Link Quality and Signal-to-Noise Ratio in 802.11 WLAN with Fading: A Time-Series Analysis

Jian Zhang Ivan Marsic

Department of ECE, Rutgers University

{jianz, marsic}@caip.rutgers.edu

***Abstract*—It is known that for multipath fading channels individual points or average signal-to-noise ratio (SNR) alone do not adequately describe the wireless channel quality. This paper uses time-series modeling to investigate the relationship between SNR and 802.11 link bandwidth which we use to define the link quality. Two models, one linear and another nonlinear, are constructed and fitted to time-series of SNR as input and link bandwidth as output. Their performance is measured in terms of the accuracy of prediction of the link bandwidth. By the linear Auto-Regressive Moving Average eXogenous variables (ARMAX) model, we show the existence of nonlinearity in the input data series, which results in high prediction error. The prediction performance is significantly improved by the nonlinear Echo State Network (ESN) model due to its ability of nonlinear re-expression of the SNR time-series and associating them with the correct link bandwidth.**

Keywords: 802.11 WLAN, link quality, time-series models

I. Introduction

Compared to wired link, wireless communication suffers from high transmission errors and its performance changes frequently and dramatically, highly depending on the channel quality, which usually exhibits great variability. The sources of variation include user mobility, environmental changes and interference. The rapidly varying channel condition results in changing packet error rate (PER), therefore making network bandwidth a highly dynamic resource. Due to the dynamic variability, the wireless link quality needs to be measured and provided to wireless applications and protocols, in order for wireless networks to most effectively utilize and manage resources.

Traditionally, signal-to-noise ratio (SNR) or carrier-to-interference ratio (CIR) is measured and reported as an indicator of wireless link quality. Most of the 802.11 WLAN cards measure and display the signal strength (usually as signal bars). There is a prior literature [8][3] on adapting network parameters, e.g., link transmission rate, based on SNR, which assumes high correlation between SNR and link quality. The theoretical relationship between the SNR and link quality, represented by bit error rate (BER), can be derived for an additive white Gaussian noise (AWGN) channel. The SNR and BER relationships for different 802.11b modulation techniques are shown in Fig. 1(a). The link bandwidth is defined in this paper as the maximal throughput that can be achieved over the link layer. It can be determined from BER, as in Fig. 1(b), for packets of a fixed length (adjusted for the overhead of PHY and MAC layers). We use the link bandwidth to represent the link quality. In one of our datasets collected for DBPSK, shown in Fig. 1(c), the link bandwidth values are distributed around the theoretical DBPSK curve. Our previous work [10] on estimating the 802.11 link bandwidth from measured SNR is also based on above assumption.

However, the above assumption of a non-fading AWGN channel can be invalid in reality, especially in typical office and home environments with fading links. The data collected in MIT Roofnet [1] show inconsistency of the relationship between PER and SNR or transmitter-receiver distance. They conclude that the observed large number of links with moderate error rates is probably due to multipath fading rather than signal attenuation [1]. In [2], the authors claim that, although an average error rate curve for their multipath scenario has been obtained by averaging over a large number of randomly generated channels, most of their observed PER is significantly lower than this average curve. Similar situation is observed in our experimental results, especially those collected in office environments with high transmission rate (CCK 11) which tends to be more sensitive to multipath. In Fig. 2, the measured link bandwidth values are distributed in the gray zone ranging from about 0 up to 5Mbps, although SNR's are all above 40dB, which by the non-fading curve should yield high link bandwidths.

