

Introduction to cryptography and classical ciphers



CYBER
CHALLENGE.IT



CYBERSECURITY
NATIONAL
LABORATORY

License & Disclaimer

2

License Information

This presentation is licensed under the
Creative Commons BY-NC License



To view a copy of the license, visit:

<http://creativecommons.org/licenses/by-nc/3.0/legalcode>

Disclaimer

- We disclaim any warranties or representations as to the accuracy or completeness of this material.
- Materials are provided “as is” without warranty of any kind, either express or implied, including without limitation, warranties of merchantability, fitness for a particular purpose, and non-infringement.
- Under no circumstances shall we be liable for any loss, damage, liability or expense incurred or suffered which is claimed to have resulted from use of this material.

Outline

3

- Role of cryptography
- The shared key model
- Historical cyphers
- Substitution techniques
- The Vigenère cipher
- Permutation techniques

Introduction

4

- Cryptography is the practice and study of techniques for secure (secret) communication in the presence of third parties called opponents.
- More generally, cryptography concerns the construction and analysis of protocols that prevent third parties from reading or altering private messages.
- Various aspects of information security, such as **confidentiality** and data **integrity**, **authentication** and **non-repudiation**, are at the heart of modern cryptography.

Cryptography is ubiquitous

5

- Passwords hashing
- Secure credit-card transactions over the internet
- Encrypted WiFi
- Disk encryption
- Digitally signed software updates
- Bitcoin
- ...

Some Key terms

6

- ❑ *Encryption algorithm*: transforms a plain text (comprehensible to a human or machine) into an encrypted text (ciphertext, incomprehensible).
- ❑ *Decryption algorithm*: takes the ciphertext and returns the plaintext.
- ❑ Each encryption algorithm must have a correspondent "inverse" decryption algorithm.
- ❑ The pair encryption/decryption algorithms is also called *encryption scheme*.

Cryptographic keys

7

- The encryption algorithm usually takes a *cryptographic key* as input.
- The cryptographic key:
 - is an information (a parameter) that determines the functional output of a cryptographic algorithm.
 - Is essential for determining the transformation of plain text into cipher text and cipher text into plain text.

Cryptography

8

- Two main methods:
 - *Symmetric key* - Single key
 - *Public/Private key* - Double key

Symmetric key cryptography

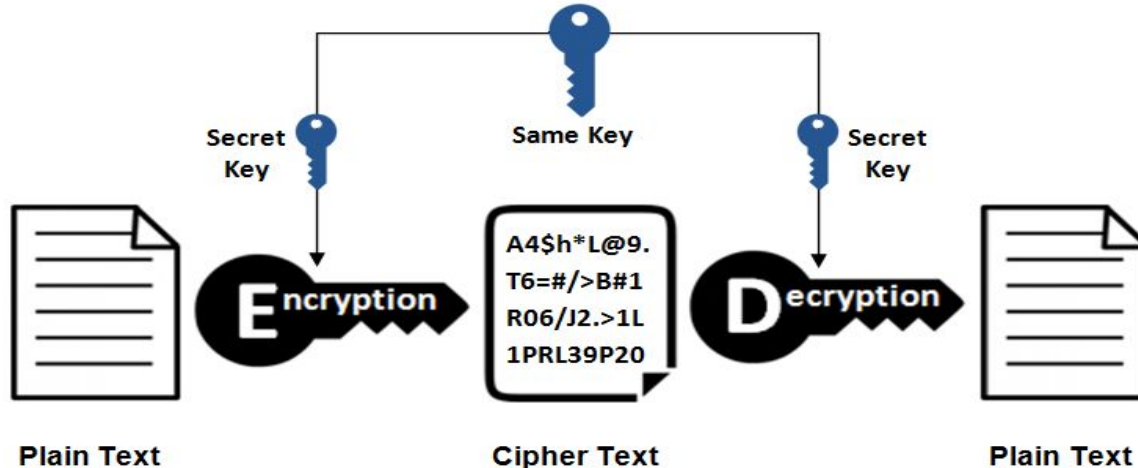
9

- Also called *conventional* or *single key cryptography*.
- Universal technology to guarantee the confidentiality of transmitted or stored data.
- The **only type of encryption** in use before **late 1970s** when double key encryption was introduced.
- It remains the most common of the two types of encryption techniques.

