

CSE 127

Computer Security

Spring 2007

Covert channels & Watermarking
Stefan Savage

Covert channels

- Similar problem, but here its not by accident
- The adversary is trying to “create” a hidden side channel by **embedding secrets** in **normal communications**
- Thought experiment
 - ◆ You are Martha Stewart’s stockbroker and you want to send her a message:
 - » 0 = “Sell Google now!”
 - » 1 = “Buy Google now!”
 - ◆ But the SEC is watching you... how might you do it?

Kinds of covert channels

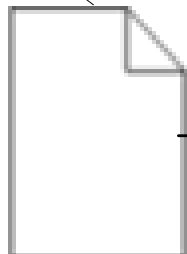
- Steganography
- Protocol encoding
- Timing encoding
- General issues:
 - ◆ How hard to discover embedded message?
 - ◆ How much “bandwidth” in the embedded channel?

Steganography: Overview

- Information Hiding
 - ◆ Intent to promote confidentiality and integrity of information by hiding it
- General Process
 - ◆ Hide sensitive information
 - » in cover medium (frequently images, or sometimes audio)
 - ◆ Combine with stego-system
 - ◆ Result is stego-medium

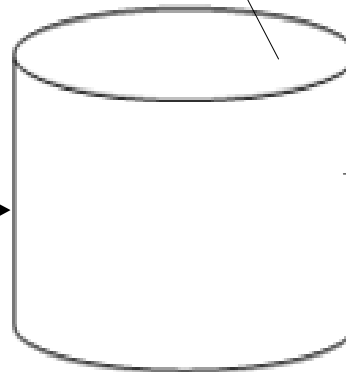
General Framework

What you want
to conceal



Information

The operation or
method use to
conceal data in
the cover medium



Steganographic
System

Cover medium with
concealed data



Stego-Medium



Cover Medium

Cryptography vs Steganography

- Steganography != cryptography
- Cryptography
 - ◆ Make message unreadable
 - » Anyone can look at it, but they can't decipher it
- Steganography
 - ◆ Make message hidden
 - » Hidden in plain sight, no one knows it exists
- Working Together
 - ◆ Encrypt message and hide it
 - » using crypto and stego respectively

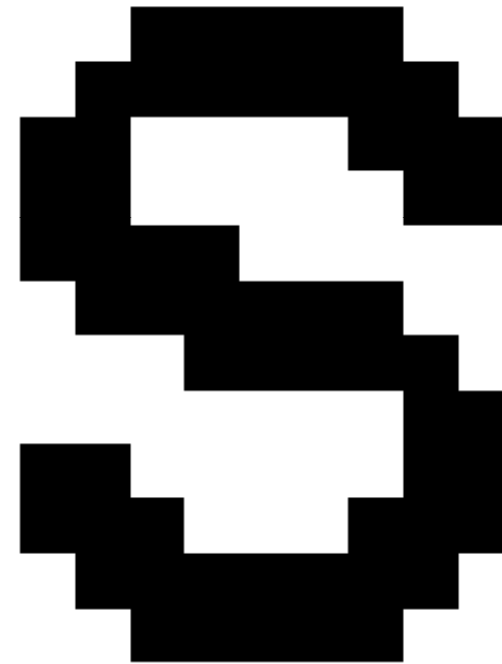
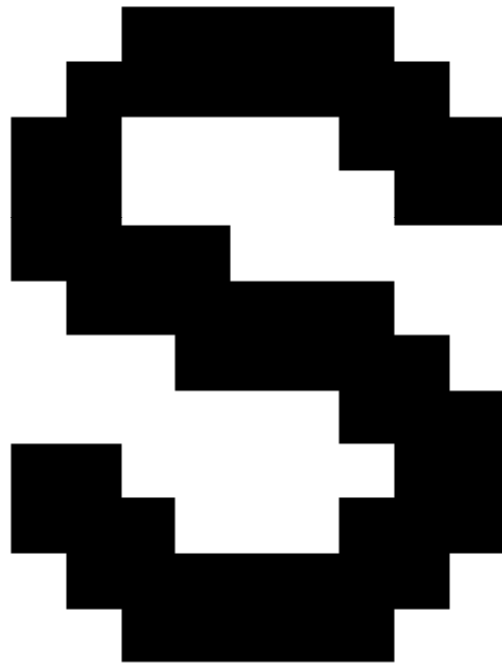
Steganographic Techniques in Digital Images

