

Data Masking: A New Approach for Data Hiding?

Regunathan Radhakrishnan Mehdi Kharrazi

Nasir Memon

Polytechnic University
Brooklyn, NY 11201, USA

Email: regu@vip.poly.edu, mehdi@vip.poly.edu, memon@vip.poly.edu

Abstract

It is well known that encryption provides secure channels for communicating entities. However, due to lack of covertness on these channels, an eavesdropper can identify encrypted streams through statistical tests and capture them for further cryptanalysis. Hence, the communicating entities can use steganography to achieve covertness. In this paper we propose a new form of multimedia steganography called *data masking*. Instead of embedding a secret message into a multimedia object, as in traditional multimedia steganography, we process the entire secret message using an inverse Wiener filter to make it look like a multimedia object itself. Thereby we foil an eavesdropper who is primarily applying statistical tests to detect encrypted communication channels. We show that our approach can potentially give a covert channel capacity which is an order of magnitude higher than traditional steganography.

I. INTRODUCTION

Rapid developments in cryptography, cryptographic standards, and relaxed export controls on encryption schemes have made cryptographic software widely available, user friendly and in some cases transparent to end-users. The relative ease with which secure channels of communication can be created between two parties using cryptography, coupled with the explosion in network traffic, has posed a problem to the surveillance operations typically carried out by law enforcement and intelligence agencies. In order to keep up with the growth in communication traffic and secure channels, massive eavesdropping campaigns, such as *Carnivore* and *Echelon*, have been developed. *Carnivore*, for example, is a combination of hardware and software, which collects all electronic packets that are sent to and from the Internet Service Provider where it is installed. Following this "collection" procedure, several filters are applied to isolate emails with certain keywords which could be further analyzed off-line. These filters could potentially include the performance of statistical tests to isolate encrypted emails for off-line cryptanalysis.

Pervasive eavesdropping like that performed by Carnivore leads to the need for covert channels, in which communicating entities hide their communication channel itself from eavesdroppers. *Steganography* refers to the science of "invisible" communication. Unlike cryptography, where the goal is to secure communications from an eavesdropper, steganographic techniques strive to hide the very presence of the message itself from an observer [?]. As broadband technologies improve bandwidth at the last-mile, multimedia content, such as still images, audio and video, have gained increasing popularity on the Internet. Given the high degree of redundancy present in a digital representation of multimedia content (despite compression), there has been an increased interest in using multimedia content for the purpose of steganography. Indeed many such techniques have been devised in the past few years. For an excellent survey of such techniques, the reader is referred to [?]. Briefly, the general model for steganography, illustrated in Figure ??, where we have Alice wishing to send a secret message m to Bob. In order to do so, she "embeds" m into a *cover-object* c , to obtain the *stego-object* s . The stego-object s is then sent through the public channel. Wendy who examines all messages in the channel, should not be able to distinguish in any sense between cover-objects (objects not containing any secret message) and stego-objects (objects containing a secret message). In this context, *steganalysis* refers to the body of techniques that aid Wendy in distinguishing between cover-objects and stego-objects. It should be noted that Wendy has to make this distinction without any knowledge of the secret key which Alice and Bob may be sharing and sometimes even without any knowledge of the specific algorithm that they might be using for embedding the secret message.

Fig. 1. Framework for Secret Key Passive Warden Steganography. Alice embeds secret message in cover image (left). Wendy the warden checks if Alice's image is a stego-image (center). If she cannot determine it to be so, she passes it on to Bob who retrieves the hidden message based on secret key (right) he shares with Alice.

There are two problems with multimedia steganography that seem to have been ignored in the literature. Firstly, in multimedia steganography, the size of the covert message is relatively much smaller than the size of the multimedia object (the stego object) that carries the covert message. In fact recent developments in steganalysis techniques have demonstrated that embedding a secret message in a multimedia object, like say

a still image, causes certain statistically discernible artifacts that reveal the presence of the secret message thereby exposing the covert communication channel [?],[?],[?],[?]. Such developments further limit the number of bits that can safely be hidden in a multimedia object from sophisticated steganalysis techniques. This leads to a cover object that is one or more orders of magnitude larger than the secret message, reducing the effective bandwidth of covert communication channels.

