Lecture 2: Classical Ciphers

Substitution ciphers. Transposition cipher.

Nargiza Tazabekova

SDU University

Cryptography course – Fall 2025

Introduction

Classical ciphers are the historical foundations of cryptography. They were designed to ensure secrecy in communication, long before modern computational tools existed.

Key Ideas

- Encryption was based on simple transformations of the alphabet.
- Security relied on keeping the method secret, not only the key.
- Classical ciphers illustrate basic principles that inspired modern cryptography.

Simple Substitution Ciphers

Definition

Let $\mathcal{A}$ be an alphabet of q symbols and $\mathcal{M}$ be the set of all strings of length t over $\mathcal{A}$. Let $\mathcal{K}$ be the set of all permutations on the set $\mathcal{A}$. For each $e \in \mathcal{K}$ define an encryption transformation E_e as:

$$E_e(m) = (e(m_1)e(m_2)\cdots e(m_t)) = (c_1c_2\cdots c_t) = c,$$

where $m = (m_1m_2\cdots m_t) \in \mathcal{M}$.

To decrypt $c = (c_1c_2\cdots c_t)$ compute the inverse permutation $d = e^{-1}$ and

$$D_d(c) = (d(c_1)d(c_2)\cdots d(c_t)) = (m_1m_2\cdots m_t) = m.$$

E_e is called a **simple substitution cipher** or a **mono-alphabetic substitution cipher**.

Substitution Ciphers

Mono-alphabetic substitution cipher

A substitution cipher replaces each symbol of the plaintext with another symbol from the same alphabet.

Substitution Ciphers

Mono-alphabetic substitution cipher

A substitution cipher replaces each symbol of the plaintext with another symbol from the same alphabet.

Example 1

Caesar Cipher

$$E_k(x) = (x + k) \bmod 26$$

with key k. For $k = 3$, $\texttt{HELLO} \mapsto \texttt{KHOOR}$.

Substitution Ciphers

Mono-alphabetic substitution cipher

A substitution cipher replaces each symbol of the plaintext with another symbol from the same alphabet.

Example 1

Caesar Cipher

$$E_k(x) = (x + k) \bmod 26$$

with key k. For $k = 3$, `HELLO` $\mapsto$ `KHOOR`.

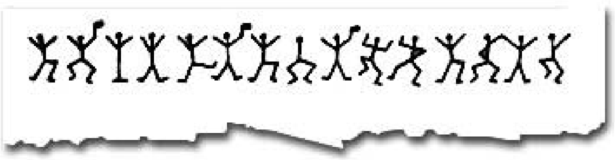
Example 2

General Monoalphabetic Cipher Any permutation of the alphabet can serve as a key. For the English alphabet, the keyspace is $26! \approx 4 \times 10^{26}$.

Substitution Ciphers: Dancing Men Cipher

Arthur Conan Doyle, *The Adventure of the Dancing Men* (1903)

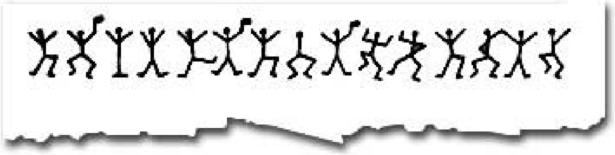
A series of mysterious stick figures were used to encode English letters. Each unique drawing corresponded to a letter of the alphabet.



Substitution Ciphers: Dancing Men Cipher

Arthur Conan Doyle, *The Adventure of the Dancing Men* (1903)

A series of mysterious stick figures were used to encode English letters. Each unique drawing corresponded to a letter of the alphabet.



In the story, Sherlock Holmes identifies repeated patterns and frequencies, then maps symbols to letters. This illustrates the vulnerability of substitution ciphers to frequency analysis.

Substitution Ciphers: Exercise

Ciphertext

19 20 8 8 19 21 15 20 23 19 6 14 21 24 21 7 8 12 21 7 6 6 23 5 24 7
19 6 6 16 21 7 6 4 8 20 24 8 6 8 12 21 15 5 23 7 21 24 8 12 21 1 21
22 21 23 21 8 12 21 17 4 23 20 6 4 1 17 12 20 19 7 23 21 24 6 26
5 19 20 17 21 1 12 21 12 5 7 24 21 25 21 23 1 21 21 24 5 17 5 8
6 23 5 23 5 18 18 20 8 8 12 5 8 7 20 7 24 6 8 1 21 21 11 8 6
18 21 6 24 12 20 1 22 5 10

Substitution Ciphers: Exercise

Ciphertext

19 20 8 8 19 21   15 20 23 19   6 14 21 24 21 7   8 12 21   7 6 6 23   5 24 7
19 6 6 16 21 7   6 4 8   20 24 8 6   8 12 21   15 5 23 7 21 24   8 12 21 1 21
22 21 23 21   8 12 21   17 4 23 20 6 4 1   17 12 20 19 7 23 21 24   6 26
5 19 20 17 21   1 12 21   12 5 7   24 21 25 21 23   1 21 21 24   5   17 5 8
6 23   5   23 5 18 18 20 8   8 12 5 8   7 20 7   24 6 8   1 21 21 11   8 6
18 21   6 24   12 20 1   22 5 10

Decryption

Plaintext (from *Alice's Adventures in Wonderland*):
Little girl opened the door and looked out into the garden. These were the curious children of Alice. She had never seen a cat or a rabbit that did not seem to be on his way.

