Active Spread-Spectrum Steganalysis for Hidden Data Extraction

Ming Li  
Dept. of Electrical Engineering  
State University of New York at Buffalo  
Buffalo, NY 14260  
miling@buffalo.edu

Michel Kulhandjian  
Dept. of Electrical Engineering  
State University of New York at Buffalo  
Buffalo, NY 14260  
mkk6@buffalo.edu

Dimitris A. Pados  
Dept. of Electrical Engineering  
State University of New York at Buffalo  
Buffalo, NY 14260  
pados@buffalo.edu

Stella N. Batalama  
Dept. of Electrical Engineering  
State University of New York at Buffalo  
Buffalo, NY 14260  
batalama@buffalo.edu

Michael J. Medley  
Air Force Research Laboratory/RITF  
525 Brooks Rd  
Rome, NY 13441  
michael.medley@rl.af.mil

ABSTRACT

This paper considers the problem of blind active spread-spectrum (SS) steganalysis defined as the extraction of hidden data with no prior information. We first develop a multi-signature iterative generalized least-squares (M-IGLS) core procedure to seek unknown messages hidden in image hosts via multi-signature direct-sequence spread-spectrum embedding. Neither the original host nor the embedding signatures are assumed available. Then, cross-correlation enhanced M-IGLS (CC-M-IGLS), a procedure described herein in detail that is based on statistical analysis of repeated independent M-IGLS processing of the host, is seen to offer most effective hidden message recovery. In fact, experimental studies show that the proposed CC-M-IGLS active SS steganalysis algorithm can achieve probability of error close to what may be attained with known embedding signatures and host autocorrelation matrix.

Categories and Subject Descriptors
D.2.11 [Software Engineering]: Software Architectures—Information hiding

General Terms
Security, Algorithms, Theory

Keywords
Blind detection, covert communications, data hiding, spread-spectrum embedding, steganalysis, steganography, water-marking

1. INTRODUCTION

Steganography, which literally means “covered writing” in Greek, is the process of hiding data under a cover medium (also referred to as host), such as image, video, or audio [1]-[3]. The basic purpose of steganography is to establish covert communication between trusting parties. While other data hiding applications (such as watermarking [4]-[6]) have their own individual requirements, the broad common objective of most steganographic applications is a satisfactory trade-off between hidden data resistance to noise/disturbance (robustness), information delivery rate (payload), and low host distortion for concealment purposes.

Steganalysis, which is the countermeasure technology to steganography, aims to discover the presence and/or extract the content of the secret data. Accordingly, steganalysis can be classified into two categories [7], passive and active. The primary task of passive steganalysis is to decide the presence or absence of hidden messages in given media objects. In contrast, active steganalysis refers to the effort of extracting the actual hidden data. While passive steganalysis is being intensively investigated in the past few years [9]-[17], active steganalysis is a relatively new branch of research. To our best knowledge, there seems to have been little attempt in developing active steganalysis methods that can blindly extract the secret data.

In this work, we focus our attention on active spread-spectrum (SS) steganalysis. In particular, we aim to recover blindly secret data hidden in image hosts via (multi-signature) direct-sequence SS embedding [18]-[25]. Neither the original host nor the embedding signatures (spreading sequences) are known (fully blind SS steganalysis). In blind active SS steganalysis the unknown host acts as a source of interference/disturbance to the data to be extracted and, in a way, the problem parallels blind signal separation (BSS) applications as they arise.

