

Watermarking: Device Control 1

Add value, rather than restrict use.

- In the radio for payed customer.
 - It is music or commercial now.
- Automatically turn on Dolby FM decoder.
 - It is Dolby FM signal or not.
- Interactive TV.
 - Command a toy with light sensor
- Imperceptible bar code for mobile phone in
 - Magazine advertisements, packaging, tickets ...

Watermarking: Device Control 2

Legacy Enhancement? Device Control!

- Digital watermarking in analog communication for new device
 - International air traffic control system.
- Synchronizing audio and video signals
 - A highly compressed audio signal within the video signal as a reference.
- Embed the lyrics directly into the songs
 - MediaSync

2.2 Applications of Steganography

Applications of Steganography

Motivation

- Benign: Lovers, Gay
- Political: Dissidents.
- Criminal: Organized crime or terrorism.

Successful steganography is not detectable, so many of the most successful applications may never become public.

Steganography: for Dissidents

Public and wide surveillance: Everyone is monitored and aware of this.

- Encryption: indication of conceal.
 - arrest and get decryption key through torture.
- Anonymous remailer:
 - A server forwarding message to receiver after removing the origin.
 - Use of an anonymous remailer is a clear sign.
 - Arrest and interrogate A, but B is safe.
 - Block the remailer.
- Steganography:
 - Basic and benign communications cannot be forbidden.

Steganography: for Criminal

Conceal and restricted surveillance: People is usually unaware of being monitored.

- Encryption: indication of conceal.
 - Infer co-conspirators from the recipients.
- Anonymous remailer:
 - Law enforcement agencies access the server.
- Steganography:
 - The computer with stego program is seized.

Easier in this situation!

2.3 Properties of Watermarking Systems

Properties of Watermarking Systems

Embedding:

- Effectiveness, fidelity, and payload.

Detection:

- Blind or informed detection, false positive behavior, and robustness.

Security feature:

- Security and the use of secret keys.

Embedding Effectiveness

Detection immediately after embedding.

Why cannot 100%?

- Cannot be successfully watermarked within certain fidelity constraints.

Measurement

- Analytically.
- Estimated empirically from samples.

Fidelity

Definition:

- Perceptual similarity between the original and watermarked version.
- At the point at which they are presented to a consumer after transmission.

Depends on

- Transmission quality.
- Requirements on quality.

Data Payload

Number of bits in a unit or a work.

- Typical units:
 - Pixels (image), per second (audio), per frame (video).
- Different requirements:
 - Copy control applications: 4-8 bits.
 - Broadcast monitoring: 24 bits.
- Zero-bit watermarks:
 - A watermark is present or not.
 - N bits encodes $2^N + 1$ possible outputs.

Blind or Informed Detection

- Blind: No any information related to the original.
 - Public watermarking systems.
 - Copy control etc..
- Informed: Requires access to the original, unwatermarked work.
 - Private watermarking systems.
 - Transaction-tracking etc..

False Positive Rate

- False positive
 - The detection of a watermark that does not actually exist.
- Probability
 - Fixed work, randomly selected watermarks:
 - Transaction tracking.
 - Fixed watermark, randomly selected works:
 - Copy control
- Rate: **Presentation: 7.1-7.2**
 - Proof of ownership: 10^{-6} .
 - Copy control: 10^{-12} .

Robustness

- Ability to detect the watermark after common signal processing operations
 - Image: filtering, compression, distortions ...
 - Video: recording, change of frame rate ...
 - Audio: DJ mixture ...
- Need only survive the operations between embedding and detection.
 - Broadcast monitoring: lossy compression, analog transmission, additive noise ...
- Even irrelevant or undesirable.
 - Fragile watermark in authentication.

Security 1

Ability to resist hostile attacks.

From the viewpoint of watermarking

- Unauthorized removal: active attacks
- Unauthorized embedding: active attacks
- Unauthorized detection: passive attack

About the viewpoint of cover work

- Active attacks: modify the cover work.
- Passive attacks: do not modify ...

It is not important for some applications (e.g. enhanced functionality to consumers).

Security 2

Unauthorized removal: prevent the watermark from being detected.

- Elimination attacks: No watermark at all.
 - Does not necessarily mean reconstructing the original, unwatermarked work
- Masking attacks: Still there, but undetectable EASILY.
 - Could still be detected by a more sophisticated detector.
- Collusion attack: combines several copies of a given work with different watermarks.
 - In transaction tracking

Security 3

Unauthorized embedding (forgery)

- Content Authentication
- Proof of Ownership
- ...

Security 4

Unauthorized detection (passive attacks), can be broken down into three levels of severity.

- 1 Detects and deciphers the message.
- 2 Detect the watermark and distinguish it from another.
- 3 To determine that a watermark is present.

More concern in steganography than in watermarking, but

- Broadcast monitoring company: I embed it for my use only.

Cipher and Watermark Keys

Need to read chapter 3 and 10.

Modification and Multiple Watermarks

- Modify watermarks is useful in some applications (e.g. copy control):
 - copy-once to copy-no-more
 - make a copy of a broadcast for the noncommercial purpose.
 - risk of modifying copy-once to copy-freely.
- Multiple watermarks:
 - Embed/Remove the second watermark for copy-no-more.
 - Transactional watermarks: owner, distributor, purchaser.

