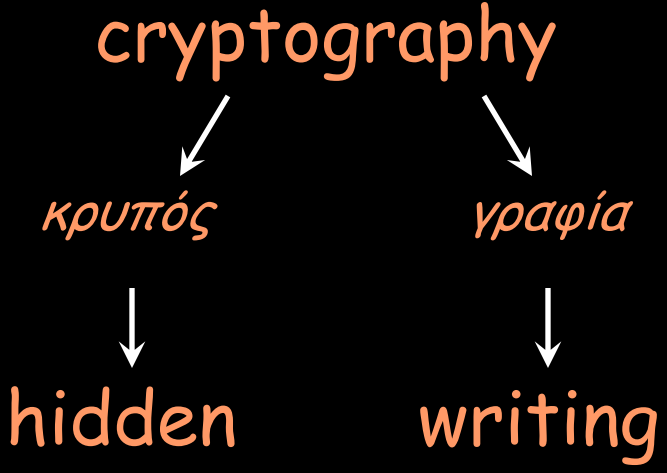
Distributed Systems

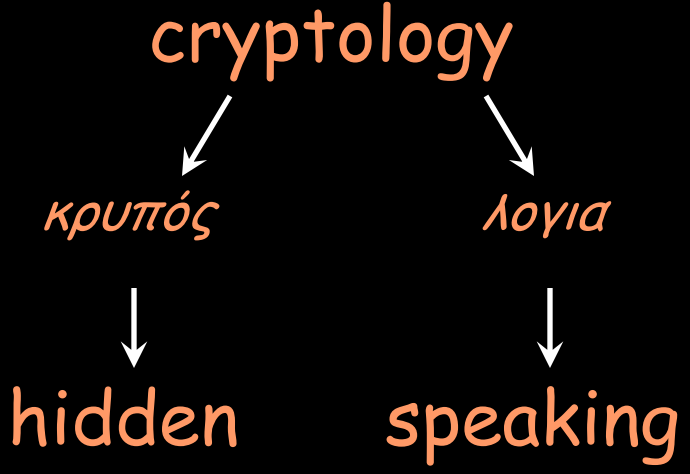
Steganography

Paul Krzyzanowski
pxk@cs.rutgers.edu



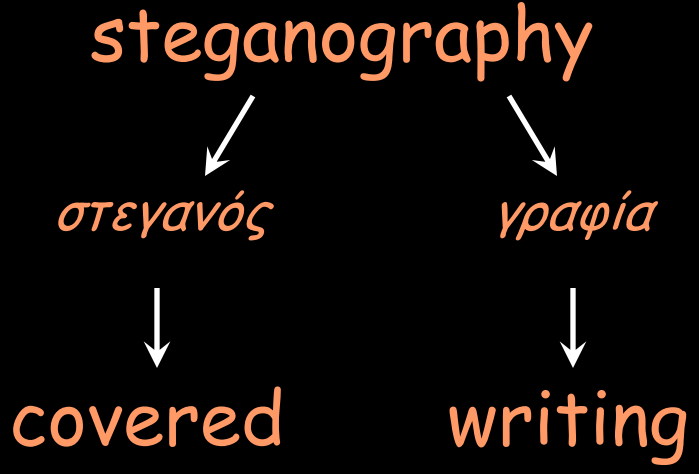
A secret manner of writing, … Generally, the art of writing or solving ciphers.

— *Oxford English Dictionary*



1967 D. Kahn, *Codebreakers* p. xvi, Cryptology is the science that embraces cryptography and cryptanalysis, but the term ‘cryptology’ sometimes loosely designates the entire dual field of both rendering signals secure and extracting information from them.

— *Oxford English Dictionary*



The art of secret (hidden) writing

Steganography

Art and science of communicating in a way that hides the existence of a message

signal or pattern imposed on content

- persistent under transmission
- not encryption
  - original image/file is intact
- not fingerprinting
  - fingerprinting leaves separate file describing contents

Classic techniques

- Invisible ink (1st century AD - WW II)
- Tatoo message on head
- Overwrite select characters in printed type in pencil
  - look for the gloss
- Pin punctures in type
- Microdots (WW II)
- Newspaper clippings, knitting instructions, XOXO signatures, report cards, …

Motivation

- Steganography received little attention in computing
- Renewed interest because of industry desire to protect copyrighted digital work
  - audio
  - images
  - video
  - Text
- Detect counterfeiter, unauthorized presentation, embed key, embed author ID
- Steganography ≠ Copy protection

Null Cipher

- Hide message among irrelevant data
- Confuse the cryptoanalyst

Null Cipher

- Hide message among irrelevant data
- Confuse the cryptoanalyst

Big rumble in New Guinea.
The war on
celebrity acts should end soon.
Over four
big ecstatic elephants replicated.