In another interpretation of active steganalysis, the steganalyst manipulates the embedded data, such as introducing noise, in hopes of destroying the secret message (if any) [8].
This paper considers the problem of blind active spread-spectrum (SS) steganalysis defined as the extraction of hidden data with no prior information. We first develop a multisignature iterative generalized least-squares (M-IGLS) core procedure to seek unknown messages hidden in image hosts via multi-signature direct-sequence spread-spectrum embedding. Neither the original host nor the embedding signatures are assumed available. Then, cross-correlation enhanced MIGLS (CC-M-IGLS), a procedure described herein in detail that is based on statistical analysis of repeated independent M-IGSL processing of the host, is seen to offer most effective hidden message recovery. In fact, experimental studies show that the proposed CC-M-IGLS active SS steganalysis algorithm can achieve probability of error close to what may be attained with known embedding signatures and host autocorrelation matrix.
in the fields of array processing, biomedical signal processing, and code-division multiple-access (CDMA) communication systems. Under the assumption that the embedded secret messages are independent identically distributed (i.i.d.) random sequences and independent to the cover host, independent component analysis (ICA) -one particular family of BSS methods- may be utilized to approach the hidden data extraction problem \cite{7,26}. However, ICA-based BBS algorithms degrade rapidly in the presence of correlated signal interference as is the case in SS image embedding. In \cite{27}, Gkizeli et al. developed an iterative generalized least squares (IGLS) procedure to blindly recover unknown messages hidden in image hosts via SS embedding. The algorithm has low complexity and remarkably good recovery performance. However, the scheme is designed solely for single-signature SS embedding where messages are hidden with one signature only. Realistically, a steganographer would favor multi-signature SS embedding to increase security and payload rate. The work in \cite{27} is not generalizable to the multi-signature case.

In this paper, we develop a new multi-signature iterative generalized least squares (M-IGLS) SS steganalysis algorithm for hidden data extraction. For improved recovery performance, in particular for small hidden messages that pose the greatest challenge, we propose an algorithmic upgrade referred to as cross-correlation enhanced M-IGLS (CC-M-IGLS). CC-M-IGLS relies on statistical analysis of independent M-IGLS executions on the host and experimental studies indicate that can achieve hidden data recovery with probability of error close to what may be attained with known embedding signatures and known original host autocorrelation matrix.

The rest of the paper is organized as follows. In Section 2 we present the signal model for the multi-signature SS embedding procedure and formulate the problem of active SS steganalysis. After developing the hidden data extraction algorithms in Section 3, experimental studies are presented in Section 4. Finally, some concluding remarks are drawn in Section 5.

The following notation is used throughout the paper. Boldface lower-case letters indicate column vectors and boldface upper-case letters indicate matrices; \( \mathbb{R} \) denotes the set of all real numbers; \( (\cdot)^T \) is the transpose operator; \( \mathbf{I}_L \) is the \( L \times L \) identity matrix; \( \text{sgn}\{\cdot\} \) denotes zero-threshold quantization and \( \mathbb{E}\{\cdot\} \) represents statistical expectation. Finally, \( |\cdot|, \|\cdot\|, \|\cdot\|_F \) are the scalar magnitude, vector norm, and matrix Frobenius norm, respectively.

2. MULTI-SIGNATURE SS EMBEDDING AND STEGANALYSIS PROBLEM FORMULATION

Consider a host image \( \mathbf{H} \in \mathcal{M}^{N_1 \times N_2} \) where \( \mathcal{M} \) is the finite 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 \( M \) local non-overlapping blocks of size \( \frac{N_1 N_2}{M} \). Each block, \( \mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_M \), is to carry \( K \) hidden information bits coming -potentially- from \( K \) distinct messages. Embedding is performed in a 2-D transform domain \( \mathcal{T} \) (such as the discrete cosine transform, a wavelet transform, etc.). After transform calculation and vectorization (for example by conventional zig-zag scanning), we obtain \( \mathcal{T}(\mathbf{H}_m) \in \mathbb{R}^{8 \times 8}, m = 1, 2, \ldots, M \). From the transform domain vectors \( \mathcal{T}(\mathbf{H}_m) \) we choose a fixed subset of \( L \leq 8 \times 8 \) coefficients (bins) to form the final host vectors \( \mathbf{x}(m) \in \mathbb{R}^L, m = 1, 2, \ldots, M \). It is common and appropriate to avoid the dc coefficient (if applicable) due to high perceptual sensitivity in changes of the dc value.

The autocorrelation matrix of the host data \( \mathbf{x} \) is an important statistical quantity for our developments and is defined as \( \mathbf{R}_x \equiv \mathbb{E}\{\mathbf{x}\mathbf{x}^T\} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{x}(m)\mathbf{x}(m)^T. \) 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. For example, \( 8 \times 8 \) DCT with 63-bin host data formation (excluding only the dc coefficient) for the \( 256 \times 256 \) gray-scale Baboon image in Fig. 1(a) gives the host autocorrelation matrix \( \mathbf{R}_x \) in Fig. 1(b).

