Sum-SINR/sum-capacity optimal multisignature spread-spectrum steganography†

Lili Wei, Dimitris A. Pados, Stella N. Batalama and Michael J. Medley

Department of Electrical Engineering, State University of New York at Buffalo, Buffalo, NY 14260;
Air Force Research Laboratory, 525 Brooks Road, Rome, NY 13441

ABSTRACT

For any given digital host image or audio file (or group of hosts) and any (block) transform domain of interest, we find an orthogonal set of signatures that achieves maximum sum-signal-to-interference-plus-noise ratio (sum-SINR) spread-spectrum message embedding for any fixed embedding amplitude values. We also find the sum-capacity optimal amplitude allocation scheme for any given total distortion budget under the assumption of (colored) Gaussian transform-domain host data. The practical implication of the results is sum-SINR, sum-capacity optimal multiuser/multisignature spread-spectrum data hiding in the same medium. Theoretically, the findings establish optimality of the recently presented Gkizeli-Pados-Medley multisignature eigen-design algorithm.

Keywords: Authentication, covert communications, multiuser data hiding, signal-to-interference-plus-noise ratio, spread spectrum, steganography, sum capacity, sum SINR, watermarking.

1. INTRODUCTION

Digital steganography, from the Greek stegano-graphia meaning covered writing, is the process of embedding a digital “secret” signal (hidden message) in another digital signal called “cover” or “host”1. Unlike general watermarking applications2–3, steganography attempts to establish covert communication between trusting parties and imposes the requirement of concealing the existence of the embedded message.

Determining the embedding process is a crucial step in the design of a steganographic system. Host-carrier properties and distortion, payload, message detector design and recovery performance depend directly on how the message is inserted in the host data. The broad common objective of most steganographic applications is a satisfactory tradeoff between host distortion for concealment purposes and information delivery rate. Message embedding can be performed either directly in the time (audio) or spatial (image) domain4–8 or in a transform domain (for example, for images we may consider full-frame discrete Fourier transform (DFT)9–12, full-frame discrete cosine transform (DCT)13, block DFT or DCT14–18, or wavelet transforms19–21). Direct embedding in the original host signal domain may be desirable for system complexity purposes, while embedding in a transform domain may take advantage of the particular transform domain properties22 and enables the powerful notion of spread-spectrum (SS) data hiding when the secret signal is spread over a wide range of host frequencies23–25. Spread-spectrum steganography parallels, to that extend, modern developments in spread-spectrum communications theory and practice26.

In this paper, we focus our attention on the emerging concept of multiuser (multisignature) steganography where multiple messages (or parts thereof) are hidden with different embedding signatures in the same...
medium with, potentially, different intended recipients. The theoretical challenges of multiuser steganography parallel problems encountered in the fast-rising field of code-division multiple-access (CDMA) communications. In the present work, we find the orthonormal set of signatures that offers maximum sum-signal-to-interference-plus-noise ratio (sum-SINR) embedding in arbitrary transform domains with any given embedding amplitude values. Moreover, for any given total host distortion budget we present a power (amplitude) allocation scheme that maximizes the Shannon sum-capacity of the multiuser steganographic system. The described power allocation algorithm is optimal under the assumption that the transform-domain host data behave as (colored) Gaussian distributed. These theoretical findings establish optimality of the recently presented Gkizeli-Pados-Medley multisignature eigen-design algorithm under the general requirement of an orthogonal multiuser signature set. Experimental studies and comparisons included herein demonstrate the new theoretical results.

The rest of the paper is organized as follows. Section II presents the model and our notation for the multiuser steganographic system. In Section III, we identify and describe the optimal signature set design and power allocation scheme. Section IV is devoted to experimental demonstrations. A few concluding remarks are drawn in Section V.