- Stego with lossless compression images
 - ◆ Binary Images
 - ◆ Least Significant Bit Insertion
 - ◆ Kurak-McHugh Method
- Lossy compression image stego-systems

Binary Images

- Two color images
 - ◆ only two color options (black = 1 & white = 0)
- Process
 - ◆ Divide the image into a set of $n \times m$ matrices
 - » i.e. if image is 10 by 20 pixels, then make ten 4 by 5 matrices
 - ◆ Each matrix represents a bit
 - » The sum of the indices mod 2
 - » Add/remove pixels in each $n \times m$ matrix as necessary to form secret message
- Example, digitally transmitted fax
- Highly prone to transmission errors

Example (2 15x15 matrices)



- Image on left: 61/225 black pixels = 1 bit
- Image on right: 60/225 black pixels = 0 bit

Red Green Blue (RGB) Model

- 24-bit color representation
 - ◆ 16,777,216 unique colors
- Uses a 3-tuple
 - ◆ Components are 8 bits
 - ◆ Represent amount of red, green, & blue
- Useful since to the human eye true red (255, 0, 0) and 1-bit away from true red (254, 0, 0) look virtually identical.

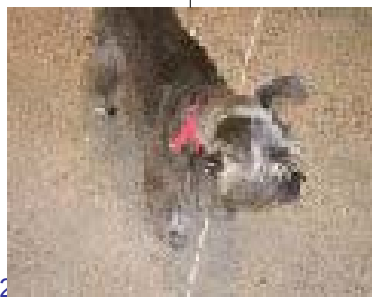
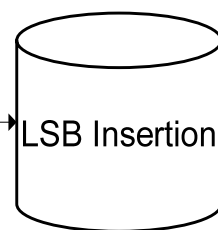
Least Significant Bit Insertion

- Example formats,
 - ◆ Graphic Interchange Format (GIF)
 - ◆ Bitmap (BMP)
- Each pixel can store 3 bits
 - ◆ A pixel is RGB (8 bits, 8 bits, 8 bits)
 - ◆ Hidden messages are fraction of the size of cover image, restricted by number of LSBs in image
- Process
 - ◆ Modify last bit of each component (the least significant bit) to desired bit
 - ◆ Result is nearly identical image to original, human eye cannot detect

Example

- Given 3 pixels
 $(113, 040, 210)_{10} \Rightarrow (1110001, 0101000, 11010010)_2$
 $(113, 041, 209)_{10} \Rightarrow (1110001, 0101001, 11010001)_2$
 $(113, 040, 205)_{10} \Rightarrow (1110001, 0101000, 11001101)_2$
- Insert the letter A by modifying LSB
 - ♦ ASCII representation of A is
 - » Decimal: 97
 - » Binary: 01100001
- Result
 $(1110000\mathbf{0}, 0101001\mathbf{1}, 1101001\mathbf{1})_2 \Rightarrow (112, 041, 211)_{10}$
 $(1110000\mathbf{0}, 0101000\mathbf{0}, 1101000\mathbf{0})_2 \Rightarrow (112, 040, 208)_{10}$
 $(1110000\mathbf{0}, 0101001\mathbf{1}, 11001101)_2 \Rightarrow (112, 041, 205)_{10}$

Example



May 5, 2012

Color Difference



Kurak-McHugh Method

- Adaptation of LSB insertion
 - ◆ Hide images of same size
- Process
 - ◆ Store most significant bits of hidden image in least significant bits of cover image
- Loss of Quality
 - ◆ Stego-image not as good as cover image
 - » sometimes you can tell they are different
 - ◆ Extracted image is not as good

Example

- Idea
 - ♦ 11100**010** (cover, LSB in bold)
 - ♦ **101**01011 (hidden, MSB in bold)
 - ♦ 11100**101** (stego)
- Loss of quality

Example



Figure 3: Landsat image of RDU



Figure 5: An F-16 aircraft



Figure 4: RDU contaminated with F-16



Figure 6: F-16 extracted from RDU

Example



■ The stego-image (i.e., after the hiding process):



Ma



■ The image extracted from the stego-image:



Stego for compressed images

- JPEG
 - ◆ Different color space
 - » YCbCr, not RGB
 - ◆ Discrete cosine transformation
 - ◆ Hide info in LSB of DCT coefficients matrix
 - ◆ Only changes low frequency components of image... human visual system largely insensitive

Surviving image transformation

- Rotation
- Translation
- Scaling
- Cropping
- Compressing
- Noise addition
- Etc...