2.1 Multi-signature SS Embedding

The \( K \) distinct message bit sequences \( \{b_k(m)\}_{m=1}^{M} \), \( k = 1, 2, \ldots, K \), \( b_k(m) \in \{-1\} \), are hidden in the transform-domain host vectors \( \{\mathbf{x}(m)\}_{m=1}^{M} \) via additive SS embedding by means of \( K \) spreading sequences (signatures) \( \mathbf{s}_k \in \mathbb{R}^L \).
amplitude-including embedding signatures. Then, we can further rewrite SS embedding as
\[ y(m) = \sum_{k=1}^{K} b_k(m) v_k + z(m) \quad (6) \]
\[ = V b(m) + z(m), \quad m = 1, \ldots, M, \quad (7) \]
where \( V \triangleq [v_1, \ldots, v_K] \in \mathbb{R}^{L \times K} \) is the amplitude-including signature matrix and \( b(m) \in \{ \pm 1 \}^{K \times 1} \) is the vector of bits embedded in the \( m \)th host block. For notational simplicity, we can write the whole stego image data as one matrix
\[ Y = V B + Z \quad (8) \]
where \( Y \triangleq [y(1) \ldots y(M)] \in \mathbb{R}^{L \times M}, B \triangleq [b(1) \ldots b(M)] \in \{ \pm 1 \}^{K \times M}, \) and \( Z \triangleq [z(1) \ldots z(M)] \in \mathbb{R}^{L \times M}. \)

3. **ACTIVE STEGANALYSIS FOR HIDDEN DATA EXTRACTION**

If \( Z \) were to be modeled as Gaussian distributed, the joint maximum-likelihood (ML) estimator of \( V \) and detector of \( B \) would be
\[ \hat{V}, \hat{B} = \arg \min_{V \in \mathbb{C}^{L \times K}, B \in \{ \pm 1 \}^{K \times M}} \| R_x^{\frac{1}{2}} (Y - V B) \|^2_F \quad (9) \]
where multiplication by \( R_x^{\frac{1}{2}} \) can be interpreted as prewhitening of the compound observation data. If Gaussianity of \( Z \) is not to be invoked, then (9) is simply referred to as the joint generalized least-squares (GLS) solution \(^6\) of \( V \) and \( B \).

3.1 **Multi-signature Iterative Generalized Least-Squares Procedure**

The global GLS-optimal message matrix \( \hat{B} \) in (9) can be computed independently of \( \hat{V} \) by exhaustive search over all possible choices under the criterion function \( \| R_x^{\frac{1}{2}} Y P_B \|^2_F \),
\[ \hat{B} = \arg \min_{B \in \{ \pm 1 \}^{K \times M}} \| R_x^{\frac{1}{2}} Y P_B \|^2_F \quad (10) \]
where \( P_B \triangleq I - B (B B^T)^{-1} B \). Exhaustive search has, of course, complexity exponential in \( KM \) (total size of hidden messages in bits). We consider this cost unacceptable and attempt to reach a quality approximation of the solution of (10) (or (9), to that respect) by alternating generalized least-squares estimates of \( V \) and \( B \), iteratively, as described below.

Pretend \( B \) is known; the generalized least-squares estimate
\(^6\)Generalized-least-squares solutions are weighted least-squares (WLS) solutions with optimal weighting matrices, here \( R_x^{-\frac{1}{2}} \), that yield the lowest variance of the estimation error \([31],[34]\).
of $\mathbf{V}$ is

$$
\hat{\mathbf{V}}_{\text{GLS}} = \arg \min_{\mathbf{V} \in \mathbb{R}^{K \times M}} \| \mathbf{R}_z \odot \mathbf{V} - \mathbf{Y} \|_F^2
$$

$$
\mathbf{B}_{\text{IGLS}} = \arg \min_{\mathbf{B} \in \mathbb{R}^{K \times M}} \| \mathbf{R}_z \odot \mathbf{B} \|_F^2
$$

It is always possible but the challenge is to ensure that the iterative scheme will lead us to a reliable estimate of $\mathbf{B}$ over the real field.

