

On Channel State Inference and Prediction Using Observable Variables in 802.11b Networks

Shirish Karande, Syed Ali Khayam, Yongju Cho, Kiran Misra,
Hayder Radha, Jaegon Kim and Jin-Woo Hong

Abstract- Performance of cross-layer protocols that recommend the relay of corrupted packets to higher layers can be improved significantly by accurately inferring/predicting the bit error rate (BER) in the packets. Here, inference refers to estimating the BER in an already received packet, while prediction specifically refers to anticipating the BER in a future packet. This paper presents a measurement based study of 802.11b WLANs that analyzes the utility of observable variables in channel state inference (CSI) and prediction (CSP). The first part of the paper investigates the utility of SSR and background traffic (BT) intensity ρ as *observable* side-information for CSI. Our results show significant utility for SSR indications and the feasibility of utilizing ρ for additional improvements. The second part of the paper utilizes the proposed CSI mechanism for BER prediction (i.e., the CSP aspect). We observe that the BER of temporally adjacent packets are correlated and thus BER of a current packet, by itself or on the basis of link-specific temporal correlation model, can provide efficient prediction. However since BER of a packet is not an *observable variable*, the above prediction mechanism cannot be realized practically without significant processing overheads. Thus we propose to estimate the channel state for the current packet using observable side-information and this estimate is subsequently used as an input to a BER based predictor. Our analysis and simulations based on an extensive set of actual 802.11 traces show that the proposed methods can provide accurate CSI/CSP under a variety of realistic channel conditions

Index Terms— Channel State Estimation, Wireless Local Area Networks, Cross-layer Protocols

I. INTRODUCTION

THIS paper presents a study based on an extensive set of actual 802.11b link-level bit error traces and associated

Manuscript received September 28, 2006.

S. Karande is with the Electrical and Computer Engineering Department, Michigan State University, East Lansing, MI 48824 USA. (phone: 517-355-3769; fax: 517-353-1980; e-mail: karandes@egr.msu.edu).

S. Khayam, K. Misra and H. Radha are with the Electrical and Computer Engineering Department, Michigan State University, East Lansing, MI 48824 USA. (e-mail: {khayamsy, misrakir, radha}@egr.msu.edu).

Y. Cho was with Electronics and Telecommunications Research Institute, Daejeon, Korea, 305-350. He is now with the Electrical and Computer Engineering Department, Michigan State University, East Lansing, MI 48824 USA. (e-mail: yongjucho@etri.re.kr).

J. Kim and J. Hong are with Electronics and Telecommunications Research Institute, Daejeon, Korea, 305-350. (e-mail: {jgkim,jwhong}@etri.re).

actual measurements of Signal to Silence Ratio (SSR) and Background Traffic (BT). The motivating applications for this work are the recently proposed cross-layer protocols for delivery of multimedia, in particular video [1]-[9]. It is crucial to elaborate upon this motivation: Unlike the traditional wired Internet based communication, the number of packet drops due to bit errors can be substantial in wireless networks. Bandwidth hungry multimedia systems can be adversely affected by such packet drops. Therefore many recent multimedia related studies [1]-[9] have recommended the development of cross-layer protocols that do not discard partially damaged packets. Relay of corrupted packets to higher layers can lead to significant improvements in the video throughput. However, the efficacy of information recovery from a corrupted packet is a function of the bit error rate (BER) in the packet. Hence, accurate channel state inference (CSI) and prediction (CSP) can provide substantial capacity gains.

Utility of channel awareness in improving capacity is a well researched area and has been demonstrated by many theoretical- and simulation-based studies, most of which focus on the physical layer. However unlike the physical layer, the channel observed at the link/MAC layer is discrete, and each individual bit does not have a signal strength associated with it. This makes the task of practically providing channel state information (CHSI) about the corruption levels in a packet significantly harder. Therefore, the authors in a prior work started identifying practical mechanisms that can provide CHSI at the link-layer [4], [14]. The work in this paper is a continuation of [4], [14]. We defer detailed discussion on how the present work relates with our previous studies to the related work section.

