A Wavelet Based Long Range Signal Strength Prediction in Wireless Networks

Xiaobo Long and Biplab Sikdar
Electrical, Computer and System Engineering
Rensselaer Polytechnic Institute, 110 8th Street, Troy NY 12180

Abstract— **Prediction of rapidly time varying fading channel conditions enables adaptive data transmissions in wireless systems, which in turn improves the quality of service for end users and reduces the power consumption for data transmissions. Most of the existing long range prediction methods for fast fading in wireless networks use autoregressive (AR) models and make the assumption that the input fading signal is stationary and wireless channel parameters vary slowly [1]. In this paper, we provide a method to predict the non-stationary received signal strength in a more realistic and fast varying wireless environment, using multiresolution wavelet analysis. We first use discrete wavelet decomposition (DWT) to decompose the signal strength trace into components at different scales, then use AR and linear regression models to predict small, medium and large scale fading components respectively, and finally synthesize the output signal of our prediction algorithm. By properly choosing the wavelet basis, we map the non-stationary signal strength trace into stationary wavelet detail coefficients and use them as input to the AR predictor at different scales. Longer prediction range is easily achieved by choosing the appropriate maximum decomposition scale, while still achieving low prediction error. Our experimental results shows that our wavelet based algorithm outperforms existing time-domain AR prediction methods in terms of both prediction accuracy and computational complexity.**

I. Introduction

In wireless networks, the fading wireless channel is a time-varying (TV) system. This is not only because the channel is susceptible to noise, interference, multipath fading and shadow fading, but also because the impediments in the channel change over time in unpredictable ways due to user movement that cause path loss fading in the received signal strength. Multipath fading (also called small-scale fading) causes changes in the signal strength within the order of one wavelength. Shadow (medium-scale) fading is influenced by the spatial movements in the order of tens of wavelengths and creates random variations in the average power of the received signal. Path loss (large-scale fading) is caused by spatial movements in the order of hundreds of wavelengths making the average power level vary in power-law fashion with path length. The above three fading components are mutually independent of each other. The received signal strength, which is the sum of the above fading components in wireless networks is usually a non-stationary process. Although small scale fading can be assumed to be a stationary Rayleigh process in many wireless propagation models, in some severe fading environments, it is non-stationary and "worse than rayleigh" [3].

Prediction methods for fast fading channels as TV systems have received considerable attention in literature and many schemes exist that can predict small scale fading channel coefficients, at a time horizon of several milliseconds. However, in many scenarios and applications, predictions on a longer time scale, such as that in the order of hundreds of milliseconds may be desirable. For example, dynamic route selection methods based on link quality in multi-hop wireless networks require abundant time to explore and select new paths when the signal to noise ratio on the previous route degrades. In such scenarios, large scale fading needs to be predicted, in addition to small scale fading. Further more, many popular methods to predict fast fading employ adaptive algorithms, which assume that the system parameters are slowly varying. If the parameters vary too fast, adaptive algorithms are not able to track the TV system's time evolution [9]. To overcome this problem, multiresolution wavelet analysis provides a natural way to represent signal at different time and frequency resolutions. By expansion of the signal into wavelet basis and analyzing the signal in different scales, both local details (small scale fading) as well as the global trend of the signal (medium and large scale fading) can be captured.

Our aim is to provide an accurate, low-complexity, on-line mechanism for long range prediction based on past signal strength measurements. We first use DWT to decompose the signal strength trace into different scales components. We then apply an AR model on wavelet detail coefficients at each scale, predict the approximation coefficients using a linear regression model, then synthesize the predicted output signal. Our proposed algorithm has the following advantages: First, by appropriately choosing the wavelet basis, we can map a non-stationary signal into stationary detail coefficients. Although AR models may work poorly with non-stationary data, our algorithm doesn't need the assumption of stationarity, making the proposed method robust in realistic mobile radio environments. Second, since many of the wavelets coefficients are zero, they can represent the TV system with fewer parameters and make the prediction algorithm more computationally efficient. Last but not the least, longer range prediction can be easily achieved by choosing larger value for the maximum decomposition scale, while still achieving low prediction error. We demonstrate the above observation through experimental results and compare the prediction error of our scheme with existing time-domain AR based prediction methods.

