

Cryptography

Problems Addressed by Cryptography

- Key Agreement
- Secure Communication

Key Agreement

- Exchanging a key among two parties to access confidential message

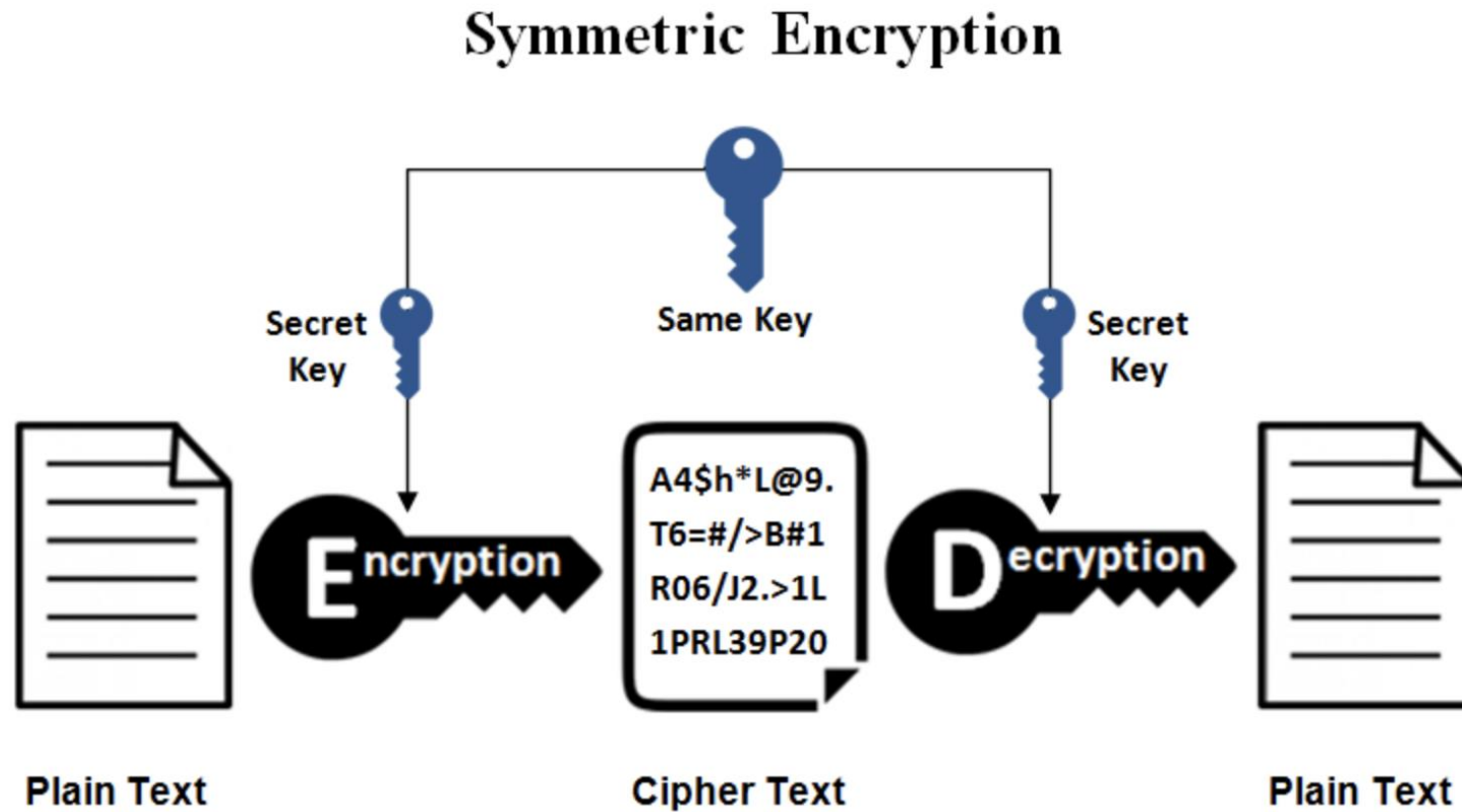
Secure Communication

- There is challenge to maintain
 - Confidentiality
 - Integrity

Types of Cryptographic Primitives

- Symmetric Key primitives (private-key cryptography)
 - Same Key (k) used at both ends
 - Computationally efficient
 - Key Agreement a Big issue
- Asymmetric Key primitives (public-key cryptography)
 - Different Key (k) used at both the ends
 - Computationally inefficient
 - No key agreement required