2. SYSTEM MODEL AND NOTATION

Consider a host image \( \mathbf{H} \in \mathcal{M}^{N_1 \times N_2} \) where \( \mathcal{M} \) is the image alphabet and \( N_1 \times N_2 \) is the image size in pixels. Without loss of generality, the image \( \mathbf{H} \) is partitioned* into \( P \) local blocks of size \( N_1 \times N_2/P \) pixels. Under the multiuser steganographic model, each block \( \mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_P \) is to carry \( K \) hidden information bits, one for each different user potentially. Embedding is performed in a real 2-D transform domain \( \mathcal{T} \). After transform calculation for each block and conventional zig-zag scanning vectorization, we obtain \( \mathcal{T}((\mathbf{H}_p)) \in \mathbb{R}^{N_1 \times N_2/P}, \ p = 1, 2, \ldots, P \). From the transform domain vectors \( \mathcal{T}((\mathbf{H}_p)) \) we choose a fixed subset of \( L \leq N_1 \times N_2/P \) coefficients (bins) to form the final host vectors \( \mathbf{x}_p \in \mathbb{R}^L, \ p = 1, 2, \ldots, P \); (for example, it is common and appropriate to exclude the dc coefficient from the host \( \mathbf{x}_p \)).

\( K \)-user/signature embedding is carried out in the transform domain by

\[
\mathbf{y} = \sum_{i=1}^{K} A_i b_i \mathbf{s}_i + \mathbf{x} + \mathbf{n}
\]

where \( b_1, b_2, \ldots, b_K \in \{ \pm 1 \} \) are the individual message bits embedded simultaneously in \( \mathbf{x} \) with corresponding amplitudes \( A_i > 0 \) and signatures \( \mathbf{s}_i \in \mathbb{R}^L, \parallel \mathbf{s}_i \parallel = 1, \ i = 1, 2, \ldots, K \); \( \mathbf{n} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_L) \) accounts in the model for possible external white Gaussian noise\(^1\) with \( \mathbf{I}_L \) being the size-\( L \) identity matrix.

The autocorrelation matrix of the transform-domain host data \( \mathbf{x} \) is an important statistical quantity for our developments and is defined as follows

\[
\mathbf{R}_x \triangleq E \{ \mathbf{x} \mathbf{x}^T \} = \frac{1}{P} \sum_{p=1}^{P} \mathbf{x}_p \mathbf{x}_p^T
\]

where \( E \{ \cdot \} \) denotes statistical expectation (here, with respect to \( \mathbf{x} \) over the image blocks) and \( \{ \cdot \}^T \) is the transpose operator. It is easy to verify that in general \( \mathbf{R}_x \neq \alpha \mathbf{I}_L, \alpha > 0 \); that is, \( \mathbf{R}_x \) is not constant-value diagonal or “white” in field language.

Under a statistical bit independence assumption across messages, the mean-square distortion of the original image due to the inserted whole multimage is

\[
\mathcal{D} = E \left\{ \left\| \sum_{i=1}^{K} A_i b_i \mathbf{s}_i \right\|^2 \right\} = \sum_{i=1}^{K} A_i^2.
\]

*Arbitrary segmentation of \( \mathbf{H} \), or part of it, into potentially overlapping blocks is certainly possible as well.

\(^1\)Additive white Gaussian noise is frequently viewed as a suitable model for quantization errors, channel transmission disturbances and/or image processing attacks.
The contribution of each individual embedded message bit $i$ to the composite signal is $A_ib_is_i$ and the mean-square distortion to the original host data $x$ due to the embedded message $i$ alone is

$$
\mathcal{D}_i = E\left\{ \|A_ib_is_i\|^2 \right\} = A_i^2, \quad i = 1, 2, \ldots, K.
$$

Assume, for example, that given $y$, message $j$, $j \in \{1, 2, \ldots, K\}$, is the message of interest. With signal of interest $A_jb_js_j$ and respective total disturbance $\sum_{i=1}^{K} A_ib_is_i + x + n$ from (1), the linear filter that operates on $y$ and offers maximum SINR at its output can be calculated as

$$
\mathbf{w}_{\text{maxSINR},j} = \arg \max_{\mathbf{w}} \frac{E\left\{ |\mathbf{w}^T (A_jb_js_j)|^2 \right\}}{E\left\{ |\mathbf{w}^T \left( \sum_{i=1,i\neq j}^{K} A_ib_is_i + x + n \right)|^2 \right\}} = \mathbf{R}_j^{-1}s_j
$$

where $\mathbf{R}_j$ denotes the “exclude-j” received data autocorrelation matrix, that is the autocorrelation matrix of the disturbance to message $j$, defined as

$$
\mathbf{R}_j \triangleq E\left\{ \left( \sum_{i=1,i\neq j}^{K} A_ib_is_i + x + n \right) \left( \sum_{i=1,i\neq j}^{K} A_ib_is_i + x + n \right)^T \right\} = \mathbf{R}_x + \sigma^2 \mathbf{I}_L + \sum_{i=1,i\neq j}^{K} A_i^2s_is_i^T.
$$

