# Wavelet Transform-based Network Traffic Prediction: A Fast On-line Approach

Hong Zhao[1] and Nirwan Ansari[2]

[1] Fairleigh Dickinson University, 1000 River Road T-MU1-01, Teaneck, NJ 07666, USA
[2] New Jersey Institute of Technology, University Heights, Newark, NJ 07102

High speed network traffic prediction is essential to provision QoS for multimedia applications while keeping bandwidth utilization high. Wavelet transform is a powerful technique for analyzing time domain signals. When combined with LMS, wavelet based predictor can achieve better performance than time domain predictor for MPEG-4 VBR videos and self-similar traffic. However, the computational complexity in predicting each wavelet coefficient is high. In this paper, LMK (Least Mean Kurtosis), which uses the negated kurtosis of the error signal as the cost function, is first proposed to estimate wavelet coefficients; then, by analyzing the wavelet coefficients of two consecutive data sets, Reduced Computation Complexity Wavelet LMK (RCCWLMK) is proposed to reduce the computational complexity. Simulation results for a wide range of MPEG-4 videos and network self-similar traffic show that RCCWLMK not only incurs smaller prediction error, but also reduces the computational complexity greatly.

*Keywords:* multiscale analysis, traffic prediction, MPEG-4 videos

## 1. Introduction

Traffic prediction plays an important role in network management which is concerned with controlling and allocating network bandwidth, reducing delay, and minimizing congestion in networks. The main goal of network management is to provide QoS (Quality of Service) guarantees and make efficient use of network resources. Traffic prediction, when combined with dynamic bandwidth allocation, can provide an effective solution to both achieve network efficiency and provision QoS guarantees [1]. The rapid growth of emerging multimedia applications has exerted tremendous burden on network resources, in particular those on local area networks. The self-similar nature of current network traffic, which exhibits both SRD (Short-Range Dependence) and LRD (Long-Range Dependence) in its correlated structure, has exacerbated the task of network traffic predication. A number of traffic prediction algorithms have been proposed in the literature, including neural network approaches, Normalized Least Mean Square (NLMS) approaches and wavelet transform based approaches. These existing algorithms can be broadly classified into two categories: time domain and wavelet domain approaches.

Prediction based on neural networks is a Training Based (TB) technique [2]. Neural networks are composed of simple elements operating in parallel. Connecting those elements in various ways leads to different architectures of neural networks. A neural network can be trained to perform a particular function by adjusting the values of the connections among elements [?]. The training phase is based on the history of the traffic measurements called the training data set. The neural network is then fed by the last observations in order to predict the future value. The main drawback of this approach is the amount of data required for the proper training of neural networks. In this paper, we focus on Non-Training Based (NTB)) approaches.

In the time domain, the non-training based NLMS approaches [4] [5] have been widely used

due to their simplicity and relatively good performance. In addition, they do not require any prior knowledge of the video statistics nor do they assume stationarity. Thus, they are suitable for on line VBR video prediction. However, one problem associated with NLMS is its slow convergence. For VBR video traffic characterized by frequent scene changes, this slow convergence may result in an extended period of intractability, and thus causing excessive cell loss during scene changes.

Wavelet transform is a powerful technique for analyzing time domain signals. The basic idea of wavelet transform based approaches for prediction is to first decompose the original signal into wavelet subspaces, and then, instead of predicting the original traffic directly, this approach predicts the coefficients in the wavelet domain. The predicted values of the original traffic can be constructed based on the predicted coefficients. Several wavelet domain prediction methods have been proposed in literature. These methods decompose the original traffic into wavelet coefficients and scaling coefficients at different time scales, and then predict each coefficient independently by using the NLMS algorithm. The predicted values of the original traffic can be reconstructed based on the predicted coefficients [6][7][8]. As wavelet transform reduces the eigenvalue spread of the correlation matrix of the input data, fast convergence and smaller prediction error could be achieved. However, this method introduces a complexity in the order of *(MN)*, where *M* and *N* are the wavelet filter length and the adaptive filter order, respectively. For the wavelet filter length of even as small as, 4, 8, and 16 coefficients, the computation complexity will be in the order of 4*N*, 8*N*, and 16*N*, respectively. In this paper, the Reduced Computation Complexity Wavelet Transform based Least Mean Kurtosis (RCCWLMK) predictor is proposed. By exploiting the characteristics of input traffic and redundancy in computation of wavelet coefficients, the computation complexity is reduced. In addition, LMK, which uses the negated kurtosis of the error signal as the cost function, is proposed to predict the wavelet coefficients. The rest of the paper is organized as follows. In Section 2, wavelet transform based adaptive filter is briefly reviewed. The fast RCCWLMK is proposed in Section 3; Section 4 presents

simulation results and performance analysis for both MPEG videos and network traffic. Finally, concluding remarks are included in Section 5.

## 2. Wavelet Transform-based Adaptive Filters

Wavelet transform is similar to Fourier transform, but with a completely different merit function. The main difference is that while Fourier transform decomposes signals into sines and cosines–functions localized in the Fourier space, wavelet transform uses functions that are localized in both real and Fourier spaces. Depending on the merit functions used in computation, there are many types of wavelet transform. Among which, the orthonormal discrete wavelet transform decomposes a signal into mutually orthogonal set of wavelets. In the orthogonal discrete wavelet transform, the quadrature mirror filter (QMF) pair, $\phi$ and $\psi$, form an orthonormal basis of the Hilbert space $L^2(R)$. One simple and fast wavelet basis is the Haar-basis DWT with QMF pair [9]:

$$\phi_{2^j}[n] = \begin{cases} 2^{-j/2} & (n = 0, .., 2^j - 1) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$\psi_{2^j}[n] = \begin{cases} 2^{-j/2} & (n = 0, ..., 2^j - 1) \\ -2^{-j/2} & (n = 2^{j-1}, ...2^j - 1) \\ 0 & \text{otherwise} \end{cases}$$
$$(2)$$