The second problem is that the common interpretation of the warden-based framework when used for multimedia steganography described has been that the warden examines the cover-object (or stego-object) perceptually *and* statistically. However, perceptual tests, which may have been practical on low volume communication channels, would not be practical with high volume communication channels. Such channels must rely on automated statistical tests for detecting potential stego objects and only then may apply (if at all) some form of perceptual tests to the selected candidates. Indeed, large scale eavesdropping operations like Carnivore rely first on statistical analysis to isolate "interesting messages". Hence, if Alice and Bob can evade statistical tests performed by Wendy, they may not have to undergo any perceptual test at all! So for example, if the stego object is statistically indistinguishable from an image, its perceptual quality does not matter at all. Can this fact be exploited to improve covert communications in any sense? In this paper we show that this is indeed possible.

Specifically, we propose a novel technique, which we call *Data Masking*, for multimedia steganography. We again look at the problem of Alice and Bob wishing to communicate covertly, but we assume that the warden Wendy is an entity like Carnivore and will first examine messages statistically. Only if the statistical tests indicate a possibility of covert communication, would she then examine the message further. Hence instead of selecting a cover object and embedding a covert message in it like traditional steganography, we mask the secret message to make it appear like a normal multimedia object under typical statistical examinations.

To demonstrate our approach we assume the covert message is encrypted and hence a random bit stream. Since a random bit string can be easily identified by Wendy using simple statistical tests, we process the random bit stream to make it appear like a audio signal or image, in a statistical sense. This process removes randomness from the encrypted message, and makes it difficult for Wendy to distinguish it from a typical

audio stream or image. We refer to the process of converting a random bitstream into audio and image as “audio datamasking” and “image datamasking” respectively. The rest of the paper is organized as follows: in the following section we describe the proposed system for audio and image datamasking. Section 3 presents some experimental results and possible improvements and we conclude in Section 4 with future work.

II. PROPOSED SYSTEM

If Alice wants to send an encrypted message to Bob, the warden Wendy would be able to detect such a message as an encrypted stream since it would exhibit properties of randomness. In order for a secure channel to achieve covertness, it is necessary to preprocess the encrypted stream at the end points to remove randomness such that the resulting stream defeats statistical tests for randomness and the stream is reversible at the other end. In this section, we propose schemes that make the processed encrypted streams appear like more correlated multimedia streams (audio and image).

A. Audio Datamasking

We propose the following two methods for audio datamasking.

- Audio datamasking using spectral factorization.
- Audio datamasking using LPC analysis-synthesis.

Fig. 2. Audio datamasking using spectral factorization

1) *Audio datamasking using spectral factorization:* Audio datamasking using spectral factorization is illustrated in Fig.???. Let us consider the cipher stream as samples from a Wide Sense Stationary (WSS) Process, E . We would like to transform this input process with high degree of randomness to another stationary process, A , with more correlation between samples by using a linear filter, H . It is well known that Power spectrum of a input WSS Process, $A(w)$, to a linear time invariant system and Power spectrum of the corresponding output Process, $E(w)$ are related by the following equation:

$$E(w) = |H(w)|^2 A(w) \quad (1)$$

If $E(w)$ is a white noise process, then $H(w)$ is the whitening filter or Wiener filter. Since the encrypted stream is random, its power spectral density is flat and resembles the power spectral density of a white noise process. Then, the desired Wiener filter can be obtained by spectral factorization of $(E(w)/A(w))$ followed by selection of poles and zeros to obtain the minimum phase solution for $H(w)$. Since the factor $(E(w)/A(w))$ can have arbitrary shape, it would require a filter of very high order for realization.

2) *Audio datamasking using LPC analysis-synthesis:* Audio datamasking can also be performed using LPC Analysis/Synthesis as shown in Figure ???. The LPC Analysis filter for reference Audio clip, A , is obtained as follows. Let $X_0, X_1, X_2 \dots X_{N-1}$ represent N previous samples of the reference audio clip. The goal is to obtain the filter coefficients $h_0, h_1, h_2 \dots h_{N-1}$ such that $\sum E((X_i - \widehat{X}_i)^2)$ is minimized. Here \widehat{X}_i is the predicted value of the current sample based on N previous samples in the reference audio and is defined as $\sum_{k=0}^{(N-1)} h_k X_k$.

