Lecture 2: Introduction to Cryptography

Stephen Huang

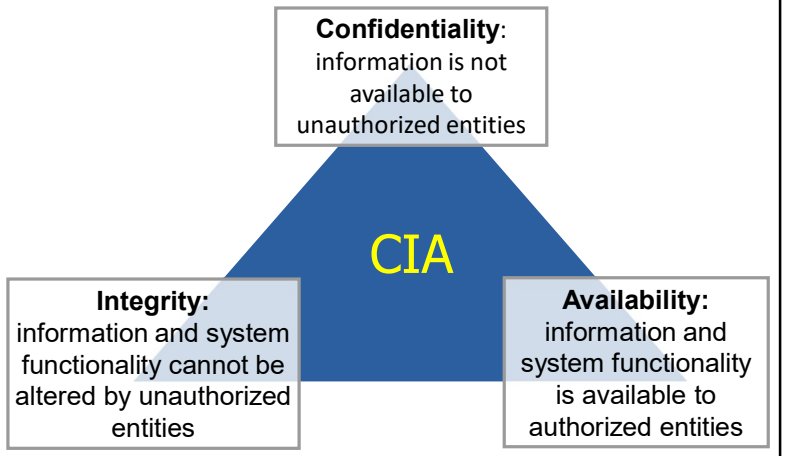
Content

1. Attacker
   - What should we assume about them?
2. Cryptography
   - Classic cryptography (and why it doesn't work),
   - Toward perfect security (but don't get your hopes up; it is not practical)

Cryptography

- This is not a course in cryptography. Cryptography alone can be a course, and probably more.
- Our point here will be to give some intuitions about:
  - what are the fundamental concepts of cryptography;
  - how is it used as a tool for security;
  - how effective is it in that regard?
- Nevertheless, it is a significant part of the course due to its importance.

Reminder: Security Objectives



CIA

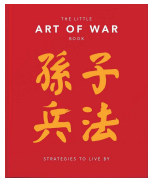
The CIA Triad		
What Is the CIA?		
Confidentiality	**Integrity**	**Availability**
The information is safe from accidental or intentional disclosure.	The information is safe from accidental or intentional modification or alteration.	The information is available to authorized users when needed.
Example		
I send you a message, and no one else knows what that message is.	I send you a message, and you receive exactly what I sent you (without any modification)	I send you a message, and you are able to receive it.
What's The Purpose of the CIA?		
Data is not disclosed	Data is not tampered	Data is available
How Can You Achieve the CIA?		
e.g., Encryption	e.g., Hashing, Digital signatures	e.g., Backups, redundant systems
Opposite of CIA		
Disclosure	Alteration	Destruction

https://preview.redd.it/xegh56kbrk751.png

1. Attacker

- "It is said that if you know your enemies and know yourself, you will not be imperiled in a hundred battles;
- if you do not know your enemies but do know yourself, you will win one and lose one;
- if you do not know your enemies nor yourself, you will be imperiled in every single battle."

— Sun Tzu (~500 BC), Art of War



Is it secure?

Is this secure?

Or is this?

It depends…

Attacker Model

- We can define security only with respect to an attacker model
  - What the attacker can do (Capability)
  - What the attacker knows (Information)
  - What the attacker wants to achieve (Objective)
  - …
- *Example:* none of the commonly used cryptographic primitives can withstand an attack with unlimited computational power

Attacker Types

Type	Motivation	Capabilities	Objective
"Script-kiddies"	reputation gain	techniques and software developed by others	integrity (*e.g.*, website defacement) or availability
Cybercriminals	financial gain	expertise/investment in finding/purchasing vulnerabilities, infrastructure to support their operations	confidentiality (*e.g.*, stealing financial information) or integrity (*e.g.*, ransomware)
Industrial espionage	information gain	targeted attacks, ample resources	confidentiality (*e.g.*, learning trade secret)
Cyberwarfare	information gain or causing damage	heavy investments in multiple area of security, targeted attacks	confidentiality (*e.g.*, espionage) or integrity/availability (*e.g.*, sabotage)

Attacker Model

- We can define security only with respect to an attacker model
  - What the attacker can do (Capability)
  - What the attacker knows (Information)
  - What the attacker wants to achieve (Objective)
  - …
- *Example:* none of the commonly used cryptographic primitives can withstand an attack with unlimited computational power
- It is generally better to overestimate the attacker's capabilities, knowledge, and determination than to underestimate them.