The discrete wavelet transform (DWT)[1] is defined as follows:

$$W_v[2^J, \frac{k}{2^J}] = \sum_n f(n)\phi_{2^J}(n - \frac{k}{2^J}), \quad (3)$$

$$W_w[2^j, \frac{k}{2^j}] = \sum_n f(n)\psi_{2^j}(n - \frac{k}{2^j}), j = 1, 2, ...J,$$
$$(4)$$

where $J$ is the maximum level of decomposition, $f(n)$ is the original signal, $W_v$ is the approximation representation of the signal, and $W_w$ is the detailed representation at various decomposition levels. Signals are represented by a weighed sum of approximation and detailed coefficients. Owing to orthonormality, the signal can be reconstructed by the inverse discrete

wavelet transform (IDWT). In the adaptive filtering scheme using wavelet transform, the projections of the input signal onto the above mentioned orthogonal subspaces are used as inputs to a linear combiner. The weights of the linear combiner can hence be updated by the LMS algorithm. For the input vector $X(n)$ of length $N$, the transformed wavelet coefficient vector can be written as follows [10]:

$$\mathbf{W}(n) = \mathbf{R}\mathbf{X}(n), \qquad (5)$$

where $\mathbf{R}$ is the wavelet transform matrix composed from wavelet basis parameters, $\mathbf{X}(n)$ is the input vector, and $\mathbf{W}(n)$ is the wavelet transformed data vector or wavelet coefficients which include approximation and detail coefficients. Instead of predicting the original signal, the wavelet based adaptive filter predicts approximation coefficients and detail coefficients independently at each scale by using a linear filter as shown in Figure 1.

$$\hat{W}_{n+1,i} = f\left(W_{n,i}, W_{n-1,i}, \dots W_{n-M+1,i}\right),$$

$$i = 1, 2 \dots N, \qquad (6)$$

where $f$ represents the linear predictor, $M$ is the length of the linear predictor, and $N$ is the number of wavelet coefficients that is dependent on the length of wavelet filters. The predicted values of the original traffic can be reconstructed based on the predicted wavelet coefficients through inverse wavelet transform. The prediction of one of those wavelet coefficients can be expressed as follows:

$$\hat{W}_{n+1,i} = \sum_{k=0}^{M-1} h_{n,i}(k) W_{n-k,i}. \qquad (7)$$

Eq. (7) shows that wavelet coefficients $\hat{W}_{n+1,i}$, $i = 1, 2 \dots, N$, are to be predicted based on the past wavelet coefficients, $W_{n,i}, W_{n-1,i}, \dots, W_{n-M+1,i}$, obtained in the past as inputs. $\mathbf{h}_{n,i}$ is the prediction coefficient vector which is time varying and updated by minimizing the mean square error $\xi$

$$\xi = \mathbf{E}[e_{n,i}^2]. \qquad (8)$$

The prediction error is

$$e_{n,i} = W_{n+1,i} - \hat{W}_{n+1,i} \qquad (9)$$

$\mathbf{W}_{n,i}, \mathbf{h}_{n,i}$ are defined as follows, where $\mu$ is the step size and $M$ is the order of the predictor.

$$\mathbf{W}_{n,i} = [W_{n,i}, W_{n-1,i} \dots W_{n-M+1,i}], \qquad (10)$$

$$\mathbf{h}_{n,i} = [h_{n,i}(1), h_{n,i}(2) \dots h_{n,i}(M)]. \qquad (11)$$

The LMS predictor minimizes the mean square error by adaptively adjusting the coefficient vector $\mathbf{h}_{n,i}$. In the normalized LMS algorithm, if we use one-step linear predictor, $\mathbf{w}_{n,i}$ is updated by

$$\mathbf{h}_{n+1,i} = \mathbf{h}_{n,i} + \frac{\mu \mathbf{W}_{n,i} e_{n,i}}{\|\mathbf{W}_{n,i}\|^2}, \qquad (12)$$
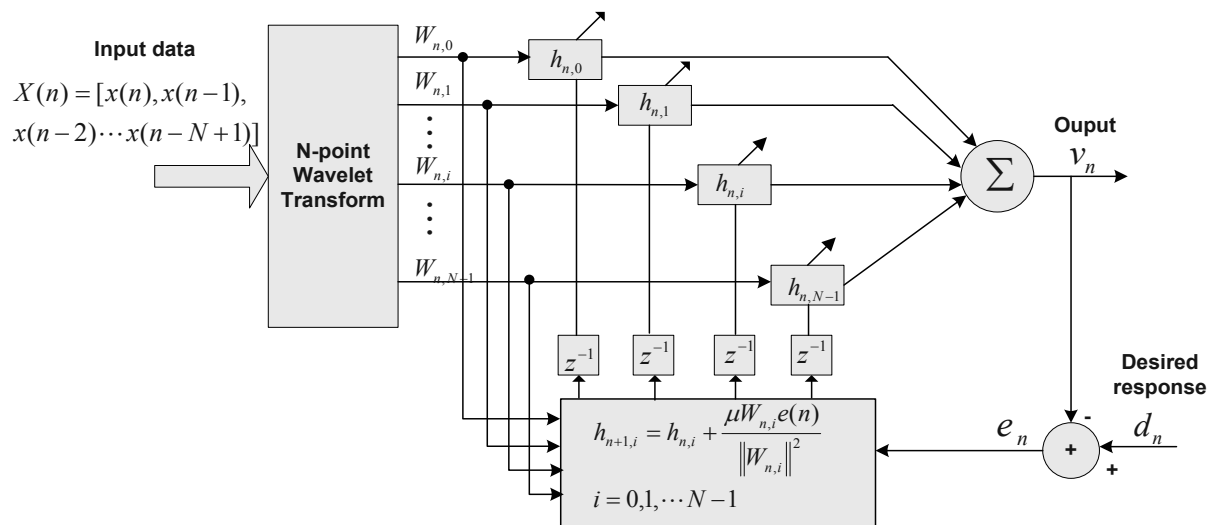


*Figure 1.* Wavelet based predictor.

where $\mu$, $0 < \mu < 2$, controls the adjustment step size. Since at time $n$ the value of $W_{n+1,i}$ is not available to compute $e_{n,i}$, $e_{n-1,i}$ is used instead. As the wavelet transform approximately diagonalizes the correlation matrix of the input data, it permits independent adaption of each of the filter coefficients. In addition, it reduces the eigenvalue spread of the correlation matrix of the input data. As a result, the wavelet predictor achieves fast convergence and smaller prediction error. It is thus expected to perform well in tracking rapid changes in a signal [10]. The only drawback is that the computation complexity is high. In the next section, we will propose the RCCWLMK algorithm in predicting MPEG videos and self-similar network traffic.

