Model Based Steganography with Precover

Tomáš Denemark and Jessica Fridrich, Department of ECE, SUNY Binghamton, NY, USA, {tdenema1,fridrich}@binghamton.edu

Abstract

It is widely recognized that steganography with side-information in the form of a precover at the sender enjoys significantly higher empirical security than other embedding schemes. Despite the success of side-informed steganography, current designs are purely heuristic and little has been done to develop the embedding rule from first principles. Building upon the recently proposed MiPOD steganography, in this paper we impose multivariate Gaussian model on acquisition noise and estimate its parameters from the available precover. The embedding is then designed to minimize the KL divergence between cover and stego distributions. In contrast to existing heuristic algorithms that modulate the embedding costs by 1–2|e|, where e *is the rounding error, in our model-based approach the sender should modulate the steganographic Fisher information, which is a loose equivalent of embedding costs, by (1–2|e|)^2. Experiments with uncompressed and JPEG images show promise of this theoretically well-founded approach.*

Introduction

Steganography is a privacy tool in which messages are embedded in inconspicuous cover objects to hide the very presence of the communicated secret. Digital media, such as images, video, and audio are particularly suitable cover sources because of their ubiquity and the fact that they contain random components, the acquisition noise. On the other hand, digital media files are extremely complex objects that are notoriously hard to describe with sufficiently accurate and estimable statistical models. This is the main reason for why current steganography in such empirical sources [3] lacks perfect security and heavily relies on heuristics, such as embedding "costs" and intuitive modulation factors. Similarly, practical steganalysis resorts to increasingly more complex high-dimensional descriptors (rich models) and advanced machine learning paradigms, including ensemble classifiers and deep learning.

Often, a digital media object is subjected to processing and/or format conversion prior to embedding the secret. The last step in the processing pipeline is typically quantization. In side-informed steganography with precover [21], the sender makes use of the unquantized cover values during embedding to hide data in a more secure manner. The first embedding scheme of this type described in the literature is the embedding-while-dithering [14] in which the secret message was embedded by perturbing the process of color quantization and dithering when converting a true-color image to a palette format. Perturbed quantization [15] started another direction in which rounding errors of DCT coefficients during JPEG compression were used to modify the embedding algorithm. This method has been advanced through a series of papers [23, 24, 29, 20], culminating with approaches based on advanced coding techniques with a high level of empirical security [19, 18, 6].

Side-information can have many other forms. Instead of one precover, the sender may have access to the acquisition oracle (a camera) and take multiple images of the same scene. These multiple exposures can be used to estimate the acquisition noise and also incorporated during embedding. This research direction has been developed to a lesser degree compared to steganography with precover most likely due to the difficulty of acquiring the required imagery and modeling the differences between acquisitions. In a series of papers [10, 12, 11], Franz et al. proposed a method in which multiple scans of the same printed image on a flat-bed scanner were used to estimate the model of the acquisition noise at every pixel. This requires acquiring a potentially large number of scans, which makes this approach rather labor intensive. Moreover, differences in the movement of the scanner head between individual scans lead to slight spatial misalignment that complicates using this type of side-information properly. Recently, the authors of [7] showed how multiple JPEG images of the same scene can be used to infer the preferred direction of embedding changes. By working with quantized DCT coefficients instead of pixels, the embedding is less sensitive to small differences between multiple acquisitions.

Despite the success of side-informed schemes, there appears to be an alarming lack of theoretical analysis that would either justify the heuristics or suggest a well-founded (and hopefully more powerful) approach. In [13], the author has shown that the precover compensates for the lack of the cover model. In particular, for a Gaussian model of acquisition noise, precover-informed rounding is more secure than embedding designed to preserve the cover model estimated from the precover image assuming the cover is "sufficiently non-stationary." Another direction worth mentioning in this context is the bottom-up model-based approach recently proposed by Bas [2]. The author showed that a high-capacity steganographic scheme with a rather low empirical detectability can be constructed when the process of digitally developing a RAW sensor capture is sufficiently simplified. The impact of embedding is masked as an increased level of photonic noise, e.g., due to a higher ISO setting. It will likely be rather difficult, however, to extend this approach to realistic processing pipelines.

Inspired by the success of the multivariate Gaussian model in steganography for digital images [25, 17, 26], in this paper we adopt the same model for the precover and then derive the embedding rule to minimize the KL divergence between cover and stego distributions. The side-

information is used to estimate the parameters of the acquisition noise and the noise-free scene. In the next section, we review current state of the art in heuristic side-informed steganography with precover. In the following section, we introduce a formal model of image acquisition. In Section "Side-informed steganography with MVG acquisition noise", we describe the proposed model-based embedding method, which is related to heuristic approaches in Section "Connection to heuristic schemes." The main bulk of results from experiments on images represented in the spatial and JPEG domain appear in Section "Experiments." In the subsequent section, we investigate whether the public part of the selection channel, the content adaptivity, can be incorporated in selection-channel-aware variants of steganalysis features to improve detection of side-informed schemes. The paper is then closed with Conclusions.

The following notation is adopted for technical arguments. Matrices and vectors will be typeset in boldface, while capital letters are reserved for random variables with the corresponding lower case symbols used for their realizations. In this paper, we only work with grayscale cover images. Precover values will be denoted with $x_{ij} \in \mathbb{R}$, while cover and stego values will be integer arrays c_{ij} and s_{ij}, $1 \le i \le n_1$, $1 \le j \le n_2$, respectively. The symbols $[x]$, $\lceil x \rceil$, and $\lfloor x \rfloor$ are used for rounding and rounding up and down the value of x. By $\mathcal{N}(\mu,\sigma^2)$, we understand Gaussian distribution with mean μ and variance σ^2. The complementary cumulative distribution function of a standard normal variable (the tail probability) will be denoted $Q(x) = \int_x^\infty (2\pi)^{-1/2} \exp\left(-z^2/2\right) \mathrm{d}z$. Finally, we say that $f(x) \approx g(x)$ when $\lim_{x\to\infty} f(x)/g(x) = 1$.

Prior art in side-informed steganography with precover

All modern steganographic schemes, including those that use side-information, are implemented within the paradigm of distortion minimization. First, each cover element c_{ij} is assigned a "cost" ρ_{ij} that measures the impact on detectability should that element be modified during embedding. The payload is then embedded while minimizing the sum of costs of all changed cover elements, $\sum_{c_{ij}\neq s_{ij}} \rho_{ij}$. A steganographic scheme that embeds with the minimal expected cost changes each cover element with probability

$$\beta_{ij} = \frac{\exp(-\lambda\rho_{ij})}{1+\exp(-\lambda\rho_{ij})}, \tag{1}$$

if the embedding operation is constrained to be binary, and

$$\beta_{ij} = \frac{\exp(-\lambda\rho_{ij})}{1+2\exp(-\lambda\rho_{ij})}, \tag{2}$$

for a ternary scheme with equal costs of changing c_{ij} to $c_{ij} \pm 1$. Syndrome-trellis codes [8] can be used to build practical embedding schemes that operate near the rate–distortion bound.

For steganography designed to minimize costs (embedding distortion), a popular heuristic to incorporate a precover value x_{ij} during embedding is to modulate the costs based on the rounding error $e_{ij} = c_{ij} - x_{ij}$, $-1/2 \le e_{ij} \le 1/2$ [23, 29, 20, 18, 19, 6, 24]. A binary embedding scheme modulates the cost of changing $c_{ij} = [x_{ij}]$ to $[x_{ij}] + \mathrm{sign}(e_{ij})$ by $1-2|e_{ij}|$, while prohibiting the change to $[x_{ij}] - \mathrm{sign}(e_{ij})$:

$$\rho'_{ij}(\mathrm{sign}(e_{ij})) = (1-2|e_{ij}|)\rho_{ij} \tag{3}$$

$$\rho'_{ij}(-\mathrm{sign}(e_{ij})) = \Omega, \tag{4}$$

where $\rho_{ij}(u)$ is the cost of modifying the cover value by $u \in \{-1,1\}$, ρ_{ij} are costs of some additive embedding scheme, and Ω is a large constant. This modulation can be justified heuristically because when $|e_{ij}| \approx 1/2$, a small perturbation of x_{ij} could cause c_{ij} to be rounded to the other side. Such coefficients are thus assigned a proportionally smaller cost because $1-2|e_{ij}| \approx 0$. On the other hand, the costs are unchanged when $e_{ij} \approx 0$, as it takes a larger perturbation of the precover to change the rounded value.

