A General Framework for the Structural Steganalysis of LSB Replacement

Andrew Ker

adk@comlab.ox.ac.uk

Royal Society University Research Fellow
Oxford University Computing Laboratory

7th Information Hiding Workshop
8 June 2005

A General Framework for the Structural Steganalysis of LSB Replacement

Outline of Presentation

- Detection of LSB Replacement steganography
- Analysis of “structural properties” of LSB operations: extend from pairs of samples (already known and exploited) to triplets (novel)
- Experimental results for new detector

For more on the general framework itself, analysis of “structural properties” of LSB operations for groups of arbitrary size, and some details, read the paper.

LSB Replacement

- Extremely simple spatial-domain embedding method: secret payload overwrites least significant bits of cover.

- Can be performed without specialist stego software.

- Visually imperceptible but highly vulnerable to statistical analysis.

- Nonetheless, not reliably detectable if hidden payload is short enough (of the order of 0.01 secret bits per cover byte).

LSB Replacement

- Extremely simple spatial-domain embedding method: secret payload overwrites least significant bits of cover.

- Can be performed without specialist stego software.

```
perl -n0777e '$_=unpack"b*",$_;split/(\s+)/,<STDIN>,5;
              @_[8]=~s{.}{$&&v254|chop()&v1}ge;print@_'
              <input.pgm >output.pgm stegotext
```

- Visually imperceptible but highly vulnerable to statistical analysis.

- Nonetheless, not reliably detectable if hidden payload is short enough (of the order of 0.01 secret bits per cover byte).

LSB Replacement

- Extremely simple spatial-domain embedding method: secret payload overwrites least significant bits of cover.

- Can be performed without specialist stego software.

```
perl -n0777e '$_=unpack"b*",$_;split/(\s+)/,<STDIN>,5;
              @_[8]=~s{.}{$&&v254|chop()&v1}ge;print@_'
              <input.pgm >output.pgm stegotext
```

- Visually imperceptible but highly vulnerable to statistical analysis.

  Structural property: *even cover samples can only be incremented*
  odd cover samples can only be decremented

- Nonetheless, not reliably detectable if hidden payload is short enough (of the order of 0.01 secret bits per cover byte).

Detection Literature

1. “Signal processing”-style detectors

Not specific to LSB Replacement
Not very sensitive

Detection Literature

1. “Signal processing”-style detectors

2. “First generation” structural detectors

e.g.

Chi-square	[Westfeld]
Raw Quick Pairs	[Fridrich]

Make use of structural properties of LSB replacement on individual pixels
Not very sensitive

Detection Literature

1. “Signal processing”-style detectors

2. “First generation” structural detectors

3. “Second generation” structural detectors

e.g.

RS	[Fridrich *et al*]
Pairs	[Fridrich *et al*]
Sample Pairs a.k.a. Couples	[Dumitrescu *et al*] [Ker]
Difference Histogram	[Zhang & Ping]
Least Squares Sample Pairs	[Lu *et al*]

Make use of structural properties of LSB replacement on (mostly) ***pairs of*** *pixels*
All estimate the amount of hidden data
Seem to have a lot in common

Detection Literature

1. “Signal processing”-style detectors

2. “First generation” structural detectors

3. “Second generation” structural detectors

e.g.

RS	[Fridrich *et al*]
Pairs	[Fridrich *et al*]
Sample Pairs a.k.a. Couples	[Dumitrescu *et al*] [Ker]
Difference Histogram	[Zhang & Ping]
Least Squares Sample Pairs	[Lu *et al*]

“Almost Couples” Steganalysis

We look at adjacent pairs of pixel values, and the effects of LSB operations on them.

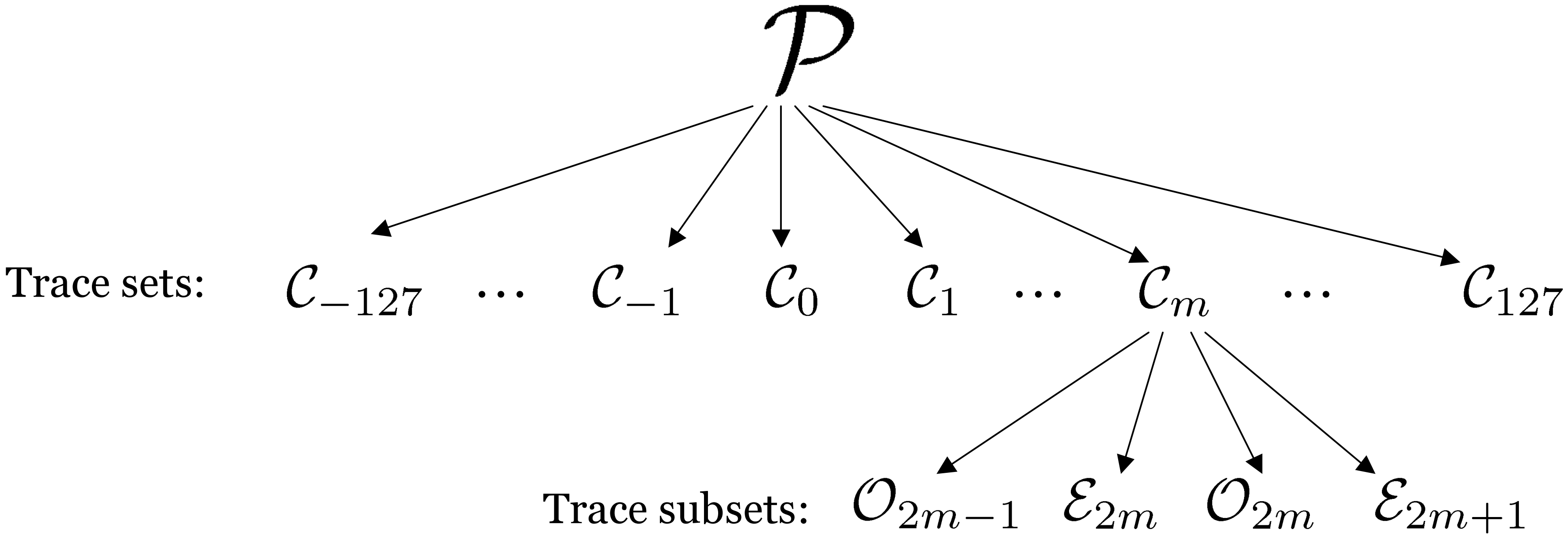
Definitions (sets of pairs)

$\mathcal{P}$	all pairs (x, y) used in the analysis
$\mathcal{C}_m$	values divide by two to give a pair of the form $(u, u + m)$
$\mathcal{E}_m$	pairs of the form $(x, x + m)$ where x is even
$\mathcal{O}_m$	pairs of the form $(x, x + m)$ where x is odd

e.g. if 66 and 72 are the values of two adjacent pixels then (66,72) is in $\mathcal{P}$, $\mathcal{C}_3$ and $\mathcal{E}_6$

Trace Sets

$\mathcal{P}$	all pairs (x, y) used in the analysis
$\mathcal{C}_m$	values divide by two to give a pair of the form $(u, u + m)$
$\mathcal{E}_m$	pairs of the form $(x, x + m)$ where x is even
$\mathcal{O}_m$	pairs of the form $(x, x + m)$ where x is odd



