Lightweight Models for Prediction of Wireless Link Dynamics in Wireless/Mobile Local Area Networks

Liang Cheng and Ivan Marsic
Department of Electrical and Computer Engineering,
Rutgers — The State University of New Jersey
94 Brett Rd., Piscataway, NJ 08854-8058
{chengl,marsic}@caip.rutgers.edu

Abstract — **Linear time-series analysis of session throughput in wireless/mobile local area networks shows that the dynamics of wireless link bandwidth in such networks can be predicted to a useful degree from past values by using Autoregressive and Windowed-Mean models. These lightweight models are suitable for adaptive-application deployment on mobile devices.**

I. INTRODUCTION

Wireless/mobile local area networks (LANs) are increasingly involved in our daily life with the proliferation of mobile communication devices and wireless connectivity. Because of the convenient communications and increasing bandwidth supported by wireless/mobile LANs, more and more multimedia applications are targeting mobile data communications. It is generally observed that the throughput of a multimedia application streaming data in wireless/mobile LANs fluctuates from time to time. This is because that the throughput reflects dynamics of wireless links, i.e. changing available link bandwidth and link latency. In this paper, therefore, we focus on the throughput of a data streaming session as the performance index of wireless link dynamics in wireless/mobile LANs. It is useful for a multimedia session to predict the dynamics of wireless links so that its performance can be improved by adapting to the dynamics [1].

Modeling for prediction is a widely studied subject, as the related work discussed in Section 6. However, most models are not computationally lightweight, which is critical to the application deployment on mobile devices in wireless/mobile LANs. Moreover, to our best knowledge, there is few existing research on the feasibility of using lightweight models for prediction of wireless link dynamics in wireless/mobile LANs. We choose time series analysis here since it is well studied for data-based "black-box" prediction and it includes several lightweight models.

In the next section, the experiment scenarios for measuring the session throughput for the study of wireless link dynamics are presented. Section 3 describes the data collection and discusses the relationship between the session throughput and wireless link dynamics. In Section 4, we analyze the collected data, study the modeling techniques, and choose appropriate models for prediction of the wireless link dynamics in wireless/mobile LANs. Section 5 evaluates selected models based on their prediction performance. Section 6 discusses related works and Section 7 presents conclusions.

II. EXPERIEMENT SCENARIOS

Experiments have been designed to measure session throughput for the study of wireless link dynamics in wireless/mobile LANs. Considering the proliferation of audio and video applications [2], we study the throughput of data streaming sessions across wireless/mobile LANs.

A. Experimental Networks

Figure 1 illustrates the abstract layout of our experimental wireless/mobile LANs. The LANs are campus networks and are used by multiple users simultaneously. They include Proxim RangeLAN [3] and Sony WirelessLAN [4]. Proxim RangeLAN operates at 2.4 to 2.483 GHz using spread spectrum frequency hopping with media access protocol OpenAir and delivers data traffic up to 1.6 Mbps. Sony WirelessLAN is an IEEE 802.11b [5] wireless network operating at 2.4 GHz radio frequency band using direct sequence spread spectrum (DSSS) and delivers up to 11 Mbps which is comparable to wired Ethernet.

B. Data Streaming Source and Sink

The data stream source or sender, as shown in Figure 1, is a Sun Sparc10 workstation with Solaris 2.6 platform. The data stream sinks or receivers include a mobile laptop and a wired personal computer, both with Windows NT platform installed. The sender sends datagram traffic in a multicast fashion so that both receivers can receive the same data traffic. A stable traffic generator, which overcomes the randomness caused by the scheduling mechanisms of operating systems and coarse granularity of timers, is used at the source to generate stable and smooth data streams at fixed rates. Figure 2(a) shows the wired-receiver-side traces of different data traffic rates the traffic generator generates. The receivers receive the data stream and record the number of bits received every second as the session throughput.

