

Steganography

COMP 522

What to protect

Level	What to protect	Method
3	Existence of message	Steganography
2	Metadata of message	Privacy-enhancing technologies
1	Content of message	Encryption
0	Nothing	None

Table by I.A. Goldberg

Metadata of message here is: the sender, the recipient, the time the message was sent, or the length of the message, etc

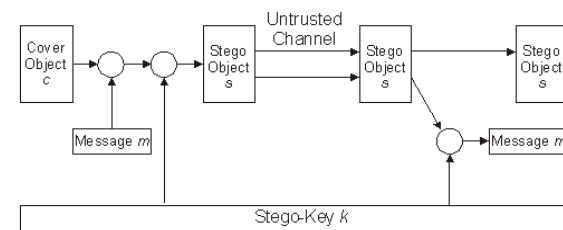
COMP 522

Steganography and information hiding

- **Steganography**, derived from “covered writing” in Greek
- It includes the methods of secure communications that conceal the **very existence** of the message
- Examples (non-digital): invisible ink, microdots, etc

COMP 522

Steganography in digital world



Picture by C.Shoemaker

COMP 522

Digital watermarking

Digital watermarking:

- aim is to embed an amount of information that could not be removed or altered without making the cover object entirely unusable
- adds additional requirement of **robustness** as compared **with** steganography
- Can be used for copyright protection

COMP 522

Texts as cover objects

Text as a cover object:

Apparently neutral's protest is thoroughly discounted and ignored. Isman hard hit. Blockade issue affects pretext for embargo on by products, ejecting suets and vegetable oils.

(Real example of the text sent by a spy in WWII)

COMP 522

Texts as cover objects

Text as a cover object:

*A

pparently neutral's protest **is** thoroughly **dis**counted **and** **ign**ored. **Is**man hard hit. **Bl**ockade **iss**ue **af**fects **pre**text **fo**r **emb**argo **on** **by** **pro**ducts, **e**jecting **su**ets **and** **ve**getable **oil**s*

Taking the second letter in each word gives the message: **Pershing sails from NY June 1**

COMP 522

Images as cover objects

LSB (Least Significant Bit) substitution method:

- Least significant bits used to store characteristics of particular pixels of an image (cover object) are modified to store a message
- Colours and lightness of pixels of obtained image may differ slightly from original cover image, but both images looks identically to human eye.
- Easy to implement, but not too robust methods
- Transformations of images may easily destroy the message (watermark)

COMP 522

Images as cover objects



Watermarked image (LSB substitution)



Watermark recovered

Note: watermark is embedded as the image, not the plain text to improve robustness

COMP 522

Advantages and disadvantages of LSB

Advantages of LSB

- easy to implement
- has high capacity

Disadvantages of LSB

- is not robust
- message is easy to detect:
 - A message insertion introduces distortion to the statistical properties of image which never naturally appear

COMP 522

Stochastic modulation method

Simple variant:

- Before embedding a message a randomly chosen pixels are altered by changing their intensities (= a number between 0 and 255) by +1 or -1;
- For a parameter p in $[0;1]$ a pixel intensity is increased/decreased by 1 with probability p ; it is left unchanged with probability $1-2p$;
- Then LSB method is used

Provide more protection against detection of the message

COMP 522

Stochastic modulation method

Improved method (J.Fridrich, M.Goljan):

The idea:

- take a cover image and add a “noise” modulated by a message bits
- “noise” actually means pseudo-noise here , that is a sequence of pseudo-random values, which can be generated deterministically given a secret initial value (key)
- If initial value (key) is known then generation of pseudo-noise can be repeated (used for extraction of the message)

COMP 522

Stochastic modulation

- Simple implementation;
- High capacity;
- Low embedding and extraction complexity
- Embedding noise can have arbitrary characteristics and may approximate the noise of a given device => high security

COMP 522

Transform space algorithms

- **Jsteg** algorithm (D.Upham) uses specifics of JPEG image format
- For each colour component JPEG image format uses *discrete cosine transform* (DCT)
- DCT is used by JPEG to transform consecutive 8 by 8 pixel blocks of the image to 64 DCT coefficients each:
$$F(u, v) = \frac{1}{4} C(u) C(v) \left[\sum_{x=0}^7 \sum_{y=0}^7 f(x, y) * \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \right]$$
 where $C(x) = 1/\sqrt{2}$, if $x = 0$ and $C(x) = 1$ otherwise

COMP 522

Transform space algorithms

- JPEG: after quantization DCT coefficients are stored;
- Jsteg algorithm:
 - Replace sequentially the least-significant bit of discrete cosine transform coefficients with the message data
- Gives better protection (as many others TS algorithms) against visual attacks

COMP 522

Audio (video) files as cover objects

- LBS can be used, but it introduces a significant noise to audio data ;
- A message may be encoded in audio signal *phase*, replacing original phase with a reference phase representing a hidden message; more difficult to implement;
- *Spread spectrum* method: encoded data spread across the maximum range of frequencies; difficult to detect hidden message;
- Video objects (files, streams) can be used for hiding information as well;

COMP 522

Network packets as cover objects

- Steganography within TCP/IP:
 - Insert data within TCP and IP protocol headers
 - IP identifier, TCP initial sequence number, least significant bit of TCP timestamp, IP flags.
 - Relatively easy to detect naive embedding by anomaly detection in TCP/IP fields
- One can prevent easy detection by taking into account the properties of concrete implementations of TCP/IP (Murdoch, Lewis, 2005)

COMP 522

Redundancy

- Steganography is applicable to any data objects that contain redundancy;
- Redundancy is used to hide the presence of the embedded message
- On the other hand redundancy may be removed during data compression
- One may combine data compression and message embedding: MP3stego by F.Petitcolas

COMP 522