$\mathcal{P}$

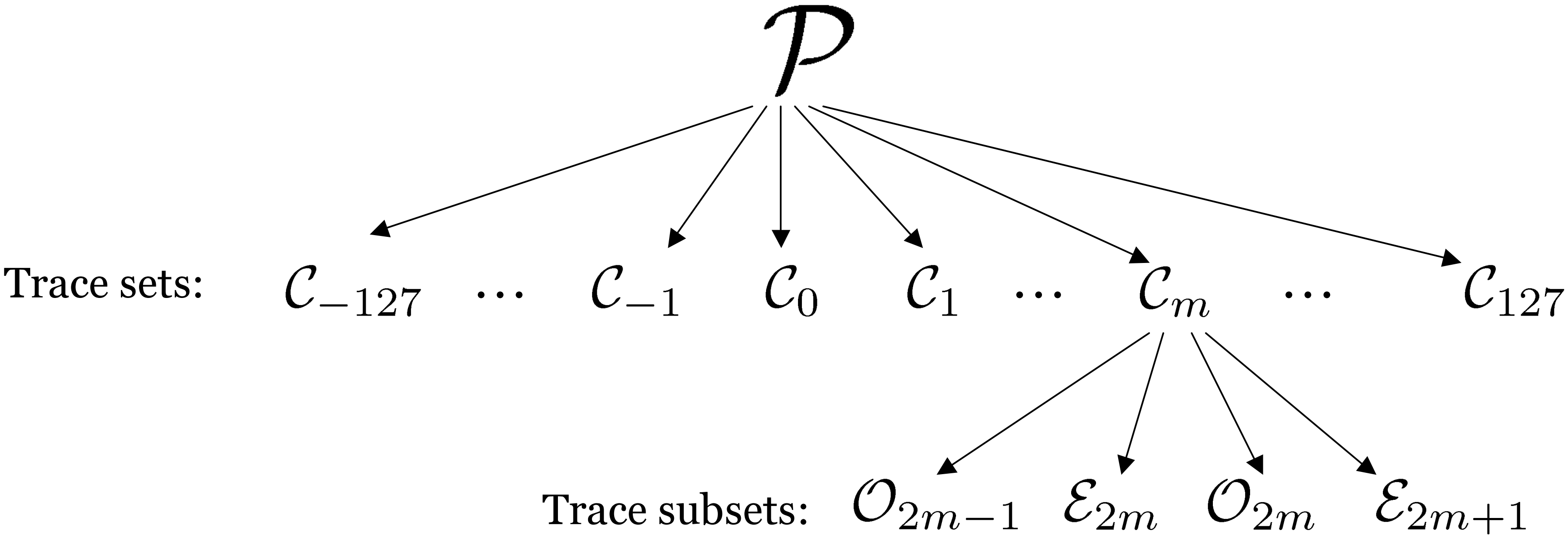
Trace sets: $\mathcal{C}_{-127}$ $\cdots$ $\mathcal{C}_{-1}$ $\mathcal{C}_0$ $\mathcal{C}_1$ $\cdots$ $\mathcal{C}_m$ $\cdots$ $\mathcal{C}_{127}$

Trace subsets: $\mathcal{O}_{2m-1}$ $\mathcal{E}_{2m}$ $\mathcal{O}_{2m}$ $\mathcal{E}_{2m+1}$

Trace Sets

Structural Property:

LSB replacement moves pairs between trace subsets,
but the trace sets are fixed.



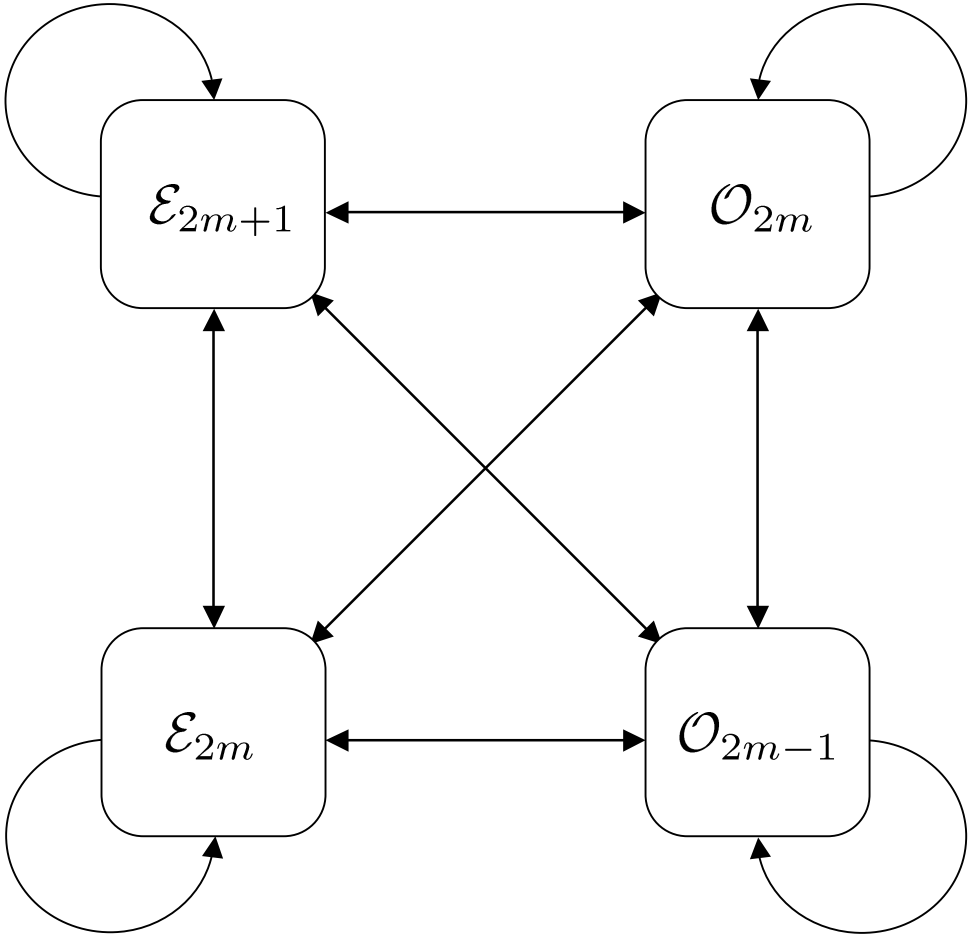
$\mathcal{P}$

Trace sets: $\mathcal{C}_{-127}$ $\cdots$ $\mathcal{C}_{-1}$ $\mathcal{C}_0$ $\mathcal{C}_1$ $\cdots$ $\mathcal{C}_m$ $\cdots$ $\mathcal{C}_{127}$

Trace subsets: $\mathcal{O}_{2m-1}$ $\mathcal{E}_{2m}$ $\mathcal{O}_{2m}$ $\mathcal{E}_{2m+1}$

