January 2009

Wavelet Decompositions for Quantitative Pattern Matching

Bruce Kessler
Western Kentucky University, bruce.kessler@wku.edu

Recommended Citation
Bruce Kessler, "Wavelet Decompositions for Quantitative Pattern Matching", *Proceedings of the 42nd Annual Hawaii International Conference on System Sciences* (CD-ROM), January 5-8, 2009, Computer Society Press, (2009) 10 pages.

Multiwavelets for Quantitative Pattern Matching

Bruce Kessler
Western Kentucky University
bruce.kessler@wku.edu

Abstract

The purpose of this paper is to provide an introduction to the concepts of wavelets and multiwavelets, and explain how these tools can be used by the analyst community to find patterns in quantitative data. Three multiwavelet bases are introduced, the GHM basis from [3], a piecewise polynomial basis with approximation order 4 from [2], and a smoother approximation-order-4 basis developed by the author in previous work [6]. The technique of using multiwavelets to find patterns is illustrated in a traffic-analysis example.

1. Introduction

The concept of wavelets and wavelet analysis is a relatively new idea in the science of analysis, which has been dominated over the last 200 years by Fourier analysis. Even newer is the concept of multiwavelets, which have greater flexibility for handling boundaries, symmetries, and anti-symmetries. The purpose of this paper is to provide the network security community an overview of the concept, and some of the bases that are available for their use.

1.1 Multiresolution Analyses

A function ϕ that satisfies the dilation equation

$$\phi(x) = \sqrt{2}\sum_{n\in\mathbb{Z}} c_n \phi(2x-n) \tag{1}$$

for some sequence of coefficients c_n is said to be *refinable*. A simple example of a refinable function would be the characteristic function over $[0,1)$,

$$\chi_{[0,1)} = \begin{cases} 1 & \text{for } 0 \le x < 1, \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

shown in Figure 1, since for $\phi = \chi_{[0,1)}$,

$$\phi(x) = \sqrt{2}\left(\frac{1}{\sqrt{2}}\phi(2x) + \frac{1}{\sqrt{2}}\phi(2x-1)\right).$$

A refinable function ϕ that has the property that the set

$$\{\phi(x-n) : n \in \mathbb{Z}\}$$

is a linearly independent set is called a *scaling function.* A scaling function ϕ where the dilation equation (1) is satisfied by a finite sequence c_n and

$$\int_{\mathbb{R}} \phi(x)\phi(x-n)\,dx = \begin{cases} 1 & \text{for } n = 0, \\ 0 & \text{for } n \in \mathbb{Z},\ n \ne 0 \end{cases}$$

is called an *orthogonal* scaling function. The function $\phi = \chi_{[0,1)}$ defined in (2), along with its integer translates, is called the *Haar basis*, and is an orthogonal scaling function due to its very short support.

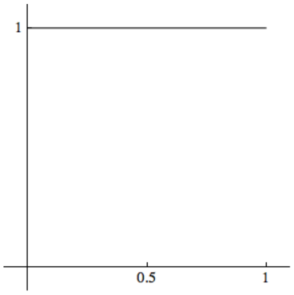

Figure 1: The orthogonal scaling function $\phi = \chi_{[0,1)}$.

A *multiresolution analysis* (MRA) of square-integrable functions defined on $\mathbb{R}$ (typically denoted $L^2(\mathbb{R})$) is a set of linear spaces (V_p) that satisfy the following criteria:

- $\cdots \supset V_{-2} \supset V_{-1} \supset V_0 \supset V_1 \supset V_2 \cdots$,
- $\overline{\bigcup_{p\in\mathbb{Z}} V_p} = L^2(\mathbb{R})$,
- $\bigcap_{p\in\mathbb{Z}} V_p = \{0\}$,
- $f \in V_0$ iff $f(2^{-j}\cdot) \in V_j$ for $j \in \mathbb{Z}$, and
- there exists a function ϕ whose integer translates form a basis (minimal spanning set) for the space V_0.

Based on these criteria, scaling functions can clearly be used to generate MRA's. The orthogonal complement of V_0 in V_{-1} is typically denoted W_0, and a function ψ whose integer translates form a basis for this space are called *wavelets*. The wavelet associated with the Haar basis is shown in Figure 2.

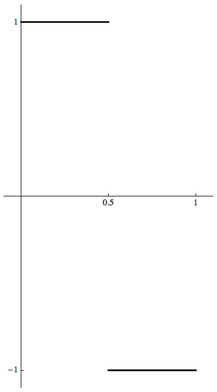

Figure 2: The wavelet associated with the Haar basis.

If we define similar spaces W_j as the orthogonal complement in V_{j-1} in V_j for all integer j, called *wavelet spaces*, then we may decompose a function in V_0 into increasingly "smoother" approximations, and keep the error in the wavelet spaces. An simple example of this process is shown in Figures 3 through 6 using the Haar basis.