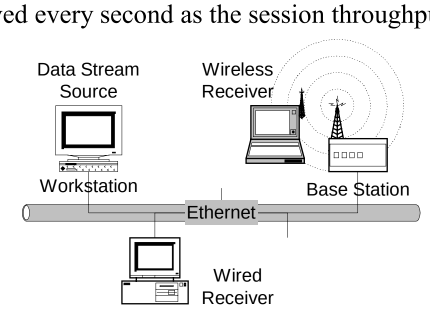


Fig. 1. The abstract layout of experimental networks

C. Mobility Scenarios of the Mobile Receivers

The mobile laptop travels with varying velocity around the base station (or access point) along the hallway and enters the offices and the labs. (If the velocity is zero, the mobile laptop is motionless.) The maximum distance between the laptop and the base station reaches 100 feet. Generally there is no line-of-sight communication between the base station and the mobile laptop.

III. Data Collection

Based on the above-described experimental scenarios, we categorize the experiments into two groups, namely the RangeLAN group and the 802.11b group. For each group, several data sets of the throughputs recorded at the mobile receivers with various streaming rates have been collected. Each data set contains 10,000 values of the throughput, with one value recorded every second. We recorded data sets of throughputs of different sessions with various session sending rates. The time of the experiments for the collection of the data sets is randomly distributed over several weeks. In order to keep the illustrations clear, not all data sets and sample values are plotted in the graphs shown in this paper. We have observed that a fraction of all the data sets and sample values in each data set well represents the whole data sets in terms of statistical characteristics. Figures 2(b) shows typical representatives of the general results.

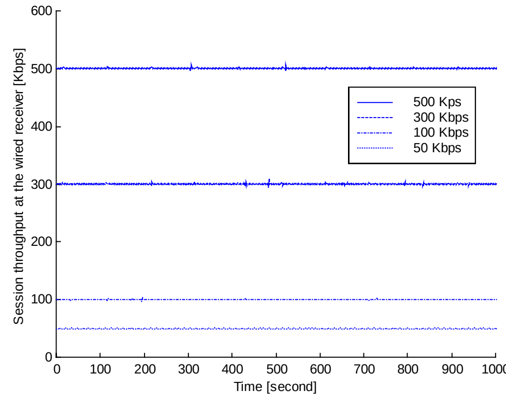
A. Session Throughput vs. Wireless Link Dynamics

Comparing Figure 2(a) to Figures 2(b), it shows that the data traffic received at the wired host is very different from that received at the mobile hosts. The session throughputs shown in Figure 2(b) change dynamically in a wide range. The difference is solely caused by the dynamics of wireless link since the mobile and the wired receivers are in the same multicast streaming session and the only difference is the link type of the last link to the receivers, i.e., wireless vs. wired. It can be concluded that, in our experimental wireless/mobile LANs, the correct prediction of session throughput implies the correct prediction of wireless link dynamics.

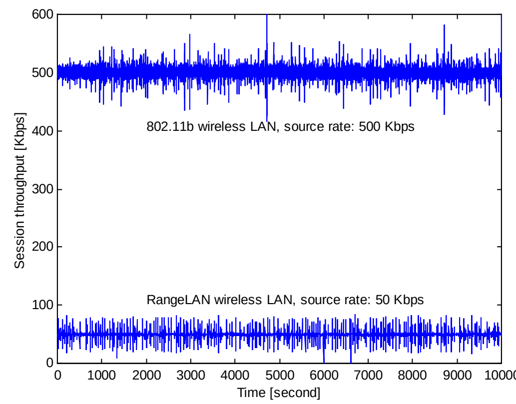
IV. Analysis and Modeling

A. Analysis

We examined autocorrelation function of each data set. The autocorrelation quantifies how well a throughput value at time t is linearly correlated with its corresponding throughput value at time $t+\Delta$, which in turn shows how well the value at time t predicts the value at time $t+\Delta$. The value of autocorrelation function ranges between -1 and 1. The closer the value to 1, the better linear correlation of the value at time t and time $t+\Delta$. The time series analysis shows that the past throughput value has a strong influence on the future throughput value. Figure 3 shows the autocorrelation function to a lag of up to 120 seconds (2 minutes) for the data set in Figure 2(b). Notice that in both scenarios the values of session throughput are strongly correlated.



(a) Data streaming traffic received at the wired receiver



(b) Data streaming traffic received at the mobile receiver

Fig. 2. Data streaming traffic received at the receivers

The above statistical properties are presented in all the data sets, irrespective of the velocity of the mobile receivers. This implies that it is feasible to use linear time series models to predict the dynamics of the wireless link, which is reflected by the session throughputs. We believe that other technologies, e.g., Bluetooth [6] and HyperLan [7], in wireless/mobile LANs, should demonstrate similar statistical properties to those found in this research.

