

Lecture 1

Network Security

Introduction

Chapter 1

Learning Objective

- Introduce the security requirements
 - confidentiality
 - integrity
 - availability
- Describe the X.800 security architecture for OSI

Network Security Requirements

Network Security

- Definition: The protection afforded to an automated information system in order to attain the application objectives to preserving the **integrity, availability, and confidentiality** of information system resources (includes hardware, software, firmware, information/data, and telecommunications).
 - NIST Computer Security Handbook

Confidentiality

- **Data confidentiality:** Assures that private or confidential information is not made available or disclosed to **unauthorized** individuals;
- **Privacy:** Assures that individual's control or influence **what information** related to them may be collected and stored and **by whom** and **to whom** that information may be disclosed
- i.e., student grade information

Integrity

- **Data integrity:** Assures that data (both stored and in transmitted packets) and programs are changed only in a **specified** and **authorized** manner;
- **System integrity:** Assures that a system performs its intended function in an unimpaired manner, free from deliberate or inadvertent **unauthorized** manipulation of the system
- i.e., a hospital patient's allergy information

Availability

- **Availability:** Assures that systems work promptly, and service is not denied to authorized users, ensuring **timely** and **reliable** access to and use of information
- i.e., denial of service attack

Other security requirements

- **Authenticity**
- **Accountability**
 - traceable data source,
 - fault isolation
 - intrusion detection and prevention,
 - recovery and legal action
 - system must keep records of their activities to permit later forensic analysis to trace security breaches or to aid in transaction disputes

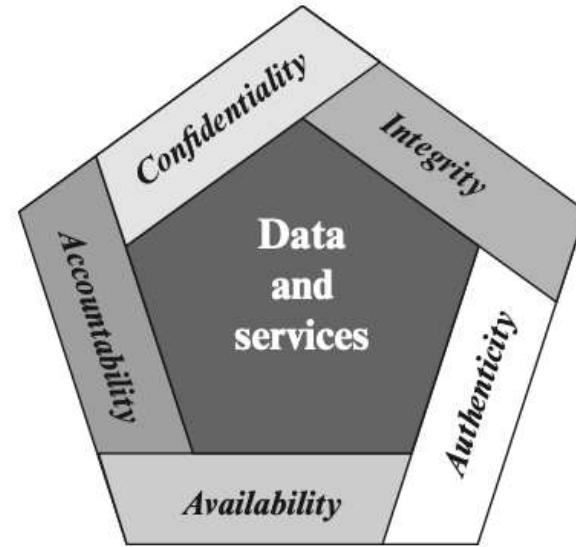


Figure 1.1 Essential Network and Computer Security Requirements

Question

- What security requirements does a blockchain system have achieved?

Project

- **Task1: OnDemand Professor Q&A Bot**

- Your task is to build a Q&A Bot over private data that answers questions about the network security course using the open-source alternatives to ChatGPT that can be run on your local machine. Data privacy can be compromised when sending data over the internet, so it is mandatory to keep it on your local system.
- Your Q&A Bot should be able to understand user questions and provide appropriate answers from the local database, then the citations should be added (**must be accomplished**) if the response is from the internet, then the web references should be added.
- Train your bot using network security lecture slides, network security textbook, and the Internet.
- By using Wireshark capture data for Step 4's of the LLM workflow shown in Figure 1. Provide detailed explanations of the trace data. Also, Maintain a record of Step 1's prompt and its mapping to the trace data in Step 4's.

- **Task2: Quiz Bot**

- Your task is to build a quiz bot based on a network security course using the open-source alternatives to ChatGPT that can be run on your local machine. Data privacy can be compromised when sending data over the internet, so it is mandatory to keep it on your local system.
- Two types of questions should be offered by the bot: randomly generated questions and specific topic questions and the answers should be pulled from the network security database. Train your bot using network security quizzes, lecture slides, network security textbook, and the Internet.
- The quiz must include multiple-choice questions, true/false questions, and open-ended questions.
- Finally, the bot should be able to provide feedback on the user's answers.

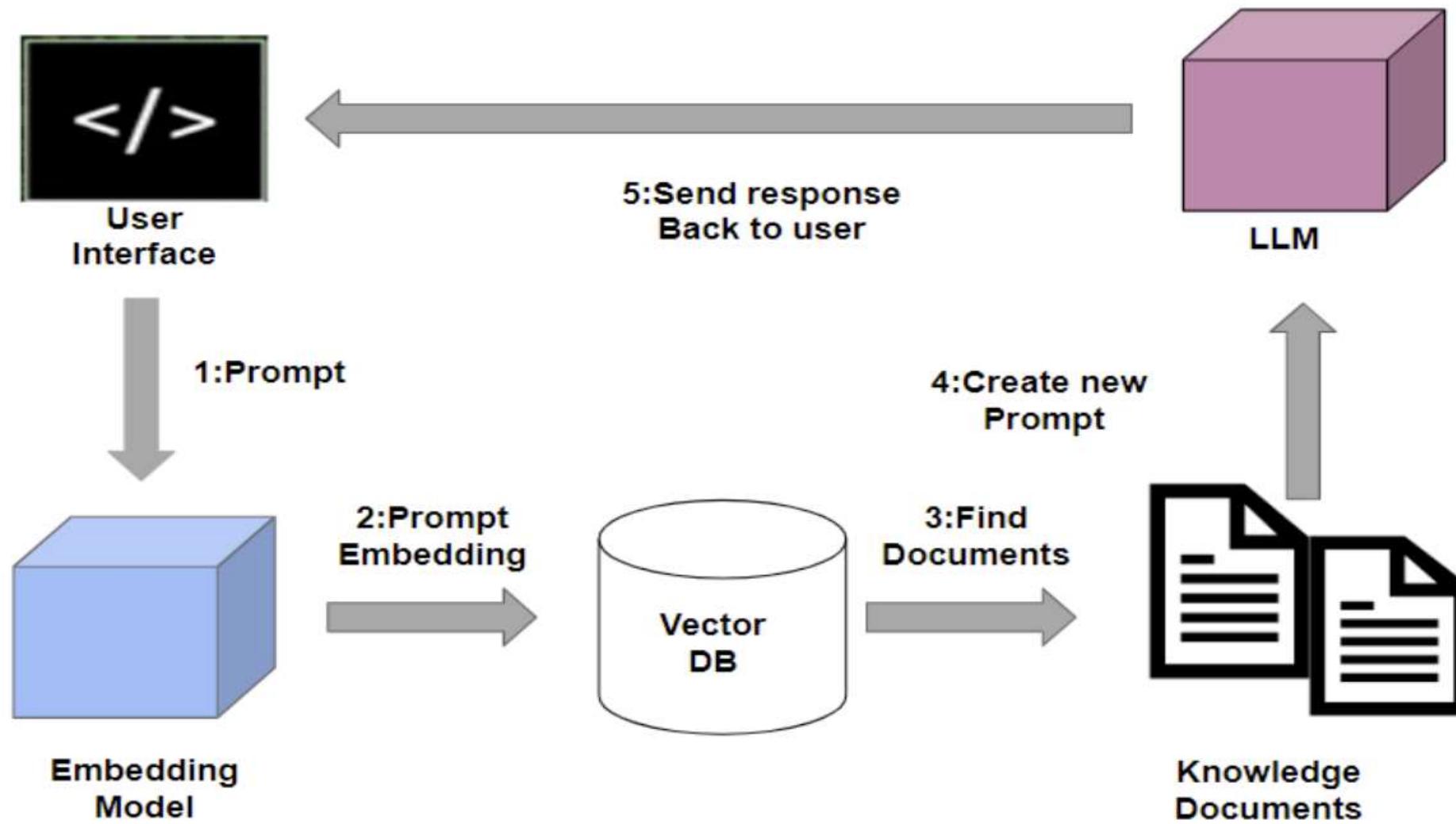


Figure 1: LLM workflow

Form Project Groups

- Form a group (**no submission**)
 - The project will be assigned as a group project. Six students will be considered a group for a project. Here is the link to enter your project group member names.
[FALL 2023 CS5342 PROJECT GROUP NAMES.xlsx](#)
 - It is recommended that one person in the group fills in the form to avoid multiple entries and submits project files on Blackboard. If you cannot get a group, contact to TA.
 - Deadline to submit your group members is 11:59PM on Sept 8th, 2023

Email communication

- Email subject: course# _course name _reason
 - such as “CS 5342_network security_late submission”

Lecture 2

Outline

- Review
- OSI Security Architecture
 - Attack model

Review

- Security requirements
 - Integrity
 - Availability
 - Confidentiality
 - Authenticity
 - Accountability

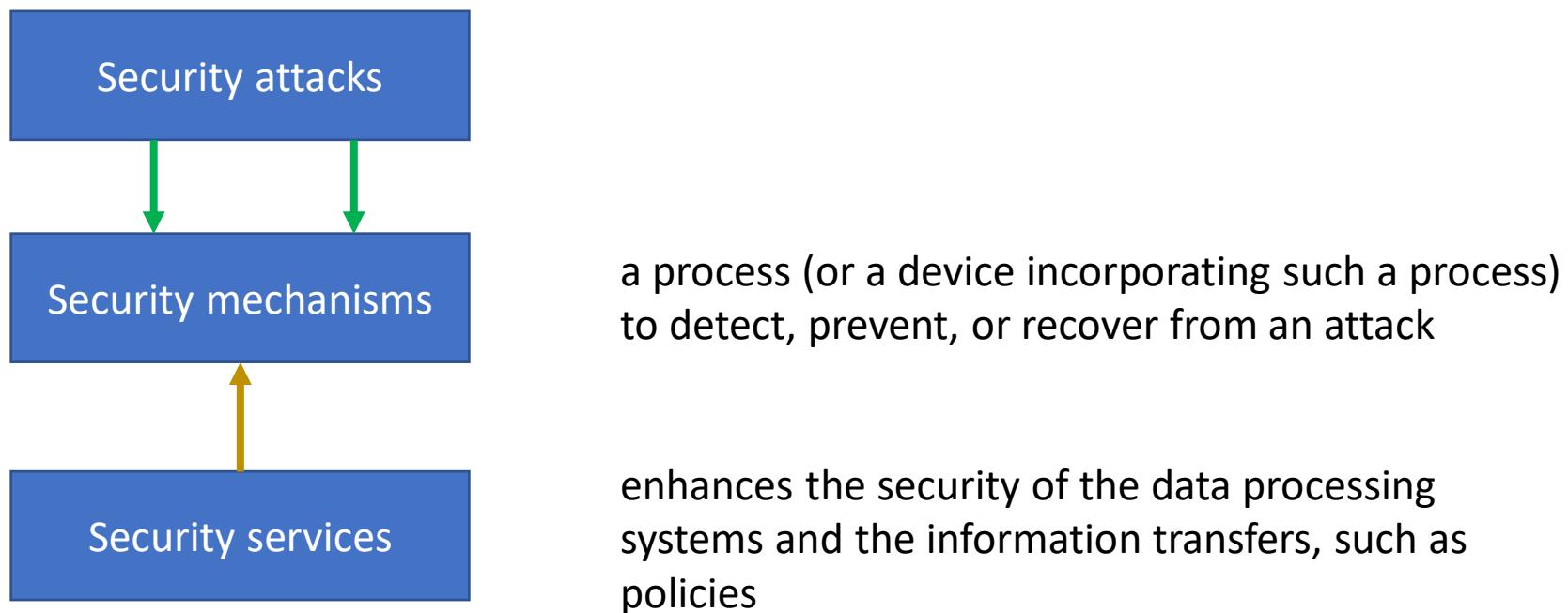
Challenges to achieve a secure system

- The mechanisms used to meet those requirements can be quite **complex**, and understanding them may evolve rather **subtle** reasoning
- When developing security mechanisms, must always consider **potential attacks**
- Sometimes, security mechanisms are **counterintuitive**
- **Where** to use them?
- Involve **more** than a particular algorithm or protocol
- **No agreement** on security for complex and heterogeneous systems i.e. trusts on data in different countries
- etc.

OSI Security Architecture

OSI Security Architecture

- International Telecommunication Union – Telecommunication (ITU-T) recommends X.800
- Security Architecture for Open Systems Interconnection (OSI)
 - Defines a systematic way of defining and providing security requirements
 - Used by IT managers and vendors in their products

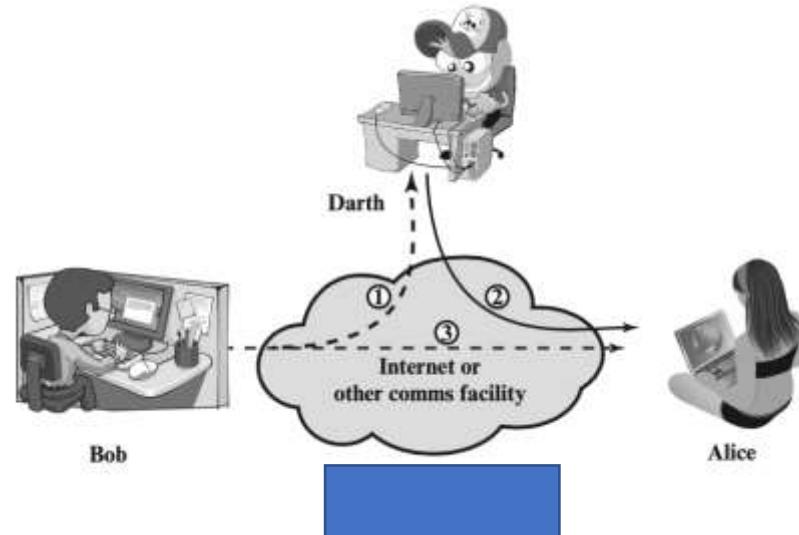
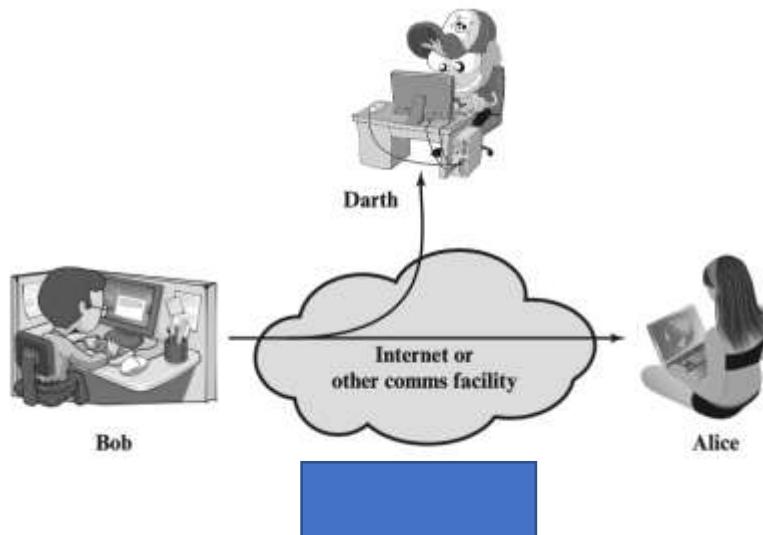


Other Security Architectures

- OWASP - Open Web Application Security Project
 - web application security
 - OWASP foundation
- NIST, Cybersecurity Framework
 - <https://www.nist.gov/cyberframework>
 - [VIRTUAL WORKSHOP #2](#) | February 15, 2023 (9:00 AM – 5:30 PM EST). Join us to discuss potential significant updates to the CSF as outlined in the soon-to-be-released CSF Concept Paper.
 - [https://www.nist.gov/news-events/events/2023/02/journey-nist-cybersecurity-framework-csf-20-workshop-2](#)

Security attack

- **Definition:** any action that compromises the security of information owned by an organization
- Two types of security attacks
 - Passive attack
 - Active attack



Passive attack

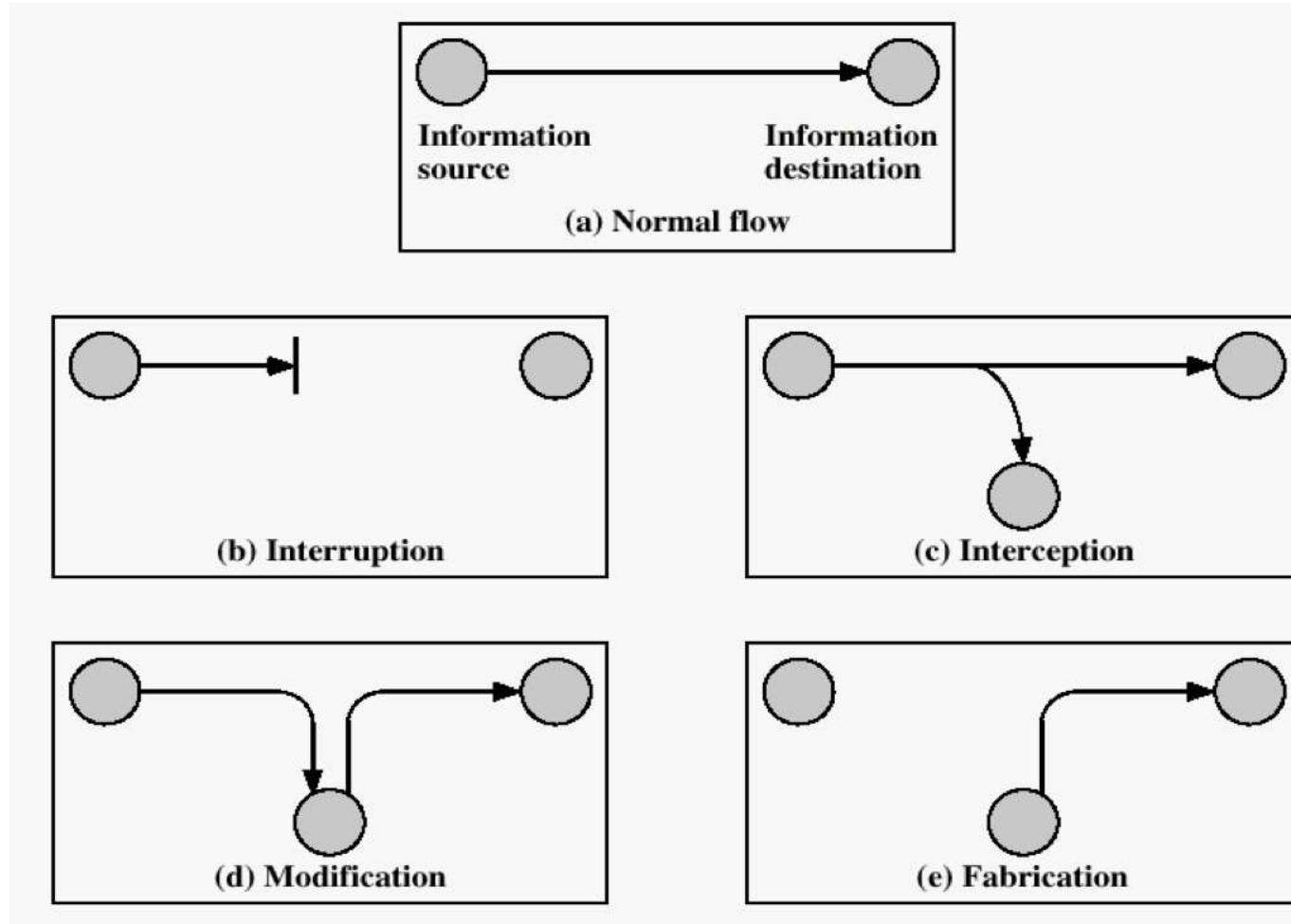
- i.e. eavesdropping on or monitoring of transmissions
- Goal: obtain information being transmitted
 - release of message contents
 - traffic analysis – a promiscuous sniffer
- Very difficult to detect – no alteration of the data
- But easy to prevent, **why?**

Active attack

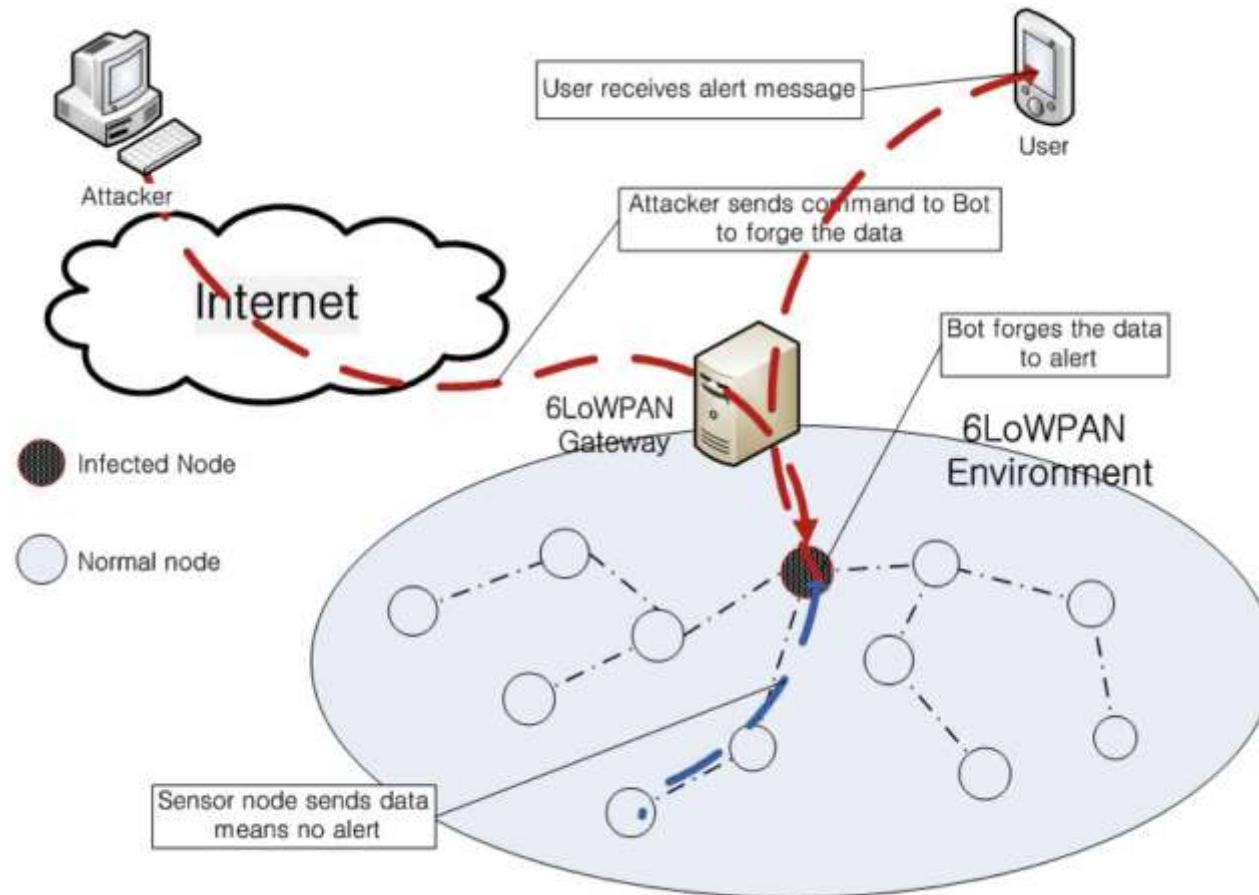
- active attack includes:
 - replay
 - Modification of messages
 - Denial of service
 - Masquerade