Substitution Ciphers: Exercise

Ciphertext without spaces

19 20 8 8 19 21 15 20 23 19 6 14 21 24 21 7 8 12 21 7 6 6 23 5 24 7 19 6 6 16 21 7 6 4 8 20 24 8 6 8 12 21 15 5 23 7 21 24 8 12 21 1 21 22 21 23 21 8 12 21 17 4 23 20 6 4 1 17 12 20 19 7 23 21 24 6 26 5 19 20 17 21 1 12 21 12 5 7 24 21 25 21 23 1 21 21 24 5 17 5 8 6 23 5 23 5 18 18 20 8 8 12 5 8 7 20 7 24 6 8 1 21 21 11 8 6 18 21 6 24 12 20 1 22 5 10

Substitution Ciphers: Exercise

Ciphertext without spaces

19 20 8 8 19 21 15 20 23 19 6 14 21 24 21 7 8 12 21 7 6 6 23 5 24 7 19 6 6 16 21 7 6 4 8 20 24 8 6 8 12 21 15 5 23 7 21 24 8 12 21 1 21 22 21 23 21 8 12 21 17 4 23 20 6 4 1 17 12 20 19 7 23 21 24 6 26 5 19 20 17 21 1 12 21 12 5 7 24 21 25 21 23 1 21 21 24 5 17 5 8 6 23 5 23 5 18 18 20 8 8 12 5 8 7 20 7 24 6 8 1 21 21 11 8 6 18 21 6 24 12 20 1 22 5 10

Frequencies of Numbers

Number	1	4	5	6	7	8	10	11	12	14	15	16
Count	6	3	10	13	9	14	1	1	9	1	2	1
Number	17	18	19	20	21	22	23	24	25	26		
Count	4	3	6	9	22	2	9	9	1	1		

English Letter Frequencies

Typical Distribution (in %)

E – 12.7	T – 9.1	A – 8.2	O – 7.5
I – 7.0	N – 6.7	S – 6.3	H – 6.1
R – 6.0	D – 4.3	L – 4.0	C – 2.8
U – 2.8	M – 2.4	W – 2.4	F – 2.2
G – 2.0	Y – 2.0	P – 1.9	B – 1.5
V – 1.0	K – 0.8	J – 0.15	X – 0.15
Q – 0.10	Z – 0.07		

Observation

By comparing ciphertext frequencies with the typical English distribution, one can begin guessing the substitution scheme.

Substitution Ciphers: Main weakness

Even though the keyspace is large, substitution ciphers are vulnerable.

- The frequency distribution of letters is preserved.
- Statistical analysis (e.g., 'E' is most frequent in English) can reveal the substitution.

Homophonic Substitution Ciphers

Definition

To each symbol $a \in \mathcal{A}$, associate a set $H(a)$ of strings of length t, with the restriction that the sets $H(a)$, $a \in \mathcal{A}$, be pairwise disjoint. A **homophonic substitution cipher** replaces each symbol a in a plaintext message block with a randomly chosen string from $H(a)$.
To decrypt a string c of t symbols, one must determine an $a \in \mathcal{A}$ such that $c \in H(a)$. The key for the cipher consists of the sets $H(a)$.

Homophonic Substitution Ciphers

Definition

To each symbol $a \in \mathcal{A}$, associate a set $H(a)$ of strings of length t, with the restriction that the sets $H(a)$, $a \in \mathcal{A}$, be pairwise disjoint. A **homophonic substitution cipher** replaces each symbol a in a plaintext message block with a randomly chosen string from $H(a)$.
To decrypt a string c of t symbols, one must determine an $a \in \mathcal{A}$ such that $c \in H(a)$. The key for the cipher consists of the sets $H(a)$.

Example

Let $\mathcal{A} = \{a, b\}$, with $H(a) = \{00, 10\}$ and $H(b) = \{01, 11\}$.

For messages of length 2, the codomain consists of the following disjoint sets:

$$\begin{aligned} aa &\longmapsto \{0000, 0010, 1000, 1010\}, & ab &\longmapsto \{0001, 0011, 1001, 1011\}, \\ ba &\longmapsto \{0100, 0110, 1100, 1110\}, & bb &\longmapsto \{0101, 0111, 1101, 1111\}. \end{aligned}$$

Polyalphabetic Substitution Ciphers

Definition

A *polyalphabetic substitution cipher* is a block cipher with block length t over an alphabet $\mathcal{A}$ having the following properties:

- The key space $\mathcal{K}$ consists of all ordered sets of t permutations $(p_1, p_2, \ldots, p_t)$, where each permutation p_i is defined on $\mathcal{A}$.
- Encryption of the message $m = (m_1 m_2 \cdots m_t)$ under the key $e = (p_1, p_2, \ldots, p_t)$ is given by

$$E_e(m) = (p_1(m_1)p_2(m_2) \cdots p_t(m_t)).$$

- The decryption key associated with $e = (p_1, p_2, \ldots, p_t)$ is

$$d = (p_1^{-1}, p_2^{-1}, \ldots, p_t^{-1}).$$

Vigenère Cipher: Example 1

Setup

Let $\mathcal{A} = \{A, B, C, \ldots, Z\}$ and $t = 3$. Choose $e = (p_1, p_2, p_3)$ where:

- p_1 maps each letter to the letter 3 positions to its right,
- p_2 maps each letter 7 positions to its right,
- p_3 maps each letter 10 positions to its right.