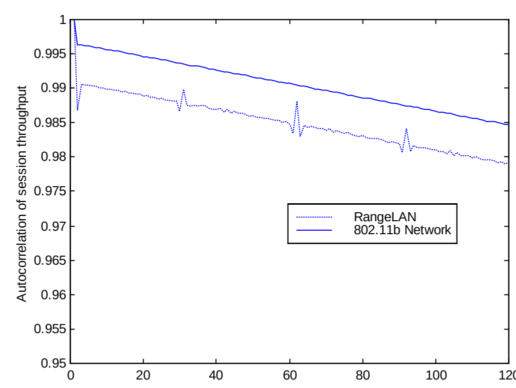


Fig. 3. Autocorrelation function of data sets in Figure 2(b)

B. Modeling

There exist a number of models for linear time series analysis, only AR (Autoregressive), MA (Moving Average), and ARMA (Autoregressive Moving Average) models [8] are studied here considering the lightweight implementation of modeling and prediction functionality in mobile devices.

1. ARMA Models

If a data set with time series value $\{X_t\}$ can be fit to an ARMA model, then it can be described as follows:

$\{X_t\}$ is an ARMA(p,q) process if $\{X_t\}$ is stationary and if for every t

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q} \quad (1)$$

where $\{Z_t\}$ is a white noise sequence WN($0,\sigma^2$). It is convenient to use a concise form as

$$\phi(B)\, X_t = \theta(B) Z_t \quad (2)$$

where $\phi(\cdot)$ and $\theta(\cdot)$ are the p^{th} and q^{th} degree polynomials as

$$\phi(x) = 1 - \phi_1 x - \ldots - \phi_p x^p \quad (3)$$

$$\theta(x) = 1 + \theta_1 x + \ldots + \theta_q x^q \quad (4)$$

and B is the backward shift operator ($B^j X_t = X_{t-j}$, $B^j Z_t = Z_{t-j}$, j=0,±1,±2…). The time series model is said to be an autoregressive model of order p or AR(p) if $\theta(\cdot)\equiv 1$ and a moving average model of order q or MA(q) if $\phi(\cdot)\equiv 1$.

2. AR Models

Note that deploying MA and ARMA models is a much more difficult proposition for a system designer since fitting time series data to them takes a non-deterministic amount of time. Instead of a linear system, fitting a MA or ARMA model present us with a quadratic system. Thus AR models are highly desirable since they can be fit to data in a deterministic amount of time. For example, in an AR model with p-order using the Yule-Walker technique [8], the autocorrelation function is computed to a maximum lag p and then a p-wide Toeplitz system of linear equations is solved. Even for relatively large values of p, this can be done almost instantaneously. The evaluation results presented below demonstrate that AR models are suitable for modeling the session throughput of wireless data networks. We also compare AR models with simple models such as MEAN and WM (Windowed Mean) [9].

3. Simple Models

The MEAN model has $X_t = \mu$, so the future values of the time sequence are predicted to be the mean. The WM model simply predicts the next sequence value to be the average of the previous w values, a simple windowed mean. Note that WM subsumes an even simpler model: LAST model as “predict the next value be the same as the last one”, i.e., w=1.

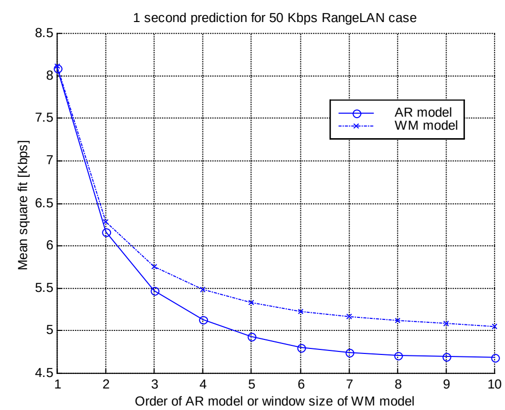
V. Evaluation

We collect additional data sets other than those collected in Section 2 and use all of them for evaluating different models. One-step-ahead and multi-step-ahead predictions are used to evaluate the correctness of different models, such as AR(p), MEAN, WM(w) models. In the case of sampling frequency equal to 1 second, m-step-prediction means predicting the throughput value at time instance that is m seconds ahead of current time instance. The comparison index is the mean square fit (MSF), which is expressed as follows.