The rest of the paper is organized as follows: Section II presents the related work. Section III presents the experimental setup. Section IV describes the proposed methodology for long range signal strength prediction. In Section V we compare our prediction results with predictions from existing AR models. Finally, Section VI presents the concluding remarks.

II. Related Work

The deterministic channel model and an AR signal model with its parameter estimation schemes for predicting the mobile radio channel are compared in [1]. Based on realistic simulation data and measurements, the authors show that the AR model performs best. In [2], the authors present a mechanism for reliable prediction of fast fading channel coefficients several milliseconds in advance using an adaptive AR model. They assume that the signal is stationary with slowly varying parameters. In [4] authors show that the better performance of the AR based prediction algorithms is achieved due to their lower sampling rate. However, the prediction for more realistic non-stationary data is not improved significantly by the lower sampling rate. What's more, the iterative AR models used in their method have the problem of error propagation for prediction steps larger than one. In [5] an adaptive channel prediction algorithm using Kalman filtering is proposed. A mechanism based on support vector machines and nonlinear regression for long range prediction of fading channels is proposed in [7]. However, the prediction range in all the above methods is only about a wavelength and cannot be used in applications such as dynamic route selection in multi-hop networks where predictions of tens of wavelength ahead are desirable. A prediction algorithm based on multi-layer perception (MLP) is proposed in [6]. However, this requires measured and pre-processed channel data consisting of upto 6072752 patterns to be first developed for a given site, making it computationally complex for on-line deployments.

III. Signal Strength Measurement Methodology

In this section we outline the methodology applied to obtain the traces of the signal strength. It is well known that the performance of a wireless system depends on the environment in which it operates. This dependence on the environment mainly comes from the variation in radio channel behavior in different sites. One of the main aims of this paper is to develop a non site-specific prediction mechanism for wireless link quality and validate its performance in different environments. For this reason, measurements were carried out in multiple, diverse locations. Both outdoor and indoor scenarios were considered in our measurements and were conducted in various buildings and locations in the RPI campus. The indoor measurements were carried out in three different buildings. The first is the Johnsson Engineering Center which primarily consists of rooms for faculty and space for laboratories. In the floors of this building where the measurements were conducted, concrete walls were the main cause of signal obstruction and attenuation. The second building was the campus library where the large number of metallic bookshelves were the primary source of attenuation and shadowing. The third indoor setting was the Student Union dinning hall where there were lesser obstructions. In addition to these, outdoor measurements were also conducted at various locations in the campus. In addition to the measurements carried out at the university campus, a set of measurements was also carried in home settings, in an apartment. In all these measurement scenarios, multiple traces for the signal strength were collected as the user walked around inside the building or outside. More than 40 signal strength measurement traces were collected with the receiver moving at walking speed at 13 different environments.

Signal strength measurements were done using a LINKSYS Wireless-G Broadband Router as the access point (AP) and IBM T42 laptop, running Linux Feroda core 5, with built in PH12127-E IBM 802.11a/b/g Wireless LAN Mini PCI adapter as receiver. The signal strength measurements were directly provided from the card by the `madwifi`-0.9.2 driver used for the card. The driver uses RSSI as the basic measure for signal strength which is converted to dBm. The driver assumes a constant noise level of -96dBm since this is the thermal noise for 20MHz OFDM signals, plus an additional 5dBm noise from the amplifiers. The SNR levels are then obtained by SNR(dBm)=Signal(dBm)-Noise(dBm). The actual signal strength measurements were conducted while the laptop received packets from the AP. The packets were from an UDP video data stream transmitted at a data rate of 54Mbps in 802.11g wireless network. We collected signal strength measurement every 0.01 seconds.