We can calculate the exact maximum output SINR value attained when we use the filter $\mathbf{w}_{\text{maxSINR},j}$ to

$$
\text{SINR}_j = A_j^2s_j^T \left( \mathbf{R}_x + \sigma^2 \mathbf{I}_L + \sum_{i=1,i\neq j}^{K} A_i^2s_is_i^T \right)^{-1} s_j = A_j^2s_j^T \mathbf{R}_j^{-1}s_j.
$$

Viewing $\text{SINR}_j$ in (7) as a function of $s_j$ that can be potentially further maximized, in eq. (36), an iterative multisignature design algorithm is described with no known/guaranteed optimality upon convergence. The signature $s_j$ is set at the minimum-eigenvalue eigenvector of $\mathbf{R}_j$, for each $j$ from 1 through $K$. Then, all disturbance autocorrelation matrices are updated and the recalculation of the signature set continues until numerical convergence is observed or for a predetermined number of cycles. In the following section of this paper, for any $K \leq L$, we present a one-shot sum-SINR and sum-capacity optimal signature set design algorithm that operates directly on the transform-domain host data autocorrelation matrix $\mathbf{R}_x$ of (2).

### 3. OPTIMAL MULTISIGNATURE EMBEDDING

We begin with a tedious -yet all important- algebraic manipulation of the maximum $\text{SINR}$ expression for user $j$, $j = 1, 2, \ldots, K$, in (7). Repeated $K - 1$ times application of the Matrix Inversion Lemma (also known as Woodbury’s Identity) on the inverse-matrix term of the equation leads to

$$
\text{SINR}_{j} = \text{SINR}^{(1)}_{j} - \sum_{n=1}^{K-1} T_j^{(n)}
$$

where

$$
\text{SINR}^{(1)}_{j} \triangleq A_j^2s_j^T (\mathbf{R}_x + \sigma^2 \mathbf{I}_L)^{-1} s_j,
$$

$$
T_j^{(n)} \triangleq \begin{cases} 
\frac{A_j^2|\rho|^{(n)}_j^2}{\left(1/A_{n+1}^2\right)+s_{n+1}^T (\mathbf{R}_x + \sigma^2 \mathbf{I}_L + \sum_{i=1,i\neq j}^{n} A_i^2s_is_i^T)^{-1} s_{n+1}}, & j \leq n \\
\frac{A_j^2|\rho|^{(n)}_j^2}{\left(1/A_{n}^2\right)+s_{n}^T (\mathbf{R}_x + \sigma^2 \mathbf{I}_L + \sum_{i=1}^{n-1} A_i^2s_is_i^T)^{-1} s_{n}}, & j > n 
\end{cases}
$$

Proc. of SPIE Vol. 6982 69820D-3
\[ \rho_{j}^{(n)} = \begin{cases} s_{j}^{T} \left( R_{x} + \sigma^{2}I_{L} + \sum_{i=1, i \neq j}^{n} A_{i}^{2} \sigma_{i}^{2} s_{i}^{T} \right)^{-1} s_{n+1}, & j \leq n \\ s_{j}^{T} \left( R_{x} + \sigma^{2}I_{L} + \sum_{i=1, i \neq j}^{n-1} A_{i}^{2} \sigma_{i}^{2} s_{i}^{T} \right)^{-1} s_{n}, & j > n. \end{cases} \] 

(11)

Let \( \{q_{1}, q_{2}, \ldots, q_{L}\} \) be the \( L \) eigenvectors of \( R_{x} \) with corresponding eigenvalues \( \lambda_{1} \geq \lambda_{2} \geq \cdots \geq \lambda_{L} \). Further examination of the formulas in (8)-(11) reveals the findings presented in the form of Proposition 1 below. The proof is omitted due to lack of space.