Example: two points communication

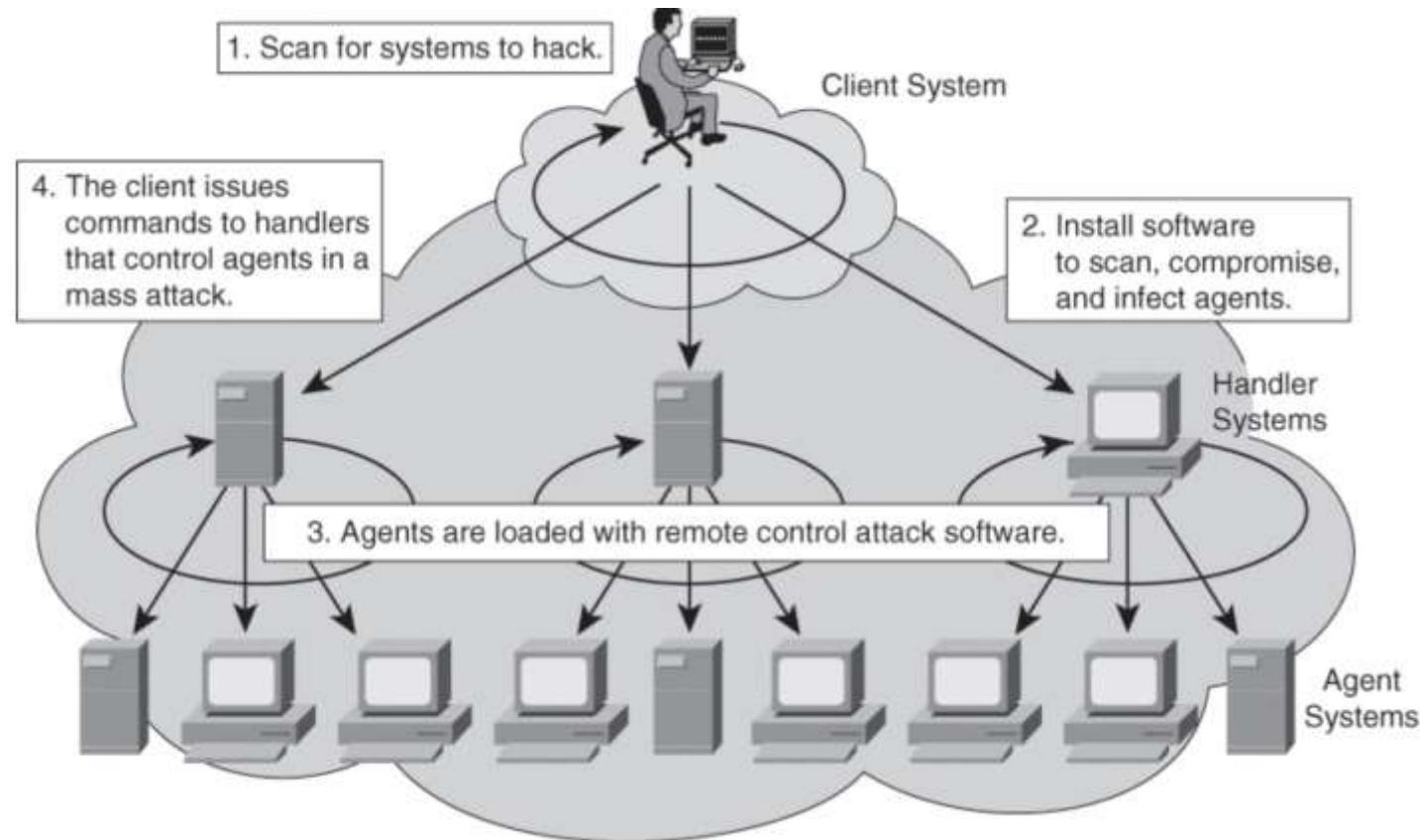
- Generic types of attacks



Example of modification attack in 6LoWPAN



Example: a group of attackers

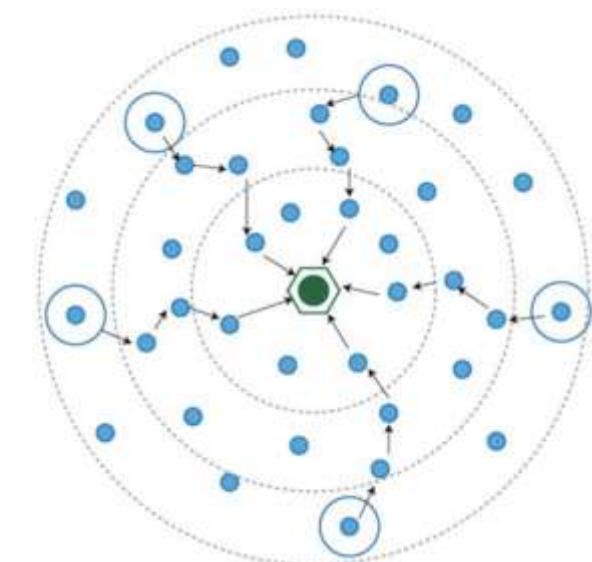


Know Your Threat Model

- **Threat model:** A model of who your attacker is and what resources they have
- One of the best ways to counter an attacker is to attack their reasons

Example: adversary model

- “The adversary is assumed to be intelligent and has limited number of resources. Before capturing the nodes, it exploits the various vulnerabilities of the networks. It knows the topology of the network, routing information. It aims to capture the sink node so as to disrupt the whole traffic. If it is not able to capture the sink node, it will capture the nearby nodes of the sink. It tries to disrupt the whole traffic of the network with minimum number of captured nodes. It is also assumed that the adversary tends to attack more on the nodes closer to the data sink than nodes that are far away”



Lecture 3

Story...

- The bear race
- **Takeaway:** Even if a defense is not perfect, it is important to always stay on top of best security measures



I don't have to outrun the bear. I just have to outrun you

Design in security from the start

- When building a new system, include security as part of the design considerations rather than patching it after the fact
 - A lot of systems today were not designed with security from the start, resulting in patches that don't fully fix the problem!
- Keep these security principles in mind whenever you write code!

Human Factors

- The users
 - Users like convenience (ease of use)
 - If a security system is unusable, it will be unused
 - Users will find way to subvert security systems if it makes their lives easier
- The programmers
 - Programmers make mistakes
 - Programmers use tools that allow them to make mistakes (e.g. C and C++)
- Everyone else
 - Social engineering attacks exploit other people's trust and access for personal gain

Summary for Chapter 1

- Have learned:
 - Security requirements
 - Attack models
 - X.800 secure architecture, security services, mechanisms

Supplementary materials

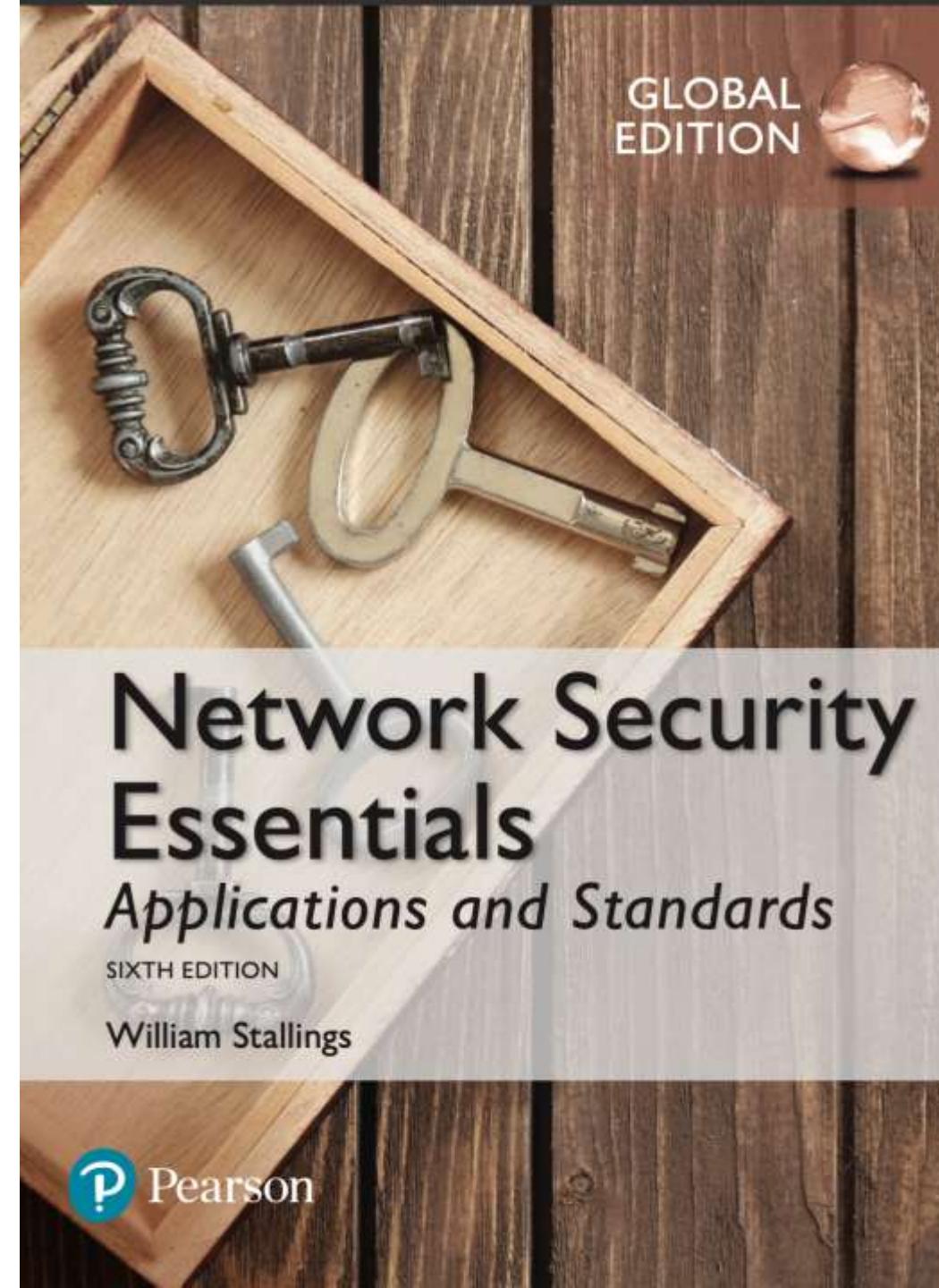
- Internet Security Glossary, v2 – produced by Internet Society
<https://datatracker.ietf.org/doc/html/rfc4949>
- X.800 – OSI network security
https://www.itu.int/rec/dologin_pub.asp?lang=f&id=T-REC-X.800-199103-I!!!PDF-E&type=items

GLOBAL
EDITION



Review Questions

- William Stallings (WS), “Network Security Essentials”, 6th Global Edition
- RQ 1.1 - 1.3
- Prob 1.5



Network Security

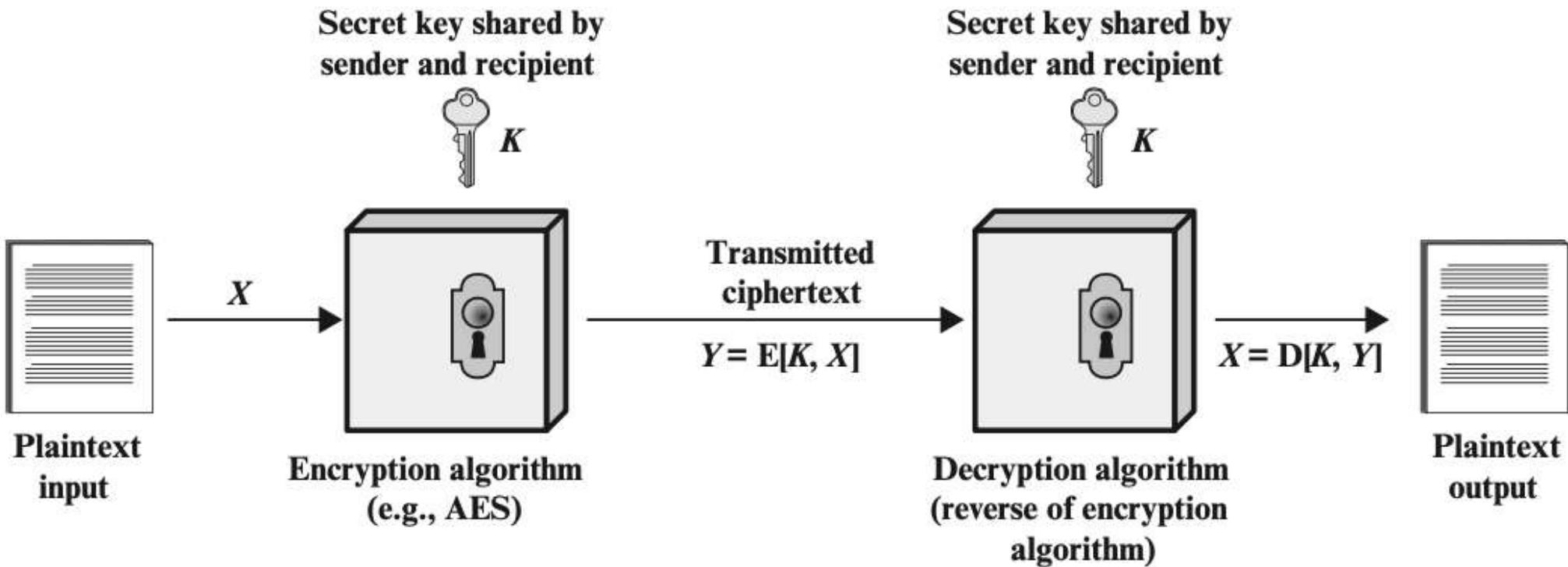
Chapter 2

Symmetric encryption

- Sender and recipient share a common/same key
- Was the only type of cryptography, prior to invention of public-key in 1970's

Symmetric Encryption Principles

Simplified model of symmetric encryption



Symmetric encryption

- Has five ingredients
 - **Plaintext:** the original message or data
 - **Encryption algorithm:** performs various substitutions and transformations on the plaintext
 - **Secret key**
 - **Ciphertext:** the coded message
 - **Decryption algorithm:** takes the ciphertext and the same secret key and produces the original plaintext

Other basic terminology

- **cipher** - algorithm for transforming plaintext to ciphertext
- **encipher (encrypt)** - converting plaintext to ciphertext
- **decipher (decrypt)** - recovering plaintext from ciphertext
- **cryptography** - study of encryption principles/methods
- **cryptanalysis (codebreaking)** - the study of principles/ methods of deciphering ciphertext *without* knowing key

Requirements

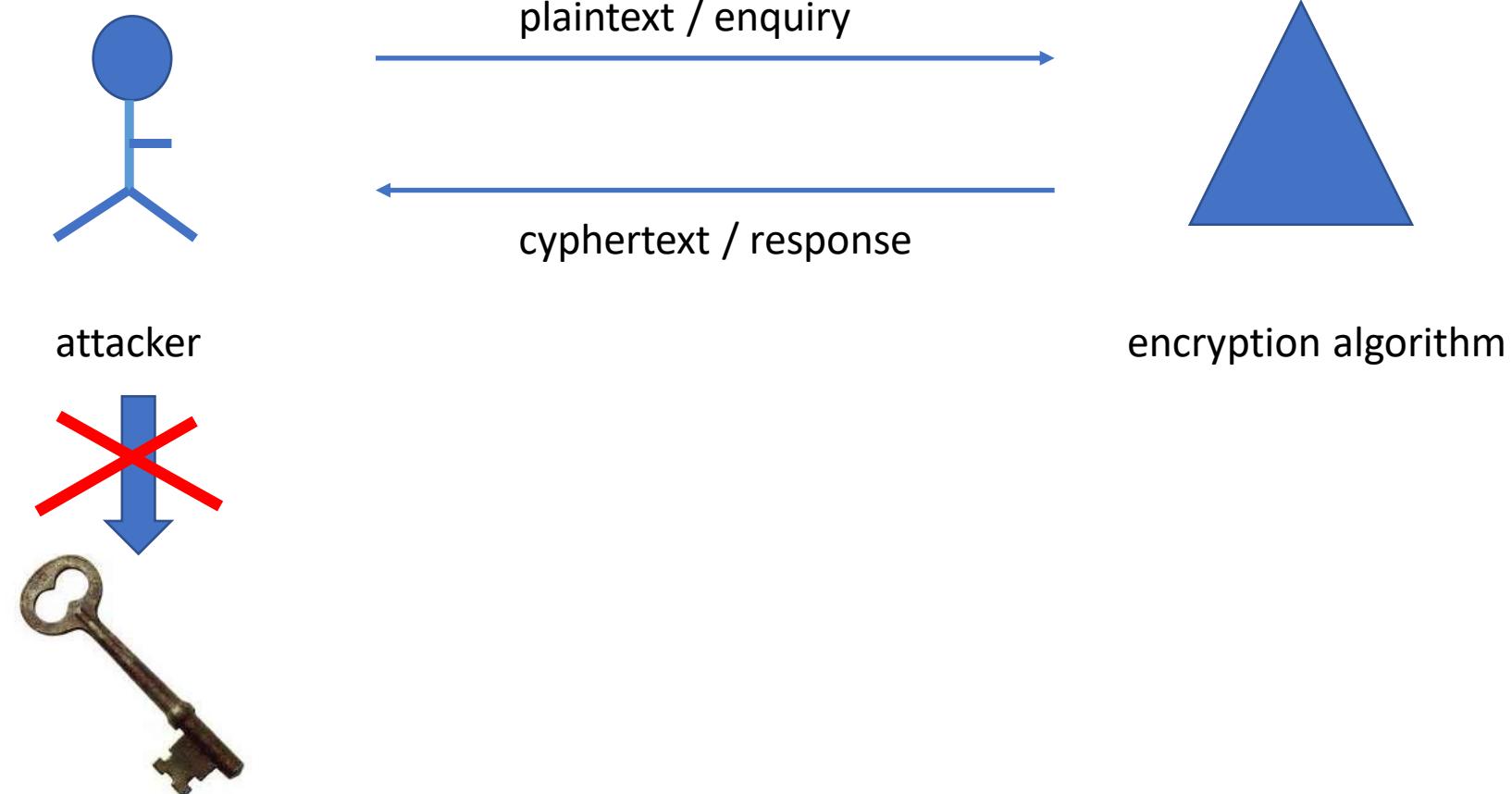
- Two requirements for secure use of symmetric encryption:
 - a strong encryption algorithm
 - a secret key known only to sender / receiver
- $Y = E_k(X)$
- $X = D_k(Y)$
- assume encryption algorithm is known
- the security of symmetric encryption depends on the secrecy of the key
- implies a secure channel to distribute key

Lecture 4

Requirements

- Two requirements for secure use of symmetric encryption:
 - a strong encryption algorithm
 - a secret key known only to sender / receiver
- $Y = E_k(X)$
- $X = D_k(Y)$
- assume encryption algorithm is known
- the security of symmetric encryption depends on the secrecy of the key
- implies a secure channel to distribute key

A strong encryption algorithm

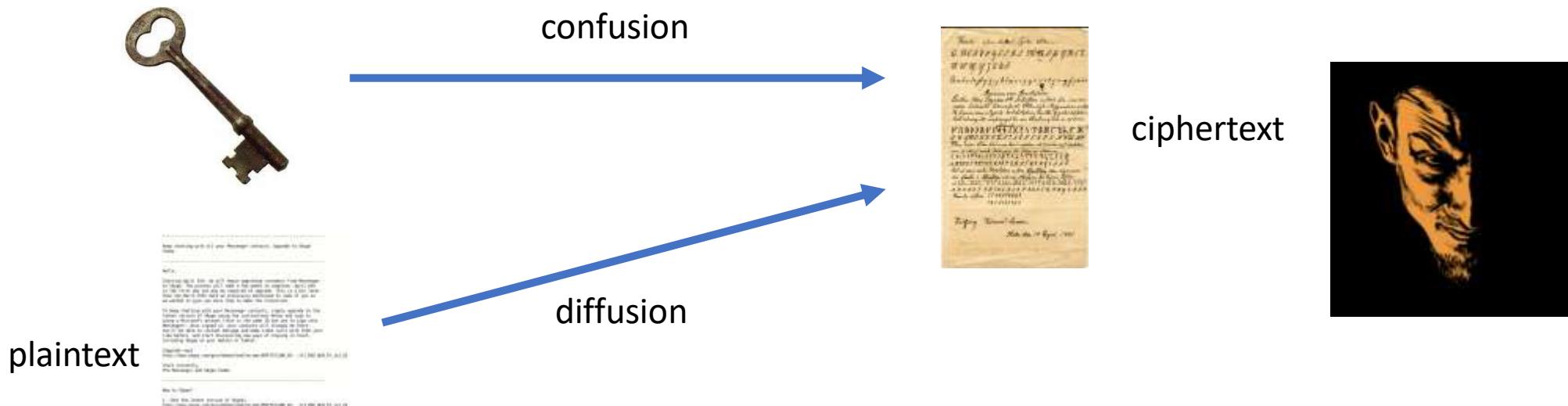


Secure Encryption Scheme

- **Unconditional security**
 - no matter how much computer power is available, the cipher cannot be broken since the ciphertext provides insufficient information to uniquely determine the corresponding plaintext
- **Computational security**
 - the cost of breaking the cipher exceeds the value of the encrypted information;
 - or the time required to break the cipher exceeds the useful lifetime of the information

Desired characteristics

- Cipher needs to completely obscure statistical properties of original message
- more practically Shannon suggested combining elements to obtain:
 - Confusion – how does changing a bit of the key affect the ciphertext?
 - Diffusion – how does changing one bit of the plaintext affect the ciphertext?



