



Basic Ciphers

Ahmet Burak Can

Hacettepe University

abc@hacettepe.edu.tr



Basic Terminology in Cryptography – I

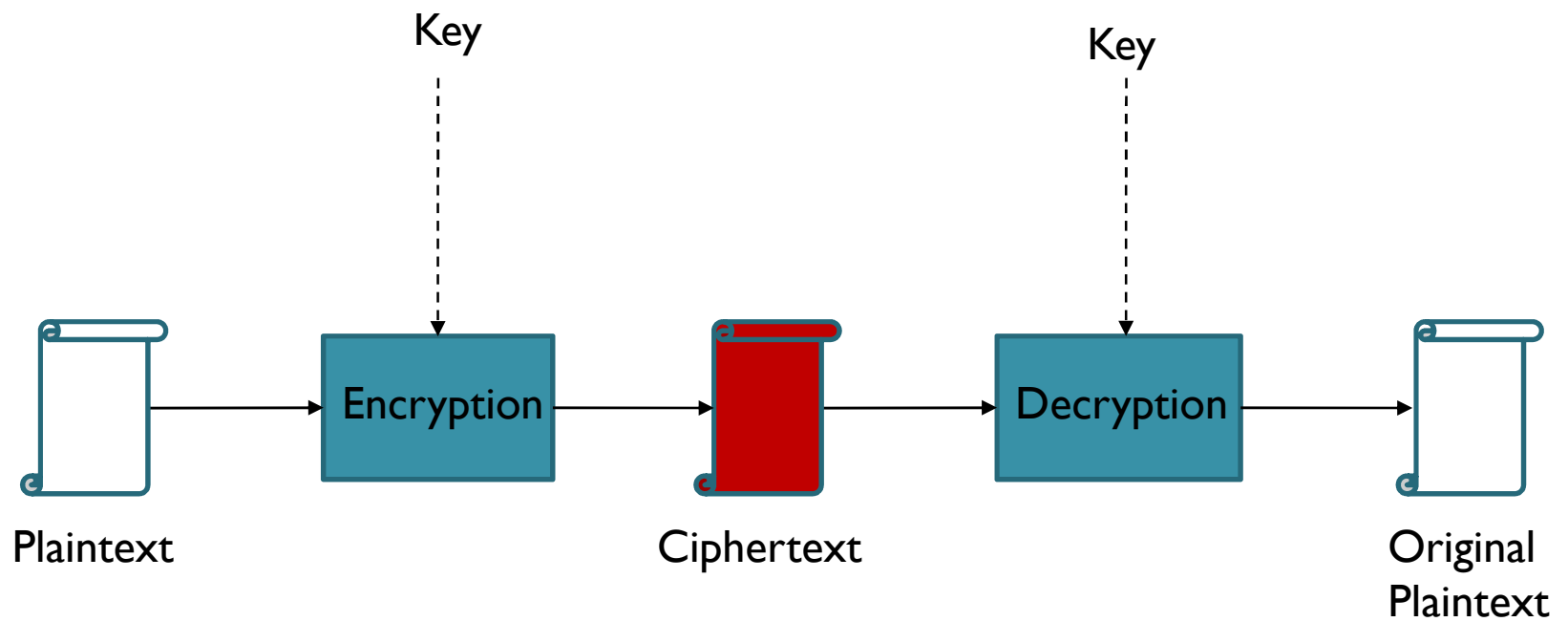
- **Cryptography**: the study of mathematical techniques related to aspects of providing information security services.
- **Cryptanalysis**: the study of mathematical techniques for attempting to defeat information security services.
- **Cryptology**: the study of cryptography and cryptanalysis.