Beyond Digital Images

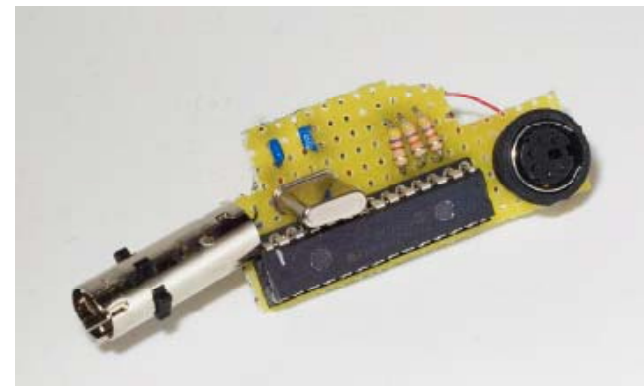
- In addition to digital images, steganography can be used in
 - ◆ Text
 - » Distortion technique
 - ◆ Audio
 - » Discrete Fourier Transformation
 - ◆ Video
- Hide data in “noise” from compression, redundant data, or patterns in format

Other covert channels

- Protocol headers (lots of options)
- Order of accesses
- Timing between accesses
- In general, non-deterministic communications
 - ◆ Options due to randomness or user choice
 - ◆ Tradeoffs: how hard to detect vs bandwidth of channel
 - ◆ Defenses: detection of signal vs adding noise to covert channel

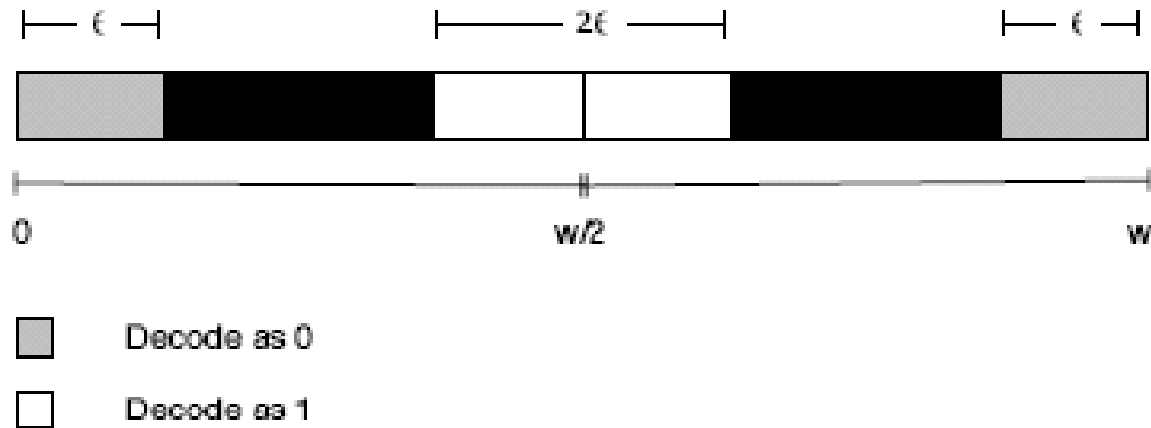
Jitterbugs

- Context
 - ◆ Timing-based covert channels for interactive applications (e.g. telnet/ssh) identified for some time
 - ◆ But require compromising system
 - ◆ Can be detected by comparing time of input via keyboard vs time of output at network
- Alternatively: embed covert channel into keyboard itself