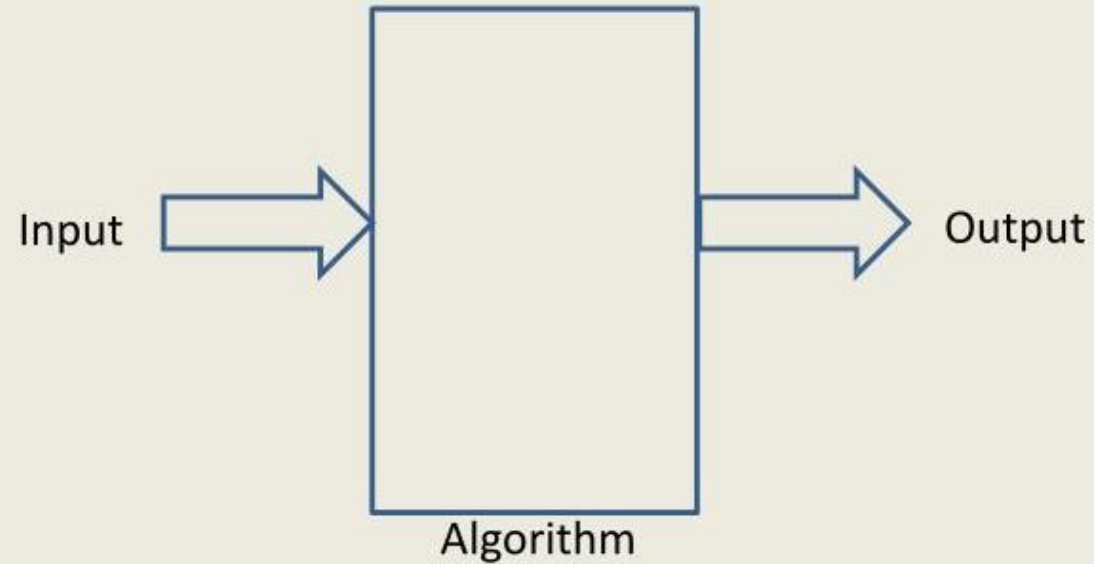
Symmetric Key Cryptography



Cryptographic Algorithms

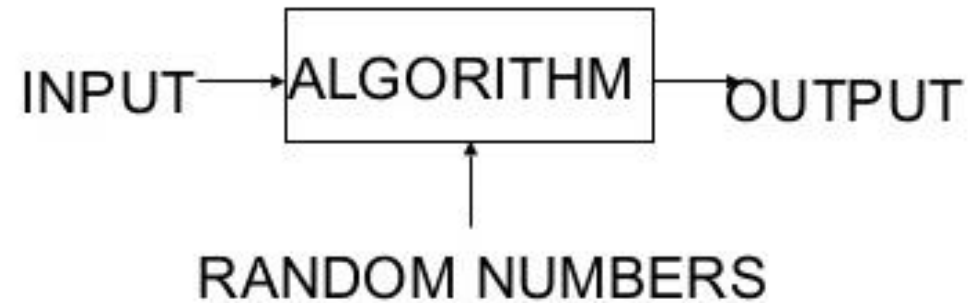
- Deterministic
- Randomized

Deterministic Algorithm



- The **output** as well as the **running time** are functions only of the input.

Randomized Algorithms



- In addition to input, algorithm takes a source of random numbers and makes random choices during execution;
- Behavior can vary even on a fixed input;

Algorithms in Symmetric Key Cryptography

- GEN
 - To Generate a randomized key (K)
- ENC
 - To convert message (plaintext) to Ciphertext using key (K)
- DEC
 - To convert Ciphertext to message (plaintext) using key (K)

Properties expected from a Secure Cipher

- Correctness

- $DEC_k(\underline{ENC_k(m)}) := m$

- Privacy

- Ciphertext **c** should not reveal anything about **m**

Challenges to formalize Privacy

- Definition 1

- An Encryption process is secure if the ciphertext does not reveal the underlying key.
- Consider an ENC algorithm which always outputs plaintext as the ciphertext.

- Definition 2

- An Encryption process is secure if the ciphertext does not reveal the underlying plaintext.
- Consider an ENC algorithm where the first 1% of the ciphertext is the same as the first 1% of the plaintext.

Challenges to formalize Privacy Definition

- Definition 3

- An Encryption process is secure **if the ciphertext does not reveal any character of the underlying plaintext.**
- Consider an ENC algorithm where the ciphertext reveals whether the underlying plaintext is less than or greater than a certain value.

- Definition 4

- An Encryption process is secure **if the ciphertext does not reveal any meaningful information about the underlying plaintext.**
- The notion of the meaningful information varies from application to application.

Challenges to formalize Privacy Definition

- Definition 5

- An Encryption process is secure **if the ciphertext does not help to compute any function of the underlying plaintext.**
- Precisely what we expect from a secure cipher. But there are certain loopholes in the above definition.
 - How to formalize whether a given ciphertext helps to compute a given function of the underlying plaintext?
 - What is the underlying adversary / attack model?
 - Is the adversary passive or malicious?
 - Does the adversary have access to any kind of “additional Information”?

Attack Models

- Ciphertext Only Attack (COA)
- Known plaintext Attack (KPA)
- Chosen plaintext Attack (CPA)
- Chosen Ciphertext Attack (CCA)

Note: In all attack models, the goal of the adversary is to compute some function of the underlying plaintext from the ciphertext.

Ciphertext Only Attack (COA)

- Simplest possible attack
- Attacker have access to the ciphertext
- No other additional knowledge is available other than the ciphertext.