A ternary version of this embedding strategy [6] allows modifications both ways with costs:

$$\rho'_{ij}(\mathrm{sign}(e_{ij})) = (1-2|e_{ij}|)\rho_{ij} \tag{5}$$

$$\rho'_{ij}(-\mathrm{sign}(e_{ij})) = \rho_{ij}. \tag{6}$$

Some embedding schemes do not use costs and, instead, minimize statistical detectability. In MiPOD [25], the embedding probabilities β_{ij} are derived from their impact on the cover multivariate Gaussian model by solving the following equation for each pixel ij:

$$\beta_{ij} I_{ij} = \lambda \ln \frac{1-2\beta_{ij}}{\beta_{ij}}, \tag{7}$$

where $I_{ij} = 2/\hat{\sigma}_{ij}^4$ is the Fisher information with $\hat{\sigma}_{ij}^2$ an estimated variance of the acquisition noise at pixel ij, and λ is a Lagrange multiplier determined by the payload size. To incorporate the side-information, the sender first converts the embedding probabilities into costs and then modulates them as in (3) or (5). This can be done by reversing the formula for optimal embedding probabilities for ternary cost-based schemes (2):

$$\rho_{ij} = \ln\left(1/\beta_{ij} - 2\right). \tag{8}$$

When reversing (2) λ can be set to 1 because multiplying costs by a positive scalar does not change the embedding scheme.

Modeling acquisition

An image $\mathbf{x}$ acquired using an imaging sensor has two components – the true scene $\mathbf{t}$ and acquisition imperfections (noise) $\mathbf{n}$:

$$\mathbf{x} = \mathbf{t} + \mathbf{n}(\mathbf{t}, \boldsymbol{\theta}). \tag{9}$$

While the scene $\mathbf{t}$ is deterministic, the acquisition noise $\mathbf{n}$ is stochastic in nature. It depends on $\mathbf{t}$ because, for example, the variance of the photonic (shot) noise linearly

depends on light intensity (the so-called heteroscedastic noise model [9]). Other random components of $\mathbf{n}$, including readout and electronic noise depend on the particular sensor. We exclude from $\mathbf{n}$ imperfections that are consistent from picture to picture (of the same scene under the same conditions and camera settings), which include the photo-response non-uniformity and dark current. Demosaicking, color correction, and additional filtering applied to the acquired image either in the camera or during post-production introduce dependencies into spatially neighboring samples of $\mathbf{n}$, turning it into a random field parametrized by $\boldsymbol{\theta}$, a vector encompassing the properties of the imaging hardware, camera settings, as well as the processing pipeline.

Fundamentally, the steganographic capacity of $\mathbf{x}$ is the entropy $H(\mathbf{n})$. However, embedding a payload of this size undetectably is generally possible only in very special cases when the processing pipeline is drastically simplified. Recently, Bas [2] has showed that RAW images acquired using a monochrome sensor (a sensor not equipped with a color filter array) with no spatial filtering applied to them can carry a rather large payload with a very low level of empirical detectability. The method is called "steganography by cover source switching" because the stego image statistically resembles an image acquired with a higher sensor gain setting (ISO). In general, however, constructing a steganographic method capable of embedding $H(\mathbf{n})$ bits is likely infeasible in virtually all practical situations because of the daunting complexity and non-stationarity of the random field $\mathbf{n}$. More seriously, many elements of the processing pipeline are unknown to the sender as most digital camera manufacturers use proprietary demosaicking algorithms with local content-dependent rules for interpolating the missing colors as well as proprietary content-adaptive algorithms for denoising and sharpening.

Side-informed steganography with MVG acquisition noise

It is intuitively clear that incorporating even partial information about the noise-free scene $\mathbf{t}$ and the statistical properties of $\mathbf{n}$ at the sender should improve security. Indeed, the sender has a fundamental advantage because, in contrast to the Warden, she may have access to the acquisition oracle (the digital camera) as well as the scene being imaged. For example, she can utilize the RAW sensor capture or multiple acquisitions of the same scene or even "manufacture" the side-information prior to embedding by subjecting the precover image to an information-reducing operation, such as quantization, downsampling, format conversion, and/or reduction of the dynamic range.

In this section, we adopt a tractable model of the acquisition noise estimated from the available side-information and derive the embedding rule by minimizing the KL divergence between cover and stego images. The precover is modeled as an array of $n = n_1 \times n_2$ independent but not necessarily identically distributed random variables $X_{ij},\ i = 1,\dots,n_1,\ j = 1\dots,n_2$. The error introduced by color interpolation is also included in the acquisition noise since we view demosaicking as part of acquisition. Therefore, X_{ij} will naturally have a larger variance in textured and noisy regions where color interpolation algorithms are less accurate. Since a commonly adopted model of acquisition noise inherent to the sensor is the Gaussian distribution (the heteroscedastic model [9, 28]), we adopt this model in this paper as well:

$$X_{ij} \sim \mathcal{N}(t_{ij}, \sigma_{ij}^2), \tag{10}$$

where t_{ij} is the precover value that would be registered by the sensor in the absence of acquisition noise. Finally, we assume that the ijth cover element c_{ij} is obtained by rounding a specific realization of X_{ij}. This is a simplification because in practice X_{ij} will be constrained to a finite dynamic range. Note that the rounding encompasses more general uniform scalar quantizers.

We constrain ourselves to the simplest case when the sender has one precover x_{ij}, which will be used to obtain an estimate of the true scene, $\hat{t}_{ij}$, as well as the acquisition noise variance, $\hat{\sigma}_{ij}$. We note that the methodology proposed here can be extended to more general types of side-information, such as multiple precovers / covers. This possibility to postponed to our future work.

The embedding operation considered in this paper is binary, meaning that each precover element x_{ij} will be either rounded to $c_{ij} = [x_{ij}]$ or to $\bar{c}_{ij} = [x_{ij}] + \text{sign}(x_{ij} - [x_{ij}])$ with the convention that when x_{ij} is an integer, $\bar{c}_{ij} \in \{c_{ij} - 1, c_{ij} + 1\}$ is chosen uniformly randomly.

Denoting the stego image elements with s_{ij}, the embedding will change c_{ij} to $\bar{c}_{ij}$ with probability β'_{ij}:

$$\Pr\{s_{ij} = c_{ij}\} = 1 - \beta'_{ij} \tag{11}$$

$$\Pr\{s_{ij} = \bar{c}_{ij}\} = \beta'_{ij}, \tag{12}$$

effectively embedding $h_2(\beta'_{ij})$ nats, where $h_2(x) = -x\ln x - (1-x)\ln(1-x)$ is the binary entropy function.

Denoting the closed interval $\mathcal{I}_{ij} = [c_{ij}, \bar{c}_{ij}]$ when $\bar{c}_{ij} > c_{ij}$ and $\mathcal{I}_{ij} = [\bar{c}_{ij}, c_{ij}]$ when $\bar{c}_{ij} \le c_{ij}$, the embedding is designed to minimize the KL divergence between cover and stego distributions conditioned on precover values $X_{ij} \in \mathcal{I}_{ij}$. WLOG, in what follows we will assume that $\bar{c}_{ij} > c_{ij}$. This conditional probability for the cover $c_{ij} = [X_{ij}]$ is

$$\Pr\{[X_{ij}] = c_{ij} | X_{ij} \in \mathcal{I}_{ij}\} = \frac{p_{ij}(c_{ij})}{p_{ij}(c_{ij}) + p_{ij}(\bar{c}_{ij})} \tag{13}$$

$$\Pr\{[X_{ij}] = \bar{c}_{ij} | X_{ij} \in \mathcal{I}_{ij}\} = \frac{p_{ij}(\bar{c}_{ij})}{p_{ij}(c_{ij}) + p_{ij}(\bar{c}_{ij})} \tag{14}$$

where

$$p_{ij}(c_{ij}) = Q\left(\frac{c_{ij} - \hat{t}_{ij}}{\hat{\sigma}_{ij}}\right) - Q\left(\frac{c_{ij} + 1/2 - \hat{t}_{ij}}{\hat{\sigma}_{ij}}\right) \tag{15}$$

$$p_{ij}(\bar{c}_{ij}) = Q\left(\frac{c_{ij} + 1/2 - \hat{t}_{ij}}{\hat{\sigma}_{ij}}\right) - Q\left(\frac{c_{ij} + 1 - \hat{t}_{ij}}{\hat{\sigma}_{ij}}\right) \tag{16}$$

with $Q(x)$ the tail probability of a standard normal random variable $\mathcal{N}(0,1)$.