Ways to achieve

- Symmetric Encryption:
 - substitution / transposition / hybrid
- Asymmetric Encryption:
 - Mathematical hardness - problems that are efficient to compute in one direction, but inefficient to reverse by the attacker
 - Examples: Modular arithmetic, factoring, discrete logarithm problem, Elliptic Logs over Elliptic Curves

Project

- TA Name: Lin, Yu
- Email: Yu.Lin@ttu.edu
- Form a group (no submission)
 - The project will be assigned as a group project. Six students will be considered a group for a project. Here is the link to enter your project group member names.
[FALL 2023 CS5342 PROJECT GROUP NAMES.xlsx](#)
 - It is recommended that one person in the group fills in the form to avoid multiple entries and submits project files on Blackboard. If you cannot get a group, contact the TA.
- Deadline to submit your group members is 11:59 PM on Sept 8th, 2023

Lecture 5

Review

- Strong Encryption Algorithm
 - Confusion
 - Diffusion

Symmetric Block Encryption

Block cipher

- the most commonly used symmetric encryption algorithms
- input: fixed-size blocks (Typically 64, 128 bit blocks), output: equal size blocks
- provide secrecy and/or authentication services
- Data Encryption Standard (DES), triple DES (3DES), and the Advanced Encryption Standard (AES)s
- Usually employ Feistel structure

Feistel Cipher Structure

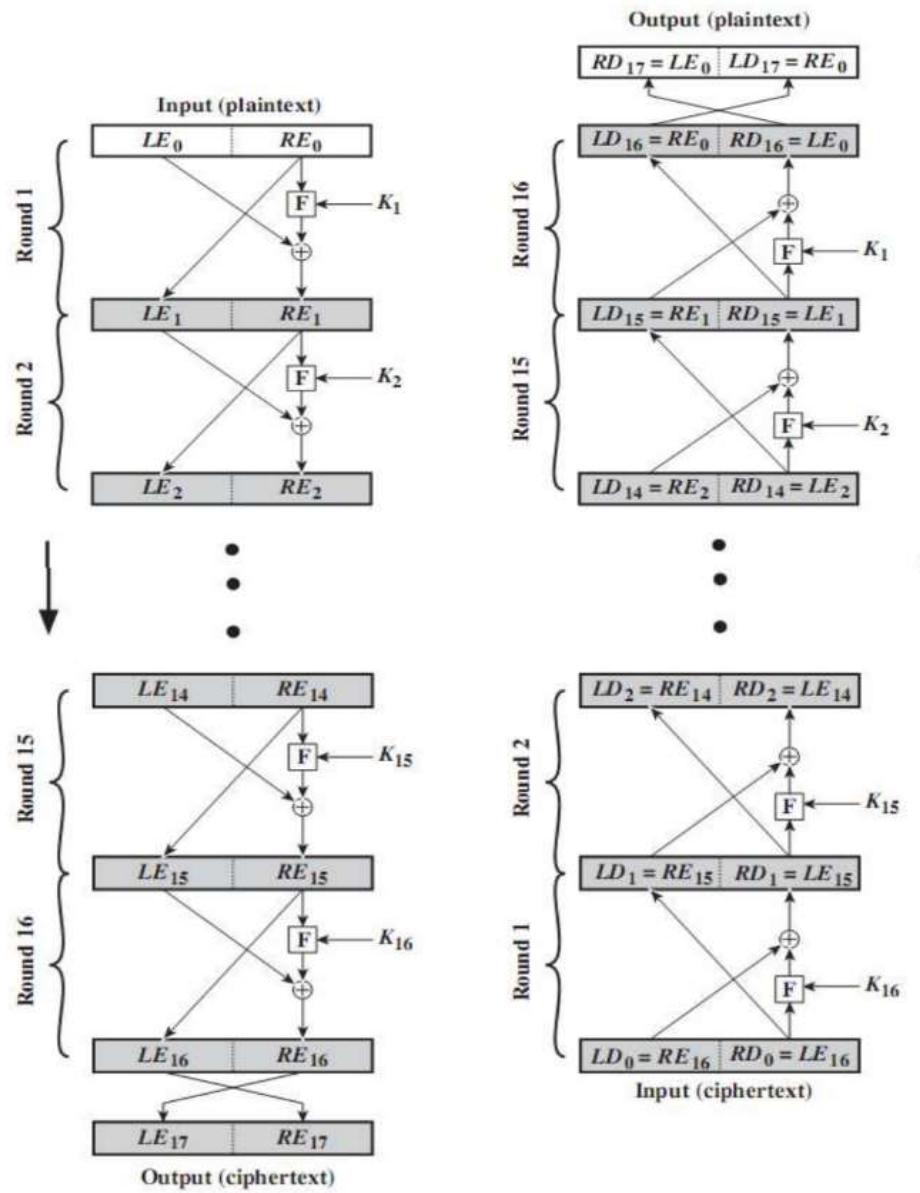
Feistel Cipher Structure

- most symmetric block ciphers are based on a **Feistel Cipher Structure**
- based on the two primitive cryptographic operations
 - *substitution* (S-box)
 - *permutation* (P-box)
- provide *confusion* and *diffusion* of message

Feistel Cipher Structure

- Horst Feistel devised the **feistel cipher** in the 1973
 - based on concept of invertible product cipher
- partitions input block into two halves
 - process through multiple rounds which
 - perform a substitution on left data half
 - based on round function of right half & subkey
 - then have permutation swapping halves
- implements Shannon's substitution-permutation network concept

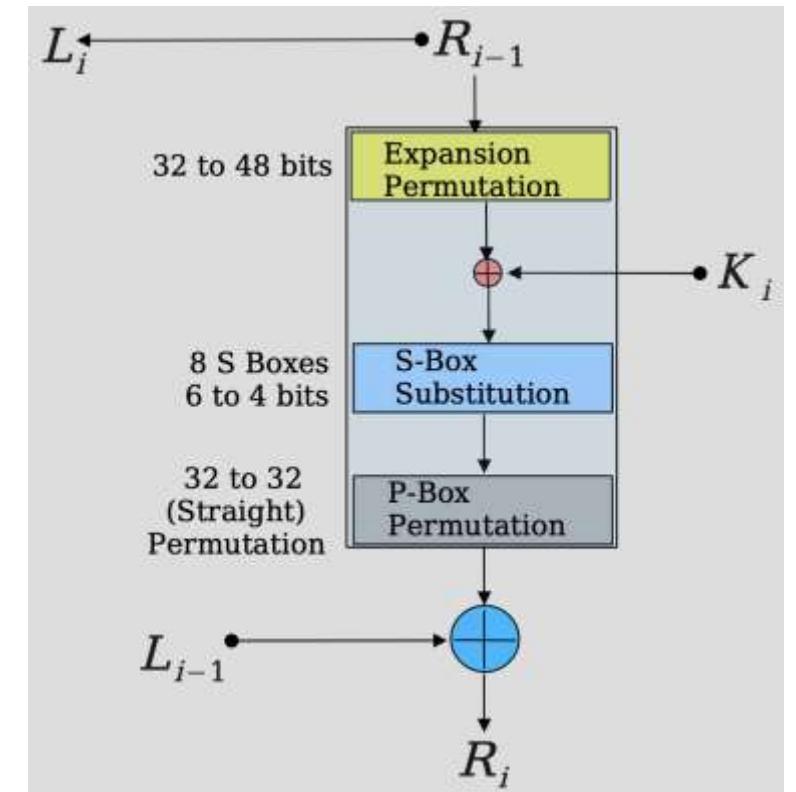
Feistel Encryption and Decryption



Encryption

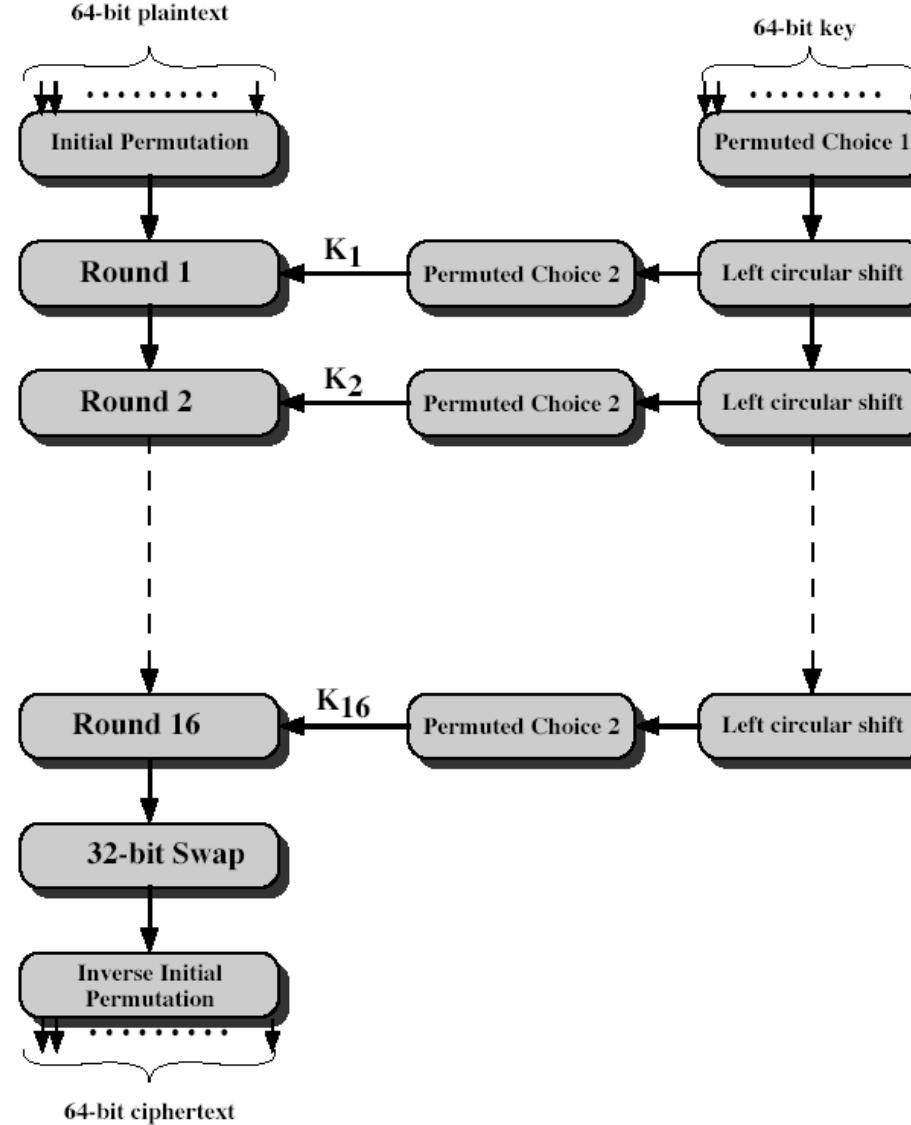
$$L_i = R_{i-1}$$

$$R_i = L_{i-1} \oplus F(R_{i-1}, K_i)$$



DES encryption

- 64 bits plaintext
- 56 bits effective key length



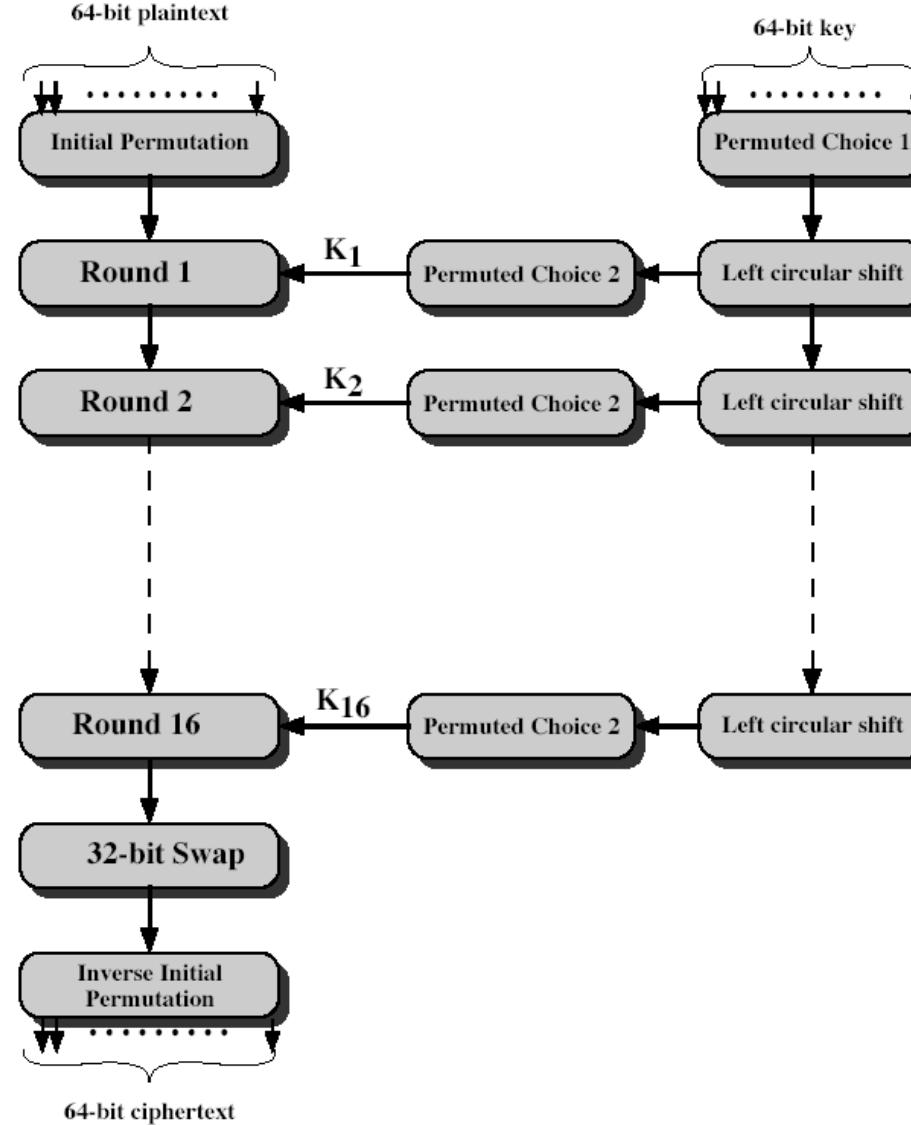
Lecture 6

Outline

- DES
- 3DES
- AES

DES encryption

- 64 bits plaintext
- 56 bits effective key length



DES Weakness

- short length key (56 bits) is not secure enough. Brutal force search takes short time.

Triple DES (3DES)

$$C = E(K_3, D(K_2, E(K_1, P)))$$

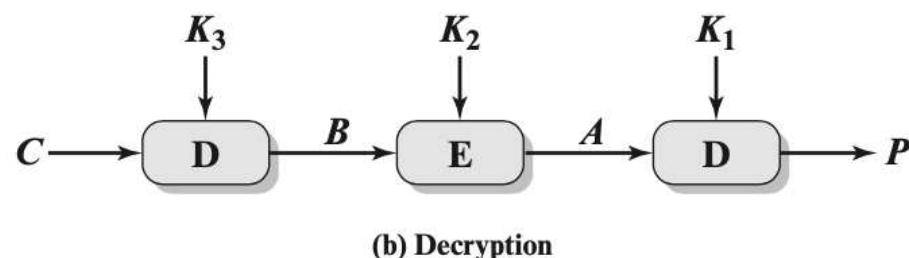
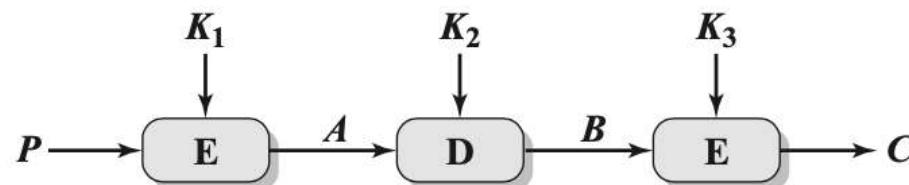
where

C = ciphertext

P = plaintext

$E[K, X]$ = encryption of X using key K

$D[K, Y]$ = decryption of Y using key K



Decrypting with the wrong key will further convolute the output

3DES

- Triple DES with three different keys – brute-force complexity 2^{168}
- 3DES is the FIPS-approved symmetric encryption algorithm
- Weakness: slow speed for encryption

FIPS – Federal Information Processing Standards. The United States' Federal Information Processing Standards are publicly announced standards developed by the National Institute of Standards and Technology for use in computer systems by non-military American government agencies and government contractors

AES

- clearly a replacement for DES was needed
 - have theoretical attacks that can break it
 - have demonstrated exhaustive key search attacks
- can use Triple-DES – but slow with small blocks
- US NIST issued call for ciphers in 1997
- 15 candidates accepted in Jun 98
- 5 were short-listed in Aug-99
- Rijndael was selected as the AES in Oct-2000
- issued as FIPS PUB 197 standard in Nov-2001

Criteria to evaluate AES

- General security
- Software implementations
- Restricted-space environments
- Hardware implementations
- Attacks on implementations
- Encryption versus decryption
- Key agility
- Other versatility and flexibility
- Potential for instruction-level parallelism

[Cryptographic Standards and Guidelines | CSRC \(nist.gov\)](https://csrc.nist.gov)

AES Specification

- symmetric block cipher
- 128-bit data, 128/192/256-bit keys
- stronger & faster than Triple-DES
- provide full specification & design details
- both C & Java implementations
- NIST have released all submissions & unclassified analyses

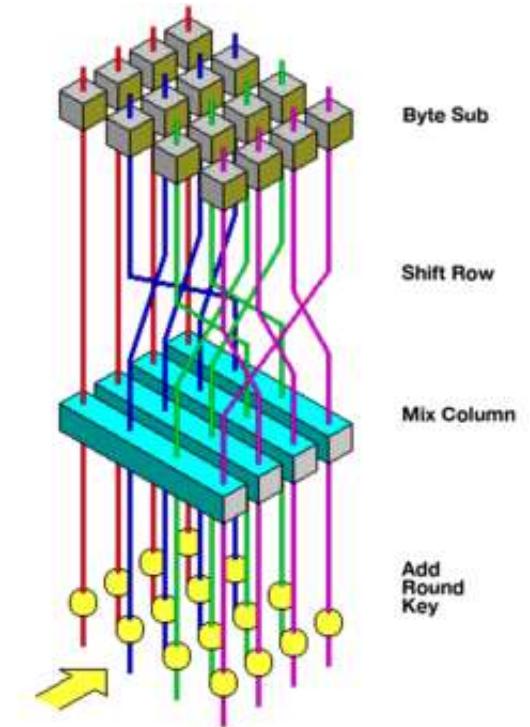
<https://csrc.nist.gov/CSRC/media/Projects/Cryptographic-Standards-and-Guidelines/documents/aes-development/Rijndael-ammended.pdf>

The AES Cipher - Rijndael

- designed by Rijmen-Daemen in Belgium
- has 128/192/256 bit keys
- an **iterative** rather than **feistel** cipher
 - treats data in 4 groups of 4 bytes
 - operates an entire block in every round
- designed to be:
 - resistant against known-plaintext attacks
 - speed and code compactness on many CPUs
 - design simplicity

Rijndael

- processes data as 4 groups of 4 bytes (state) = 128 bits
- has 10/12/14 rounds in which state undergoes:
 - byte substitution (1 S-box used on every byte)
 - shift rows (permute bytes row by row)
 - mix columns (alter each byte in a column as a function of all of the bytes in the column)
 - add round key (XOR state with key material)
- 128-bit keys – 10 rounds, 192-bit keys – 12 rounds, 256-bit keys – 14 rounds



Project

- Task 1: G1 to G5
- Task 2: G6 to G10
- We will create submission portal on BB.
- The deadline for the first round of submissions is 10/13/2023 at 11:59 PM.