Our experimental methodology allows us to associate three link quality measures with each individual packet: ρ represents the BT overheard between two received packets in terms of packets/sec, SSR measures the Signal to Silence Ratio (SSR) for the first few microseconds (μs) and we assume that a link-level checksum can provide us with a binary indication (Z) of whether a received packet is corruption free or not; which is a viable assumption in cross-layer protocols based on an underlying 802.11 MAC that supports CRC checksum. We refer to ρ , SSR and Z as

freely observable variables, because we do not need to transmit any additional pilot/parity bits (beyond what is available in a standard 802.11b implementation) to observe these variables.

The first contribution of this paper is to develop a *link-invariant* correlation model that relates the freely observable variables to the BER in an individual packet. We subsequently exhibit the utility of such correlation models in accurately estimating the BER in a corrupted packet by only observing ρ , SSR and Z associated with that packet. Based on extensive experiments using actual 802.11 traces, we show that utilizing ρ and SSR (on a packet-by-packet basis) can provide significantly improved performance over merely utilizing some average statistic along with the binary indicator Z. In particular we observe that SSR indications by themselves, despite the presence of varying intensity of BT, can provide robust CSI.

The second part of our contribution utilizes the above developed channel state inference mechanism for prediction. We measure the efficacy of a simplistic approach which assumes that the channel state of temporally adjacent packets does not vary significantly. Such a simplistic approach was surprisingly efficient on a number of links. However, often significant benefits can be achieved by developing link-specific correlation models that utilize the BER of a current packet to predict the BER in a subsequent packet. A fundamental hurdle in practically utilizing such a temporal correlation model emerges from the fact that in real-life communications the BER in a corrupted packet is not freely observable. Exact value of a BER can be determined only by transmitting additional parity bits and conducting a complete error recovery. Such processing over-heads are often infeasible, especially in the context of delay-sensitive real-time applications. We show that the channel state estimation mechanism developed in the first part of this paper can be efficiently combined with a link-specific temporal correlation model for BER to practically provide improved channel state prediction.

The structure of the paper is as follows: Section II describes our experimental methodology. In Section III we evaluate the feasibility and utility of correlation models for CSI. In Section IV we utilize this CSI mechanism for CSP. In Section V we provide a brief review of the related work in this field. We summarize the key conclusions of this work and identify important tasks for future work in Section VI.

II. TRACE COLLECTION METHODOLOGY

For this study, we consider two wireless setups, as shown in Figure 1, to collect the error traces. In Setup A five wireless receivers were used to simultaneously collect error traces on an 802.11b WLAN. One receiver was placed within clear line-of-sight (LoS) of the access point (AP), while the remaining four receivers were placed at different locations in a room