Therefore, individual SNR points alone do not adequately describe the wireless channel quality for 802.11 links with fading channels, i.e., the two-dimensional scatter plot curve is not sufficient to model their relationship. Realizing the limitations of a simple correlation, [2] suggests a substitute indicator for the prediction of PER, the computation of which, however, requires ideal channel estimation in PHY. In this paper, we propose a time-series modeling method, assuming that

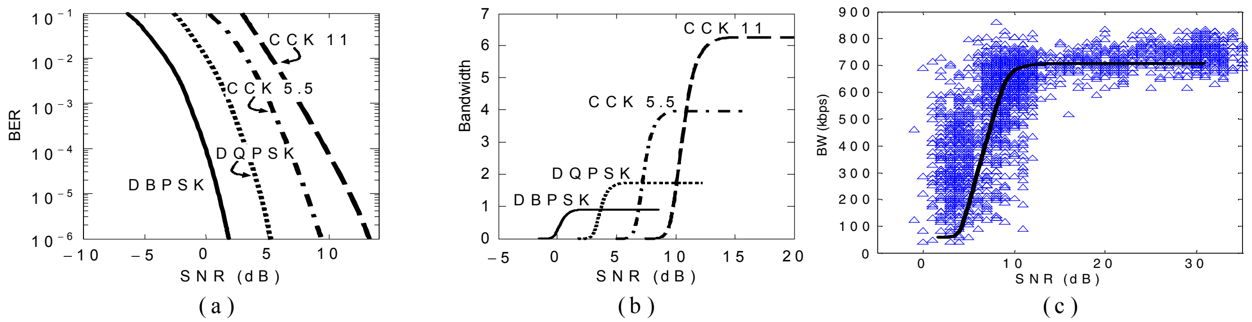

Fig. 1: SNR vs. bandwidth relationship for AWGN channel. (a) *Q*-functions for different modulation schemes. (b) Corresponding SNR-BW theoretical relationships. (c) Measured SNR-BW for DBPSK.

SNR time series may provide more information on link quality than single points of SNR. Instead of a single-point SNR, the current and historical SNR's and corresponding bandwidths are treated as two time-series signals and their relationship is modeled as a time series transfer function. The model is identified and fitted to the SNR signal and corresponding link bandwidth in a training dataset. This model is tested on other data, collected in the same environment, to determine if there exists regularity in the relationship between SNR and link bandwidth. Then, we can use the model to predict the link bandwidth in similar scenarios.

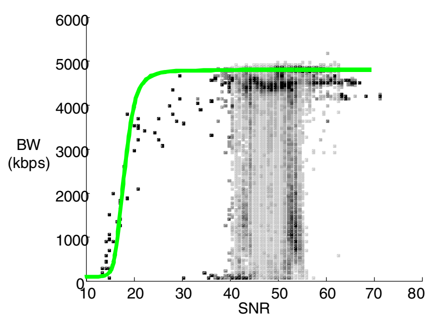

Fig. 2: Distribution of a dataset on BW-SNR space.

In the rest of the paper, we describe two models, a linear and a nonlinear one. In section III, we show that the nonlinearity of the signals makes the linear model difficult to fit to the training dataset, while the nonlinear model in section IV is able to extract the features and regularities of the input sequences of SNR and associate them with correct output link bandwidth.

II. Data collection

To build and test the model, we need to collect sufficient amount of data of SNR and link bandwidth. We ignore the issues of multi-user medium sharing on the bandwidth measurement since here we only deal with the channel quality. Therefore, a single transmitter is connected to a single receiver in 802.11 ad-hoc mode and both are located in different offices. The modulation rate is fixed to CCK 11. A UDP flow is started with the rate high enough to saturate the link.

The sampling period of the series is 1 second. The receiver records an SNR point for each received frame. To obtain the SNR over 1-second intervals, we average the SNR's of 20 randomly selected frames that were received during that second. If less than 20 frames were received in a second, all of them are used. For each 1-second interval, the link bandwidth is also calculated as the product of the frame size and the number of the received frames.