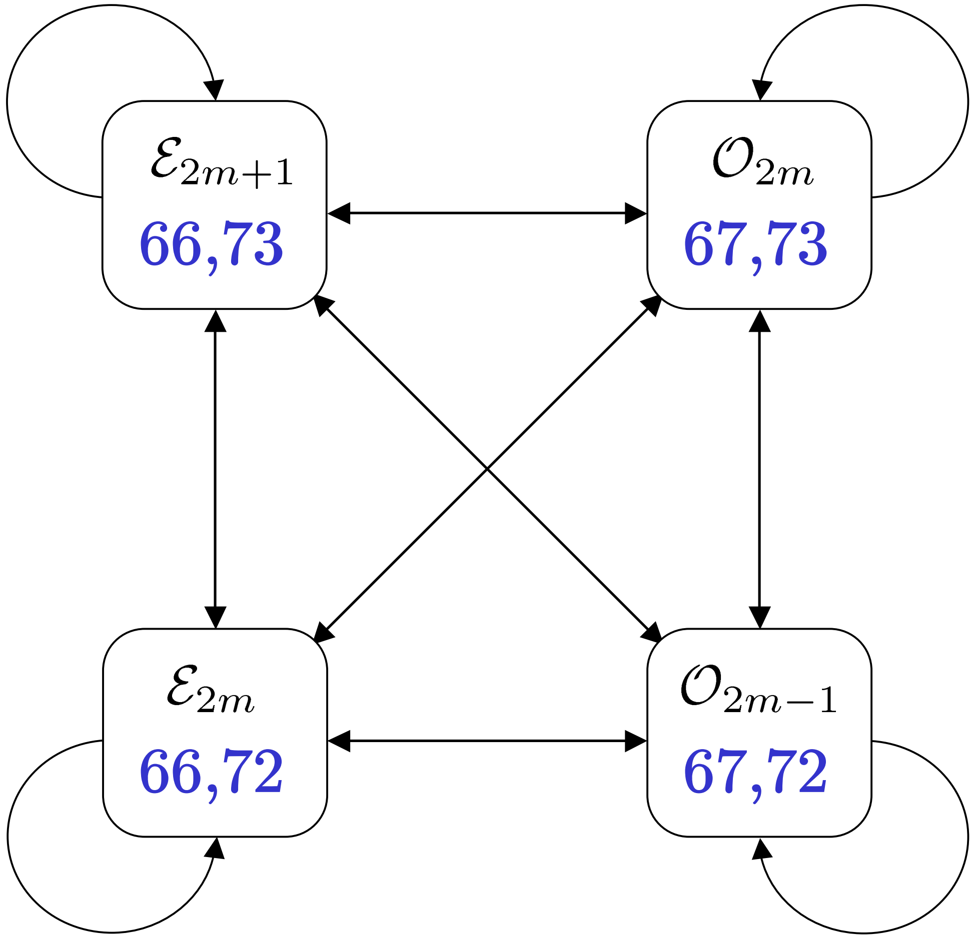
The Transition Process

Fix m. How are the trace subsets of $\mathcal{C}_m$ affected by LSB operations?