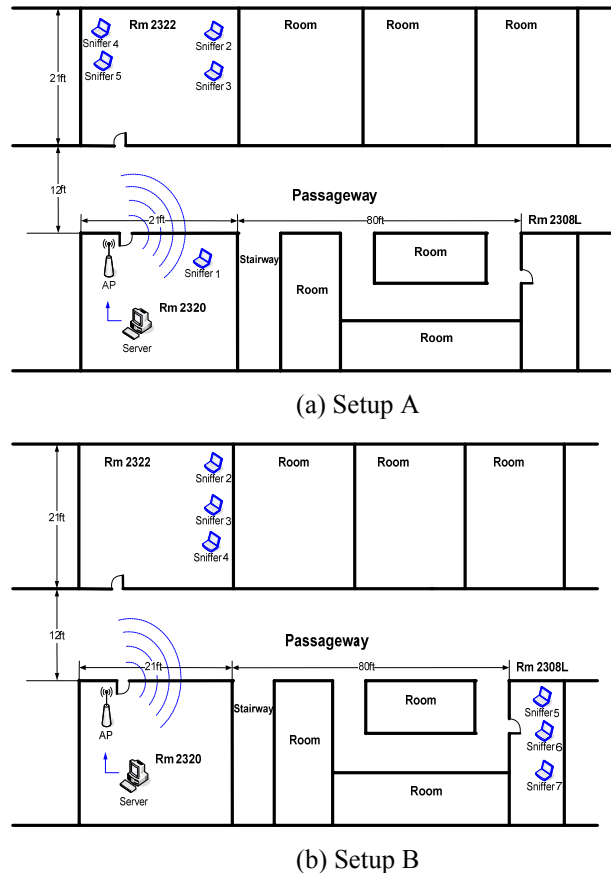


Figure 1 Topologies used for wireless trace collection

across the hallway. In setup B six receivers were used to simultaneously collect error traces. Three receivers were placed in a room across the hallway, while three receivers were placed (at an extreme edge of the network) in a room 100 feet down the hallway. A wired sender was used to send multicast packets with a predetermined payload on the wireless LAN; multicasting disabled MAC layer retransmissions. Each experiment comprised of one million packets with a payload of 1,000 bytes each. At the physical layer, the auto rate selection feature of the AP was disabled and for each experiment the AP was forced to transmit at a fixed data rate. Each trace collection experiment was repeated for 5.5 and 11 Mbps physical layer (PHY) data rates. For each PHY data rate and for each SETUP, we collected traces for two distinct packet transmission rates. The transmission rate is controlled by adjusting the time interval t between packets. In setup A we collected traces at 500 Kbps and 1024 Kbps, while in Setup B we collected traces at 750 Kbps and 900 Kbps. For ease of notation, we prefer to label the traces by

Table I Trace numbering key for 5.5Mbps and 11Mbps

		Client					
		1	2	3	4	5	6
Setup A	500 Kbps	1	2	3	4	5	-
	1024 Kbps	18	19	20	21	22	-
Setup B	750 Kbps	6	7	8	9	10	11
	900 Kbps	12	13	14	15	16	17

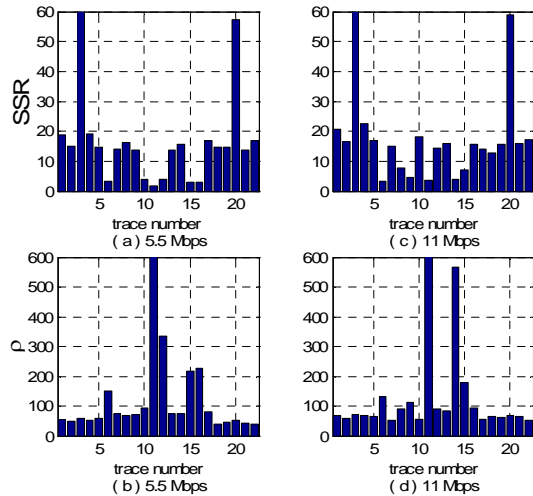


Figure 2 Average statistics for the 802.11b trace data set.

their PHY data rate and single number. Table I describes the numbering we shall use for the remainder of the work.

The receivers' MAC layer device drivers were modified to pass corrupted packets to higher layers. To capture packets at high transmission rates, packet dissectors were implemented inside the device drivers. These packet dissectors ensured that only packets pertinent to our wireless experiment are processed, while all other packets are simply dropped. In addition to a packet's header and payload information, for each packet three additional parameters were logged at the receivers:

- Background traffic (**B**): A four byte number representing the total number of background packets observed between two trace packets;
- Signal strength (**S**) for the received packet: A one byte number representing the signal strength in dBm;
- Silence Value (**N**) for the received packet: A one byte number which can be said to be representing the noise + interference strength in dBm.

We use the following defining equations to associate a SSR and BT intensity with each packet:

$$SSR = S - N \quad \text{and} \quad \rho = B/t \quad (1)$$

It is important to note that in typical 802.11b wireless receivers, the silence value **N** is measured just before the packet reception starts and the signal strength is measured for only the first few μs (~ 10). Thus the SSR indication associated with each packet is only an approximate measure of the link-quality experienced by a packet transmission. Similarly, also note that in measuring **B** we do not differentiate between packets on the basis of their size, signal strength, PHY rate etc. Thus we acknowledge that our measurement of the background traffic intensity, ρ , is relatively coarse and can be improved upon. Nevertheless, we shall see that even such coarse measurements are sufficient to

provide accurate channel prediction. Incorporating improved measurements of ρ is a part of our future work.