Null Cipher

- Hide message among irrelevant data
- Confuse the cryptoanalyst

Big rumble in New Guinea.
The war on
celebrity acts should end soon.
Over four
big ecstatic elephants replicated.

Bring two cases of beer.

BBC News 27 April 2006

Judge creates own Da Vinci code

The judge who presided over the failed Da Vinci Code plagiarism case at London's High Court hid his own secret code in his written judgement.

Seemingly random italicised letters were included in the 71-page judgement given by Mr Justice Peter Smith, which apparently spell out a message.

Mr Justice Smith said he would confirm the code if someone broke it.

"I can't discuss the judgement, but I don't see why a judgement should not be a matter of fun," he said.

Italicised letters in the first few pages spell out "Smithy Code", while the following pages also contain marked out letters.

http://news.bbc.co.uk/go/pr/fr/-/1/hi/entertainment/4949488.stm

Chaffing & Winnowing

- Separate good messages from the bad ones
- Stream of unencoded messages with signatures
  - Some signatures are bogus
  - Need key to test

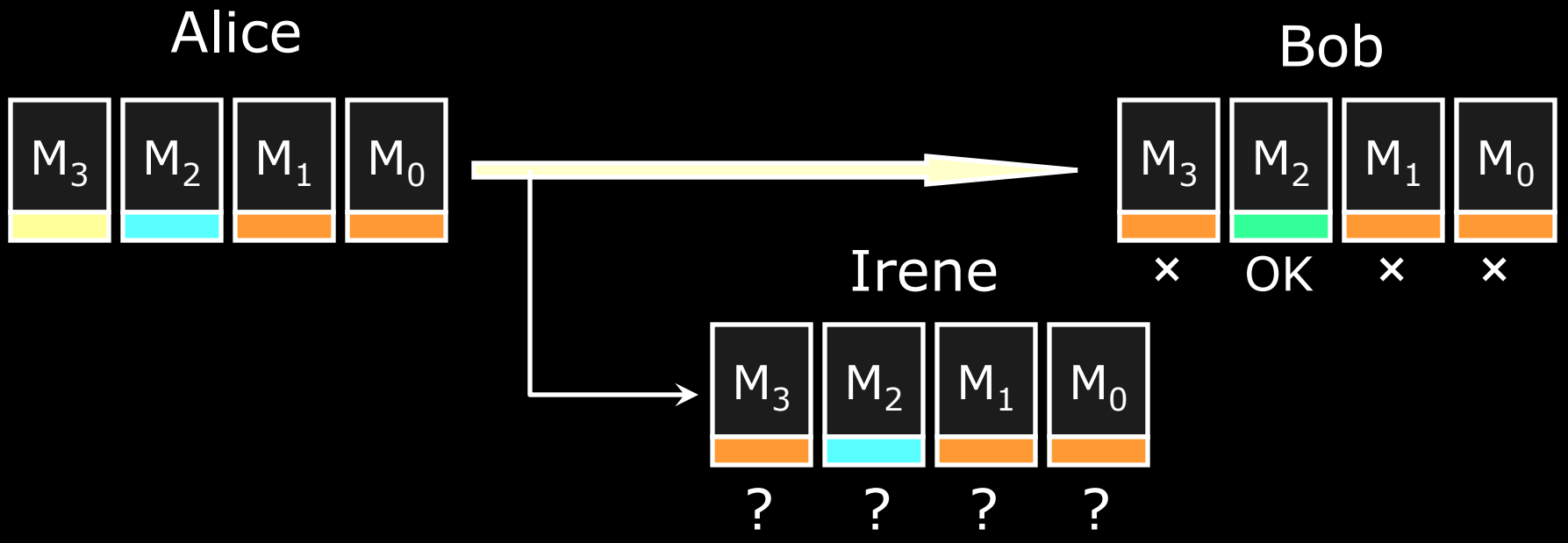


Image watermarking

- Spatial domain watermarking
  - bit flipping
  - color separation
- Frequency domain watermarking
  - embed signal in select frequency bands (e.g. high frequency areas)
  - apply FFT/DCT transform first