IV. Methodology

A. Overview

In the proposed mechanism, a sliding window of size M is defined that consists of the last M signal strength measurements, $y = [y_1, \cdots, y_M]$. We predict future samples based on y. Our prediction algorithm consists of three steps. The first step is to obtain wavelet coefficients of y using DWT. The wavelet basis is carefully chosen so that we can map the non-stationary signal strength series into stationary detail coefficients. The maximum decomposition scale is chosen according to the desired prediction range. The second step is prediction of wavelet coefficients. We apply an AR model to predict the wavelet detail coefficients at each scale and employ a linear regression model to predict approximation coefficients. Both the AR and linear regression models are chosen according to the nature of the wavelet detail and approximation coefficients respectively. The third step is then to synthesize the predicted output signal from the predicted wavelet coefficients.

B. Wavelet Decomposition

Before we employ any prediction model, we first use DWT to map the measured data y to its wavelet coefficients. Our motivation is that using DWT we can decompose a signal strength trace into its three independent components: multipath fading, shadow fading and path loss fading. Wavelet transform provides the time-frequency representation of the signal at different scales. A signal can be presented by its approximation at any scale (octave) J where $1 \leq J \leq J_{MAX}$ ($J_{MAX} = \log_2(M)$ is the maximum decomposition scale and

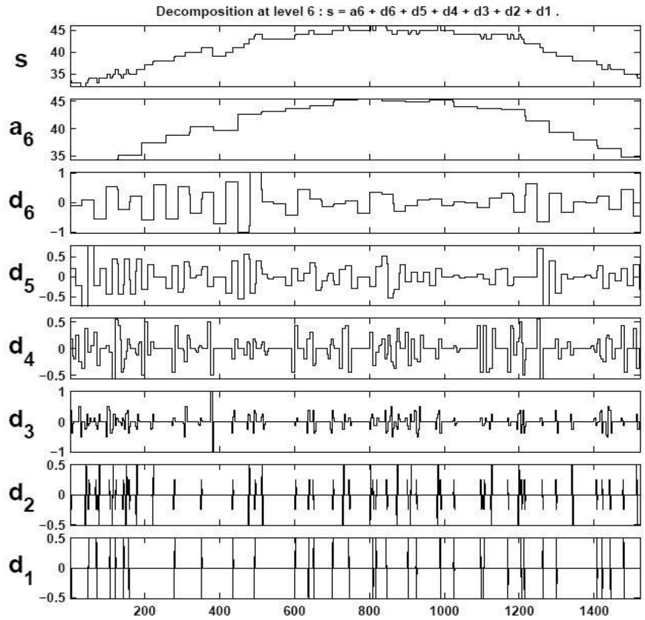

Fig. 1. Wavelet decomposition at level 6: $s = a_6 + d_6 + d_5 + d_4 + d_3 + d_2 + d_1$

is determined by the length, M, of the time series), plus all the details at lower scales $j, 1 \leq j \leq J$. The wavelet decomposition formula is given by

$$y = \text{approx}_J + \sum_{j=1}^{J} \text{details}_j$$
$$= \sum_k a_x(J,k)\phi_{J,k} + \sum_{j=1}^{J}\sum_k d_x(j,k)\varphi_{j,k} \quad (1)$$

where $a_x(j,k)$ and $d_x(j,k)$ are the wavelet transform approximation and detail coefficients, respectively, at scale j and time k. $\phi_{J,k}$ is the wavelet function transformed from the mother wavelet function ϕ and $\varphi_{j,k}$ is the scale function. As an example, Figure 1 shows that our experimental fading signal strength trace s may be decomposed into an approximation at octave $J = 6$ (a_6) plus all the details at octave $j = 1, \cdots, 6$ ($d_1, \cdots, d_6$). For our experimental data, there are 1524 samples in each trace, which covers about 50 meters walking distance, so our sampling rate is 0.033m/sample. At 2.437GHz frequency (channel 6 in IEEE 802.11g), the radio wavelength is 0.1231m. At octave 5, the signal trace is represented by only 32 samples (1.56m/sample), which means that the time resolution of the signal approximation at octave 5 is of the order of tens of the wavelength, which corresponds to the shadow fading scale. For octaves 1-3, the time resolution of the signal is of the order of a wavelength, and corresponds to multipath fading. Octave 4 corresponds to the transition region between medium and small scales. Large scales are octaves larger than 6 and correspond to path loss fading.