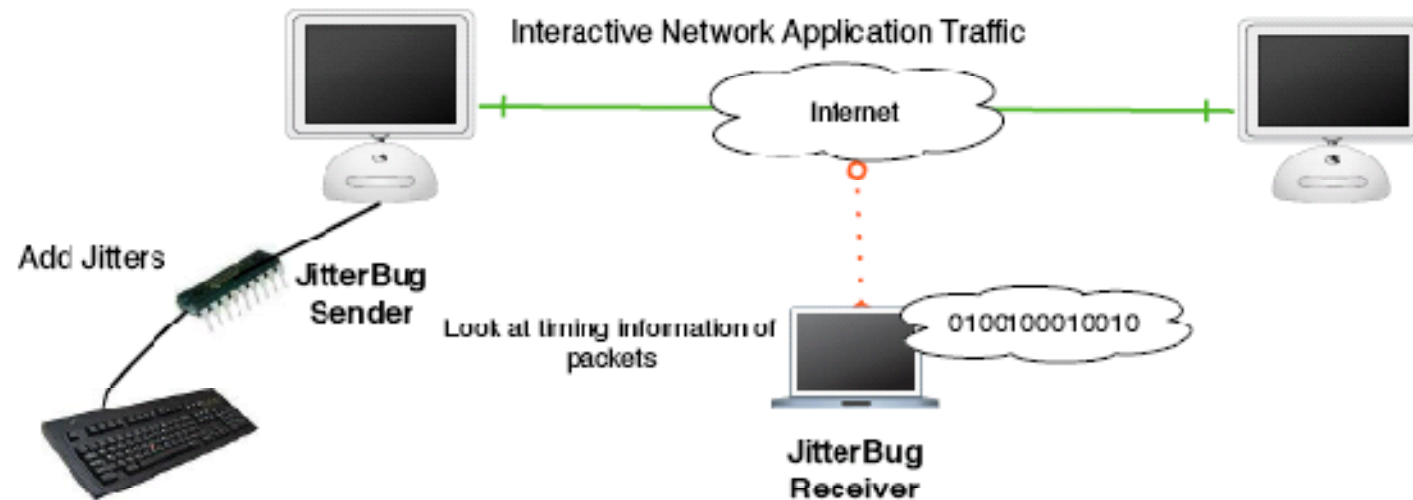


Hiding data in key-stroke inter-arrival times

- Algorithm
 - ◆ If want to send 0 then add no delay
 - ◆ If want to send 1 then add $w/2$ delay



Model of use



Lots of details

- Framing
- Error correction, etc
- Bottom line: $< 1\%$ error across the world for low bandwidth (typing speed)

Aside: Watermarking

- Similar idea to steganography but...
- Message isn't hidden, but is **well-known**
- Attaches meta-data to object that is hard to remove

Steganography vs. Watermarking

- Steganography
 - ◆ Trying to send secret message encoded in cover medium
- Watermarking
 - ◆ Trying to **maintain** secret encoded in cover medium
 - ◆ Adversary can transform the medium
 - » Possible attacks
 - Rotation, translation, scaling transforms (RST)
 - Compression
 - Color/frequency requantization
 - Non-linear transformation (play and record)
 - ◆ Challenge: maintain perceptual transparency while making the watermark difficult to remove
 - ◆ Alternatively: make it difficult to remove watermark without violating perceptual transparency

What's a physical watermark?

- Innovation of the 13th century (Italy)
 - ◆ Process to reduce pulp thickness in paper making
 - ◆ Used to identify different paper producers/lines
- Today used against forgery & for branding



\$100 bill



Security stationery

Digital watermarks

- Similar idea
 - ◆ Embed some digital label into a digital object to identify it
- Why do we care about this?

Copying is the problem

- Binary data is trivial to copy
- But it can also be valuable
 - ◆ Pictures, video, songs, programs, secret nuclear weapon documents, etc
- How do you make digital objects available and manage the copying problem?

Why watermarks?

- Other solutions
 - ◆ Labeled content
 - » Append a label (keyed one-way hash) to the front of the object identifying it?
 - » Problems?
 - ◆ Content identification
 - » Maintain database of copyrighted objects (via content hash function perhaps) and recognize them?
 - If $h(x)$ is in database, then it is copyrighted
 - E.g. Digimarc Marcspider
 - » Problems?

Digital watermarks for copyright identification

- Operation
 - ◆ Embed (O, I): outputs watermarked version of object O , with information I contained in it
 - ◆ Retrieve (O'): examines a watermarked object O' and outputs the embedded information I
- Simply check if file has I and that I is valid
 - ◆ Proof that you have a file that I own
 - ◆ Go crawl Web and find your files
- Problems?