Basic Terminology in Cryptography – 2

- **Encryption (encipherment)**: the process of transforming information (plaintext) using an algorithm (cipher) to make it unreadable to anyone except those possessing special knowledge
- **Decryption (decipherment)**: the process of making the encrypted information readable again
- **Key**: the special knowledge shared between communicating parties
- **Plaintext**: the data to be concealed.
- **Ciphertext**: the result of encryption on the plaintext

Encryption & Decryption





Breaking Ciphers - I

- There are different methods of breaking a cipher, depending on:
 - the type of information available to the attacker
 - the interaction with the cipher machine
 - the computational power available to the attacker

Breaking Ciphers - 2

- **Ciphertext-only attack:** The cryptanalyst knows only the ciphertext. Sometimes the language of the plaintext is also known.
 - The goal is to find the plaintext and the key.
 - Any encryption scheme vulnerable to this type of attack is considered to be **completely insecure**.
- **Known-plaintext attack:** The cryptanalyst knows one or several pairs of ciphertext and the corresponding plaintext.
 - The goal is to find the key used to encrypt these messages or a way to decrypt any new messages that use that key.

Breaking Ciphers - 3

- **Chosen-plaintext attack** : The cryptanalyst can choose a number of messages and obtain the ciphertexts for them
 - The goal is to deduce the key used in the other encrypted messages or decrypt any new messages using that key.
- **Chosen-ciphertext attack**: Similar to the chosen-plaintext attack, but the cryptanalyst can choose a number of ciphertexts and obtain the plaintexts.



Today's Ciphers

- Shift Cipher
- Transposition Cipher
- Mono-alphabetical Substitution Cipher
- Polyalphabetic Substitution Ciphers
- Rotor Machine
- Enigma

Shift Cipher

- A substitution cipher
- The Key Space:
 - [1 .. 25]
- Encryption given a key K:
 - each letter in the plaintext P is replaced with the K'th letter following corresponding number (shift right)
- Decryption given K:
 - shift left
- History: K = 3, Caesar's cipher

Shift Cipher: An 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 |
| 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 |

P = CRYPTOGRAPHYISFUN

K = 11

C = NCJAVZRCCLASJTDQFY

C \rightarrow 2 $2 + 11 \bmod 26 = 13 \rightarrow$ N

R \rightarrow 17 $17 + 11 \bmod 26 = 2 \rightarrow$ C

...

N \rightarrow 13 $13 + 11 \bmod 26 = 24 \rightarrow$ Y



Shift Cipher: Cryptanalysis

- Can an attacker find K?
 - YES: exhaustive search,
 - key space is small (≤ 26 possible keys)
 - the attacker can search all the key space in very short time
- Once K is found, very easy to decrypt

Transposition Cipher

- Write the plaintext horizontally in fixed number columns and read vertically to encrypt.
 - The ancient Spartans used a form of transposition cipher
- Example:
 - P = 'meet me near the clock tower at twelve midnight tonite'

```
m e e t m
e n e a r
t h e c l
o c k t o
w e r a t
t w e l v
e m i d n
i g h t t
o n i t e
```

- C = 'metowteioenhcewmgneekreihitactalddtmrlotvnte'



Transposition Cipher: Cryptanalysis

- Can an attacker decrypt a transposed text?
 - Do exhaustive search on number of columns
 - Since the key space is small, the attacker can search all the key space in very short time
- Once the number of columns is guessed, very easy to decrypt

General Mono-alphabetical Substitution Cipher

- The key space: all permutations of $\Sigma = \{A, B, C, \dots, Z\}$
- Encryption given a key π :
 - each letter X in the plaintext P is replaced with $\pi(X)$
- Decryption given a key π :
 - each letter Y in the ciphertext P is replaced with $\pi^{-1}(Y)$