The ijth pixel in the stego image is modeled as a discrete random variable S_{ij} with range $\{c_{ij}, \bar{c}_{ij}\}$ with pmf

$$\Pr\{S_{ij} = c_{ij} | X_{ij} \in \mathcal{I}_{ij}\} = \frac{q_{ij}(c_{ij})}{q_{ij}(c_{ij}) + q_{ij}(\bar{c}_{ij})}, \tag{17}$$

$$\Pr\{S_{ij} = \bar{c}_{ij} | X_{ij} \in \mathcal{I}_{ij}\} = \frac{q_{ij}(\bar{c}_{ij})}{q_{ij}(c_{ij}) + q_{ij}(\bar{c}_{ij})}, \tag{18}$$

where

$$q_{ij}(c_{ij}) = (1-\beta'_{ij})p_{ij}(c_{ij}) + \beta'_{ij}p_{ij}(\bar{c}_{ij}) \tag{19}$$

$$q_{ij}(\bar{c}_{ij}) = \beta'_{ij}p_{ij}(c_{ij}) + (1-\beta'_{ij})p_{ij}(\bar{c}_{ij}). \tag{20}$$

Denoting for compactness $r^{\mathrm{L}}_{ij} = p_{ij}(c_{ij})$ and $r^{\mathrm{R}}_{ij} = p_{ij}(\bar{c}_{ij})$, a straightforward derivation gives the following KL divergence between ijth cover and stego elements conditioned on $X_{ij} \in \mathcal{I}_{ij}$:

$$\begin{aligned}
d_{ij} &= D_{\mathrm{KL}}\left(C_{ij} \middle\| S_{ij} \middle| X_{ij} \in \mathcal{I}_{ij}\right) && (21)\\
&= \frac{r^{\mathrm{L}}_{ij}}{r^{\mathrm{L}}_{ij}+r^{\mathrm{R}}_{ij}} \log \frac{r^{\mathrm{L}}_{ij}}{(1-\beta'_{ij})r^{\mathrm{L}}_{ij} + \beta'_{ij}r^{\mathrm{R}}_{ij}} && (22)\\
&\quad + \frac{r^{\mathrm{R}}_{ij}}{r^{\mathrm{L}}_{ij}+r^{\mathrm{R}}_{ij}} \log \frac{r^{\mathrm{R}}_{ij}}{\beta'_{ij}r^{\mathrm{L}}_{ij} + (1-\beta'_{ij})r^{\mathrm{R}}_{ij}} && (23)\\
&= -\frac{r^{\mathrm{L}}_{ij}}{r^{\mathrm{L}}_{ij}+r^{\mathrm{R}}_{ij}} \log \left(1 - \beta'_{ij} + \beta'_{ij}\frac{r^{\mathrm{R}}_{ij}}{r^{\mathrm{L}}_{ij}}\right) && (24)\\
&\quad - \frac{r^{\mathrm{R}}_{ij}}{r^{\mathrm{L}}_{ij}+r^{\mathrm{R}}_{ij}} \log \left(1 - \beta'_{ij} + \beta'_{ij}\frac{r^{\mathrm{L}}_{ij}}{r^{\mathrm{R}}_{ij}}\right) && (25)\\
&\doteq \frac{{\beta'_{ij}}^2}{2(r^{\mathrm{L}}_{ij}+r^{\mathrm{R}}_{ij})}\\
&\quad \times \left(r^{\mathrm{L}}_{ij}\left(1 - \frac{r^{\mathrm{R}}_{ij}}{r^{\mathrm{L}}_{ij}}\right)^2 + r^{\mathrm{R}}_{ij}\left(1 - \frac{r^{\mathrm{L}}_{ij}}{r^{\mathrm{R}}_{ij}}\right)^2\right) && (26)\\
&= \frac{{\beta'_{ij}}^2}{2}\frac{(r^{\mathrm{L}}_{ij} - r^{\mathrm{R}}_{ij})^2}{r^{\mathrm{L}}_{ij}+r^{\mathrm{R}}_{ij}}\left(\frac{1}{r^{\mathrm{L}}_{ij}} + \frac{1}{r^{\mathrm{R}}_{ij}}\right) && (27)\\
&\doteq \frac{1}{2}{\beta'_{ij}}^2 I'_{ij}, && (28)
\end{aligned}$$

where

$$I'_{ij} = \frac{(r^{\mathrm{L}}_{ij} - r^{\mathrm{R}}_{ij})^2}{r^{\mathrm{L}}_{ij}r^{\mathrm{R}}_{ij}} \tag{29}$$

is the Fisher information obtained by expanding the logarithms using Taylor series w.r.t. β' at $\beta' = 0$ and keeping only the leading term.

The total KL divergence between the cover and stego objects $C = (C_{ij})$, $S = (S_{ij})$, $i = 1, \ldots, n_1$, $j = 1 \ldots, n_2$, is the sum

$$D_{\mathrm{KL}}(C||S) = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} d_{ij} \doteq \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \frac{1}{2}{\beta'_{ij}}^2 I'_{ij}. \tag{30}$$

The actual embedding change rates β'_{ij} are determined by minimizing $D_{\mathrm{KL}}(C||S)$ under the payload constraint

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} h_2(\beta'_{ij}) = \alpha n, \tag{31}$$

where α is expressed in nats per pixel. Similar to MiPOD [25], the proposed scheme minimizes the sum of pixels' Fisher information weighted by squared change rates as in this case the embedding "cost" relates to statistical detectability. Note that minimizing the KL divergence makes the design optimal only against an omniscient Warden who knows the exact actions of the embedder, including the rounding errors e_{ij}. When the problem of embedding and detection are formulated withing a game-theoretical framework when both the sender and the Warden randomize their strategies of how to distribute (detect) the payload, the sender should also minimize a weighted sum of squared change rates to operate at the Nash equilibrium of a zero-sum game when the payoff is defined as Warden's test power for a bounded false alarm [22].

The optimization problem can be approached in a standard manner using the method of Lagrange multipliers by solving the equations $\partial L/\partial\beta'_{ij} = 0$, where

$$L = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \frac{1}{2}{\beta'_{ij}}^2 I'_{ij} - \lambda'\left(h_2(\beta'_{ij}) - \alpha n\right). \tag{32}$$

This leads to n non-linear equations for β'_{ij}

$$\beta'_{ij} I'_{ij} = \lambda' \log \frac{1-\beta'_{ij}}{\beta'_{ij}}, \tag{33}$$

for each ij, which can be solved numerically, for example, by a binary search over $\lambda' \in [0, \infty]$ in a similar manner as described in [17]. This embedding algorithm will be called side-informed MiPOD (SI-MiPOD).

Extension to JPEG domain

The proposed scheme is extended to work with JPEG images in a straightforward manner. The independence imposed on the noise $\mathbf{n}$ and linearity of the DCT allows us to easily compute the variance of each non-rounded DCT coefficient given the estimated variances of pixels $\hat{\sigma}_{ij}$. Thus, the same MVG model of acquisition noise can be applied to non-rounded DCT coefficients.

Given an 8×8 block of pixel values $z_{ij} \in \{0, \ldots, 255\}$, $0 \le i, j \le 7$, in an uncompressed image $\mathbf{z}$, the klth DCT coefficient d_{kl}, $0 \le k, l \le 7$, before rounding is a linear combination of pixel values

$$d_{kl} = 1/q_{kl} \sum_{i,j=0}^{7} f^{(k,l)}_{ij} z_{ij}, \tag{34}$$

or in matrix form symbolically $\mathbf{d} = \mathrm{DCT}(\mathbf{z})$, where

$$f^{(k,l)}_{ij} = \frac{w_k w_l}{4} \cos\frac{\pi k(2i+1)}{16} \cos\frac{\pi l(2j+1)}{16}, \tag{35}$$

$w_0 = 1/\sqrt{2}$, $w_k = 1$ for $k > 0$, and q_{kl} is the JPEG quantization matrix. With the adopted acquisition model (10), the precover DCT coefficients d_{kl} are independent samples from an array of Gaussian variables $\mathcal{N}(\mu_{kl}, \sigma^2_{kl})$. The parameters of the Gaussian acquisition noise can be estimated as

$$\hat{\sigma}^2_{kl} = 1/q^2_{kl} \sum_{i,j=0}^{7} \left(f^{(k,l)}_{ij}\right)^2 \hat{\sigma}^2_{ij}, \tag{36}$$

where $\hat{\sigma}^2_{ij}$ are pixel variances estimated in the spatial domain using, e.g., the variance estimator used in MiPOD. With a single precover, the mean is estimated as

$$\hat{\mu}_{kl} = d_{kl}. \tag{37}$$

Having estimated the MVG parameters, the proposed side-informed embedding rounds d_{kl} to either $[d_{kl}]$ or $[d_{kl}] + \text{sign}(e_{kl})$, $e_{kl} = d_{kl} - [d_{kl}]$ as described above. This embedding algorithm will be called side-informed J-MiPOD (SI-J-MiPOD).

