

# Writing on Wet Paper

本身不是白纸  $\rightarrow$  打理  $\rightarrow$  写字  $\rightarrow$  干  $\rightarrow$  分不清新写的和原来的。

- Cover image (paper)  $x$ : has wet region.
- Only allowed to slightly modify the dry part.
- The received image (paper)  $y$  dries.
- Where the message is written?

# An Equivalent Well-Known Problem

大容量做ECC

In information theory: writing in memory with defective cells.

- Writer known the location of stuck cells.
- Reader do not know that.
- How to correctly read that?
- How to write as many bits as possible?

A special case of the Gel'fand-Pinsker channel.

# The Idea, Matrix Embedding Again!

Message  $\mathbf{m} \in \{0, 1\}^m$  in  $y \in \mathbb{Z}^n$  via a shared parity matrix  $\mathbf{D} \in \{0, 1\}^{m \times n}$ :

$$\mathbf{D}_{m \times n} \mathbf{y}_{n \times 1} = \mathbf{m}_{m \times 1}.$$

XOR is addition modulo 2.

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 16 \\ 69 \\ 35 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

# Wet Paper

In  $\mathbf{x}$ , we only change part of it.

- Dry part:  $\mathbf{x}[j], j \in \mathcal{J} \subset \{1, \dots, n\}$ . ex. 4.5, 5.5

- Can be changed.

- Wet part:  $\mathbf{x}[j], j \notin \mathcal{J}$ . ex. 6.1, 8.8

- Cannot be changed, i.e. fixed.

Thus the change  $\mathbf{v} = \mathbf{y} - \mathbf{x}$  has the property:

$$\mathbf{v}[j] = 0, j \notin \mathcal{J}.$$

# A Constrained Equation

Under the constraints:

$$\mathbf{v}[j] = 0, j \notin \mathcal{J}.$$

Solving the following equation.

$$\mathbf{D}\mathbf{y} = \mathbf{m}$$

$$\mathbf{D}(\mathbf{x} + \mathbf{v}) = \mathbf{m}$$

$$\mathbf{D}\mathbf{v} = \mathbf{m} - \mathbf{D}\mathbf{x}.$$

# Removing the Known Values

Using a permutation matrix  $\mathbf{P}$  to sort fixed  $\mathbf{v}[j]$  to the end.

$$\mathbf{P}\mathbf{v} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{|\mathcal{J}|} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{u} \\ 0 \end{pmatrix}$$

Handwritten red annotations:

- 干燥 (dry) with an arrow pointing to  $u_1$
- 干燥 (dry) with an arrow pointing to  $u_{|\mathcal{J}|}$
- 干燥 (dry) with an arrow pointing to the bottom zeros

# Continue

考试 不会解方程，写到 H 即停 or 简单数字.

$$\mathbf{D}\mathbf{v} = \mathbf{m} - \mathbf{D}\mathbf{x} = \mathbf{z}$$

$$(\mathbf{D}\mathbf{P}^{-1})(\mathbf{P}\mathbf{v}) = \mathbf{z}$$

$$(\mathbf{H} \quad \mathbf{K}) \begin{pmatrix} \mathbf{u} \\ 0 \end{pmatrix} = \mathbf{z}$$

$$\mathbf{H}_{m \times |\mathcal{J}|} \mathbf{u} = \mathbf{z}.$$

线性  
mass 了 无解  $\Leftrightarrow$  水印  
信息点, dry

Choosing the solution with the minimal number of changes.

# Acceleration

- Gaussian elimination:  $O(|\mathcal{J}|^3)$ .
- Matrix LT Process: much faster.



