From Blind to Quantitative Steganalysis

Tomáš Pevný\textsuperscript{1}, Jessica Fridrich\textsuperscript{2}, Andrew D. Ker\textsuperscript{3}

\textsuperscript{1}GIPSA-Lab, INPG, France
\textsuperscript{2}Binghamton University, SUNY, USA
\textsuperscript{3}University of Oxford, UK

19th January 2009
Outline

1 Motivation

2 Methodology

3 Experiments
   - General results
   - Detailed results for Jsteg and nsF5
   - Comparison to previous art

4 Conclusion and Future directions
1 Motivation

2 Methodology

3 Experiments
   - General results
   - Detailed results for Jsteg and nsF5
   - Comparison to previous art

4 Conclusion and Future directions
Steganalysis

- *Steganalysis* detects presence of secret message.
- *Steganalyzer* is a binary detector (classifier).

Quantitative steganalysis

- *Quantitative steganalysis* estimates number of embedding changes (length of message).
- *Quantitative steganalyzer* is an estimator.
Advantages of Quantitative Steganalysis

- provide the steganalyst with further information (estimate of message length).
- useful for forensic analysis (message is encrypted).
- important in pooled steganalysis.\(^a\)
- allow a finer control of false positive and false negative rate in targeted blind steganalysis.
- alleviate problems with dependence of the steganalyzer on message length in the training set.\(^b\)

\(^b\) Cancelli et al., *A Comparative Study of ±1 Steganalyzers*, 2008.
Outline

1 Motivation

2 Methodology

3 Experiments
   - General results
   - Detailed results for Jsteg and nsF5
   - Comparison to previous art

4 Conclusion and Future directions
Methodology

Assumption

- Steganographic features used in blind steganalysis *react predictably* to the number of embedding changes.
- Identify relationship between *feature vector* and *change rate*

First two most significant components of merged features of nsF5 identified by Partial Least Square.
Quantitative Steganalysis by Regression

Problem

- We seek a function $\psi : \mathcal{X} \mapsto [0, 1]$ revealing relationship between location of feature vector and change rate ($\mathcal{X}$ is the feature space).

- Function $\psi$ is learned from a set of examples
  \[ \{(x_1, y_1), \ldots, (x_l, y_l)\}, \]
  \[ x_i \in \mathcal{X} \text{ features of stego image with change rate } y_i. \]

- Construction of a quantitative steganalyzer is a regression problem, for which many tools are available.

- This work utilizes
  - linear ordinary least-square regression,
  - support vector regression.
Advantages over Prior Art

Prior art
Quantitative steganalyzers are built from heuristic principles and always rely on full knowledge of embedding algorithm.

Advantages of proposed method

Cookie cutter approach:
1. Find feature set detecting the stego algorithm.
2. Create set of training examples \((x_i, y_i)\).
3. Use regression to learn dependence between features and change rate.

The knowledge of embedding mechanism is not needed.
Outline

1 Motivation

2 Methodology

3 Experiments
   - General results
   - Detailed results for Jsteg and nsF5
   - Comparison to previous art

4 Conclusion and Future directions
Experimental Settings

- Quantitative steganalyzers for 8 steganographic methods: JP Hide&Seek, Jsteg, MBS1, MMx, F5 with removed shrinkage (nsF5), OutGuess, Perturbed Quantization (PQ), and Steghide.

- Quantitative steganalyzers were constructed by
  - linear ordinary least-square regression (OLS)
  - support vector regression (SVR).

- Single-compressed JPEGs with quality factor 80, all created from 9163 raw images evenly divided into training/testing set.

- Random payload between zero and maximum for given image and algorithm was embedded into images.

- 274 “calibrated merged features” augmented by number of non-zero DCTs.
Outline

1 Motivation

2 Methodology

3 Experiments
   - General results
   - Detailed results for Jsteg and nsF5
   - Comparison to previous art

4 Conclusion and Future directions
Detection Accuracy of MB1 and MMx

Figure: Estimated versus true relative change rate of SVR quantitative steganalyzers of MB1 and MMx.
### Experimental Results