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

BECAUSE \rightarrow AZDBJSZ



General Substitution Cipher: Cryptanalysis

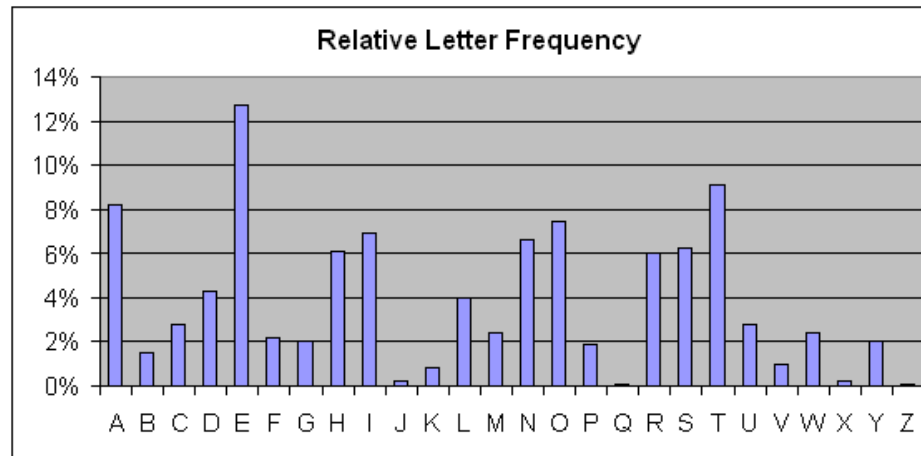
- Exhaustive search is infeasible
 - for the letter A, there are 26 probabilities
 - for the letter B, there are 25 probabilities
 - for the letter C, there are 24 probabilities
 - ... and so on
- Key space size is $26! \approx 4 \times 10^{26}$



Cryptanalysis of Substitution Ciphers: Frequency Analysis

- Basic ideas:
 - Each language has certain features: frequency of letters, or of groups of two or more letters.
 - Substitution ciphers preserve the language features.
 - Substitution ciphers are vulnerable to frequency analysis attacks.
- History of frequency analysis:
 - Earliest known description of frequency analysis is in a book by the ninth-century scientist al-Kindi
 - Rediscovered or introduced from the Arabs in the Europe during the Renaissance

Frequency Features of English



- Vowels, which constitute 40 % of plaintext, are often separated by consonants.
- Letter A is often found in the beginning of a word or second from last.
- Letter I is often third from the end of a word.
- Letter Q is followed only by U
- Some words are more frequent, such as **the, and, at, is, on, in**



Cryptanalysis using Frequency Analysis

- The number of different ciphertext characters or combinations are counted to determine the frequency of usage.
- The cipher text is examined for patterns, repeated series, and common combinations.
- Replace ciphertext characters with possible plaintext equivalents using known language characteristics.
- **Frequency analysis made substitution cipher insecure**

Improve the Security of Substitution Cipher

- Using nulls
 - e.g., using numbers from 1 to 99 as the ciphertext alphabet, some numbers representing nothing are inserted randomly
- Deliberately misspell words
 - e.g., “Thys haz thi ifekkt off diztaughting thi ballans off frikwenseas”
- Homophonic substitution cipher
 - each letter is replaced by a variety of substitutes
- These make frequency analysis more difficult, but not impossible



Summary

- Shift ciphers are easy to break using brute force attacks, they have small key space.
- Substitution ciphers preserve language features and are vulnerable to frequency analysis attacks.



Polyalphabetic Substitution Ciphers

- Main weaknesses of monoalphabetic substitution ciphers
 - each letter in the ciphertext corresponds to only one letter in the plaintext letter
- Idea for a stronger cipher (1460's by Alberti)
 - use more than one cipher alphabet, and switch between them when encrypting different letters
 - Developed into a practical cipher by Vigenère (published in 1586)

The Vigenère Cipher

- **Definition:**

- Given m , a positive integer, $P = C = (\mathbb{Z}_{26})^n$, and $K = (k_1, k_2, \dots, k_m)$ a key, we define:

- **Encryption:**

- $E_k(p_1, p_2 \dots p_m) = (p_1 + k_1, p_2 + k_2 \dots p_m + k_m) \pmod{26}$

- **Decryption:**

- $D_k(c_1, c_2 \dots c_m) = (c_1 - k_1, c_2 - k_2 \dots c_m - k_m) \pmod{26}$

Example:

Plaintext: C R Y P T O G R A P H Y

Key: L U C K L U C K L U C K

Ciphertext: N L A Z E I I B L J J I



Security of Vigenère Cipher

- Vigenere masks the frequency with which a character appears in a language: one letter in the ciphertext corresponds to multiple letters in the plaintext. Makes the use of frequency analysis more difficult.
- Any message encrypted by a Vigenere cipher is a collection of as many shift ciphers as there are letters in the key.