C. Wavelet Basis and Maximum Decomposition Scale

Choosing the proper wavelet basis and the maximum decomposition scale is essential in our algorithm design. Assume that signal y is decomposed into n_j wavelet coefficients, $d(j,k), k = 1, \cdots, n_j$ at each octave j. Authors of [8] show that for each j, $d(j, \cdot)$ is a stationary, short range dependent process with zero mean, provided that

$$N \geq (\alpha - 1)/2 \quad (2)$$

where N is the number of vanishing moments of wavelet mother function ϕ and α is the estimated scale-invariance parameter which can be obtained by a LRD estimator tool. More details about the meaning of α and its estimation can be found in [8]. In our experiments, N is carefully chosen so that Equation 2 is always satisfied. For example, for the trace s shown in Figure 1, estimated α is 0.85 and N is chosen to be 4. In our design, orthogonal Daubechies wavelets that offer a high number of vanishing moments from 4 to 20 are chosen.

Let M be the length of historical data sequence y, which is the input to our prediction algorithm. We decompose y into $d(j,k), k = 1, \cdots, m_j$, where m_j is the number of detail coefficients at octave j. Then at octave j, $m_j = M/2^j$ and $m_{j+1} = m_j/2$. If we do prediction for detail coefficients at each scale, at octave j, $d(j, m_j + step_j) = F(d(j,1), ...d(j, m_j))$ where F is some prediction function and $step_j$ is the prediction range at octave j. Let the maximum decomposition scale be J. At least one data should be predicted at scale J. i.e. $step_J \geq 1$. Since $m_{j+1} = m_j/2$, in order to synthesize the data from prediction of wavelet coefficients, the number of wavelet coefficients that need to be predicted at scale j is at least 2^{J-j}. The prediction range L of output prediction data is therefore

$$L \geq 2^J \quad (3)$$

By increasing J, we can enhance the prediction range to a desirable value. The maximum value J is determined by the length of x since $J_{max} = \lfloor log_2 M \rfloor$, where function $\lfloor \cdot \rfloor$ rounds the element toward zero. In our experiments, we choose $L = 2, 4, 8, 16, 32, 64$. $L = 64$ corresponds to a prediction range in the order of tens of wavelength ahead and is desirable for long range prediction applications.

D. Prediction Algorithm

After we obtain the wavelet coefficients, we then predict the future coefficients based on historical ones. We use an iterative AR model to predict the detail coefficients. Prediction step at scale j is $step_j = 2^{J-j}$. The Minimum Mean Square Error (MMSE) prediction of the future one step sample $\hat{d}(j, m_j + 1)$ based on previous samples $d(j,1), \cdots, d(j, m_j)$ is given by

$$\hat{d}(j, m_j + 1) = \sum_{k=1}^{p} a_k d(j,k) \quad (4)$$

where $1 \leq p \leq m_j$ is the AR order, and the optimal coefficients a_k are determined by orthogonal principle, which leads to Yule-Walker equations and can be solved by the Levinson-Durbin recursion [10].

For prediction steps more than one, an iterative AR model is employed. The prediction $\hat{d}(j, m_j + 1)$ obtained from the previous step is used together with $d(j,1), \cdots, d(j, m_j)$ as input of the AR model to predict $\hat{d}(j, m_j + 2)$. Iteratively, $\hat{d}(j, m_j + n)$ where $n = 2, \cdots, step_j$ is estimated based on previous samples $d(j, 1), \cdots, d(j, m_j)$ and previous estimated samples $\hat{d}(j, m_j + 1), \cdots, \hat{d}(j, m_j + n - 1)$.

