Part I
Introduction and Classical Cryptography

Chapter 1
Introduction

Cryptography and Modern Cryptography

- **Cryptography** : “ the art of writing or solving codes.”

——Concise Oxford English Dictionary

- **Historically accurate**:

Cryptography nowadays encompasses much more than codes.

Art->science.

Military organizations and governments->everyone.

The Setting of Private-Key Encryption

- Security of all classical encryption schemes relied on a secret—a key—shared by the communicating parties in advance and unknown to the eavesdropper.

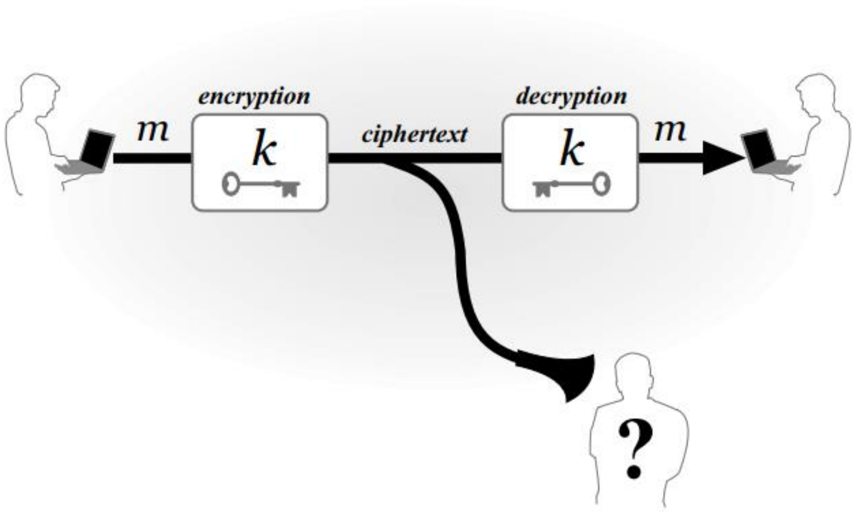

FIGURE 1.1: One common setting of private-key cryptography (here, encryption): two parties share a key that they use to communicate securely

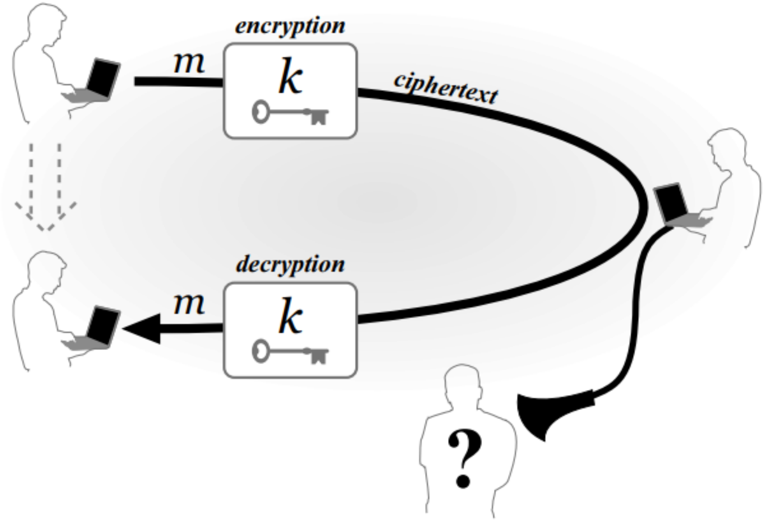

FIGURE 1.2: Another common setting of private-key cryptography (again, encryption): a single user stores data securely over time.

The syntax of encryption

- **a private-key encryption**
- =a message space M e.g.01字符串全体$\{0,1\}^*$
- +a procedure for generating keys (Gen) probabilistic algorithm

  key space K
- +a procedure for encrypting (Enc)

  input : k m output : c $:= Enc_k(m)$
- +a procedure for decrypting (Dec).

  input : k c output : $m := Dec_k(c)$

correctness requirement:

$$Dec_k\left(Enc_k(m)\right) = m$$

Keys and Kerckhoffs' principle

- Should we keep the decryption algorithm Dec secret, too?
- 19th century, Auguste Kerckhoffs: No.

- **Kerckhoffs' principle**:

The cipher method must not be required to be secret, and it must be able to fall into the hands of the enemy without inconvenience.

Three Reasons

- It is significantly easier for the parties to maintain secrecy of a short key than to keep secret the (more complicated) algorithm they are using.
- It will be much easier to change a key than to replace an encryption scheme.
- It is significantly easier for users to all rely on the same encryption algorithm/software (with different keys) than for everyone to use their own custom algorithm.

