Lecture slides of the course

Information hiding & secret sharing

COURSE INTRODUCTION

Phạm Trọng Nghĩa

ptnghia@fit.hcmus.edu.vn

Main topics in this course

- Steganography
- Digital watermarking
- Secret sharing

Main topics in this course

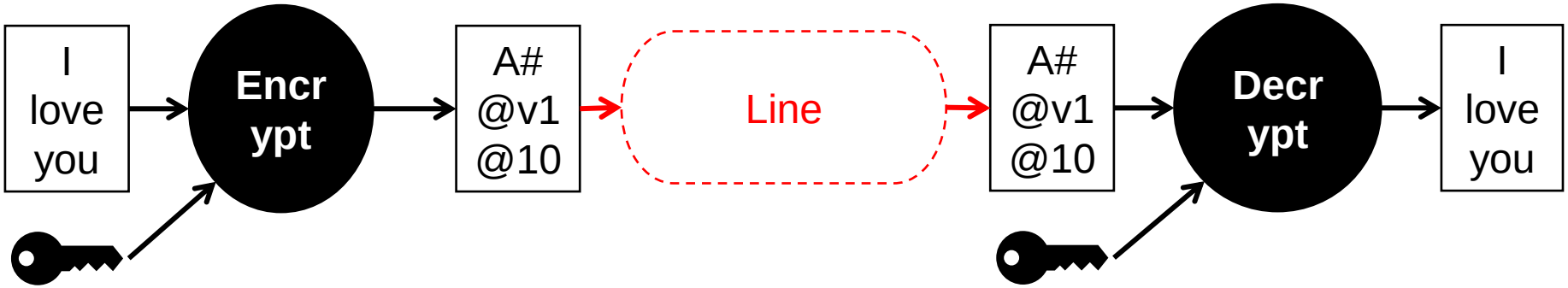
- **Steganography**
- Digital watermarking
- Secret sharing

The need for secure communication

Suppose **Bob** wants to send a message to **Alice** over the Internet. How is the information transmitted securely?

Solution 1: **encrypt** the information to be transmitted

The attacker knows the existence of the info, but it is difficult to know the content of the info



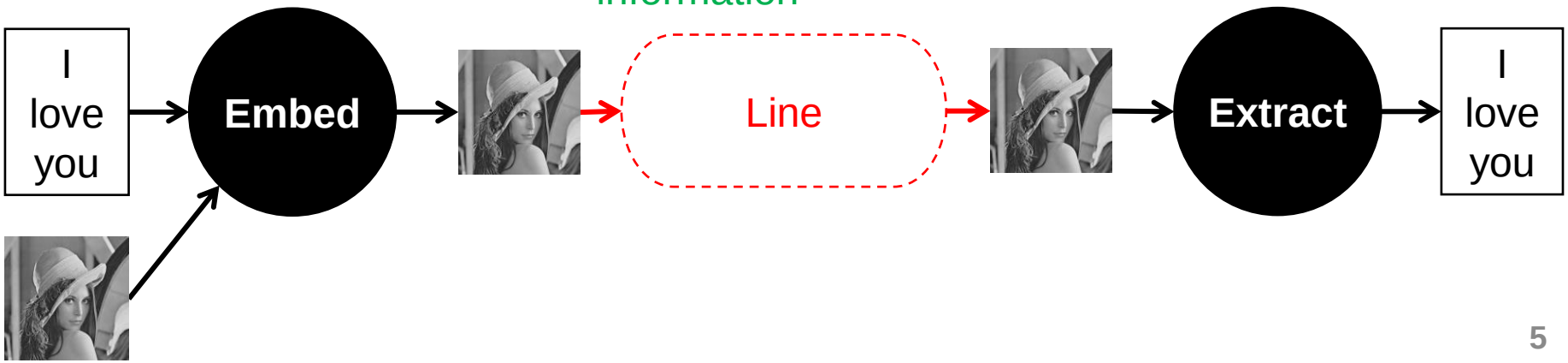
The need for secure communication

Suppose **Bob** wants to send a message to **Alice** over the Internet. How is the information transmitted securely?