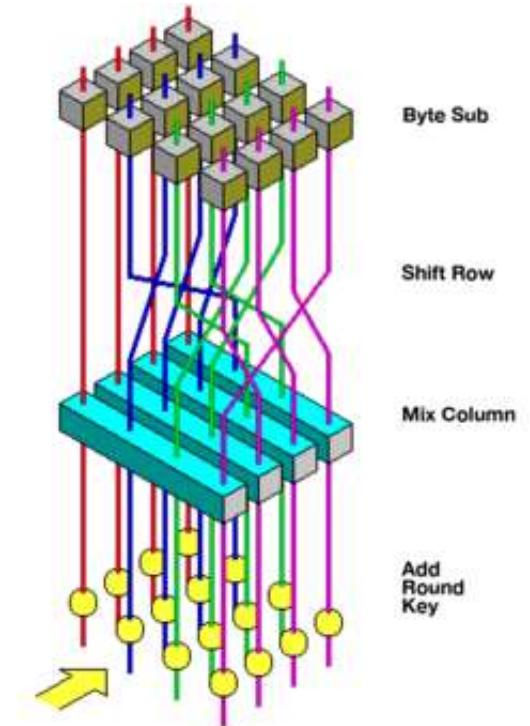
Lecture 7

Outline

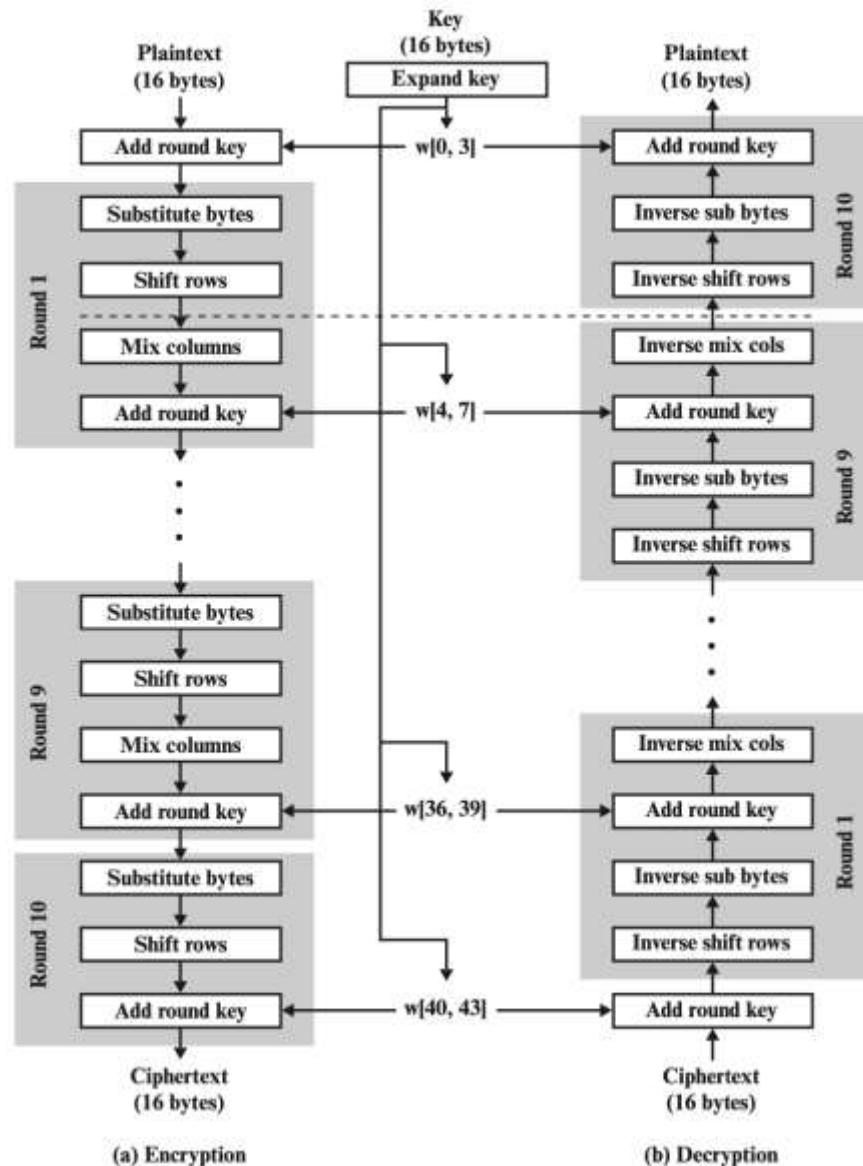
- AES
- Random number

Rijndael

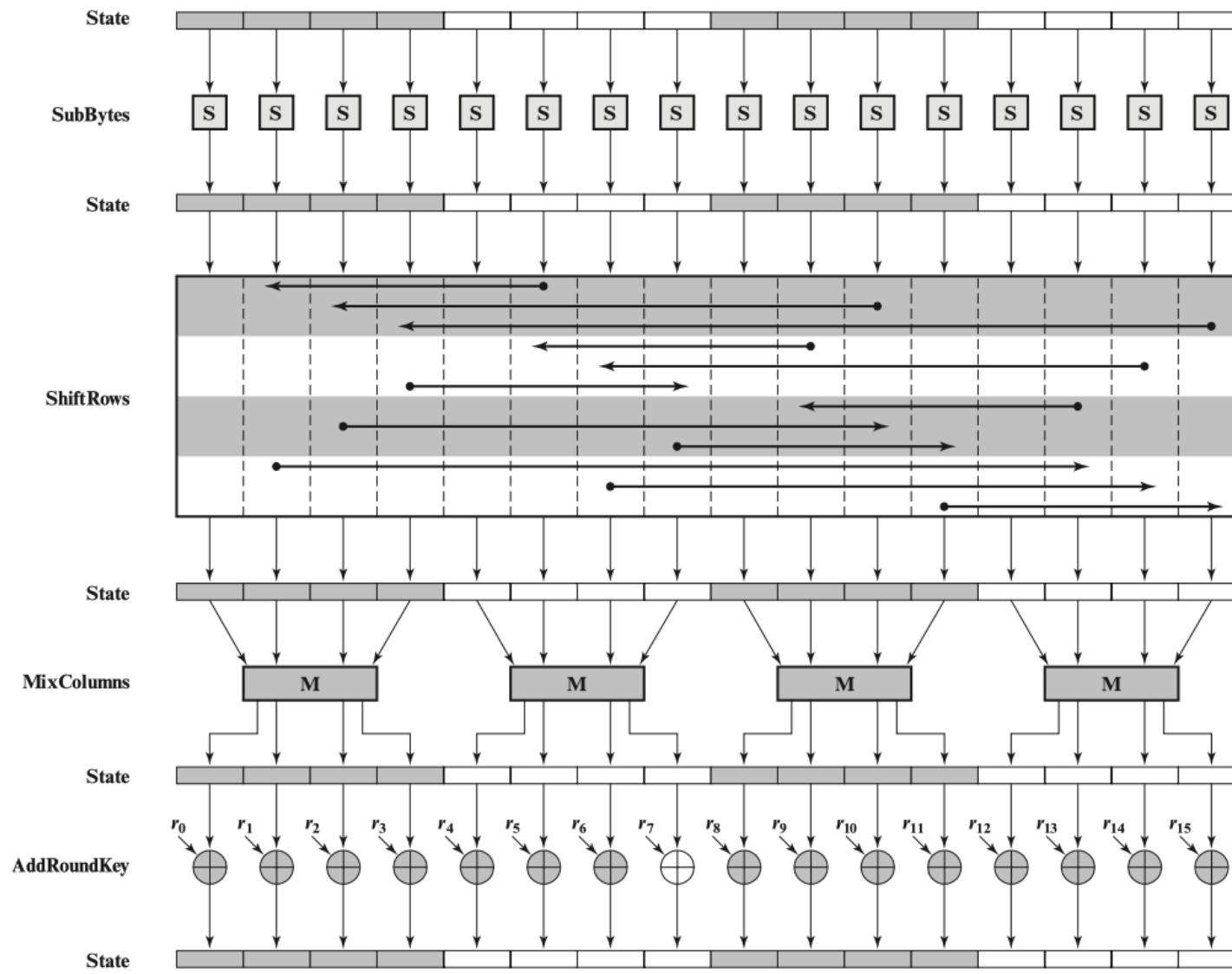
- processes data as 4 groups of 4 bytes (state) = 128 bits
- has 10/12/14 rounds in which state undergoes:
 - byte substitution (1 S-box used on every byte)
 - shift rows (permute bytes row by row)
 - mix columns (alter each byte in a column as a function of all of the bytes in the column)
 - add round key (XOR state with key material)
- 128-bit keys – 10 rounds, 192-bit keys – 12 rounds, 256-bit keys – 14 rounds



AES Encryption and Decryption



AES encryption round



AES pros

- Most operations can be combined into XOR and table lookups - hence very fast & efficient

Take-home Exercises

- Find an AES code to encrypt a text (A), then decrypt it and check whether the original text (A) equals the decrypted text (B). Whether A = B?
- Compare the decryption time with different key lengths, and with DES and 3DES.
 - Suggestions: find a large A file. Run decryption a couple of times and take the average.

Reading materials

- [FIPS 197, Advanced Encryption Standard \(AES\) \(nist.gov\)](#)

Lecture 8

Random and Pseudorandom Numbers

When to use random numbers?

- Generation of a stream key for symmetric stream cipher
- Generation of keys for public-key algorithms
 - RSA public-key encryption algorithm (described in Chapter 3)
- Generation of a symmetric key for use as a temporary **session key**
 - used in a number of networking applications, such as Transport Layer Security (Chapter 5), Wi-Fi (Chapter 6), e-mail security (Chapter 7), and IP security (Chapter 8)
- In a number of key distribution scenarios
 - Kerberos (Chapter 4)

Two types of random numbers

- True random numbers:
 - generated in non-deterministic ways. They are not predictable and repeatable
- Pseudorandom numbers:
 - appear random, but are obtained in a deterministic, repeatable, and predictable manner

Properties of Random Numbers

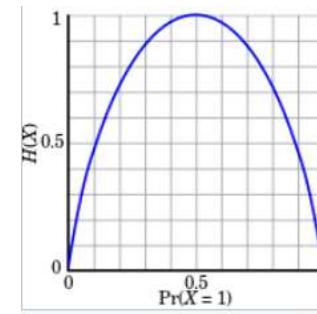
- Randomness
 - Uniformity
 - distribution of bits in the sequence should be uniform
 - Independence
 - no one subsequence in the sequence can be inferred from the others
- Unpredictable
 - satisfies the "next-bit test"

Entropy

- A measure of uncertainty
 - In other words, a measure of how unpredictable the outcomes are
 - **High entropy** = unpredictable outcomes = desirable in cryptography
 - The uniform distribution has the highest entropy (every outcome equally likely, e.g. fair coin toss)
 - Usually measured in bits (so 3 bits of entropy = uniform, random distribution over 8 values)

$$H = - \sum_i p_i \log_2(p_i)$$

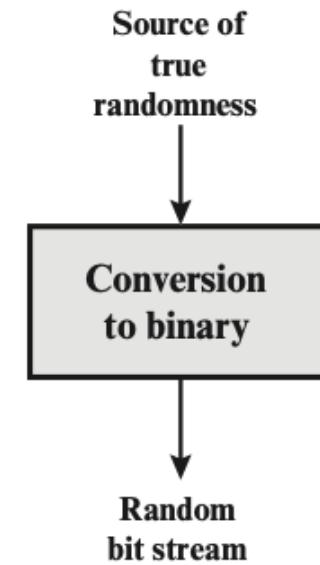
Entropy of an information source



Lecture 9

True random numbers generators

- Several sources of randomness – natural sources of randomness
 - decay times of radioactive materials
 - electrical noise from a resistor or semiconductor
 - radio channel or audible noise
 - keyboard timings
 - disk electrical activity
 - mouse movements
 - Physical unclonable function (PUF)
- Some are better than others



(a) TRNG

Combining sources of randomness

- Suppose r_1, r_2, \dots, r_k are random numbers from different sources.

E.g.,

r_1 = electrical noise from a resistor or semiconductor

r_2 = sample of hip-hop music on radio

r_3 = clock on computer

$$b = r_1 \oplus r_2 \oplus \dots \oplus r_k$$

If any one of r_1, r_2, \dots, r_k is truly random, then so is b

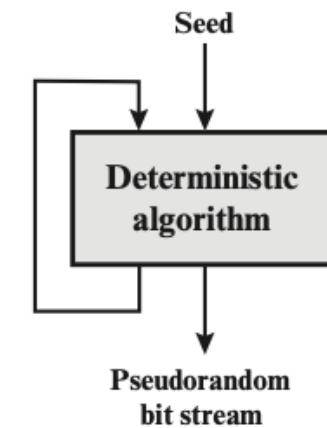
Many poor sources + 1 good source = good entropy

Pseudorandom Number Generators (PRNGs)

- True randomness is expensive
- **Pseudorandom number generator (PRNGs):** An algorithm that uses a little bit of true randomness to generate a lot of random-looking output
 - Also called **deterministic random bit generators (DRBGs)**
- PRNGs are deterministic: Output is generated according to a set algorithm
 - However, for an attacker who can't see the internal state, the output is *computationally indistinguishable* from true randomness

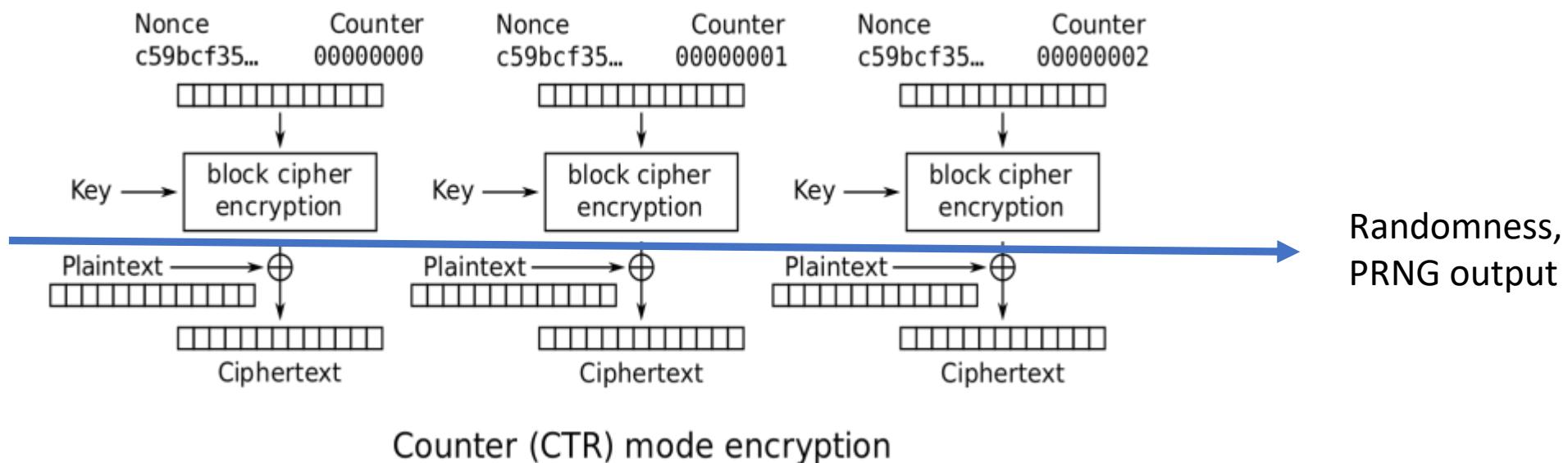
PRNG: Definition

- A PRNG has two functions:
 - PRNG.Seed(randomness): Initializes the internal state using the entropy
 - Input: Some truly random bits
 - PRNG.Generate(n): Generate n pseudorandom bits
 - Input: A number n
 - Output: n pseudorandom bits
 - Updates the internal state as needed
- Properties
 - **Correctness:** Deterministic
 - **Efficiency:** Efficient to generate pseudorandom bits
 - **Security:** Indistinguishability from random
 - **Rollback resistance:** cannot deduce anything about any previously-generated bit



Example construction of PRNG

- Using block cipher in Counter (CTR) mode:
- If you want m random bits, and a block cipher with E_k has n bits, apply the block cipher m/n times and concatenate the result:
- $\text{PRNG.Seed}(K \mid \text{IV}) = E_k(\text{IV}, 1) \mid E_k(\text{IV}, 2) \mid E_k(\text{IV}, 3) \dots E_k(\text{IV}, \text{ceil}(m/n))$,
 - \mid is concatenation
 - Initialization vector (IV) / Nonce – typically is random or pseudorandom



PRNG: Security

- Can we design a PRNG that is truly random?
- A PRNG cannot be truly random
 - The output is deterministic given the initial seed
- A secure PRNG is computationally indistinguishable from random to an attacker
 - Game: Present an attacker with a truly random sequence and a sequence outputted from a secure PRNG
 - An attacker should be able to determine which is which with probability ≈ 0
- Equivalence: An attacker cannot predict future output of the PRNG

Create pseudorandom numbers

- Truly random numbers are impossible with any program!
- However, we can generate seemingly random numbers, called pseudorandom numbers
- The function `rand()` returns a non-negative number between 0 and `RAND_MAX`
- For C, it is defined in `stdlib.h`

PRNGs: Summary

- True randomness requires sampling a physical process
- PRNG: An algorithm that uses a little bit of true randomness to generate a lot of random-looking output
 - Seed(entropy): Initialize internal state
 - Generate(n): Generate n bits of pseudorandom output
- Security: computationally indistinguishable from truly random bits

Lecture 10

Stream Ciphers

Stream Ciphers

- process the message bit by bit (as a stream)
- typically have a (pseudo) random **stream key**
- combined (XOR) with plaintext bit by bit
- randomness of **stream key** completely destroys any statistically properties in the message
 - $C_i = M_i \text{ XOR } \text{StreamKey}_i$
- what could be simpler!!!!
- but must never reuse stream key
 - otherwise, can remove effect and recover messages, $M \oplus K \oplus K = M$

How to generate Stream Key?

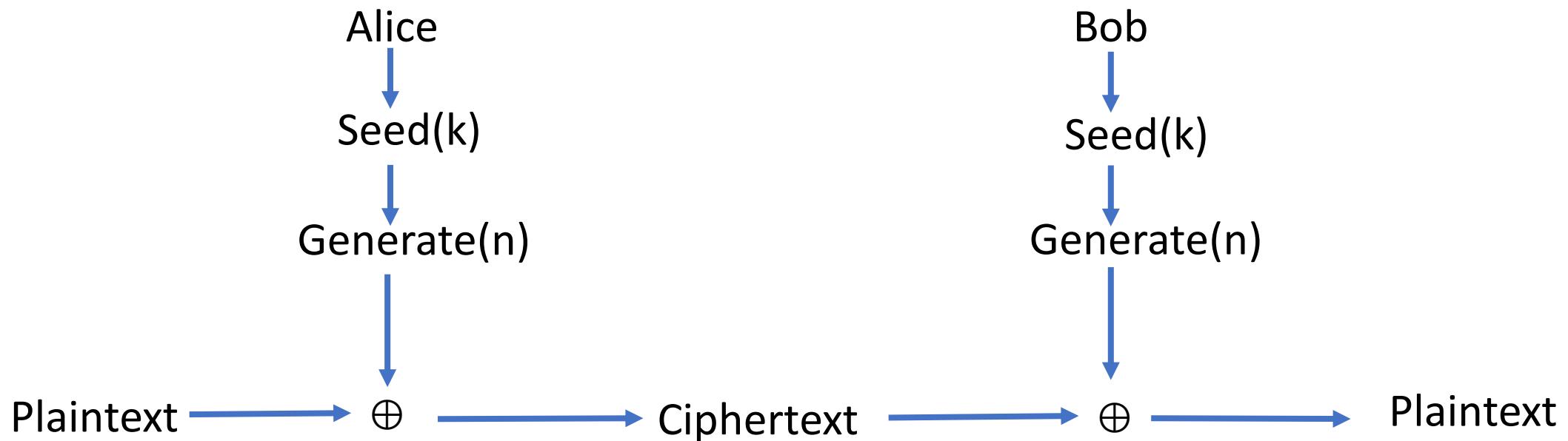
- How to generate Stream Key?

Stream Ciphers

- Idea: replace “rand” by “pseudo rand”
- Use Pseudo Random Number Generator
 - A secure PRNG produces output that looks indistinguishable from random
 - An attacker who can’t see the internal PRNG state can’t learn any output
- PRNG: $\{0,1\}^s \rightarrow \{0,1\}^n$
 - expand a short (e.g., 128-bit) random seed into a long (typically unbounded) string that “looks random”
- Secret key is the seed
 - Basic encryption method: $E_{key}[M] = M \oplus \text{PRNG}(key)$

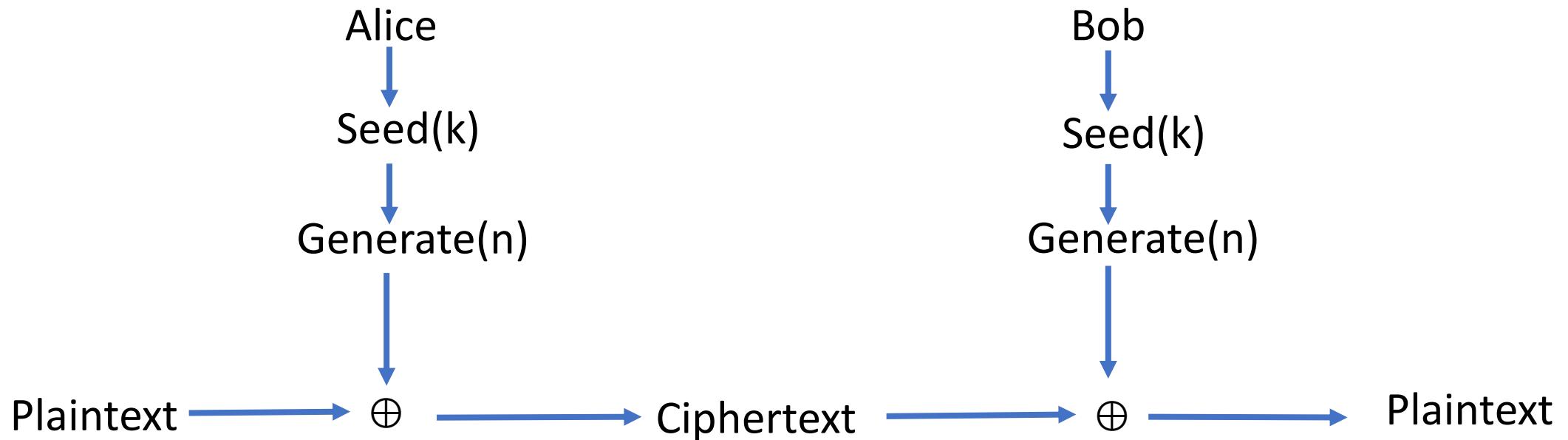
Stream Ciphers

- Protocol: Alice and Bob both seed a secure PRNG with their symmetric secret key, and then use the output as the key for stream key



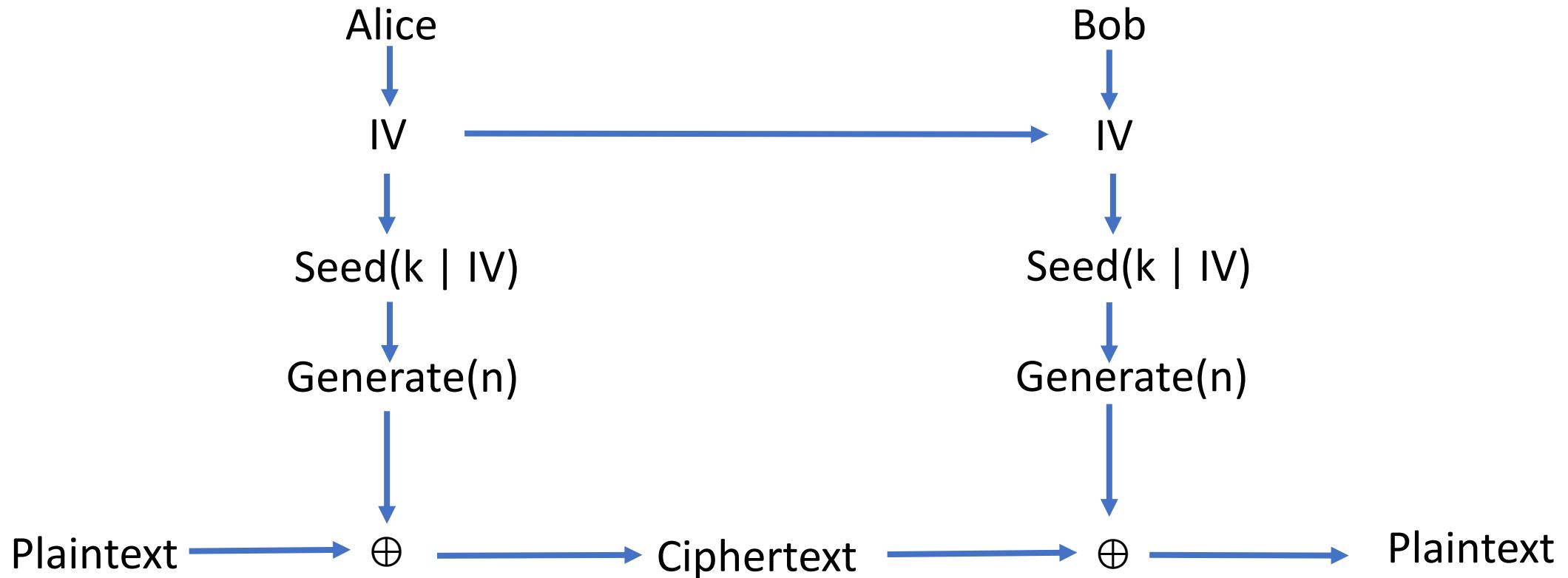
Stream Ciphers: Encrypting Multiple Messages

- How do we encrypt multiple messages without key reuses?