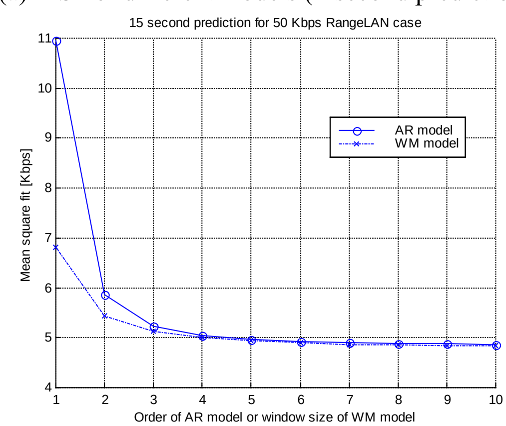
Suppose X is the vector of the observed values, and the X^{pred} is the vector of the predicted values. Then MSF is:

$$\left\| X - X^{pred} \right\| \,/\, \mathrm{sqrt}(\mathrm{length}(X))$$

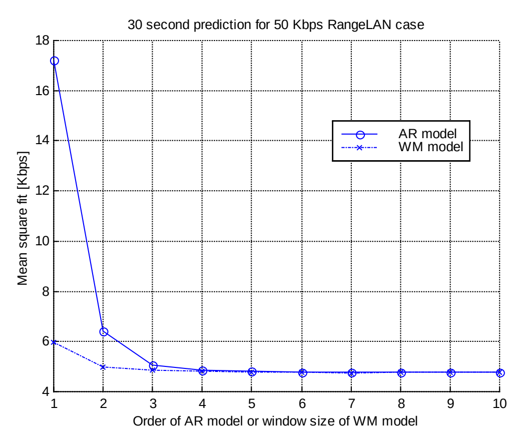
where $\|\cdot\|$ is the norm operator.



(a) MSF of different models (1-second prediction)



(b) MSF of different models (15-second prediction)



(c) MSF of different models (30-second prediction)

Fig. 4. MSF of different models in multi-step prediction

In this research, the norm of a vector is the square root of the sum of the squares of individual elements in the vector, and the vector length equals 10,000.

Figure 4(a) shows the mean square fits of AR(p) models with different orders of p and WM(w) models with different window size w, when predicting one-second-ahead session throughput in the 50 Kbps case of RangeLAN. The mean square fit of the MEAN model in this case is 4.7979. Note that unlike the WM model, the MEAN model has to use an infinite buffer to store history data as the time increases, which is infeasible in practical modeling and prediction.

Using the same evaluation methods described above, we studied the prediction performance of AR models and WM models in the other group of data sets, i.e., the cases of 802.11b networks. The results show that the prediction errors by all models decrease as the order of the models increases. Moreover, when the models' order is high enough, e.g., 10, different models show similar prediction accuracy and the nominal prediction errors are in the range of 10%.

Figures 4(b) and 4(c) show the performance of AR models and WM models in multi-step-prediction cases for the 50 Kbps session in RangeLAN. The figures show the 15 seconds prediction and the 30 seconds prediction, respectively. We observe that performances of the AR models and WM models are almost the same when their orders are high enough, e.g., 5. This result remains true in all experimental scenarios.

VI. Related Work

There is research in predicting computing resources, including host and network parameters [9][10] in wired networks. Parameters that have been monitored and predicted include: usage of computer resources (CPU, memory), throughput or available throughput of a communication path, latency of a data link, etc. There also exists research on wired networks using statistical models to study the data traffic behavior in both wide-area networks [11] and local-area networks [12]. Linear time series models are used in predicting both long-term and short-term Internet conditions and traffic behaviors. The results are used to predict network performance in support of real-time services and applications.