Connection to heuristic schemes

We now contrast the derived embedding rule with heuristic binary side-informed embedding with precover (see Eq. (3)). When only the precover value x_{ij} is available as side-information, $\hat{t}_{ij} = x_{ij}$ is a minimum variance unbiased estimate under our acquisition model. To obtain a closed-form expression, we work in the limit of small and large values of $\hat{\sigma}_{ij}$ for both schemes.

Model-based SI-MiPOD

Without loss on generality, we will assume that $e_{ij} \geq 0$. Recalling the formula for p_{ij} (15) and Taylor expansion of $Q(x) \approx \frac{1}{2} - \frac{1}{\sqrt{2\pi}}x + \frac{1}{6\sqrt{2\pi}}x^3 + \mathcal{O}(x^5)$ at $x = 0$, for large $\hat{\sigma}_{ij}$

$$\begin{aligned} r^{\mathrm{L}}_{ij} &= Q\left(\frac{[x_{ij}] - x_{ij}}{\hat{\sigma}_{ij}}\right) - Q\left(\frac{[x_{ij}] + 1/2 - x_{ij}}{\hat{\sigma}_{ij}}\right), \\ &= Q\left(\frac{-e_{ij}}{\hat{\sigma}_{ij}}\right) - Q\left(\frac{-e_{ij} + 1/2}{\hat{\sigma}_{ij}}\right) && (38) \\ &\doteq \frac{1}{2\hat{\sigma}_{ij}\sqrt{2\pi}} - \frac{1}{6\hat{\sigma}^3_{ij}\sqrt{2\pi}}\left(\frac{3}{2}e^2_{ij} - \frac{3}{4}e_{ij} + \frac{1}{8}\right) && (39) \end{aligned}$$

and

$$\begin{aligned} r^{\mathrm{R}}_{ij} &= Q\left(\frac{[x_{ij}] + 1/2 - x_{ij}}{\hat{\sigma}_{ij}}\right) - Q\left(\frac{[x_{ij}] + 1 - x_{ij}}{\hat{\sigma}_{ij}}\right), \\ &= Q\left(\frac{-e_{ij} + 1/2}{\hat{\sigma}_{ij}}\right) - Q\left(\frac{-e_{ij} + 1}{\hat{\sigma}_{ij}}\right) && (40) \\ &\doteq \frac{1}{2\hat{\sigma}_{ij}\sqrt{2\pi}} - \frac{1}{6\hat{\sigma}^3_{ij}\sqrt{2\pi}}\left(\frac{3}{2}e^2_{ij} - \frac{9}{4}e_{ij} + \frac{7}{8}\right). && (41) \end{aligned}$$

Thus,

$$(r^{\mathrm{L}}_{ij} - r^{\mathrm{R}}_{ij})^2 \doteq \frac{(1/2 - e_{ij})^2}{16\hat{\sigma}^6_{ij}2\pi} \tag{42}$$

and

$$1/r^{\mathrm{L}}_{ij} \doteq 1/r^{\mathrm{R}}_{ij} \doteq 2\hat{\sigma}_{ij}\sqrt{2\pi} \tag{43}$$

for large $\hat{\sigma}_{ij}$. Finally, the Fisher information (29) becomes

$$I'_{ij} \doteq \frac{1}{4\hat{\sigma}^4_{ij}}\left(1/2 - |e_{ij}|\right)^2. \tag{44}$$

This expression for the FI (44) has been written for the general case when e_{ij} can be both negative and positive. It informs us that, for large noise variance, the rounding error should modulate the Fisher information by $(1/2 - |e_{ij}|)^2$, which achieves a similar effect as multiplying the costs by $1 - 2|e_{ij}|$ in heuristic cost-based side-informed embedding schemes. We remark that the approximation (44) is fairly accurate as long as $\hat{\sigma}_{ij} \gtrsim 2$ for all values of e_{ij}.

For a small $\hat{\sigma}_{ij}$, the terms r^{L}_{ij} and r^{R}_{ij} in (29) can be simplified using the asymptotic expression for $Q(x) \approx \frac{1}{\sqrt{2\pi}x}e^{-x^2/2}$ for large x:

$$\begin{aligned} r^{\mathrm{L}}_{ij} &= Q\left(\frac{[x_{ij}] - x_{ij}}{\hat{\sigma}_{ij}}\right) - Q\left(\frac{[x_{ij}] + 1/2 - x_{ij}}{\hat{\sigma}_{ij}}\right), \\ &\doteq 1 - \frac{\exp\left(-(-e_{ij})^2/(2\hat{\sigma}^2_{ij})\right)}{\sqrt{2\pi}\left(e_{ij}/\hat{\sigma}_{ij}\right)} \\ &\quad - \frac{\exp\left(-(1/2 - e_{ij})^2/(2\hat{\sigma}^2_{ij})\right)}{\sqrt{2\pi}\left((1/2 - e_{ij})/\hat{\sigma}_{ij}\right)} && (45) \\ &\doteq 1. && (46) \end{aligned}$$

Following the same steps,

$$r^{\mathrm{R}}_{ij} \doteq \frac{\exp\left(-(1/2 - e_{ij})^2/(2\hat{\sigma}^2_{ij})\right)}{\sqrt{2\pi}(1/2 - e_{ij})/\hat{\sigma}_{ij}}, \tag{47}$$

and finally

$$\begin{aligned} I'_{ij} &= \frac{\left(r^{\mathrm{R}}_{ij} - r^{\mathrm{L}}_{ij}\right)^2}{r^{\mathrm{L}}_{ij} r^{\mathrm{R}}_{ij}} \doteq \frac{1}{r^{\mathrm{R}}_{ij}} && (48) \\ &\doteq \sqrt{2\pi}\frac{1/2 - e_{ij}}{\hat{\sigma}_{ij}} \exp\left((1/2 - e_{ij})^2/(2\hat{\sigma}^2_{ij})\right). && (49) \end{aligned}$$

For small $\hat{\sigma}_{ij}$, we will derive an approximate solution to (33) by taking its logarithm

$$\ln \beta'_{ij} + \ln I'_{ij} \doteq \ln \lambda' + \ln\left(-\ln\left(\beta'_{ij}\right)\right), \tag{50}$$

from which $\beta'_{ij} \doteq \lambda'/I'_{ij}$ and thus

$$\beta'_{ij} \doteq \frac{\lambda'}{I'_{ij}} = \frac{1}{\sqrt{2\pi}}\frac{\hat{\sigma}_{ij}}{1/2 - e_{ij}} \exp\left(-(1/2 - e_{ij})^2/(2\hat{\sigma}^2_{ij})\right). \tag{51}$$

When $\hat{\sigma}_{ij} \ll 1/2 - e_{ij}$, the value of β'_{ij} will rapidly approach zero.