Caesar's cipher

begin the attack now

+3 -3

EHJLQWKHDWWDFNQRZ

The shift cipher

- M={finite sequences of integers from{0,⋯,25}}

$$Enc_k(m_1 \dots m_l) = c_1 \dots c_l \ \ where\ c_i = [(m_i + k) mod 26]$$
$$Dec_k(c_1 \dots c_l) = m_1 \dots m_l\ where\ m_i = [(c_i - k) mod 26]$$

brute-force/exhaustive-search attack

Sufficient key-space principle

- *Any secure encryption scheme must have a key space that is sufficiently large to make an exhaustive-search attack infeasible.*
- 2^{70}

The mono-alphabetic substitution cipher

- For example:

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
X	E	U	A	D	N	B	K	V	M	R	O	C	Q	F	S	Y	H	W	G	L	Z	I	J	P	T

- K={bijections of the alphabet} $|K| = 26! \approx 2^{88}$
- ~~Brute-force attack.~~

The mono-alphabetic substitution cipher

- Permutation : one to one.

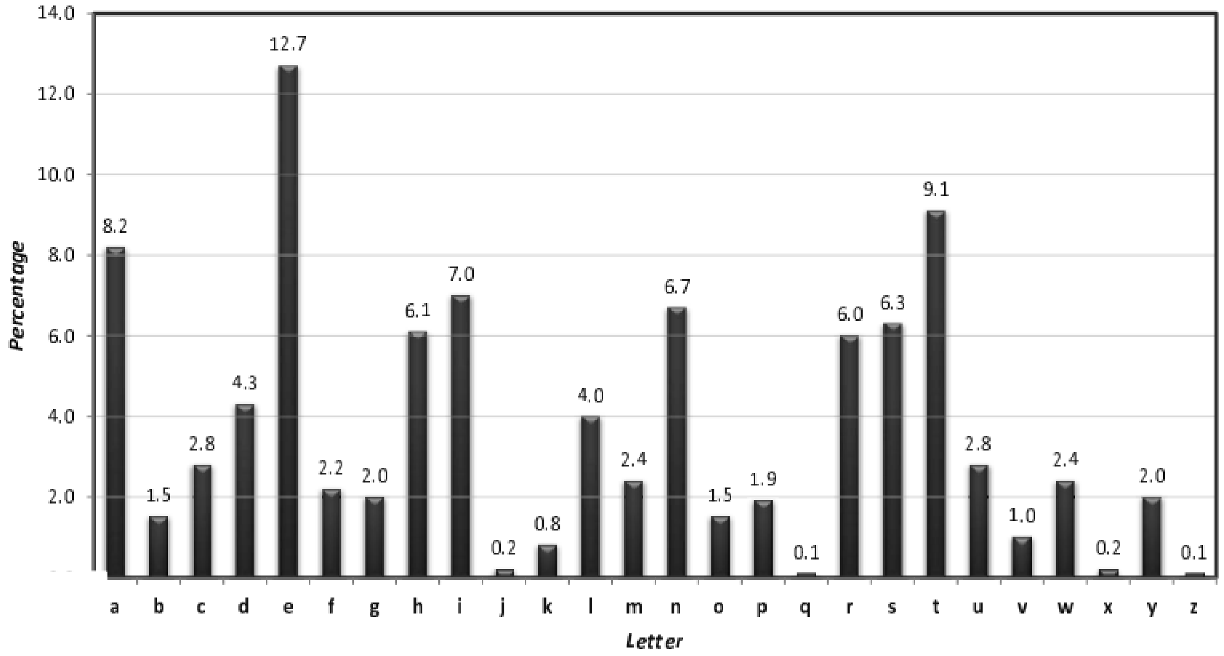

FIGURE 1.3: Average letter frequencies for English-language text.

The shift cipher

- M={finite sequences of integers from{0,⋯,25}}

$$Enc_k(m_1 \dots m_l) = c_1 \dots c_l \;\; where\; c_i = [(m_i + k) mod 26]$$
$$Dec_k(c_1 \dots c_l) = m_1 \dots m_l \; where\; m_i = [(c_i - k) mod 26]$$