Examples of Attacker Capabilities

- Communications security: access to communication channel

passive (*e.g.*, eavesdropping, traffic analysis)

active (*e.g.*, tampering, replaying, interrupting, masquerading)

Examples of Attacker Capabilities

- System security*:* access to a system
  - **remote**: without any prior authorization (*e.g.*, via Internet or other public network)
  - **local**: with some prior authorization (*e.g.*, using an unprivileged user account)
  - **physical access**

Safe Assumptions for Knowledge

Attacker May Know
- algorithms
- system design
- implementation
- configuration
- …

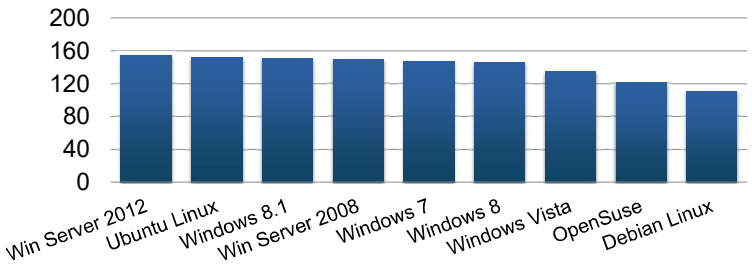
Attacker Cannot Know
- truly random values

algorithms?
system design?
implementation?
configuration?
true random values?

Security by Obscurity

- Providing security by keeping the design or implementation of a system secret.
- Security experts, researchers, standards bodies, etc., generally **reject** this idea
- Obscurity may slow down an attack but cannot stop it
  - if we think of an idea, an attacker might also think of it
  - an attacker may try its attack for many possible design and implementation choices
- A false sense of security may be very dangerous

Security by Obscurity



number of vulnerabilities in the *National Vulnerability Database* for 2015

Bug Bounty

Platform Programs

Program Name	Start Date	Last Updated	End Date	Eligible Entries	Bounty Range
Microsoft Hyper-V	2017-05-31	2020-04-13	Ongoing	Critical remote code execution, information disclosure and denial of services vulnerabilities in Hyper-V	Up to $250,000 USD
Microsoft Windows Insider Preview	2017-07-26	2020-08-27	Ongoing	Critical and important vulnerabilities in Windows Insider Preview	Up to $100,000 USD
Microsoft Applications and On-Premises Servers	2021-03-24	2022-04-05	Ongoing	Critical and important vulnerabilities in Microsoft Applications and On-Premises Servers	Up to $30,000 USD
Windows Defender Application Guard	2017-07-26	2017-07-26	Ongoing	Critical vulnerabilities in Windows Defender Application Guard	Up to $30,000 USD
Microsoft Edge (Chromium-based)	2019-08-20	2021-10-21	Ongoing	Critical, important, and moderate vulnerabilities in Microsoft Edge (Chromium-based) Dev, Beta, and Stable channels	Up to $30,000 USD
Microsoft 365 Insider	2017-03-15	2023-01-20	Ongoing	Vulnerabilities on Microsoft 365 Insider	Up to $15,000 USD

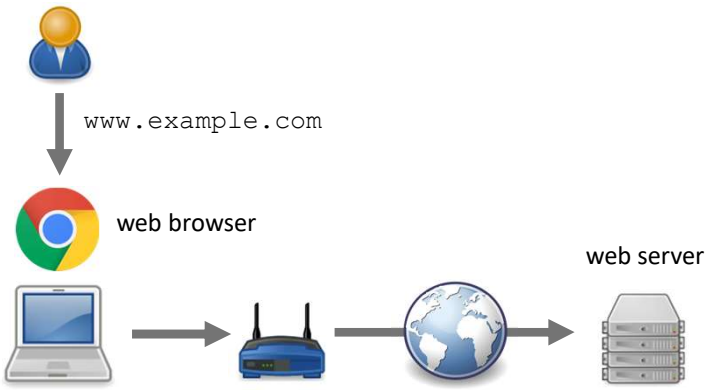
Transparency

- In some cases, openness and transparency can actually improve security.
- By allowing independent security researchers to analyze and scrutinize the system, potential vulnerabilities can be identified and addressed before they are exploited by malicious actors.