Heuristic SI-MiPOD

To obtain closed-form expressions, we will again consider the case of a small $\hat{\sigma}_{ij}$ and large $\hat{\sigma}_{ij}$. As described in the above section on prior art, the heuristic side-informed MiPOD starts with MiPOD's embedding costs ρ_{ij} derived in (8), which are then modulated by the rounding error $\rho'_{ij} = \rho_{ij}\left(1-2\left|e_{ij}\right|\right)$. The side-informed embedding probabilities are

$$\beta'_{ij} = \frac{e^{-\lambda'\rho'_{ij}}}{1+2e^{-\lambda'\rho'_{ij}}} = \frac{1}{2+e^{\lambda'\rho'_{ij}}}, \tag{52}$$

where λ' is a Lagrange multiplier determined from the payload constraint $\sum_{ij} h_2(\beta'_{ij}) = \alpha n$. The embedding can be reinterpreted as MiPOD with Fisher information

$$\begin{aligned} I'_{ij} &= \frac{\lambda'}{\beta'_{ij}} \ln\left(1/\beta'_{ij} - 2\right) \\ &= \lambda'(2+e^{\lambda'\rho'_{ij}})\lambda'\rho'_{ij}. \end{aligned} \tag{53}$$

First, we carry out the analysis for large $\hat{\sigma}_{ij}$. To this end, we rewrite Eq. (7) determining MiPOD's embedding change rates as

$$\exp(\beta_{ij} I_{ij}/\lambda) = 1/\beta_{ij} - 2. \tag{54}$$

Recalling that $I_{ij} = 2/\hat{\sigma}_{ij}^4 \ll 1$ for large $\hat{\sigma}_{ij}$, the left side can be approximated using Taylor expansion as

$$1 + \frac{\beta_{ij} I_{ij}}{\lambda} + \frac{\beta_{ij}^2 I_{ij}^2}{2\lambda^2} = 1/\beta_{ij} - 2 \tag{55}$$

which is equivalent to

$$\frac{1}{2\lambda^2} y_{ij}^3 + \frac{1}{\lambda} y_{ij}^2 + 3y_{ij} = I_{ij} \tag{56}$$

where $y_{ij} = I_{ij}\beta_{ij}$. For small I_{ij}, the first-order solution to this cubic equation is $y_{ij}^{(1)} \doteq \frac{I_{ij}}{3}$ and the second-order solution

$$y_{ij}^{(2)} = \frac{I_{ij}}{3} - \frac{1}{3\lambda}\left(\frac{I_{ij}}{3}\right)^2, \tag{57}$$

which gives us $\beta_{ij} \doteq \frac{1}{3} - \frac{I_{ij}}{27\lambda}$. Substituting this approximation to Eq. (8) and keeping only the leading term gives $\rho_{ij} \doteq \frac{I_{ij}}{3\lambda}$ and for the Fisher information (53)

$$\begin{aligned} I'_{ij} &\propto \rho'_{ij} = \rho_{ij}(1-2|e_{ij}|) \\ &\propto I_{ij}(1-2|e_{ij}|). \end{aligned} \tag{58}$$

For small $\hat{\sigma}_{ij}$, I_{ij} will be large and $\beta_{ij} \doteq \lambda/I_{ij}$ small, as derived in the previous section, and therefore, $\rho_{ij} \approx \ln(I_{ij}/\lambda - 2)$. Substituting this result into (52) and recalling that $I_{ij} = 2\hat{\sigma}_{ij}^{-4}$:

$$\begin{aligned} \beta'_{ij} &= \frac{1}{2+\exp(\lambda'\rho'_{ij})} \\ &\doteq \exp(-\lambda'\rho'_{ij}) = \exp(-\lambda'\rho_{ij}(1-2|e_{ij}|)) \\ &\doteq \exp\left(-\lambda'\ln(I_{ij}/\lambda - 2)(1-2|e_{ij}|)\right) \\ &= C(\lambda, e_{ij})\hat{\sigma}_{ij}^{4\lambda'(1-2|e_{ij}|)}, \end{aligned} \tag{59}$$

where $C(\lambda, e_{ij})$ does not depend on $\hat{\sigma}_{ij}$.

Comparison of model-based and heuristic MiPOD

In this section, we compare the properties of heuristic SI-MiPOD and model-based SI-MiPOD using the above approximations and experimentally.

For $\hat{\sigma}_{ij} \gtrsim 2$, the Fisher information in model-based SI-MiPOD is modulated by $\left(1-2\left|e_{ij}\right|\right)^2$ (44) while in the heuristic SI-MiPOD by $\left(1-2\left|e_{ij}\right|\right)$ (58). Therefore, model-based SI-MiPOD embeds a higher payload in pixels with large noise variance.

For small $\hat{\sigma}_{ij}$, the embedding probabilities β'_{ij} for model-based SI-MiPOD rapidly approach zero for all values of e_{ij} (51), while for the heuristic SI-MiPOD the behavior of the probabilities $\beta'_{ij}(\hat{\sigma}_{ij})$ depends on the exponent $4\lambda'\left(1-2\left|e_{ij}\right|\right)$ (59). As $e_{ij} \to 1/2$, the exponent approaches zero and the embedding probability as a function of $\hat{\sigma}_{ij}$ should become concave. Summarizing both observations, we conclude that the model-based SI-MiPOD is more adaptive to the acquisition noise variance $\hat{\sigma}_{ij}^2$ than heuristic SI-MiPOD.

All quantitative conclusions reached above are now confirmed on an artificial precover x_{ij} with 5×64 pixels with all 5×64 combinations of values of five rounding errors $e_i \in \{0, 0.125, 0.25, 0.375, 0.495\}$ and 64 variances linearly spaced between 0.01 and 64. WLOG, we set the precover values to $x_{ij} = \hat{t}_{ij} = e_i$.

First, the embedding change rates β_{ij} were computed for payload α nats as explained in the section on prior art (also, see [25, 17]), converted to costs ρ_{ij} using Eq. (8), and modulated $\rho'_{ij} = (1-2e_i)\rho_{ij}$ as in Eq. (3). The modulated costs were then used to obtain the change rates via $\beta'_{ij} = \exp(-\lambda'\rho'_{ij})/(1+\exp(-\lambda'\rho'_{ij}))$ with λ' determined by the same payload α.

Figure 1 shows the embedding change probability β'_{ij} for $\alpha = 0.3$ nats as a function of the variance σ_j^2 for each e_i with the five curves corresponding to five values of e_i, $i = 1, \ldots, 5$. The image on the left is for the heuristic binary SI-MiPOD (see [6] and Eq. (3)) while the image on the right corresponds to the model-based SI-MiPOD. The lines positioned lower correspond to lower e_i. Note that both schemes embed maximum possible entropy when $e \to 1/2$. In agreement with our analysis, for small variance, model-based SI-MiPOD is more conservative (embeds with smaller change rates) than heuristic SI-MiPOD. On the other hand, for large variance, the model-based version embeds with larger change rates. In other words, model-based SI-MiPOD is more content adaptive than its heuristic counterpart. While this may make the model-based approach more vulnerable to attacks utilizing the selection channel, such attacks are much more difficult to implement for the Warden because she does not have access to the rounding errors. More on this topic appears in Section "Public vs. private side-information and adaptivity."

Experiments

In this section, we provide the results and interpretation of all experiments in both spatial and JPEG domains. We start with the description of our image sources. To con-

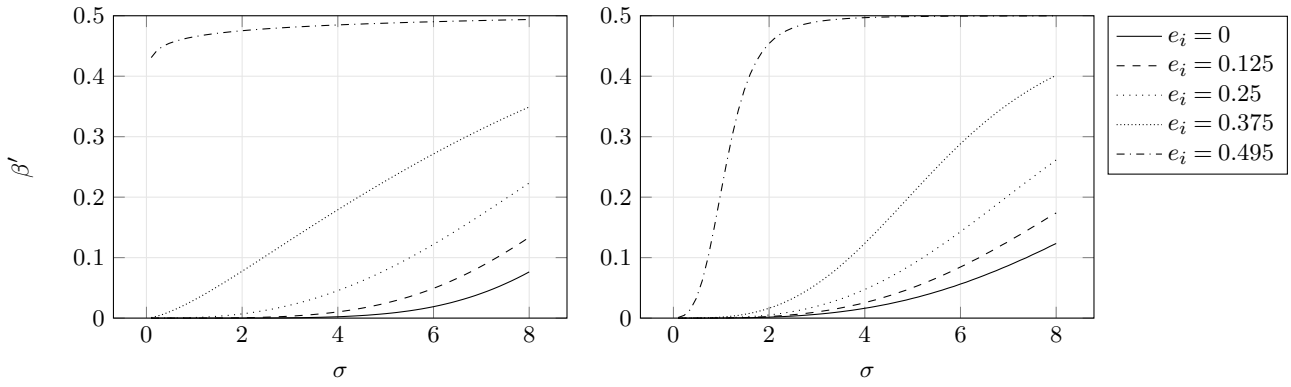

Figure 1. *Embedding change probability β' as a function of variance σ^2 on a synthetic cover for $\alpha = 0.3$ nats using heuristic side-informed binary MiPOD (left) and model-based binary MiPOD (right).*

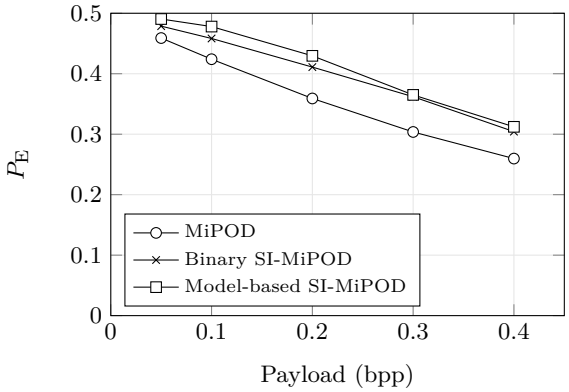

Figure 2. *Security of MiPOD and its heuristic and model-based SI versions with side-information in the form of precover obtained by converting BOSSColor images to grayscale.*

trast the performance of heuristic side-informed schemes with the model-based versions, we include experiments on uncompressed as well as JPEG images.