Using the orthogonality principle (Hilbert space projection theorem), N equations (called Yule-Walker equations) can be set up to solve for the optimal filter coefficients in the MMSE (Minimum Mean Square Error) sense. Then, the inverse of the LPC analysis filter so designed, can be used to filter the noise-like cipher stream to remove randomness from cipher stream and transform it into a reference audio-like waveform that has more correlation between samples.

Fig. 3. Audio Data Masking Using LPC Analysis/Synthesis

With the knowledge of filter coefficients the receiver can reconstruct the cipher stream from the reference audio, as in the Inverse Wiener filtered cipher stream.

B. Image Datamasking

Since the source model for speech is well understood, LPC Analysis-synthesis can be used for datamasking. The all-pole synthesis filter is known to model the human vocal tract filter. However, the source model

for images is not well understood and hence we propose a different scheme for image datamasking as shown in Fig. ??.

Fig. 4. Proposed System for Image Datamasking

As in audio datamasking, we use the prediction error to mask the encrypted stream. We use a causal neighborhood of a pixel to predict its value. The causal neighborhood consists of W, NW, N, NE (west, northwest, north, northeast) pixels. According to MED predictor, the predicted value of current pixel is $\min(W, N)$ if $NW \geq \max(W, N)$ or it is $\max(W, N)$ if $NW \leq \min(W, N)$ or it is $W + N - NW$ otherwise.

A Huffman codebook, that is constructed for the Laplacian prediction error probability density function (PDF), is used to decode the input random encrypted stream. Then the PDF of Huffman decoded encrypted samples would also have a Laplacian PDF. Hence, one can modify each pixel so that prediction error at that pixel is matched with the Huffman decoded encrypted stream sample.

Note that in the proposed system for image datamasking, there are two parameters that the communicating parties should exchange beforehand. The first is a Huffman codebook designed for coding prediction errors with a Laplacian PDF and the second is a threshold to select the range prediction error values to be mapped to encrypted stream samples. For instance, if the chosen threshold is M , there would be $2M + 1$ Huffman codewords mapping $\{-M \dots 0 \dots M\}$. This threshold controls the tradeoff between datamasking capacity and quality of the data masked image. By choosing the threshold small, one would select only smooth regions of the image to hide information and hence the quality of the datamasked image would be better. With the knowledge of these two parameters, the receiver would map the prediction error at each pixel to the random encrypted stream.

In the following section, we present some of our experimental results to evaluate the performance of audio and image data masking.

III. EXPERIMENTAL RESULTS

We chose an AES (Advanced Encryption Standard) [?] encrypted stream as input to our system. The system's desired output is to generate a multimedia object from this stream. We use any audio clip or image as a reference to guide the system in the process.

A. Results on audio datamasking

We picked a speech clip of duration 2 seconds as our reference audio, A . We read every seven bits from the encrypted stream and append a random LSB and interpret it as a sample from a random process, E . Figures ?? and ?? compare the encrypted stream and reference audio clip in time domain and frequency domain respectively.

Fig. 5. Comparison of reference audio clip and encrypted stream in time domain

The reference audio is then segmented into frames of length 1024 samples, such that the time duration is short enough for the stationarity assumption to hold. Then for each frame of audio, an equal number of bytes from the encrypted stream are read. A LPC Analysis filter of order 31 was designed, the inverse of which was used to filter the encrypted stream. We thus have a time-varying inverse Wiener filter shaping every 1024 samples of the encrypted stream to match the reference audio frame.

Fig. 6. Comparison of Power Spectral Densities (PSDs) of reference audio clip and encrypted stream

Figure ?? shows the results of inverse Wiener filtering on encrypted stream corresponding to 2 seconds of reference audio clip and its PSD. It is clear that the filtered signal has more correlation than the encrypted signal and would potentially defeat most statistical tests for randomness. However, the filtered signal does not match the reference audio waveform exactly and hence the resulting waveform was noisier when listened

to (perceptual test). If the shaping of the encrypted stream perfectly matched that of reference audio for each frame, then the inverse Wiener filtered encrypted stream would also sound like the reference audio.