brute-force/exhaustive-search attack

An improved attack on the shift cipher

- $p_i\,, 0 \leq p_i \leq 1.$
- $\sum_{i=0}^{25} p_i^2 \approx 0.065$

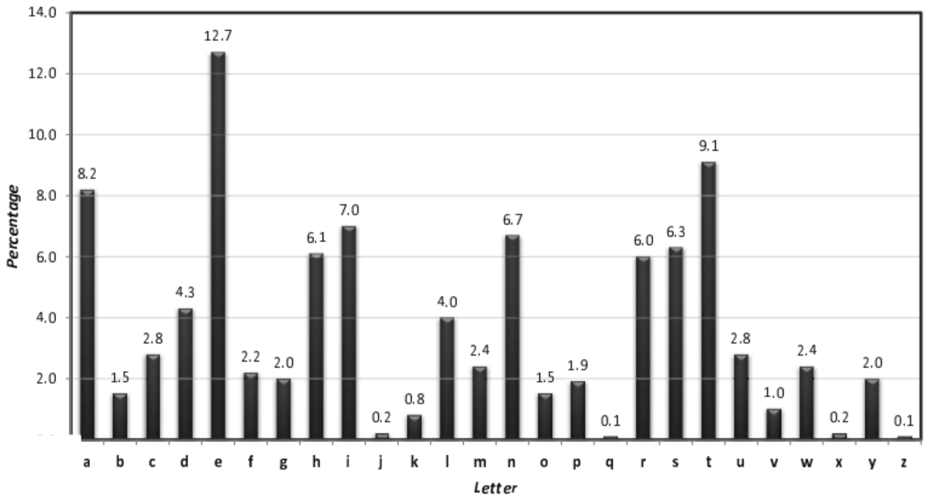

FIGURE 1.3: Average letter frequencies for English-language text.

- $q_i\,, 0 \leq q_i \leq 1.$
- $I_j = \sum_{i=0}^{25} p_i \cdot q_{i+j}$
- $j \in \{0, \ldots, 25\}$

The Vigenere (poly-alphabetic shift) cipher

Plaintext:	`tellhimaboutme`
Key (repeated):	`cafecafecafeca`
Ciphertext:	`VEQPJIREDOZXOE`

The character frequencies of the ciphertext are “smoothed out”!

Principles of Modern Cryptography

- Formal Definitions.
- Art->Science.

- *If you don’t understand what you want to achieve, how can you possibly know when (or if) you have achieved it?*

- “Secure”?

What should a secure encryption scheme guarantee?

- It should be impossible for an attacker to recover the key.
- It should be impossible for an attacker to recover the entire plaintext from the ciphertext.
- It should be impossible for an attacker to recover any character of the plaintext from the ciphertext.

- The “right” answer: regardless of any information an attacker already has, a ciphertext should leak no additional information about the underlying plaintext.

A threat model

- Specifying what "power" the attacker is assumed to have, but does not place any restrictions on the adversary's strategy.

- Ciphertext-only attack

  just a ciphertext (or multiple ciphertexts)
- Known-plaintext attack

  learn one or more plaintext/ciphertext pairs
- Chosen-plaintext attack

  plaintext/ciphertext pairs for plaintexts of its choice
- Chosen-ciphertext attack

  the decryption of ciphertexts of its choice

Precise Assumptions

- Most modern cryptographic constructions cannot be proven secure unconditionally.

- Validation of assumptions.
- Comparison of schemes.
- Understanding the necessary assumptions.

Why not simply assume that the construction itself is secure?

- An assumption that has been tested for several years is preferable to a new.
- There is a general preference for assumptions that are simpler to state.
- Low-level assumptions can typically be used in other constructions.
- Low-level assumptions can provide modularity.

Chapter 2
Perfectly Secret Encryption

“Classical” cryptography

- Before the revolution in cryptography that took place in the mid-1970s and 1980s

- **Perfectly secret**

Provably secure even against an adversary with unbounded computational power.

- Encryption scheme : **M**, Gen, Enc(probabilistic), Dec.
- $\Pr[M = m]$
- $\Pr[M = m | C = c]$
- $\Pr[Enc_K(m) = c | K = k]$

Bayes' Theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Example : Shift Cipher

- **K**={0,⋯,25} $\Pr[K = k] = \frac{1}{26}$
- $\Pr[\mathrm{M} = \mathrm{a}] = 0.7$ and $\Pr[\mathrm{M} = z] = 0.3$

- What is the probability that the ciphertext is B?