Known Plaintext Attack (KPA)

- Attack has access to several (plaintext, ciphertext) pairs under the same k
 - E.g. the first word in an email is usually “hello” / “dear”, etc

Chosen Plaintext Attack (CPA)

- Attacker gets “encryption oracle” service --- active attack
 - Gets **encryption** of the plain-texts of **its choice**, without the knowledge of sender / receiver

Chosen Ciphertext Attack (CCA)

- It is the strongest possible attack model
- Attacker gets “encryption oracle” plus “decryption oracle” service.
 - Gets decryption of ciphertexts of its choice.

Keys and Kerckhoffs' Principle

- To main security, key should be definitely a secret
- What about ENC and DEC algorithm?
 - More security by keeping them private too?
- Kerckhoffs' Principle:
 - “A cryptosystem should be secure even if everything about the system, except the key, is a public knowledge”.

Importance of Kerckhoffs' Principle

- Maintaining the privacy of a key is relatively easier.
 - Key size \approx 100 bits, Program size : 1000 times larger
 - Algorithms can be leaked, reverse engineered
- Easy to replace a key if the key is exposed.
- Infeasible to assign a secret pair of algorithm for every pair of parties.
- Published designs undergo public scrutiny and so likely to be more secure.

Note: **Dangerous to use a Proprietary Encryption Scheme**

Classical Cypher & their Cryptoanalysis

Classical Cipher

- Shift cipher
- Mono-alphabetic substitute cipher
- Poly-alphabetic substitution (Vigenere) cipher

Shift Cipher

- Plain-text and cipher-text character $\in \{a, \dots, z\}$
 - Encryption: shift each plain-text character by k positions “forward”
 - Decryption: shift each cipher-text character by k positions “backward”
 - $k \in \{0, \dots, 25\}$ and randomly selected by the key-generation algorithm
- Mathematical interpretation of the shift cipher
 - Interpret the set $\{a, \dots, z\}$ as $\{0, \dots, 25\}$
 - $K = \{0, \dots, 25\}$ and $M = C = \text{set of strings over } \{0, \dots, 25\}$
- GEN
 - $k \in_R K$
- ENC
 - Input: $m_i \in M$ and k
 - Process: $c_i := (m_i + k) \bmod 26$
 - Output: $c_i \in C$
- DEC
 - Input: $c_i \in C$ and k
 - Process: $m_i := (c_i - k) \bmod 26$
 - Output: $m_i \in M$

Cryptoanalysis of Shift Cipher

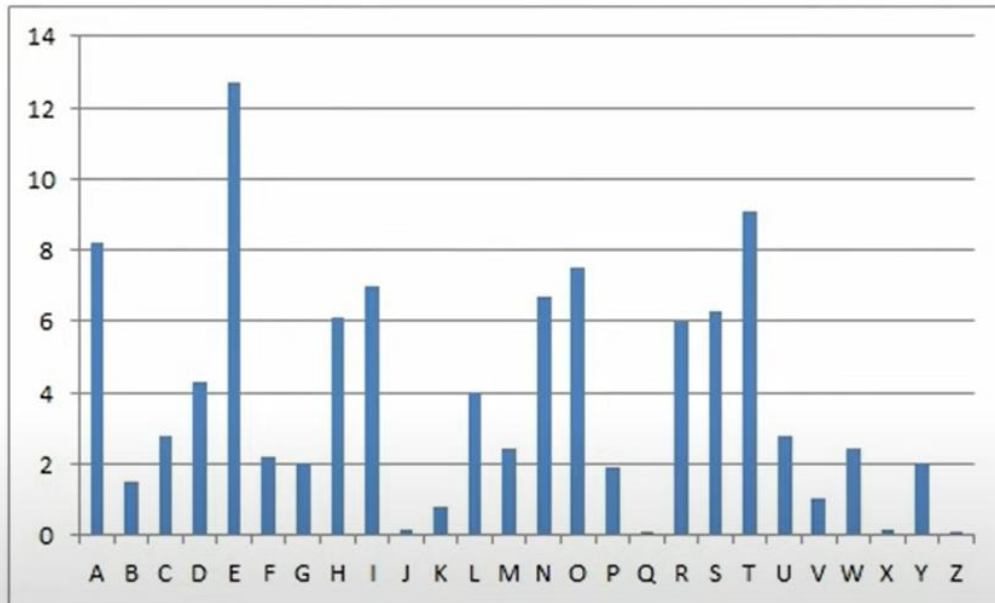
- Plaintext: $m = (m_1, \dots, m_l)$
- Ciphertext: $c_i = (c_1, \dots, c_l)$
- Attack model: Ciphertext Only Attack (COA)
- Information known to attacker
 - Ciphertext
 - Process through which the ciphertext is generated i.e. $c_i := (m_i + k) \bmod 26$
- Attack
 - An attacker can try to decrypt c *with all possible k* (brute force)
 - Easy to mount: only 26 candidate keys
- Learning
 - **Sufficient key-space principle:** Any secure cipher must have a key-space that is not vulnerable to exhaustive search.