Symmetric key cryptography

10

- Requires that both sender and recipient know the same key.
- An issue is how they do share it without meeting.



Symmetrical cryptographic systems

11

The cryptographic schemes are differentiated by:

- The *kind of operation* to transform the plaintext (P) into ciphertext (C):
 - *Replacements*: each element of P is replaced with another element
 - *Transpositions*: swappings are made between elements of P.
- The *way of “processing” P*:
 - *Block ciphers*: P is processed by the algorithm “one block at a time”.
 - *Flow ciphers*: P is processed considering a single element at a time.

Private-key encryption

12

- A symmetric-key *encryption scheme* is defined by a message space \mathcal{M} and three algorithms (**Gen**, **Enc**, **Dec**):
 - **Gen** (**key-generation algorithm**): outputs $k \in \mathcal{K}$
 - **Enc** (**encryption algorithm**): takes a key k and a message $m \in \mathcal{M}$ as input and outputs the ciphertext c : $c \leftarrow \text{Enc}_k(m)$
 - **Dec** (**decryption algorithm**): takes the key k and the ciphertext c as input and outputs m : $m := \text{Dec}_k(c)$

For all $m \in \mathcal{M}$ and k output by **Gen**, we have: $\text{Dec}_k(\text{Enc}_k(m)) = m$

Cryptanalysis

13

- ▣ *Cryptanalysis* is a set of techniques to test the robustness of the algorithm and of the key by trying to infer the key from the available ciphertext.

More in lecture CR_1.2

Encryption techniques

14

- Two basic mechanisms are applied to plaintext to obtain ciphertext:
 - *substitutions*
 - *permutations*
- The composition of these two techniques is at the basis of all modern symmetric encryption/decryption algorithms.

Substitution Techniques

15

- An encryption technique based on substitution relies on **replacing** each element of the plaintext with another element from the same alphabet.
- Replacements should be **reversible**, i.e., it must be possible to go back to the initial text.
- For each element in the plaintext, the **key determines** the element to be used to **substitute** it in the corresponding ciphertext.

Caesar's cipher

16

- The first known replacement cipher encryption is due to **Julius Caesar**.
- The algorithm requires replacing each letter of the alphabet with the one which is "three hops" away; when the alphabet ends it starts counting again.

| | | | | | | | | | | | | | |
|------------|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PLAINTEXT | a | b | c | d | e | f | g | h | i | j | k | l | m |
| CIPHERTEXT | D | E | F | G | H | I | J | K | L | M | N | O | P |
| PLAINTEXT | n | o | p | q | r | s | t | u | v | w | x | y | z |
| CIPHERTEXT | Q | R | S | T | U | V | W | X | Y | Z | A | B | C |

attack
becomes
dwwdfn

Caesar's cipher

17

□ Another example:

Plaintext: meet me after the toga party

Ciphertext: PHHW PH DIWHU WKH WRJD SDUWB

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|----|----|----|
| A | B | C | D | E | F | G | H | I | J | K | L | M |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |

| | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|----|----|
| N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |

With the association on the right, the number coding of ciphertext C in correspondence of the number p is given by:

$$C = E(3,P) = (P+3) \bmod 26.$$

Modular arithmetic

18

- $x = y \bmod N$ if and only if N divides $x-y$
 - $x \bmod N$ = the remainder when x is divided by N
 - the unique value $y \in \{0, \dots, N-1\}$ such that $x = y \bmod N$
- Examples:
 - $[25 \bmod 10] = [35 \bmod 10]$
 - $25 \neq [35 \bmod 10]$
 - $5 = [35 \bmod 10]$

Shift Ciphers

19

Generalizing Caesar's cipher

- $\mathcal{M} = \{\text{strings over lowercase English alphabet}\}$
- **Gen**: choose uniform $k \in \{0, \dots, 25\}$
- **Enc** _{k} ($m_1 \dots m_t$): output $c_1 \dots c_t$, where
$$c_i := [m_i + k \bmod 26]$$
- **Dec** _{k} ($c_1 \dots c_t$): output $m_1 \dots m_t$, where
$$m_i := [c_i - k \bmod 26]$$