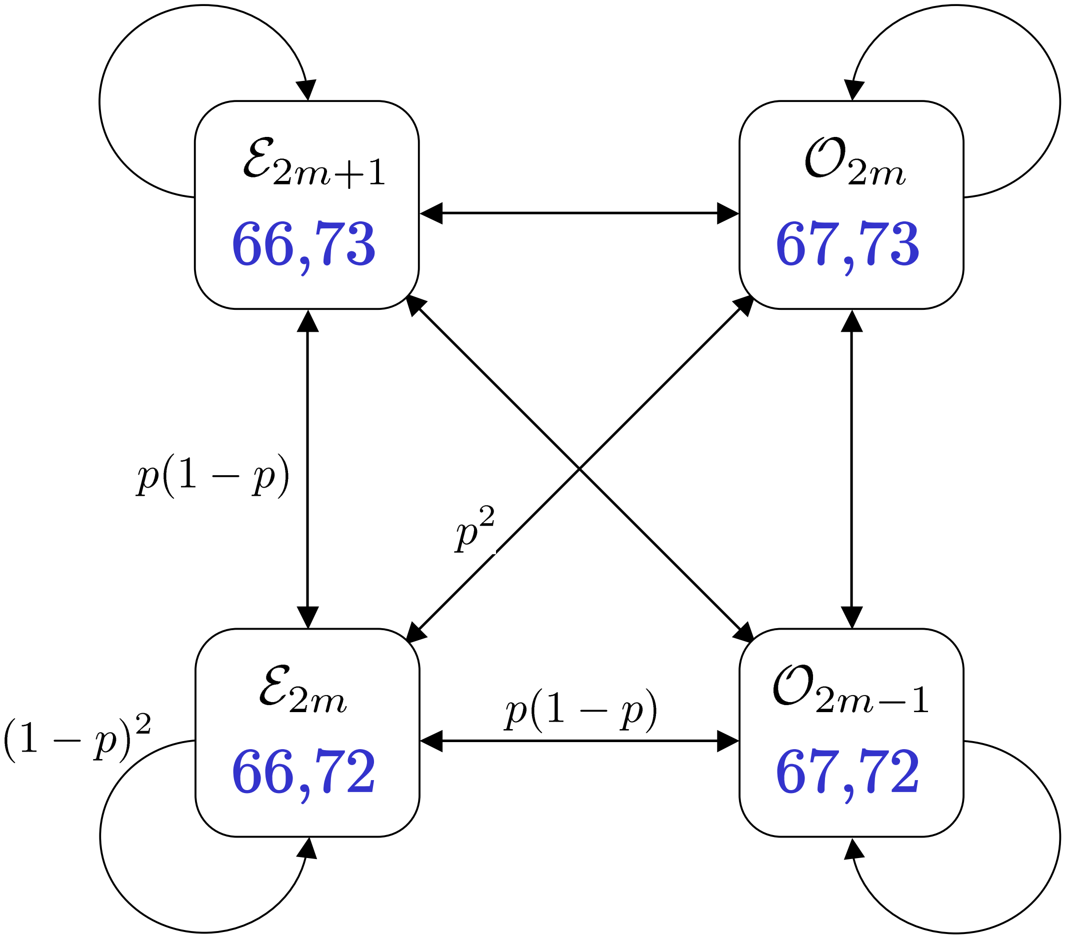
The Transition Process

Example: some pairs for m=3



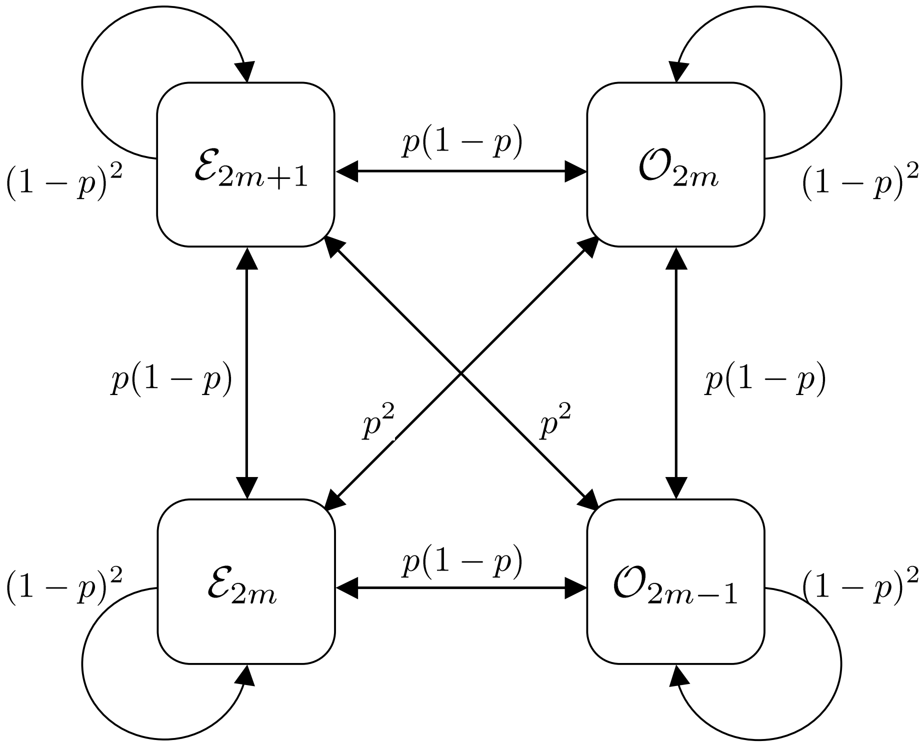
The Transition Process

When LSBs are flipped at random, with probability p



The Transition Process

Fix a cover of size N. Embed a random message of length $2pN$.



Define

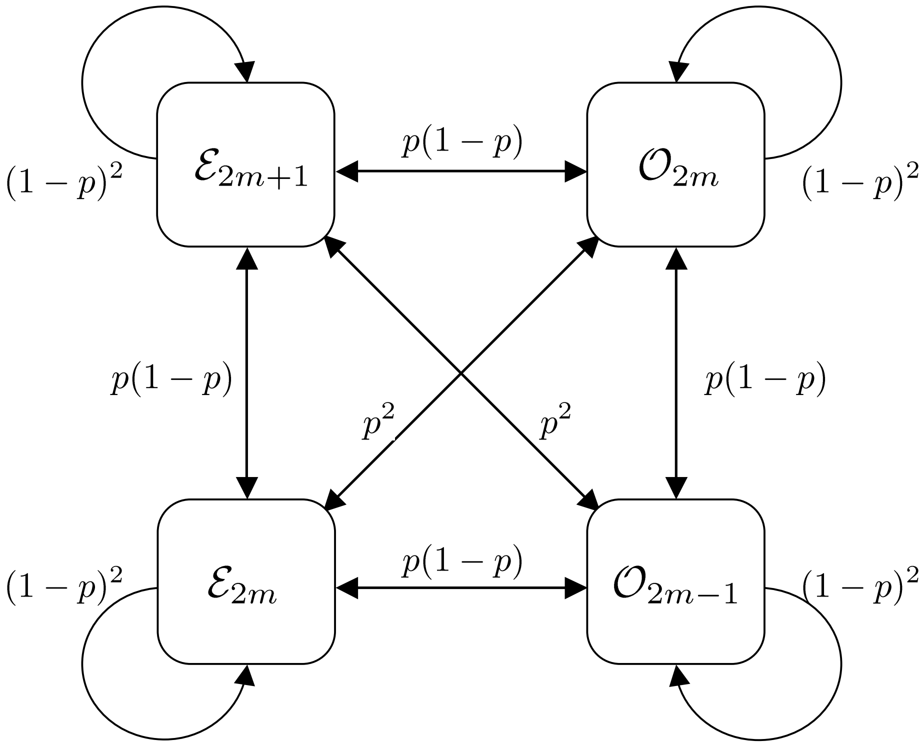
$$
\begin{aligned}
e_m &= \text{\#pairs in } \mathcal{E}_m \text{ in cover}\\
o_m &= \text{\#pairs in } \mathcal{O}_m \text{ in cover}\\
e'_m &= \text{\#pairs in } \mathcal{E}_m \text{ after embedding}\\
o'_m &= \text{\#pairs in } \mathcal{O}_m \text{ after embedding}
\end{aligned}
$$

Then

$$e'_{2m} = (1-p)^2 e_{2m} + p(1-p)o_{2m-1} + p(1-p)e_{2m+1} + p^2 o_{2m}.$$

The Transition Process

Fix a cover of size N. Embed a random message of length $2pN$.



Define

$$\begin{aligned} e_m &= \text{\#pairs in } \mathcal{E}_m \text{ in cover} \\ o_m &= \text{\#pairs in } \mathcal{O}_m \text{ in cover} \\ e'_m &= \text{\#pairs in } \mathcal{E}_m \text{ after embedding} \\ o'_m &= \text{\#pairs in } \mathcal{O}_m \text{ after embedding} \end{aligned}$$

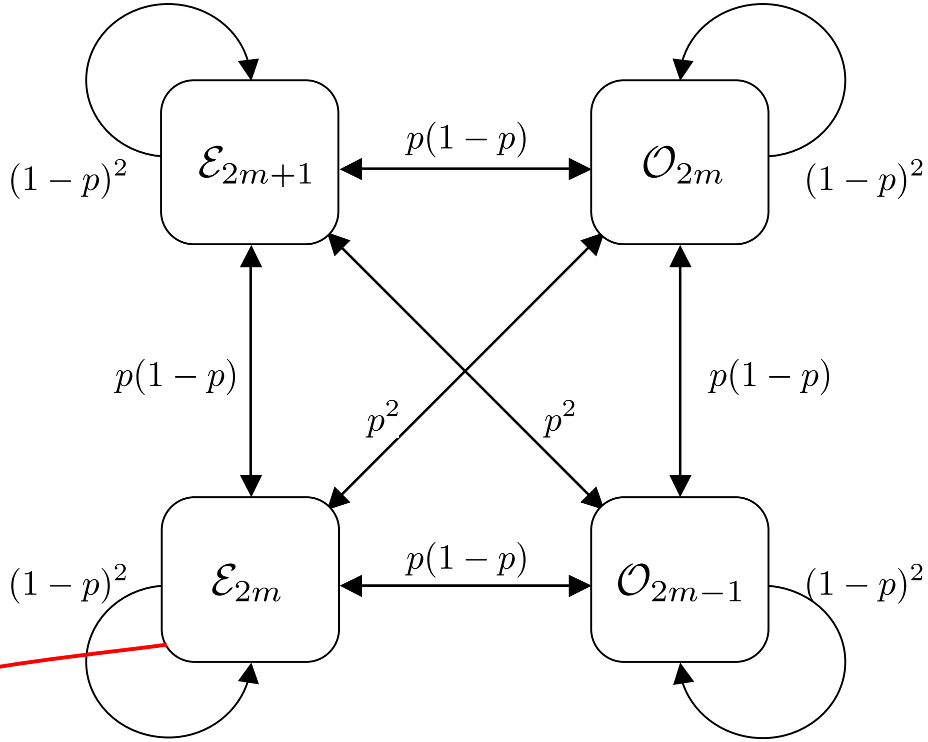
Then

$$e'_{2m} = (1-p)^2 e_{2m} + p(1-p) o_{2m-1} + p(1-p) e_{2m+1} + p^2 o_{2m}.$$

(really, the expectation of the random variable)

The Transition Process

Fix a cover of size N. Embed a random message of length $2pN$.