## 3. The Reduced Computation Complexity WLMK Algorithm

The main motivation for basing our predictor on the wavelet method is that wavelet transform can reduce complex temporal correlation in the MPEG-4 videos and self similar traffic to short range dependence in the wavelet domain. There is a wide variety of popular wavelet algorithms [11] including Daubechis Wavelets, Mexican Hat Wavelets, and Haar Wavelets. The Haar wavelet[12] is adopted in this paper as it is conceptually simple, fast, and memory efficient. The approximation and detail coefficients are estimated based on a higher order statistics Least Mean Kurtosis.

Let $X(n)$ be the time series representation of a video traffic

$$\mathbf{X}(n) = [x(n), x(n-1), ..., x(n-N+1)]. \quad (13)$$

The Haar wavelet transform of $\mathbf{X}(n)$ at the decomposition level of 3 is used here as shown in Figure 2. The output decomposition structure contains approximation coefficients $CA$ and detail coefficients $CD$, which are organized in the following cell structure:

$$\mathbf{W}(n) = [CA3, CD3, CD2, CD1]. \quad (14)$$

Here, $CA3$ is a vector, representing approximation coefficients at level 3. $CD3$, $CD2$, and $CD1$ are vectors of detail coefficients at levels 3, 2, and 1, respectively. The length of these vectors depends on the length of the wavelet filter. For example, the 16-point Haar wavelet transform will have 16 coefficients, which are organized in $[CA3, CD3, CD2, CD1]$ of the sizes $[2, 2, 4, 8]$ as shown in Figure 3 for data set $X(n)$. We propose to use LMK to estimate these coefficients:

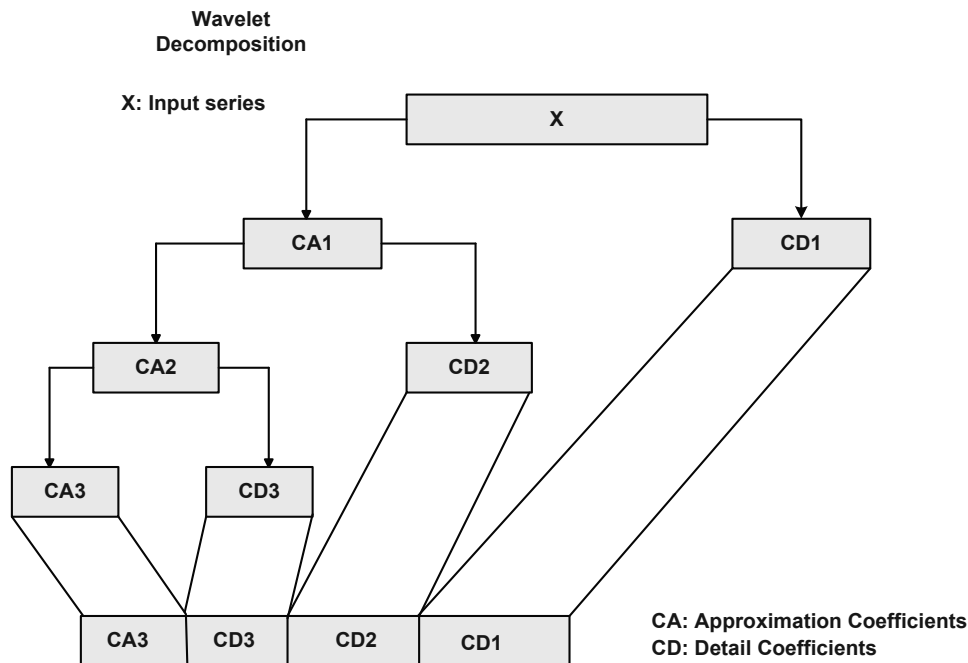$$\hat{W}_{n+1,i} = \sum_{l=0}^{p-1} h_{n,i}(l) W_{n-l,i}, \quad (15)$$
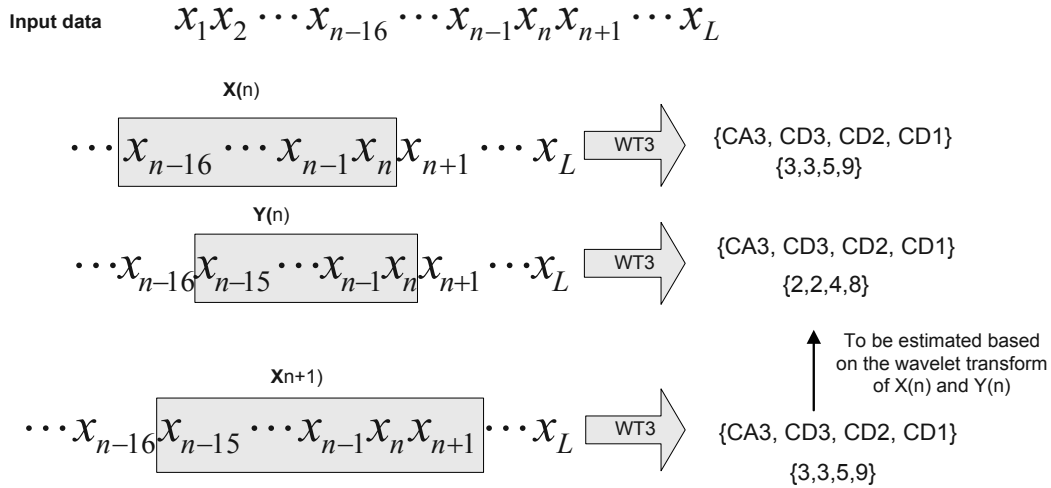


*Figure 2.* Wavelet decomposition.

**X**(n)

$\cdots \boxed{x_{n-16} \cdots x_{n-1} x_n} x_{n+1} \cdots x_L$  [WT3]  {CA3, CD3, CD2, CD1}
{3,3,5,9}

**Y**(n)

$\cdots x_{n-16} \boxed{x_{n-15} \cdots x_{n-1} x_n} x_{n+1} \cdots x_L$  [WT3]  {CA3, CD3, CD2, CD1}
{2,2,4,8}

To be estimated based
on the wavelet transform
of X(n) and Y(n)

**X**(n+1)

$\cdots x_{n-16} \boxed{x_{n-15} \cdots x_{n-1} x_n x_{n+1}} \cdots x_L$  [WT3]  {CA3, CD3, CD2, CD1}
{3,3,5,9}

*Figure 3.* Reduction of redundancy in the calculations of wavelet coefficients.