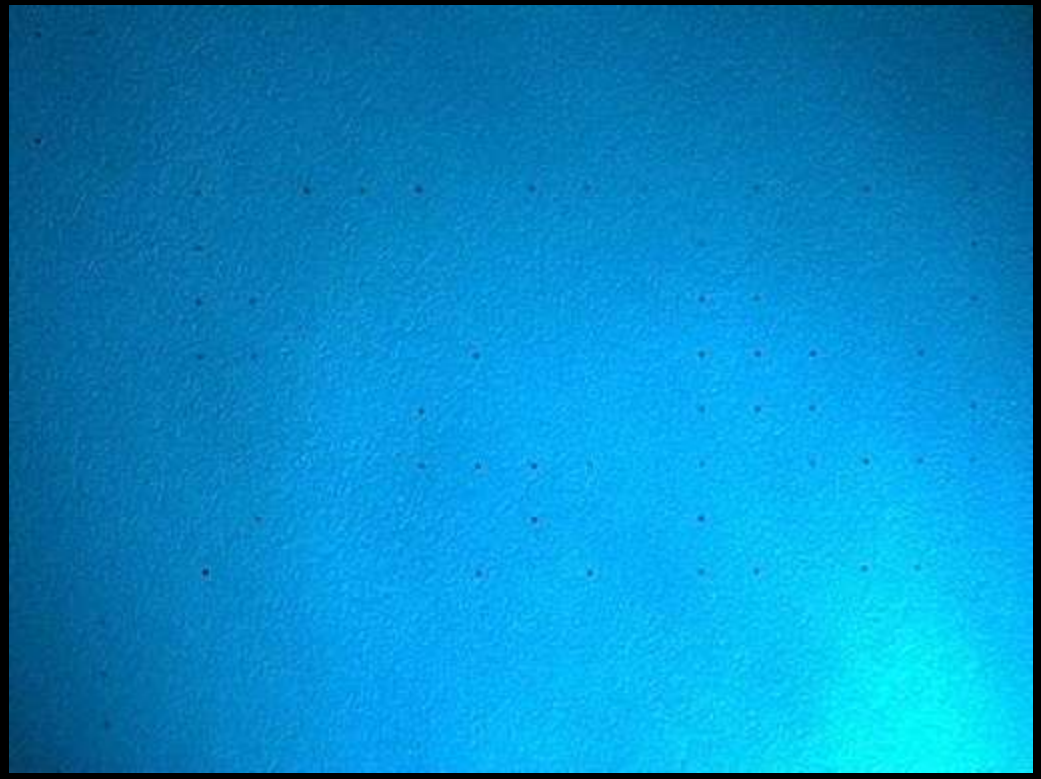
  - e.g. Digimarc
  - watermark should alter the least perceptible bits
    - these are the same bits targeted by lossy image compression software