Define

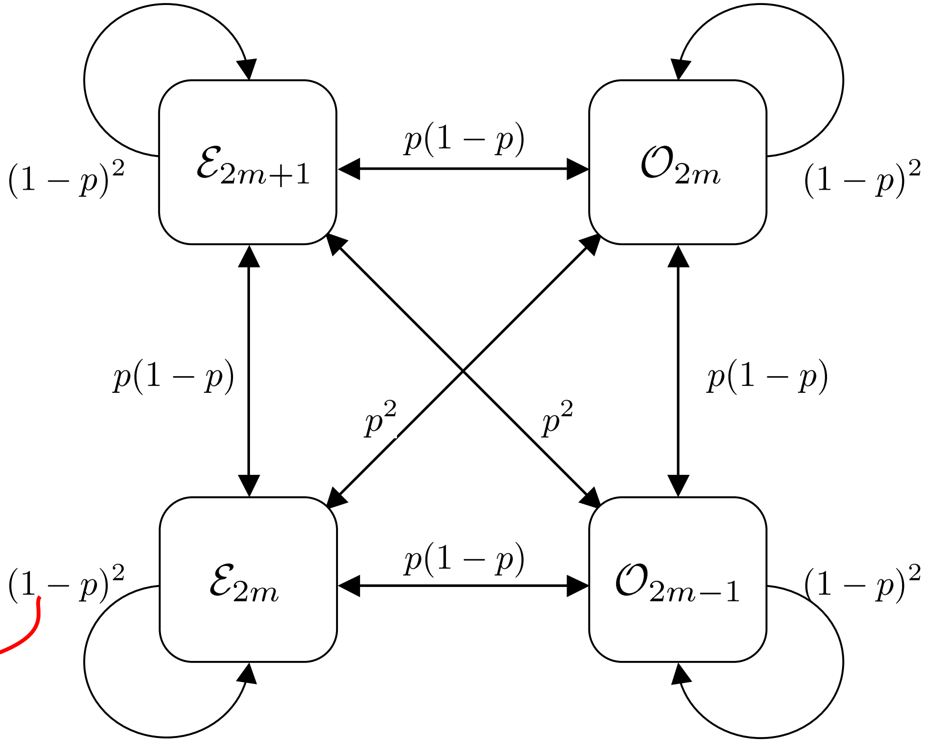
$$\begin{aligned} e_m &= \#\text{pairs in } \mathcal{E}_m \text{ in cover} \\ o_m &= \#\text{pairs in } \mathcal{O}_m \text{ in cover} \\ e'_m &= \#\text{pairs in } \mathcal{E}_m \text{ after embedding} \\ o'_m &= \#\text{pairs in } \mathcal{O}_m \text{ after embedding} \end{aligned}$$

Then

$$e'_{2m} = (1-p)^2 e_{2m} + p(1-p)o_{2m-1} + p(1-p)e_{2m+1} + p^2 o_{2m}.$$

The Transition Process

Fix a cover of size N. Embed a random message of length $2pN$.



Define

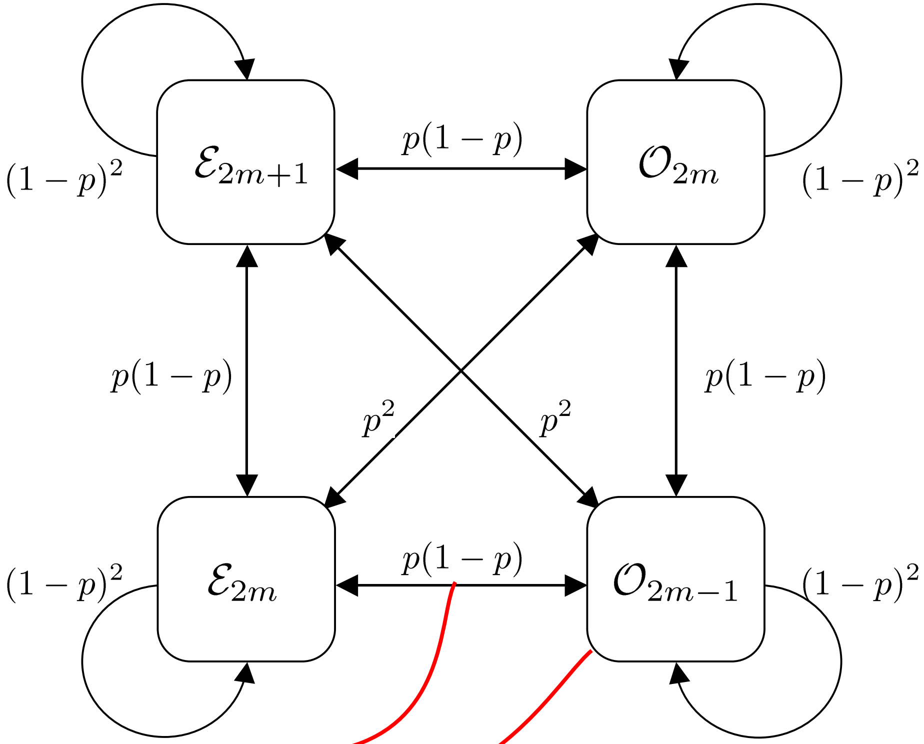
$$\begin{aligned} e_m &= \#\text{pairs in } \mathcal{E}_m \text{ in cover} \\ o_m &= \#\text{pairs in } \mathcal{O}_m \text{ in cover} \\ e'_m &= \#\text{pairs in } \mathcal{E}_m \text{ after embedding} \\ o'_m &= \#\text{pairs in } \mathcal{O}_m \text{ after embedding} \end{aligned}$$

Then

$$e'_{2m} = \textcolor{red}{(1-p)^2} e_{2m} + p(1-p) o_{2m-1} + p(1-p) e_{2m+1} + p^2 o_{2m}.$$

The Transition Process

Fix a cover of size N. Embed a random message of length $2pN$.



Define

$$\begin{aligned} e_m &= \#\text{pairs in } \mathcal{E}_m \text{ in cover} \\ o_m &= \#\text{pairs in } \mathcal{O}_m \text{ in cover} \\ e'_m &= \#\text{pairs in } \mathcal{E}_m \text{ after embedding} \\ o'_m &= \#\text{pairs in } \mathcal{O}_m \text{ after embedding} \end{aligned}$$

Then

$$e'_{2m} = (1-p)^2 e_{2m} + p(1-p)o_{2m-1} + p(1-p)e_{2m+1} + p^2 o_{2m}.$$

The Transition Process

We derive:

$$\underset{\overset{\uparrow}{\textbf{stego}}}{\begin{pmatrix} e'_{2m} \\ o'_{2m-1} \\ e'_{2m+1} \\ o'_{2m} \end{pmatrix}} = \begin{pmatrix} (1-p)^2 & p(1-p) & p(1-p) & p^2 \\ p(1-p) & (1-p)^2 & p^2 & p(1-p) \\ p(1-p) & p^2 & (1-p)^2 & p(1-p) \\ p^2 & p(1-p) & p(1-p) & (1-p)^2 \end{pmatrix} \underset{\overset{\uparrow}{\textbf{cover}}}{\begin{pmatrix} e_{2m} \\ o_{2m-1} \\ e_{2m+1} \\ o_{2m} \end{pmatrix}}$$