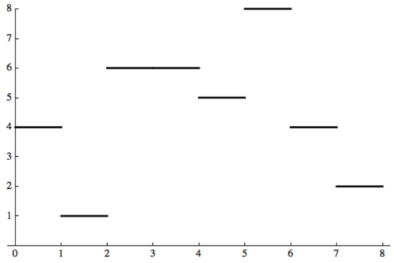

Figure 3: A signal in the space V_0 generated by integer translates of the Haar basis.

Since $\psi \in W_{-1}$, by necessity ψ will satisfy a dilation equation of its own,

$$\psi(x) = \sqrt{2} \sum_{n \in \mathbb{Z}} d_n \psi(2x - n),$$

for some sequence of coefficients d_n. For example, the Haar wavelet ψ illustrated in Figure 2 satisfies the equation

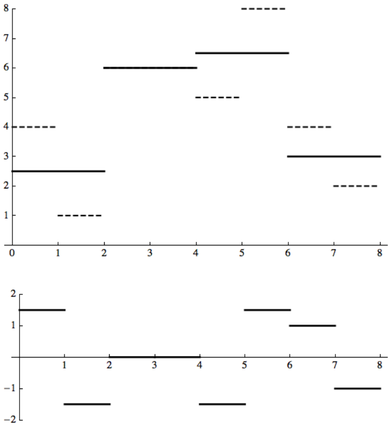

Figure 4: A best approximation in V_1 of the signal in Figure 3 at top, and the error in W_1 at bottom.

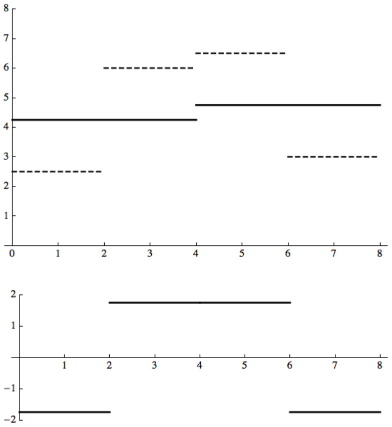

Figure 5: A best approximation in V_2 of the signal in Figure 4 at top, and the error in W_2 at bottom.

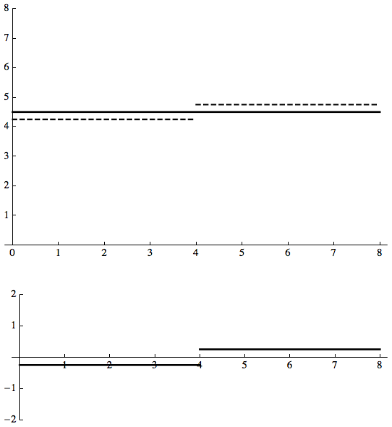

Figure 6: A best approximation in V_3 of the signal in Figure 5 at top, and the error in W_3 at bottom.

$$\psi(x) = \sqrt{2}\left(\frac{1}{\sqrt{2}}\phi(2x) - \frac{1}{\sqrt{2}}\phi(2x-1)\right).$$

In the case where the MRA is generated by an orthogonal scaling vector, the coefficients c_n and d_n can be used to construct perfect-reconstruction filters for decomposing and reconstructing the original function in V_0.

The *approximation order* of a scaling function refers to the highest degree polynomials that, when limited to a finite interval, are present in the space V_0 generated by the scaling function. A scaling function of approximation order k is able to reproduce degree $k-1$ polynomials exactly. For example, the scaling function $\phi = \chi_{[0,1)}$ generates a space that contains degree zero (constant) functions over a finite interval $[a, b]$, a, b integers, so we say that ϕ has approximation order 1. The approximation order of a scaling function will be an important deciding factor when we start to consider bases for applications.

1.2 Wavelets vs. Fourier Analysis

The fast Fourier transform (FFT) is a marvelous tool for isolating the frequencies in a discrete, uniformly-sampled signal, but it has its limitations. The FFT is a computationally efficient method of calculating the coefficients a_n and b_n, $n = 0, \ldots, N-1$, needed to represent a sampled signal $\{y_k\}_{k=0}^{N-1}$ as the sum

$$y_k = \sum_{n=0}^{N-1}\left(a_n \cos\left(\frac{2\pi}{N}nk\right) + b_n \sin\left(\frac{2\pi}{N}nk\right)\right)$$

for each $k = 0, \ldots, N-1$. Typically, one looks at the moduli of the various complex numbers $a_n + b_n i$, called the *power spectrum*, for an indication that the signal has a component that oscillates n times over the length of the signal. A modulus of zero means that that frequency is not present in the signal, while a larger modulus means that more of that frequency is present in the signal. This is particularly useful when analyzing sound signals, especially if trying to sort and denoise certain frequencies out of the signal. The FFT provides no information about *where* the frequencies occur in the signal, just how much is present.

For example, the two data sets shown in Figure 7 will have different Fourier transforms, but with the exact same power spectrum, shown in Figure 8. To combat this shortcoming, a *windowed* Fourier transform is used, but this in turn limits the frequencies that can be measured. Also, if the same data is added to a set of linear data, both the Fourier transform and its power spectrum will be radically different, as shown in Figure 9.