$$e_{n,i} = \hat{W}_{n,i} - W_{n,i} = \mathbf{h}_{n,i}^T \mathbf{W_{n,i}} - \mathbf{W_{n,i}}, \quad (16)$$

where $\mathbf{h}_{n,i}$, $\mathbf{W}_{n,i}$ are vectors of adaptive filter coefficients and input wavelet coefficients, respectively. Let

$$\mathbf{h}_{n,i} = [h_{n,i}(1), h_{n,i}(2), \cdots h_{n,i}(M)]^T,$$

$$\mathbf{W}_{n,i} = [W_{n,i}, W_{n-1,i}, \cdots W_{n-M+1,i}]^T.$$

In the LMK algorithm, the cost function is defined to be the negated kurtosis [14]:

$$J_{LMK}(\mathbf{h}) = 3E^2[e_{n,i}^2] - E[e_{n,i}^4]$$

$$= 3[E(\mathbf{h}_{n,i}^T \mathbf{W}_{n,i} - W_{n,i}^2]^2 - E[\mathbf{h}_{n,i}^T \mathbf{W}_{n,i} - W_{n,i}]^4.$$
$$(17)$$

The adaptive filter coefficients are then updated as follows:

$$\mathbf{h}_{n+1,i} = \mathbf{h}_{n,i} + \frac{\mu[3G(n) - e_{n,i}^2]e_{n,i}\mathbf{W}_{n,i}}{(\mathbf{W}_{n,i}^T \mathbf{W}_{n,i})^2}. \quad (18)$$

$$G(n) = \beta G(n-1) + (1-\beta)e_{n,i}^2. \quad (19)$$

Here, $\mu$ is the step size, and $\beta$ is a forgetting factor that controls the memory of the error power estimator. Since this prediction is based on a higher order statistics rather than the second order statistics used in the LMS algorithm, our earlier work [15] demonstrated that it achieves better performance than the LMS algorithm. In designing the wavelet transform-domain adaptive filter, LMK could be used to improve the prediction accuracy. It has been shown that the wavelet transform introduces a computational complexity which is relatively large as compared to time domain prediction methods, even when the very short wavelet filters are used. The computational complexity of WLMK is primarily attributed to the matrix computation in implementing the LMK method. We can see from Figure 1 that the number of $N$ LMK predictors are needed to predict $N$ coefficients. Thus, it is clear that for the number of wavelet coefficients, even as small as 4 or 8, the number of 4 or 8 LMK adaptive filters are needed [13]. This is generally not acceptable. In the following, we propose the reduced computation complexity WLMK (RCCWLMK) method to mitigate this problem.

For $N$-point Haar wavelet transform, the data input vector, $\mathbf{X}(n) = [x(n), x(n-1)...x(n-N)]$, is updated at each new iteration by letting one data sample to enter and one to leave. At iteration $n$, the input data vector $\mathbf{X}(n) = [x(n), x(n-1)...x(n-16)]$ is decomposed into the wavelet domain at level 3, and the output wavelet coefficients in the structure $[CA3, CD3, CD2, CD1]$ contain 3 approximation coefficients in $CA3$, and 3, 5, and 9 detail coefficients in $CD3$, $CD2$, $CD1$, respectively. One 0 is inserted in detail coefficient vectors $CD3$, $CD2$, $CD1$ as the number of data in $\mathbf{X}(n)$ is not even. Thus, there are 17 effective wavelet coefficients for $\mathbf{X}(n)$. In this case, 17 LMK predictors are needed to predict wavelet coefficients for reconstructing data vector $\mathbf{X}(n+1) = [x(n+1), x(n-1)...x(n-15)]$ in order to finally acquire future data $x(n+1)$. By

analyzing the wavelet coefficients of input data vector $\mathbf{Y}(n) = [x(n), x(n-1), ...x(n-15)]$, we found that there is only one approximation coefficient difference in $CA3$ between $\mathbf{X}(n+1)$ and $\mathbf{Y}(\mathbf{n})$; we have the same coefficients in $CD3, CD2, CD1$ except that 0 is inserted in detail coefficients for $\mathbf{X}(n+1)$. Thus, instead of predicting 17 coefficients to get $x(n+1)$, we only need to implement one LMK filter to obtain the approximation coefficients in $CA3$ as the other wavelet coefficients for $\mathbf{X}(n+1)$, which are available at iteration $n$, could be obtained through $\mathbf{Y}(n)$. This greatly reduces the computational complexity in WLMK and, more importantly, increases the prediction accuracy as the other wavelet coefficients, which are not predicted based on past values, are available at iteration $n$. Owing to these two merits, this algorithm is more practical and accurate for traffic prediction. Figure 3 shows the procedure of reducing the computation complexity in calculating wavelet coefficients.

## 4. Simulations and Performance Analysis of RCCWLMK

Simulations are conducted on a wide range of video traces including movies, cartoons, sport events and news. These video traces are 60 minutes long each, and are readily available from video library [16]. Simulations are also conducted on self-similar traffic generated by FARIMA. The predictions for the above traffic are achieved by the time-domain LMK and our proposed wavelet domain RCCWLMK using the one-step-ahead prediction. For a fair performance comparison, both LMK and RCCWLMK use the same order of 12 predictors. The parameters for both methods are adjusted once, based on the video traffic *StarWars IV* and kept the same for the other traffic traces. The reciprocal of the signal to noise ratio is used as the performance metric:

$$SNR_W^{-1} = \frac{\sum e_n^2}{\sum W_n^2}, \qquad (20)$$

$$SNR_T^{-1} = \frac{\sum (x(n) - \hat{x}(n))^2}{\sum x(n)^2}, \qquad (21)$$

where $W_n$ is the wavelet coefficient, and $x(n)$ is the time series representation of a video trace or self-similar traffic. $SNR_W^{-1}$ is used to compare and evaluate the performance of predicting wavelet coefficients while $SNR_T^{-1}$ is used to evaluate the performance of predicting the original video trace and self similar traffic. As both $SNR_W^{-1}$ and $SNR_T^{-1}$ are proportional to the square of prediction error to the signal strength ratio, the smaller the better.

### 4.1. Simulations on MPEG-4 Videos

The performance of the proposed RCCWLMK in predicting the approximation coefficient for *StarWarIV* is $SNR_W^{-1} = 0.0014$, and $SNR_W^{-1} = 0.0135$ for the *Jurassic Park* video trace, which were shown in [17]. The approximation coefficients are fairly well predicted. The original video traffic is thus accurately reconstructed based on the predicted wavelet coefficients. The performance is $SNR_T^{-1} = 0.0181$ for *StarWar*, and $SNR_T^{-1} = 0.0135$ for *Jurassic Park*. In Figures 4 and 6, the forecasted values by using WLMK appear very close to the actual values, even during scene changes.