III. Linear ARMAX modeling

In a single-input, single-output linear causal Auto-Regressive Moving Average eXogenous variables (ARMAX) model [7], the relationship between output series $\boldsymbol{y_t}$ and the input series $\boldsymbol{x_t}$ can be described through a linear filter as:

$$y_t = \nu_0 x_t + \nu_1 x_{t-1} + \nu_2 x_{t-2} + \cdots + n_t \tag{1}$$

where $\boldsymbol{n_t}$ is a noise series of the system that is independent of the input series $\boldsymbol{x_t}$. Eq. (1) can be rewritten as:

$$y_t = \nu(B) \cdot x_t + n_t \tag{2}$$

where $\nu(B) = \sum_{j=0}^{\infty} \nu_j B^j$ is called the transfer function of the filter, and $\boldsymbol{B}$ is the backward shift operator. The goal of the modeling is to identify and estimate the transfer function $\boldsymbol{\nu(B)}$ and the noise $\boldsymbol{n_t}$ based on the available information of the input series $\boldsymbol{x_t}$ and the output series $\boldsymbol{y_t}$. In our case, these correspond to the SNR and the link bandwidth, respectively. As seen in Eq. (1), the current bandwidth output of the ARMAX model is not only a function of the current SNR point as in Fig. 1(b), but also depends on the current and past SNR points.

To build an ARMAX model, the input series x_t is assumed to follow some ARMA(p, q) model, i.e., x_t satisfies

$$\phi_x(B)x_t = \theta_x(B)\alpha_t \tag{3}$$

where α_t is a white noise series, and $\phi(\cdot)$ and $\theta(\cdot)$ are the pth and qth degree polynomials as

$$\phi_x(B) = 1 - \phi_1 B - \cdots - \phi_p B^p \tag{4}$$

$$\theta_x(B) = 1 - \theta_1 B - \cdots - \theta_q B^q \tag{5}$$

The original SNR signal has a non-constant mean and time varying variance. It has to be processed and transformed to a stationary series before modeling and fitting. To stabilize the variance, we have tried three different kinds of Box-Cox's power transformations, logarithmic, square root and reciprocal transformations [7]. The square root transformation is chosen for the best performance. Then, we difference the transformed signal to make the mean nearly constant. The original SNR series and the one after preprocessing are shown in Fig. 3.

We compute the sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF) of the preprocessed series to identify the orders of p and q. By comparing the Akaike information criterion (AIC) of the models with different combinations of p and q, we find that the best fitted model is an ARMA(2,6) model, like so

$$\begin{aligned}(1-0.131B-0.541B^2)x_t &= (1-0.603B-0.566B^2 \\ &+ 0.189B^3 + 0.034B^4 + 0.053B^5 - 0.050B^6)\alpha_t\end{aligned} \tag{6}$$

The variance of the noise α_t is 0.0126.

Brock-Dechert-Scheinkman (BDS) test [2] is performed on the residuals of the model for different embedding dimension m and 'near' distance ε. The large values of BDS test in Table I show that the model rejects the i.i.d. hypothesis for the residual series, which suggests some hidden structures in the series, such as nonlinearity and/or nonstationarity. Although there may exist nonlinear transformations other than the power transformations we have tried, which can better represent such nonstationarity, it is difficult to find them since they can be found only by exhaustive trial-and-error search.

Table I. BDS test results

	$m = 2$	$m = 3$	$m = 4$
$\varepsilon = 0.5\sigma$	35.7942	49.1564	63.8780
$\varepsilon = \sigma$	32.6483	40.5832	46.2860
$\varepsilon = 1.5\sigma$	28.1427	32.9581	35.5951
$\varepsilon = 2\sigma$	24.9336	28.0945	29.5435

Due to the nonlinearity in the signals, the linear ARMAX model shows high error on the prediction output. In addition, the input SNR series and the output link-bandwidth series have to be transformed to stationary series for the ARMAX modeling. The output has to be postprocessed to reverse the preprocessing operations, which demands some unavailable prior knowledge in prediction process, such as the initial value of the output series for reversing the differencing.

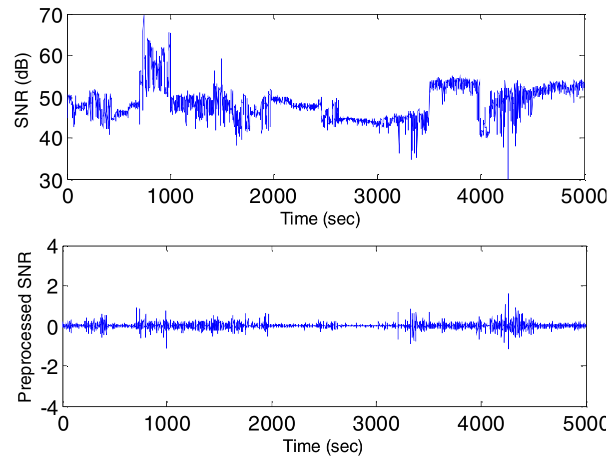

Fig. 3: (a) The original SNR series. (b) The transformed and differenced series.