Vigenère Cipher: Example 1

Setup

Let $\mathcal{A} = \{A, B, C, \ldots, Z\}$ and $t = 3$. Choose $e = (p_1, p_2, p_3)$ where:

- p_1 maps each letter to the letter 3 positions to its right,
- p_2 maps each letter 7 positions to its right,
- p_3 maps each letter 10 positions to its right.

Encryption If

$$m = \text{THI SCI PHE RIS CER TAI NLY NOT SEC URE},$$

then

Vigenère Cipher: Example 1

Setup

Let $\mathcal{A} = \{A, B, C, \ldots, Z\}$ and $t = 3$. Choose $e = (p_1, p_2, p_3)$ where:

- p_1 maps each letter to the letter 3 positions to its right,
- p_2 maps each letter 7 positions to its right,
- p_3 maps each letter 10 positions to its right.

Encryption If

$$m = \text{THI SCI PHE RIS CER TAI NLY NOT SEC URE},$$

then

$$c = E_e(m) = \text{WOS VJS SOO UPC FLB WHS QSI QVD VLM XYO}.$$

Vigenère Cipher: Example 2

The following message was encrypted using the Vigenère cipher with the key **WORD**. Your task is to decrypt it.

Ciphertext

```
ISVW ISRW PVVS WFBJ WHV
```

For decryption: $P_i = (C_i - K_i) \bmod 26$.

Vigenère Cipher: Example 2

The following message was encrypted using the Vigenère cipher with the key **WORD**. Your task is to decrypt it.

Ciphertext

ISVW ISRW PVVS WFBJ WHV

For decryption: $P_i = (C_i - K_i) \bmod 26$.

Ciphertext and Key Alignment

C:	I	S	V	W	I	S	R	W	P	V	V	S	W	F	B	J	W	H	V
K:	W	O	R	D	W	O	R	D	W	O	R	D	W	O	R	D	W	O	R
P:	M	E	E	T	M	E	A	T	T	H	E	P	A	R	K	G	A	T	E

Plaintext

MEET ME AT THE PARK GATE

Vigenère Cipher: Analysis

Why it was considered strong:

- Same letter may be encrypted differently, depending on the key letter.
- Frequency analysis is less straightforward.

Weaknesses:

- Repetition in the key leads to periodic patterns in the ciphertext.
- Methods such as Kasiski's test or index of coincidence reveal key length.

Thus, the Vigenère cipher, though much stronger than Caesar, is still breakable with systematic analysis.

Simple Transposition Cipher

Definition

Consider a symmetric-key block encryption scheme with block length t.
Let $\mathcal{K}$ be the set of all permutations on the set $\{1, 2, \dots, t\}$.
For each $e \in \mathcal{K}$ define the encryption function

$$E_e(m) = (m_{e(1)} m_{e(2)} \cdots m_{e(t)}),$$

where $m = (m_1 m_2 \cdots m_t) \in \mathcal{M}$, the message space.
The set of all such transformations is called a **simple transposition cipher**. The decryption key corresponding to e is the inverse permutation $d = e^{-1}$. To decrypt $c = (c_1 c_2 \cdots c_t)$, compute

$$D_d(c) = (c_{d(1)} c_{d(2)} \cdots c_{d(t)}).$$

Transposition Cipher: Example

Plaintext:

`SECRET MESSAGES ARE HARD TO CRACK`

Choose block size 5 and the key permutation

$$e : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 5 & 2 \end{pmatrix}$$

Encryption: divide into blocks of 5:

`SECRE   TMESS   AGESA   REHAR   DTOCR   ACKXX`

Apply e:

Transposition Cipher: Example

Plaintext:

SECRET MESSAGES ARE HARD TO CRACK

Choose block size 5 and the key permutation

$$e : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 5 & 2 \end{pmatrix}$$

Encryption: divide into blocks of 5:

SECRE TMESS AGESA REHAR DTOCR ACKXX

Apply e:

Ciphertext: CESER STEMS GAESA RARHE TDRCO KACXX

Transition to Modern Cryptography

Classical ciphers illustrate:

- Substitution and permutation as fundamental tools.
- The concept of key-based transformations.
- The necessity of resisting frequency analysis and statistical attacks.

Lesson: Security must depend on the secrecy of the key, not on the secrecy of the algorithm.

Any questions?