Figures 5 and 7 show the LMK performance in the time domain prediction $SNR_T^{-1} = 0.0544$ for *StarWar* and $SNR_T^{-1} = 0.0425$ for *Jurassic Park*. It can be seen from these figures that our proposed RCCWLMK is more accurate than LMK in predicting the frame size of video traces. This is attributed to the fact that wavelet transform diagonalizes the correlation matrix of the input data, and thus we can use linear correlation to approximate the wavelet coefficients to exploit the long range dependence in video traffic. In addition, the wavelet transform reduces the eigenvalue spread of the correlation matrix of the input data, thus resulting in fast convergence. Figure 8 shows the rate of convergence in predicting wavelet coefficients. Figure 9 shows the rate of convergence in the time domain prediction. Note that the rate of convergence has improved tremendously by using RCCWLMK, and thus it tracks scene changes better.
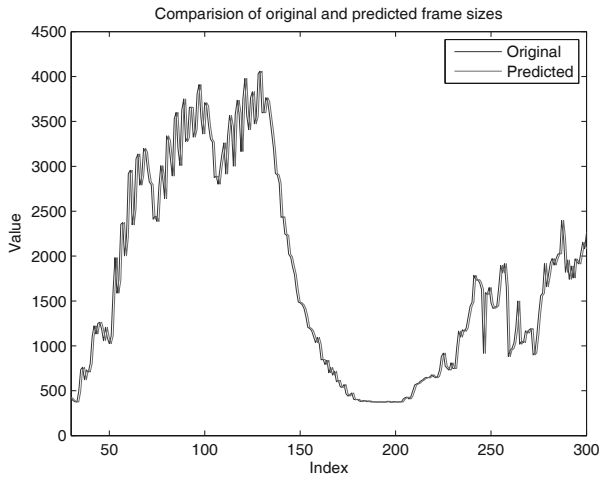
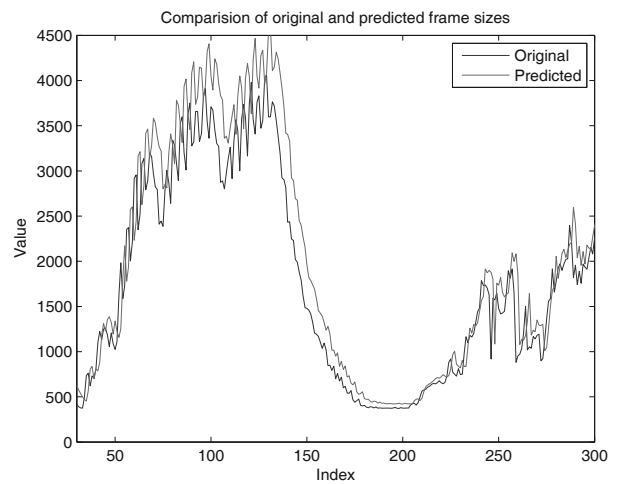*Figure 4.* Original and predicted I frame size of *StarWarIV* by RCCWLMK.



*Figure 5.* Original and predicted I frame size of *StarWarIV* by the time domain LMK.
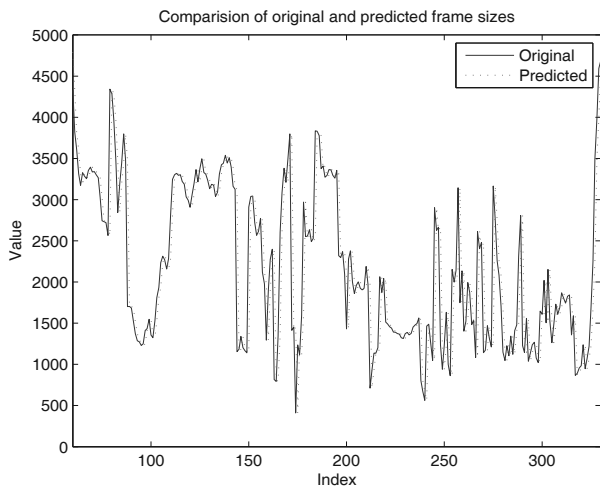


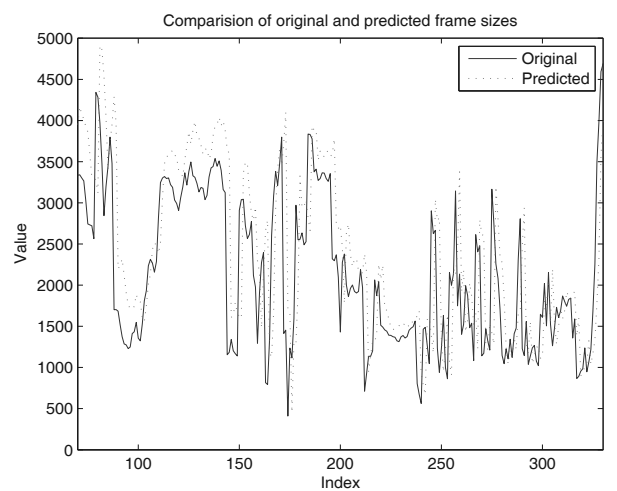*Figure 6.* Original and predicted I frame size of *Jurassic Park* by RCCWLMK.



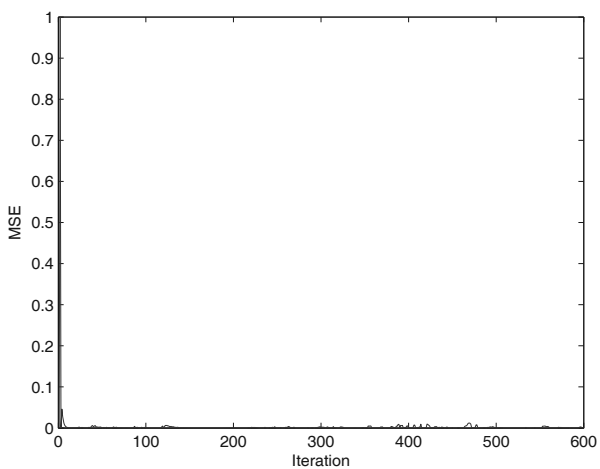*Figure 7.* Original and predicted I frame size of *Jurassic Park* by the time domain LMK.



