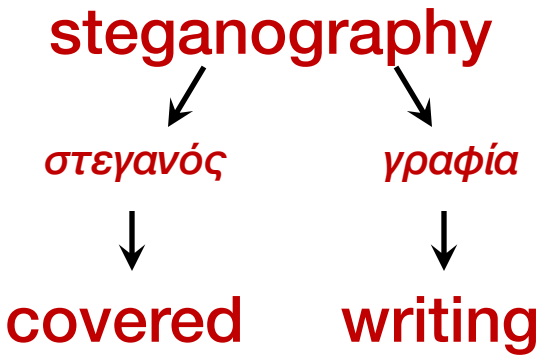
CS 419: Computer Security

Week 9: Steganography

Paul Krzyzanowski

November 3, 2020



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)
- Tattooed message on head
- Overwrite select characters in printed type in pencil
- Pin punctures in type
- Microdots (early 20th century)
- Newspaper clippings, knitting instructions, XOXO signatures, report cards, …

Motivation

- **Steganography received little attention in computing**
- **Renewed interest because of industry's desire to protect copyrighted digital work**
  - Audio, images, video, documents
- **Detect counterfeiter, unauthorized presentation, embed key, embed author ID**
- **Also useful for forensics: enemies may use steganography to conceal their messages**
  - Communication, stolen data, botnet controls

Steganography ≠ Copy protection

INDEPENDENT

Isis and al-Qaeda sending coded messages through eBay, pornography and Reddit

Kashmira Gander – Monday 2 March 2015 19:29 GMT

Isis and al-Qaeda members are communicating with each other via coded messages hidden on websites including eBay, Reddit, and inside pornographic photos, according to a new book.

Gordon Thomas, who has sources inside Israel's Mossad spy agency, has revealed that the organisation's cyber warfare department's most skilled cryptologists mastered a technique known as steganography, which is used to to conceal secret information within a digital file. The spies found that al-Qaeda had used the technique to hide messages in goods offered for sale on eBay, according to extracts from *Gideon's Spies: The Secret History of the Mossad* published by *The New York Post*.

Null Cipher

Hide message among irrelevant data

Confuse the cryptanalyst

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 cryptanalyst

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.

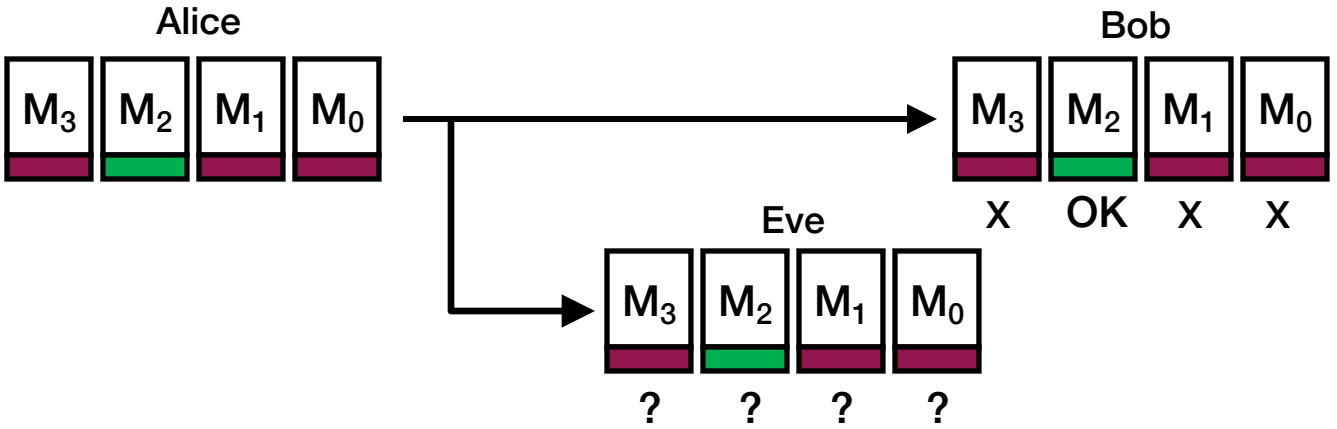
"I can't discuss the judgement, but I don't see why a judgement should not be a matter of fun," he said.