Solution 2: **hide** the information to be transmitted in another message

This approach is called **steganography** (**covered writing**) and is considered a branch of **data/information hiding**.

The attacker is not even aware of the existence of the information



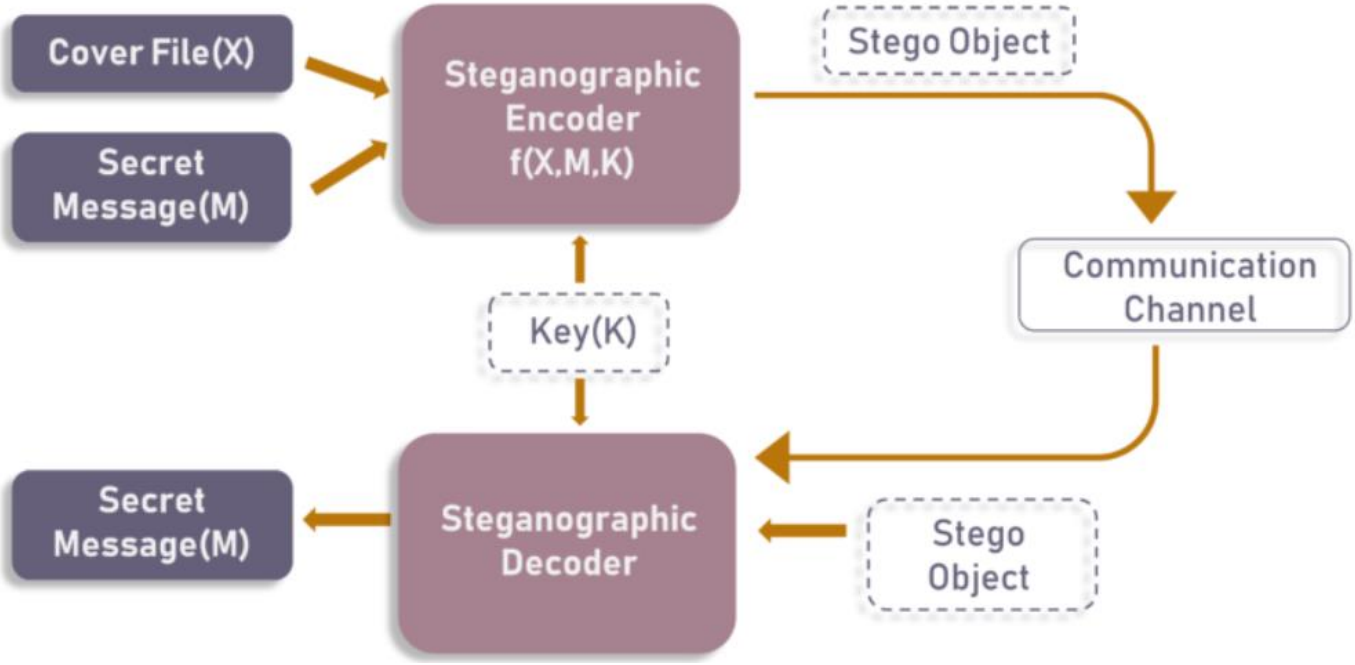
The need for secure communication

Suppose **Bob** wants to send a message to **Alice** over the Internet. How is the information transmitted securely?

Solution 3: **encrypt** the information to be transmitted and then **hide** it in another message

Steganography

- **Definition**: **Steganography** is the art and science of embedding secret messages in a cover message in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message



Steganography: Historical Background

- **Steganography** is the practice of concealing a secret message behind a normal message.
- It stems from two Greek words:
  - ***Steganos****:* means covered
  - ***Graphia****:* means writing.
- Steganography is an ancient practice, being practiced in various forms for thousands of years to keep communications private

Steganography: Historical Background

- The first use of steganography can be traced back to 440 BC when ancient Greece, people wrote messages on wood and covered it with wax, that acted as a covering medium
- Romans used various forms of Invisible Inks, to decipher those hidden messages light or heat were used
- During World War II the Germans introduced microdots, which were complete documents, pictures, and plans reduced in size to the size of a dot and were attached to normal paperwork
- Null Ciphers were also used to hide unencrypted secret messages in an innocent looking normal message