Limits of Shift Ciphers

20

- ❑ From a cryptanalytical point of view, shift ciphers are rather weak, for two main reasons:
 - ❑ There are only 25 possible shifts to try
 - ❑ It is easy to see when the used key is the right one.
- ❑ Shift cyphers can be easily attacked by exhaustive searches.
- ❑ Their weakness is due to the fact that a specific substitution is used (e.g., a shift of 3)

Shift ciphers

21

- ❑ A cipher can be strengthened by allowing **arbitrary substitutions**, by introducing the possibility of not just shifting letters but arbitrarily changing them.
- ❑ For a given set of n elements, we have $n!$ permutations; if $S = \{a, b, c\}$ there are six permutations of S : abc , acb , bac , bca , cab , cba .
- ❑ With shift cipher we have instead only 3 possibility (the string “ abc ” could be mapped only to “ abc ”, “ bca ” and “ cab ”
- ❑ If message are encrypted by moving the plaintext according to one of the (**$26!$**) **possible permutations** it would be (much) more difficult to carry out brute force attacks.

Shift ciphers

22

- Given the following permutation of the 26 letters

ABCDEFGHIJKLMNOPQRSTUVWXYZ

DXUTNAVWKZFQGSIOYJBPLHCERM

- The plaintext «arrivano rinforzi» becomes :

ARRIVANORINFORZI

DJJKHDSIJKSAIJMK

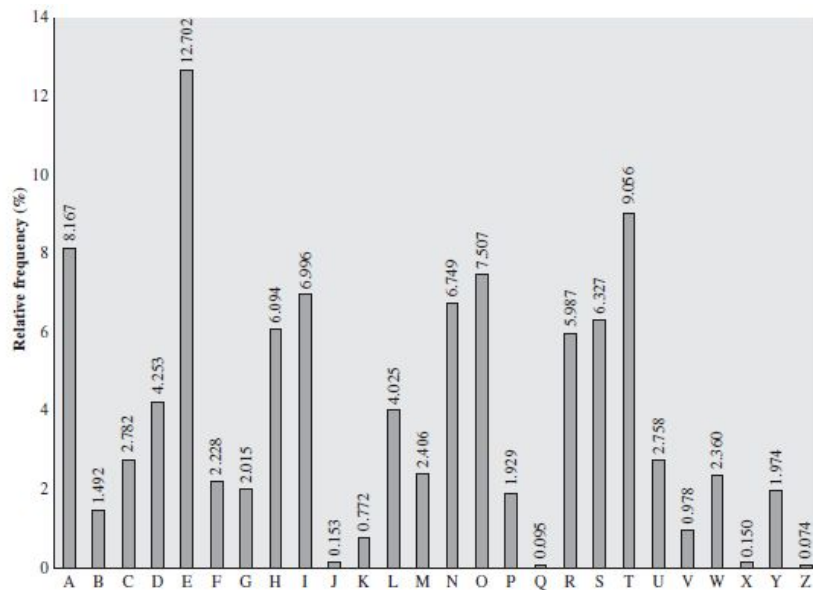
- The chosen permutation would be the key
- A brute force attack would require trying, in the worst case, 26! keys, instead of 25 keys as in the case of shift cipher.