Example : Shift Cipher

$$\begin{aligned}\Pr[M = \mathtt{a} \wedge K = 1] &= \Pr[M = \mathtt{a}] \cdot \Pr[K = 1] \\ &= 0.7 \cdot \left(\frac{1}{26}\right).\end{aligned}$$

$$\Pr[M = \mathtt{z} \wedge K = 2] = 0.3 \cdot \left(\tfrac{1}{26}\right)$$

$$\begin{aligned}\Pr[C = \mathtt{B}] &= \Pr[M = \mathtt{a} \wedge K = 1] + \Pr[M = \mathtt{z} \wedge K = 2] \\ &= 0.7 \cdot \left(\frac{1}{26}\right) + 0.3 \cdot \left(\frac{1}{26}\right) \;=\; 1/26.\end{aligned}$$

Example : Shift Cipher

- What is the probability that the message a was encrypted, given that we observe ciphertext B?

$$\begin{aligned}\Pr[M = \mathtt{a} \mid C = \mathtt{B}] &= \frac{\Pr[C = \mathtt{B} \mid M = \mathtt{a}] \cdot \Pr[M = \mathtt{a}]}{\Pr[C = \mathtt{B}]} \\ &= \frac{0.7 \cdot \Pr[C = \mathtt{B} \mid M = \mathtt{a}]}{1/26}.\end{aligned}$$

$$\Pr[C = \mathtt{B} \mid M = \mathtt{a}] = 1/26$$

$$\Pr[M = \mathtt{a} \mid C = \mathtt{B}] = 0.7. = \Pr[M = a]$$

Perfect secrecy

DEFINITION 2.3 *An encryption scheme* (Gen, Enc, Dec) *with message space* $\mathcal{M}$ *is* perfectly secret *if for every probability distribution over* $\mathcal{M}$*, every message* $m \in \mathcal{M}$*, and every ciphertext* $c \in \mathcal{C}$ *for which* $\Pr[C = c] > 0$*:*

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

The ciphertext reveals nothing about the underlying plaintext, and the adversary learns absolutely nothing about the plaintext that was encrypted.

Perfect secrecy

- $\Pr[Enc_K(m) = c] = \Pr[Enc_K(m') = c]$
- The probability distribution of the ciphertext does not depend on the plaintext.

LEMMA 2.4 *An encryption scheme* (Gen, Enc, Dec) *with message space* $\mathcal{M}$ *is perfectly secret if and only if Equation (2.1) holds for every* $m, m' \in \mathcal{M}$ *and every* $c \in \mathcal{C}$.

$$\Pr[C = c \mid M = m] = \Pr[\mathsf{Enc}_K(M) = c \mid M = m] = \Pr[\mathsf{Enc}_K(m) = c],$$

Proof of Lemma 2.4

- Prove : $\Pr[Enc_K(m) = c] = \Pr[Enc_K(m') = c]$

  $\Rightarrow \Pr[M = m|C = c] = \Pr[M = m]$
- Assume $\Pr[M = m] > 0$

$$\begin{aligned}
\Pr[M = m \mid C = c] &= \frac{\Pr[C = c \mid M = m] \cdot \Pr[M = m]}{\Pr[C = c]} \\
&= \frac{\Pr[C = c \mid M = m] \cdot \Pr[M = m]}{\sum_{m' \in \mathcal{M}} \Pr[C = c \mid M = m'] \cdot \Pr[M = m']} \\
&= \frac{\delta_c \cdot \Pr[M = m]}{\sum_{m' \in \mathcal{M}} \delta_c \cdot \Pr[M = m']} \\
&= \frac{\Pr[M = m]}{\sum_{m' \in \mathcal{M}} \Pr[M = m']} \quad = \quad \Pr[M = m],
\end{aligned}$$

Perfect (adversarial) indistinguishability

- Another equivalent definition of perfect secrecy.

The adversarial indistinguishability experiment $\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi}$:

1. *The adversary* $\mathcal{A}$ *outputs a pair of messages* $m_0, m_1 \in \mathcal{M}$.
2. *A key* k *is generated using* Gen, *and a uniform bit* $b \in \{0,1\}$ *is chosen. Ciphertext* $c \leftarrow \mathsf{Enc}_k(m_b)$ *is computed and given to* $\mathcal{A}$. *We refer to* c *as the* challenge ciphertext.
3. $\mathcal{A}$ *outputs a bit* b'.
4. *The output of the experiment is defined to be* 1 *if* $b' = b$, *and* 0 *otherwise. We write* $\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1$ *if the output of the experiment is* 1 *and in this case we say that* $\mathcal{A}$ succeeds.

Perfect (adversarial) indistinguishability

DEFINITION 2.5 *Encryption scheme* $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *with message space* $\mathcal{M}$ *is* perfectly indistinguishable *if for every* $\mathcal{A}$ *it holds that*

$$\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right] = \frac{1}{2}.$$

LEMMA 2.6 *Encryption scheme* Π *is perfectly secret if and only if it is perfectly indistinguishable.*

Example

- Vigenere cipher is not perfectly indistinguishable.
- The period is chosen uniformly in {1,2}.