UV Watermarking

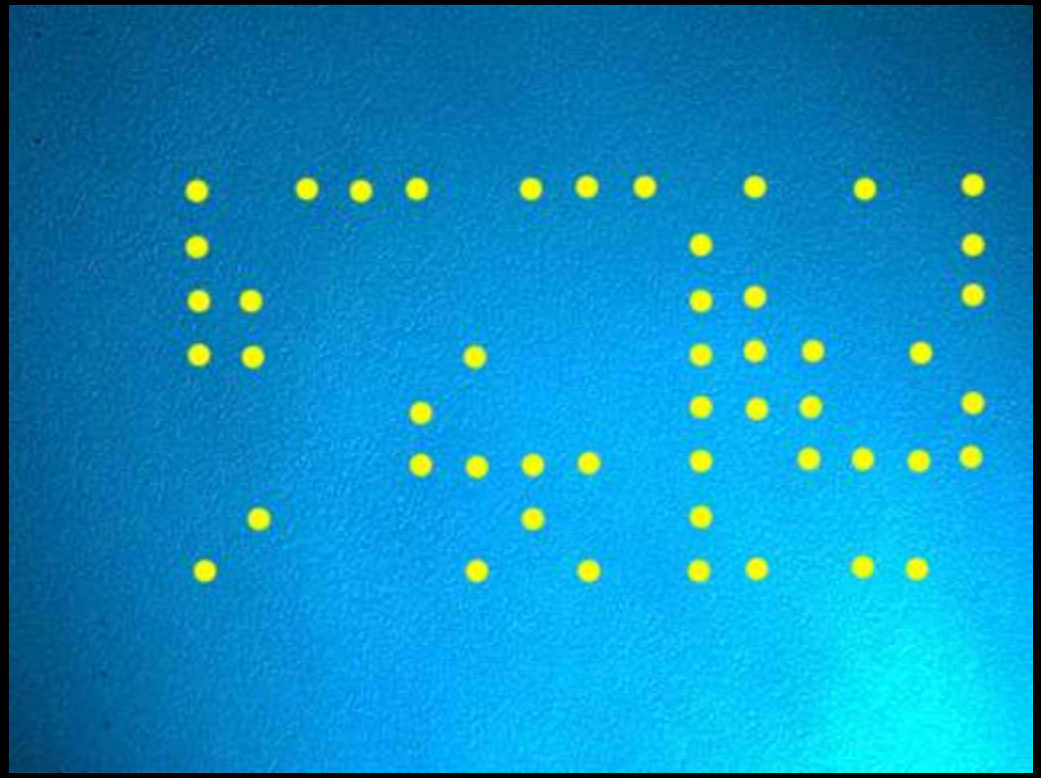
Machine ID codes in laser printers

See http://www.eff.org/Privacy/printers/

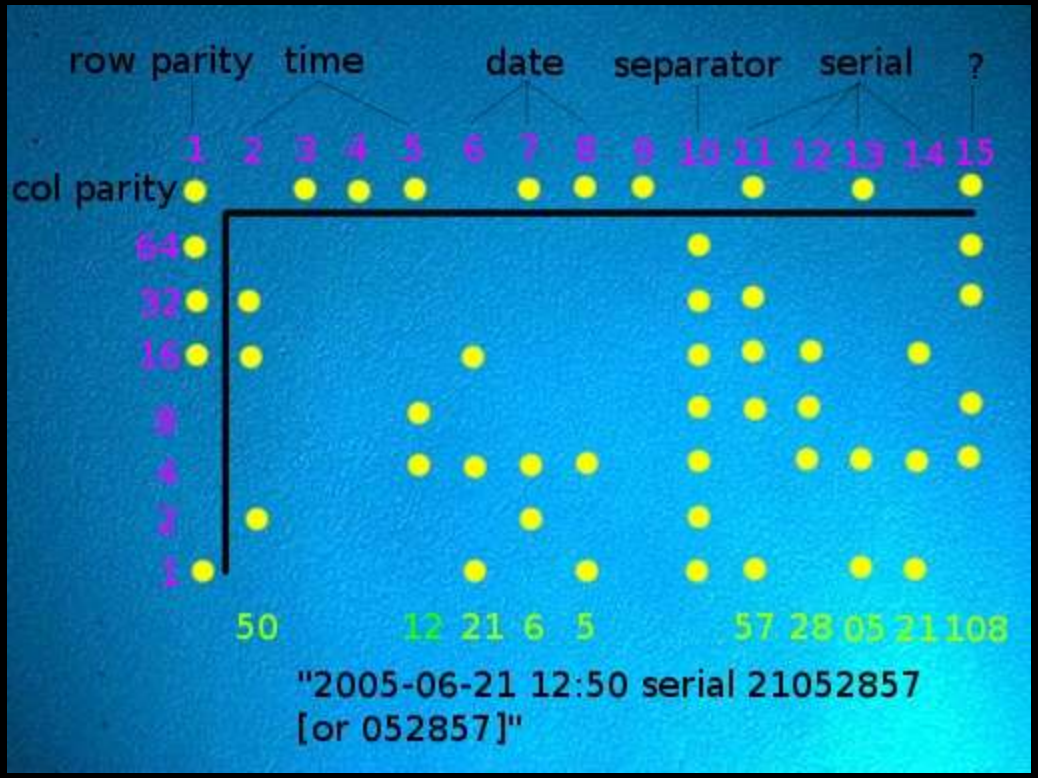
Machine ID codes in laser printers



Machine ID codes in laser printers



Machine ID codes in laser printers

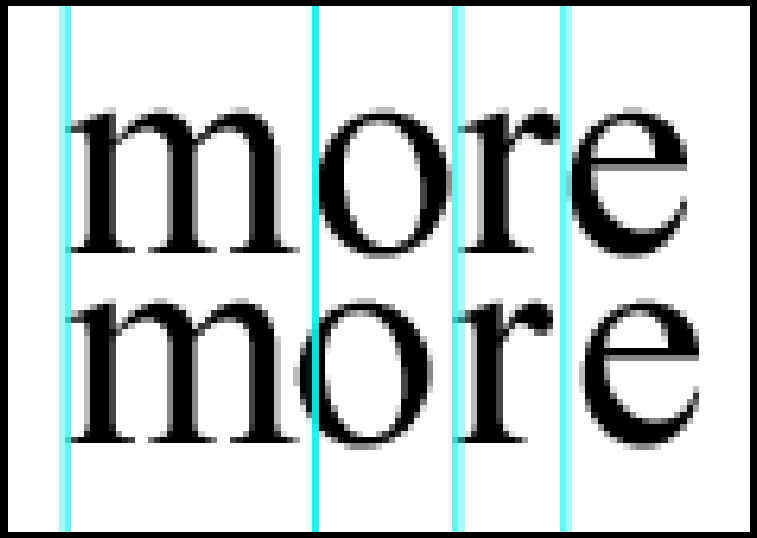


Text

- Text lines shifted up/down (40 lines text $\Rightarrow$ 2^{40} codes)
- word space coding
- character encoding - minor changes to shapes of characters