Recall that signal y is decomposed into an approximation at octave J, $a(J, \cdot)$ plus all the details at octave $j = 1, \cdots, J$ ($d(1, \cdot), \cdots, d(J, \cdot)$). We also need to predict the approximation coefficients $a(J, \cdot)$, which represents the trend of signal and can not be simply assumed to be stationary. For $J = 6$, approximation $a(J, \cdot)$ represents large scale fading which makes the average power level of signal strength vary in power-law fashion with path length. We use a linear regression model to predict $a(J, \cdot)$ in our algorithm.

Let the previous approximation coefficients for segment data y be $z = (a(J, 1), \cdots, a(J, n))$, where $n = M/2^J$ and M is the length of y. A linear regression model, $\widehat{Z}$=α+βW, is used to fit the data $z = a(J, 1), \cdots, a(J, n)$ with $w = [1, \cdots, n]$. The parameters of this regression model are given by

$$\beta = \frac{\Sigma_{wz} - \overline{w} \cdot \overline{z}}{\Sigma_w^2 - n\overline{w}^2} \tag{5}$$

$$\alpha = \overline{z} - b\overline{w} \tag{6}$$

where Σ_{wz} is the cross covariance of variable W and Z, $\overline{w}$ and $\overline{z}$ are the means of W and Z respectively, and Σ_w^2 is the variance of W. The predicted coefficient $a(J, n+p) = z(n+p)$ is then given by

$$\widehat{z}(n + p) = \alpha + \beta(n + p) \tag{7}$$

where p is the prediction step and chosen to be one. $p = 1$ is sufficient for prediction range of approximation coefficients, since it leads to the final prediction range $L = 2^J$ desirably long by choosing proper J. Besides, since the sample size of approximation coefficients $M/2^J$ is small, it's hard to achieve high prediction accuracy and at the same time, achieve a long prediction range based on very limited previous samples. Our experimental results show that $p > 1$ here will lead to significant decrease in prediction accuracy.

We then use reversed DWT to synthesize the signal from predicted wavelet coefficients to obtain final prediction results. For perfect reconstruction, the synthesis filters are identical to the analysis filters except for a time reversal. Our prediction algorithm is given in Algorithm 1.

Algorithm 1 Prediction Algorithm

J: `maximum wavelet decomposition scale;`
n_j: `number of` $d(j, \cdot)$ `at scale j;`
n_{J+1}: `number of` $a(J, \cdot)$;
L: `prediction range;`
x: `signal strength trace;`
x_k: k_{th} `signal strength measurement;`
$\widehat{x}_{k+L}$: `prediction of` $(k + L)_{th}$ `sample;`
w: `sliding window size;`
$k = w$;{k is a loop control index};
PREDICTION
while (1) **do**
  read new incoming measurement x_k;
  $y = [x_{k-w+1}, \cdots, x_k]$;
  DWT decomposition of $y = d_1 +, \cdots, +d_J + a_J$;
  $j = 1$;
  repeat
    $m = 1$;
    repeat
      use AR model to predict $d(j, n_j + m)$ based on $d(j, 1), \cdots, d(n_j + m - 1)$;
      $m + +$;
    until $m > 2^{J-j}$
    $j + +$;
  until $j > J$
  use linear regression model to predict $a(j, n_{J+1} + 1)$ based on $a(J, 1), \cdots, a(J, n_{J+1})$;
  synthesis wavelet coefficients to get $\widehat{x}_{k+L}$;
  $k + +$;
end while

V. Performance Evaluation

In this section we design and report on a set of measurement based experiments to validate our proposed prediction methodology. We also compare our methodology with one of the most popular schemes that have been proposed in literature. The performance metrics used are prediction accuracy and computational complexity. We use the normalized mean square error (NMSE), defined as

$$NMSE = \frac{1}{m} \sum_{j=1}^{m} \frac{\sum_{i=1}^{n} (x_i - \widehat{x}_i)^2}{\sum_{i=1}^{n} x_i^2} \tag{8}$$

to measure the prediction accuracy. In Equation (8) n is the sample size and x_i and $\widehat{x}_i$ for $i = 1 \ldots n$ are the measured and predicted data, respectively. m represents the number of traces collected at a given location with the user following the same path and $m = 3$ in our experiments.