Mono-alphabetic Substitution Cipher

- Map each plain-text character to an arbitrary cipher-text character in a one-to-one fashion
 - Key: A secret permutation (determined by the key generation algorithm)
- Is mono-alphabetic substitution cipher secure?
 - Brute-force attack is impractical
 - Number of candidate keys = $(26!) \approx 2^{88}$

Cryptoanalysis of Mono-alphabetic substitution Cipher

- Frequency Analysis: applicable when plaintext space is a natural language.
 - Idea: exploit the redundancy present in the underlying natural language.



Average English letter frequency

Bigram	Percentage	Bigram	Percentage
TH	3.15	HE	2.51
AN	1.72	IN	1.69
ER	1.54	RE	1.48
ES	1.45	ON	1.45
EA	1.31	TI	1.28
AT	1.24	ST	1.21
EN	1.20	ND	1.18

Average English
bigram frequency

THE, ING, AND, HER, ERE, ENT, THA, NTH,
WAS, ETH, FOR

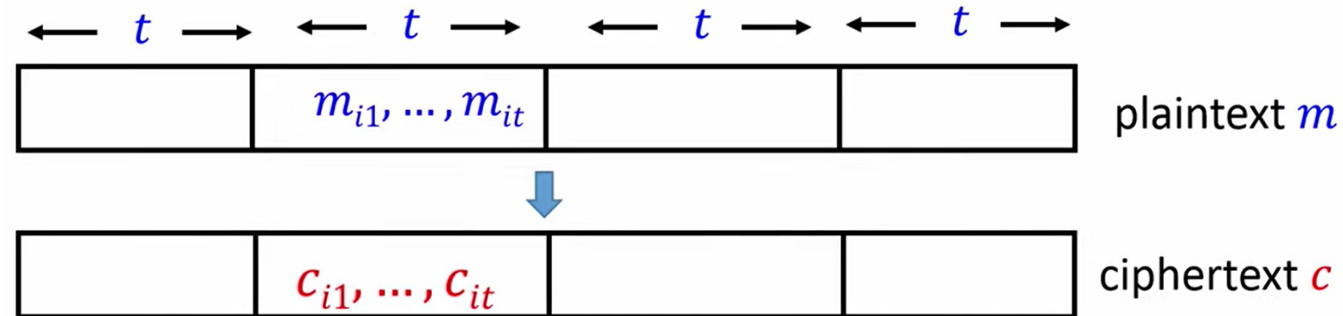
Popular English trigram in
decreasing order

Cryptoanalysis of Mono-alphabetic substitution Cipher

- Most frequently occurring character/ bigram/ trigram in the ciphertext, corresponds to most frequently occurring character / bigram/ trigram in the plaintext

Poly-alphabetic Substitution (Vignere) Cipher

- Idea: invoke multiple instances of shift cipher
 - In each instance, a plain-text character is mapped to a different ciphertext character.
- $M = C = \{0, \dots, 25\}^*$ $K = \{0, \dots, 25\}^t$, t randomly chosen by GEN
- Key-generation algorithm: output a uniformly random key $k = (k_1, \dots, k_t)$
- Encryption:



$$c_{ij} = (m_{ij} + k_j) \bmod 26$$

Example: Vigenere Cipher

- Key $k = CIPHER = (2,8,15,7,4,17)$ $t = 6$
- Plaintext $m = thiscryptosystemisnotsecure$

← 6 →						← 6 →						← 6 →						← 6 →						← 3 →		
19 7 8 18 2 17						24 15 19 14 18 24						18 19 4 12 8 18						13 14 19 18 4 2						20 17 4		
+ mod 26						+ mod 26						+ mod 26						+ mod 26						+ mod 26		
2 8 15 7 4 17						2 8 15 7 4 17						2 8 15 7 4 17						2 8 15 7 4 17						2 8 15		
21 15 23 25 6						8 0 23 8 21 22 15						20 1 19 19 12 9						15 22 8 25 8 19						22 25 19		

Cryptoanalysis of Vigenere Cipher

- Two stage approach
 - stage 1: Determine the length t of the unknown key k
 - Kasiski's method, index of coincidence method
 - Stage 2: Try to determine the characters k_1, k_2, \dots, k_t of the key k
- Stage II
 - t independent instances of letter frequency analysis

Learnings from cryptoanalysis of Classical Ciphers

- Can be broken by launching a ciphertext-only attack
 - They can be even badly broken through stronger attack models
- Sufficient Key-space principle
 - Key space should be sufficiently large to make brute-force attack infeasible
- Designing secure cipher a tough task
-

Classical vs Modern Cryptography

- Classical cryptography was an art
 - No scientific foundation --- end result : disaster
- Modern Cryptography
 - Strong scientific foundations and principles
- Principle 1
 - Formal security definition
- Principle 2
 - Precisely stating any assumption used in construction
- Principle 3
 - Rigorous proof of security

OSI Security Architecture

An International Standard

Introduction

- Security architecture for OSI offers a systematic way of defining security requirements and characterizing the approaches to achieve these requirements.
- It was developed as an international standard.