Vigenere Cipher: Cryptanalysis

- Find the **length of the key**.
 - Divide the message into that many shift cipher encryptions.
 - Use frequency analysis to solve the resulting shift ciphers.
- Vigenère cipher is vulnerable: once the key length is found, a cryptanalyst can apply frequency analysis.
- How to Find the Key Length?
 - For Vigenere, as the length of the keyword increases, the letter frequency shows less English-like characteristics and becomes more random.
 - Two methods to find the key length:
 - Kasisky test
 - Index of coincidence (Friedman)

Kasisky Test

- Two identical segments of plaintext will be encrypted to the same ciphertext, if they occur in the text at the distance Δ , ($\Delta \equiv 0 \pmod{m}$), m is the key length).
- Algorithm:
 - Search for pairs of identical segments of length at least 3
 - Record distances between the two segments: $\Delta_1, \Delta_2, \dots$
 - m divides $\gcd(\Delta_1, \Delta_2, \dots)$

| | |
|-----|---|
| PT | T H E S U N A N D T H E M A N I N T H E M O O N |
| Key | K I N G K I N G K I N G K I N G K I N G K I N G |
| CT | D P R Y E V N T N B U K W I A O X B U K W W B T |

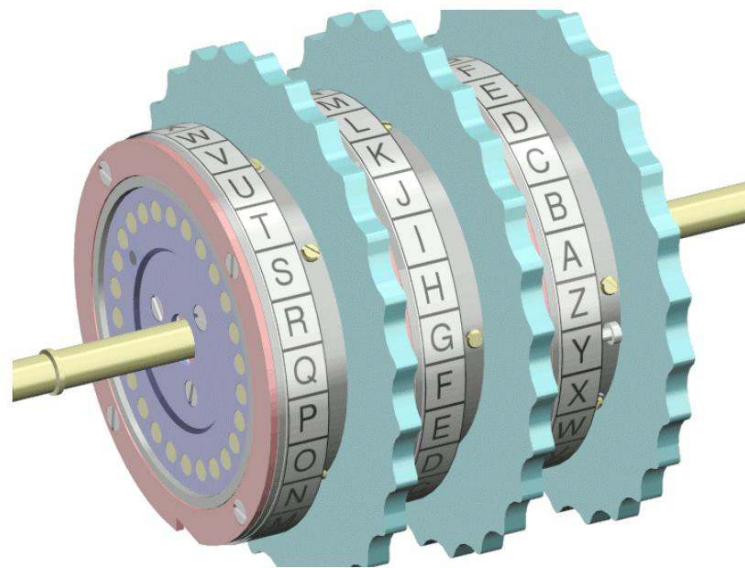


Rotor Machines-I

- Basic idea: if the key in Vigenere cipher is very long, then the attacks won't work
- Implementation idea: multiple rounds of substitution
- A machine consists of multiple cylinders
 - each cylinder has 26 states, at each state it is a substitution cipher: the wiring between the contacts implements a fixed substitution of letters
 - each cylinder rotates to change states according to different schedule changing the substitution

Rotor Machines-2

- A m-cylinder rotor machine has 26^m different substitution ciphers
 - $26^3 = 17576$
 - $26^4 = 456,976$
 - $26^5 = 11,881,376$



Enigma Machine

- Patented by Scherius in 1918
 - Came on the market in 1923, weighted 50 kg (about 110 lbs), later cut down to 12kg (about 26 lbs)
 - It cost about \$30,000 in today's prices
 - 34 x 28 x 15 cm
- Widely used by the Germans from 1926 to the end of second world war
 - First successfully broken by Polish in the thirties by exploiting the repeating of the message key and knowledge of the machine design)
 - During the WW II, Enigma was broken by **Alan Turing** (1912 - 1954) in the UK intelligence. He was an english mathematician, logician and cryptographer, father of modern computer science.