Stream Ciphers: Encrypting Multiple Messages

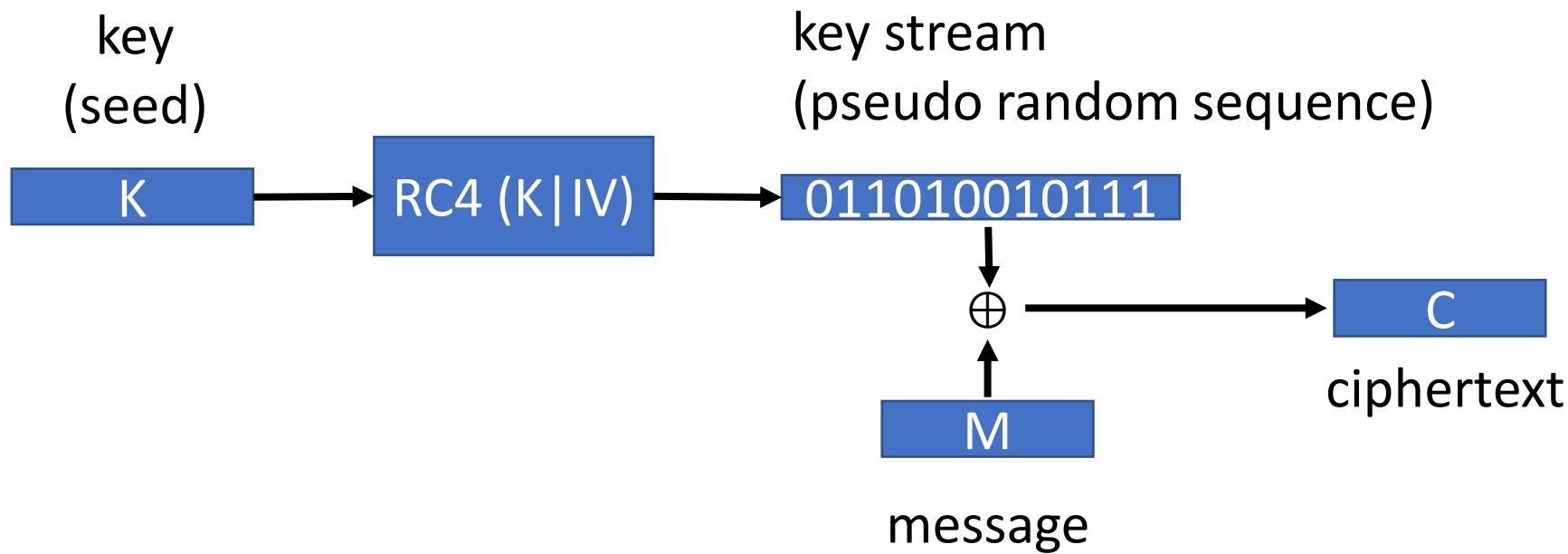
- Solution: For each message, seed the PRNG with the key and a random IV, concatenated ("|"). Send the IV with the ciphertext



Real-world example: RC4

- A proprietary cipher designed in 1987
- Extremely simple but effective!
- Very fast - especially in software
- Easily adapts to any key length, byte-oriented stream cipher
- Uses that permutation to scramble input info processed a byte at a time
- Widely used (web SSL/TLS, wireless WEP, WPA)

RC4 Stream Cipher



RC4 Key Schedule

- starts with an array S of numbers: 0...255
- use key to well and truly shuffle
- S forms internal state of the cipher
- given a key k of length L bytes

```
/* Initialization */
for i = 0 to 255 do
    S[i] = i;
    T[i] = K[i mod keylen];

/* Initial Permutation of S */
j = 0;
for i = 0 to 255 do
    j = (j + S[i] + T[i]) mod 256;
    Swap (S[i], S[j]);
```

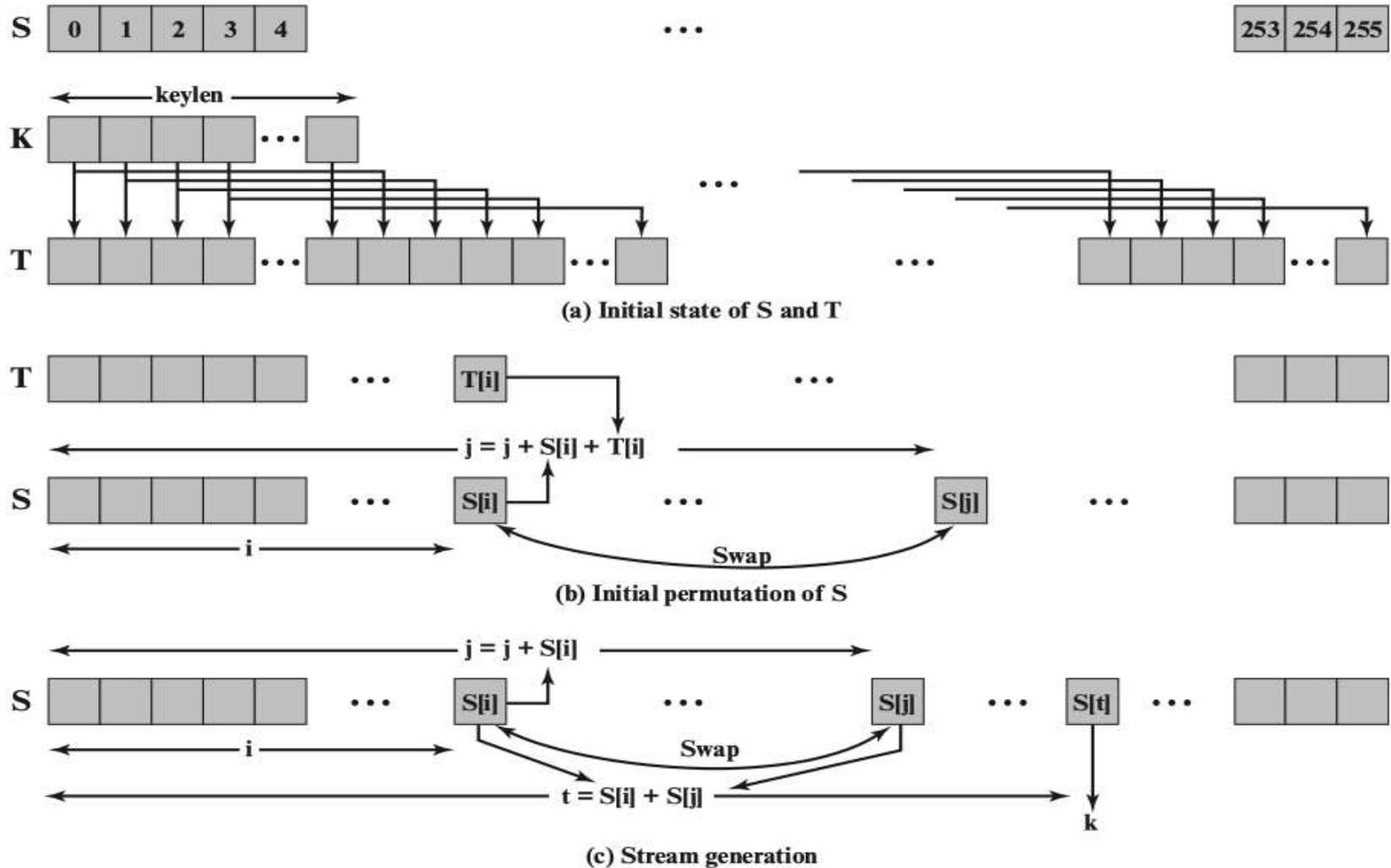
Throw away T & K, retain S

RC4 Encryption

- encryption continues shuffling array values
- sum of shuffled pair selects "stream key" value
- XOR with next byte of message to en/decrypt

```
i = j = 0
for each message byte Mi
    i = (i + 1) (mod 256)
    j = (j + S[i]) (mod 256)
    swap(S[i], S[j])
    t = (S[i] + S[j]) (mod 256)
    Ci = Mi XOR S[t]
```

RC4



Lecture 10

Stream Ciphers

Stream Ciphers

- process the message bit by bit (as a stream)
- typically have a (pseudo) random **stream key**
- combined (XOR) with plaintext bit by bit
- randomness of **stream key** completely destroys any statistically properties in the message
 - $C_i = M_i \text{ XOR StreamKey}_i$
- what could be simpler!!!!
- but must never reuse stream key
 - otherwise, can remove effect and recover messages, $M \oplus K \oplus K = M$

How to generate Stream Key?

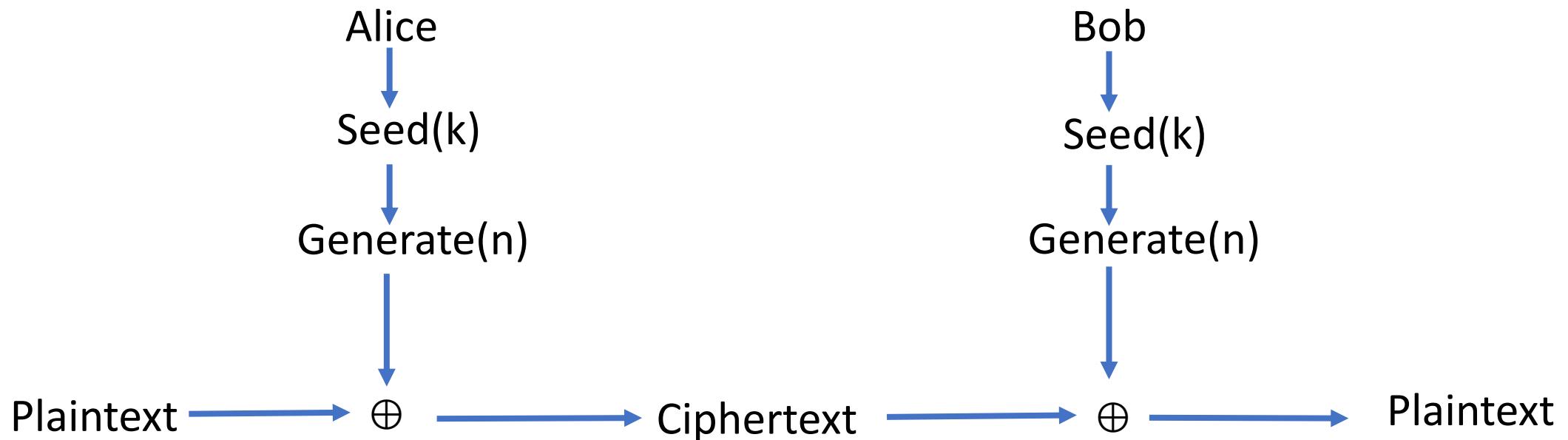
- How to generate Stream Key?

Stream Ciphers

- Idea: replace “rand” by “pseudo rand”
- Use Pseudo Random Number Generator
 - A secure PRNG produces output that looks indistinguishable from random
 - An attacker who can’t see the internal PRNG state can’t learn any output
- PRNG: $\{0,1\}^s \rightarrow \{0,1\}^n$
 - expand a short (e.g., 128-bit) random seed into a long (typically unbounded) string that “looks random”
- Secret key is the seed
 - Basic encryption method: $E_{key}[M] = M \oplus \text{PRNG}(key)$

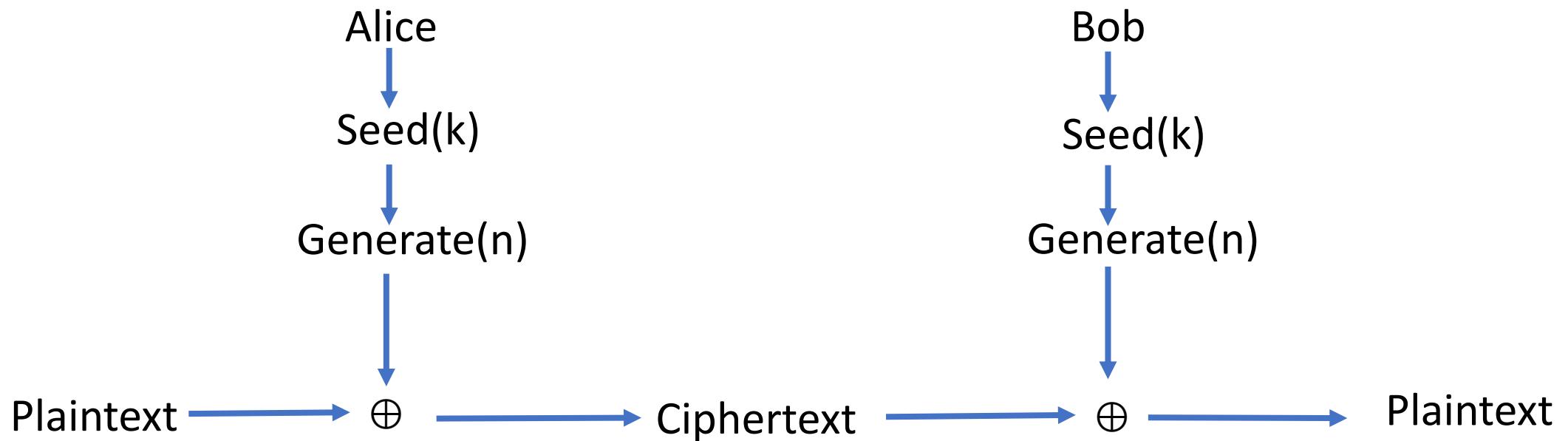
Stream Ciphers

- Protocol: Alice and Bob both seed a secure PRNG with their symmetric secret key, and then use the output as the key for stream key



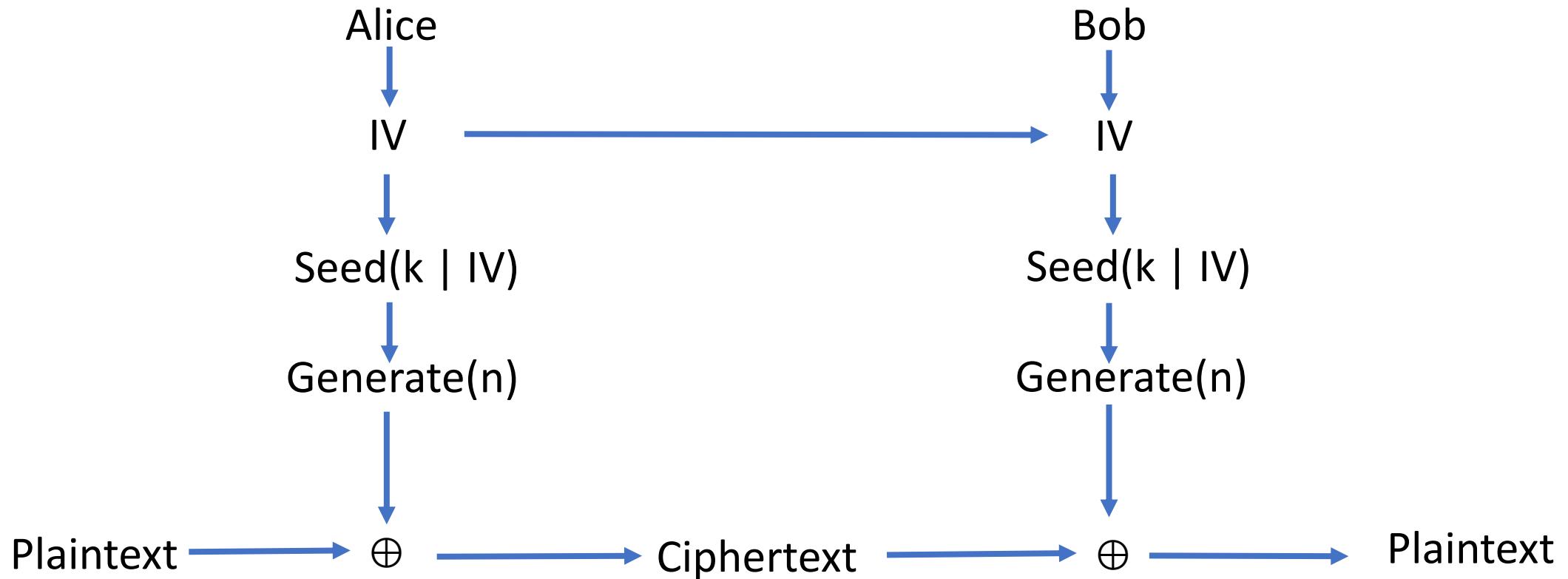
Stream Ciphers: Encrypting Multiple Messages

- How do we encrypt multiple messages without key reuses?



Stream Ciphers: Encrypting Multiple Messages

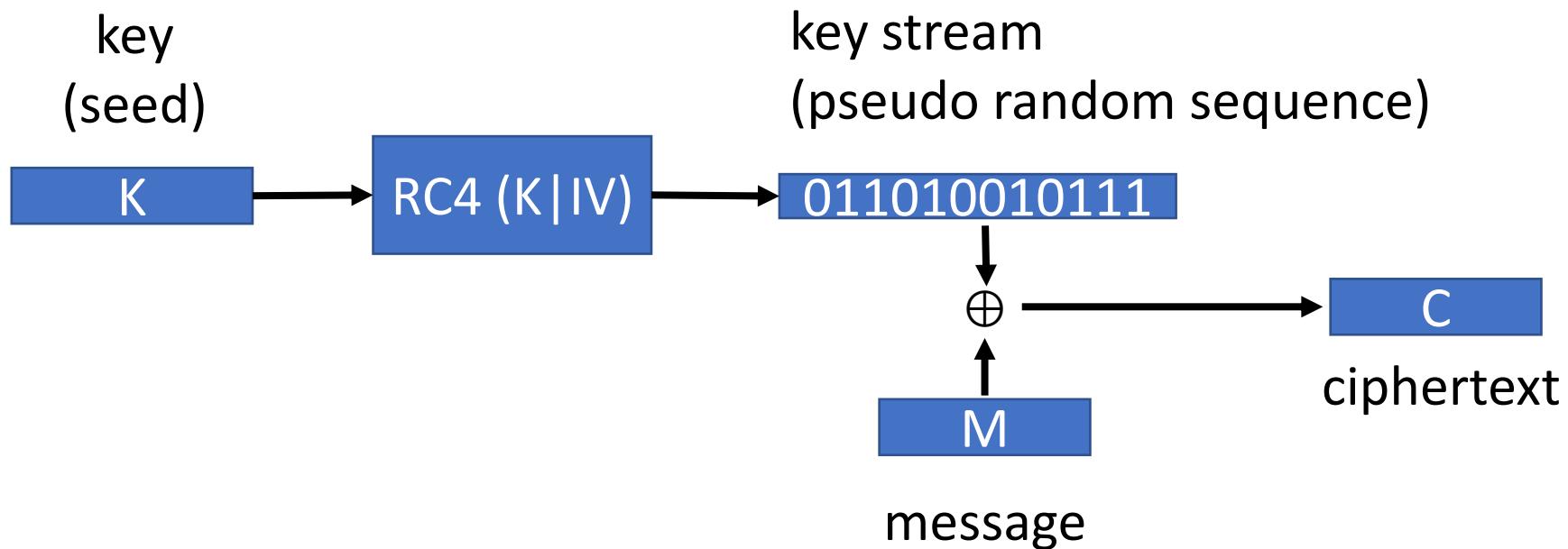
- Solution: For each message, seed the PRNG with the key and a random IV, concatenated (“|”). Send the IV with the ciphertext



Real-world example: RC4

- A proprietary cipher designed in 1987
- Extremely simple but effective!
- Very fast - especially in software
- Easily adapts to any key length, byte-oriented stream cipher
- Uses that permutation to scramble input info processed a byte at a time
- Widely used (web SSL/TLS, wireless WEP, WPA)

RC4 Stream Cipher



RC4 Key Schedule

- starts with an array S of numbers: 0...255
- use key to well and truly shuffle
- S forms internal state of the cipher
- given a key k of length L bytes

```
/* Initialization */
for i = 0 to 255 do
    S[i] = i;
    T[i] = K[i mod keylen];

/* Initial Permutation of S */
j = 0;
for i = 0 to 255 do
    j = (j + S[i] + T[i]) mod 256;
    Swap (S[i], S[j]);
```

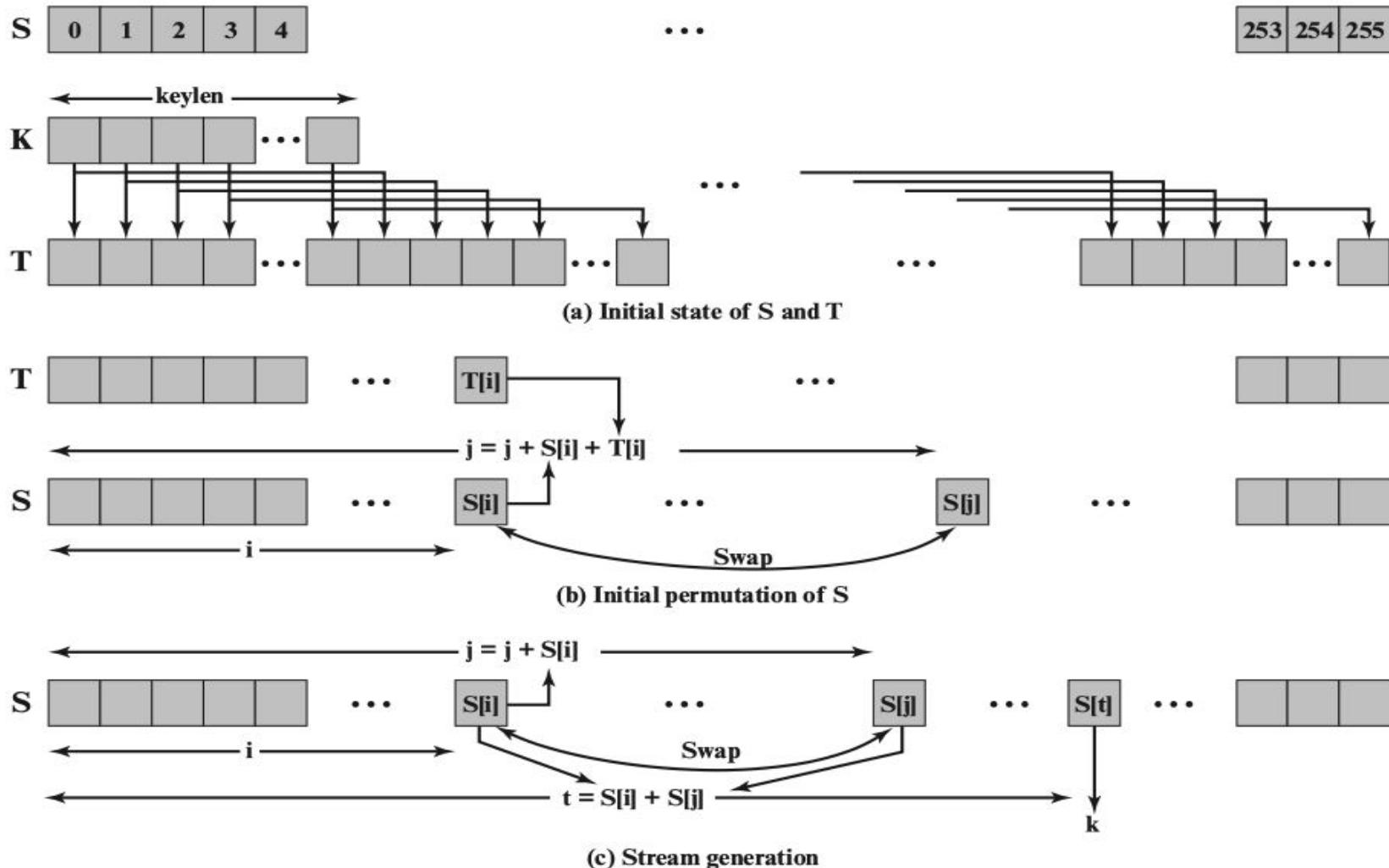
Throw away T & K, retain S

RC4 Encryption

- encryption continues shuffling array values
- sum of shuffled pair selects "stream key" value
- XOR with next byte of message to en/decrypt

```
i = j = 0
for each message byte Mi
    i = (i + 1) (mod 256)
    j = (j + S[i]) (mod 256)
    swap(S[i], S[j])
    t = (S[i] + S[j]) (mod 256)
    Ci = Mi XOR S[t]
```

RC4



Lecture 11

Summary – Chapter 2

- Symmetric block cipher
 - DES, 3DES
 - AES
- Random number
 - true random number
 - pseudorandom number
- Stream cipher
- The security of symmetric encryption depends on the secrecy of the key

Homework 1 - individual

- For Chapter 1 & 2
- Deadline: Oct. 2 (Monday), 11:59 pm
- We will use the blackboard submission time as your final timestamp
- 10% penalty per day for late submission

Network Security

Chapter 3

Public-Key Cryptography and Message Authentication

Public-Key Cryptography

Conventional cryptography

- traditional **private/secret/single-key** cryptography uses **one key**
- shared by both sender and receiver
- if this key is disclosed communications are compromised
- also is **symmetric**, parties are equal

Pros and cons

- Pros:
 - Encryption is fast for large amounts of data
 - Provide the same level of security with a shorter encryption key
 - By now, it's unbreakable to quantum computing
- Cons
 - Key distribution assumes a secure channel
 - Does not protect sender from receiver forging a message & claiming it's sent by sender
 - It does not scale well for large networks. It requires a separate key for each pair of communicating parties, which can result in a large number of keys to manage and protect.