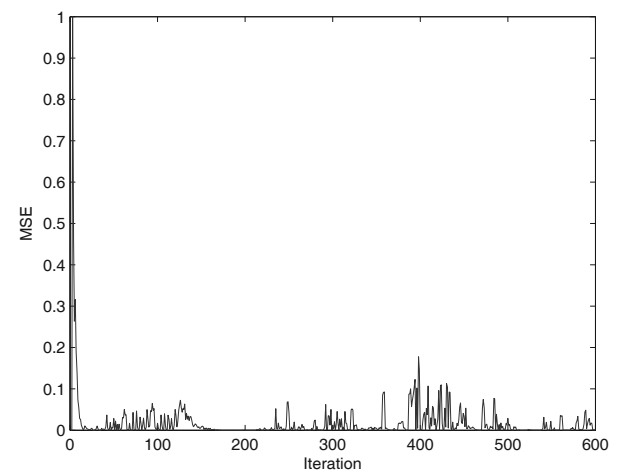*Figure 8.* Convergence property of RCCWLMK on *StarWarIV* trace.



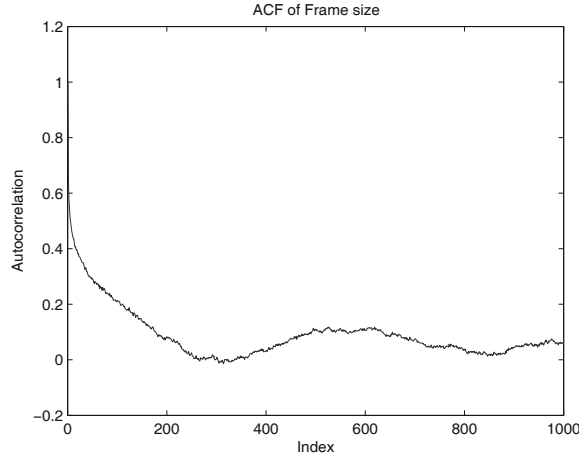*Figure 9.* Convergence property of LMK on *StarWarIV* trace.

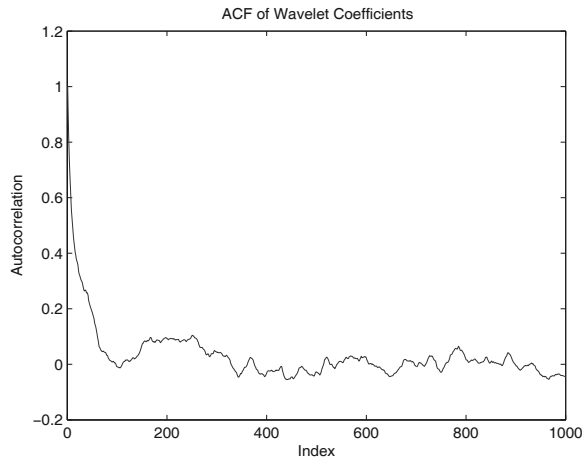*Figure 10.* ACF of P frames of *StarWarIV* trace.



*Figure 11.* ACF of Wavelet coefficients on *StarWarIV* trace.

Figures 10 and 11 show the normalized autocorrelations of the frame size traffic and its wavelet coefficients. In Figure 10, it can be seen that the temporal correlation of the traffic decays very slowly; the correlation of the corresponding wavelet coefficients shown in Figure 11, on the other hand, decays rapidly to zero, and hence demonstrating that RCCWLMK is short range dependent. Although the video traffic exhibits both SRD and LRD correlation structures, the corresponding wavelet coefficients are all SRD. The simulations show that the SRD wavelet coefficients can be better approximated by LMK than by LMS as LMK is based on a higher order statistics rather than the second order statistics used in the LMS algorithm [15]. Table 1 shows that RCCWLMK outperforms LMK for all the video traces including movies, cartoons, sport events and news. The average performance improvement is around 63.3%. The improvement is larger for video traces with high quality than the ones with medium quality. Here, H stands for high quality video. An another important feature about RCCWLMK is the reduction in the computational load, which has been achieved by exploiting the redundancy in the wavelet coefficients of two successive data sets in each new iteration. Similar results are obtained for other video traces.

| Type | Sequence | RCCWLMK | LMK | Improvement |
|---|---|---|---|---|
| Movies | Star Wars IV | 0.0181 | 0.0544 | 66.7% |
| | Star Wars IV (H) | 0.0161 | 0.0492 | 67.3% |
| | Jurassic Park | 0.0135 | 0.0425 | 68.2% |
| | Silence of the Lambs | 0.0325 | 0.1108 | 70.7% |
| | Mr. Bean | 0.0228 | 0.0614 | 62.9% |
| | Star Trek | 0.0294 | 0.0866 | 66.1% |
| | Die Hard III | 0.0355 | 0.0922 | 61.5% |
| Cartoons | Simpsons | 0.0671 | 0.1082 | 38% |
| | Aladdin | 0.0357 | 0.0821 | 56.5% |
| | Futurama | 0.0219 | 0.0489 | 55.2% |
| Sports | Alpin ski | 0.0390 | 0.1002 | 61.1% |
| | Soccer(H) | 0.0167 | 0.0513 | 67.5% |
| | Formula 1 | 0.0223 | 0.0678 | 67.1% |
| | Formula 1 (H) | 0.0177 | 0.0548 | 67.7% |
| News | ARD News (H) | 0.0098 | 0.0316 | 69.0% |
| | N3 Talk (H) | 0.0153 | 0.0469 | 67.4% |

*Table 1.* Performance Comparison of RCCWLMK and LMK using $SNR_T^{-1}$ for different types of video traces.

## 4.2. Simulations on Self-similar Traffic

We also apply the RCCWLMK to estimate the self-similar traffic generated by the FARIMA model. In our simulations[18], the Hurst parameter of the generated self-similar traffic is set as $H = 0.8$; this experiment is repeated 50 times and the resulting ensemble averaged MSE (Mean Square Error) is shown in Figure 12. Here, the iteration represents the iteration index. The performance of RCCWLMK is compared with time domain LMS (Least Mean Square). Note that RCCWLMK produces significantly smaller prediction error than LMS does. As we used a training sequence for the first 30 data points in RCCWLMK, the prediction er-
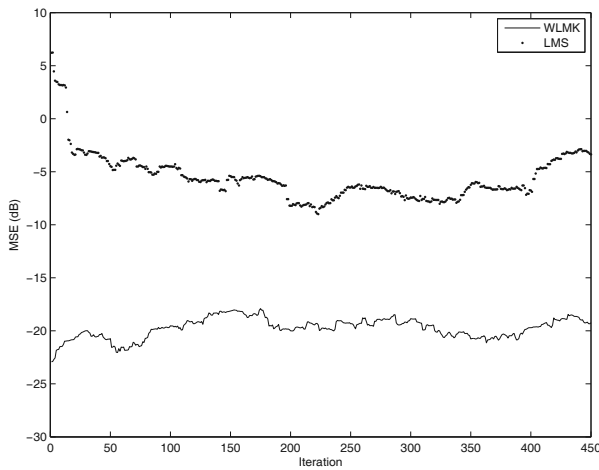
rors become a little bit larger than those after the training sequence.