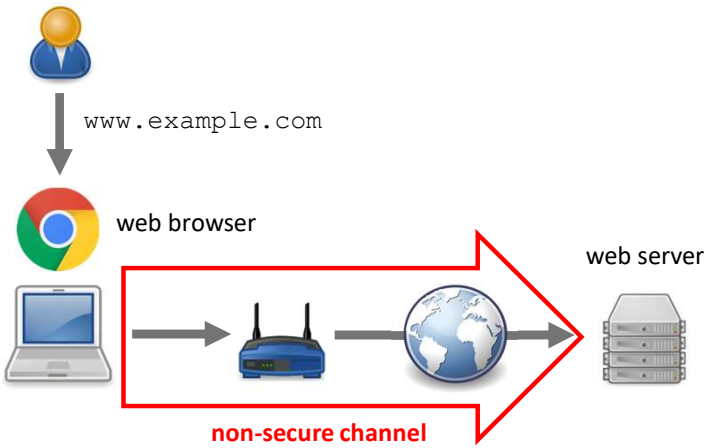
2. Cryptography

- Secure communication in the presence of adversaries.

Communications Security

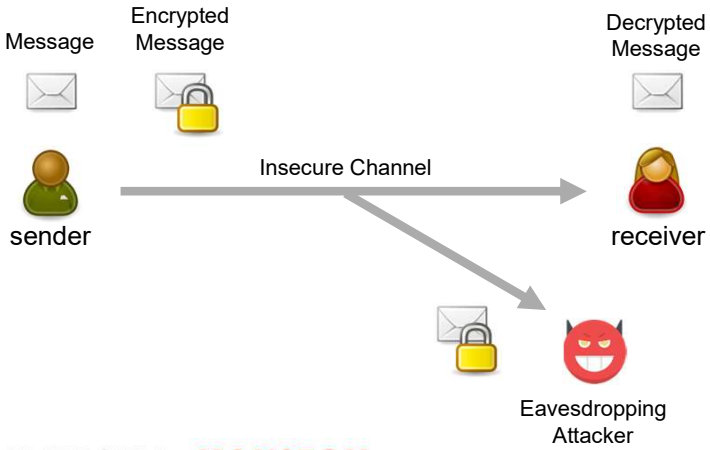


Communication Security



Communication Security

Communication Security



Cryptography

- Etymology

  crypto + graphy
  κρυπτός + γράφειν
  "secret" + "writing"

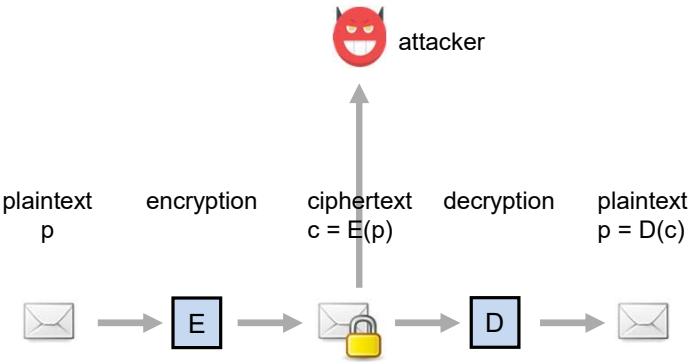
- Traditionally: confidentiality → encryption

- Nowadays
  - integrity → message authentication
  - non-repudiation → digital signatures

Encryption

- Basic model
  - p: plaintext (or P)
  - E: encryption algorithm/cipher
  - c: ciphertext (or C)
  - $D = E^{-1}$: decryption algorithm/cipher
- Attacker's goal: recover the plaintext for a given ciphertext.
- Cryptography means the practice of using encryption to conceal text.
- Cryptanalysis is the attempt to extract the meaning of encrypted messages.