Steganography vs Cryptography

	STEGANOGRAPHY	CRYPTOGRAPHY
Definition	It is a technique to hide the existence of communication	It’s a technique to convert data into an incomprehensible form
Purpose	Keep communication secure	Provide data protection
Data Visibility	Never	Always
Data Structure	Doesn’t alter the overall structure of data	Alters the overall structure of data
Key	Optional, but offers more security if used	Necessary requirement
Failure	Once the presence of a secret message is discovered, anyone can use the secret data	If you possess the decryption key, then you can figure out original message from the ciphertext

Steganography Techniques

- Depending on the nature of the cover object (actual object in which secret data is embedded), steganography can be divided into five types:
  - **Text Steganography**
  - **Image Steganography**
  - Video Steganography
  - **Audio Steganography**
  - Network Steganography

Text Steganography

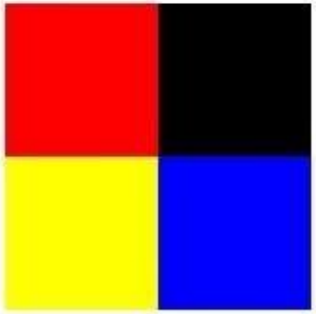
- Text Steganography is hiding information inside the text files.
  - Changing the format of existing text
  - Changing words within a text
  - Generating random character sequences
  - Using context-free grammars to generate readable texts.
- Various techniques used to hide the data in the text are:
  - Format Based Method
  - Random and Statistical Generation
  - Linguistic Method

Image Steganography

- Hiding the data by taking the cover object as the image is known as image steganography. In digital steganography, images are widely used cover source because there are a huge number of bits present in the digital representation of an image.
  - Least Significant Bit Insertion
  - Masking and Filtering
  - Redundant Pattern Encoding
  - Encrypt and Scatter
  - Coding and Cosine Transformation

Image Steganography - Example

Original Image

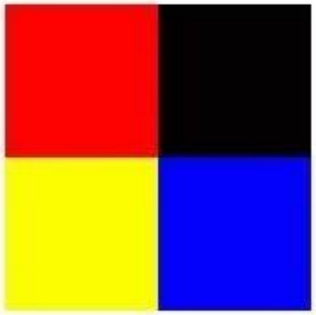


```
11111111  00000000
00000000  00000000
00000000  00000000

11111111  00000000
11111111  00000000
00000000  11111111
```

Least Significant Bit Steganography

Stego Image



```
11111101  00000011
00000010  00000001
00000000  00000010

11111100  00000011
11111101  00000001
00000001  11111100
```

c	a	t
01 10 00 11	01 10 00 01	01 11 01 00

Audio Steganography

- Hiding the data by taking the cover object as the image is known as image steganography. In digital steganography, images are widely used cover source because there are a huge number of bits present in the digital representation of an image.
  - Least Significant Bit Insertion
  - Masking and Filtering
  - Redundant Pattern Encoding
  - Encrypt and Scatter
  - Coding and Cosine Transformation

Steganography – Toy examples

- Message:

randoM capitalosis is a rarE disEase ofTen contrAcTed by careless inTernet users. tHis sad illnEss causes the aFfected peRsON To randomly capitalize letters in a bOdy oF texT. please do not confuse this disease witH a blatant attEmpt aT steganogRAPhy.

Reveals: MEET AT THE FRONT OF THE TRAP

Steganography – Toy examples

- A woman named Alice sends the following e-mail to her friend Bob

“My friend Bob,

Until yesterday I was using binoculars for stargazing. Today, I decided to try my new telescope. The galaxies in Leo and Ursa Major were unbelievable! Next, I plan to check out some nebulas and then prepare to take a few snapshots of the new comet. Although I am satisfied with the telescope, I think I need to purchase light pollution filters to block the xenon lights from a nearby highway to improve the quality of my pictures.