Figure 13 shows the performance of the proposed RCCWLMK in predicting self-similar traffic randomly generated by the FARIMA model. The performance is quantified by $SNR_T^{-1} = 0.0165$. Note that the predicted traffic values are very close to the original values, even during some bursts. For the same self-similar traffic, the LMS performance is quantified by $SNR_T^{-1} = 0.0278$. Figure 14 shows the performance of LMS. From these figures, we can see the proposed RCCWLMK is more accurate than LMS. Again, this is attributed to the fact that wavelet transform diagonalizes the correlation matrix of the input data so that we can use linear



*Figure 12.* Comparison of averaged MSE (dB) of RCCWLMK and LMS versus number of iteration for the self-similar traffic.



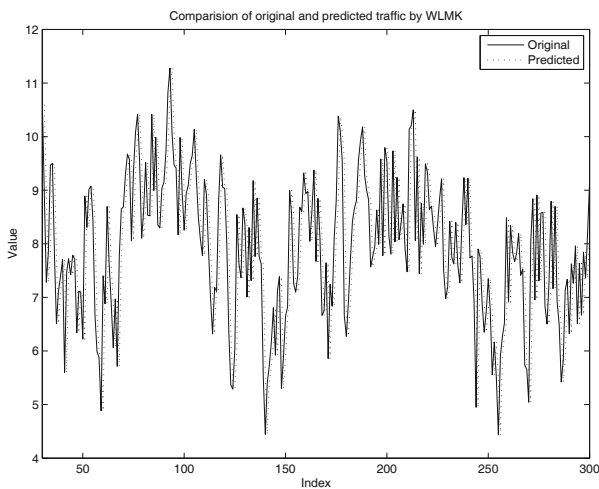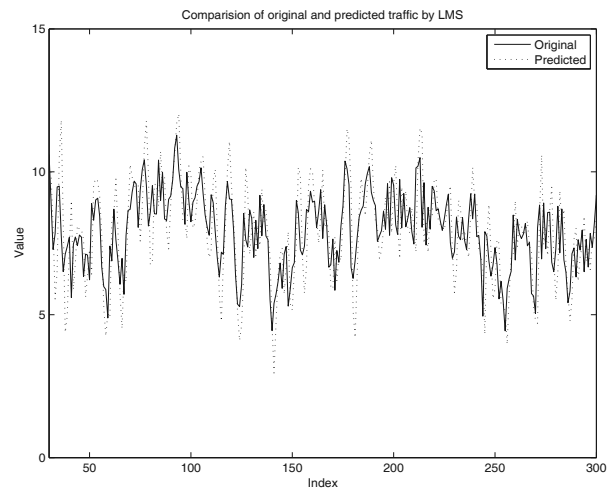*Figure 14.* Original and predicted self-similar traffic by LMS.



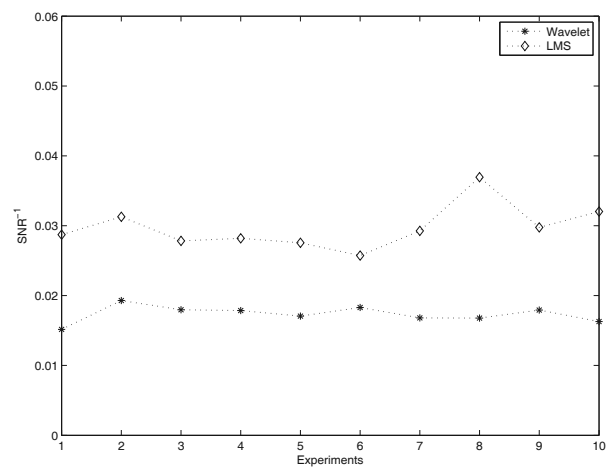*Figure 13.* Original and predicted self-similar traffic by RCCWLMK.



*Figure 15.* Comparison of the $SNR_T^{-1}$ performance of RCCLMK and LMS predictors on self similar traffic.

correlation to approximate the wavelet coefficients to exploit the long range dependence in self similar traffic.

In Figure 15, LMS and RCCWLMK are compared among ten experiments in terms of $SNR_T^{-1}$. Note that the proposed RCCWLMK incurs much smaller prediction error in all experiments than LMS does; owing to the space limit, results of only ten experiments are presented here. Another important feature about the proposed WLMK is the reduction in computational load, which has been achieved by exploiting the redundancy in the wavelet coefficients of two successive data sets in each new iteration.

## 5. Conclusions

The ability to predict traffic within a network is critical to network design and management. In this paper, RCCWLMK has been proposed to predict the MPEG-4 video traffic and self-similar traffic. The characteristics and contributions of the proposed RCCWLMK can be summarized as follows:

- Raw data traffic is first decomposed into a complete set of wavelet representation–wavelet coefficients by using the Haar algorithm at level 3.

- Future wavelet coefficients are then estimated by LMK which is based on a higher order statistics and provides better prediction performances than LMS.

- By further analyzing the wavelet coefficients of two consecutive data sets, RCCWLMK not only reduces the computation complexity, but also provides more accurate prediction, thus making this method more practical for on-line applications. For an N-point wavelet transform, one LMK adaptive filter is required, and the other wavelet coefficients can be obtained by exploiting the redundancy of wavelet transform of a set of successive data. Thus, RCCWLMK greatly reduces the computational complexity and improves the prediction accuracy.

- Simulation results show that the proposed RCCWLMK outperforms LMK on a wide range of MPEG-4 videos including movies, cartoons, sports and news, and self-similar

traffic. It reduces the complex temporal correlation into short range dependence in the wavelet domain. Fast convergence speed and smaller prediction error are achieved for both MPEG-4 videos and self-similar traffic. Especially for VBR video traffic characterized by frequent scene changes, the proposed RCCWLMK tracks scene changes much better with less computational complexity.