The multi-signature iterative generalized least-squares (M-IGLS) procedure described in Table 1 to the optimal ALS solution of (9) is not guaranteed in general. Extensive experimentation with the algorithm in Table 1 shows that the M-IGLS procedure is always possible but the challenge is to ensure that the iterative scheme will lead us to a reliable estimate of $\mathbf{B}$ over the real field.

We suggest the approximate binary message solution

$$
\hat{b}_{i,j} = \begin{cases} 
1 & \text{if } \eta_{i,j} > \epsilon \ \text{and} \ \eta_{i,j} < \epsilon \\
0 & \text{otherwise}
\end{cases}
$$

where $\epsilon$ is very low at about 0.03. Motivated by this, we introduce below Criterion 1 that classifies convergence points of the M-IGLS procedure in Table 1 as “compliant” or not based on the sample statistics of the returned data matrix $\mathbf{B}$.

**Criterion 1:** If $|\eta_{i,j}| \leq \frac{3}{\sqrt{6}}$ for all $i \neq j \in \{1, 2, \ldots, K\}$, then $(\hat{\mathbf{B}}, \hat{\mathbf{V}})$ returned by the M-IGLS procedure in Table 1 are classified as “Criterion-1-compliant.”

Table 1: Iterative generalized least-squares SS steganalysis

<table>
<thead>
<tr>
<th>Step</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$d := 0$; initialize $\hat{\mathbf{B}}^{(0)} \in { \pm 1 }^{K \times M}$ arbitrarily.</td>
</tr>
<tr>
<td>2</td>
<td>$d := d + 1$;</td>
</tr>
<tr>
<td></td>
<td>$\hat{\mathbf{V}}^{(d)} := \mathbf{Y}^{(d-1)} \mathbf{R}<em>{\text{sgn}}^{(d-1)} \mathbf{B}</em>{\text{IGLS}}^{(d-1)} \mathbf{Y}^{(d-1)}$</td>
</tr>
<tr>
<td>3</td>
<td>Repeat Step 2 until $\hat{\mathbf{B}}^{(d)} \approx \mathbf{B}_{\text{IGLS}}^{(d-1)}$.</td>
</tr>
</tbody>
</table>

is how to assess whether solutions returned by the M-IGLS procedure are reliable or not without any side information. The rest of this section is devoted to addressing this challenge.
eliminate erroneous solutions. Yet, there exists a tight cluster/region formed by 210 or so of the Criterion-1-equipped M-IGLS convergence points around the true embedding signature.

The basic idea now behind our second and final refinement of the M-IGLS blind hidden data extraction procedure is to identify and average these reliable clustered estimates. Of course, identification of the reliable estimates is not a trivial task due to our complete lack of knowledge of \( v_k \) (or \( s_k \)), \( k = 1, \ldots, K \). In this context, assume that we have \( P \) estimates of \( v_k \) denoted by \( \hat{v}_k^{(j)} \), \( j = 1, \ldots, P \), obtained by \( P \) runs of the Criterion-1-equipped M-IGLS procedure. From the example of Fig. 2, we understand that reliable estimates \( \hat{v}_k^{(j)} \) of \( v_k \) have high normalized cross-correlation (close to 1) with each other, while they will have low normalized cross-correlation with other unreliable estimates of \( v_k \). In contrast, unreliable estimates will tend to have low normalized cross-correlation with each other. Therefore, the reliability of \( \hat{v}_k^{(j)} \) may be quantified/assessed by examining the sum-cross-correlation with the other \( \hat{v}_k^{(t)} \), \( t \neq j \in \{1, \ldots, P\} \),

\[
\rho_k^{(j)} \triangleq \frac{1}{P} \sum_{t=1,t\neq j}^{P} \frac{\hat{v}_k^{(j)}\hat{v}_k^{(t)}}{\|\hat{v}_k^{(j)}\|\|\hat{v}_k^{(t)}\|}, \quad (17)
\]

A reasonable threshold value for binary reliability classification may be the average value

\[
\bar{\rho}_k \triangleq \frac{1}{P} \sum_{j=1}^{P} \rho_k^{(j)}, \quad k = 1, \ldots, K, \quad (18)
\]

utilized in the proposed Criterion 2 below.