A. Prediction Accuracy

The results from our prediction mechanism and its comparison with the time domain adaptive AR model method of [2] for six different prediction ranges is shown in Table I. $L = 2, 4, 8$ correspond to the small scale fading prediction ranges, for which [2] is designed. To further increase the prediction range, $L = 16, 32$ and 64 which correspond to medium and large scale fading are tested in our experiments. The NMSE for the two methods is tabulated for five different locations, both indoors and outdoors, and the methodology of this paper is labeled "wavelet". Parameters including size of sliding window and AR order for both methods were optimized through empirical experimental tests. The time domain AR model performs poorly if directly applied on our data traces which can't be assumed stationary. In order to make the comparison fair, we removed the trend of the data before it is input to the AR model by removing the average of the segment data along the sliding window. This average value

TABLE I

PREDICTION PERFORMANCE COMPARED WITH TIME-SERIES AR MODEL

NMSE ($\times 10^{-4}$) for trace 1						
L	2	4	8	16	32	64
AR	1.1057	2.0427	3.6634	6.6236	13	31
wavelet	0.8042	1.1754	2.2129	3.9321	8.1151	26
NMSE ($\times 10^{-4}$) for trace 2						
L	2	4	8	16	32	64
AR	1.2329	2.0217	3.2789	5.5334	10	19
wavelet	0.9985	1.7394	3.2419	5.3969	9.7372	13
NMSE ($\times 10^{-4}$) for trace 3						
L	2	4	8	16	32	64
AR	0.9897	1.6612	2.5317	3.9662	6.5691	9.5515
wavelet	0.7318	1.4207	1.9759	3.9363	6.3219	8.9550
NMSE ($\times 10^{-4}$) for trace 4						
L	2	4	8	16	32	64
AR	1.3761	2.2657	3.5118	5.9372	8.5892	14
wavelet	1.0153	1.6222	3.1375	5.4979	8.1081	8.4497
NMSE ($\times 10^{-4}$) for trace 5						
L	2	4	8	16	32	64
AR	0.9571	1.5136	2.2123	3.8785	7.4831	11
wavelet	0.8308	1.4156	2.2917	3.3489	5.9077	8.4882

is then added to the prediction result obtained from the AR model. We show that the proposed method has good prediction accuracy and outperforms the time domain AR based model for small, medium as well as large scale fading prediction.

To further analyze the prediction errors, we note that abrupt changes in signal strength are the main reasons of prediction errors. Multipath fading, shadow fading and path loss all lead to abrupt changes in signal strength in different scales. One advantage of the wavelet based prediction method is that the prediction error can be easily interpreted at different time resolution. Figure 2 shows the final prediction error introduced by the prediction error at each scale. The distance between a_J and d_J at a given prediction range is the error introduced by the prediction error of approximation coefficients a_J. The distance between d_j and d_{j-1} for $1 \leq j \leq J$ at some prediction range is the error introduced by the prediction of detail coefficients d_j. Figure 2 illustrates that the error prediction of approximation coefficients a_J contributes most compared the the small prediction errors of detail coefficients d_j. Sharp increase of error introduced by a_J occurs when $J > 3$. The prediction error of a_J for $J > 3$ can be explained by the abrupt changes in the trend of signal strength which is caused by encountering a showing object or a sudden change in the path loss due to unpredictable user movement.