Encryption



Keyed Encryption

- Often, the encryption and decryption algorithms use a key K. The key selects a specific algorithm from the family of algorithms defined by E.
- We write this dependence as:

$$c = E(p, K_E) \text{ and } p = D(c, K_D)$$

- If $K_E = K_D$, then the algorithm is called symmetric. If not, then it is called asymmetric. In general,

$$p = D(E(p, K_E), K_D)$$

- An algorithm that does not use a key is called a keyless cipher.

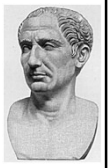
Classical Cryptography

Caesar Cipher

- Named after Julius Caesar, the first recorded user of the scheme.
- Shift each letter of the plaintext by three letters down the alphabet:

```
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
d e f g h i j k l m n o p q r s t u v w x y z a b c
```

  - *plaintext:* "meet me after the toga party"
  - *ciphertext:* "phhw ph diwhu wkh wrjd sduwb"



Caesar Cipher

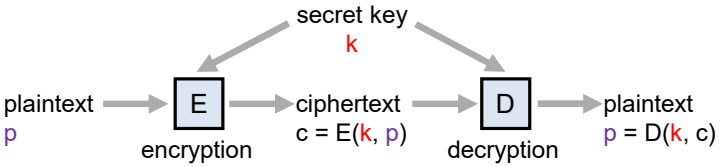
- Formally
  - assign an integer to each letter:

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

  - encryption: c = E(p) = p + 3 mod 26
  - decryption: p = D(c) = c - 3 mod 26
- Security by obscurity: confidentiality of the plaintext relies on the attacker not knowing the algorithm

Cryptographic Key

- Sender and receiver use a random secret key k
- Secret key k: chosen at random and known only by the sender and receiver.



- Attacker's goal: recover the key or some plaintext.

Types of Attack

- Ciphertext only: an attacker knows the algorithms and the given ciphertext.
- Known plaintext: an attacker also has one or more plaintext-ciphertext pairs.
- Chosen ciphertext: an attacker can also choose one or more ciphertexts (but not the given one) and obtain the corresponding plaintexts.
- Chosen plaintext: an attacker can also choose one or more plaintexts and obtain the corresponding ciphertexts.

Generalized Caesar Cipher

- Secret key: integer k randomly chosen from [1, 25]
  - encryption: c = E(k, p) = p + k mod 26
  - decryption: p = D(k, c) = c - k mod 26
- Example: with k = 5
  - plaintext: "meet me after the toga party"
  - encryption: c = p + 5 mod 26
  - decryption: c = p - 5 mod 26
  - ciphertext: "rjjy rj fkyjw ymf ytlf ufwyd"

Brute-Force Attack

- Brute-force attack: attacker tries every possible key on a given ciphertext until finding a correct translation into plaintext
- Known plaintext attack: search for k until a match p = D(k, c) is found
- Ciphertext only attack: attacker must be able to recognize the correct plaintext to find the key
  - Attacker knows that plaintext is an HTTP request

```
GET / HTTP/1.1
Host: …
Connection: keep-alive
User-Agent: …
```

  - May be more difficult for fragments of compressed plaintexts
- On average, half of all possible keys must be tried to achieve success

Affine Cipher

- Modular arithmetic
  - operations: addition + and multiplication ·, which "wrap around"
  - additive inverse -x: 0 = x + (-x) mod m
    (*example:* -5 = 21 modulo 26 since 5 + (-5) = 5 + 21 = 0 mod 26)
  - multiplicative inverse x^{-1}: 1 = x · x^{-1} mod m
    - exists only if x and m are coprime (i.e., their greatest common divisor is 1)
      (*example:* only 1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 25 are invertible modulo 26)
- Secret key: k_1 from {1, 3, …, 11, 15, …, 23, 25} and k_2 from [0, 25]
  - encryption: $c = p \cdot k_1 + k_2 \mod 26$
  - decryption: $p = (c - k_2) \cdot k_1\text{-}1 \mod 26$
  - number of possible keys = 12 · 26 = 312