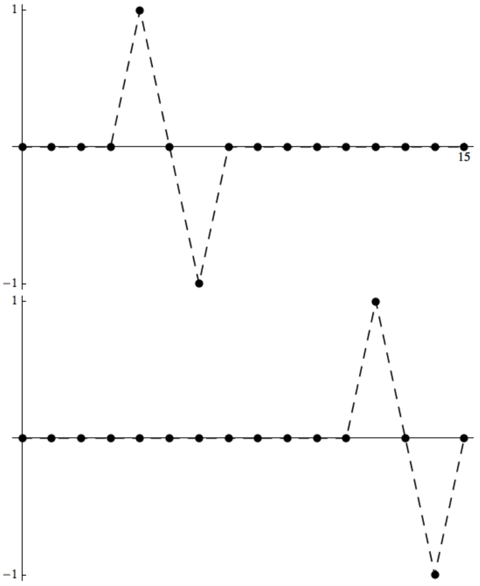

Figure 7: Two different, but similar, data sets.

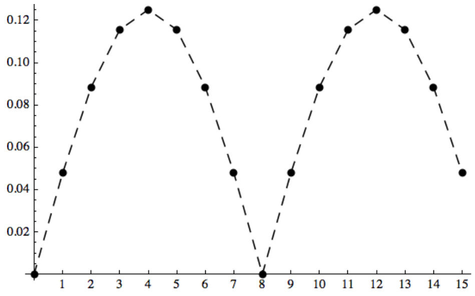

Figure 8: The power spectrum for both of the data sets in Figure 3.

With wavelet analysis, the two data sets shown in Figure 7 will have completely different wavelet decompositions, with larger magnitude coefficients in the location of the "bump". However, with the use of the correct basis (namely, a basis with approximation order 2 or above), the top set of data in Figure 7 and the data shown at the top of Figure 9 will have exactly the same wavelet decomposition. This property can be useful when trying to analyze data overlaid upon other background data.

1.3 Wavelets vs. Multiwavelets

The concept of a scaling function can be generalized to a vector of functions, called a *scaling* *vector.* A scaling vector $\Phi = (\phi_1, \ldots, \phi_r)^T$ satisfies a dilation equation of its own,

$$\Phi(x) = \sqrt{2} \sum_{n \in \mathbb{Z}} c_n \Phi(2x - n) \tag{3}$$

for some sequence of $r \times r$ matrices, and the entire set

$$\{\phi_j(x - n) : \; j \in \{1, \ldots, r\}, \, n \in \mathbb{Z}\} \tag{4}$$

is a linearly independent set. Scaling vectors are said to be *orthogonal* if the sequence of matrices c_n is finite and the components of Φ satisfy

$$\int_{\mathbb{R}} \phi_i(x)\phi_j(x - n) = \begin{cases} 1 & \text{when } i = j, \, n = 0, \\ 0 & \text{otherwise.} \end{cases}$$

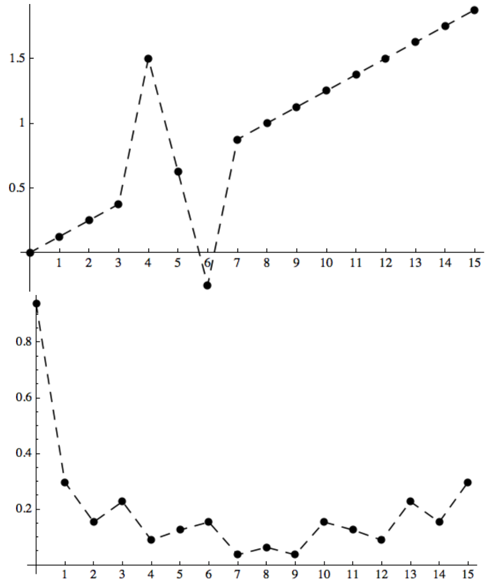

Figure 9: The same data set with added linear data, and its power spectrum.

Scaling vectors still establish MRA's, with the set given in (4) as the basis for V_0. Within this structure, there will exist a vector of functions $\Psi = (\psi_1, \ldots, \psi_r)^T$ that span the orthogonal complement of V_0 in V_{-1}, denoted W_0, that will necessary satisfy its own dilation equation

$$\Psi(x) = \sqrt{2} \sum_{n \in \mathbb{Z}} d_n \Psi(2x - n) \tag{5}$$

for some finite sequence of $r \times r$ matrices d_n. As before, when using an orthogonal scaling vector, the matrix coefficients can be used to construct perfect-reconstruction filters for the decomposition and reconstruction of the signal in V_0.