Need for OSI Security Architecture

- To assess the security needs, of an organization effectively and choose various security products and policies.
- The need for some systematic way of defining the requirements for security and characterizing the approaches to satisfy those requirements.
- This is difficult enough in a centralized data-processing environment, and with the use of local area and wide area network, the problems are compounded.

The OSI security Architecture

- Such a systematic approach is defined by ITU-T (The International Telecommunication Union – Telecommunication Standardization Sector)
- It is a United Nation (UN) sponsored agency that develops standards, called recommendations, relating to telecommunication and to Open System Interconnection (OSI) recommendations X.800, security Architecture for OSI.

Benefits

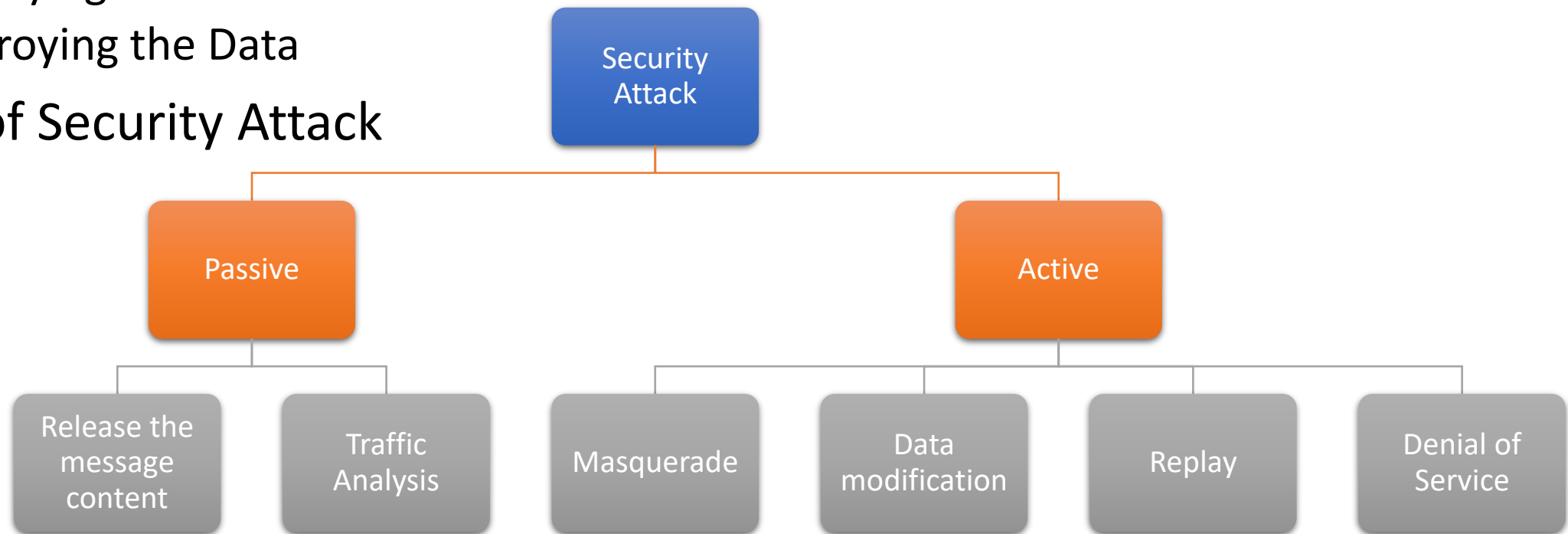
- The OSI security architecture is useful to managers as way of organization the task of providing security
- Furthermore, because this architecture was developed as international standards, computer and communications vendors have developed security feature for their products and services that relate to this structured definition of services and mechanisms.

OSI Security Architecture focus

- Security Attack
 - Any action that compromise the security of information owned by an organization.
- Security Mechanism
 - A process that is designed to detect, prevent or recover from a security attack
 - And security mechanism is a method which is used to protect your message from unauthorized entity.
- Security Services
 - Security Services is the services to implement security policies and implemented by security mechanism.

Security Attack

- Security attack is a process of gaining an access of data by unauthorized user
 - Accessing the Data
 - Modifying the Data
 - Destroying the Data
- Types of Security Attack



Security Mechanism



Encipherment

Digital Signature

Traffic padding

Notarization

Security Services

- Confidentiality
 - Ensures that the information in a computer system and transmitted information are accessible only for reading by authorized parties.
- Authentication
 - Ensures that the origin of a message or electronic document is correctly identified, with an assurance that the identity is not false
- Integrity
 - Ensures that only authorized parties are able to modify computer system assets and transmitted information
- Non-repudiation
 - Requires that neither the sender nor the receiver of a message be able to deny transmission.
- Access control
 - Requires that access to information resources may be controlled by or for the target system.
- Availability
 - Requires that computer system assets be available to authorized parties when needed.