Weakness of shift ciphers

23

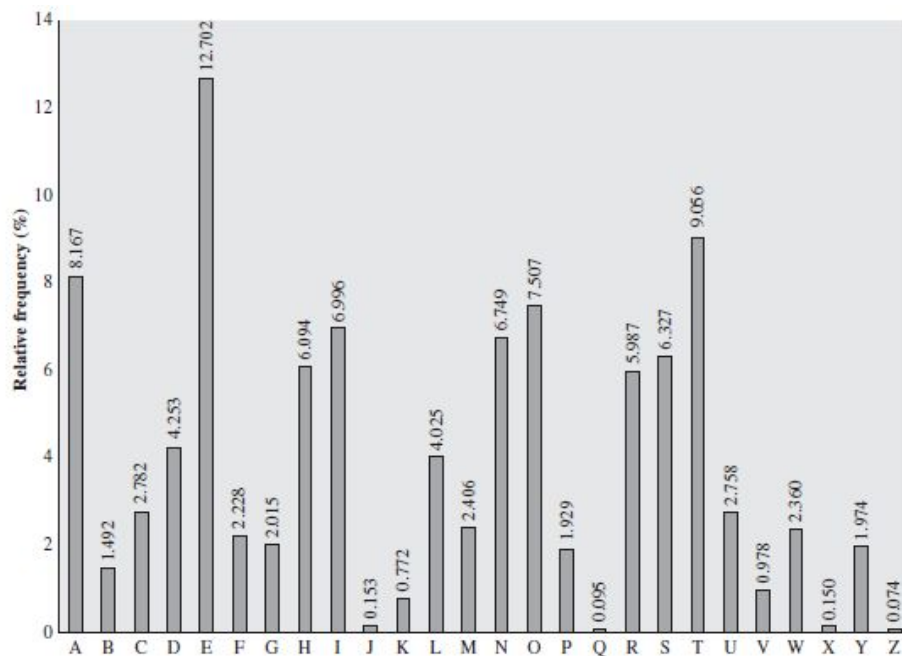
- Shift cipher are **monoalphabetic**:
“same letter of plaintext always produces same letter of ciphertext”
- Regularities of languages such as:
 - frequency of individual cipher letters
 - pairs of letters close together for common words like "of"can be exploited

Letters frequency in English



Letters frequency in English

24



Most frequent letters in languages

25

| ENGLISH | | GERMAN | | FINNISH | | FRENCH | | ITALIAN | | SPANISH | |
|---------|-------|--------|-------|---------|-------|--------|-------|---------|-------|---------|-------|
| | % | | % | | % | | % | | % | | % |
| E | 12.31 | E | 18.46 | A | 12.06 | E | 15.87 | E | 11.79 | E | 13.15 |
| T | 9.59 | N | 11.42 | I | 10.59 | A | 9.42 | A | 11.74 | A | 12.69 |
| A | 8.05 | I | 8.02 | T | 9.76 | I | 8.41 | I | 11.28 | O | 9.49 |
| O | 7.94 | R | 7.14 | N | 8.64 | S | 7.90 | O | 9.83 | S | 7.60 |
| N | 7.19 | S | 7.04 | E | 8.11 | T | 7.26 | N | 6.88 | N | 6.95 |
| I | 7.18 | A | 5.38 | S | 7.83 | N | 7.15 | L | 6.51 | R | 6.25 |

Polyalphabetic cipher

26

- ❑ **Polyalphabetic** ciphers are an alternative to monoalphabetic ones, relying on **multiple substitution alphabets**.
- ❑ Proposed to **disguise** the plaintext **letter frequency analysis**
 - ❑ If P is the most frequent letter in a ciphertext of an English plaintext, one might suspect that P corresponds to E,
 - ❑ This is not possible if **E is enciphered as different ciphertext letters at different points in the message**.
- ❑ The Vigenère cipher is the best-known example of a polyalphabetic cipher.

The Vigenère cipher

27

- The key is now a *string*, not just a character, and to each character is assigned a number according to the position in the alphabet (a:0, b:1, ..., z:25)
- To encrypt, shift each character in the plaintext by the amount dictated by corresponding character of the key
 - Wrap around in the key as needed
- Decryption just reverses the process

te11himaboutme
cafecafecafeca
veqpji redozxoe

the key is cafe

The Vigenère cipher

28

- Substitutions are dictated by the symbol in the array cells determined by the pair $\langle \text{plaintext character}, \text{key character} \rangle$
 - $\langle d, g \rangle$ yields **J**, $\langle a, c \rangle$ yields **C**, $\langle d, a \rangle$ yields **D**, ...

| | | | |
|-----|---------------------------------|--------------|-----------------|
| | abcdefghijklmnopqrstuvwxyz... | | |
| a | ABCDEFGHIJKLMNOPQRSTUVWXYZ.. | key <i>k</i> | gcaigcaigcai... |
| b | BCDEFGHIJKLMNOPQ... | | |
| c | CDEFGHIJKLMNOPQ... | plaintext | dadybeca... |
| d | DEFGHIJKLMNOPQ... | | |
| ... | ... | cyphertext | JCDGHGCI... |
| y | YZABCDEFGHIJKLMNOPQRSTUVWXYZ... | | |
| z | ZABCDEFGHIJKLMNOPQRSTUVWXYZ... | | |