Fig. 7. Inverse Wiener Filtered encrypted stream & its PSD

In order to reconstruct the encrypted stream at the receiver, knowledge of LPC analysis filter coefficients for the current frame is essential. Therefore, the filter coefficients for the current frame are appended to the encrypted stream that was shaped in the previous frame and the coefficients corresponding to the first frame were included in the header of the .wav file. Wiener filtering can then be performed for each frame by retrieving the filtering coefficients reconstructed from the previous frame to obtain the encrypted stream for the current frame. Figure ?? shows the reconstructed encrypted stream and its difference from the original. The maximum difference between the samples of original and reconstructed encrypted stream is within 0.015. Since each sample was interpreted as a byte from the encrypted stream and scaled by $(1/64)$ and shifted by -1 , a bit flip in the encrypted stream can occur only if the error is greater than $(1/64)$. Therefore, if the error is within $(1/64)$ or (0.0156) , it is possible to reconstruct the encrypted stream without any bit error which is critical for proper decryption.

Fig. 8. Reconstructed encrypted stream & its difference from original encrypted stream

Now that we are convinced that the proposed scheme would mask encrypted data from statistical tests, let us evaluate the performance of this scheme as a audio data hiding technique. If we assume that each filter coefficient occupies a word, we have an overhead of 124 bytes (4×31) for every 1024 bytes of encrypted stream. The effective payload, $(1024 - 124 - 128) = 772$ bytes, can be thought of as embedding capacity for this scheme per frame of audio. It has been observed that the quantization error introduced by a precision of 8 bits per sample of the filtered encrypted stream would result in an error greater than $(1/64)$. Therefore, we

	[A]	[B]
Analyzer 1	66	-
Analyzer 2	66	-
Analyzer 3	66	-
Analyzer 4	66	-
Analyzer 5	66	-

TABLE I

PERFORMANCE OF EXISTING AUDIO STEGANALYSIS TECHNIQUES ON DATAMASKED AUDIO STREAMS; [A]: NUMBER OF INPUT DATAMASKED AUDIO STREAMS; [B]: NUMBER OF AUDIO STREAMS DETECTED BY THE STEGANALYZER;

use 16 bit precision for each sample in the filtered encrypted stream and assume that the transmitted inverse Wiener filtered stream was received without any distortion. We can compare our scheme's embedding capacity to that of LSB embedding, in which one bit is embedded in the LSB per sample of audio. The proposed scheme hides 6176 bits (8×772) per 1024 bytes of audio data whereas LSB embedding (which also has low robustness) can hide 1024 bits per 1024 bytes of audio data. Therefore, the embedding capacity of proposed scheme is 6.0313 times that of the capacity of LSB embedding. In order to achieve an embedding capacity of 6176 bits per 1024 bytes using LSB embedding technique, it would require changing at least 6 bits of the 8 bits in each sample! Note that in practice, steganalysis techniques given in [?],[?],[?] can detect the presence of LSB embedding even when 10% of the bits in the cover message have been flipped (in some cases this can be as low as 2%). So we potentially have an order of magnitude larger capacity as compared to traditional multimedia steganography. Of course, we gain this by assuming that warden does not examine the message perceptually, but only statistically.

In order to see what kind of statistical tests can detect the datamasked audio signals, we tested them with a number of existing audio steganalysis tools. Table 1 summarizes the performance of some steganalysis techniques on 66 datamasked audio streams of total duration 29 minutes.

B. Results on image datamasking

A reference image was chosen and a Huffman codebook was designed for the prediction error PDF. Fig. shows the original reference image and the PDF of the prediction error. We used a threshold of 50 on the absolute value of prediction error and hence there were 101 huffman codewords to map the encrypted stream to prediction errors. Fig. shows the datamasked image and its prediction error PDF.

Original Reference Image

Prediction error histogram

Fig. 9. Original reference image and its prediction error histogram

Datamasked Image

Prediction error histogram

Fig. 10. Datamasked image and its prediction error histogram

M = 25

M = 15

Fig. 11. Datamasked images with M = 25 and M = 15 with corresponding embedding capacities of 4.56 & 3.65 bpp

The embedding capacity of this datamasking scheme for the chosen image was found to be 5.47 bits per byte of image stream. Fig. ?? shows the tradeoff between embedding capacity and quality of the datamasked image for two different values of M.