B. Computational Complexity

One advantage of our approach is its lower computational complexity. The traditional AR model is operated on the finest scale while our approach can operate on low-pass filtered and sub-sampled data. Given a data of length M, at scale j the data points to be processed are of length $M/2^j$, implying a 2^j to 1 complexity reduction. What's more, when $d(j,k) = 0$ for $1 \leq k \leq M/2^j$, estimation of AR parameters becomes trivial. Our experimental results show that as the sliding window moves, the DWT decomposition of segment signal strength measurement y results in $d(j,k) = 0$ for many k, $1 \leq k \leq n_j$, at scales $j = 1, 2, 3$. In our experiments, 40.04%, 15.65%,

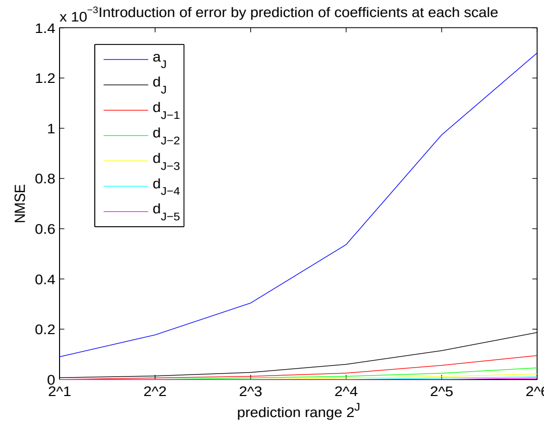

Fig. 2. Error introduced by each scale prediction

7.03% of $d(1,\cdot), d(2,\cdot)$ and $d(3,\cdot)$ are zero respectively. Therefore there are much fewer AR parameters to be estimated and this makes the prediction algorithm more computationally efficient. The time domain AR model based prediction method, on the contrary, does not have this advantage.

VI. CONCLUSIONS

We provide an accurate, low-complexity, on-line prediction mechanism for long range prediction of non-stationary signal strength traces, using wavelet based signal analysis tools. Our method can predict at long ranges (several tens of wavelength ahead) and outperforms time domain AR based channel prediction models in terms of both the accuracy and the computational complexity. Our predictor can be used in adaptive transmission applications such as dynamic route selection in multi-hop networks, where predictions are needed many hundreds of milliseconds in advance.

REFERENCES

[1] S. Semmelrodt and R. Kattenbach, "Investigation of Different Fading Forecast Schemes for Flat Fading Radio Channels," *Proceedings of IEEE VTC,* pp. 149-153, Orlando, FL, Oct. 2003.

[2] A. Duel-Hallen, "Long Range Prediction of Fading Signal," *IEEE Signal Processing Magazines,* pp. 62-75, vol. 17, no. 3, May 2000.

[3] I. Sen, D.W. Matolak, W. Xiong, " Wireless Channels that Exhibit "Worse than Rayleigh" Fading: Analytical and Measurement Results," *Proceedings of IEEE MILCOM,* pp. 1-7, Washington, DC, October 2006.

[4] T. Eyceoz, S. Hu and A. Duel-Hallen, "Performance Analysis of Long range Prediction for Fast fading Channels," *Proceedings of the 33rd CISS,* Baltimore, MD, March 1999.

[5] A. Heidari, D. McAvoy, A.K. Khandani, "Adaptive Long-Range Prediction of Mobile Fading," *Proceedings of IEEE Communications 23rd Biennial Symposium,* pp. 3984- 3989, vol. 4, December 2003.

[6] J. Klaue and A. Aguiar, "Robust Real-Time Channel Prediction Based on Inaccurate Instantaneous Measurements: an Approach," *Proceedings of IEEE Workshop on Signal Processing Advances for Wireless Communications,* New York, NY, June 2005.

[7] L.Piroddi and W. Spinelli, "Long-Range Nonlinear Prediction: A Case Study," *Proceedings of IEEE Decision and Control 42nd Conference,* pp. 219-222, June 2006.

[8] P. Abry and D. Veitch, "Wavelet Analysis of Long Range Dependent Traffic," *Transactions of IEEE Information Theory,* pp. 2-15, vol.44, no.1, January 1998.

[9] M.K. Tsatsanis and G.B. Giannakis, "Time-Varying System Identification And Model Validation Using Wavelets," *Transactions of IEEE Signal Processing,* pp. 3512-3523, vol.41, no.12, December 1993.

[10] M.H. Hayes, "Statistical Digital Signal Processing and Modeling," New York Wiley, 1996.