Cheers,

Alice”

Steganography – Toy examples

“My friend Bob,

Until yesterday I was using binoculars for stargazing. Today, I decided to try my new telescope. The galaxies in Leo and Ursa Major were unbelievable! Next, I plan to check out some nebulas and then prepare to take a few snapshots of the new comet. Although I am satisfied with the telescope, I think I need to purchase light pollution filters to block the xenon lights from a nearby highway to improve the quality of my pictures.

Cheers,

Alice”

Step 1: Listing the first letters of all words

mfbuyiwubfstidttmnttgilaumwuniptcosnatpttafsotncaiaswttitintplpftbt xlfanhtitqompca.

Steganography – Toy examples

Step 1: Listing the first letters of all words

mfbuyiwubfstidttmnttgilaumwuniptcosnatpttafsotncaiaswttitintplpftbt xlfanhtitqompca.

Step 2: π = 3.141592653589793

Reads the message from the extracted sequence of letters by putting down the third letter in the sequence, then the next first letter, the next fourth letter, etc. The resulting message is:

buubdlupnpsspx

Steganography – Toy examples

Step 1:

mfbuyiwubfstidttmnttgilaumwuniptcosnatpttafsotncaiaswttitintplpftbt

xlfanhtitqompca.

Step 2: **buubdlupnpsspx**

Step 3: replaces each letter with the letter that precedes it in the alphabet and deciphers the secret message

attack tomorrow.

Main topics in this course

- Steganography
- **Digital watermarking**
- Secret sharing

The need for copyright protection for digital works

Let's say you take a good photo and put it on the internet. Others will be able to easily copy your photo, perhaps even say it's theirs, and use it for their own benefit. How to prevent this problem?

One solution: **digital watermarking** (this is another branch of data hiding)

- Embed a watermark that is the owner's information in the work
- Requirement: robust - it must be difficult for an attacker to remove or edit the watermark

N

The need for copyright protection for digital works

Let's say you are an e-book publisher, you sell multiple copies of a book file to many customers. A customer can arbitrarily create copies, send them to other people;

In the end, the result was that the book file appeared on the internet and was freely available to anyone to download. How to prevent this problem?

One solution: **digital watermarking**

- When selling a copy to a customer, **embed** a code specific to that customer. If the book file is leaked online, the publisher can extract the code to know which customer the source of the leak is.
- Requirements: **robust** and **invisible**

The need for copyright protection for digital works

- In the digital world, a watermark is a **pattern of bits** inserted into a digital media that can **identify the creator** or authorized users.
- The digital watermark—unlike the traditional printed, visible watermark—is designed to be invisible to viewers.
- The bits embedded into an image are scattered all around to avoid identification or modification.
- Therefore, a digital watermark must be robust enough to survive the detection, compression, and other operations that might be applied upon a document.

Three criteria

- **Robustness:** the ability of the hidden message to **remain undamaged** even if the stego–media undergoes transformation, scaling and blurring, cropping and various other techniques.
- **Invisibility**: measured by the **similarity** of the stego image and its corresponding cover image
- **Capacity**: **how much data** can be hidden in the cover carrier