There are several advantages to using multiwavelets for analysis over single wavelets. The Haar scaling function defined in (2) and shown in Figure 1 is the only single scaling function to have any symmetry/anti-symmetry properties. However, as we shall see in the next section, we may build scaling vectors such that each function in the vector has symmetry/anti-symmetry properties. Also, each single scaling function ϕ must necessarily satisfy the condition that

$$\int_{\mathbb{R}} \phi(x) \, dx \neq 0.$$

This is not true for all of the functions in a scaling vector, but for only one of the functions. (The anti-symmetric members of the scaling vector would not satisfy this criteria, for example, so not all of the scaling vector functions could be anti-symmetric.) And lastly, for the class of scaling vectors that we will illustrate in this paper, it is very easy to build boundary functions and still maintain orthogonality. With the exception of the Haar basis, if you truncate a single scaling function supported on $[a, b]$ to a shorter integer-length interval, the function will no longer maintain its orthogonality to other truncated functions. The scaling vectors shown in Section 2, all of which are supported over an interval no larger than $[-1, 1]$, have the property that they are orthogonal (or in the last basis shown, can be made orthogonal with a simple procedure) when restricted both to the interval $[-1, 0]$ and $[0, 1]$, meaning that the functions, when truncated at $x = 0$ and normalized, are still part of an orthogonal basis set on a bounded region.

Multiwavelets do have one disadvantage over the single wavelet constructions: the filters can not generally be applied directly to the raw data without losing approximation order. (Much work has been done in creating *balanced* scaling vectors that can be applied directly to data. See [1], [7], and [8] for more details.) In order to maintain the polynomial order of the data, one has to first convert the raw data to basis coefficients in the V_0 space, called *prefiltering*, and if reconstructing the data, convert the basis coefficients back into data, called *postfiltering*. Ideally, a prefilter will preserve the l^2-norm of the data, called an *orthogonal* prefilter, as well as preserve the polynomial order of the data, but orthogonal prefilters that preserve approximation order of the basis are sometimes hard to find. (They also have issues on bounded regions. See [5] for an excellent introduction to prefiltering.) Luckily, an orthogonal prefilter is more necessary in image compression applications than in pattern recognition. We are able to use a relatively simple non-orthogonal prefilter, called a *quasi-interpolation* prefilter, that maps the data to basis coefficients in V_0 so that the combination of basis functions interpolates data sampled from polynomials up to the approximation order of the basis, and comes close to interpolating non-polynomial data.

2. Useful Bases

There are a number of orthogonal multiwavelet bases that have been developed. (See [9] and [10] for some other examples.) As previously mentioned, the ones presented here have the advantage of being easily applied to bounded data, but they also have another advantage. The following bases all generate spaces that contain *spline spaces*; that is, piecewise polynomial spaces over integer knots that meet certain match-up conditions at the integers. We use the notation $\mathcal{S}_d^r(\mathbb{Z})$ to indicate the spline space of polynomials of degree d that have equal r^{th} derivatives at integer knots, where the 0^{th} derivative is understood to be the function value. Thus, while we will refer to the approximation order k of the following bases, it it important to realize that the V_0 space generated by the bases will also contain piecewise polynomials of degree $k-1$ that are either continuous at integer values, or in the case of the last basis, have one derivative at integer values.

2.1 GHM Basis

The GHM basis first appeared in [3], and was reconstructed from a macroelement perspective in [4], and are illustrated in Figure 10. This scaling vector $\Phi = (\phi_1, \phi_2)^T$ has approximation order 2, but the V_0 space that it generates also includes the spline space $\mathcal{S}_1^0(\mathbb{Z})$.

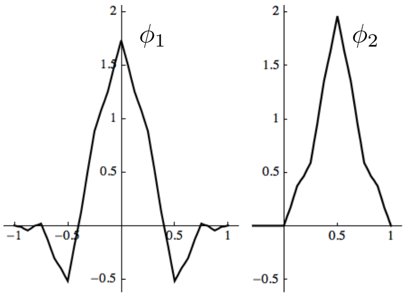

Figure 10: The GHM scaling vector.

The scaling vector Φ satisfies equation (3) with the following matrix coefficients:

$$c_{-2} = \begin{bmatrix} 0 & -\frac{1}{20} \\ 0 & 0 \end{bmatrix} \qquad c_{-1} = \begin{bmatrix} -\frac{3\sqrt{2}}{20} & \frac{9}{20} \\ 0 & 0 \end{bmatrix}$$

$$c_0 = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{9}{20} \\ 0 & \frac{3\sqrt{2}}{10} \end{bmatrix} \qquad c_1 = \begin{bmatrix} -\frac{3\sqrt{2}}{20} & -\frac{1}{20} \\ \frac{4}{5} & \frac{3\sqrt{2}}{10} \end{bmatrix}.$$

The function values needed to interpolate data or otherwise develop prefilters are

$$\phi_1(-1) = \phi_1(1) = \phi_2(0) = \phi_2(1) = 0,$$

$$\phi_1\left(-\frac{1}{2}\right) = \phi_1\left(\frac{1}{2}\right) = -\frac{3\sqrt{3}}{10}, \text{ and } \phi_2\left(\frac{1}{2}\right) = \frac{4\sqrt{6}}{5}.$$

Orthonormal basis elements for the left and right boundaries of a data set are constructed by keeping the right and left half of ϕ_1, respectively, and normalizing; that is,

$$\phi_1^L = \sqrt{2}\,\phi_1\chi_{[0,1]} \text{ and } \phi_1^R = \sqrt{2}\,\phi_1\chi_{[-1,0]},$$

respectively. Their decomposition filters reflect this normalization in all but the entries corresponding to dilated versions of these functions.