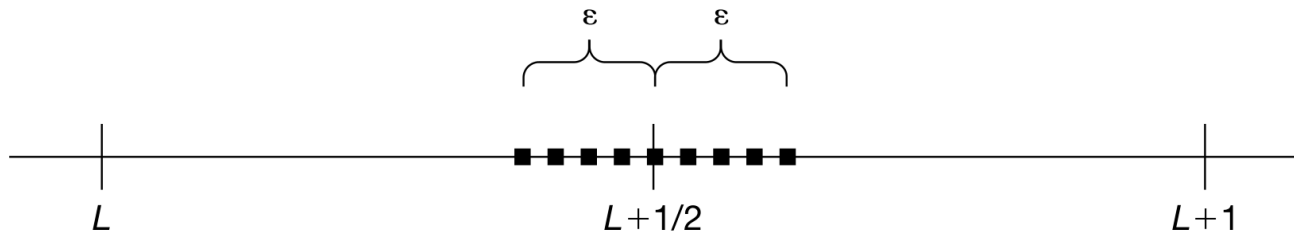
# Perturbed Quantization

带扰动的量化

One of the most secure steganographic schemes known today.

$\epsilon$  少, set 小 的  $m$  少  
根据  $m$  长度 设置  $\epsilon$ .

$$J = \{j | j \in \{1, \dots, n\},$$
$$\mathbf{u}[j] \in [L + 0.5 - \epsilon, L + 0.5 + \epsilon], L \in \mathbb{Z}\}.$$



# **Digital Watermarking and Steganography**

**by Ingemar Cox, Matthew Miller, Jeffrey Bloom, Jessica Fridrich, Ton Kalker**

## **Chapter 13. Steganalysis**

Lecturer: Jin HUANG

As to steganography, Shannon declared that

- “systems are primarily a psychological problem” and did not consider them further.

Similar in the counter part: Steganalysis.

# Detection

Two-class classification.

Error:

- False alarm.
- False detection.  $\Leftrightarrow$  false negative

# Basic Strategy

- Explicitly model the probability distribution.
  - likelihood ratio test.
- High dimension leads to big challenge.
- Low dimensional space: features.
  - Histogram.
  - Fourier transform of the intensity histogram.
  - ...

# Region Overlapping

Overlap of

- Cover work region.
- Stego Work region.

Boundary separation: hyperplane (SVM).

# Targeted Steganalysis

找技术专门针对某种算法

Know the steganographic algorithm.

For example, LSB, a naive method:

- Increase the noise.
- The sum of the absolute values of the differences is large.
  - Between all pairs of neighboring pixels.
- Not reliable: texture ...

Histogram! Even estimate the length of the message.

# Blind Steganalysis

Steganographic algorithm is unknown.

- Find a feature is sensitive to **all** methods.
  - We do not know.
  - Stimulate a new class of steganographic algorithm.



# Training

Data set:

并非完全 Blind. 数据集中有已知技术

- Cover works and Stego works from many known steganalysis algorithms.
  - If one of them is used, Bingo!
  - It may work for similar/related unseen ones.
- Cover works only: one-class learning.
  - What is the normal ones looks like.
  - Do not need to be retrained when new embedding methods appear.

# System Attacks

An example:

- 2006, a much more advanced encryption algorithm in DVD.
- Within months, the system had been broken.

# System Attacks

An example:

- 2006, a much more advanced encryption algorithm in DVD.
- Within months, the system had been broken.
- ... read the secret key from memory.

# A Stupid Case

In early implementation of F5

- JPEG header: “JPEG Encoder Copyright 1998, James R. Weeks and BioElectroMech”.

# Stego Key

和 work 大小有关

E.g. using dictionary attack.

- Follow the possible random walk path
  - Check the histogram (skip the correction).
  - Even read the message!
  - Or fake a message.

# Forensic Steganalysis

伪造隐写分析.

The warden can reliably determine the communication.

Block may not be good:

偷听

- Will alert Alice and Bob.
- No enough resource (political or economical).

So

- Try to read the message etc.

# Attack

- Stego work only attack.
- Known cover attack.
- Known the algorithm attack.
  - Dictionary attack.

# Cover Work Choosing

有差异  
原图上传 → 压缩 → 解压 → 原图.

Spatial domain LSB on a decompressed JPEG.

- Spatial vs DCT: many to one. 原图 (多) → 压缩后图.
- Re-compress it into JPEG.
- Decompress it back: the original cover work.
- Shorter message is more easy to detected!

图要大, 但不能过大, 否则



# Large Cover Work is not Good

For a fixed relative message length

- Features on smaller image is usually more noise.
  - Smaller image has less correlation to analysis.
- Noisy cover is better
  - Small noise over large one.