There is little existing research on modeling and predicting network parameters in wireless/mobile data networks. Noble et al. [13] studied the agility and stability of exponentially weighted moving average (EWMA) methods for estimation of available bandwidth in the context of mobile networking. Variations of exponential model have been used and studied in [9][13] for the estimation/prediction of network parameters as well as in TCP for round-trip time estimation [14]. It can be expressed as follows:

$$x^p_t = \alpha x_{t-1} + (1-\alpha)x^p_{t-1} \tag{5}$$

where x^p_t is the predicted value of x at time instance t, and x_{t-1} is the actual value at time instance $t-1$. This is a special form of AR model for prediction. From Eq. (5),

$$x^p_{t-1} = \alpha x_{t-2} + (1-\alpha)x^p_{t-2} \tag{6}$$

Substitute Eq. (6) to Eq. (5):

$$x^p_t = \alpha x_{t-1} + \alpha(1-\alpha)\, x_{t-2} + (1-\alpha)^2 x^p_{t-2} \tag{7}$$

Perform similar substitutions recursively, then:

$$x^p_t = \alpha x_{t-1} + \ldots + \alpha(1-\alpha)^{j-1}x_{t-j} + (1-\alpha)^{j+1}x^p_{t-j-1} \tag{8}$$

We can ignore the last item once j becomes large enough, say, N, since $\alpha<1$. Therefore:

$$x^p_t = \alpha x_{t-1} + \ldots + \alpha(1-\alpha)^{N-1}x_{t-N} \tag{9}$$

VII. Conclusions

Our research is, to our best knowledge, the first study to demonstrate that wireless link dynamics in wireless/mobile LANs is predictable to a useful degree from past behavior by using linear time series techniques, such as AR and WM models. They are computationally lightweight to be deployed in mobile devices with limited computing capabilities.

Acknowledgments

This research is supported by Cisco Systems, Inc., NSF KDI Contract No. IIS-98-72995, US Army CECOM Contract No. DAAB07-00-D-G505, and by the Rutgers Center for Advanced Information Processing (CAIP).

References

[1] J. Bolliger and T. Gross, "A framework-based approach to the development of network-aware applications," *IEEE Transactions on Software Engineering*, 24(5), pp. 376–390, May 1998.

[2] M. H. Willebeek-LeMair, K. G. Kumar, and E. C. Snible, "Bamba, audio and video streaming over the Internet," 42(2), pp. 269–280, *IBM Journal of Research and Develpment - Multimedia systems*, 1998.

[3] Proxim Inc., http://www.proxim.com/.

[4] Sony Computing Products, http://www.ita.sel.sony.com /products/pc/notebook/accessories/wlan.html.

[5] B. O'Hara and A. Petrick, *The IEEE 802.11 Handbook: A Designer's Companion*, IEEE Press, 1999

[6] The Official Bluetooth SIG Website, http://www.bluetooth.com/

[7] R. LaMaire, A. Krishna, P. Bhagwat, J. Panian, "Wireless LANs and Mobile Networking: Standards and Future Directions," *IEEE Communications Magazine*, 34(8), pp. 86–94, Aug. 1996

[8] G.E.P. Box, G.M. Jenkins, and G. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd ed., Prentice Hall, 1994

[9] R. Wolski, *Dynamically forecasting network performance using the Network Weather Service*, UCSD Technical Report TR-CS96-494, Computer Science and Engineering Department, University of California, San Diego, January 1998.

[10] P. Dinda and D. O'Hallaron, "Host Load Prediction Using Linear Models," *Cluster Computing*, 3(4), 2000.

[11] V. Paxson and S. Floyd, "Wide-area traffic: the failure of Poisson modeling," *IEEE/ACM Transactions on Networking*, 3(3), pp. 226–244, Jun. 1995.

[12] H. J. Fowler and W. E. Leland, "Local Area Network Traffic Characteristics, with Implications for Broadband Network Congestion Management," *IEEE JSAC*, 9(7), pp. 1139–1149, Sept. 1991.

[13] M. Kim and B. D. Noble, "Mobile network estimation," *Proceedings of Seventh ACM Conference on Mobile Computing and Networking (MobiCom'2001)*, pp. 298–309, Rome, Italy, Jul. 2001.

[14] L. Peterson and B. Davie, *Computer Networks: A Systems Approach*, 2nd ed., Morgan-Kaufmann Publisher, 1999.