The multiwavelet $\Psi = (\psi_1, \psi_2)^T$ associated with the GHM scaling vector is illustrated in Figure 11, and satisfies the equation (5) with the following matrix coefficients:

$$d_{-2} = \begin{bmatrix} 0 & \frac{1}{20} \\ 0 & \frac{\sqrt{2}}{20} \end{bmatrix} \qquad d_{-1} = \begin{bmatrix} \frac{3\sqrt{2}}{20} & -\frac{9}{20} \\ \frac{3}{10} & -\frac{9\sqrt{2}}{20} \end{bmatrix}$$

$$d_0 = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{9}{20} \\ 0 & \frac{9\sqrt{2}}{20} \end{bmatrix} \qquad d_1 = \begin{bmatrix} \frac{3\sqrt{2}}{20} & \frac{1}{20} \\ -\frac{3}{10} & -\frac{\sqrt{2}}{20} \end{bmatrix}.$$

Only truncated and normalized versions of the symmetric ψ_1 are needed at the boundaries,

$$\psi_1^L = \sqrt{2}\,\psi_1\chi_{[0,1]} \text{ and } \psi_1^R = \sqrt{2}\,\psi_1\chi_{[-1,0]}$$

on the left and right, respectively.

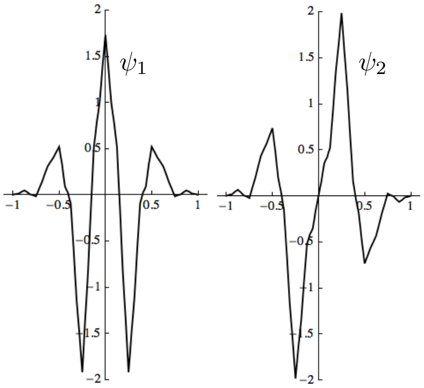

Figure 11: The GHM multiwavelet.

2.2 Approximation-Order-4 Basis

The groundwork for this basis was laid in [2], although the scaling vector was not explicitly constructed in that paper. The scaling vector $\Phi = (\phi_1, \phi_2, \phi_3, \phi_4)^T$ is illustrated in Figure 12, and has approximation order 4. The V_0 space that it generates also includes the spline space $\mathcal{S}_3^0(\mathbb{Z})$. In fact, one may notice that ϕ_2 and ϕ_3 are quadratic and cubic polynomials, respectively, restricted to $[0,1]$ and normalized.

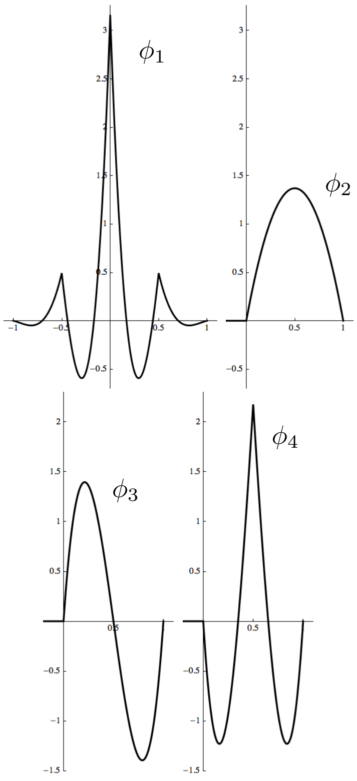

Figure 12: An approximation-order-4 scaling vector.

The scaling vector Φ satisfies the dilation equation (3) just as with the previous GHM scaling vector. However, the four 4×4 matrix coefficients are too large to provide here, and are available instead from the author's website, www.wku.edu/˜bruce.kessler, as are the function values needed to interpolate data or otherwise develop prefilters.

As before, orthonormal basis elements for the left and right boundaries of a data set are constructed by keeping the right and left half of ϕ_1, respectively, and normalizing; that is,

$$\phi_1^L = \sqrt{2}\,\phi_1\chi_{[0,1]} \text{ and } \phi_1^R = \sqrt{2}\,\phi_1\chi_{[-1,0]},$$

respectively. Their decomposition filters reflect this normalization in all but the entries corresponding to dilated versions of these functions.

The multiwavelet $\Psi = (\psi_1, \psi_2, \psi_3, \psi_4)^T$ associated with this scaling vector is illustrated in Figure 13, with the matrix solutions to the dilation equation (5) also available on the author's website. Again, only truncated and normalized versions of the symmetric ψ_1 are needed at the boundaries,

$$\psi_1^L = \sqrt{2}\,\psi_1\chi_{[0,1]} \text{ and } \psi_1^R = \sqrt{2}\,\psi_1\chi_{[-1,0]}$$

on the left and right, respectively.