Inspired "*rotor machines*" such as Enigma used during World War II

The Vigenère cipher and its weakness

29

□ Robustness

- If keys are 14-character strings over the English alphabet, then the **key space has size $26^{14} \approx 2^{66}$**
- If variable length keys are used, then the key space is even bigger
- Brute-force search is infeasible

□ Believed secure for many years ...

- One weakness is the **repeating nature of the key**.
- If one guesses the key's length, the cipher text can be treated as interwoven Caesar ciphers, which can be broken individually.

Weaknesses of ENIGMA

30

- ❑ The operators, rather than picking a random key would sometimes pick three consecutive letters from the Enigma keyboard such as **QWE** or **BNM**.
- ❑ The predictable message keys became known as **cillies**, probably the initials of the operator's girlfriend, CIL, that was sometimes used.
- ❑ Before cracking Enigma the hard way, it became routine for the cryptanalysts to **try out the cillies**, and they would be sometimes successful.
- ❑ After building up a vast library of decrypted messages, Turing noticed that many of them conformed to **a rigid structure**.
- ❑ He could then predict part of the contents of an undeciphered message, based on when it was sent and its source:
 - ❑ **Germans sent a regular enciphered weather report shortly after 6 A.M. each day.**
 - ❑ **An encrypted message intercepted at 6:05 A.M. could contain the word "wetter"**

Variant of Vigenère cipher

31

Key: LOVEBIRD

| | | | | | | | | |
|-------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Plaintext: | BIRDS | LOVEW | HEATB | READB | ROWNR | ICEAN | DAWON | DERFU |
| Key : | LOVEB | IRDLO | VEBIR | DLOVE | BIRDL | OVEBI | RDLOV | EBIRD |
| Ciphertext: | MWMHT | TFYPK | CIBBS | UPOYF | SWNQC | WXIBV | UDHCI | HFZWX |

Original: key (LOVEBIRD) is written repeatedly below the plaintext and this is changed as dictated by the key.

Key: LOVEBIRD

| | | | | | | | | |
|-------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Plaintext: | BIRDS | LOVEW | HEATB | READB | ROWNR | ICEAN | DAWON | DERFU |
| Key : | LOVEB | IRDMP | WFCJS | ENQXG | DKTFO | RYHEL | UGPSZ | IFMVH |
| Ciphertext: | MWMHT | TFYQL | DJCT | VRQAH | UYPSF | ZALEY | XGLGM | LJAB |

Variant: the key is also shifted by 1 at each repetition:

LOVEBIRD becomes
MPWFCJS (new key)

Permutation Techniques

32

With permutations techniques, ciphertext is obtained by *changing the order of the letters*.

If the plaintext is **written in diagonal** sequences and then read in line, the message "meet me after the toga party" becomes:

m e m a t r h t g p r y
e t e f e t e o a a t

If read by line, the above message becomes

MEMATRHTGPRY**ETEFETEOAAT**

Rail fence technique

33

Plain text is written downwards and diagonally on successive "rails" of an imaginary fence, moving down till the bottom rail is reached and then up till the top rail is reached, continuing until the whole plaintext is written out.

```
W . . . E . . . C . . . R . . . L . . . T . . . E
. E . R . D . S . O . E . E . F . E . A . O . C .
. . A . . . I . . . V . . . D . . . E . . . N . .
```

WECRLTEERDSOEFEAOCAIVDEN

Other permutation techniques

34

An alternative is to write in a square the message line by line, and read it column by column by setting the order of reading between the columns

Key: 4 3 1 2 5 6 7

Plaintext: a t t a c k p

o s t p o n e

d u n t i l t

w o a m x y z

When columns are read in the order imposed by the key, we have

TTNAAPTMTSUOAODWCOIXKNLYPETZ

Introduction to cryptography and classical ciphers



CYBER
CHALLENGE.IT



CYBERSECURITY
NATIONAL
LABORATORY