The Transition Process

We derive:

$$\underset{\substack{\uparrow \\ \textbf{stego}}}{\begin{pmatrix} e'_{2m} \\ o'_{2m-1} \\ e'_{2m+1} \\ o'_{2m} \end{pmatrix}} = \begin{pmatrix} (1-p)^2 & p(1-p) & p(1-p) & p^2 \\ p(1-p) & (1-p)^2 & p^2 & p(1-p) \\ p(1-p) & p^2 & (1-p)^2 & p(1-p) \\ p^2 & p(1-p) & p(1-p) & (1-p)^2 \end{pmatrix} \underset{\substack{\uparrow \\ \textbf{cover}}}{\begin{pmatrix} e_{2m} \\ o_{2m-1} \\ e_{2m+1} \\ o_{2m} \end{pmatrix}}$$

Inverting,

$$\underset{\substack{\uparrow \\ \textbf{cover}}}{\begin{pmatrix} \hat{e}_{2m} \\ \hat{o}_{2m-1} \\ \hat{e}_{2m+1} \\ \hat{o}_{2m} \end{pmatrix}} = \frac{1}{(1-2p)^2} \begin{pmatrix} (1-p)^2 & -p(1-p) & -p(1-p) & p^2 \\ -p(1-p) & (1-p)^2 & p^2 & -p(1-p) \\ -p(1-p) & p^2 & (1-p)^2 & -p(1-p) \\ p^2 & -p(1-p) & -p(1-p) & (1-p)^2 \end{pmatrix} \underset{\substack{\uparrow \\ \textbf{stego}}}{\begin{pmatrix} e'_{2m} \\ o'_{2m-1} \\ e'_{2m+1} \\ o'_{2m} \end{pmatrix}}$$

A Model for Covers

In continuous covers, we believe that

$$e_m \approx o_m$$

because the number of pairs differing by m should not be correlated with parity of the values.

Technical difficulty: provides no distinction between covers and stego images when m is even. So only consider the case of odd m*.*

Framework

1. Determine (expectation of) macroscopic properties of stego image, given cover and p
2. Invert: determine (estimate of) macroscopic properties of cover, given stego image and p
3. Form model for macroscopic properties of covers $e_{2m+1} \approx o_{2m+1}$
4. Given a suspect image, estimate p as whichever implies the best cover fit

Framework

1. Determine (expectation of) macroscopic properties of stego image, given cover and p
2. Invert: determine (estimate of) macroscopic properties of cover, given stego image and p
3. Form model for macroscopic properties of covers $e_{2m+1} \approx o_{2m+1}$
4. Given a suspect image, estimate p as whichever implies **the best cover fit**

Define error $\epsilon_m = \hat{e}_{2m+1} - \hat{o}_{2m+1}$ as a function of p

Minimize $|\sum \epsilon_m|$ or $\sum \epsilon_m^2$

Framework

1. Determine (expectation of) macroscopic properties of stego image, given cover and p
2. Invert: determine (estimate of) macroscopic properties of cover, given stego image and p
3. Form model for macroscopic properties of covers $e_{2m+1} \approx o_{2m+1}$
4. Given a suspect image, estimate p as whichever implies **the best cover fit**

Define error $\epsilon_m = \hat{e}_{2m+1} - \hat{o}_{2m+1}$ as a function of p

Minimize $|\sum \epsilon_m|$ or $\sum \epsilon_m^2$

Apart from some minor differences, leads to Dumitrescu's "Sample Pairs" estimator [IHW'02] a.k.a. "Couples"

Leads to "Least Squares Sample Pairs" estimator [Lu et al, IHW'04]

“Triples” Analysis

Now the extension to larger sample groups seems relatively straightforward.

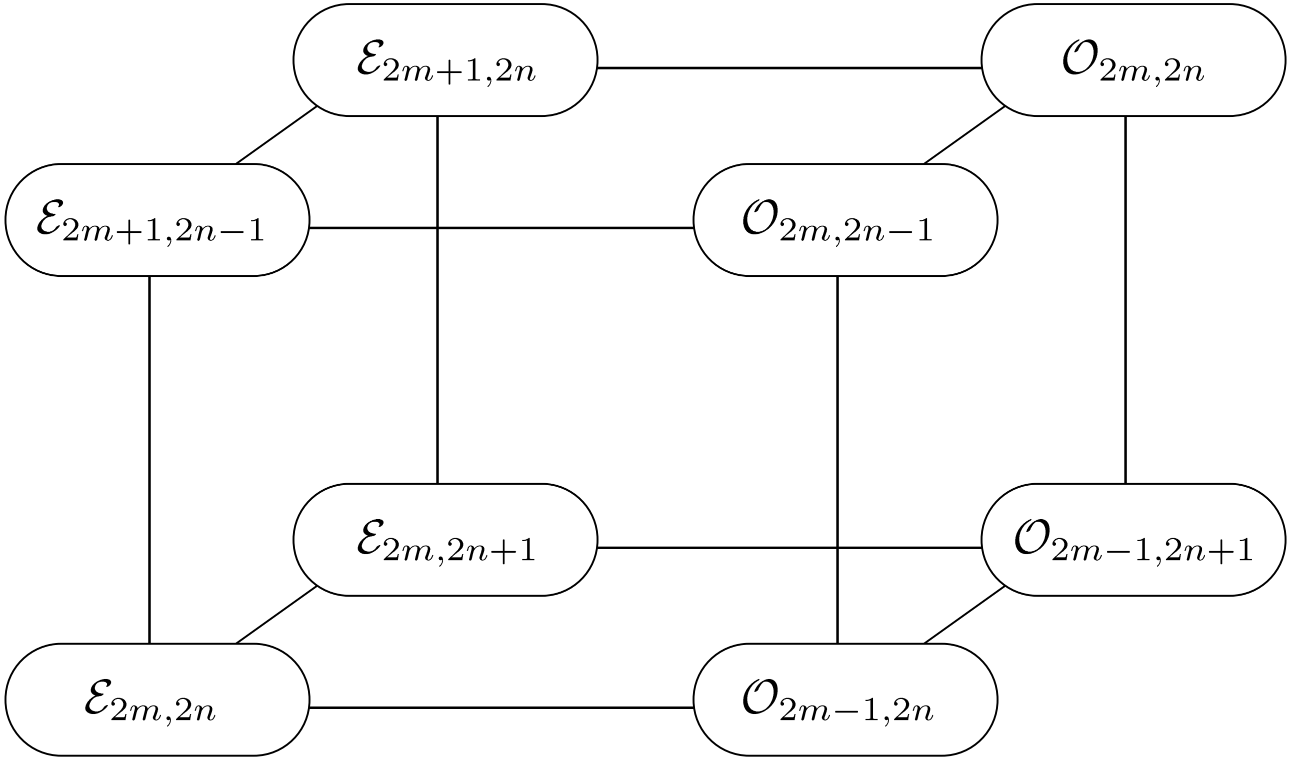
Definitions (sets of triples)

$\mathcal{T}$	all triples (x, y, z) used in the analysis e.g. all adjacent triples
$\mathcal{C}_{m,n}$	values divide by two to give a triple of the form $(u, u + m, u + m + n)$
$\mathcal{E}_{m,n}$	triples of the form $(x, x + m, x + m + n)$ where x is even
$\mathcal{O}_{m,n}$	triples of the form $(x, x + m, x + m + n)$ where x is odd

Each trace set $\mathcal{C}_{m,n}$ is fixed by LSB operations, and decomposes into 8 trace subsets which are affected by LSB operations.