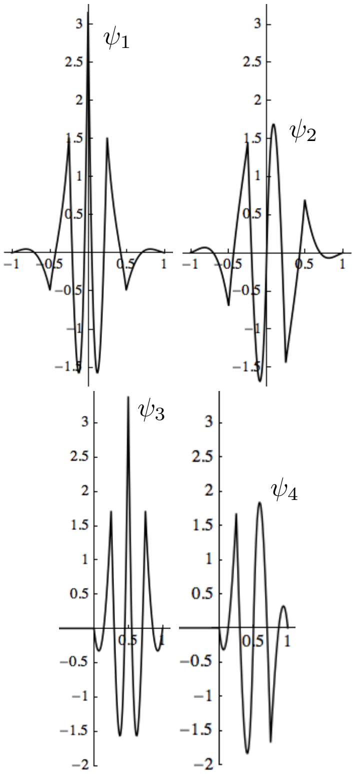

Figure 13: The multiwavelet associated with the scaling vector in Section 2.2.

2.3 Differentiable Basis with Approximation Order 4

Each of the bases shown in the previous sections include a continuous spline space, but the V_0 space in each case also contains much more than that. With each basis, V_0 will also contain functions that look like the left and right halves of the ϕ_1 function in that particular scaling vector. Hence, each approximation space, even at the lowest resolution, will contain functions that have at least one non-differentiable "corner". This may not be desirable, especially if the points of interest in the original signal happen to fall on these corners at every resolution (for example, at the midpoint of the data set). In this and many other cases, a differentiable basis (i.e., has no corners) would be more useful.

A scaling vector was developed by the author in [6] that has approximation order 4, and generates a space that contains the spline space $\mathcal{S}_3^1(\mathbb{Z})$, that is, the space of piecewise cubic polynomials that are both continuous and differentiable at the integer knots. This scaling vector $\Phi = (\phi_1, \phi_2, \phi_3, \phi_4)^T$ is illustrated in Figure 14. As with the basis discussed in Section 2.2, the matrix coefficients that satisfy the dilation equation (3) for Φ are available on the author's website, as are the function values needed to interpolate data or otherwise develop prefilters. Left- and right-hand versions of both ϕ_1 and ϕ_2 can be created to handle bounded data.

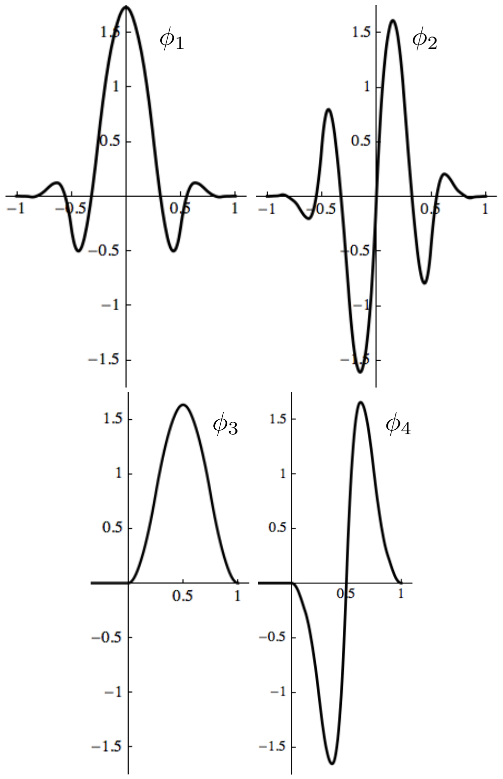

Figure 14: A differentiable approximation-order-4 scaling vector.

The multiwavelet $\Psi = (\psi_1, \psi_2, \psi_3, \psi_4)^T$ associated with this scaling vector is illustrated in Figure 15, with the matrix solutions to the dilation equation (5) also available on the author's website. The website also contains the matrices needed for filter construction of the left and right boundary wavelets.

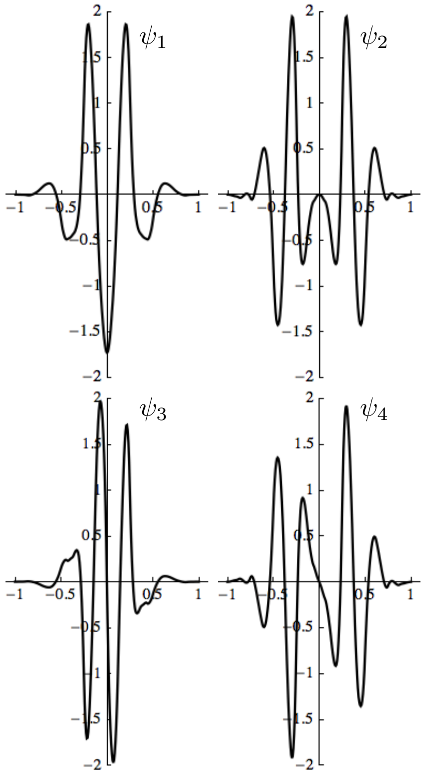

Figure 15: The multiwavelet associated with the scaling vector in Section 2.3.

3. Pattern Matching

The premise behind using wavelets for pattern matching is that the wavelet decomposition for the pattern for which you are searching and for the pattern added to data sampled from a polynomial of degree less than the approximation order of your scaling vector will be equal. For example, the wavelet decompositions for both data sets shown in Figure 16 will be identical when using either of the bases discussed in Sections 2.2 and 2.3. It is necessary that the pattern not be in the lowest-resolution approximation space, so that it will not have a wavelet decomposition with only 0 coefficients. This problem is rare, but if it occurs, it is remedied by either continuing the decomposition to an even lower-resolution space, or by using a different scaling vector of lower approximation order.