Image sources

Two primary data sources will be used: BOSSbase 1.01 with 10,000 512×512 *color* images, called BOSSColor, and the same database of *grayscale* images, which will be addressed as BOSSbase. All images were all taken in the RAW format by seven different cameras, downsampled and cropped to the final size of 512×512 pixels. The script used for the conversion and processing is also available from the same web site as the database itself [1]. To create BOSSColor, we modified the script to skip the conversion from RGB to grayscale applied when creating BOSSbase.

Spatial domain (BOSSColor)

Figure 2 shows the results of our first experiment on real imagery. BOSSColor images were converted to grayscale using the formula $x = 0.2989R + 0.5870G + 0.1140B$, producing non-rounded precover values x_{ij}, which would then be rounded to 8bit integers $c_{ij} = [x_{ij}]$ to obtain the cover source of 8-bit grayscale 512×512 images. Estimates of pixels' noise variance $\hat{\sigma}^2_{ij}$ were obtained from precover grayscale values x_{ij} using the same variance estimator as in MiPOD (Section V in [25]). Because the selection channel depends on the rounding error e_{ij}, it is not clear how to utilize selection-channel-aware feature sets. Thus, we carry out steganalysis using the spatial rich model (SRM) [16] with the low-complexity linear classifier [4] as a classifier. The empirical detectability was measured using the minimal total probability of error, $P_{\mathrm{E}} = \min_{P_{\mathrm{FA}}}(P_{\mathrm{MD}} + P_{\mathrm{FA}})/2$, where P_{MD} and P_{FA} are missed-detection and false-alarm rates.

Alongside the above proposed side-informed technique with $\hat{t}_{ij} = x_{ij}$ and $\hat{\sigma}^2_{ij}$, we also tested the (non side-informed) MiPOD itself on grayscale cover images with pixel values $c_{ij} = [x_{ij}]$ and variances $\hat{\sigma}^2_{ij}$, and its binary side-informed version with costs modulated as described in (3). We note that all three embedding schemes were simulated using an embedding simulator.

The model-based binary SI-MiPOD improves the (ternary) MiPOD by 3–5% and the heuristic SI-MiPOD by up to 2%. No improvement over the heuristic method was observed for payloads larger than 0.3 bpp. We expect that further gain may be obtained by searching for the best parameters of the variance estimator in our model.

JPEG domain

We start with a note that with the introduction of model-based side-informed JPEG steganography (SI-J-MiPOD) as described at the end of the previous section, there now exists an equivalent embedding algorithm that does not employ side-information, which we call J-MiPOD. It starts from a JPEG cover image that is decompressed to the spatial domain (without rounding or clipping), pixel variances are estimated in the spatial domain using MiPOD's variance estimator, and the corresponding variances of DCT coefficients are computed using (36). MiPOD is then directly applied as described in [25] to quantized DCT coefficients of the cover JPEG. That is, the coefficients are changed by ± 1 with equal probabilities that are computed to minimize the KL divergence

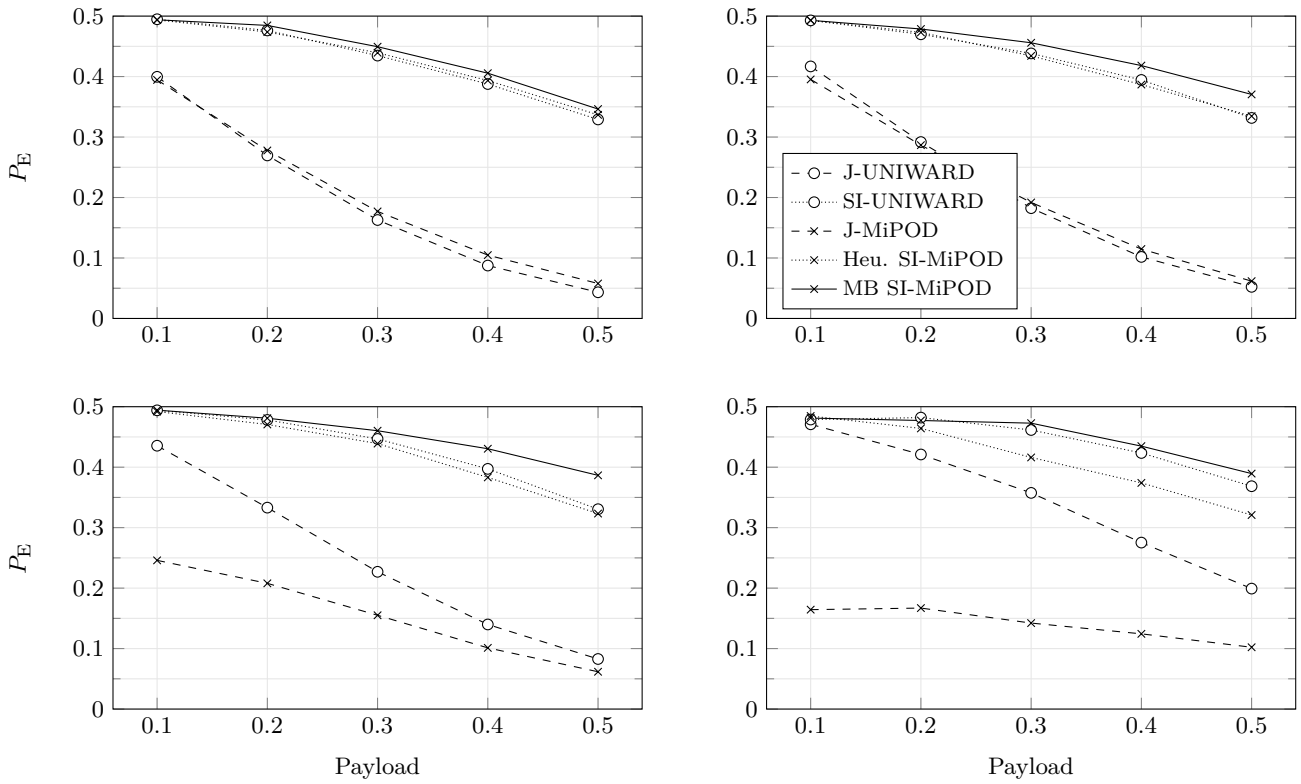

Figure 3. *Security of two non side-informed and three side-informed embedding schemes as a function of payload in bpnzac on BOSSbase for JPEG quality factors 65, 75, 85, and 95.*

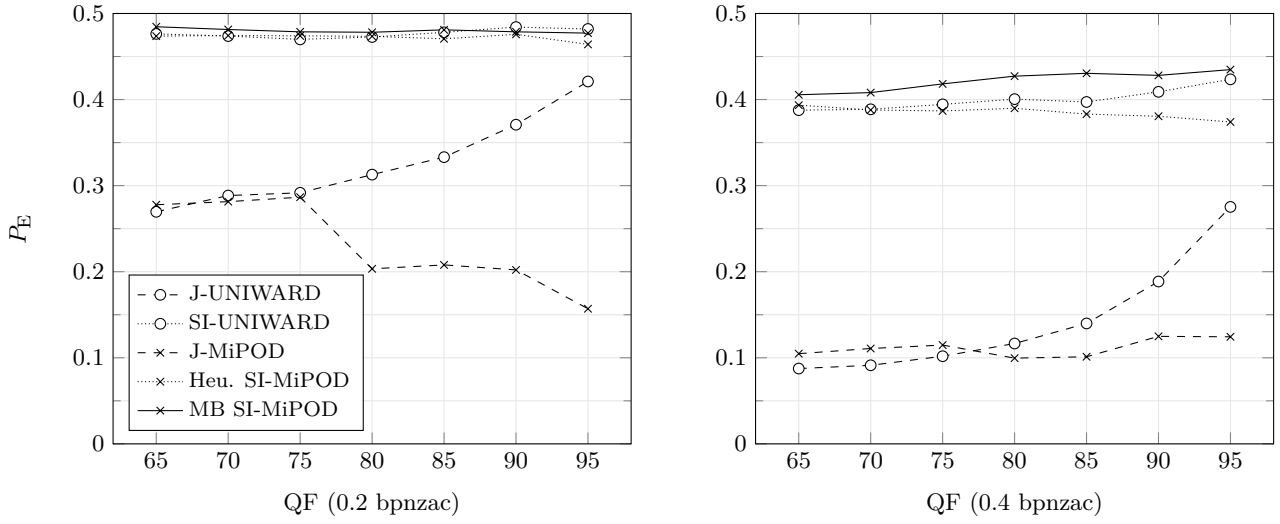

Figure 4. *Security of two non side-informed and three side-informed embedding schemes on BOSSbase as a function of JPEG quality factor for relative payload 0.2 bpnzac (left) and 0.4 (right).*

between cover and stego DCT coefficients under a payload constraint.