**Criterion 2:** Let \( \hat{v}_k^{(j)} \) be the estimates of \( v_k \) returned by \( P \) arbitrary initializations of the Criterion-1-equipped M-IGLS procedure of Table 1, \( k = 1, \ldots, K \), \( j = 1, \ldots, P \). If \( \bar{\rho}_k \geq \bar{\rho}_k \), then \( \hat{v}_k^{(j)} \) is considered a reliable estimate of the \( v_k \); otherwise we declare it as unreliable.

<table>
<thead>
<tr>
<th>Table 2: Cross-correlation Enhanced M-IGLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>For ( j := 1 ) to ( P )</td>
</tr>
<tr>
<td>1) Execute M-IGLS of Table 1 with arbitrary initialization and obtain estimates ( \hat{v}_k ), ( k = 1, \ldots, K ).</td>
</tr>
<tr>
<td>2) If ( \hat{v}_k ) is Criterion-1-compliant,</td>
</tr>
<tr>
<td>( \hat{v}_k^{(j)} := \hat{v}_k ), ( k = 1, \ldots, K );</td>
</tr>
<tr>
<td>else go to 1).</td>
</tr>
<tr>
<td>End</td>
</tr>
<tr>
<td>For ( k := 1 ) to ( K )</td>
</tr>
<tr>
<td>3) Identify reliable estimates for ( v_k ) according to Criterion 2.</td>
</tr>
<tr>
<td>4) Calculate the average over all reliable estimates ( \overline{v}_k ) by (19).</td>
</tr>
<tr>
<td>End</td>
</tr>
<tr>
<td>5) Set ( \overline{V} \triangleq [\overline{v}_1, \ldots, \overline{v}_K] ).</td>
</tr>
<tr>
<td>6) Execute M-IGLS of Table 1 with initialization</td>
</tr>
<tr>
<td>( \overline{B}^{(0)} = \text{sgn}\left{ \left( \overline{V}^{T} \overline{R}_y^{-1} \overline{V} \right)^{-1} \overline{V}^{T} \overline{R}_y^{-1} \overline{Y} \right} ).</td>
</tr>
<tr>
<td>Finally, we average our reliable (according to Criterion 2) estimates of the effective signatures ( v_k ) to produce one last high-quality initialization of the M-IGLS algorithm of Table 1. Let ( S_k ) denote the set of all reliable estimates of ( v_k ) according to Criterion 2 and let (</td>
</tr>
</tbody>
</table>

\[
\overline{V} \triangleq [\overline{v}_1, \ldots, \overline{v}_K] \quad \text{where} \quad \overline{v}_k = \frac{1}{|S_k|} \sum_{j \in S_k} \hat{v}_k^{(j)}, \quad k = 1, \ldots, K, \quad (19)
\]

i.e. \( \overline{v}_k \) is the average over all reliable estimates of \( v_k \) according to Criterion 2. We execute M-IGLS in Table 1 a final time initialized at \( \overline{B}^{(0)} = \text{sgn}\left\{ \left( \overline{V}^{T} \overline{R}_y^{-1} \overline{V} \right)^{-1} \overline{V}^{T} \overline{R}_y^{-1} \overline{Y} \right\} \). |

We call M-IGLS with both Criteria 1 and 2 incorporated, Cross-Correlation enhanced M-IGLS (CC-M-IGLS) and summarize the complete procedure in Table 2.

| Figure 2: Histogram of normalized cross-correlation between \( \hat{v}_1 \) and \( v_1 \) (256 \times 256 Baboon image, 8 \times 8 DCT), \( L = 63, K = 4, D_k = 31.5\text{dB}, k = 1, 2, \ldots, 4, \sigma^2_\text{w} = 3\text{dB}; \hat{v}_1 \) returned by Table 1 M-IGLS steganalysis procedure). |