Other applications

- Fingerprinting
 - ◆ Embed unique identifier into each copy of object (identifies source).
 - » Hugely successful in dealing with Academy-associated piracy
- Copy-resistance or integrity verification
 - ◆ Fragile watermark that doesn't withstand copying/transformation
- Rights management
 - ◆ Limit acceptable uses (where, when, how, by whom)
- Authentication
 - ◆ Watermark is used to detect modification applied to cover work
 - » Example : Checking for fraud passport photographs

Authentication

- ◆ Watermark is used to detect modification applied to cover work



Different kinds of watermarks

- Visible vs invisible
 - ◆ Is the watermark clearly evident?
- Fragile vs robust
 - ◆ How well does the watermark survive transformation
- Public vs private (blind vs non-blind)
 - ◆ Do you need the original to implement retrieve?

Visible watermarks

- Useful in special situations
 - ◆ E.g. samples
- Degrades object
- Obvious what needs to be removed
- Easily detected by people, but harder for computers
- Generally, more interest in *invisible watermarks* (transparently embedded)



Inherent tradeoffs for invisible watermarks

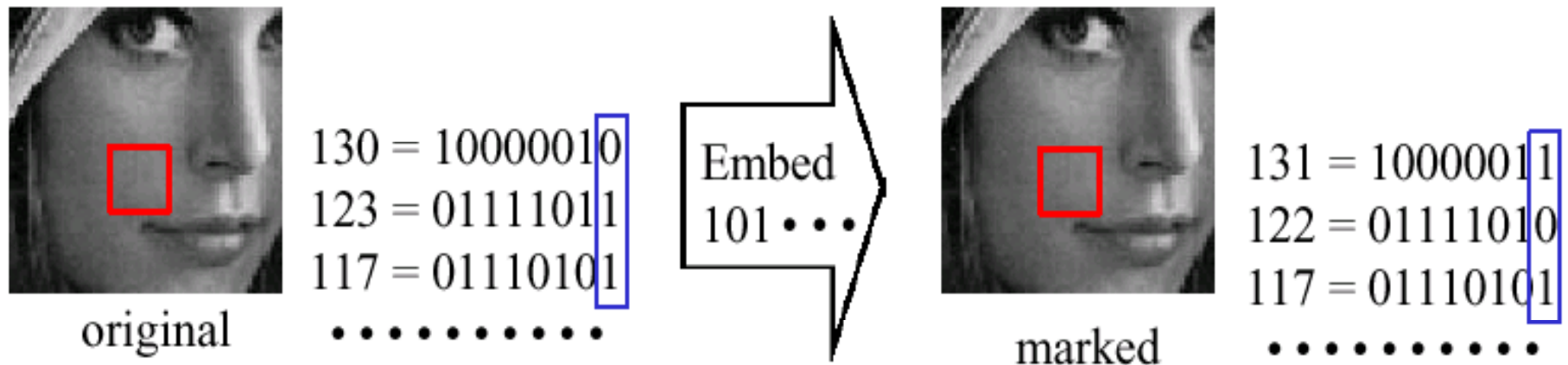
- Generally, robustness increases if the energy in the watermark is larger
- Perceptual quality generally decreases as energy in watermark is larger
- Greater robustness and perceptual quality tend to involve greater overhead

Early work: Jordan-Kutter

- Image watermarking
 - ◆ Embed watermark in LSB of blue plane in the spatial domain
 - ◆ Search different rotations, scales, translations for watermark (RST)
- Problems?

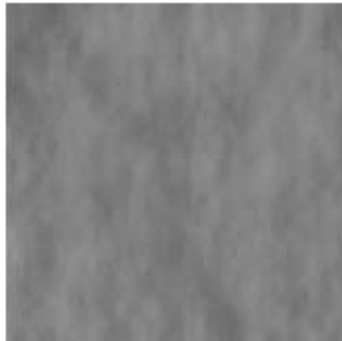
Least Significant Bit Modulation

- Imperceptible: modify only LSBs
- Secure
- Not Robust: random change of LSBs



Transform-domain watermarks

- Watermark added to frequency coefficients of image (e.g. for JPEG)
- Place mark in perceptually important regions
 - ◆ Lower-frequency DCT areas
- Spreads encoding across object
- Strong error correction in encoding