## References

[1] X. WANG, S. JUNG, J. S. MEDITCH, "Dynamic bandwidth allocation for VBR video traffic using adaptive wavelet prediction", *ICC98 (Interational Conference on Communications)*, vol. 1, pp. 549–553, (1998).

[2] M. F. ZHANI, H. ELBIAZE, "Analysis and prediction of real network traffic", *Journal of Networks*, vol. 4, No. 9, pp. 855–865, (2009).

[3] G. RUTKA, "Neural network models for Internet traffic prediction", *Electronics and Electrical Engineering*, No. 4(68), pp. 55–58, (2006).

[4] A. M. ADAS, "Using adaptive linear prediction to support real-time VBR video under RCBR network service model," *IEEE/ACM Transaction on Networking*, vol. 6, pp. 635–644, (1998).

[5] L. LANFRANCHI, B. BING, "Long-term prediction of MPEG video traffic for broadband cable Networks," *Sarnoff Symposium 2010*.

[6] X. WANG, X. SHAN, "A wavelet-based method to predict Internet traffic," *Interational Conference on Communications, Circuits and Systems and West Shin Exposition*, pp. 690–694, (2002).

[7] S. ATTALLAH, "The wavelet transform-domain LMS algorithm: A more practical approach", *IEEE Transaction on Circuits and Systems–Analog and Digital Signal Processing*, vol. 47, No. 3 March, (2000).

[8] S. AHMAD, A. POPOOLA, K. AHMAD, "Wavelet-based multiresolution forecasting", *Unis Technical Report*, June 2005.

[9] M. VETTERLI, J. KOVACEVIC, Wavelet and subband coding, *Pretence Hall*, 1995.

[10] N. ERDOL, F. BASBUG, "Wavelet transform based adaptive filters: analysis and new results", *IEEE Transactions on Signal Processing*, vol. 44, No. 9, pp. 2163–2171, (1996).

[11] C. S. BURRUS, H. GUO, R. A. GOPINATH, "Introduction to Wavelets and Wavelet Transforms: A Primer", *Printence Hall*, ISBN:0-13-489660-9.

[12] I. KAPLAN, "Applying the Haar wavelet transform to time series information",
http://www.bearcave.com/misl/misltech/wavelets/haar.html.

[13] S. ATTALLAH, M. NAJJIM, "A fast wavelet transform domain LMS algorithm", *Proc. Intl. Conf. on Accoustics, Speech, and Signal Processing*, pp. 1343–1346, (1996).

[14] O. TANRIKULU, A. G. CONSTANTINIDES, "Least-mean Kurtosis: A novel high-order statistics based adaptive filltering algorithm," *IEEE Electron. Lett.*, pp. 189–190, (1994).

[15] H. ZHAO, N. ANSARI, Y. Q. SHI, "Network traffic prediction using Least Mean Kurtosis", *IEICE Transaction on communication*, vol. E89-B, No. 5, pp. 1672–1674, (2006).

[16] F. H. P. FILTZEK, M. REISSLEIN, "MPEG-4 and H.263 video traces for network performance Evaluation (Extended Version)", video traces at `http://www.tkn.tu-berlin.de/research/trace/ltvt.html`.

[17] H. ZHAO, "A practical wavelet domain LMK algorithm for predicting multimedia traffic", *Interational Conference on Communications* (2008).

[18] H. ZHAO, "Multiscale analysis and prediction of network traffic", *International Performance Computing and Communications Conference* (2009).

*Contact addresses:*
Hong Zhao
Fairleigh Dickinson University
1000 River Road T-MU1-01
Teaneck, NJ 07666, USA
e-mail: `zhao@fdu.edu`

Nirwan Ansari
New Jersey Institute of
Technology University Heights
Newark, NJ 07102
e-mail: `nirwan.ansari@njit.edu`

HONG ZHAO received the B.S., M.S., and PhD from Taiyuan University of Technology, Xian Jiaotong University, and New Jersey Institute of Technology, respectively, all in Electrical Engineering. She is an Assistant Professor of Electrical and Computer Engineering at Fairleigh Dickinson University, New Jersey, US. Her research interests include multimedia streaming over networks, cross-layer design for wireless networks, digital signal processing and network steganography and steganalysis. She has served as a TPC member, technical paper reviewer and book reviewer for IEEE conferences, journal magazines and book publishers. She also serves as Chair of the IEEE North Jersey Computer Society Chapter.

NIRWAN ANSARI received the B.S.E.E. (summa cum laude with a perfect GPA) from the New Jersey Institute of Technology (NJIT), Newark, in 1982, the M.S.E.E. degree from University of Michigan, Ann Arbor, in 1983, and the PhD degree from Purdue University, West Lafayette, IN, in 1988. He joined NJIT's Department of Electrical and Computer Engineering as Assistant Professor in 1988, became a tenured Associate Professor in 1993, and has been a Full Professor since 1997. He has also assumed various administrative positions at NJIT. He was Visiting (Chair) Professor at several universities. He authored Computational Intelligence for Optimization (Springer, 1997) with E.S.H. Hou and edited Neural Networks in Telecommunications (Springer, 1994) with B. Yuhas. He has contributed over 400 technical papers, over one third of which were published in widely cited refereed journals/magazines. He has also guest-edited a number of special issues, covering various emerging topics in communications and networking. His current research focuses on various aspects of broadband networks and multimedia communications.

Prof. Ansari has served on the Editorial Board and Advisory Board of eight journals, including as a Senior Technical Editor of the IEEE Communications Magazine (2006-2009). He has served the IEEE in various capacities such as Chair of the IEEE North Jersey Communications Society (COMSOC) Chapter, Chair of the IEEE North Jersey Section, Member of the IEEE Region 1 Board of Governors, Chair of the IEEE COMSOC Networking Technical Committee Cluster, Chair of the IEEE COMSOC Technical Committee on Ad Hoc and Sensor Networks, and Chair/Technical Program Committee Chair of numerous conferences/symposia. Some of his recognitions include an IEEE Fellow, the 2008 NJIT Excellence in Teaching Award in Outstanding Professional Development, a 2008 IEEE MGA Leadership Award, the 2009 NCE Excellence in Teaching Award, a couple of best paper awards (IC-NIDC 2009 and IEEE GLOBECOM 2010), a 2010 Thomas Alva Edison Patent Award, and designation as an IEEE Communications Society Distinguished Lecturer (2006-2009, two terms). He was also granted over fifteen US patents.