Adversary $\mathcal{A}$ does:

1. Output $m_0 = \texttt{aa}$ and $m_1 = \texttt{ab}$.
2. Upon receiving the challenge ciphertext $c = c_1c_2$, do the following: if $c_1 = c_2$ output 0; else output 1.

Example

$$\begin{aligned}
&\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right] \\
&= \frac{1}{2} \cdot \Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1 \mid b = 0\right] + \frac{1}{2} \cdot \Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1 \mid b = 1\right] \\
&= \frac{1}{2} \cdot \Pr[\mathcal{A} \text{ outputs } 0 \mid b = 0] + \frac{1}{2} \cdot \Pr[\mathcal{A} \text{ outputs } 1 \mid b = 1],
\end{aligned}$$

Example

(1) a key of period 1 is chosen,

(2) a key of period 2 is chosen, and both characters of the key are equal.

$$\Pr[\mathcal{A} \text{ outputs } 0 \mid b = 0] = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{26} \approx 0.52.$$

Example

- A key of period 2 is chosen and the first character of the key is one more than the second character of the key.

$$\Pr[\mathcal{A} \text{ outputs } 1 \mid b=1] = 1 - \Pr[\mathcal{A} \text{ outputs } 0 \mid b=1] = 1 - \frac{1}{2} \cdot \frac{1}{26} \approx 0.98.$$

$$\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right] = \frac{1}{2} \cdot \left(\frac{1}{2} + \frac{1}{2} \cdot \frac{1}{26} + 1 - \frac{1}{2} \cdot \frac{1}{26}\right) = 0.75 > \frac{1}{2},$$

The One-Time Pad

- 1917, Vernam patented a perfectly secret encryption scheme called the one-time pad.

CONSTRUCTION 2.8

Fix an integer $\ell > 0$. The message space $\mathcal{M}$, key space $\mathcal{K}$, and ciphertext space $\mathcal{C}$ are all equal to $\{0,1\}^\ell$ (the set of all binary strings of length ℓ).

- Gen: the key-generation algorithm chooses a key from $\mathcal{K} = \{0,1\}^\ell$ according to the uniform distribution (i.e., each of the 2^ℓ strings in the space is chosen as the key with probability exactly $2^{-\ell}$).
- Enc: given a key $k \in \{0,1\}^\ell$ and a message $m \in \{0,1\}^\ell$, the encryption algorithm outputs the ciphertext $c := k \oplus m$.
- Dec: given a key $k \in \{0,1\}^\ell$ and a ciphertext $c \in \{0,1\}^\ell$, the decryption algorithm outputs the message $m := k \oplus c$.

The one-time pad encryption scheme.

The One-Time Pad

- **THEOREM 2.9** The one-time pad encryption scheme is perfectly secret.
- **Proof:**

$$\Pr[M = m \mid C = c] = \frac{\Pr[C = c \mid M = m] \cdot \Pr[M = m]}{\Pr[C = c]}$$

The One-Time Pad

$$\begin{aligned}
\Pr[C = c \mid M = m'] = \Pr[\mathsf{Enc}_K(m') = c] &= \Pr[m' \oplus K = c] \\
&= \Pr[K = m' \oplus c] \\
&= 2^{-\ell},
\end{aligned}$$

$$\begin{aligned}
\Pr[C = c] &= \sum_{m' \in \mathcal{M}} \Pr[C = c \mid M = m'] \cdot \Pr[M = m'] \\
&= 2^{-\ell} \cdot \sum_{m' \in \mathcal{M}} \Pr[M = m'] \\
&= 2^{-\ell},
\end{aligned}$$

Limitations of Perfect Secrecy

THEOREM 2.10 *If* (Gen, Enc, Dec) *is a perfectly secret encryption scheme with message space* $\mathcal{M}$ *and key space* $\mathcal{K}$, *then* $|\mathcal{K}| \geq |\mathcal{M}|$.

Proof:

Assume |K|<|M|. Consider the uniform distribution over M and let c ∈ C be a ciphertext that occurs with non-zero probability.

$$\mathcal{M}(c) \stackrel{\text{def}}{=} \{m \mid m = \mathsf{Dec}_k(c) \text{ for some } k \in \mathcal{K}\}.$$

|M(c)| ⩽ |K|<|M|, there is some $m_0 \in M$ such that $m_0 \notin M(c)$.

$$\Pr[M = m' \mid C = c] = 0 \neq \Pr[M = m'],$$