Public-Key Cryptography

- In public-key schemes, each person has two keys
 - **Public key:** Known to everybody
 - **Private key:** Only known by that person
 - Keys come in pairs: every public key corresponds to one private key
- Uses number theory
 - Examples: Modular arithmetic, factoring, discrete logarithm problem, Elliptic logs over Elliptic Curves
 - Contrast with symmetric-key cryptography (uses XORs and bit-shifts)
- Messages are numbers
 - Contrast with symmetric-key cryptography (messages are bit strings)

Lecture 12

Public-key Cryptography

- **Benefit:** No longer need to assume that Alice and Bob already share a secret
- **Drawback:** Much slower than symmetric-key cryptography
 - Number theory calculations are much slower than XORs and bit-shifts

Reading materials

- [Encryption: Strengths and Weaknesses of Public-key Cryptography](#)
- [Public-key cryptography is a public invention due to Whitfield Diffie & Martin Hellman at Stanford Uni in 1976](#)

Public-key cryptography

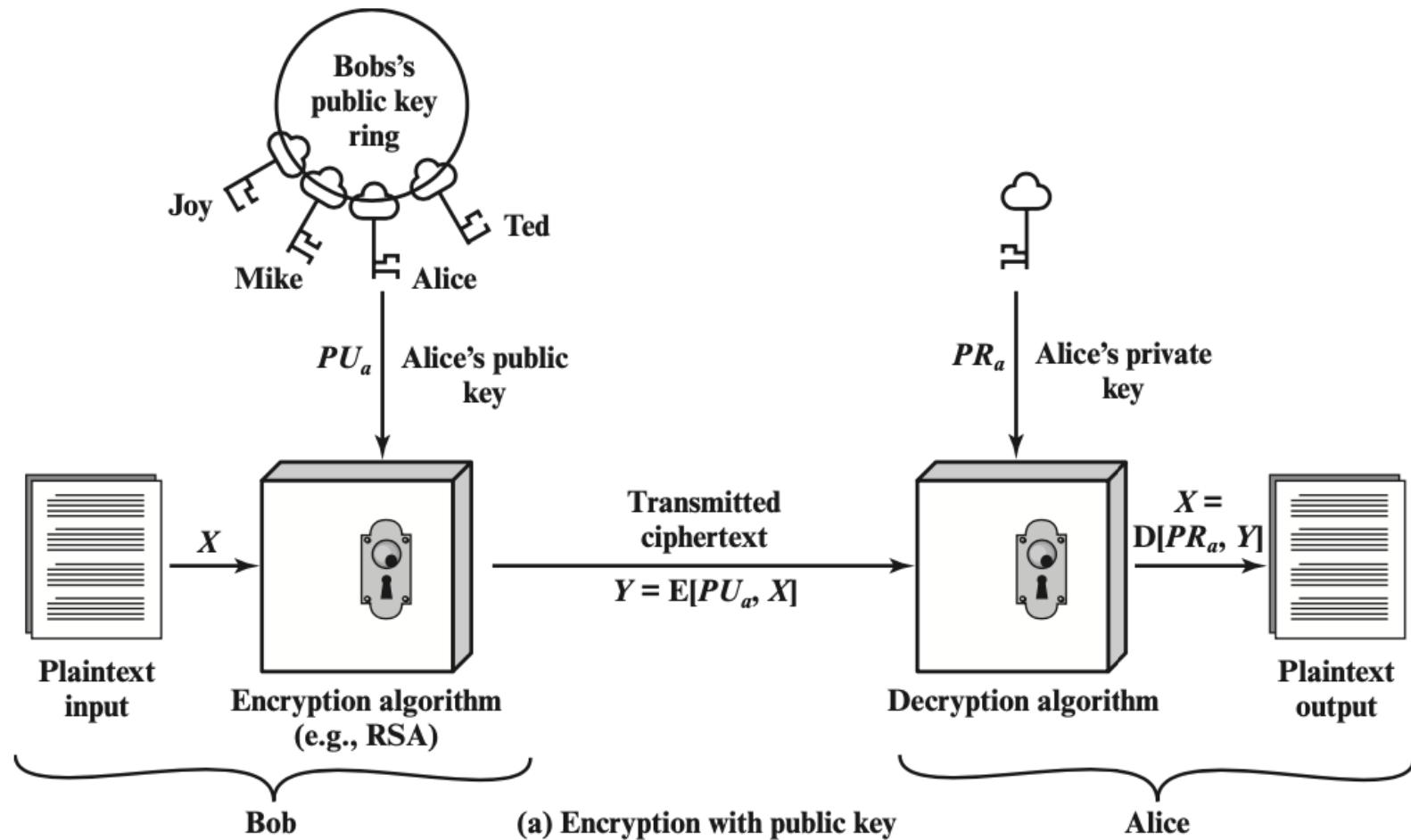
- **public-key/two-key/asymmetric** cryptography involves the use of **two keys**:
 - a **public-key**, which may be known by anybody, and can be used to **encrypt messages**, and **verify signatures**
 - a **private-key**, known only to the recipient, used to **decrypt messages**, and **sign (create) signatures**
- is **asymmetric** because
 - Not the same key
 - those who encrypt messages or verify signatures **cannot** decrypt messages or create signatures

Public-Key Encryption

- Everybody can encrypt with the public key
- Only the recipient can decrypt with the private key



Public-Key Cryptography - Encryption



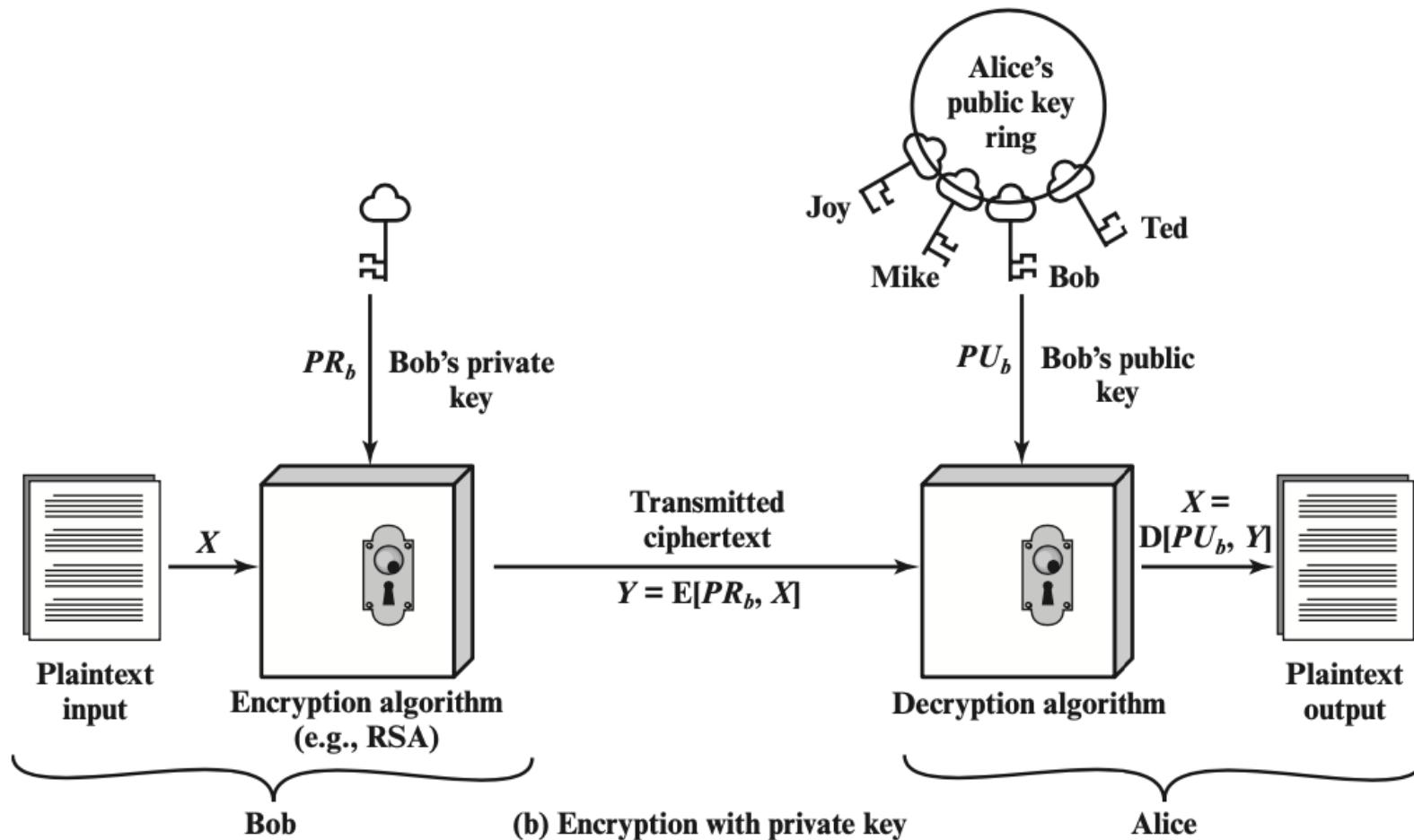
Encryption steps

- step1: generate a pair of keys
- step2: keep the private key / secret key (SK) and distribute the public key (PK) – place PK in a public register or other accessible file
- step3: Bob encrypts the message with Alice's PK
- step4: upon receiving the ciphertext (CT), Alice decrypt CT with SK

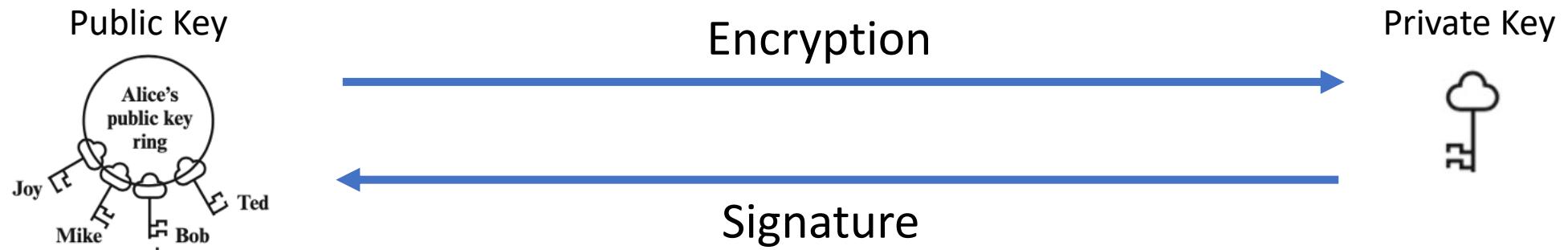
Public-Key Encryption: Definition

- Three parts:
 - $\text{KeyGen}() \rightarrow PK, SK$: Generate a public/private keypair, where PK is the public key, and SK is the private (secret) key
 - $\text{Enc}(PK, M) \rightarrow C$: Encrypt a plaintext M using public key PK to produce ciphertext C
 - $\text{Dec}(SK, C) \rightarrow M$: Decrypt a ciphertext C using secret key SK
- Properties
 - **Correctness**: Decrypting a ciphertext should result in the message that was originally encrypted
 - $\text{Dec}(SK, \text{Enc}(PK, M)) = M$ for all $PK, SK \leftarrow \text{KeyGen}()$ and M
 - **Efficiency**: Encryption/decryption should be fast
 - **Security**: 1. Alice (the challenger) just gives Eve (the adversary) the public key, and Eve doesn't request encryptions. Eve cannot guess out anything; 2. computationally infeasible to recover M with PK and ciphertext

Public-Key Cryptography - Signature



Review



Public-Key application

- can classify uses into 3 categories:
 - **encryption/decryption** (provide secrecy)
 - **digital signatures** (provide authentication)
 - **key exchange** (of session keys)
- some algorithms are suitable for all uses; others are specific to one
- Either of the two related keys can be used for encryption, with the other used for decryption

Algorithm	Encryption/Decryption	Digital Signature	Key Exchange
RSA	Yes	Yes	Yes
Diffie–Hellman	No	No	Yes
DSS	No	Yes	No
Elliptic curve	Yes	Yes	Yes

Security of Public Key Schemes

- Keys used are **very large** (>512bits)
 - like private key schemes brute force **exhaustive search** attack is always theoretically possible
- Security relies on a large enough difference in **difficulty** between easy (en/decrypt) and hard (cryptanalyze) problems
 - more generally the hard problem is known, it's just made too hard to do in practice
- Requires the use of **very large numbers**, hence is **slow** compared to private/symmetric key schemes

Public-Key Cryptography Algorithm (RSA)

RSA Public-key encryption

- by Rivest, Shamir & Adleman of MIT in 1977
- currently the “work horse” of Internet security
 - most public key infrastructure (PKI) products
 - SSL/TLS: certificates and key-exchange
 - secure e-mail: PGP, Outlook,
- based on exponentiation in a finite (Galois) field over integers modulo a prime
 - exponentiation takes $O((\log n)^3)$ operations (easy)
- security due to cost of factoring large integer numbers
 - factorization takes $O(e^{\log n \log \log n})$ operations (hard)
- uses large integers (eg. 1024 bits)

RSA key setup

- each user generates a public/private key pair by:
 - selecting two large primes at random - p, q
 - computing their system modulus $n=p \cdot q$
 - note $\phi(n) = (p-1)(q-1)$
 - selecting at random the encryption key e
 - where $1 < e < \phi(n)$, $\gcd(e, \phi(n)) = 1$
 - solve following equation to find decryption key d
 - $ed \equiv 1 \pmod{\phi(n)}$
 - publish their public encryption key: $pk=\{e,n\}$
 - keep secret private decryption key: $sk=\{d,p,q\}$

Key Generation	
Select p, q	p and q both prime, $p \neq q$
Calculate $n = p \times q$	
Calculate $\phi(n) = (p-1)(q-1)$	
Select integer e	$\gcd(\phi(n), e) = 1; 1 < e < \phi(n)$
Calculate d	$de \pmod{\phi(n)} = 1$
Public key	$KU = \{e, n\}$
Private key	$KR = \{d, n\}$

RSA example

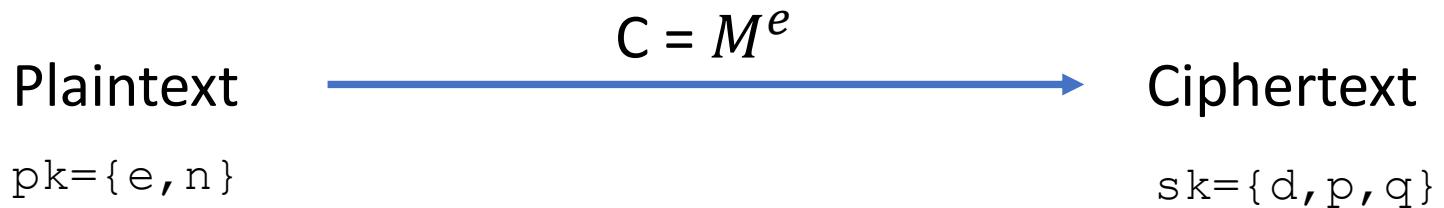
1. Select primes: $p=17$ & $q=11$
2. Compute $n = pq = 17 \times 11 = 187$
3. Compute $\phi(n) = (p-1)(q-1) = 16 \times 10 = 160$
4. Select e : $\gcd(e, 160) = 1$; choose $e=7$
5. Determine d : $de \equiv 1 \pmod{160}$ and $d < 160$ Value is $d=23$ since $23 \times 7 = 161 = 10 \times 160 + 1$
6. Publish public key $pk = \{7, 187\}$
7. Keep secret private key $sk = \{23, 17, 11\}$

Key Generation

Select p, q	p and q both prime, $p \neq q$
Calculate $n = p \times q$	
Calculate $\phi(n) = (p-1)(q-1)$	
Select integer e	$\gcd(\phi(n), e) = 1; 1 < e < \phi(n)$
Calculate d	$de \pmod{\phi(n)} = 1$
Public key	$KU = \{e, n\}$
Private key	$KR = \{d, n\}$

RSA use

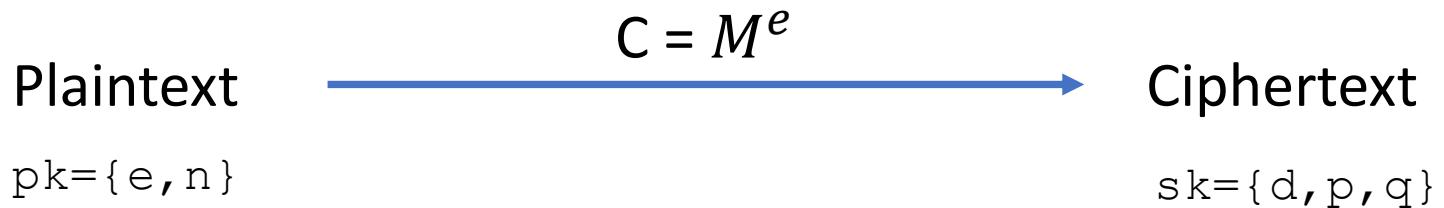
- to encrypt a message M the sender:
 - obtains **public key** of recipient $pk = \{e, n\}$
 - computes: $C = M^e \pmod{n}$, where $0 \leq M < n$
- to decrypt the ciphertext C the owner:
 - uses their **private key** $sk = \{d, p, q\}$
 - computes: $M = C^d \pmod{n}$
- note that the message M must be smaller than the modulus n (block if needed)



Lecture 14

RSA use

- to encrypt a message M the sender:
 - obtains **public key** of recipient $pk = \{e, n\}$
 - computes: $C = M^e \pmod{n}$, where $0 \leq M < n$
- to decrypt the ciphertext C the owner:
 - uses their **private key** $sk = \{d, p, q\}$
 - computes: $M = C^d \pmod{n}$
- note that the message M must be smaller than the modulus n (block if needed)



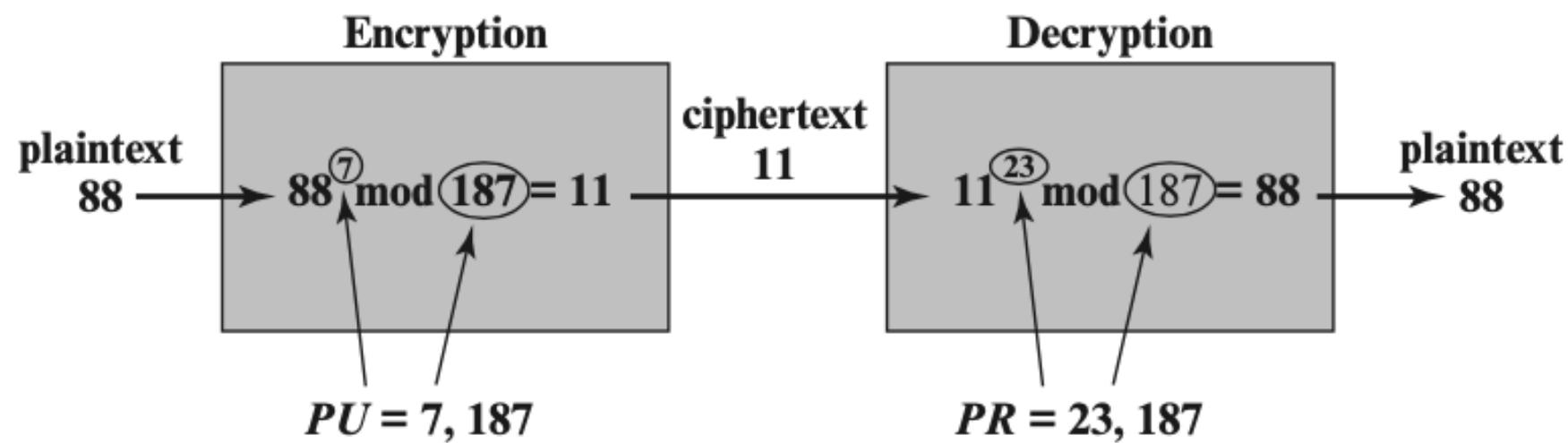
RSA example continue

- sample RSA encryption/decryption is:
- given message $M = 88$ ($88 < 187$)
- encryption:

$$C = 88^7 \bmod 187 = 11$$

- decryption:
- $$M = 11^{23} \bmod 187 = 88$$

Example of RSA algorithm



RSA key generation

- users of RSA must:
 - determine two primes at random - p, q
 - select either e or d and compute the other
- primes p, q must not be easily derived from modulus $n=p \cdot q$
 - means must be sufficiently large
 - typically guess and use probabilistic test
- exponents e, d are inverses, so use Inverse algorithm to compute the other