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**
  - Easy for someone who has the key, difficult for someone who does not
- **Stream of un-encoded messages with signatures or MACs**
  - Some signatures are bogus
  - Need to have the key to test



Steganography in images

Spatial domain

- Bit flipping
- Color separation

Frequency domain

- Apply FFT/DCT transform first
- Embed signal in select frequency bands
- Alter the least perceptible bits to avoid detection
  - But watch out: these are the same bits targeted by lossy image compression software (such as jpeg)

Just the picture

With the Declaration of Independence embedded

Differences

There are differences – but you don't notice them in the photo

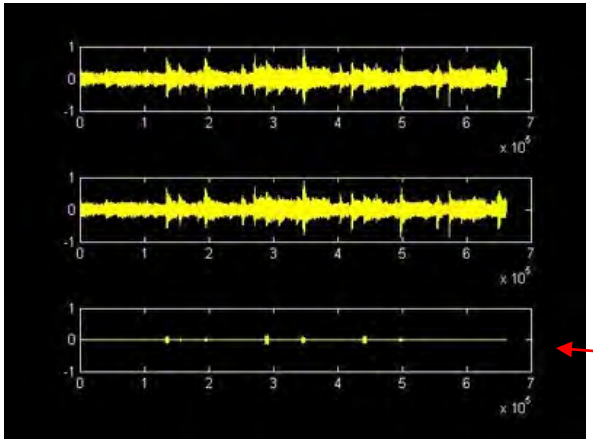
Video

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

Audio

Perceptual coding

- Inject signal into areas that will not be detected by humans
- May be obliterated by compression