Considering the above inadequacies of the ARMAX modeling, we turn to nonlinear time-series modeling. The Recurrent Neural Network (RNN) technique, more specifically Echo State Network (ESN) [4][5], is chosen due to its ability for nonlinear modeling and prediction. The ESN combines various linear and nonlinear operations and automatically tunes the corresponding parameters based on the training input/output series. The preprocessing is simple, as is the postprocessing.

IV. Nonlinear ESN modeling

A. ESN model

ESN is an efficient black-box modeling method for nonlinear predication [5]. The idea is to construct a model with a series of linear operations (weighting and summation), nonlinear operations, and delay operations such that it mimics a given empirical dataset.

An ESN is a network of artificial neurons. A neuron is a basic computational unit that computes some function, usually nonlinear, of the weighted sum of inputs from other units or an external source. Its output, in turn, can be served as input to other units. ESN has both feed-forward and feedback connections. For example, in Fig. 5, the output or activation of unit 1 is updated according to

$$x_1(t) = f\left(w_1 \cdot x_1(t-1) + w_2 \cdot x_2(t) + w_3 \cdot x_3(t)\right) \tag{7}$$

where w_1, w_2, w_3 are weights assigned to the three inputs of unit 1. Its inputs, $x_1(t-1)$, $x_2(t)$, and $x_3(t)$, are delayed output feedback from itself, current output of unit 2 and current output of unit 3, respectively. The output of internal units is called state.

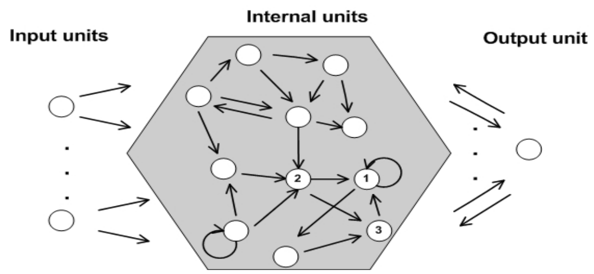

Fig. 5: Structure of an echo state network.

In Fig. 5, the input signals are introduced by the input layer to the internal layer. The internal units update their states at each time step as in Eq. (7). The output of the ESN is then decided by the states as follows

$$y(t) = f^{out}\left(\sum_{i=1}^{N} w_i^{out} \cdot x_i(t)\right) \quad (8)$$

Eqs. (7) and (8) decide the relationship between the input SNR and the output bandwidth in our ESN model. They can be viewed as re-expression and association, respectively, similar to power transformation and fitting in ARMAX. By Eq. (7), the input signal is transformed or re-expressed, by the internal neurons, as the states which expose principal patterns hidden in the input series. This mechanism provides richer nonlinear expression than the power transformation. By the ESN training algorithm, the output weights in Eq. (8) are updated automatically so that the revealed patterns are associated with the desired output. This is done by adjusting the output weights $\boldsymbol{w}_i^{out}$ so that the error $\boldsymbol{e}(\boldsymbol{n})$ in Eq. (9) is minimized, in the mean square error sense, i.e., the difference between the output of the model and the desired output is minimized

$$e(t) = (f^{out})^{-1}(y_{desired}(t)) - (f^{out})^{-1}\left(\sum_{i=1}^{N} w_i^{out} \cdot x_i(t)\right) \quad (9)$$

where $(\boldsymbol{f}^{out})^{-1}(\cdot)$ is inverse function of $\boldsymbol{f}^{out}(\cdot)$, and $\boldsymbol{y}_{desired}$ is the desired output. After training, ESN can start doing prediction when supplied by a real-time SNR input signal.