Steganography vs watermarking

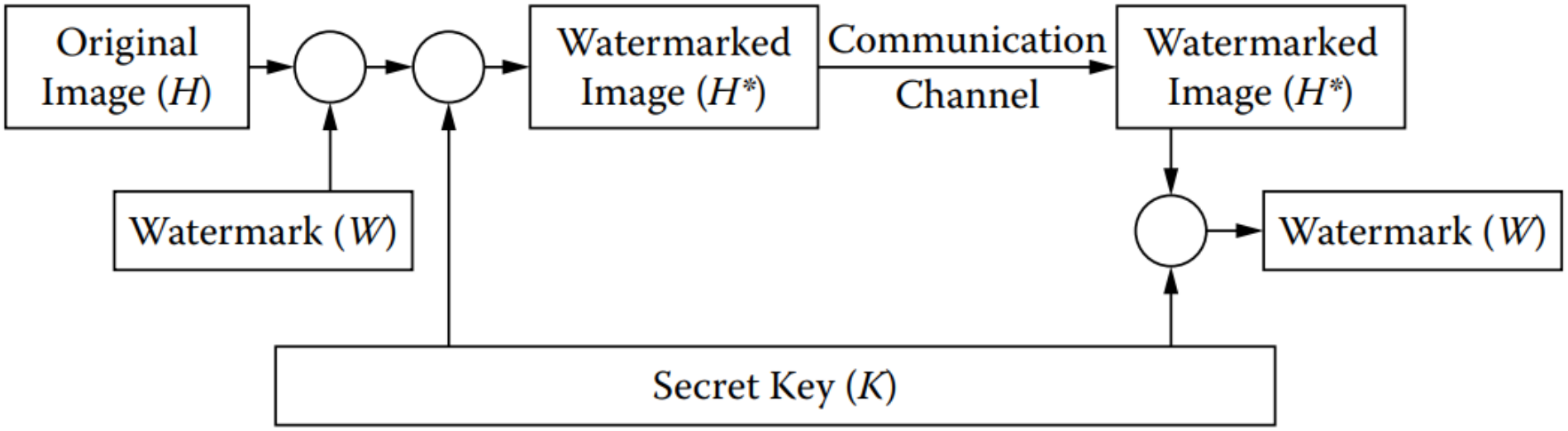
Among the 3 criteria: invisiblility - robustness – capacity

- Steganography often places heavy emphasis on **invisibility** and **capacity**
- Watermarking often places heavy emphasis on **robust** and **invisibility**

Steganography vs watermarking

- The watermarked image must look indistinguishable from the original image; if a watermarking system distorts the host image to some point of being perceptible, it is of no use.
- An ideal watermarking system should embed a **large amount of information** perfectly **securely**, but with **no visible degradation** to the host image.
- The embedded watermark should be robust with invariance to intentional (e.g., noise) or unintentional (e.g., image enhancement, cropping, resizing, or compression) attacks.
- Many researchers have been focusing on security and robustness

General digital watermarking system



Main topics in this course

- Steganography
- Digital watermarking
- **Secret sharing**

The need for secret sharing

Suppose you want to store a secret information (password, rocket launch code, ...). In addition to wanting confidential information not to be known by others, you also want confidential information to not be lost

Solution 1: store a single copy in one place (for example, on a computer)

- Secret information is difficult for others to know
- But it's easy to lose: for example, the computer breaks down

The need for secret sharing

Suppose you want to store a secret information (password, rocket launch code, ...). In addition to wanting confidential information not to be known by others, you also want confidential information to not be lost

Solution 2: store a copy in many place

- The more copies you save, the harder it will be for confidential information to be lost
- But it increases the risk of secret information being known by others (the attacker only needs to get the secret information in one storage location).

The need for secret sharing

Suppose you want to store a secret information (password, rocket launch code, ...). In addition to wanting confidential information not to be known by others, you also want confidential information to not be lost

Solution 3: **secret sharing** - divide secret information into ***n*** parts and save in ***n*** places, in order to recover the original information it is necessary to have at least any ***k*** part ($k \leq n$ and specified by the user, ect. $k = \frac{n}{2}$)

- Secret information is difficult to lose: only when there are more than k part lost, the original secret be lost
- Secret information is also difficult for others to know: an attacker must get at least part from k parts to be able to get secret information.

The need for secret sharing

Sharing secret information can also be used when there is a need to share a secret information with a group of people

- We wish that only a certain number of people would be able to recover secret information
- We also expect that if a certain number of people have a problem, it is still possible to recover secret information from the rest.

Course assesment

- **Labs: (30%)**
  - Programing language: Python
  - Working on IPython Notebook
- **Seminar (30%)**
  - Group work, 2-3 student each group
  - Research on data hiding topics and give oral presentation
- **Final exam (40%)**
  - Written and MCQ

Course regulations

- For any kind of cheating and plagiarism, students will be graded 0 for the course. The incident is then submitted to the school and university for further review.
- Students are encouraged to form study groups to discuss on the topics. However, individual work must be done and submitted on your own.
- Students absent for final exam are considered unqualified for course completion.
- Students must accumulate at least 10% of course credits for lab work.