

Computer Security
 16. Steganography, Watermarking, & Content Protection

Paul Krzyzanowski
 Rutgers University
 Fall 2019

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 1

1

steganography

```

    graph TD
      A[steganography] --> B[στεγανός]
      A --> C[γραφία]
      B --> D[covered]
      C --> E[writing]
    
```

The art of secret (hidden) writing

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 2

2

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

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 3

3

Classic techniques

- Invisible ink (1st century AD - WW II)
- Tattooed message on head
- Overwrite select characters in printed type in pencil
 - look for the gloss
- Pin punctures in type
- Microdots (early 20th century)
- Newspaper clippings, knitting instructions, XOXO signatures, report cards, ...

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 4

4


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

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 5

5

 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*.

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 6

6

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!

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 7

7

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.

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 8

8



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.

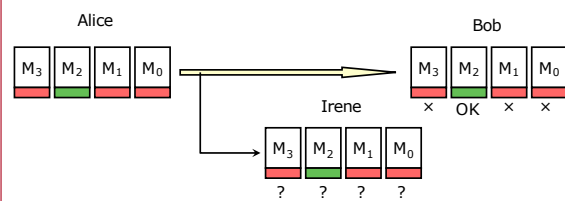
Italicised letters in the first few pages spell out "Smithy Code", while the following pages also contain marked out letters.

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 9

9

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



December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 10

10

Steganography in images

- Spatial domain**
 - Bit flipping
 - Color separation
- Frequency domain**
 - Embed signal in select frequency bands (e.g., high frequency areas)
 - Apply FFT/DCT transform first
- 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)

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 11

11



December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 12

12



13

Video

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

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 14

14

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

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 15

15

Machine ID codes in laser printers

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

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 16

16

Machine ID codes in laser printers

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 17

17

Machine ID codes in laser printers

row parity time date separator serial 7

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

col parity

64
32
16
8
4
2
1

00 12 21 4 3 77 08 05 21 2005


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

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

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 18

18

UV Watermarking




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

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 19

19

Text

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

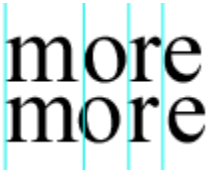


December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 20

20

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

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 21

21

Text-based steganography

“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*

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 22

22

Watermarking vs. Steganography

Both techniques hide 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

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 23

23

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

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 24

24

Content Protection

December 10, 2019

CS 419 © 2019 Paul Krzyzanowski

25

25

Content protection

- Digital content is simple to copy and distribute
 - Software, music, video, documents
- That's not always good
 - How do software companies & artists make a living if their content is freely distributed on a large scale?
 - Maintain revenue streams
 - Enforce distribution rights (e.g., video available in the U.S. first)
 - How do organizations keep their documents secure?
 - Enforce confidentiality & protect trade secrets?

How can we make illegal content access difficult?

December 10, 2019

CS 419 © 2019 Paul Krzyzanowski

26

26

DRM

- Content industry (movies, music, documents) asked for technical solutions to the content distribution problem
- This led to **digital rights management (DRM)**
 - Protection of content
 - Definition on how it can be played and copied
- Not just documents & movies:
 - Printer cartridges
 - John Deere tractors
 - Keurig coffeemakers
 - RFID connections enforce use of Keurig-branded K-cups

December 10, 2019

CS 419 © 2019 Paul Krzyzanowski

27

27

Associate software with a computer

Find unique characteristics of a machine

1. CPU serial number (early microprocessors didn't have these)
2. Add a dongle (USB hardware key)
3. Create a unique ID based on PC's configuration
4. Install software in a way that cannot be copied (e.g., mark blocks as bad)
 - Used on early PCs but not viable with modern operating systems

But

- You can go through the software with a debugger & remove checks
 - This becomes harder as software gets bigger ... but not impossible

December 10, 2019

CS 419 © 2019 Paul Krzyzanowski

28

28

Copy or execution protection

- On-device or on-installation checks
 - Software is configured to check a computer ID or license key when run
 - May validate online via a server
- Continuous or periodic network checks
 - Software must contact an on-line license server & identify itself and the computer to run
 - Subscription services do this: Adobe, Autodesk, Microsoft
- Timebombs
 - Software ceases to function if it's found to be illegally installed
 - Illegal in some places

All checks can be defeated

Goal: balance technical difficulties, user convenience, and legal repercussions

December 10, 2019

CS 419 © 2019 Paul Krzyzanowski

29

29

Cloud software

Ultimate protection

- Company provides **both** the computing platform and the software
 - And you don't have physical access to the platform
- If your subscription expires, you cannot use the platform

December 10, 2019

CS 419 © 2019 Paul Krzyzanowski

30

30

Documents & Books

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 31

31

Documents

- Trusted readers & content management
 - Microsoft Office, Adobe PDF, AutoCAD
- E-book readers
 - EPUB (default format for Apple)
 - MOBI (Mobipocket, purchased by Amazon in 2005)
 - AZW, AZW3 (Amazon Kindle – similar to MOBI)
 - PDF (Adobe)