In most cases, the pattern will not be overlaid upon perfectly polynomial data, and so, the wavelet decompositions will not match exactly. However, if the pattern is prominent enough in the signal so that the *root mean square error* ($RMSE$) of the new decomposition $\{d_i\}$ when compared to the decomposition $\{d_i^*\}$ of the original pattern, as defined by

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(d_i - d_i^*)^2}{N}},$$

is sufficiently small, then the pattern will still be detected.

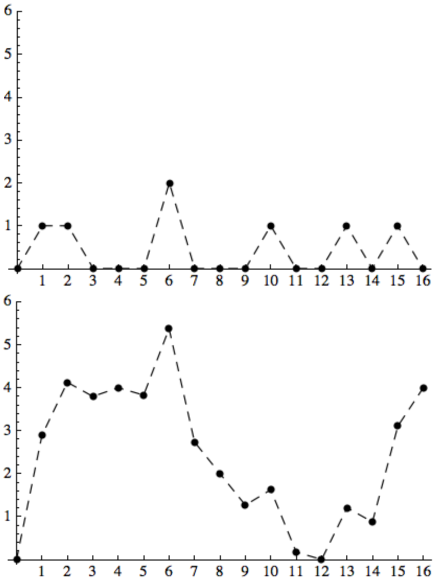

Figure 16: Pattern data (top), and the same data obscured by cubic polynomial data (bottom).

When using a single scaling function and wavelet are used in applications, each successively smoother approximation of the data stretches the support of the basis functions, effectively drawing more data from outside the region where the pattern actually occurs. By using the scaling vectors and multiwavelets mentioned in Section 2, we are able to analyze data on a bounded region, and ignore data outside of that region.

3.1 When to Use Wavelets

Wavelet decompositions can be used to find patterns in any type of data, but there are no advantages to using them for exact pattern matching. For example, if searching for the word "the" in a text file or in packet data, we have to search for the exact ASCII values 116, 104, and 101 in succession. Changing those values, even by a constant amount, changes the word. Also, to apply the wavelet bases mentioned in Section 2, we need $2^n + 1$ data values for some integer $n > 1$ for the DGM and the approximation-order-4 basis, and $2^n + 2$ data values for some integer $n > 1$ for the basis in Section 2.3, so we would have to pad the target pattern with spaces. This type of search can be conducted much more efficiently using other search methods, such as Bloom filters, etc.

However, for quantitative data like packet traffic dependent upon time, multiwavelets have the potential to spot patterns that are not immediately noticable to the human eye, such as in the bottom graph of Figure 16. The following section gives an example of how multiwavelets can be used to detect a "low and slow" pattern caused by data exfiltration.

3.2 Traffic Analysis Example

Suppose that we have a corrupted computer, and we are aware through previous experience that this particular piece of malevolent code tends to "sneak" data out under the "TCP" protocol in the following pattern:

- a single packet of size 256 bytes, followed approximately 1 second later by another packet of size 512 bytes,
- one second later, four packets of size 256 bytes are sent with a quarter second gap between the packets, and
- one quarter-second later, three packets of size 512 bytes are sent, followed one-quarter second later by one packet of size 256 bytes.

The accumulated bytes sent with respect to time in seconds are shown in Figure 17.

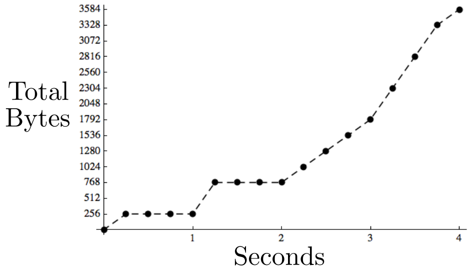

Figure 17: Accumulated bytes for the example pattern.

The packets described above have been inserted into a set of network traffic packet collected from the author's laptop over roughly 4 minutes of usage, starting at $t = 240$ seconds. The total accumulated outgoing bytes are illustrated in Figure 18. To apply the basis looking for this specific pattern, we accumulate the bytes on quarter-second intervals, and generate a wavelet decomposition for a sliding 4-second window, which is compared to the original pattern's decomposition with the $RMSE$. The $RMSE$ for the different starting times are shown in Figure 19. Note the low value of the $RMSE$ at $t = 240$ seconds, indicating a close match to our target pattern at that time.

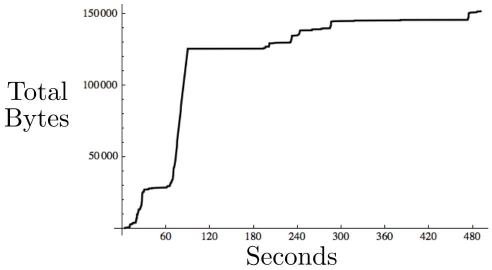

Figure 18: Accumulated outgoing bytes of a sample usage period plus the packets in our pattern.