The “Triples” Transition Process

Trace subsets of $\mathcal{C}_{m,n}$:



A triple moves along i edges with probability $p^i(1-p)^{3-i}$

The “Triples” Transition Process

We derive

$$\underset{\substack{\uparrow \\ \textbf{stego}}}{\begin{pmatrix} e'_{2m,2n} \\ \vdots \\ o'_{2m,2n} \end{pmatrix}} = T_3 \underset{\substack{\uparrow \\ \textbf{cover}}}{\begin{pmatrix} e_{2m,2n} \\ \vdots \\ o_{2m,2n} \end{pmatrix}}$$

where

$$T_3 = \begin{pmatrix}
(1-p)^3 & p(1-p)^2 & p(1-p)^2 & p^2(1-p) & p(1-p)^2 & p^2(1-p) & p^2(1-p) & p^3 \\
p(1-p)^2 & (1-p)^3 & p^2(1-p) & p(1-p)^2 & p^2(1-p) & p(1-p)^2 & p^3 & p^2(1-p) \\
p(1-p)^2 & p^2(1-p) & (1-p)^3 & p(1-p)^2 & p^2(1-p) & p^3 & p(1-p)^2 & p^2(1-p) \\
p^2(1-p) & p(1-p)^2 & p(1-p)^2 & (1-p)^3 & p^3 & p^2(1-p) & p^2(1-p) & p(1-p)^2 \\
p(1-p)^2 & p^2(1-p) & p^2(1-p) & p^3 & (1-p)^3 & p(1-p)^2 & p(1-p)^2 & p^2(1-p) \\
p^2(1-p) & p(1-p)^2 & p^3 & p^2(1-p) & p(1-p)^2 & (1-p)^3 & p^2(1-p) & p(1-p)^2 \\
p^2(1-p) & p^3 & p(1-p)^2 & p^2(1-p) & p(1-p)^2 & p^2(1-p) & (1-p)^3 & p(1-p)^2 \\
p^3 & p^2(1-p) & p^2(1-p) & p(1-p)^2 & p^2(1-p) & p(1-p)^2 & p(1-p)^2 & (1-p)^3
\end{pmatrix}$$

T_3 is invertible as long as $p \neq 0.5$.

Cover Image Assumptions

In the case of pairs of samples, the cover image assumption was

$$e_m \approx o_m$$

(which only provides discrimination between cover and stego images for odd m).

In the case of triples of samples, we have a number of plausible assumptions (which we omit discussion of here). The most useful is

$$e_{m,n} \approx o_{m,n}$$

(glossing over some other details).

Applying the Framework

1. Determine (expectation of) macroscopic properties of stego image, given cover and p
2. Invert: determine (estimate of) macroscopic properties of cover, given stego image and p
3. Form model for macroscopic properties of covers
4. Given a suspect image, estimate p as whichever implies **the best cover fit**

Given p, the estimated deviations from the cover assumptions include:

$$\epsilon_{m,n} = \hat{e}_{2m+1,2n+1} - \hat{o}_{2m+1,2n+1}$$

The total square error is

$$S(p) = \sum_{m,n} \epsilon_{m,n}^2$$

Find minimum point to estimate p.

Experimental Results

Compared the methods of RS, Sample Pairs, Least Squares SP, Triples

- as an estimator of p
- as a discriminator between covers and stego images

Simulated steganography and measured performance in large (3000-20000) sets of (colour) cover images of various types:

- bitmaps (scanned images);
- decompressed JPEGs (some originally scanned, some from digital cameras).

(it is necessary to repeat tests with different types of covers, as the results can be very different)

Experimental Results

Compared the methods of RS, Sample Pairs, Least Squares SP, Triples

- **as an estimator of p**
- as a discriminator between covers and stego images

Summary:

- In the case of uncompressed bitmap covers, Triples estimate has 10-20% smaller errors.
- In the case of covers with compression artefacts, Triples estimate has up to **10 times** smaller errors.

Experimental Results

Compared the methods of RS, Sample Pairs, Least Squares SP, Triples

- as an estimator of p
- **as a discriminator between covers and stego images**

Hypothesis test to distinguish the two cases:

$$\begin{aligned} H_0: &\quad p = 0 \\ H_1: &\quad p = p_1 > 0 \end{aligned}$$

Reliability as a Discriminator

Moral of [Ker, IHW'04]:

Estimators for the hidden message length may not be optimal for the discrimination problem.

It can be better to use a discriminating statistic which simply measures how well the cover assumptions have been met.

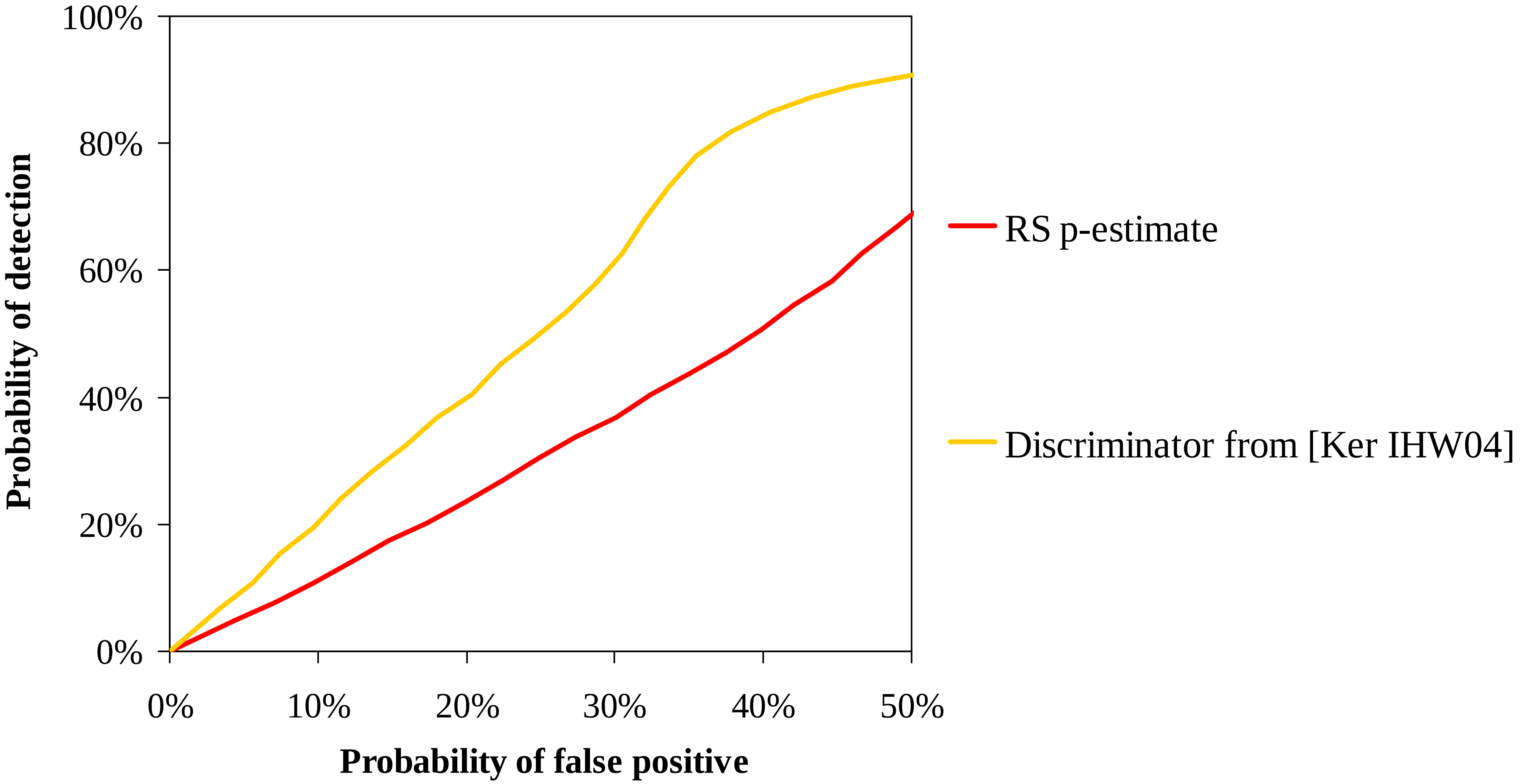
Recall $S(p) = \sum_{m,n} \epsilon_{m,n}^2$