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 32

32

MOBI/AZE formats

- MobiPocket
 - Acquired by Amazon in 2005 and used in Kindle
 - Reverse engineered & source published
- Uses a PC1 symmetric cipher with a 128-bit key
 - key = encrypted with temp key
 - temp key = Encrypt device ID with global_secret_key
- Main weakness
 - Device ID is 7 alphanumeric chars, only upper-case, followed by '\$'
- Kindle AZW format
 - Essentially Mobipocket with a device ID that ends with * instead of \$

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 33

33

Media

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 34

34

Apple FairPlay

- AAC & MP4 files
- Content encrypted via AES
- Encrypted master key stored in the MP4 container file
- User key decrypts master key
 - User obtains user key when device authorizes with Apple's servers
- Has been reverse-engineered

user_key

↓

E _{user_key} (master_key)	E _{master_key} (content)
------------------------------------	-----------------------------------

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 35

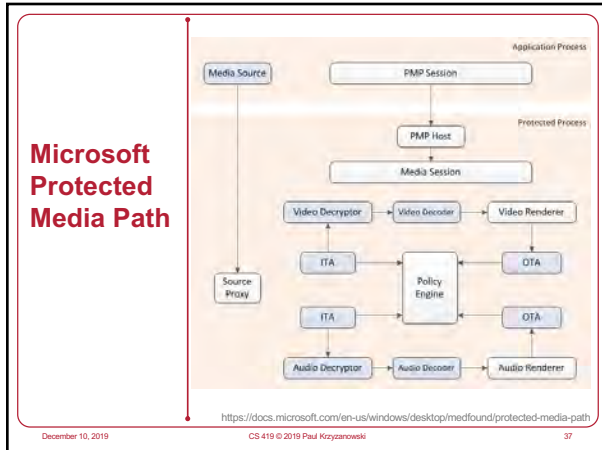
35

Microsoft Protected Environment (PE)

- Enable protected content to flow through the OS in a protected manner.
 - All components in the protected environment must be **trusted**.
 - **Trusted component** = components signed by Microsoft, including kernel modules
 - OS stops DRM-restricted content from playing while unsigned software is running
- Content flows through trusted components:
 - **Protected Media Path** (includes *Protected Video Path*)
- Media source specifies the rights for using the content
 - Play, transfer, etc.
- Final output
 - Decrypted, uncompressed video frames travel on a physical connector to the display device
 - Providers may require protection in this area, such as the use of High-Bandwidth Digital Content Protection (HDCP) or DisplayPort Content Protection (DPCP)

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 36

36



37

Broadcast Video

- Content is encrypted
- Key is transmitted via **Entitlement Management Messages (EMM)**
- The trick is to send the key so that only allowed users can get it

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 38

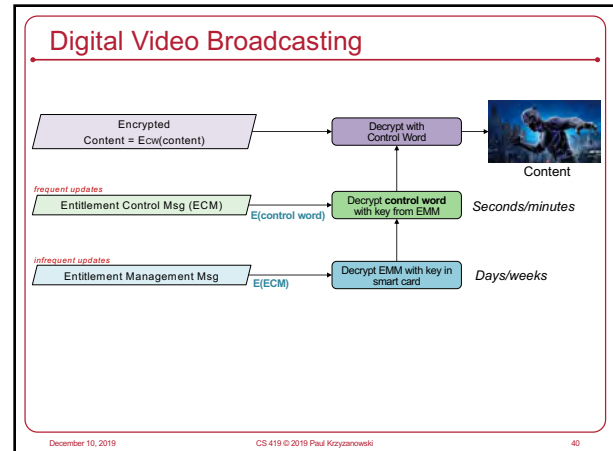
38

Digital Video Broadcasting

- **Relies on trusted hardware**
 - Data stream is decrypted with smart cards containing subscriber info
- Source content is encrypted with a 48-bit secret key (**key = control word**)
 - Control word may change several times per minute
 - Control word is encrypted & sent to all subscribers as part of an **Entitlement Control Message (ECM)**
- Key to decrypt the ECM is sent to each subscriber as an **Entitlement Management Message (EMM)**
 - Sent at less frequent intervals (several days to several weeks)
 - Encrypted per subscriber for their smart card

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 39

39



40

CableCARD

- Card device to allow customers to access digital cable TV channels on generic devices
- Identifies and authorizes subscriber
- Receives **EMM (Entitlement Management Messages)** for premium channels
- Decodes encrypted digital cable signal
 - Performs conditional access logic & decryption
 - Provides an MPEG-2 media stream to the host
 - Tuner and MPEG decoder are part of the host equipment

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 41

41

CableCARD

- CableCARD did not provide host device certification for two-way communication
- Deployment of proprietary set-top boxes is far bigger than CableCard
- Next (possible) successor: **AllVid**
 - Universal adapter for all types of pay TV and interactive program guides
 - Can communicate to any device with a screen
 - Endorsed by Google, Best Buy, Mitsubishi, Sony Electronics, TiVo
 - *Not endorsed by cable companies*

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 42

42

DVD Content Scrambling System (CSS)