4. EXPERIMENTAL STUDIES

A technically firm and keen measure of quality of an active steganalysis solution is the difference in the bit-error-rate (BER) experienced by the intended recipient and the steganalist. The intended recipient in our studies may be using any of the following three message recovery methods: (i) Standard signature matched-filtering (MF) with the known signatures \( s_k \), \( k = 1, \ldots, K \), (ii) sample-matrix inversion MMSE (SMI-MMSE) filtering with known signatures \( s_k \) and estimated true host autocorrelation matrix \( \hat{R}_y \) (see (3)); (iii) ideal MMSE filtering with known signatures \( s_k \) and known true host autocorrelation matrix \( R_y \) which serves as the ultimate performance bound reference for all methods. In terms of blind active steganalysis (neither \( s_k \) nor \( R_y \) known), we will examine (iv) the developed M-IGLS algorithm in Table 1 alone and (v) CC-M-IGLS of Table 2 with \( P = 20 \) Criterion-1 runs. Finally, the performance of two typical ICA-based blind signal separation (BSS) algorithms, (vi) FastICA [35], and (vii) JADE [36], will also be included in the studies for comparison purposes.
We first consider as a host example the gray-scale 512 × 512 “Baboon” image. We perform 8 × 8 block DCT embedding by (1) over all bins except the dc coefficient with $K = 4$ distinct arbitrary signatures $s_k \in \mathbb{R}^{63}$, $k = 1, \ldots, K$. The hidden message embedded by each signature is $512^2 / 8^2 = 4,096$ bits long. The per-message mean square distortion due to each embedded message is set to be the same for all messages, i.e. $D_k = A_k^2 = D$, $k = 1, \ldots, 4$. For the sake of generality, we also incorporate white Gaussian noise of variance $\sigma_n^2 = 3$dB. Fig. 3 shows the average BER (over all $K = 4$ messages) of all methods (i) through (vi) listed above as a function of the host distortion per message. While the independent/principal-component methods (FastICA and JADE) are failing to carry out effective active SS image steganalysis, to our satisfaction CC-M-IGLS blind active steganalysis is rather close in BER performance to the ideal MMSE detector bound where both the embedding signatures and the clean host autocorrelation matrix $R_x$ are perfectly known. It could be argued that for this host and rather large size of $M = 4,096$ bits per message, CC-M-IGLS offers a moderate gain only in comparison with M-IGLS of Table 1 by itself.

In Fig. 4, however, we repeat the exact same experimental study on the smaller 256 × 256 version of the Baboon image Fig. 1(a) with $K = 4$ hidden messages of length only $256^2 / 8^2 = 1,024$ bits per message. CC-M-IGLS now provides dramatic performance improvement over M-IGLS which surely justifies the extra computational cost and extraction delay. At the same time, comparing with Fig. 3, the gap between CC-M-IGLS and ideal MMSE increases as the hidden message size (use of signature, individually) decreases.

For additional experimental validation, the studies of Fig. 3 and Fig. 4 are repeated on the familiar “Boat” image (shown in Fig. 5) in its 512 × 512 and 256 × 256 gray-scale versions (Fig. 6 and Fig. 7, correspondingly). Identical conclusions are drawn regarding the effectiveness of CC-M-IGLS blind active steganalysis.

Finally, to examine the behavior of CC-M-IGLS under increased-density small-message hiding, we consider the 256 × 256 gray-scale “F-16 Aircraft” image (shown in Fig. 8) with $K = 4$ or $K = 8$ hidden messages of length 1Kbit each. Recovery performance plots are given in Fig. 9 and Fig. 10, correspondingly. An encompassing conclusion over all executed experiments is that CC-M-IGLS remains a most effective technique to extract blindly hidden messages, while extraction becomes more challenging as the length of hidden messages (use of an embedding signature) decreases or the number of hidden messages (number of used signatures) increases.