Perfect Security

Introduction

- In 1949, Claude Shannon published a paper entitled “**Communication Theory of Secrecy Systems**” in the Bell Systems Technical Journal.
- This paper had a great influence on the scientific study of cryptography.
- Claude Shannon is also known as father of information theory.

Perfect Secrecy

- The notion of perfect security is also called as unconditional security, information –theoretic security.
- The attack model considered in the definition of perfect security is **Ciphertext-only attack**
- It is assumed that the attacker is computationally unbounded
- Informal definition
 - “irrespective of any prior info”. The attack has about m , the cipher-text c should not leak **no additional information** about the plain-text.

Basic Definition & properties

- $\Pi = (GEN, ENC, DEC)$, A publicly known cipher
- For the attacker, Π induces a probability distribution on \mathcal{M}, \mathcal{K} and \mathcal{C}
 - \mathcal{M} : *Plaintext space*
 - \mathcal{K} : *Key space*
 - \mathcal{C} : *Ciphertext space*
- Probability distribution on \mathcal{K} : *induced by GEN*
 - Almost always a uniform distribution
 - $\Pr[\mathbf{K} = k]$: *probability that GEN output the key k* \rightarrow typically $\frac{1}{|\mathcal{K}|}$


Basic Definition & properties

- Probability distribution on \mathcal{M} : *induced by any prior information about the underlying plaintext*
 - Ex: plaintext can be “Attack” with prob 0.7 or “Retreat” with prob. 0.3
 - $\Pr[\mathbf{M} = m]$: *probability that underlying plaintext is m*
- Probability distribution on \mathcal{C} : determined by the probability distribution over M, K and the steps of ENC.
 - $\Pr[\mathbf{C} = c]$: probability that ENC outputs the ciphertext c


Formal Definition

- AN encryption scheme (GEN,ENC,DEC) over a plaintext space \mathcal{M} is perfectly-secure if for every probability distribution over \mathcal{M} and \mathcal{K} every plaintext $m \in \mathcal{M}$ and every ciphertext $c \in \mathcal{C}$, the following holds:

- $$\Pr [\mathbf{M} = m \mid \mathbf{C} = c] = \Pr [\mathbf{M} = m]$$



Posteriori probability that m
is encrypted in c



a-priori probability that m
might be communicated

Observing the cipher-text c **does not change** the
attacker's knowledge about the distribution of plaintext

Alternate Definitions

- Original definition: Probability of knowing a plain-text remains the same before and after seeing the cipher-text.

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

- Alternate Definition:

- For every probability distribution over \mathcal{M} and \mathcal{K} , every plain-text $m_0, m_1 \in \mathcal{M}$ and every cipher-text $c \in \mathcal{C}$, the following holds

$$\Pr[\mathbf{C} = c \mid \mathbf{M} = m_0] = \Pr [\mathbf{C} = c \mid \mathbf{M} = m_1]$$

- Meaning: probability distribution of cipher-text is independent of plain-text

Proof that both definition are similar

$$\text{If } \Pr[\mathbf{M} = m \mid \mathbf{C} = c] = \Pr[\mathbf{M} = m], \quad \forall m \in \mathcal{M}, c \in \mathcal{C}$$

$$\text{Then } \Pr[\mathbf{C} = c \mid \mathbf{M} = m_0] = \Pr[\mathbf{C} = c \mid \mathbf{M} = m_1], \quad \forall m_0, m_1 \in \mathcal{M}, c \in \mathcal{C}$$

Proof: Let $m_0, m_1 \in \mathcal{M}, c \in \mathcal{C}$ be arbitrary plain-texts and cipher-text

$$\text{Given that } \Pr[\mathbf{M} = m_0 \mid \mathbf{C} = c] = \Pr[\mathbf{M} = m_0]$$

