Steganalysis using Logistic Regression

Ivans Lubenko & Andrew Ker

lubenko@comlab.ox.ac.uk adk@comlab.ox.ac.uk

Oxford University Computing Laboratory

SPIE/IS&T Electronic Imaging, San Francisco

25 January 2011

Steganalysis using Logistic Regression

Outline

- Classification in steganalysis
- Logistic regression
- Experimental methodology
- Experimental results

Classification in steganalysis

Steganalysis is a binary classification problem:

- Innocent cover objects

  vs
- Stego objects.

This is a gross approximation unless

1. the embedding method is known,
2. the payload size is known,
3. the cover source is known.

For quite practical reasons, steganalysis should probably be multi-class.

Classification in steganalysis

A steganalysis method is broadly:

- **Analytical – some property of covers is distorted in an understood way.**

  Spatial-domain bit replacement [Dumitrescu, 2003; Ker, 2005-7]
  LSB replacement in JPEGs [Kodovský, 2010]

Generally only for flawed embedding.

- **Empirical – machine learning algorithms applied to vectors of features.**

  Support vector machines and noise features [Farid, 2002]
  Support vector machines and JPEG histograms [Shi, 2006; Pevný, 2007]
  Support vector machines and pixel difference co-occurrence [Pevný, 2010]

Almost all schemes in the literature use SVMs. Is there an alternative?

Logistic regression

Not a classifier! Attempts to model class probabilities.

Predict **log-odds** as linear function of input data. Binary case:

$$\mathrm{P}(C_1|x) = \frac{1}{1 + \exp(-(w^T x + b))}$$

with coefficients $w,\ b$ learned by maximum likelihood.

$$\log L = \sum_{j=1}^{n} y_j \log \mathrm{P}(C_1|x_j) + (1 - y_j) \log(1 - \mathrm{P}(C_1|x_j)) - \lambda\|w\|.$$

regularization

Advantage 1: predicts class probabilities (though often used as a simple classifier).

Can represent problem in dual space, and use the kernel trick.

Multinomial logistic regression

Predict log-odds of all classes simultaneously:

$$\mathrm{P}(C_i|x) = \frac{\exp(w_i^T x)}{\sum_j \exp(w_j^T x)}$$

with matrix of w_j's learned by maximum likelihood.

Advantage 2:

- *Time complexity for k classes is $O(k)$,*
- *Time complexity for SVMs simulating multi-class by all pairs is $O(k^2)$.*

Experimental methodology

Want to compare SVM and LR steganalysis detectors in terms of

- accuracy, and
- speed.

Extent of comparison

- Images in three sets: 'Camera', 'BOSS', 'BOWS'.
- Spatial-domain, total payload 0.5 bits per pixel:
  - LSB matching,
  - LSB replacement,
  - 2LSB replacement (2 bits per used pixel),
  - Mod-5 matching ($\log_2 5$ bits per used pixel),

  plus covers, makes up to 5 classes.
- SPAM features (686 dimensional).

Experimental methodology

Implementation

SVM	libSVM (optimized C++)
LR	'minfunc' (MATLAB)
Smooth SVM	'minfunc' (MATLAB)

… with linear or Gaussian kernel.

Training regime

- Split image set into training & testing.
- Hyperparameter optimization by cross-validation on training set.
- Final training on entire training set.
- Test on testing set.
- Compare performance differences by Student t-test.

Results: kernelized classifiers

Image set	Classification Problem	Classification accuracy		Difference
		SVM	Logistic Regression	
Camera	Binary cover v LSBM	96.20%	95.85%	Insignificant ($p > 0.05$)
	Binary cover v Mod5	97.55%	97.45%	Insignificant ($p > 0.05$)
	Binary cover v LSBR	97.15%	97.45%	Insignificant ($p > 0.05$)
	Binary cover v 2LSB	98.25%	98.45%	Insignificant ($p > 0.05$)
Camera	5 class	83.01%	80.05%	Insignificant ($p > 0.05$)
BOSS	5 class	85.05%	82.30%	Somewhat significant ($p \approx 0.01$)
BOWS	5 class	95.95%	96.32%	Insignificant ($p > 0.05$)

Results: kernelized classifiers

Image set	Classification Problem	Classification accuracy		Difference
		Smooth SVM	Logistic Regression	
Camera	Binary cover v LSBM	96.15%	95.85%	Insignificant ($p > 0.05$)
	Binary cover v Mod5	97.70%	97.45%	Insignificant ($p > 0.05$)
	Binary cover v LSBR	96.90%	97.45%	Insignificant ($p > 0.05$)
	Binary cover v 2LSB	98.35%	98.45%	Insignificant ($p > 0.05$)
Camera	5 class	80.89%	80.05%	Insignificant ($p > 0.05$)
BOSS	5 class	83.81%	82.30%	Insignificant ($p > 0.05$)
BOWS	5 class	96.59%	96.32%	Insignificant ($p > 0.05$)

Results: kernelized classifiers

Image set	Classification Problem	Classification accuracy		Final training time	
		Smooth SVM	Logistic Regression	Smooth SVM	Logistic Regression
Camera	Binary cover v LSBM	96.15%	95.85%	106s	145s
	Binary cover v Mod5	97.70%	97.45%	95s	160s
	Binary cover v LSBR	96.90%	97.45%	101s	113s
	Binary cover v 2LSB	98.35%	98.45%	85s	127s
Camera	5 class	80.89%	80.05%	2504s	1394s
BOSS	5 class	83.81%	82.30%	43003s	23446s
BOWS	5 class	96.59%	96.32%	84198s	42311s

Conclusions

- Logistic regression is a possible alternative to SVMs in steganalysis, with two potential advantages:
  - *well adapted to multinomial case,*
  - *produces class probabilities.*

- The detection accuracy seems to be similar to SVMs.

- For binary classification, the speed is similar to an SVM implemented on the same minimizer.

- For multinomial classification, the speed is superior.
  - *But this excludes the SMO algorithm for SVMs.*
  - *Needs further work to examine SMO-type algorithm for LR.*