# AN ADAPTIVE SPREAD-SPECTRUM DATA HIDING TECHNIQUE FOR DIGITAL AUDIO

*Mark Sterling, Edward L. Titlebaum, Xiaoxiao Dong, Mark F. Bocko*

University of Rochester, ECE Department
Rochester, NY 14627 USA

## ABSTRACT

In this paper we describe an application of spread spectrum techniques in audio data hiding for watermarking and steganography. The method is self-synchronizing, cover dependent, and operates in the time domain. We use a special class of frequency-hop signal know as a Welch-Costas Array. Welch-Costas Arrays have the properties of range and Doppler resolution. This allows us to recover embedded data with a matched filter. We also demonstrate a special case of an adaptive method due to Su and Girod [1].

## 1. INTRODUCTION

In spread spectrum communications, the bandwidth of a transmitted signal is increased to afford protection from interference. Spread-spectrum signals have applications in jamming protection, hidden communications, and multiple access systems. Due to the large bandwidth, reliable SNR can be achieved with relatively low transmit power.

One form of spread-spectrum is frequency-hopped spread spectrum (FHSS). In FHSS, we divide the time axis into fixed length intervals and place a sinusoid in each interval according to a pre-defined sequence of frequencies. This sequence, known as the hop pattern, may be thought of as an index into a finite set of frequencies. Typically the lowest and highest frequencies, $f_0$ and $f_1$ are given by the user, and the remaining frequencies are chosen automatically at equally spaced intermediate values. Welch-Costas Arrays are FHSS signals with good range and Doppler resolution. In addition to these correlation properties, the FHSS technique, through the selection of parameters allows us to create broadband noise-like waveforms.

A Welch-Costas Array is itself a candidate steganographic signal. We can implement a system in the following manner. Let $w_N[n]$ be an $N$ point Welch-Costas Array. This signal is defined to be 0 outside of the interval $[1, N]$. The coefficients $b_i \in (-1, +1)$ in (1) represent the embedded binary information. The signals $x[n]$, $w[n]$, and $y[n]$ are resepctively the cover, the steganographic signal, and the marked output. The constant $\eta$ is a scaling parameter—watermark strength—and $\phi \in \mathbb{Z}^+$ is an arbitrary phase shift we introduce in order to demonstrate the self-synchronizing capability of the system. Note that, unlike the spread-spectrum technique of [2], our method embeds data in the time domain.

$$w[n] = \sum_{i=1}^{M} b_i w_N[n - iN] \tag{1}$$

$$y[n] = x[n] + \eta w[n + \phi] \tag{2}$$

The embedded data can be recovered by applying a matched filter. The output of the matched filter, evaluated at the end of each frame is our estimate of the data.

$$m[n] = y[n] * w_N[-n] \tag{3}$$

$$\tilde{b}_i = \begin{cases} 1, & m[iN + \tilde{\phi}] \geq 0 \\ -1, & m[iN + \tilde{\phi}] < 0 \end{cases} \tag{4}$$

Following Su and Girod we feel that an improvement can be made. In [1] they introduce the *power-spectrum condition* for watermarking. A watermark is said to be PSC-compliant when its power spectrum is proportional to the power spectrum of the cover. Here, this can be achieved approximately in the following way. DFTs of $x[n]$ and $w[n]$ are computed in $N$ point non-overlapping blocks. The result is a set of $M$ $N$-point DFTs $X_i[k]$, $W_i[k]$, $i = 1, \ldots, M$. Each $W_i$ is weighted by the magnitude of $X_i$ and converted back to the time domain. Once this is finished, the watermark can be inserted as in (2).

$$w_{PSC,i}[n] = \text{IFFT}\{|X_i[k]|\, W_i[k]\} \tag{5}$$

Su and Girod [1] focus on the issue of robustness for watermarking. In particular, they prove that PSC-compliant watermarks are optimally resistant to a *Wiener Attack*. Given some long-standing results [3] [4] on coding for a channel with side information, the approach of (5) is preferable to that of (1) and (2) in the sense that in (5) the side information is utilized. The side information is the steganographers' knowledge of the cover. Furthermore, although (5) is not a

This work was supported by the Air Force Research Laboratory/IFEC under grant number F30602-02-1-0129

QIM approach [5], our approach is similar in that it also hides data by introdcing a cover dependent additive noise.

## 2. WELCH-COSTAS ARRAYS

As mentioned above, a Costas Array is a frequency-hop signal with good range and Doppler resolution. In radar terminology, we say that the auto-ambiguity function approaches an ideal thumbtack shape. A Welch-Costas Array is a Costas Array with a specific algebraic construction. In this paper, we do not especially exploit the Doppler resolution of the Welch-Costas Array. The time resolution, however, is an essential feature that allows us to recover the location of frame boundaries from the matched filtering.

There is a natural relationship, through the notion of a hop pattern, between FHSS signals and permutation matrices. A permutation matrix is a matrix $\mathbf{A} = (a_{ij}),\ a_{ij} \in (0, 1)$ where each column and each row contain a single 1. The rows of this matrix may be thought of as divisions in frequency and the columns may be thought of as divisions in time. When $a_{ij} = 1$ we place a sinusoidal pulse at the appropriate time shift and frequency. The hop pattern can be deduced simply from the columns of $\mathbf{A}$. The fact that $\mathbf{A}$ is a permutation matrix is equivalent to saying that only one frequency is active per time division and that all of the frequencies are visited exactly once over the total period.