Correctness of RSA

- Euler's theorem: if $\gcd(M, n) = 1$, then $M^{\phi(n)} \equiv 1 \pmod{n}$. Here $\phi(n)$ is Euler's totient function: the number of integers in $\{1, 2, \dots, n-1\}$ which are relatively prime to n . When n is a prime, this theorem is just Fermat's little theorem

$$\begin{aligned} M' &= C^d \pmod{n} = M^{ed} \pmod{n} \\ &= M^{k\phi(n)+1} \pmod{n} \\ &= [M^{\phi(n)}]^k \cdot M \pmod{n} \\ &= M \pmod{n} \end{aligned}$$

Encryption	
Plaintext:	$M < n$
Ciphertext:	$C = M^e \pmod{n}$

Attack approaches

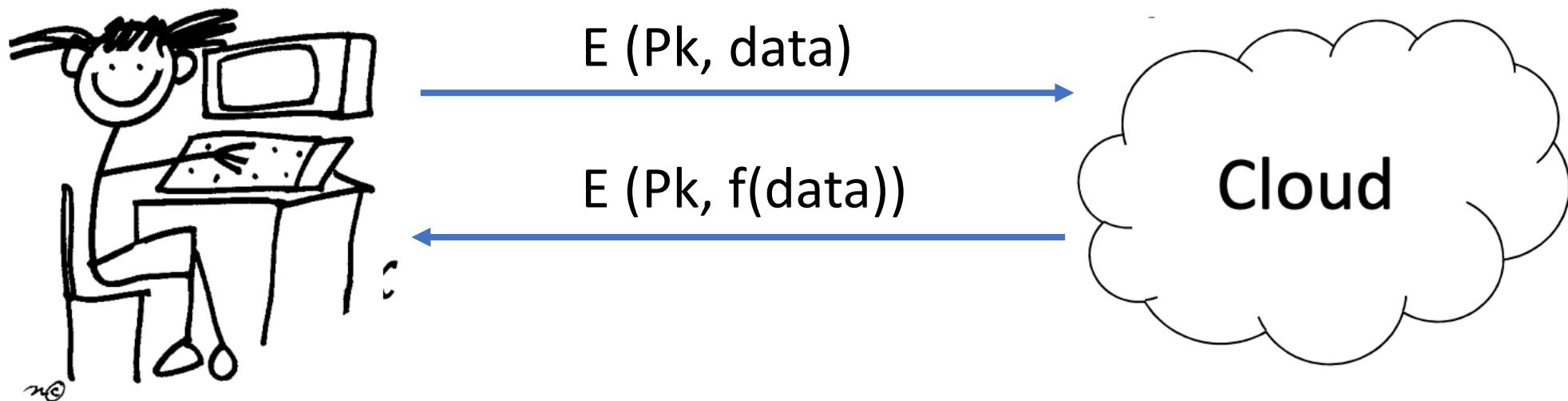
- **Mathematical attacks:** several approaches, all equivalent in effort to factoring the product of two primes. The defense against mathematical attacks is to use a large key size.
- **Timing attacks:** These depend on the running time of the decryption algorithm
- **Chosen ciphertext attacks:** this type of attacks exploits properties of the RSA algorithm by selecting blocks of data. These attacks can be thwarted by suitable padding of the plaintext, such as PKCS1 V1.5 in SSL

Homomorphic encryption

- Encryption scheme that allows computation on ciphertexts
 - i.e. a public-key encryption scheme that allows anyone in possession of the public key to perform operations on encrypted data without access to the decryption key
- Initial public-key systems that allow this for either addition or multiplication, but not both.
- Fully homomorphic encryption (FHE)

Application of homomorphic encryption

- One Use case: cloud computing
 - A weak computational device Alice (e.g., a mobile phone or a laptop) wishes to perform a computationally heavy task, beyond her computational means. She can delegate it to a much stronger (but still feasible) machine Bob (the cloud, or a supercomputer) who offers the service of doing so. The problem is that Alice does not trust Bob, who may give the wrong answer due to laziness, fault, or malice.



RSA reading materials

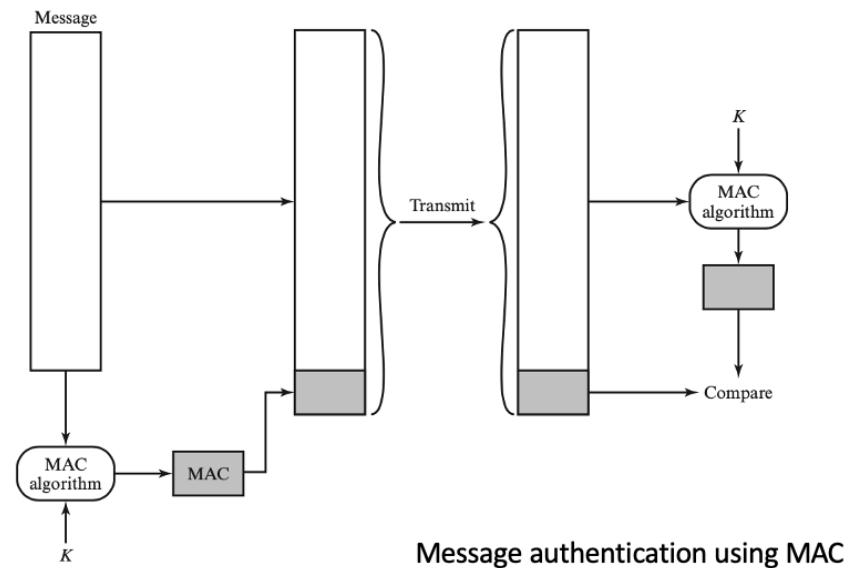
- A Method for Obtaining Digital Signatures and Public-Key Cryptosystems

Lecture 15

Message Authentication

Message authentication

- message authentication is concerned with:
 - protecting the integrity of a message
 - validating identity of originator
 - non-repudiation of origin (dispute resolution)
- then three alternative functions used:
 - message encryption - symmetric
 - message authentication code (MAC)
 - digital signature



Message encryption

- Symmetric message encryption by itself also provides a measure of authentication
- if symmetric encryption is used then:
 - receiver knows sender must have created it
 - since only sender and receiver know key used
 - know content cannot be altered

Homework 1 questions

- Q1: Symmetric Block Cypher provides authentication and confidentiality
 - Ans: True

Message encryption

- if public-key encryption is used:
 - encryption provides no confidence of sender
 - since anyone potentially knows public-key
 - so, need to recognize corrupted messages
 - however, if
 - sender **signs** message using their private-key
 - then encrypts with recipients' public key
 - have both secrecy and authentication
 - but at cost of two public-key uses on message

Reasons to avoid encryption authentication

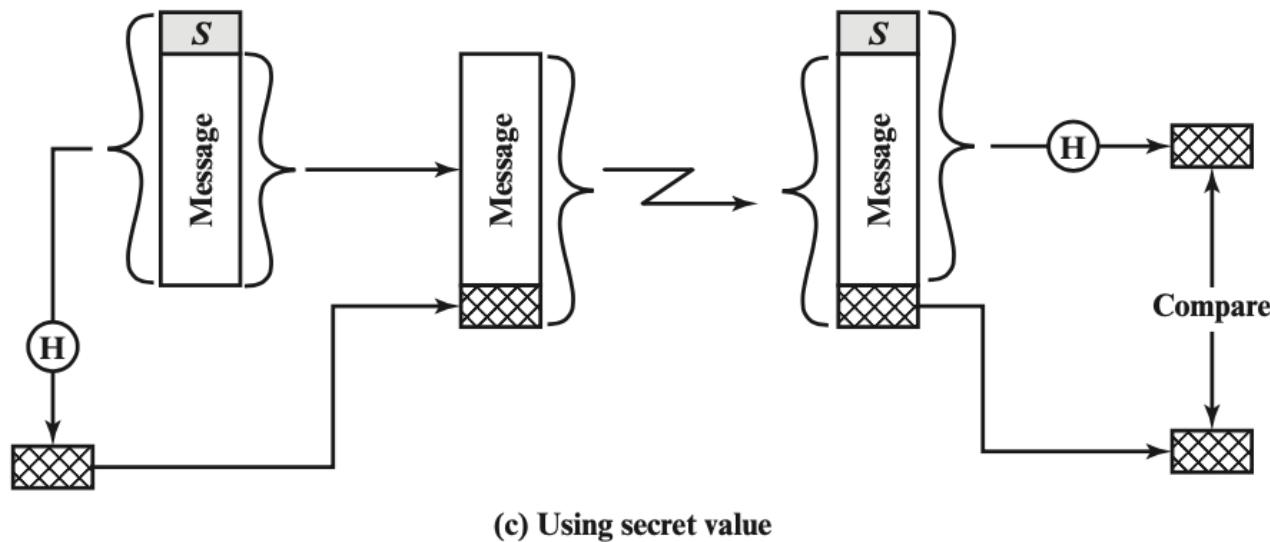
- Encryption software is quite slow
- Encryption hardware costs are nonnegligible
- Encryption hardware is optimized toward large data sizes
- An encryption algorithm may be protected by a patent

Hash Function

Hash functions

- Hash function: $h = H(M)$
 - M can be of any size
 - h is always of fixed size
 - Typically, $h \ll \text{size}(M)$

One use case - using hash function



- Initialization: A and B share a common secret, S_{AB}
- Message, M
- A calculates $MD_M = H(S_{AB} || M)$
- B recalculates MD'_M , and check
- $MD'_M = MD_M$

This scheme cannot provide authentication.

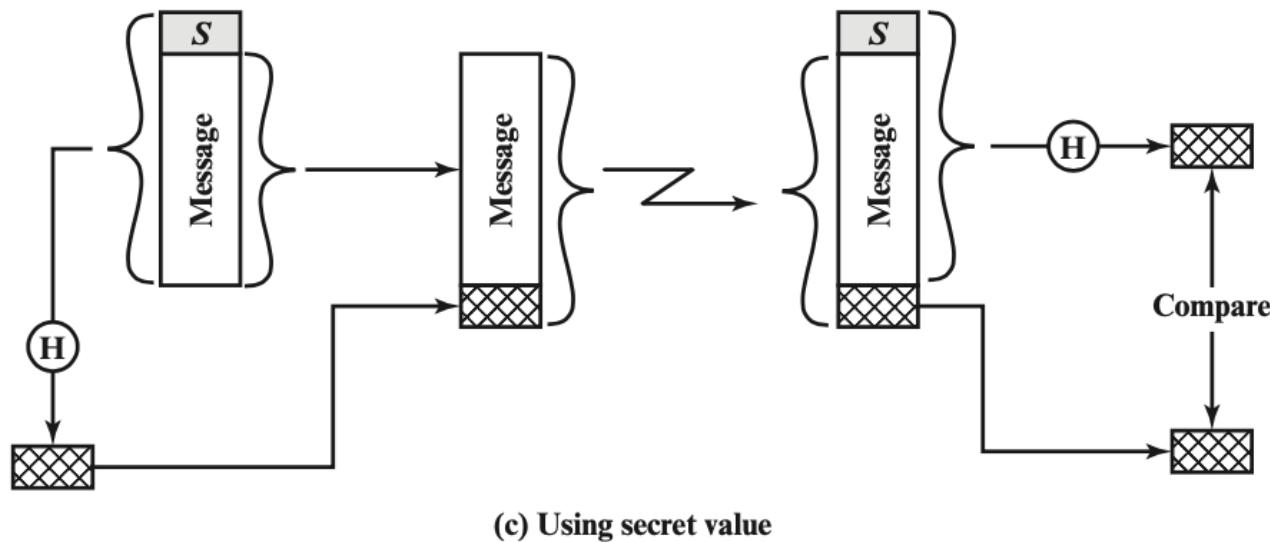
Lecture 16

Hash Function

Hash functions

- Hash function: $h = H(M)$
 - M can be of any size
 - h is always of fixed size
 - Typically, $h \ll \text{size}(M)$

One use case - using hash function



- Initialization: A and B share a common secret, S_{AB}
- Message, M
- A calculates $MD_M = H(S_{AB} || M)$
- B recalculates MD'_M , and check
- $MD'_M = MD_M$

This scheme cannot provide authentication.

Requirements for secure hash functions

- 1. can be applied to any sized message M
- 2. produces fixed-length output h
- 3. is easy to compute $h=H(M)$ for any message M
- 4. given h is infeasible to find x s.t. $H(x)=h$
 - one-way property or preimage resistance
- 5. given x is infeasible to find x' s.t. $H(x')=H(x)$
 - weak collision resistance or second pre-image resistant
- 6. infeasible to find **any pair** of x, x' s.t. $H(x')=H(x)$
 - strong collision resistance

Hash Function: Collision Resistance

- **Collision:** Two different inputs with the same output
 - $x \neq x'$ and $H(x) = H(x')$
 - Can we design a hash function with no collisions?
 - No, because there are more inputs than outputs (pigeonhole principle)
 - However, we want to make finding collisions *infeasible* for an attacker
- **Collision resistance:** It is infeasible to (i.e. no polynomial time attacker can) find any pair of inputs $x' \neq x$ such that $H(x) = H(x')$

Secure hash function

- A hash function that satisfies the first five properties is referred to as a weak hash function
- **Security:** random/unpredictability, no predictable patterns for how changing the input affects the output
 - Changing 1 bit in the input causes the output to be completely different
 - Also called “random oracle” assumption
- A message digest
 - a fixed size numeric representation of the contents of a message, computed by a hash function
- Examples: SHA-1 (Secure Hash Algorithm 1), SHA-2, SHA-3, MD5

Hash Function: Examples

- MD5
 - Output: 128 bits
 - Security: Completely broken
- SHA-1
 - Output: 160 bits
 - Security: Completely broken in 2017
 - Was known to be weak before 2017, but still used sometimes
- SHA-2
 - Output: 256, 384, or 512 bits (sometimes labeled SHA-256, SHA-384, SHA-512)
 - Not currently broken, but some variants are vulnerable to a length extension attack
 - Current standard
- SHA-3 (Keccak)
 - Output: 256, 384, or 512 bits
 - Current standard (not meant to replace SHA-2, just a different construction)

Length Extension Attacks

- **Length extension attack:** Given $H(x)$ and the length of x , but not x , an attacker can create $H(x \parallel m)$ for any m of the attacker's choosing
 - [Length extension attack - Wikipedia](#)
- SHA-256 (256-bit version of SHA-2) is vulnerable
- SHA-3 is not vulnerable

Lecture 17

Length Extension Attacks

- **Length extension attack:** Given $H(x)$ and the length of x , but not x , an attacker can create $H(x \parallel m)$ for any m of the attacker's choosing
 - [Length extension attack - Wikipedia](#)
- SHA-256 (256-bit version of SHA-2) is vulnerable
- SHA-3 is not vulnerable

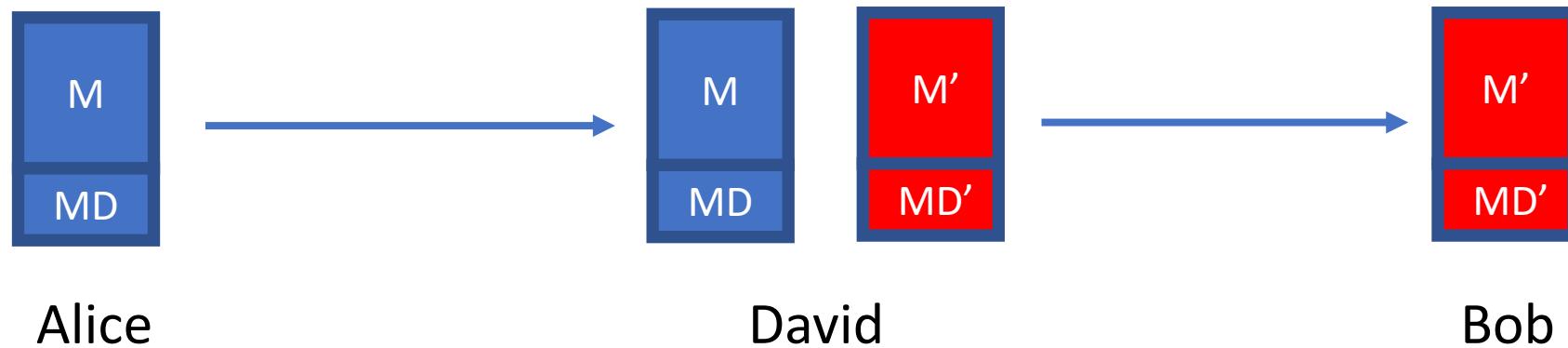
Does hashes provide integrity?

- It depends on your threat model
- Scenario
 - Mozilla publishes a new version of Firefox on some download servers
 - Alice downloads the program binary
 - How can she be sure that nobody tampered with the program?
- Idea: use cryptographic hashes
 - Mozilla hashes the program binary and publishes the hash on its website
 - Alice hashes the binary she downloaded and checks that it matches the hash on the website
 - If Alice downloaded a malicious program, the hash would not match (tampering detected!)
 - An attacker can't create a malicious program with the same hash (collision resistance)
- Threat model: We assume the attacker cannot modify the hash on the website
 - We have integrity, as long as we can communicate the hash securely

Do hashes provide integrity?

- It depends on your threat model
- Scenario
 - Alice and Bob want to communicate over an insecure channel
 - David might tamper with messages
- Idea: Use cryptographic hashes
 - Alice sends her message with a cryptographic hash over the channel
 - Bob receives the message and computes a hash on the message
 - Bob checks that the hash he computed matches the hash sent by Alice
- Threat model: David can modify the message *and the hash*
 - No integrity!

Man-in-the-middle attack



Do hashes provide integrity?

- It depends on your threat model
- If the attacker can modify the hash, hashes don't provide integrity
- Main issue: Hashes are *unkeyed* functions
 - There is no secret key being used as input, so any attacker can compute a hash on any value

Solutions

- A message digest created using a secret **symmetric key** is known as a Message Authentication Code (MAC), because it can provide assurance that the message has not been modified
- The sender can also generate a message digest and then encrypt the digest using the private key of an **asymmetric key** pair, forming a **digital signature**. The signature must then be verified by the receiver through comparing it with a locally generated digest

Hashes: Summary

- Map arbitrary-length input to fixed-length output
- Output is deterministic
- Security properties
 - One way: Given an output y , it is infeasible to find any input x such that $H(x) = y$.
 - Second preimage resistant: Given an input x , it is infeasible to find another input $x' \neq x$ such that $H(x) = H(x')$.
 - Collision resistant: It is infeasible to find another any pair of inputs $x' \neq x$ such that $H(x) = H(x')$.
- Some hashes are vulnerable to length extension attacks
- Hashes don't provide integrity (unless you can publish the hash securely)

Lecture 18

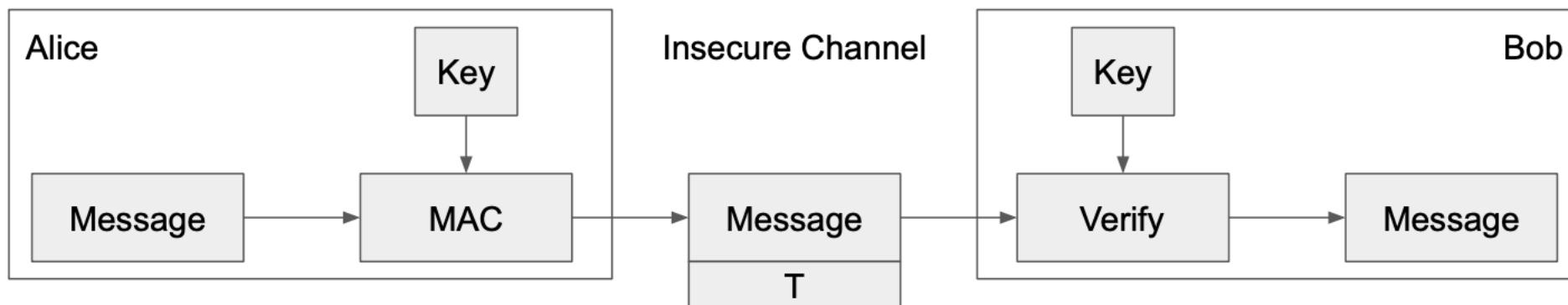
Message Authentication Code

Message authentication code (MAC)

- generated by an algorithm that creates a small fixed-sized block
 - depending on both message and some key
 - not be reversible
 - $\text{MAC}_M = F(K_{AB}, M)$
- appended to message as a signature
- receiver performs same computation on message and checks it matches the MAC
- provides assurance that message is unaltered and comes from sender

MACs: Usage

- Alice wants to send M to Bob, but doesn't want David to tamper with it
- Alice sends M and $T = \text{MAC}(K, M)$ to Bob
- Bob receives M and T
- Bob computes $\text{MAC}(K, M)$ and checks that it matches T
- If the MACs match, Bob is confident the message has not been tampered with (integrity)



MACs: Definition

- Two parts:
 - $\text{KeyGen}() \rightarrow K$: Generate a key K
 - $\text{MAC}(K, M) \rightarrow T$: Generate a tag T for the message M using key K
 - Inputs: A secret key and an arbitrary-length message
 - Output: A fixed-length **tag** on the message
- Properties
 - **Correctness:** Determinism
 - Note: Some more complicated MAC schemes have an additional $\text{Verify}(K, M, T)$ function that don't require determinism, but this is out of scope
 - **Efficiency:** Computing a MAC should be efficient
 - **Security:** existentially unforgeable under chosen plaintext attack

Existentially unforgeable

- A secure MAC is **existentially unforgeable**: without the key, an attacker cannot create a valid tag on a message
 - David cannot generate $\text{MAC}(K, M')$ without K
 - David cannot find any $M' \neq M$ such that $\text{MAC}(K, M') = \text{MAC}(K, M)$

Example: HMAC

- issued as RFC 2104 [1]
- has been chosen as the mandatory-to-implement MAC for IP Security
- Used in Transport Layer Security (TLS) and Secure Electronic Transaction (SET)