**Proposition 1** For orthonormal signature sets \( \{s_{1}, s_{2}, \ldots, s_{K}\}, K \leq L \), and corresponding fixed embedding amplitudes \( A_{1} \geq A_{2} \geq \cdots \geq A_{K} > 0 \) in (1), \( \sum_{j=1}^{K} SINR_{j}^{(1)} \) is maximized to \( \sum_{j=1}^{K} \frac{A_{j}^{2}}{\lambda_{j} - \sigma^{2} + \sigma \sigma} \) when \( s_{1}, s_{2}, \ldots, s_{K} \) are assigned as the \( K \) bottom eigenvectors of \( R_{x} \), i.e., \( s_{j} = \mathbf{q}_{L-(j-1)} \), \( j = 1, 2, \ldots, K \). At the same time, when \( s_{j} = \mathbf{q}_{L-(j-1)} \), \( j = 1, 2, \ldots, K \), \( T_{j}^{(n)} = 0 \) for every \( j = 1, 2, \ldots, K \) and \( n = 1, \ldots, K - 1 \).

We return to the problem of optimal multiuser steganography by (1) and consider the multiuser performance metrics sum SINR, defined as the sum of the individual SINRs of the \( K \) embedded messages,

\[ \text{sumSINR} \triangleq \sum_{j=1}^{K} \text{SINR}_{j}, \] 

(12)

and sum capacity, defined as the maximum sum of message coding rates at which the messages can be recovered reliably. If the transform-domain host data \( \mathbf{x} \) are assumed (colored) Gaussian, the sum capacity of the embedding scheme is\[ C_{\text{sum}} = \frac{1}{2} \sum_{j=1}^{K} \log_{2} \left( 1 + \text{SINR}_{j} \right). \] 

(13)

The following theorem, built on Proposition 1, establishes the optimal orthonormal embedding signature set.

**Theorem 1 (Optimal Multisignature Assignment)**

For orthonormal signature sets \( \{s_{1}, s_{2}, \ldots, s_{K}\}, K \leq L \), and corresponding fixed embedding amplitudes \( A_{1} \geq A_{2} \geq \cdots \geq A_{K} > 0 \) in (1), the sum SINR is maximized to \( \text{sumSINR}_{\text{max}} = \sum_{j=1}^{K} \frac{A_{j}^{2}}{\lambda_{j} - \sigma^{2} + \sigma \sigma} \) when \( s_{1}, s_{2}, \ldots, s_{K} \) are assigned as the \( K \) bottom eigenvectors of \( R_{x} \), i.e., \( s_{j} = \mathbf{q}_{L-(j-1)} \), \( j = 1, 2, \ldots, K \). If the transform-domain host data \( \mathbf{x} \) are Gaussian distributed, the same signature assignment maximizes sum capacity to \( (C_{\text{sum}})_{\text{max}} = \frac{1}{2} \sum_{j=1}^{K} \log_{2} \left( 1 + \frac{A_{j}^{2}}{\lambda_{j} - \sigma^{2} + \sigma \sigma} \right) \) bits per \( K \)-symbol embedding.

It is interesting to note that Theorem 1 establishes that the necessary condition for sum-capacity optimality under the constraint of eigenvector signature assignment in Lemma 1 of\[ 25 \] is also a sufficient condition for optimality for general orthonormal signature set design.

If we generalize our approach and view the individual amplitudes as design parameters themselves, then we can search for the optimal amplitude assignment that maximizes sum capacity subject to a total allowable distortion constraint (budget) \( D = \sum_{j=1}^{K} A_{j}^{2} \). The optimal amplitude values are derived in the theorem below. The proof follows from Theorem 1 and\[ 25 \], Lemma 2.

**Theorem 2 (Optimal Multisignature Assignment and Power Allocation)**

For orthonormal signature sets and a given total distortion budget \( D \), the sum-capacity optimal (signature, amplitude) pairs for multiuser embedding in (transform-domain) Gaussian hosts are

\[ s_{j} = \mathbf{q}_{L-(j-1)}, \quad A_{j}^{2} = \left[ -\left( \lambda_{j} - \sigma^{2} \right) + \mu \right]^{+}, \quad j = 1, 2, \ldots, K, \]
where \( [x]^+ \triangleq \max(x,0) \) and \( \mu \) is the Kuhn-Tucker coefficient chosen such that the distortion constraint \( D = \sum_{j=1}^{K} A_j^2 \) is met with equality.