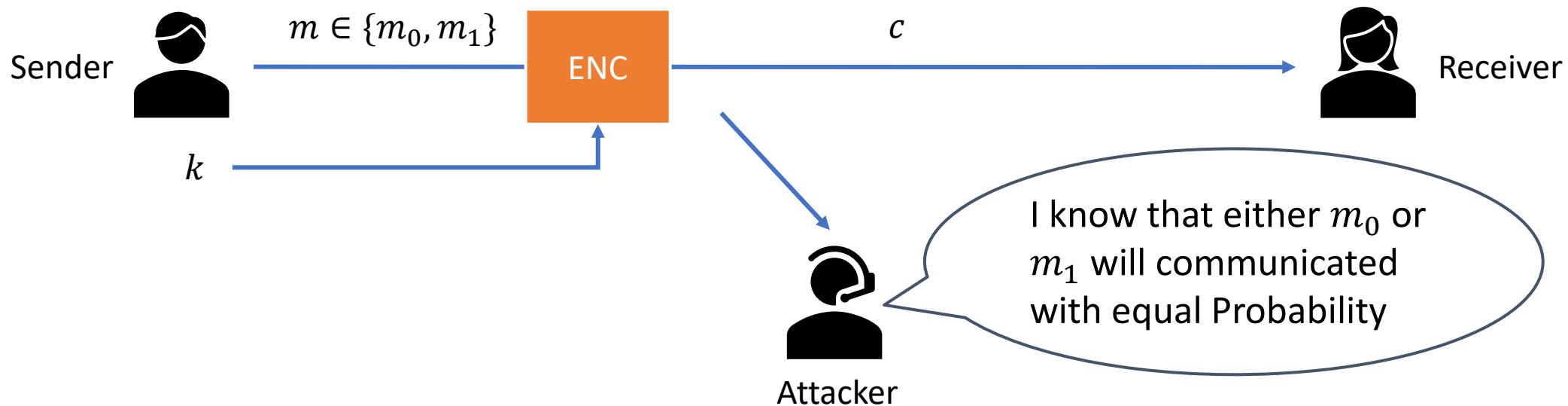
$$\frac{\Pr[\mathbf{C} = c \mid \mathbf{M} = m_0] \cdot \Pr[\mathbf{M} = m_0]}{\Pr[\mathbf{C} = c]} = \Pr[\mathbf{M} = m_0] \quad (\text{Expanding LHS by Bayes theorem})$$

$$\therefore \Pr[\mathbf{C} = c \mid \mathbf{M} = m_0] = \Pr[\mathbf{C} = c]$$

$$\text{Similarly, given that } \Pr[\mathbf{M} = m_1 \mid \mathbf{C} = c] = \Pr[\mathbf{M} = m_1]$$

$$\Pr[\mathbf{C} = c \mid \mathbf{M} = m_1] = \Pr[\mathbf{C} = c] \quad (\text{Expanding LHS by Bayes theorem and simplifying as above})$$

Second equivalent Definition

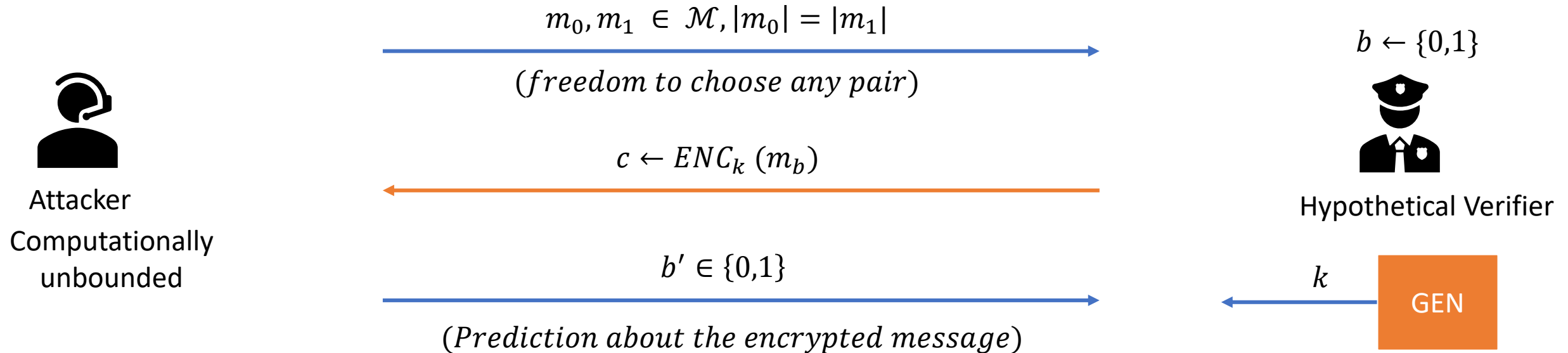


Perfect Secrecy: adversary should not get “any advantage” by seeing c

- Attacker should learn the underlying m from c only with probability $\frac{1}{2}$
- No better than guessing m

Second equivalent definition

- Definition is formalized with the help of challenge-response game (experiment)



Second Equivalent Definition

- Experiment output:
 - 1, if $b=b' \rightarrow$ interpretation : Attacker identified underlined message
 - 0, if $b \neq b' \rightarrow$ interpretation: Attacker failed to identify the underlying message