Complete discussion of Welch-Costas Arrays can be found in [6, 7, 8, 9]. The problem, as stated by J. Costas in [10] is, *Place N ones in an otherwise null N by N matrix such that each row contains a single one as does each column. Make the placement such that for all possible $x$-$y$ shift combinations of the resulting (permutation) matrix relative to itself, at most one pair of ones will coincide*. A permutation matrix for which this is true is called a Costas Array (Here, we confess to a slight abuse of terminology. For convenience, the frequency hop signal itself, the hop pattern, and the permutation matrix corresponding to the hop pattern are all referred to as a "Costas Array." The sense in which the term is intended, however, is usually clear from context ).

Suppose that $p$ is a prime number and $\alpha$ is a primitive root of $p$. A $(p-1) \times (p-1)$ permutation matrix $\mathbf{A}$ is a Welch-Costas Array if the matrix elements are such that (6) holds.

$$a_{ij} = \begin{cases} 1, & i \equiv \alpha^j \bmod p \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

Let $\phi(n)$ denote the totient function. A prime number $p$ has $\phi(\phi(p))$ primitive roots. Each of these primitive roots generates a different Welch-Costas Array. There are bounds on the number of coincidences that can occur between any two such arrays ("coincidences" as in Costas' definition).

In particular, if $\alpha_1$ and $\alpha_2$ are primitive roots of the same prime then the following equation will be true for some $n$.

$$\alpha_1 = \alpha_2^n \bmod p \tag{7}$$

Let $(\mathbf{A}_{\alpha_1}, \mathbf{A}_{\alpha_2})$ equal the number of coincidences occuring between two Welch-Costas Arrays.

$$\max_{\forall x,y \text{ shifts}} (\mathbf{A}_{\alpha_1}, \mathbf{A}_{\alpha_2}) = \begin{cases} n, & n \le \frac{p-1}{2} \\ -n \bmod p, & n > \frac{p-1}{2} \end{cases} \tag{8}$$

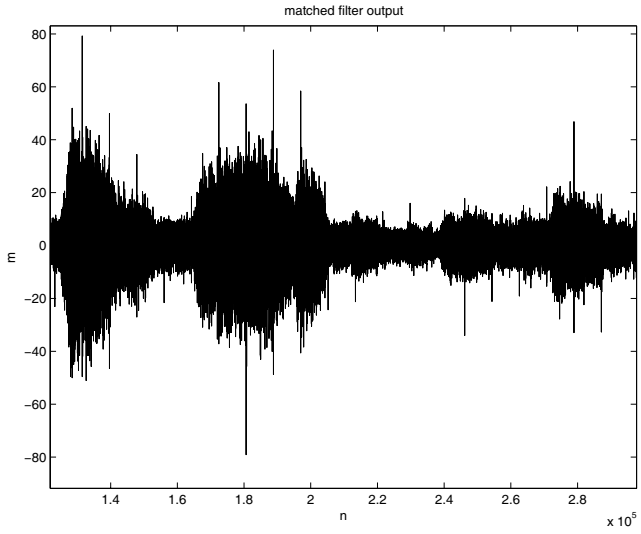


**Fig. 1**. Example of matched filter $m[n]$ with visible peaks at frame boundaries

## 3. DETECTION ALGORITHM

Our detection algorithm consists of two steps. First, the phase $\phi$ in (1) must be determined from the output of the matched filter (3). The signal $m[n]$ is passed to a sorting algorithm that retains the sample indices. The values of these indices are taken $\bmod N$ and the class which is most correlated with the large sample values is assumed to be the phase $\tilde{\phi}$. The embedded bits are found according to the rule in (4).

## 4. EXPERIMENTAL RESULTS

We conducted simulations on a set of cover files representing a variety of commercially available music. The audio clips were mono, 44.1 kHz, 16 bit PCM signals between 10 and 20 seconds long.

The results of three experimental runs are shown in Figs. 2, 3 and 4. The error rate, defined in (9), is plotted versus the signal strength $\eta$.