The amplitude/power allocation method described in Theorem 2 can be viewed as a power “waterfilling” procedure in the eigen domain of the host.

4. EXPERIMENTAL STUDIES

To carry out an experimental study of the developments presented in the previous section, we consider as a host example the familiar grayscale 512×512 “Boat” image in Fig. 1(a). We perform 8×8 block DCT embedding over all bins except the dc coefficient (\( L = 63 \)) and hide \( K = 15 \) data messages of length \( 512^2 / 8^2 = 4096 \) bits each with each message having its own individual embedding signature. For the sake of generality, we also incorporate white Gaussian noise of variance \( \sigma^2 = 3 \) dB. Fig. 2 shows the resulting sum SINR as a function of the total distortion of the host \( \mathcal{D} \) over the 12 to 32dB range when embedding is carried out with arbitrary or optimal signatures by Theorem 1 and the individual message amplitudes/distortions are fixed at \( \mathcal{D}_1, \mathcal{D}_2 = \mathcal{D}_1 - 1dB, \ldots, \mathcal{D}_{15} = \mathcal{D}_{14} - 1dB \) (1dB decrease for each successive message). Fig. 1(b) shows the Boat at \( \mathcal{D} = 20dB \) total distortion. For this example, the gain in sum SINR by the use of the optimal signature set of Theorem 1 is at least 5dB and growing as the total allowed distortion increases.

In Fig. 3, we examine the sum capacity of the suggested multiuser steganographic scheme under arbitrary signature set design, signature optimization alone by Theorem 1, and optimal signature and power allocation by Theorem 2. At 32dB total distortion, an optimized signature set (with fixed per message distortion at 1dB differences) offers a gain of about 10 bits per embedding attempt over arbitrary signature sets. About 3 more bits are added when signature optimization is combined with optimal power allocation.

To address the need of experimental verification of highest credibility, we carried out the experiments of Figs. 2 and 3 over the whole USC-SIPI database of 44 miscellaneous images. Fig. 4 shows the average sum SINR versus total distortion for the database. Fig. 5 shows the sum capacity results. The average database findings of Figs. 4 and 5 are quite similar to the individual Boat results in Figs. 2 and 3.

5. CONCLUSIONS

We considered the problem of multiuser data hiding in transform-domain hosts (images in particular herein) and identified the orthonormal signature set that offers maximum sum SINR embedding for any fixed embedding amplitude values. We showed that the set is also sum capacity optimal in terms of bits per multiuser embedding under the assumption that the transform-domain host data are Gaussian. When there is flexibility in assigning amplitudes across users under a total host distortion constraint, we derived the user amplitude values that meet the total constraint and further maximize sum capacity.

REFERENCES


Figure 1. (a) Original Boat image example (512×512 grayscale). (b) Boat image after optimal multi-signature embedding ($K = 15$ messages of size 4096 bits each, total distortion 20dB, fixed per message distortion $D_{i+1} = D_i - 1dB$, $i = 1, \cdots, K - 1$, $\sigma_n^2 = 3dB$). (c) Boat image after optimal signature and power allocation ($K = 15$ messages of size 4096 bits each, total distortion 20dB, $\sigma_n^2 = 3dB$).

Figure 2. Sum SINR versus total distortion (Boat image, $K = 15$, $D_{i+1} = D_i - 1dB$, $i = 1, \cdots, K - 1$, $\sigma_n^2 = 3dB$).
Figure 3. Sum capacity versus total distortion (Boat image, $K = 15$, $\sigma^2 = 3dB$).

Figure 4. Sum SINR versus total distortion (average findings over USC-SIPI image database$^{43}$, $K = 15$, $D_{i+1} = D_i - 1dB$, $i = 1, \cdots, K - 1$, $\sigma^2 = 3dB$).
Figure 5. Sum capacity versus total distortion (average findings over USC-SIPI image database\textsuperscript{43}, $K = 15, \sigma^2 = 3dB$).