The measure $S(0)$, i.e. observed deviation from the cover model, is **not** a good discriminator.

The measure $S(0)/S(\hat{p})$ is an excellent discriminator, measuring **how certain we are** that p is not zero.

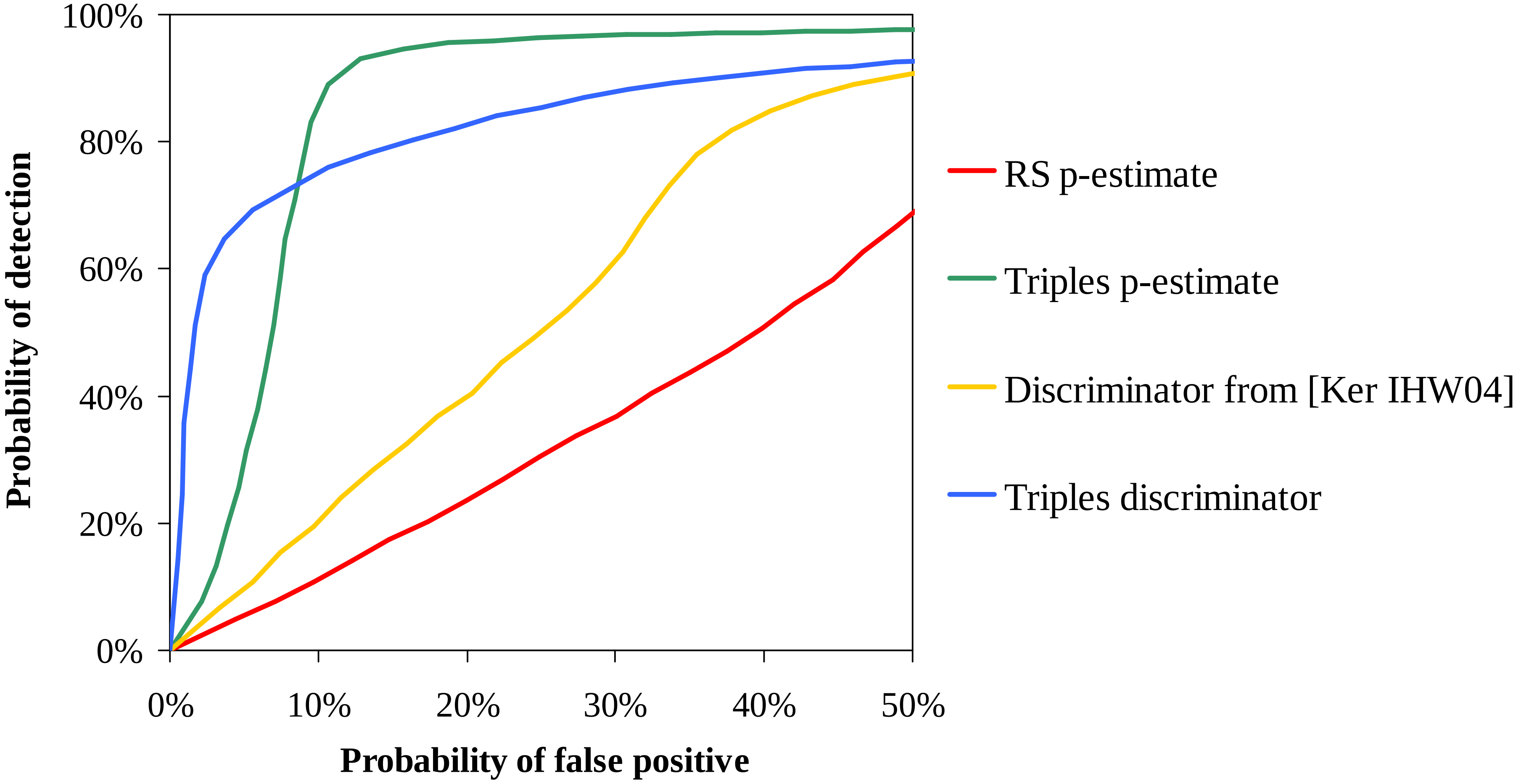
Reliability as a Discriminator

ROC curves from 3000 moderately-compressed JPEG covers. Data embedded at 0.02 bits per cover (2% of max.)



Reliability as a Discriminator

ROC curves from 3000 moderately-compressed JPEG covers. Data embedded at 0.02 bits per cover (2% of max.)



Reliability as a Discriminator

The lowest embedding rate (as percentage of maximum 1 bit per cover byte) at which less than 50% false negatives is observed with 5% false positives.

	3000 never-compressed bitmaps	10000 decompressed JPEGs
RS p-estimate		
Sample Pairs p-estimate		
Least Squares SP p-estimate		
Triples p-estimate		

Reliability as a Discriminator

The lowest embedding rate (as percentage of maximum 1 bit per cover byte) at which less than 50% false negatives is observed with 5% false positives.

	3000 never-compressed bitmaps	10000 decompressed JPEGs
RS p-estimate	5.4	8
Sample Pairs p-estimate	5.2	5.8
Least Squares SP p-estimate	6.2	2.4
Triples p-estimate	4.2	0.5

Reliability as a Discriminator

The lowest embedding rate (as percentage of maximum 1 bit per cover byte) at which less than 50% false negatives is observed with 5% false positives.

	3000 never-compressed bitmaps	10000 decompressed JPEGs
RS p-estimate	5.4	8
Sample Pairs p-estimate	5.2	5.8
Least Squares SP p-estimate	6.2	2.4
Triples p-estimate	4.2	0.5
Discriminator from [Ker IHW04]	2.8	2
Triples discriminator	5.4	0.3

Conclusions

- We have extended the analysis of "structural" properties of LSB embedding from pairs to triplets.

  Have extended to arbitrary groups, in the written paper, but also encountered some difficulties with the cover assumptions, which leaves optimal implementation incomplete.

- The detector is expressed in a new paradigm, based on inverting the effects of steganography, if the size of hidden data is known, and matching a cover model.

  This framework can encompass many – all? – other structural LSB steganography detectors.

- There is experimental evidence of improved performance, particularly in the case when the cover images were anomalous.

End

adk@comlab.ox.ac.uk