CSE 227
Computer Security
Winter 2008

Information Hiding &
Covert Channels
Stefan Savage
Quick administrative stuff

- Last class will be project presentations
  - 10mins or less
  - Motivation, what you did, evaluation
  - If you use powerpoint, please send it to me by 11am the day of class (ideally the night before)

- Final papers (~6 pages, two columns) due last day of finals week
  - Please e-mail to me

- Second to last class options
Recap: the problem

- Alice may try to leak secrets by hiding them in normal communications
- This is called a “Covert channel”

Alice might be Evil
Bob might be a foreign agent

Bent wings = 0
Straight wings = 1

Boeing

The Internet

Data

March 19, 2008

CSE 227 – Lecture 10 – Malware I
Covert channels

- 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?
Well known covert channels

- Steganography
  - Hide secrets in cover medium (frequently images) with minimal distortion

Image on left: 61/225 black pixels = 1 bit
Image on right: 60/225 black pixels = 0 bit
Kurack-McHugh method
MSB of secret in LSB of cover

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

The image extracted from the stego-image:
Covert channels

- Protocol headers (lots of options)
- Order of accesses
- Timing between accesses

- What makes a great 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

- Larger alphabets
- Framing
- Error correction, etc
- No doubt one could do better

- Bottom line: < 1% error across the world
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 stationary
Digital watermarks

- Similar idea
  - Embed some digital label into a digital object to identify it

- Is this different from a covert channel?
Big motivation

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?
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
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
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
### 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
    - Requirement of robustness against 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
There may be inherent tradeoffs

- 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
Image Watermarking

- Spatial-domain watermarking
- Transform-domain watermarking
- Perceptual-based watermarking
- Object-based watermarking

- Robust to JPEG compression
- Robust to geometric distortions
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
SDMI Challenge

- Goal: remove watermark from audio track
- Challenge overview
  - A, B, C and F: three files provided
    - Unwatermarked song
    - Same song with watermark
    - A different watermarked song
    - Challenge is to remove watermark from 3rd song
  - Oracle provided to test
Challenge A

- Human’s cannot detect short echos (<1ms)
- Watermark introduces echos periodically with amplitude a and delay t
- Look for ripples in frequency -> echos in time
  - Found between 8-16Khz

March 19, 2008
Echo parameter “hopping”
Removing the watermark

- Once you know how often echoes occur, their gain and delay, can simply extract them from source signal

- The other challenges are similar, but the observation is similar
  - Even not knowing algorithm, possible to remove watermark
  - Question about perceptual impact in general

- Historical context of this paper
Open questions

- Possible to make robust watermarks?
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

"2005-06-21 12:50 serial 21052857 [or 052857]"
Next time

- Denial-of-service