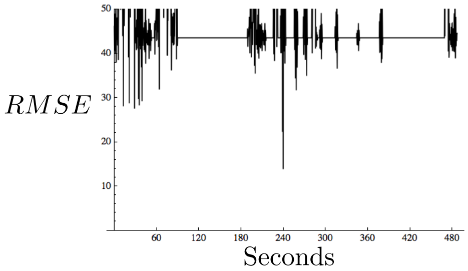

Figure 19: $RMSE$ values (less than 50) for blocks of data with different starting times. Note the low value of the $RMSE$ at $t = 240$ seconds, indicating a close match to our target pattern at that time.

In fairness, this method of finding patterns has many of the same weaknesses found in other signature-matching approaches. The pattern has to be known ahead of time, which means that this detection scheme would still be vulnerable to first-time attacks. Also, a completely innocuous data flow occurring with very similiar timing and quantities as our pattern will still cause a low $RMSE$ value, possibly causing a false-positive alert. We do claim, however, that we have a greater capacity for finding patterns in quantifiable amounts that are obscured by standard network traffic. Also, we can adjust the sensitivity of the search without changing the signature for which we are searching, by adjusting the threshold of the $RMSE$'s for which an alert is issued. Thus, statistical techniques and adaptive learning can be used to help develop optimal threshold levels for a particular pattern.

4. Conclusion

Wavelet analysis holds an as-of-yet untapped potential for analyzing usage patterns in network traffic flows, due to their ability to filter out data up to a given approximation order. Neither simple character searches nor Fourier analysis has this capability. Also, while Fourier analysis generally gives only frequency information within an analysis window, wavelet analysis gives some amount of both frequency information and the location of that frequency activity. The additional information comes at no additional computational costs, since the bases are generally applied by convolving matrix filters over the data (repeated matrix multiplication over sliding blocks of data). In particular, the multiwavelet bases introduced in this document should prove to be particularly useful in pattern-matching applications, due to

- their short support (that is, non-zero only over a small interval),
- the inclusion of splines in the approximation spaces instead of just polynomials, and
- the ability to use them on time series with left and right boundaries.

The author is currently working with computer scientists at the CyberDefense Lab (CDL) at Western Kentucky University to develop software that will utilize these bases in network security, intrusion-detection, and data-extrusion applications. Once developed, we will study the effectiveness of this type of analysis in detecting malicious behavior as compared to the more commonly used techniques, with the hope of improving detection capabilities while maintaining a low false-positive rate. The software will initially be tested in the CDL's sandbox while running attacks from its attack library with various background usage scripts being implemented, but will eventually move to less controlled network environments.

Acknowledgments: This work supported in part by the NACMAST consortium under contract EWAGSI-07-SC-0003.

References

[1] C. K. Chui and Q. T. Jiang, "Multivariate Balanced Vector-Valued Refinable Functions", in: V. W. Haussmann, K. Jetter, M.

Reimer, and J. Stöckler (Eds.), *Modern Development in Multivariate Approximation*, vol. 145, Birkhäuser Verlag, Basel (2003) pp. 71–102.

[2] G. Donovan, J. Geronimo, and D. Hardin, "Intertwining Multiresolution Analyses and the Construction of Piecewise Polynomial Wavelets", *SIAM J. Math. Anal.* **27(6)** (1996) 1791–1815.

[3] J. Geronimo, D. Hardin, and P. Massopust, "Fractal Functions and Wavelet Expansions Based on Several Scaling Functions", *J. Approx. Theory* **78:3** (1994) 373–401.

[4] D. Hardin and B. Kessler, "Orthogonal Macroelement Scaling Vectors and Wavelets in 1-D", *Arab. J. Sci. Eng. Sect. C (Theme Issue: Wavelet and Fractal Methods in Science and Engineering: Part 1)* **28:1C** (2003) 73–88.

[5] D. Hardin and D. Roach, "Multiwavelet Prefilters I: Orthogonal Prefilters Preserving Approximation Order $p \leq 2$", *IEEE Trans. Circuits Syst. II: Analog Digital Signal Proces.* **45(8)** (1998) 1106–1112.

[6] B. Kessler, "An Orthogonal Scaling Vector Generating a Space of C^1 Cubic Splines Using Macroelements", *J. Concrete Appl. Math. (Special Issue on Wavelets and Applications)* **4(4)** (2006) 393–413.

[7] J. Lebrun and M. Vetterli, "Balanced Multiwavelet Theory and Design", *IEEE Trans. Signal Process.* **46(4)** (1998) 1119–1125.

[8] J. Lebrun and M. Vetterli, "High Order Balanced Multiwavelets", in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, vol. 3, Seattle (May 1998) 1529–1532.

[9] J. Lian, "Armlets and Balanced Multiwavelets: Flipping Filter Construction", *IEEE Trans. Signal Process.* **53(5)** (2005) 1754–1767.

[10] I. W. Selesnick, "Balanced GHM-like Multiscaling Functions", *IEEE Signal Process. Lett.* **6(5)** (1999) 111–112.