In this section, we thus compare the empirical security of five embedding algorithms: two that do not use side-information and start with a JPEG cover (J-UNIWARD and J-MiPOD) and three side-informed schemes, SI-UNIWARD, heuristic SI-J-MiPOD and model-based SI-J-MiPOD as described at the end of the previous section. The detector was built with the Gabor Filter Residuals (GFR) [27] features and the low-complexity linear classifier.

Figure 3 shows the detection error P_{E} as a function of payload in bits per non-zero AC DCT coefficient (bpnzac) for four JPEG quality factors while Figure 4 shows the same detection error but as a function of JPEG quality factor for two payloads. Neither figure contains error bars for better readability. The average statistical spread of the results in terms of the standard deviation across ten database splits was 0.0029 with 91% of all spreads falling into the range 0.0010–0.0040.

It is comforting to confirm that the model-based SI-J-MiPOD is more secure than the heuristic SI-J-MiPOD, which supports the proposed theoretical approach to side-informed embedding. The gain increases with increased JPEG quality factor and becomes up to almost 7% for QF 95 and payload 0.5 bpnzac. For small payloads, all three side-informed schemes are nearly undetectable across all quality factors as evidenced in the left graph of Figure 4. For larger payloads, the Model-based SI-J-MiPOD clearly becomes the most secure tested scheme, outperforming SI-UNIWARD as well. It is also interesting to point out that the non side-informed J-MiPOD is on par with J-UNIWARD for lower quality factors (up to 75) but then J-MiPOD starts losing w.r.t. J-UNIWARD. We attribute this to the effect of compression on estimation of pixel variance because when both algorithms are fed costs and variances estimated from the precover, J-MiPOD is more secure than J-UNIWARD across all tested payloads and quality factors. We thus hypothesize that J-MiPOD might benefit from fine-tuning the variance estimator to decompressed JPEG images. Since this topic does not concern side-informed steganography, it is left to our future effort.

Public vs. private side-information and adaptivity

The selection channel in side-informed steganography is determined by both content complexity via the pixel variance $\hat{\sigma}_{ij}^2$ and by the side-information, the rounding error e_{ij}. The pixel variance is only slightly changed by the embedding itself and thus constitutes a *public* side-information available to the Warden. On the other hand, the rounding error e_{ij} is extremely difficult to estimate from the cover/stego image even for images with simple content, such as blue sky images. Although we cannot cite a source for this claim, this finding is based on our previous unpublished effort and should be entirely plausible considering the fact that, for example, it is generally impossible to obtain an accurate estimate of unquantized values of DCT coefficients in a JPEG file. The only publications related to improving the quality of JPEG decompressed images relate to visually suppressing blockiness artifacts, which is a task that is already difficult and much less demanding than estimating the unquantized DCT coefficients. In short, it is a reasonable assumption that the rounding error is a *private* side-information unavailable to the Warden.

Thus, ideally an embedding scheme should be less adaptive to content ($\hat{\sigma}_{ij}^2$) and more strongly adaptive to the rounding error e_{ij}. Since we designed the model-based SI-MiPOD to minimize the KL divergence between cover and stego distributions, we are essentially assuming an omniscient Warden who knows both $\hat{\sigma}_{ij}^2$ and e_{ij}. Based on the analysis in the above section on comparison between heuristic and model-based SI-MiPOD, the latter is more strongly adaptive to content (the acquisition noise variance) than the former. While the experiments in the previous section show that the model-based approach is less detectable when steganalyzed with the selection-channel-*un*aware SRM (GFR) features (an ignorant Warden), the situation reverses when the steganalyst utilizes the public selection channel (content). However, the model-based approach is indeed by design much less detectable w.r.t. the hypothetical omniscient Warden as long as the adopted model is good enough. Table 1 summarizes the results for side-informed embedding in the spatial domain (when converting an RGB image to grayscale) and in JPEG domain when steganalyzing with SRM (GFR) features, which corresponds to an ignorant Warden, maxSRMd2 and selection-channel-aware GFR features (SCA-GFR) [5], which simulates Warden aware of the public side-information, the content, and the omniscient Warden fully informed by both content ($\hat{\sigma}_{ij}^2$) and the private side-information, the rounding error e_{ij}.

In the spatial domain, strangely enough maxSRMd2 detects worse than SRM. On the other hand, in the JPEG domain the GFR seems to detect the embedding as reliably as the SCA variant of the features.

We note that there are really no other options for any realistic Warden who does not know the rounding errors e_{ij}. Fixing e_{ij} at some medium value is virtually the same as fixing it at any other value. This is because the rounding error is only in a multiplicative factor that modulates the costs (Eq. (3) and (5)) in the heuristic schemes as well as the Fisher information (44) in the model-based scheme.

Conclusions

Steganography with precover at the sender has come a long way. The main progress has been due to advanced coding techniques coupled with a heuristic incorporation of the precover in the embedding algorithm. The typical heuristic calls for modulating costs by $1-2|e|$, where $-1/2 < e \leq 1/2$ is the rounding error. Despite the success of side-informed schemes, such as SI-UNIWARD or UED, little has been done to design the embedding algorithm from general principles. This paper attempts to rectify this situation.

We start by adopting a multivariate Gaussian model for the precover, modeling thus the process of acquiring

Experiment	Payload	Scheme	Ignorant	Aware of σ	Aware of σ, e
RGB	0.2 bpp	Heu SI-MiPOD	0.4104±0.0031	0.4402±0.0033	0.2306±0.0020
		MB SI-MiPOD	0.4203±0.0024	0.3951±0.0072	0.3466±0.0029
	0.4 bpp	Heu SI-MiPOD	0.3002±0.0022	0.3317±0.0025	0.2065±0.0026
		MB SI-MiPOD	0.3012±0.0029	0.2696±0.0031	0.2572±0.0031
JPEG	0.2 bpnzac	Heu SI-J-MiPOD	0.4740±0.0034	0.4761±0.0029	0.4529±0.0031
		MB SI-J-MiPOD	0.4787±0.0022	0.4751±0.0023	0.4630±0.0020
	0.4 bpnzac	Heu SI-J-MiPOD	0.3868±0.0040	0.3953±0.0031	0.3618±0.0015
		MB SI-J-MiPOD	0.4182±0.0022	0.4210±0.0018	0.4038±0.0026

Table 1. Detection error when steganalyzing heuristic SI-MiPOD and model-based (MB) SI-MiPOD with SRM and maxSRMd2 and their JPEG counterparts with SRM/GFR (ignorant Warden), selection-channel-aware maxSRMd2/GFR (Warden aware of content, $\hat{\sigma}_{ij}^2$), and omniscient Warden aware of both of content $\hat{\sigma}_{ij}^2$ and rounding error e_{ij}.

a digital image using an imaging sensor. By constraining the embedding rule to be binary, the embedding change rates are derived to minimize the total KL divergence between cover and stego models estimated from the available precover while enforcing the payload constraint. In contrast to heuristic schemes, in our model-based approach the rounding error e modulates the Fisher information by multiplying it by $(1-2|e|)^2$. The resulting embedding is shown to be more adaptive to content than heuristic side-informed embedding schemes. On experiments with images represented in the spatial and JPEG domain, we demonstrate that the newly derived model-based side-informed steganography enjoys a higher level of empirical security than heuristic embedding schemes when detecting with selection-channel-*un*aware features (ignorant Warden). The same holds when steganalyzing with selection-channel-aware features fully informed by the rounding error (omniscient Warden). This is because the model-based schemes were designed to minimize the KL divergence, which implicitly assumes an omniscient Warden. With features that incorporate the knowledge of the content adaptivity, however, the model-based approach is more detectable than heuristic schemes at least in the spatial domain. Optimal embedding should thus be designed within a game-theoretic framework in which both the sender and the Warden randomize their strategies.