[1] “HMAC: Keyed-Hashing for Message Authentication”, RFC 2104, <https://datatracker.ietf.org/doc/html/rfc2104>

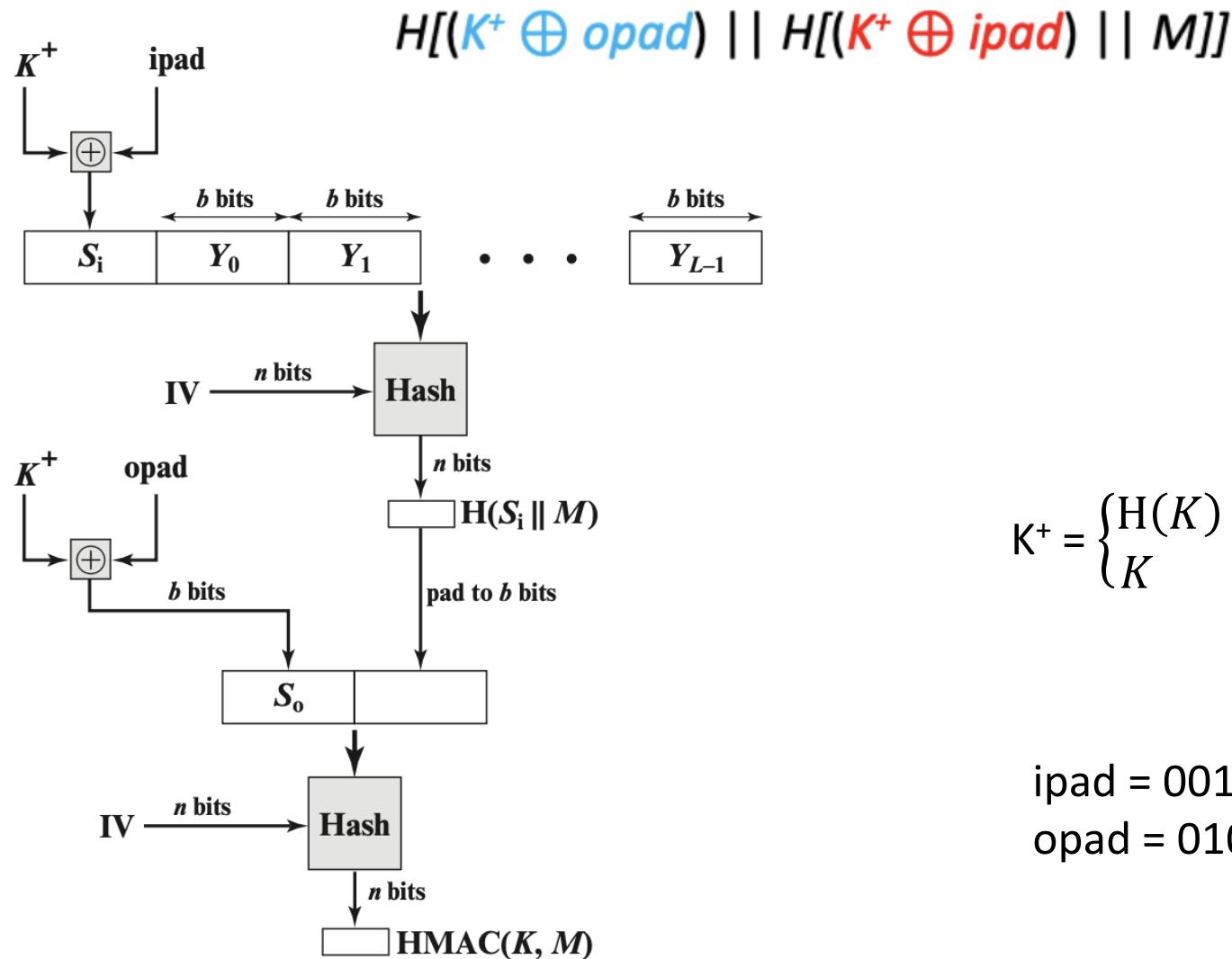
HMAC(K , M)

- will produce two keys to increase security
- If key is longer than the desired size, we can hash it first, but be careful with using keys that are too much smaller, they have to have enough randomness in them
- Output $H[(K^+ \oplus opad) \parallel H[(K^+ \oplus ipad) \parallel M]]$

Example: HMAC

- $\text{HMAC}(K, M)$:
 - Output $H[(\textcolor{blue}{K^+} \oplus \textit{opad}) \parallel H[(\textcolor{red}{K^+} \oplus \textit{ipad}) \parallel M]]$
- Use K to derive two different keys
 - \textit{opad} (outer pad) is the hard-coded byte **0x5c** repeated until it's the same length as K^+
 - \textit{ipad} (inner pad) is the hard-coded byte **0x36** repeated until it's the same length as K^+
 - As long as \textit{opad} and \textit{ipad} are different, you'll get two different keys
 - For paranoia, the designers chose two very different bit patterns, even though they theoretically need only differ in one bit

HMAC



A	B	$A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

$$K^+ = \begin{cases} H(K) & K \text{ is larger than block size} \\ K & \text{otherwise} \end{cases}$$

ipad = 00110110 , repeat b/8 times
opad = 01011100, repeat b/8 times

Figure 3.6 HMAC Structure

HMAC procedure

$$H[(K^+ \oplus opad) || H[(K^+ \oplus ipad) || M]]$$

- Step 1: Append zeros to the left end of K to create a b -bit string K^+ (e.g., if K is of length 160 bits and $b = 512$, then K will be appended with 44 zero bytes);
- Step 2: XOR (bitwise exclusive-OR) K^+ with ipad to produce the b -bit block S_i ;
- Step 3: Append M to S_i ;
- Step 4: Apply H to the stream generated in step 3;
- Step 5: XOR K^+ with opad to produce the b -bit block S_o ;
- Step 6: Append the hash result from step 4 to S_o ;
- Step 7: Apply H to the stream generated in step 6 and output the result.

Mid-term Exam

- Nov. 3, 2023 (Friday), 4:00 pm – 4:50 pm, in class
- Closed book
- Chapter 1 – 3
- Will have a review class

Lecture 19

Outline

- Review of MAC authentication
- Authenticated encryption

HMAC Properties

- $\text{HMAC}(K, M) = H[(K^+ \oplus \text{opad}) || H((K^+ \oplus \text{ipad}) || M)]$
- HMAC is a hash function, so it has the properties of the underlying hash too
 - It is collision resistant
 - Given $\text{HMAC}(K, M)$, an attacker can't learn M – *one way*
 - If the underlying hash is secure, HMAC doesn't reveal M , but it is still deterministic
- You can't verify a tag T if you don't have K
- This means that an attacker can't brute-force the message M without knowing K

MACs: Summary

- Inputs: a secret key and a message
- Output: a tag on the message
- A secure MAC is unforgeable: Even if David can trick Alice into creating MACs for messages that David chooses, David cannot create a valid MAC on a message that she hasn't seen before
 - Example: $\text{HMAC}(K, M) = H((K^+ \oplus \text{opad}) || H((K^+ \oplus \text{ipad}) || M))$
- MACs do not provide confidentiality

Do MACs provide integrity?

- Do MACs provide integrity?
 - Yes. An attacker cannot tamper with the message without being detected
- Do MACs provide authenticity?
 - It depends on your threat model
 - If only two people have the secret key, MACs provide authenticity: it has a valid MAC, and it's not from me, so it must be from the other person
 - More than one secret key, If a message has a valid MAC, you can be sure it came from *someone with the secret key*, but you can't narrow it down to one person
- Do MACs provide confidentiality?

Authenticated Encryption

Authenticated Encryption: Definition

- **Authenticated encryption (AE)**: A scheme that simultaneously guarantees confidentiality and integrity (and authenticity, depending on your threat model) on a message
- Two ways of achieving authenticated encryption:
 - Combine schemes that provide confidentiality with schemes that provide integrity
 - Use a scheme that is designed to provide confidentiality and integrity

Scratchpad: Let's design it together

- You can use:
 - An encryption scheme (e.g. AES-CBC): $\text{Enc}(K, M)$ and $\text{Dec}(K, M)$
 - An unforgeable MAC scheme (e.g. HMAC): $\text{MAC}(K, M)$
- First attempt: Alice sends $\text{Enc}(K_1, M)$ and $\text{MAC}(K_2, M)$
 - Integrity? Yes, attacker can't tamper with the MAC
 - Confidentiality? No, the MAC is not secure
- Idea 1: Let's compute the MAC on the *ciphertext* instead of the plaintext: $\text{Enc}(K_1, M)$ and $\text{MAC}(k_2, \text{Enc}(K_1, M))$
 - Integrity? Yes, attacker can't tamper with the MAC
 - Confidentiality? Yes, the MAC might leak info about the ciphertext, but that's okay
- Idea 2: Let's encrypt the MAC too: $\text{Enc}(K_1, M || \text{MAC}(K_2, M))$
 - Integrity? Yes, attacker can't tamper with the MAC
 - Confidentiality? Yes, everything is encrypted

MAC-then-Encrypt or Encrypt-then-MAC?

- Method 1: Encrypt-then-MAC
 - First compute $\text{Enc}(K_1, M)$
 - Then MAC the ciphertext: $\text{MAC}(K_2, \text{Enc}(K_1, M))$
- Method 2: MAC-then-encrypt
 - First compute $\text{MAC}(K_2, M)$
 - Then encrypt the message and the MAC together: $\text{Enc}(k_1, M \mid\mid \text{MAC}(K_2, M))$
- Which is better?
 - In theory, both are secure if applied properly
 - MAC-then-encrypt has a flaw: You don't know if tampering has occurred until after decrypting
 - Attacker can supply arbitrary tampered input, and you always have to decrypt it
 - Passing attacker-chosen input through the decryption function can cause side-channel leaks
- **Always use encrypt-then-MAC** because it's more robust to mistakes

TLS 1.0 “Lucky 13” Attack

- TLS: A protocol for sending encrypted and authenticated messages over the Internet
- TLS 1.0 uses MAC-then-encrypt: $\text{Enc}(k_1, M \parallel \text{MAC}(k_2, M))$
 - The encryption algorithm is AES-CBC
- The Lucky 13 attack abuses MAC-then-encrypt to read encrypted messages
 - Guess a byte of plaintext and change the ciphertext accordingly
 - The MAC will error, but the time it takes to error is different depending on if the guess is correct
 - Attacker measures how long it takes to error in order to learn information about plaintext
 - TLS will send the message again if the MAC errors, so the attacker can guess repeatedly
- Takeaways
 - Side channel attack: The algorithm is proved secure, but poor implementation made it vulnerable
 - Always encrypt-then-MAC
- <https://medium.com/@c0D3M/lucky-13-attack-explained-dd9a9fd42fa6>

Lecture 20

MAC-then-Encrypt or Encrypt-then-MAC?

- Method 1: Encrypt-then-MAC
 - First compute $\text{Enc}(K_1, M)$
 - Then MAC the ciphertext: $\text{MAC}(K_2, \text{Enc}(K_1, M))$
- Method 2: MAC-then-encrypt
 - First compute $\text{MAC}(K_2, M)$
 - Then encrypt the message and the MAC together: $\text{Enc}(k_1, M \mid\mid \text{MAC}(K_2, M))$
- Which is better?
 - In theory, both are secure if applied properly
 - MAC-then-encrypt has a flaw: You don't know if tampering has occurred until after decrypting
 - Attacker can supply arbitrary tampered input, and you always have to decrypt it
 - Passing attacker-chosen input through the decryption function can cause side-channel leaks
- **Always use encrypt-then-MAC** because it's more robust to mistakes

TLS 1.0 “Lucky 13” Attack

- TLS: A protocol for sending encrypted and authenticated messages over the Internet
- TLS 1.0 uses MAC-then-encrypt: $\text{Enc}(k_1, M \parallel \text{MAC}(k_2, M))$
 - The encryption algorithm is AES-CBC
- The Lucky 13 attack abuses MAC-then-encrypt to read encrypted messages
 - Guess a byte of plaintext and change the ciphertext accordingly
 - The MAC will error, but the time it takes to error is different depending on if the guess is correct
 - Attacker measures how long it takes to error in order to learn information about plaintext
 - TLS will send the message again if the MAC errors, so the attacker can guess repeatedly
- Takeaways
 - Side channel attack: The algorithm is proved secure, but poor implementation made it vulnerable
 - Always encrypt-then-MAC
- <https://medium.com/@c0D3M/lucky-13-attack-explained-dd9a9fd42fa6>

Authenticated Encryption: Summary

- Authenticated encryption: A scheme that simultaneously guarantees confidentiality and integrity (and authenticity) on a message
- First approach: Combine schemes that provide confidentiality with schemes that provide integrity and authenticity
 - MAC-then-encrypt: $\text{Enc}(K_1, M \parallel \text{MAC}(K_2, M))$
 - Encrypt-then-MAC: $\text{MAC}(K_2, \text{Enc}(K_1, M))$
 - Always use Encrypt-then-MAC because it's more robust to mistakes

Digital Signature

Digital Signatures

- NIST FIPS PUB 186-4 - the result of a cryptographic transformation of data that, when properly implemented, provides a mechanism for verifying **origin authentication, data integrity, and signatory non-repudiation**
- Based on asymmetric keys

Digital Signatures

- Asymmetric cryptography is good because we don't need to share a secret key
- Digital signatures are the asymmetric way of providing integrity/authenticity to data
- Assume that Alice and Bob can communicate public keys without David interfering

Digital Signatures: Definition

- Three parts:
 - $\text{KeyGen}() \rightarrow PK, SK$: Generate a public/private keypair, where PK is the verify (public) key, and SK is the signing (secret) key
 - $\text{Sign}(SK, M) \rightarrow sig$: Sign the message M using the signing key SK to produce the signature sig
 - $\text{Verify}(PK, M, sig) \rightarrow \{0, 1\}$: Verify the signature sig on message M using the verify key PK and output 1 if valid and 0 if invalid
- Properties:
 - **Correctness**: Verification should be successful for a signature generated over any message
 - $\text{Verify}(PK, M, \text{Sign}(SK, M)) = 1$ for all $PK, SK \leftarrow \text{KeyGen}()$ and M
 - **Efficiency**: Signing/verifying should be fast
 - **Security**: Same as for MACs except that the attacker also receives PK
 - Namely, no attacker can forge a signature for a message

Lecture 21

RSA Signature

- KeyGen():
 - Randomly pick two large primes, p and q
 - Compute $n = pq$
 - n is usually between 2048 bits and 4096 bits long
 - Choose e
 - Requirement: e is relatively prime to $(p - 1)(q - 1)$
 - Requirement: $2 < e < (p - 1)(q - 1)$
 - Compute $d = e^{-1} \bmod (p - 1)(q - 1)$
 - **Public key:** n and e
 - **Private key:** d

RSA Digital Signature Algo

Step1: Generate a hash value, or message digest, $mHash$ from the message M to be signed

Step2: Pad $mHash$ with a constant value $padding1$ and pseudorandom value $salt$ to form M'

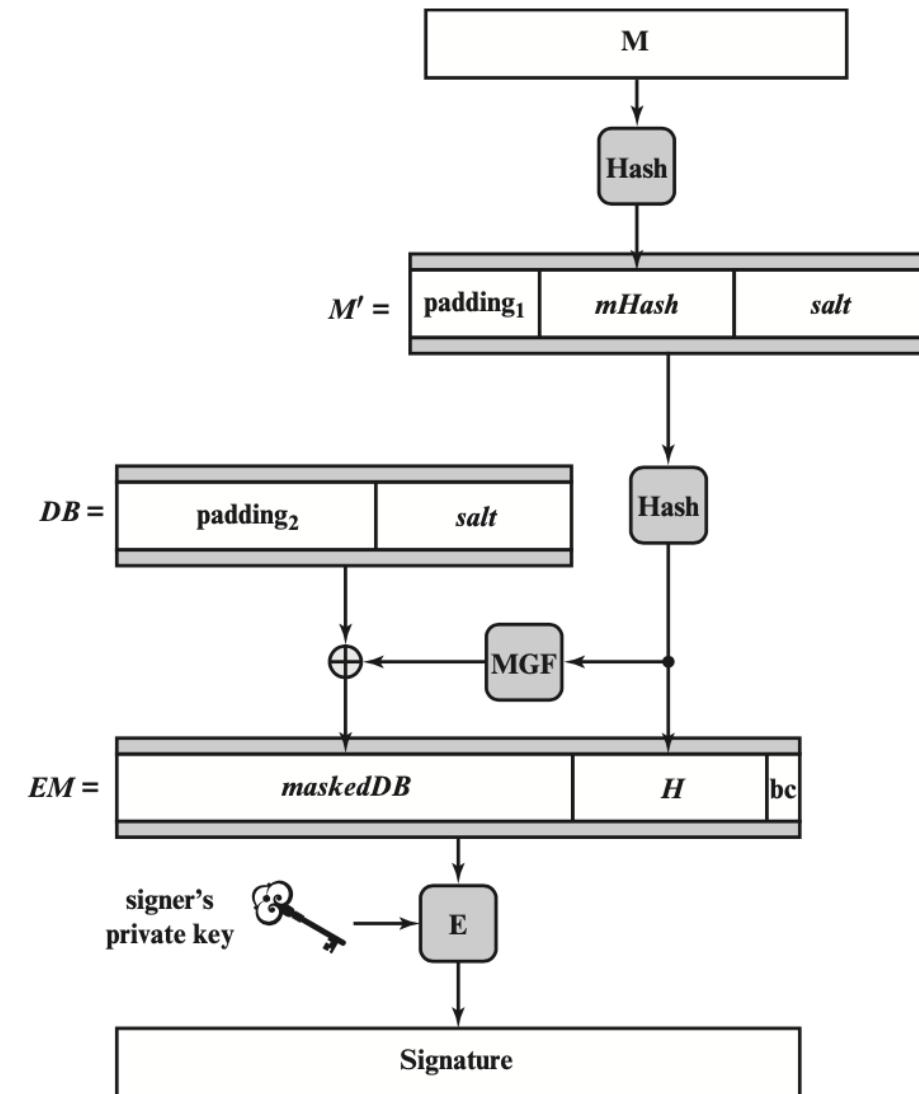
Step3: Generate hash value H from M'

Step4: Generate a block DB consisting of a constant value $padding2$ and salt

Step5: Use the mask generating function MGF, which produces a randomized out-put from input H of the same length as DB

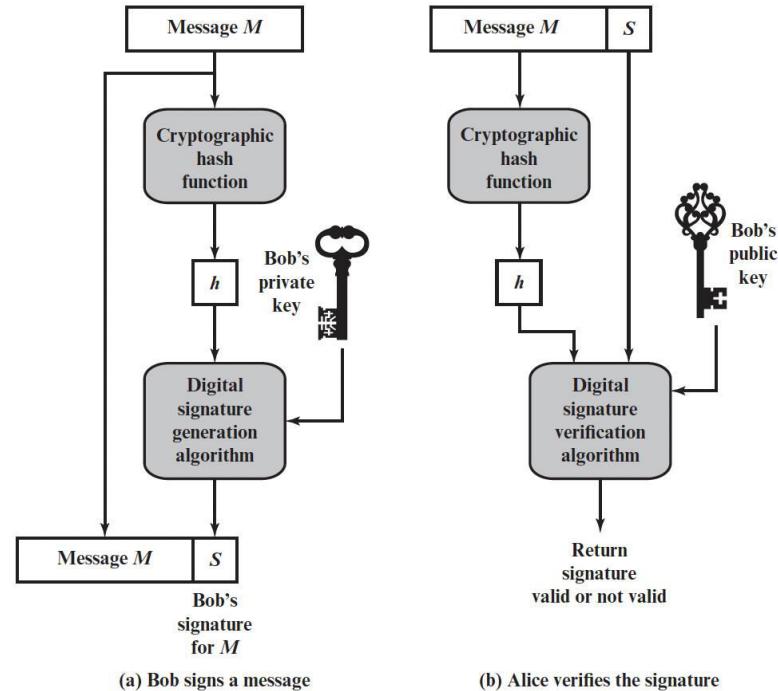
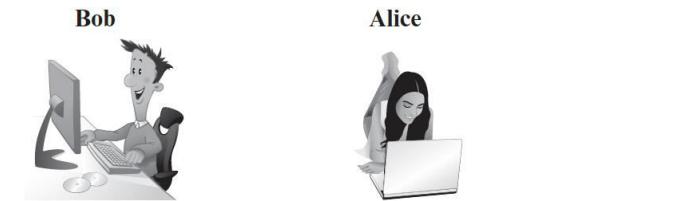
Step 6: Create the encoded message (EM) block by padding H with the hexadecimal constant bc and the XOR of DB and output of MGF

Step 7: Encrypt EM with RSA using the signer's private key



RSA Signatures

- $\text{Sign}(d, M)$:
 - Compute $H(M)^d \bmod n$
- $\text{Verify}(e, n, M, sig)$
 - Verify that $H(M) \equiv sig^e \bmod n$



RSA Signatures: Correctness

Theorem: $\text{sig}^e \equiv H(M) \pmod{N}$

Proof:

$$\text{sig}^e = [H(M)^d]^e \pmod{N} = H(M)^{ed} \pmod{N}$$

$$= H(M)^{k\phi(n)+1} \pmod{N}$$

$$= [H(M)^{\phi(n)}]^k \cdot H(M) \pmod{N}$$

$$= H(M) \pmod{N}$$

RSA Digital Signature: Security

- **Necessary hardness assumptions:**
 - **Factoring hardness assumption:** Given n large, it is hard to find primes $p, q = n$
 - **Discrete logarithm hardness assumption:** Given n large, hash , and $\text{hash}^d \bmod n$, it is hard to find d
- Salt also adds security
 - Even the same message and private key will get different signatures

Hybrid Encryption

- Issues with public-key encryption
 - Notice: We can only encrypt small messages because of the modulo operator
 - Notice: There is a lot of math, and computers are slow at math
 - Result: We don't use asymmetric for large messages
- **Hybrid encryption:** Encrypt data under a randomly generated key K using symmetric encryption, and encrypt K using asymmetric encryption
 - $\text{Enc}_{\text{Asym}}(\text{PK}, K); \text{Enc}_{\text{Sym}}(K, \text{large message})$
 - Benefit: Now we can encrypt large amounts of data quickly using symmetric encryption, and we still have the security of asymmetric encryption

Homework – no submission

- RQ: 3.1, 3.2, 3.3, 3.4, 3.6, 3.7
- Problems:
 - prove correctness of RSA digital signature
 - 3.14

Homework 2 - individual

- For Chapter 3
- Deadline: Oct. 26 (Thursday), 11:59 pm
- We will use the blackboard submission time as your final timestamp
- 10% penalty per day for late submission

Thank you!