Kinds of watermark attacks

- Removal: remove watermark from object
 - ◆ If scheme is known, then filter out watermark
 - ◆ Averaging N versions of same object with different watermarks
 - ◆ Lossy compression
 - ◆ Write new watermark (limited bandwidth of watermark channel)
- Desynchronization: prevent watermark from being recognized
 - ◆ Global geometric transforms
 - » Translation, rotation, mirroring, scaling, shearing, cropping
 - ◆ Local geometric transforms
 - » Random bending: local shifting, rotation, scaling
 - ◆ Mosaic attack
 - » Cutting the image into pieces

Physical Digital Watermarking

- Embedding digital watermark into physical objects
- Example: Xerox machine fingerprinting

Xerox DocuColor 12 Printer

Section of paper at 60x magnification

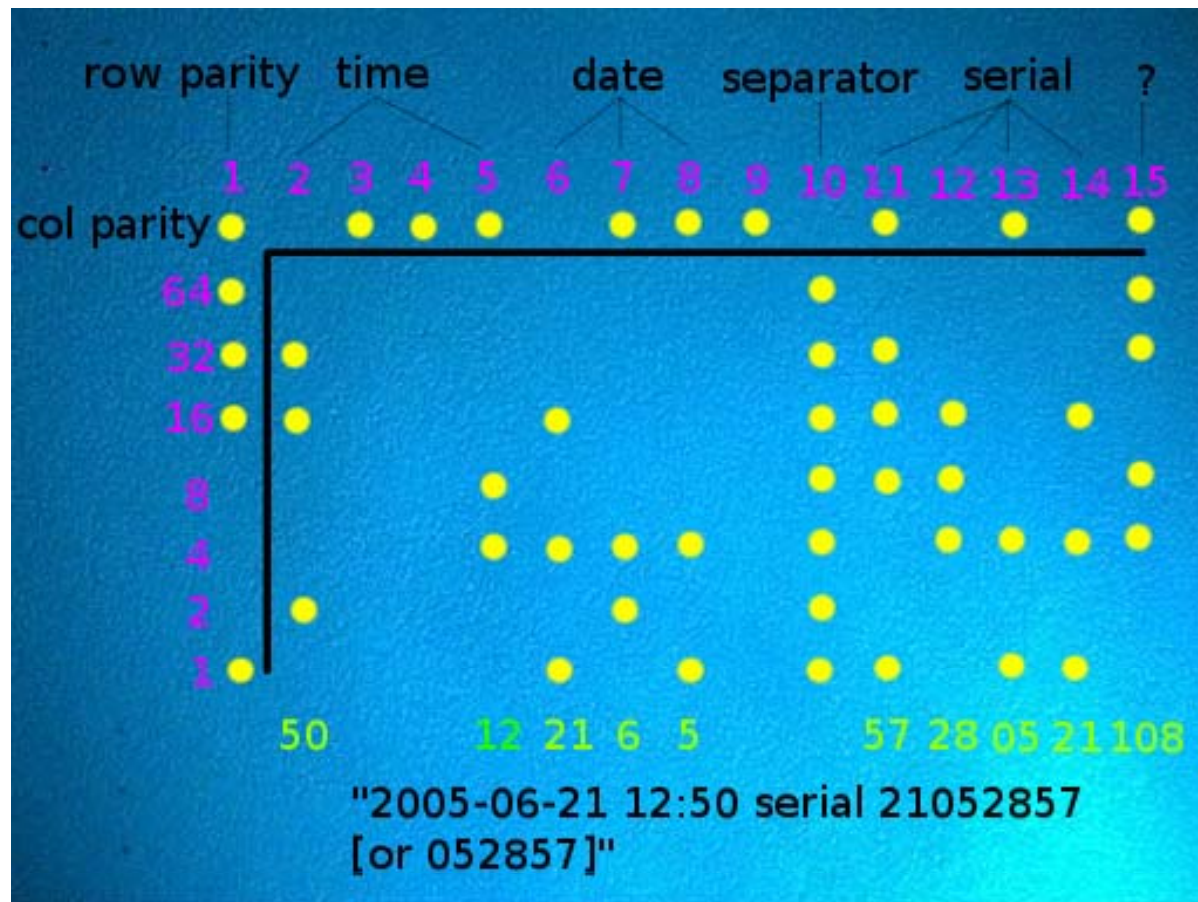


Xerox DocuColor 12 Printer

Section of paper at 10x magnification
(under blue light)



What it means



Next time

- Software security