The framework introduced in this paper lends itself to more general forms of side-information, such as multiple acquisitions of the same cover. Also, an extension to ternary side-informed schemes is necessary. Both topics are postponed to our future effort.

The code for the proposed model-based side-informed steganographic scheme is available from `http://dde.binghamton.edu/download/feature_extractors/`.

Acknowledgments

The work on this paper was partially supported by NSF grant No. 1561446 and by Air Force Office of Scientific Research under the research grant number FA9950-12-1-0124. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation there on. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied of AFOSR or the U.S. Government.

References

[1] BOSSbase 1.01. `http://agents.fel.cvut.cz/stegodata/`.

[2] P. Bas. Steganography via cover-source switching. In *IEEE International Workshop on Information Forensics and Security*, Abu Dhabi, December 4–7 2016.

[3] R. Böhme. *Advanced Statistical Steganalysis*. Springer-Verlag, Berlin Heidelberg, 2010.

[4] R. Cogranne, V. Sedighi, T. Pevný, and J. Fridrich. Is ensemble classifier needed for steganalysis in high-dimensional feature spaces? In *IEEE International Workshop on Information Forensics and Security*, Rome, Italy, November 16–19 2015.

[5] T. Denemark, M. Boroumand, and J. Fridrich. Steganalysis features for content-adaptive JPEG steganography. *IEEE Transactions on Information Forensics and Security*, 11:1736–1746, April 20 2016.

[6] T. Denemark and J. Fridrich. Side-informed steganography with additive distortion. In *IEEE International Workshop on Information Forensics and Security*, Rome, Italy, November 16–19 2015.

[7] T. Denemark and J. Fridrich. Steganography with two JPEGs of the same scene. In *IEEE ICASSP*, New Orleans, March 5–9 2017.

[8] T. Filler, J. Judas, and J. Fridrich. Minimizing additive distortion in steganography using syndrome-trellis codes. *IEEE Transactions on Information Forensics and Security*, 6(3):920–935, September 2011.

[9] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian. Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data. *IEEE Transactions on Image Processing*, 17(10):1737–1754, Oct. 2008.

[10] E. Franz. Steganography preserving statistical properties. In F. A. P. Petitcolas, editor, *Information Hiding, 5th International Workshop*, volume 2578 of Lecture Notes in Computer Science, pages 278–294, Noordwijkerhout, The Netherlands, October 7–9, 2002. Springer-Verlag, New York.

[11] E. Franz. Embedding considering dependencies between pixels. In E. J. Delp, P. W. Wong, J. Dittmann, and N. D. Memon, editors, *Proceedings SPIE, Electronic Imaging, Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, volume 6819, pages D 1–D 12, San Jose, CA, January 27–31, 2008.

[12] E. Franz and A. Schneidewind. Pre-processing for adding noise steganography. In M. Barni, J. Herrera, S. Katzenbeisser, and F. Pérez-González, editors, *Information Hiding, 7th International Workshop*, volume 3727 of Lecture Notes in Computer Science, pages 189–203, Barcelona, Spain, June 6–8, 2005. Springer-Verlag, Berlin.

[13] J. Fridrich. On the role of side-information in steganography in empirical covers. In A. Alattar, N. D. Memon, and C. Heitzenrater, editors, *Proceedings SPIE, Electronic Imaging, Media Watermarking, Security, and Forensics 2013*, volume 8665, pages 0I 1–11, San Francisco, CA, February 5–7, 2013.

[14] J. Fridrich and R. Du. Secure steganographic methods for palette images. In A. Pfitzmann, editor, *Information Hiding, 3rd International Workshop*, volume 1768 of Lecture Notes in Computer Science, pages 47–60, Dresden, Germany, September 29–October 1, 1999. Springer-Verlag, New York.

[15] J. Fridrich, M. Goljan, and D. Soukal. Perturbed quantization steganography using wet paper codes. In J. Dittmann and J. Fridrich, editors, *Proceedings of the 6th ACM Multimedia & Security Workshop*, pages 4–15, Magdeburg, Germany, September 20–21, 2004.

[16] J. Fridrich and J. Kodovský. Rich models for steganalysis of digital images. *IEEE Transactions on Information Forensics and Security*, 7(3):868–882, June 2011.

[17] J. Fridrich and J. Kodovský. Multivariate Gaussian model for designing additive distortion for steganography. In *Proc. IEEE ICASSP*, Vancouver, BC, May 26–31, 2013.

[18] L. Guo, J. Ni, and Y. Q. Shi. Uniform embedding for efficient JPEG steganography. *IEEE Transactions on Information Forensics and Security*, 9(5):814–825, 2014.

[19] V. Holub, J. Fridrich, and T. Denemark. Universal distortion design for steganography in an arbitrary domain. *EURASIP Journal on Information Security, Special Issue on Revised Selected Papers of the 1st ACM IH and MMS Workshop*, 2014:1:1–13, 2014.

[20] F. Huang, J. Huang, and Y.-Q. Shi. New channel selection rule for JPEG steganography. *IEEE Transactions on Information Forensics and Security*, 7(4):1181–1191, August 2012.

[21] A. D. Ker. A fusion of maximal likelihood and structural steganalysis. In T. Furon, F. Cayre, G. Doërr, and P. Bas, editors, *Information Hiding, 9th International Workshop*, volume 4567 of Lecture Notes in Computer Science, pages 204–219, Saint Malo, France, June 11–13, 2007. Springer-Verlag, Berlin.

[22] A. D. Ker, T. Pevný, and P. Bas. Rethinking optimal embedding. In F. Perez-Gonzales, F. Cayre, and P. Bas, editors, *4th ACM IH&MMSec. Workshop*, Vigo, Spain, June 20–22, 2016.

[23] Y. Kim, Z. Duric, and D. Richards. Modified matrix encoding technique for minimal distortion steganography. In J. L. Camenisch, C. S. Collberg, N. F. Johnson, and P. Sallee, editors, *Information Hiding, 8th International Workshop*, volume 4437 of Lecture Notes in Computer Science, pages 314–327, Alexandria, VA, July 10–12, 2006. Springer-Verlag, New York.

[24] V. Sachnev, H. J. Kim, and R. Zhang. Less detectable JPEG steganography method based on heuristic optimization and BCH syndrome coding. In J. Dittmann, S. Craver, and J. Fridrich, editors, *Proceedings of the 11th ACM Multimedia & Security Workshop*, pages 131–140, Princeton, NJ, September 7–8, 2009.

[25] V. Sedighi, R. Cogranne, and J. Fridrich. Content-adaptive steganography by minimizing statistical detectability. *IEEE Transactions on Information Forensics and Security*, 11(2):221–234, 2016.

[26] V. Sedighi, J. Fridrich, and R. Cogranne. Content-adaptive pentary steganography using the multivariate generalized Gaussian cover model. In A. Alattar and N. D. Memon, editors, *Proceedings SPIE, Electronic Imaging, Media Watermarking, Security, and Forensics 2015*, volume 9409, San Francisco, CA, February 8–12, 2015.

[27] X. Song, F. Liu, C. Yang, X. Luo, and Y. Zhang. Steganalysis of adaptive JPEG steganography using 2D Gabor filters. In P. Comesana, J. Fridrich, and A. Alattar, editors, *3rd ACM IH&MMSec. Workshop*, Portland, Oregon, June 17–19, 2015.

[28] T. H. Thai, R. Cogranne, and F. Retraint. Camera model identification based on the heteroscedastic noise model. *Image Processing, IEEE Transactions on*, 23(1):250–263, Jan 2014.

[29] C. Wang and J. Ni. An efficient JPEG steganographic scheme based on the block–entropy of DCT coefficents. In *Proc. of IEEE ICASSP*, Kyoto, Japan, March 25–30, 2012.

Author Biography

Tomáš Denemark received his MS in mathematics from the Czech Technical University in Prague in 2012 and now pursues his PhD at Binghamton University. He focuses on steganography and steganalysis.

Jessica Fridrich is Distinguished Professor of Electrical and Computer Engineering at Binghamton University. She received her PhD in Systems Science from Binghamton University in 1995 and MS in Applied Mathematics from Czech Technical University in Prague in 1987. Her main interests are in steganography, steganalysis, and digital image forensic. Since 1995, she has received 20 research grants totaling over $11 mil that lead to more than 180 papers and 7 US patents.