$$e = \frac{\sum_{i=1}^{M} \left| b_i - \tilde{b}_i \right|}{2M} \tag{9}$$

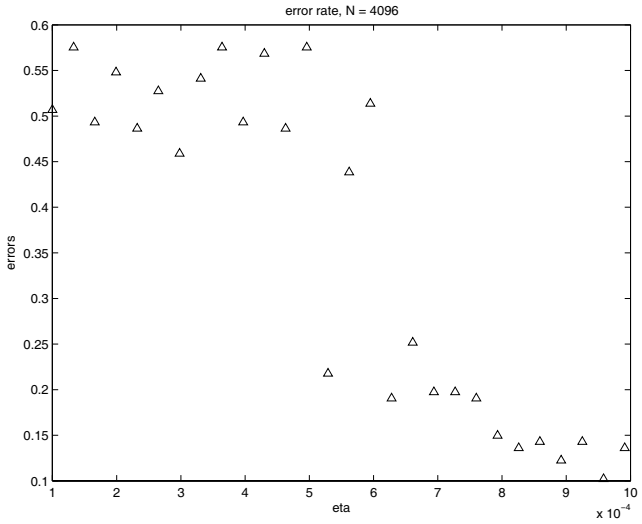


**Fig. 2**. Error rate versus $\eta$. $N = 4096$, $p = 499$, $\alpha = 7$, $f_0 = 4000$ Hz, $f_1 = 22050$ Hz

## 5. SUMMARY AND CONCLUSIONS

We have demonstrated a frequency hopped spread-spectrum technique for steganography in digital audio. The method is self-synchronizing and cover dependent. In addition, it allows a user easily specify the bandwidth of the steganographic signal.

## 6. REFERENCES


[1] J. K. Su and B. Girod, "Power-spectrum condition for energy-efficient watermarking," *IEEE Transactions on Multimedia*, vol. 4, no. 4, pp. 551–560, 2002.

[2] D. Kirovski and H. S. Malvar, "Spread-spectrum watermarking of audio signals," *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 1020–1033, 2003.


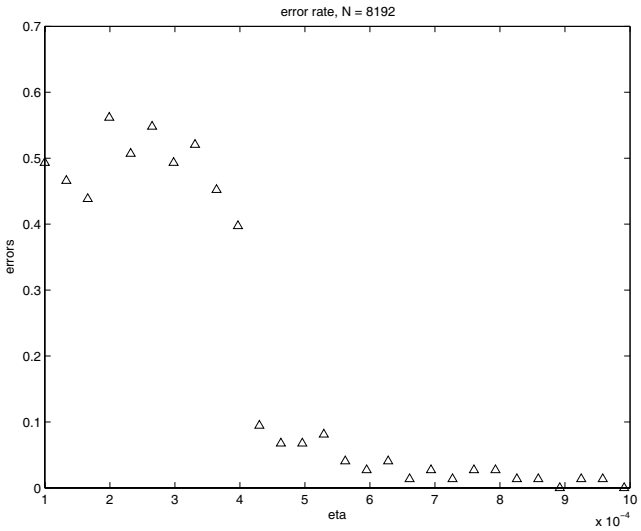


**Fig. 3**. Error rate versus $\eta$. $N = 8192$, $p = 499$, $\alpha = 7$, $f_0 = 4000$ Hz, $f_1 = 22050$ Hz

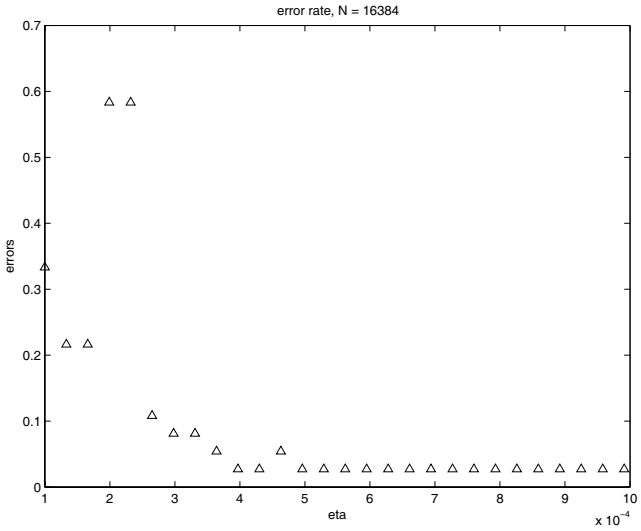


**Fig. 4**. Error rate versus $\eta$. $N = 16384$, $p = 499$, $\alpha = 7$, $f_0 = 4000$ Hz, $f_1 = 22050$ Hz


[3] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random paramters," *Problems of Control and Information Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[4] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 493–441, 1983.

[5] B. Chen and G. Wornell, "Achievable performance of digital watermarking systems," in *Proc. IEEE Int. Conf. Multimedia Comput. Syst.*, Florence, Italy, 1999, pp. 13–18.

[6] S. W. Golomb and H. Taylor, "Two dimensional syn-

chronization patterns for minimum ambiguity," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 4, pp. 600–604, 1982.

[7] S. V. Maric, I. Seskar, and E. L. Titlebaum, "On cross-ambiguity properties of welch-costas arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-30, no. 4, pp. 1063–1071, 1994.

[8] G. S. Bloom and S. W. Golomb, "Applications of numbered undirected graphs," *Proceedings of the IEEE*, vol. 65, no. 5, pp. 593–619, 1980.

[9] E. L. Titlebaum, "Frequency and time-hop coded signals for use in radar and sonar systems and multiple access communications systems," in *Proc. of the Twenty-Seventh Asilomar Conference on Signals, Systems and Communications*, Pacific Grove, CA, 1993, pp. 1096–1100.

[10] J. P. Costas, "A study of a class of detection waveforms having nearly ideal range-doppler ambiguity properties," *Proceedings of the IEEE*, vol. 72, no. 8, pp. 996–1009, 1984.