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>OLS MAE</th>
<th>OLS Bias</th>
<th>SVR MAE</th>
<th>SVR Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>JP Hide&amp;Seek</td>
<td>7.91 \cdot 10^{-03}</td>
<td>-1.70 \cdot 10^{-04}</td>
<td>5.24 \cdot 10^{-03}</td>
<td>2.41 \cdot 10^{-04}</td>
</tr>
<tr>
<td>Jsteg</td>
<td>8.38 \cdot 10^{-03}</td>
<td>-5.29 \cdot 10^{-04}</td>
<td>1.9 \cdot 10^{-03}</td>
<td>2.5 \cdot 10^{-04}</td>
</tr>
<tr>
<td>nsF5</td>
<td>8.39 \cdot 10^{-03}</td>
<td>-5.29 \cdot 10^{-04}</td>
<td>4.82 \cdot 10^{-03}</td>
<td>-2.51 \cdot 10^{-04}</td>
</tr>
<tr>
<td>MB1</td>
<td>9.07 \cdot 10^{-03}</td>
<td>3.86 \cdot 10^{-05}</td>
<td>6.63 \cdot 10^{-03}</td>
<td>-1.63 \cdot 10^{-04}</td>
</tr>
<tr>
<td>MMX</td>
<td>3.25 \cdot 10^{-03}</td>
<td>1.58 \cdot 10^{-04}</td>
<td>2.70 \cdot 10^{-03}</td>
<td>1.08 \cdot 10^{-04}</td>
</tr>
<tr>
<td>Steghide</td>
<td>3.23 \cdot 10^{-03}</td>
<td>2.60 \cdot 10^{-04}</td>
<td>2.04 \cdot 10^{-03}</td>
<td>1.80 \cdot 10^{-04}</td>
</tr>
<tr>
<td>PQ</td>
<td>5.69 \cdot 10^{-02}</td>
<td>-2.89 \cdot 10^{-03}</td>
<td>4.83 \cdot 10^{-02}</td>
<td>-3.78 \cdot 10^{-02}</td>
</tr>
<tr>
<td>OutGuess</td>
<td>2.53 \cdot 10^{-03}</td>
<td>1.51 \cdot 10^{-04}</td>
<td>2.48 \cdot 10^{-03}</td>
<td>3.67 \cdot 10^{-04}</td>
</tr>
</tbody>
</table>

**Table:** Median absolute error (MAE) and bias measured on testing images with random payload.
Outline

1 Motivation

2 Methodology

3 Experiments
   - General results
   - Detailed results for Jsteg and nsF5
   - Comparison to previous art

4 Conclusion and Future directions
Figure: Median absolute error (MAE) and bias of SVR based estimators of nsF5 and Jsteg on 21 different fixed embedding change rates.
Outline

1. Motivation
2. Methodology
3. Experiments
   - General results
   - Detailed results for Jsteg and nsF5
   - Comparison to previous art
4. Conclusion and Future directions

T. Pevný, J. Fridrich, A. Ker | From Blind to Quantitative Steganalysis
Comparison to Previous Art

**Figure:** Comparison of accuracy of SVR, Jpairs, and Weighted non-steganographic Borders attack (WB) at 21 different fixed embedding change rates on 4563 images from testing set.
Outline

1 Motivation

2 Methodology

3 Experiments
   - General results
   - Detailed results for Jsteg and nsF5
   - Comparison to previous art

4 Conclusion and Future directions
Conclusion

- A solid method to construct quantitative steganalyzer from features was presented.
- Regression is used to learn dependence between features for blind steganalysis and embedding change rate.
- Method was demonstrated on 8 JPEG stego-schemes.
- For Jsteg, accuracy is at least as good as best targeted attacks.
- Distributions of within image and between image error were estimated — same as of quantitative steganalyzers of LSB replacement.
Future Directions

- Combine existing LSB quant. steganalyzers to improve accuracy.
- Improve control of false positive/false negative rate in targeted blind steganalysis.
- Quantitative steganalysis of ±1, YASS?
<table>
<thead>
<tr>
<th>β</th>
<th>Shapiro-Wilk</th>
<th>Between $\Delta Q(Z_{cov})$</th>
<th>Within $\Delta Q(Z_{pos})$</th>
<th>Flips $\Delta Q(Z_{flip})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.63</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.025</td>
<td>3.23</td>
<td>1.52</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>3.02</td>
<td>1.91</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>0.125</td>
<td>2.79</td>
<td>2.57</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>2.87</td>
<td>3.25</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>0.375</td>
<td>3.69</td>
<td>3.56</td>
<td>0.87</td>
<td></td>
</tr>
</tbody>
</table>
## Within and Between Image Error of nsF5

<table>
<thead>
<tr>
<th>β</th>
<th>Shapiro-Wilk</th>
<th>Between</th>
<th>Within</th>
<th>Flips</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p &gt; 0.1</td>
<td>ΔQ(Z_{cov})</td>
<td>ΔQ(Z_{pos})</td>
<td>ΔQ(Z_{flip})</td>
</tr>
<tr>
<td>0</td>
<td>—</td>
<td>7.74</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.025</td>
<td>93.9%</td>
<td>6.99</td>
<td>2.81</td>
<td>0.29</td>
</tr>
<tr>
<td>0.05</td>
<td>93.9%</td>
<td>6.79</td>
<td>3.52</td>
<td>0.41</td>
</tr>
<tr>
<td>0.125</td>
<td>93.7%</td>
<td>6.93</td>
<td>4.78</td>
<td>0.62</td>
</tr>
<tr>
<td>0.25</td>
<td>94.2%</td>
<td>8.31</td>
<td>6.77</td>
<td>0.81</td>
</tr>
<tr>
<td>0.375</td>
<td>94.2%</td>
<td>10.63</td>
<td>8.47</td>
<td>0.91</td>
</tr>
</tbody>
</table>