B. Testing and estimation results

For the purpose of bandwidth prediction, we created an ESN with 400 internal units, single output (bandwidth), and single input (SNR). We do not need multiple inputs for previous SNR values since ESN, by feedback, stores and includes the historical input into the computation of current or future states and the output. Both the input and output need to be normalized to the range of [−1, 1] by shifting and scaling, which is simple and easily reversed, compared to the pre- and post-processing in the ARMAX modeling. We select the same 5000-point/second training dataset as used in the ARMAX modeling. For the purpose of effective re-expression, a sparse interconnectivity of the internal units is used for the ESN so that the internal states display a rich set of systematic variants of the input signal. It is achieved by making the spectral radius—the largest absolute value of eigenvalues of the internal-weights matrix—smaller than unity. This ‘‘reservoir of dynamics’’ feature of ESN is critical for the accuracy of nonlinear prediction. The ESN model is trained and the tuning process is repeated until the mean square error on the training data reaches desired low level. After that, we check our model by predicting the link bandwidth from the measured SNR in a different dataset with a total of 20,000 points. (Recall that both the training and estimation datasets are collected in a similar environment). The training process takes less than five minutes on a 2 GHz Pentium PC, and the prediction runs in real time.

The results are shown in Fig. 4. The output of ESN is the estimated link bandwidth, which is the dashed brighter-colored curve in the figure. The actual link bandwidth is the solid curve. The bottom row in Fig. 4(a) shows smoothed estimated bandwidth which may be more suitable for use by adaptive applications. The figure shows good agreement between the estimated and the actual values. The relative error is

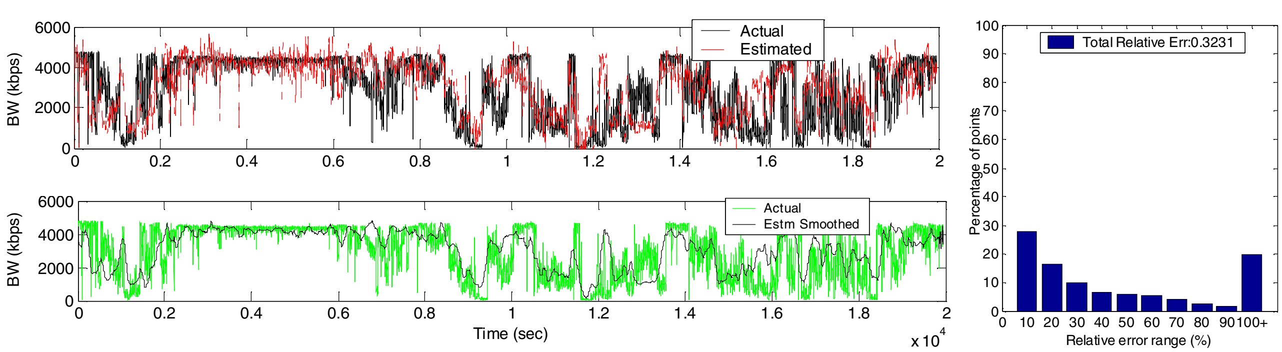

Fig. 4: (a) Comparison of actual and estimated bandwidth and estimation error by ESN model. The bottom row shows the estimated bandwidth smoothed by averaging over a window of 100 points; (b) Relative estimation error.

$$Err\ \% = |y_i - \hat{y}_i| / y_i \quad (10)$$

where y_i and $\hat{y}_i$ are the actual and estimated outputs at ith point, respectively. Fig. 4(b) shows the distribution of the estimates in different relative error ranges. It shows that around 30% of points are estimated with relative error lower than 10%. For the same dataset, the ESN model outperforms both the two-dimensional non-fading model, Fig. 1(b), and the ARMAX model. The non-fading model fails to predict the degradation of link bandwidth due to the fading since it always overestimates the bandwidth, given the high SNR input. Compared to ARMAX, the ESN model provides variety of options for re-expression of SNR series, which contributes to the performance. Despite the overall estimate accuracy improvement, on the other hand, we also notice there are around 20% of points that are estimated with high error (over 100%). The high relative error happens mainly at the points of extremely low actual bandwidth, which is an artifact of how we calculate the relative error: if the actual bandwidth y_i in the denominator of Eq. (10) is very small it yields a high error value. Another reason for those high error points could be the incompleteness of the training dataset. That is, some input/output sequences are not present in the training dataset. When they appear in the estimating dataset, their corresponding bandwidth cannot be estimated correctly. However, the bigger the training dataset is, the longer and costlier the computation. Finding a complete, yet non-redundant training dataset is still an open problem.