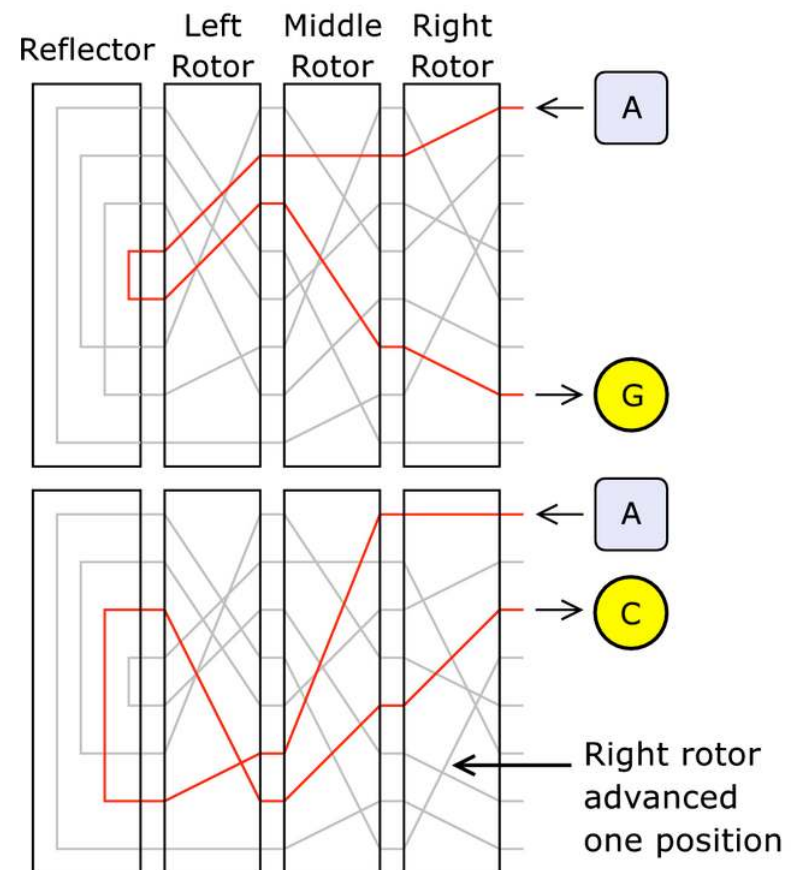
Enigma

- Use 3 scramblers (rotors): 17576 substitutions
- 3 scramblers can be used in any order: 6 combinations
- Plug board: allowed 6 pairs of letters to be swapped before the scramblers process started and after it ended.
- Total number of keys $\approx 10^{16}$
- Later versions use 5 rotors and 10 pairs of letters



Key Mapping

- A reflector enables to map a character twice with each rotor
- First rotor rotates after each key press
- Second rotor rotates after first had a complete revolution,
- and so on



Encrypting with Enigma

- Machine was designed under the assumption that the adversary may get access to the machine
- **Daily key:** The settings for the rotors and plug boards changed daily according to a **codebook** received by all operators
 - A day key has the form
 - Plugboard setting: A/L–P/R–T/D–B/W–K/F–O/Y
 - Scrambler arrangement: 2-3-1
 - Scrambler starting position: Q-C-W
- **Message key:** Each message was encrypted with a unique key defined by the position of the 3 rotors



How to Break the Enigma Machine?

- Recover 3 secrets
 - Internal connections for the 3 rotors
 - Daily keys
 - Message keys
- With 2 months of day keys and Enigma usage instructions, the Polish mathematician Rejewski succeeded to reconstruct the internal wiring



Lessons Learned From Breaking Enigma

- Keeping a machine (i.e., a cipher algorithm) secret does not help
 - The Kerckhoff's principle
 - Security through obscurity doesn't work
- Large number of keys are not sufficient
- Known plaintext attack was easy to mount
- Key management was the weakest link
- People were also the weakest link
- Even a strong cipher, when used incorrectly, can be broken



Kerckhoffs's Principle

- Auguste Kerckhoff (1835 – 1903) was a Dutch linguist and cryptographer who was professor of languages at the School of Higher Commercial Studies in Paris in the late 19th century.
- The security of a protocol should rely only on the secrecy of the keys, protocol designs should be made public (1883)
 - secrecy of a protocol does not work