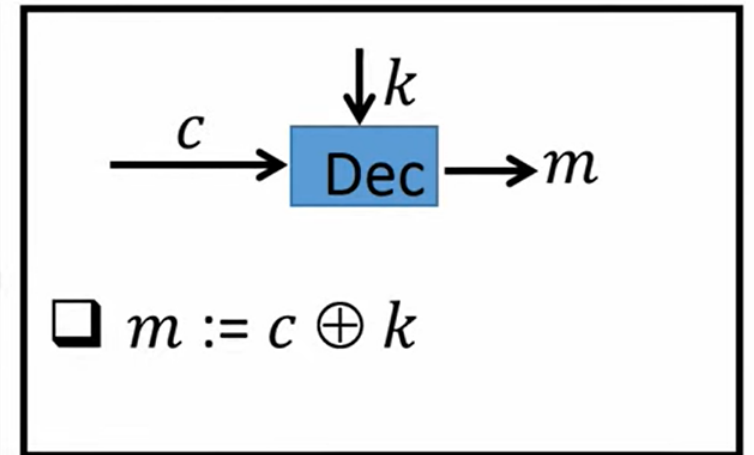
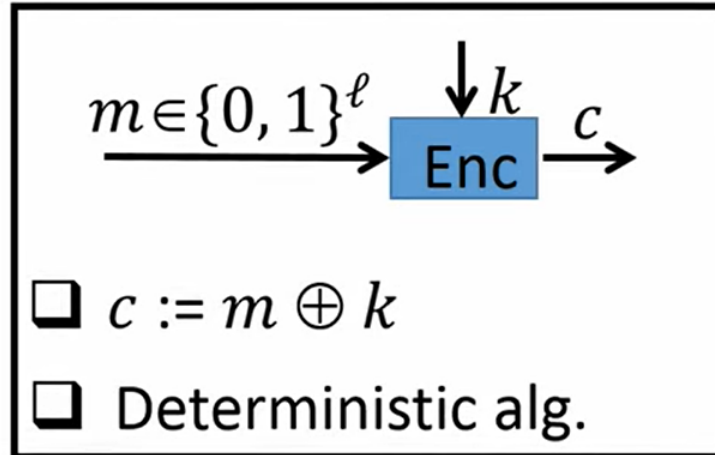
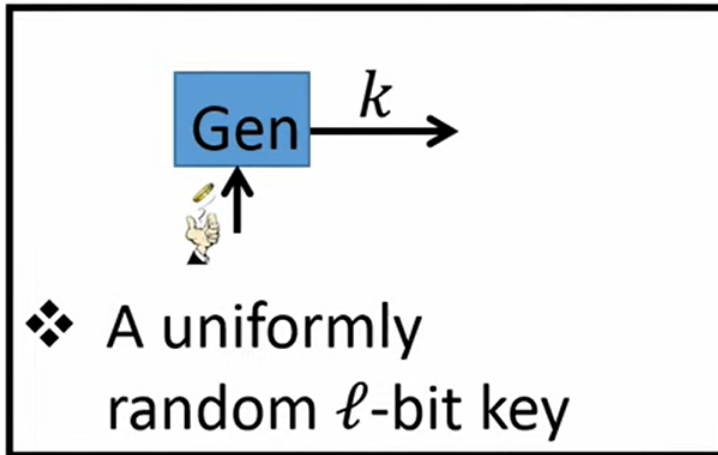
- $\Pi = (GEN, ENC, DEC)$ over \mathcal{M} is perfectly – indistinguishable if for every \mathcal{A}

$$\Pr(\text{PrivK}_{\mathcal{A}, \Pi}^{eav} = 1) = \frac{1}{2}$$

Limitations of perfect Security

Vernam Cipher (One-time pad encryption)

$$\mathcal{M} = \mathcal{K} = \mathcal{C} = \{0, 1\}^\ell$$



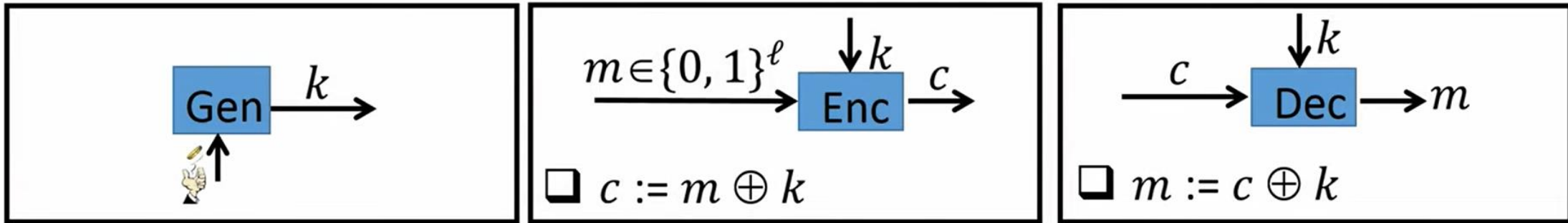
☐ The bit-wise XOR operation is defined as follows:

$$0 \oplus 0 = 0$$

$$1 \oplus 1 = 0$$

$$0 \oplus 1 = 1$$

Vernam Cipher (One-time pad encryption)



❑ **Limitation I** : Key k , should be as **large** as plaintext m

- ❖ To securely communicate 1GB file, sender and receiver need to agree on a uniformly random key of size 1GB

❑ **Limitation II** : Key k **cannot** be used to encrypt **more than one message**

- ❖ A **fresh key** for each instance of encryption

- Let key k is used to encrypt **two different messages** $m_0, m_1 \in \{0, 1\}^\ell$
- An eavesdropper on seeing $c_0 := m_0 \oplus k$ and $c_1 := m_1 \oplus k$ can compute

$$c_0 \oplus c_1 = m_0 \oplus m_1$$