The sensitivity of the ESN model to the input variations can be adjusted by configuring the spectral radius α of the internal weight matrix W. The balance and tradeoff between agility and stability can be chosen for applications with different preferences and requirements. α is intimately linked to the intrinsic timescale of the internal-state dynamics. For short-memory implementation with a small α, the agility is enhanced by sacrificing the stability; and vice versa, with big α for long-memory. The parameters of the ESN are initialized as listed in Table II.

Table II. Parameters of the ECN model

Parameter name	Value
Spectral radius	0.7
Size of internal layer	400 units
Percentage of non-zero elements in internal weight matrix	2%
Internal weights	0.1 or –0.1
Input weights	3 or –3
Backward weights	0.1 or –0.1

C. Tuning for different environments

Since the model is built on empirical data and the data is expected to depend on the environment and the type of 802.11 devices, we may need to re-train the model for different devices or fading environments (such as different buildings, indoor/outdoor). This should not be a great problem for the ESN model, since the training process is performed offline and can be finished in a short time.

V. Conclusions

In this paper, we study the relationship of SNR and bandwidth in fading environments by time-series modeling. Our experimental data confirm that individual SNR points do not adequately describe the wireless link quality with multipath. Instead, by the time-series modeling, the patterns of SNR sequences can be recognized and associated with the corresponding link quality. The process of building a linear ARMAX model indicates nonlinearity in the SNR series. The nonlinear ESN model, with its ability of providing rich re-expressions and associations of signals, achieves accurate prediction on link bandwidth. This shows that even under multipath fading, the relationship of the SNR and link bandwidth can be captured by a combination of linear, nonlinear, and delay operations.

References

[1] D. Aguayo , J. Bicket , S. Biswas , G. Judd , R. Morris, "Link-level measurements from an 802.11b mesh network", *Proc. 2004 ACM SIGCOMM Conference*, Portland, OR, 2004.

[2] W. A. Brock, W. D. Dechert, J. A. Sheinkman, "A test of independence based on the correlation dimension," SSRI no. 8702, Department of Economics, University of Wisconsin, Madison, 1987.

[3] G. Holland, N. Vaidya, P. Bahl, "A rate-adaptive MAC protocol for multi-hop wireless networks", *Proc. ACM MobiCom*, Rome, Italy, 2001.

[4] H. Jaeger, "The 'echo state' approach to analysing and training recurrent neural networks," GMD Report 148, German National Research Center for Information Technology, 2001.

[5] H. Jaeger, H. Haas, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communications," *Science*, 78-80, April 2, 2004.

[6] M. Lampe, H. Rohling, W. Zirwas, "Misunderstandings about link adaptation for frequency selective fading channels", *Proc. PIMRC Conf.*, 2002.

[7] LAN/MAN Standards Committee of the IEEE Computer Society, ANSI/IEEE Std. 802.11—1999 edition, "Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications," 1999.

[8] D. Qiao, S. Choi, "Goodput enhancement of IEEE 802.11a wireless LAN via link adaptation", *Proc. IEEE ICC*, Helsinki, Finnland, 2001.

[9] W. S. Wei, *Time Series Analysis: Univariate and Multivariate Methods*, Addison Wesley Publ. Co., Reading, MA, 1990.

[10] J. Zhang, L. Cheng, I. Marsic, "Models for non-intrusive estimation of wireless link bandwidth," *Proc. Personal Wireless Communication Conf. (PWC)*, 334-348, Venice, Italy, 2003.