5. CONCLUSIONS

In this paper we considered the problem of active blind spread-spectrum steganalysis and attempted to recover unknown messages hidden in image hosts via multi-signature spread-spectrum embedding. Neither the original host nor the embedding signatures are assumed available. We first
developed a low complexity multi-signature iterative generalized least-squares (M-IGLS) core algorithm. Cross-correlation enhanced M-IGLS (CC-M-IGLS), a procedure based on statistical analysis of repeated independent M-IGLS processing of the host, offers most effective blind hidden message recovery. In fact, experimental studies showed that CC-M-IGLS can achieve probability of error rather close to what may be attained with known embedding signatures and known original host autocorrelation matrix and present itself as an efficient countermeasure to conventional\textsuperscript{5} SS steganography.

\textsuperscript{5}In [26], Bas and Cayre present an interesting signature-based additive embedding approach different to (1) that is host-vector-by-host-vector dependent and would withstand IGLS-based active steganalysis. The embedding is, however, very sensitive to noise that would lead to high recovery error rates by intended recipients and limit the applicability to general covert communication problems.
6. REFERENCES


APPENDIX

Proof of (11)
The GLS cost function in (9) can be rewritten as
\[
J = \|R_s^{-1}Y - R_s^{-1}VB\|^2_2
= \text{tr}\{R_s^{-1}YY^T\} - \text{tr}\{R_s^{-1}YVBV^T\} - \text{tr}\{R_s^{-1}VBY^T\} + \text{tr}\{R_s^{-1}VBB^TV^T\} \tag{21}
\]
where \(\text{tr}\{\cdot\}\) denotes the trace of a matrix.

For a given message matrix \(B\), the GLS optimal estimate of \(V\) can be obtained by differentiating the cost function \(J\) with respect to \(V^T\) and setting the outcome equal to the zero matrix,
\[
\frac{\partial J}{\partial V} = -R_s^{-1}YB + R_s^{-1}VBB^TV^T = 0, \tag{22}
\]
where \(\frac{\partial J}{\partial V} = \text{tr}\{R_s^{-1}YV^T\} - \text{tr}\{R_s^{-1}V^T(VBB^TV^T)\} - \text{tr}\{R_s^{-1}YVBV^T\} + \text{tr}\{R_s^{-1}VBB^TV^T\}. \tag{23}
\]

Proof of (12)
We manipulate the GLS cost function in the form of (21) to write
\[
J = \text{tr}\{R_s^{-1}YY^T\} - \text{tr}\{V^TR_s^{-1}YVB\} - \text{tr}\{R_s^{-1}VBY^T\} + \text{tr}\{V^TR_s^{-1}VBB^TV^T\}. \tag{24}
\]

Proof of (13)
Since \(R_y = \mathbb{E}\{yy^T\} = VV^T + R_s\), by the Matrix Inversion Lemma (also known as Woodbury’s Identity [37]), we can obtain
\[
R_y^{-1} = R_s^{-1} - R_s^{-1}V(I + V^TR_s^{-1}V)^{-1}V^TR_s^{-1}V. \tag{27}
\]
Then,
\[
V^TR_y^{-1}V = V^TR_s^{-1}V - V^TR_s^{-1}V(I + V^TR_s^{-1}V)^{-1}V^TR_s^{-1}V
= V^TR_s^{-1}V(I - (I + V^TR_s^{-1}V)^{-1}V^TR_s^{-1}V)
= V^TR_s^{-1}V(I + V^TR_s^{-1}V)^{-1}\]
\[
[I + V^TR_s^{-1}V] - V^TR_s^{-1}V
= V^TR_s^{-1}V(I + V^TR_s^{-1}V)^{-1}. \tag{28}
\]
By the property of the inverse of a product of matrices [37],
\[
(V^TR_y^{-1}V)^{-1} = (I + V^TR_s^{-1}V)(V^TR_s^{-1}V)^{-1}
= (V^TR_s^{-1}V)^{-1} + I. \tag{29}
\]
We combine the results of (27) and (29) and finally obtain
\[
(V^TR_s^{-1}V)^{-1}V^TR_s^{-1}V = ((V^TR_s^{-1}V)^{-1} + I) V^T \]
\[
\left(R_s^{-1} - R_s^{-1}V(I + V^TR_s^{-1}V)^{-1}V^TR_s^{-1}V\right) \tag{30}
= (V^TR_s^{-1}V)^{-1}V^TR_s^{-1}V. \tag{31}
\]