- Stream cipher – weak – based on 25-bit key – can be broken in 2^{25} tries
- Each player has one or more manufacturer-specific keys
- Each DVD has a disk key encrypted under **each** of the manufacturer's keys
 - Goal was to produce new disks that omit a specific manufacturer's key if it leaked
 - BUT – given any key in the system, all others can be found
 - Manufacturers had an incentive to keep costs down, not use tamper-resistant hardware
- DVD players on PCs
 - PCs are an open platform – only way to "protect" the code was to obfuscate it

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 43

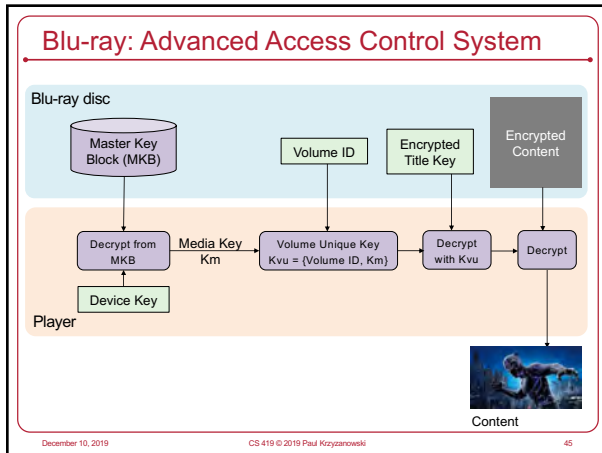
43

Blu-ray: Advanced Access Control System

- Blu-ray disc contains
 - Encrypted content**: encrypted with a **Title Key**
 - Encrypted Title Key**: Encrypted with a **Volume Unique Key (Kvu)**
 - Volume ID (VID)**: serial number of disc – will not be duplicated
 - Media Key Block (MKB)** = lots of encrypted keys (~50 GB)
 - Allows each compliant device, **using its secret device key**, to compute a Media Key
- Player contains
 - One or more secret **Device Keys**
 - 128-bit keys provided to trusted parties by the AACS org
 - Device Keys may be unique per device or – often – shared by multiple devices
- Decryption
 - Use Device Key to decrypt a Media Key (Km) from the Media Key Block (MKB)
 - Combine **Media Key & Volume ID** to get the **Volume Unique Key (Kvu)**
 - Use **Volume Unique Key** to decrypt the **Encrypted Title Key**
 - Decrypt the content with the **Title Key**

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 44

44



45

Blu-ray: Advanced Access Control System

- DVD CSS**
 - Unique encryption key for content – key is encrypted for a set of players
 - All players of a model group have the same decryption key
 - Disc contains several hundred encrypted keys, one for each licensed player model
- Blu-ray AACS**
 - Unique media key for content – key is encrypted for a player
 - Each individual player has a unique set of decryption keys
 - Licensors can revoke keys for individual players in future content
 - AACS keys compromised since 2007 – keys were found using debuggers

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 46

46

Content isn't really protected

- People built databases of media keys – so no need to decrypt the media key**
 - Do a google/bing search for **AACS KEYDB.cfg**
 - <https://gist.github.com/HenkPoley/41ed899251aa771cb1d061d49a3888e5>
 - 18 processing keys
 - 23,999 titles as of 9/3/2017
- There's also the **analog hole**

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 47

47

Widevine Content Protection

- Google's Widevine DRM Platform**
 - Used by Google Play Movies, Hulu, Spotify, Netflix, Amazon Prime Video, Disney+
 - Supported by:
 - Chromium, Firefox, Opera (not Firefox or MS Edge)
 - Most smart TVs
- Over 30 chipsets support Widevine**
 - ARM Trusted Execution Environment (TEE) handles:
 - Rights management, integrity management, firmware updates, authentication
- Google licenses the code only to approved developers, devices, and applications**
 - This is closed source software!

December 10, 2019 CS 419 © 2019 Paul Krzyzanowski 48

48

Widevine security levels

Level 1 (L1)

- All content processing, cryptography & control must be performed within the TEE
- Usually needed to access HD content

Level 2 (L2)

- Cryptography must be performed within the TEE but not video processing

Level 3 (L3)

- Used when the device does not have a TEE or processing is done outside the TEE
- Appropriate measures must be taken to protect the media stream within the host OS
- Broken in Jan 2019

December 10, 2019

CS 419 © 2019 Paul Krzyzanowski

49

49

Legal barriers: DMCA

Digital Millennium Copyright Act

Criminalizes production and dissemination of technology, devices, or services intended to circumvent measures (DRM) that control access to copyrighted works. It also criminalizes the act of circumventing an access control, whether or not there is actual infringement of copyright itself.

Without DMCA, anyone would be able to build a set-top box to decode video signals

- Just crack HDCP (High Definition Content Protection)

Also

- Licensing agreements (EULAs)
- EU's Copyright Directive

December 10, 2019

CS 419 © 2019 Paul Krzyzanowski

50

50

The end

December 10, 2019

CS 419 © 2019 Paul Krzyzanowski

51

51