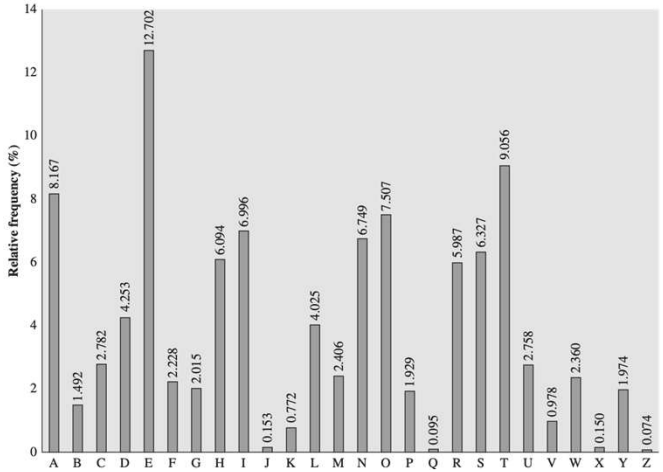
Substitution Cipher

- Secret key: permutation over the alphabet

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

  - Plaintext: "meet me after the toga party"
  - Ciphertext: "taah ta qdhaf hza hlxq wqfhm"
- Number of possible keys = 26! ≈ $4 \cdot 10^{26}$ ≈ 2^{88}
  - Relatively strong against brute-force attacks
- Vulnerable to known plaintext attacks

Cryptanalysis



Cryptanalysis

- Breaking a substitution cipher: Attacker relies on the nature of the algorithm and knowledge of the general characteristics of the plaintext
- Ciphertext:
  - `TAAHTAQDHAFHZAHLXQWQFHM`
- frequency:
  - A: 5 → E
  - H: 5 → T
  - Q: 3 → A
  - F: 2 → R
  - …
- Plaintext: "`meet me after the toga party`"

Monoalphabetic Cipher

- `TAAH TA QDHAF HZA HLXQ WQFHM`
- `TeeH Te QDHeF Hze HLXQ WQFHM`
- `Teet Te QDteF tZe tLXQ WQFtM`
- `Teet Te aDteF tZe tLXa WaFtM`
- `Teet Te aDter tZe tLXa WartM`

Playfair Cipher

- Invented in 1854 by Charles Wheatstone, but named after Lord Playfair, who promoted the use of the cipher
  - used by British forces in the Second Boer War and World War I
- Secret key: 5 x 5 table filled with letters
  - keyword + remaining letters of the alphabet

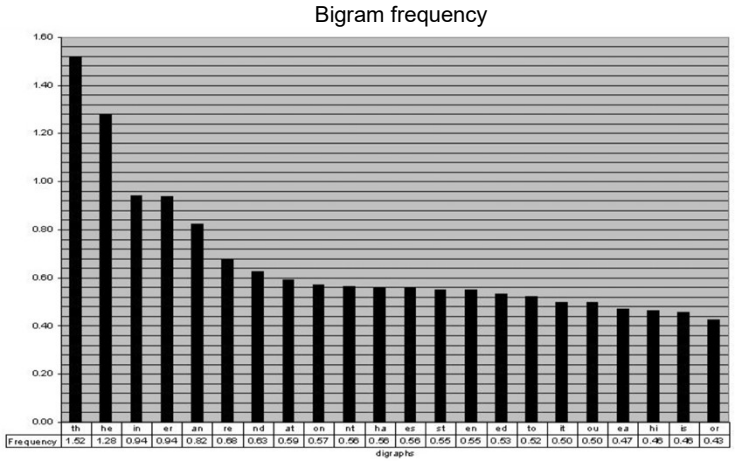
M	O	N	A	R
C	H	Y	B	D
E	F	G	I/J	K
L	P	Q	S	T
U	V	W	X	Z

Playfair Cipher

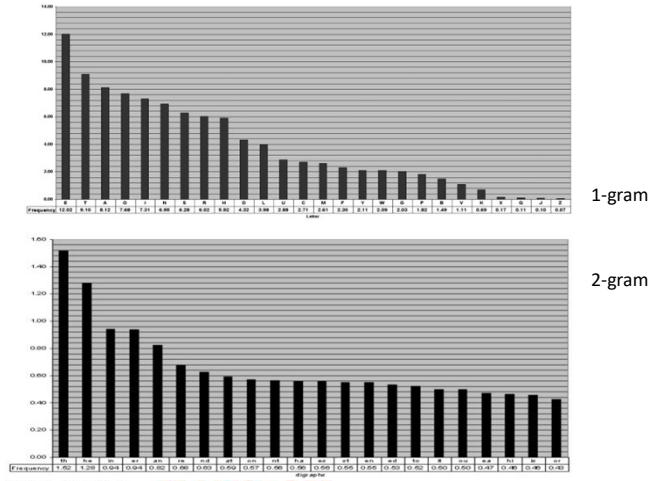
- Plaintext is encrypted two letters at a time
  - repeating plaintext letters that are in the same pair are separated by a filler (*e.g.*, X)
  - if the pair is on the same row/column, shift them right/down (with the rows/columns circularly following each other)
  - finally, replace each letter in a pair with the letter that lies in the same row but the column of the other letter (*e.g.*, plaintext pair "BP" becomes ciphertext pair "HS")

