$

ciphertext: $Y \equiv y_1, y_2, \ldots, y_6 =$ R K K R T B.

Attacks on Shift Cipher

1. Ciphertext-only: Try all possible keys ($|k| = 26$). This is known as "brute force attack" or "exhaustive search".

   Secure cryptosystems require a sufficiently large key space. Minimum requirement today is $|K| > 2^{80}$, however for long-term security, $|K| \geq 2^{100}$ is recommended.

2. Same cleartext maps to same ciphertext $\Rightarrow$ can also easily be attacked with letter-frequency analysis.

1.6.2 Affine Cipher

This cipher is an extension of the Shift Cipher ($y_i = x_i + k \mod m$).

Definition 1.6.2 Affine Cipher *Let* $P = C = Z_{26}$.

encryption: $e_k(x) = a \cdot x + b \mod x$.

key: $k = (a, b)$ *where* $a, b \in Z_{26}$.

decryption: $a \cdot x + b = y \mod 26$.

$a \cdot x = (y - b) \mod 26$.

$x = a^{-1} \cdot (y - b) \mod 26$.

restriction: $gcd(a, 26) = 1$ *in order for the affine cipher to work since* a^{-1} *does not always exist.*

Question: How is a^{-1} obtained?

Answer: $a^{-1} \equiv a^{11} \mod 26$ (the proof for this is in Chapter 6)

or by trial-and-error for the time being.