For all of the chosen M values, the embedding capacity is high compared to many of the multimedia steganographic schemes. With such a high embedding capacity, would the datamasked image escape statistical tests by existing steganalyzers?

This motivated us to test a number of datamasked images with existing steganalyzers for known steganographic techniques. Table presents the results for the same. Also, the detection performance of a generic steganalyzer on the same set of datamasked was found to be 2.81%. It can be observed that most of the steganalyzers have a low detection rate. It suggests that the proposed datamasking scheme has a different “signature” from that of existing steganographic schemes. Hence, one needs to train a new steganalyzer with datamasked images to learn the statistics of the datamasking scheme.

IV. CONCLUSIONS

In this paper, we proposed a novel solution to remove randomness from cipher streams in a reversible manner. This would make difficult for an eavesdropper to distinguish encrypted streams from other network

	[A]	[B]	[C]
F5r12	71	61	85.91%
F5r11	71	36	50.7%
Out-	71	71	100.0%
Out+	71	0	0.0%
Lsb	71	0	0.0%

TABLE II

DETECTION PERFORMANCE OF LINEAR CLASSIFIERS TRAINED WITH STEGO IMAGES OF SPECIFIC EMBEDDING TECHNIQUES USING BINARY SIMILARITY METRIC(BSM) AS FEATURES; [A]: NUMBER OF INPUT DATAMASKED IMAGE STREAMS; [B]: NUMBER OF IMAGE STREAMS DETECTED BY THE STEGANALYZER; [C]: DETECTION RATE

	[A]	[B]	[C]
F5r12	71	1	1.4%
F5r11	71	4	5.63%
Out-	71	5	7.04%
Out+	71	0	0.0%
Lsb	71	9	12.67%

TABLE III

DETECTION PERFORMANCE OF SVM CLASSIFIERS TRAINED WITH STEGO IMAGES OF SPECIFIC EMBEDDING TECHNIQUES USING BINARY SIMILARITY METRIC(BSM) AS FEATURES; [A]: NUMBER OF INPUT DATAMASKED IMAGE STREAMS; [B]: NUMBER OF IMAGE STREAMS DETECTED BY THE STEGANALYZER; [C]: DETECTION RATE

streams, thereby, providing covertness to secure channels. We have proposed schemes to generate audio streams and image streams from encrypted streams. Since perceptual tests are not feasible with such large volumes of data, we relax the requirement for perceptual transparency while maintaining the statistical properties. Our experimental results show that the proposed datamasking schemes have a “different” signature than existing steganographic schemes and was missed by many steganalyzers.

REFERENCES

- [1] G. J Simmons, Prisoner’s Problem and the Subliminal Channel (The), CRYPTO83 - Advances in Cryptology, August 22-24. 1984. pp. 51-67.
- [2] N. F Johnson, S. Katzenbeisser, “A Survey of Steganographic Techniques”, in S. Katzenbeisser and F. Petitcoals (Eds.): Information Hiding, pp. 43-78. Artech House, Norwood, MA, 2000.
- [3] N. F. Johnson, S. Jajodia, “Steganalysis of Images Created Using Current Steganography Software”, in David Aucsmith (Ed.): Information Hiding, LNCS 1525, pp. 32-47. Springer-Verlag Berlin Heidelberg 1998.
- [4] A. Westfield, A.Pfitzmann, “Attacks on Steganographic Systems”, in Information Hiding, LNCS 1768, pp. 61-76, Springer-Verlag Heidelberg, 1999.
- [5] J. Fridrich, M. Goljan and R. Du, “Reliable Detection of LSB Steganography in Color and Grayscale Images”. Proc. of the ACM Workshop on Multimedia and Security, Ottawa, CA, October 5, 2001, pp. 27-30.
- [6] I. Avcibas, N. Memon, B. Sankur, “Steganalysis Using Image Quality Metrics”, Security and Watermarking of Multimedia Contents, SPIE, San Jose, 2001.
- [7] Advanced Encryption Standard, FIPS 197, 2001.