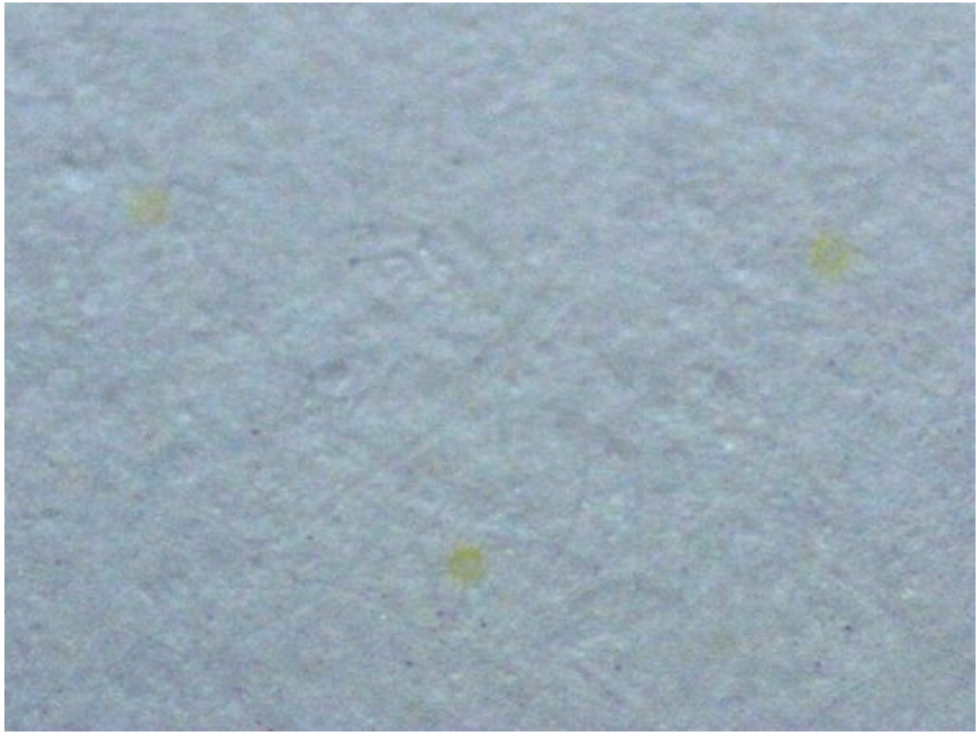


Amazon MP3 audio

Identifies where the song was purchased, not the user

Difference

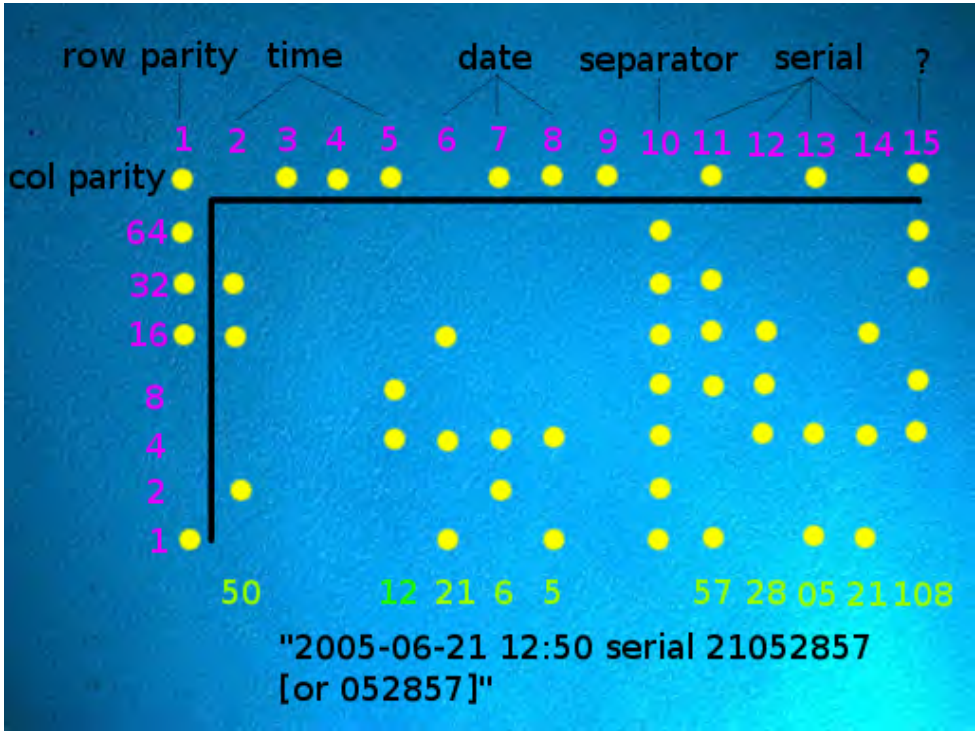
Machine ID codes in laser printers



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

Machine ID codes in laser printers

Machine ID codes in laser printers



Designed by Xerox to identify counterfeit currency and help track down counterfeiters

UV Watermarking

Also passports, hand stamps for amusement park re-entry,

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

Text-based stegranography

"Apparently, during the 1980's, British Prime Minister Margaret Thatcher became so irritated at press leaks of cabinet documents that she had the word processors programmed to encode their identity in the word spacing of documents, so that disloyal ministers could be traced."

– *Ross Anderson*
Stretching the Limits of Steganography

Watermarking vs. Steganography

Both techniques embed a message in data

Goal of steganography

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

Goal of watermarking

- Intruder cannot remove or replace the message (robustness is important)
- Doesn't have to be invisible
- Primarily 1:many communication

Watermarking applications

- **Copyright protection**
  - Embed information about owner
- **Copy protection**
  - Embed rights management information
  - But you need a trusted player
- **Content authentication**
  - Detect changes to the content

The End