Text

- Text lines shifted up/down (40 lines text $\Rightarrow$ 2^{40} codes)
- word space coding
- character encoding - minor changes to shapes of characters



- works only on "images" of text e.g., PDF, postscript

Audio

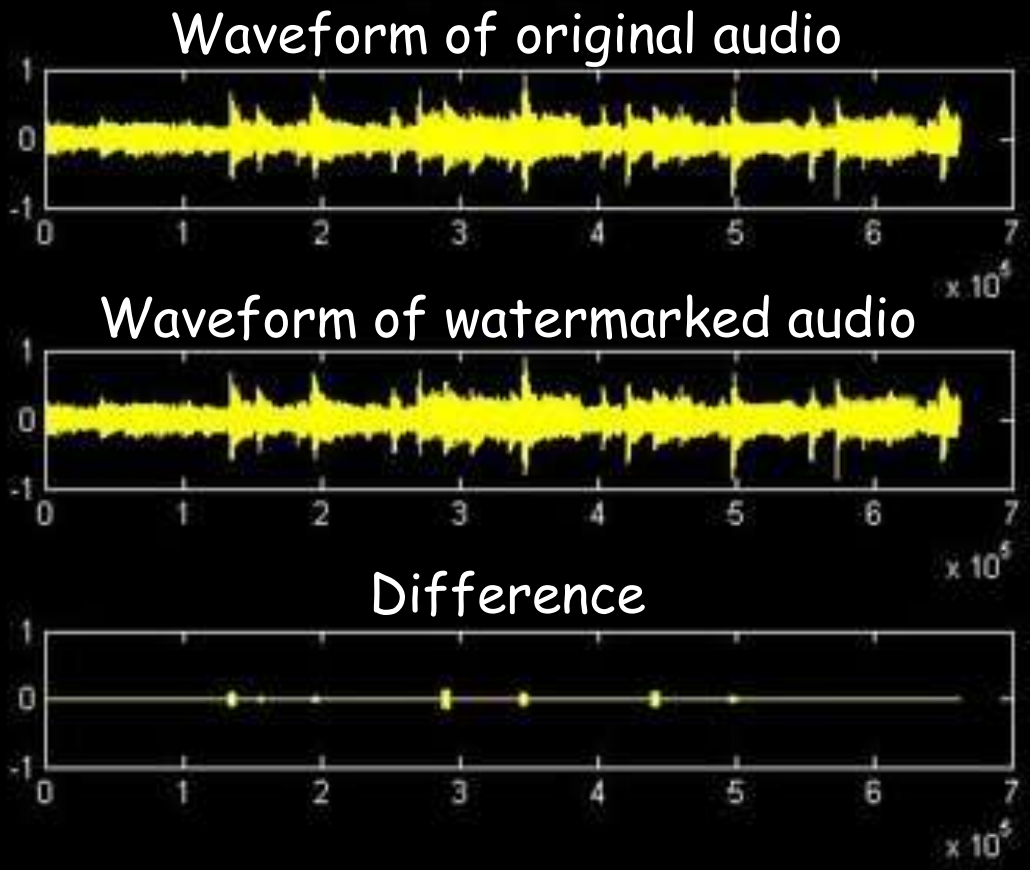
Perceptual coding

- inject signal into areas that will not be detected by humans
- may be obliterated by compression

Hardware with copy-protection

- not true watermarking - metadata present on media
- DAT
- minidisc
- presence of copy protection mechanisms often failed to give the media wide-spread acceptance

Amazon MP3 Audio



Video

- Coding still frames - spatial or frequency
- data encoded during refresh
  - closed captioning
- visible watermarking
  - used by most networks (logo at bottom-right)

Watermarking vs. Steganography

Goal of steganography

- Intruder cannot detect a message
- Primarily 1:1 communication

Goal of watermarking

- Intruder cannot remove or replace the message
- Primarily 1:many communication

The end.