Cost

Two principal issues of concern:

- Speed/Complexity.
 - Broadcast monitoring: real time.
 - Proof of ownership: takes days.
- Number of embedders and detectors.
 - Copy control: a few embedders and many detectors (in every playback device).
 - Transaction-tracking in DiVX: each player embeds a distinct watermark, a few detectors.

Other issue: special-purpose hardware devices or as software applications or plugins.

2.4 Evaluating Watermarking Systems

“Best”

- The criteria depends on application
 - Robust to rotation: useful for copy control, but not for broadcast monitoring.
- Improvement in one may help others
 - More payload: 20 bits → 10 bits.
 - Less false positive probability: 20 bits message and 10 bits checksum ($2^{-10} = 1024$).
- Related properties:
 - Embedding strength: robustness and fidelity.
 - Detection threshold: robustness and false positive probability.

Benchmarking

- CPTWG's effort to test watermarks for copy control in DVD recorders.
- Image watermarking: embed 80 bits with certain fidelity, and then distorted the image (via StirMark) to measure the robustness.
 - Some applications only use a few bits.
 - Assume adversary has no the detector.

No benchmark can ever be relevant to all watermarking systems and applications.

Scope of Testing

- Typicality of the test set: photographs, X-ray images, animation frames.
- To test false positive rates (e.g. 10^{-6})
 - Using millions samples.
 - Verify some statistical model used to predict it from less samples.

In this book: 20,000 images including photographs, but some of them are paintings and textures.

2.5 Properties of Steganographic and Steganalysis Systems

Properties of Steganographic and Steganalysis Systems

Embedding:

- Embedding effectiveness, fidelity, capacity, efficiency.

Detection:

- Blind or informed extraction, blind or targeted steganalysis, statistical undetectability, false alarm rate, robustness.

Security feature:

- Security, stego key.

Embedding Effectiveness

The probability of an error when the detector is applied immediately after embedding.

- Constraints from fidelity.
- Watermarking: Fail.
- Steganography: Choose the most appropriate cover work.
 - The message is not about the cover work.

Fidelity

Not a primary property for steganography.

- Steganalysis systems do not have access to the original cover work.

Capacity, Efficiency

- Embedding capacity
 - The maximum number of bits that can be hidden in a given cover Work.
- Steganographic capacity
 - Negligible probability of detection.
 - = watermark payload \leq embedding capacity.
 - Difficult to determine.
 - More bits, more risk.
- Embedding efficiency
 - Number of bits embedded per unit distortion.
[Read section 12.5.1.](#)

Blind or Informed Extraction

- Steganalysis systems do not have access to the original cover work.
- But A,B have.
 - Weak embedding is enough for less risk.

Blind or Targeted Steganalysis

All steganalysis algorithms can be categorized as

- Blind: Independent of the steganography method used.
- Targeted: Specific to one or more particular steganographic methods.
- System attacks: implementation weakness, e.g. insufficient stego keyspace.

Blind

Idea:

- Converting works into high-dimensional feature space.
- Clustering (cover and stego works) via machine learning.

Advantages

- Potentially detect an unknown stego scheme.
- Classifying stego works to individual steganographic methods.

Targeted

Method

- Designed for each individual steganographic method.

Advantages

- More accurate.

Statistical Undetectability 1

Stego Work should be inconspicuous

- Elusive and difficult to formalized/quantified.

Quantified properties

- Histograms and a variety of higher-order statistics.
- Compliant with the intended use of the communication channel.
- Hack/joke: Although it is risky to warden, one can use overt communication.

Statistical Undetectability 2

Assumptions

- The same to cryptographic, Kerckhoffs' Principle
 - Completely know all the methods.
 - Only the secret key is unknown.
- Different from cryptographic:
 - Presence \geq Content.
- Adversary knows statistical distribution of all valid cover works.
 - An image about food in a vehicle related conversation.

False Alarm Rate

= false positive rate in watermarking.

- Wrong detection alerts the conspirators.
- More costly verification.

Never detecting the presence of a stego work.

- 100% missing stego works.
- Trade-off between false positives and false negatives
 - Receiver operating characteristic (ROC) curves.

Robustness

Less important than watermarking.

- Usually via digital network (e.g. Internet).

Sometimes, should be considered

- Broadcasting.
- Complex transmission (spy).

Security

The ability to resist hostile attacks (active and malicious).

Three types of warden

- Passive: passively observes the communication.
 - Statistical undetectability.
- Active: distort, compress the cover work.
- Malicious: remove the message or impersonate the communicating parties.

Or = undetectability (depends on context).

Stego Key

Conjunction with a publicly known steganographic algorithm.

Two keys

- Crypto key: encrypt the message.
- Stego key: [read chapter3, section 12.4.3.](#)

Different from usual crypto key:

- Sometimes, key length is not that critical.
[read 13.2.2.](#)

2.6 Evaluating and Testing Steganographic Systems

Evaluating and Testing 1

Which one is better?

- Currently no universally accepted tests or standards.

Using blind steganalysis

- Classifier from supervised learning:
 - Training set
 - Testing set
- Empirical procedure depends on
 - Steganalysis algorithm
 - Data set.

Evaluating and Testing 2

"Practically secure" steganographic scheme

- No successful targeted attack.
- Resist current blind steganalysis.

The status of a steganographic method may suddenly change with the appearance of a new steganalysis method.

- MD5 collision
 - A malicious program with official md5 check sum.
 - <http://blog.csdn.net/singlerace/article/details/1360400>