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

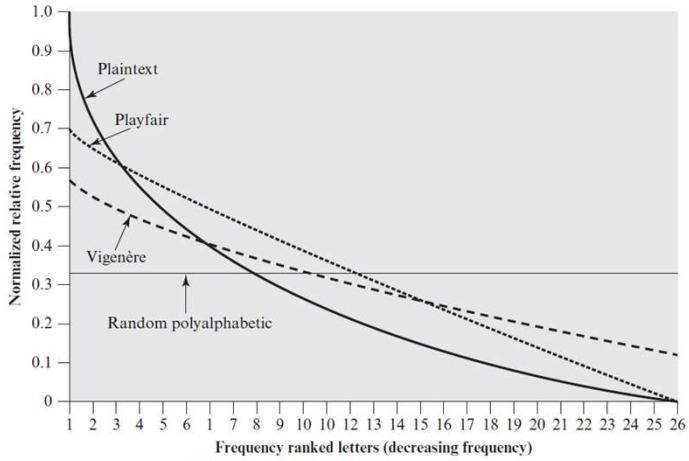
Cryptanalysis for Multiple-Letter Ciphers



Cryptanalysis



Relative Frequency of Letters



Vigenère Cipher

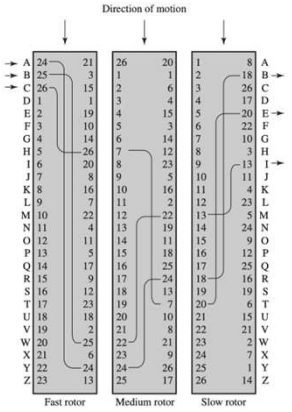
- This is the simplest polyalphabetic cipher
- The set of related monoalphabetic substitution rules consists of the 26 Caesar ciphers with shifts of 0 through 25.
- Reinvented many times through history
  - originally described by Giovan Battista Bellaso in 1553, but named after Blaise de Vigenère
  - used by, for example, the Confederate States of America in the American Civil War
  - in 1917, Scientific American characterized this system as "impossible of translation"
- Key: letters k_1, …, k_N (each corresponds to a number from [0, 25])

Vigenère Cipher

- Encryption
  - repeat the N letters of the key so that it is as long as the plaintext
  - ith letter of ciphertext: $c_i = p_i + k_i \bmod 26$
  - *example:*
    - key: `DECEPTIVEDECEPTIVEDECEPTIVE`
    - plaintext: `wearediscoveredsaveyourself`
    - ciphertext: `ZICVTWQNGRZGVTWAVZHCQYGLMGJ`
- Vulnerable to cryptanalysis

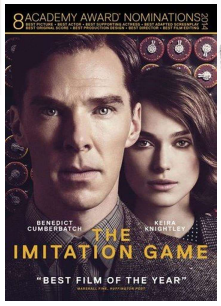
Rotor Machines

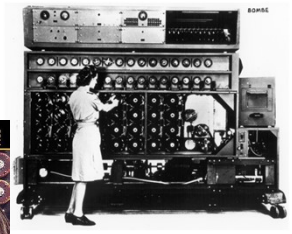
- Machines based on the rotor principle were used by both Germany (Enigma) and Japan (Purple) in World War II
- British and Polish cryptologists developed electromechanical devices, called bombes, to break Enigma



Rotor Machines

Three-rotor Enigma





US Navy Bombe

Next Topic

- Intro. to Cryptography
- Stream Ciphers
- Stream and Block Ciphers