Figure 2 shows average statistics deduced from a subset of the error traces considered in this work. These average statistics should provide a good representation of the wireless environment in which we have conducted our experiments. In Figure 2, (a) and (c) show the average value of SSR; and (b) and (d) show the average value of BT intensity ρ . It can be seen that traces 3 and 20 which correspond to the LOS client have a very good link quality and thus they rarely saw any packet corruptions. Hence we often exclude these traces from our analysis. The value of ρ was less than 50 packets/sec for most traces but for certain traces the intensity can be well above 200 packets/sec. Comparing the plots for 5.5 Mbps and 11 Mbps, it can be seen that the long-term average value of SSR for a specific link does not vary significantly, however the same cannot be said about the BT. Additionally, note that as expected the packet corruption ratio is lower when the PHY data rate is 5.5Mbps as compared with the 11Mbps traces.

III. CHANNEL STATE INFERENCE

In this section, we shall investigate the utility of observable variables as side-information to be used for CSI. By CSI, we specifically refer to a problem where we want to estimate the BER in a packet that has already been received. Accuracy of such estimates plays an important role in soft-decoding algorithms [4] and can also be important for a variety of other reactive protocols, see [16] for examples.

A. Motivation

Figure 3 shows the average value of BER θ over all the corrupted packets as a function of SSR, for the cases when there is no BT and when there is heavy BT. Note that we use the term BT with reference to a particular wireless receiver, and by that term we refer to all transmissions that take place when the medium is not being accessed by the particular receiver. Under the assumption of no *delay spread* [17], the background traffic should not impact the reception quality of a receiver. However delay spreads are commonly observable in actual deployment of 802.11b

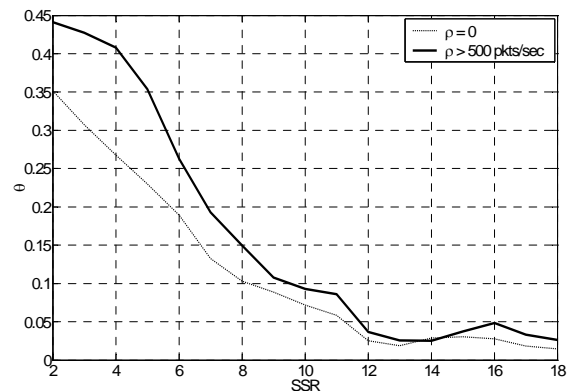


Figure 3 Impact of BT and SSR on BER, PHY rate 11Mbps.

WLANs and hence as shown by Figure 3 the amount of bit corruptions in a received packet may be impacted by the intensity of the BT. It can be clearly seen that the BER in a packet is correlated with the BT and SSR. Moreover the relationship of SSR with BER can vary in presence of BT. Thus it should be feasible to utilize SSR and ρ as side-information and additional performance benefits may be achievable by utilizing them in a joint manner. We investigate these issues in the following sub-sections.

B. Notation, Measures and Modeling Framework

The performance measures we choose are strongly influenced by the discussion on Method of Types presented in chapter 12 of [15].

Type of a packet: A received packet \bar{x} is said to be of type T_θ if the (sample) BER in the packet is θ .

We can use such a definition of type, purely because \bar{x} is made up of binary symbols and the probability distribution on these symbols is completely defined by the parameter θ . All the discussion in this section is strictly for the case of binary symbols and often we shall use θ to actually represent the probability distribution on the binary symbols.

Our goal in this paper is to infer or predict the type of a packet as accurately as possible. Additionally we assume that the side-information Z is able to accurately determine the type T_0 , i.e. the packets with no errors ($\theta=0$). Hence we concentrate all our analysis only on packets with type $T_\theta \neq T_0$. Therefore, we can represent the cost of estimating/predicting the type of a packet as $\hat{\theta}$ by

$$f(\theta, \hat{\theta}) = D(\theta \parallel \hat{\theta}), \quad (2)$$

where $D(\theta \parallel \hat{\theta})$ is the Kullback-Leibler (KL) divergence [15]. Thus $f(\theta, \hat{\theta})$ can be interpreted as a ‘‘loss in capacity’’ by assuming that the binary process \bar{x} is governed by the *estimated* Bernoulli parameter $\hat{\theta}$ when it is actually governed by the *true* parameter θ . This capacity loss interpretation can occur, for example, if additional information needs to be transmitted to compensate for the distance $D(\theta \parallel \hat{\theta})$. Further, it is important to note that the loss in capacity takes place regardless of whether the estimated $\hat{\theta}$ is pessimistic or optimistic relative to the true BER value θ , which corresponds to the packet type T_θ .

Model Training:

A dataset \mathfrak{A} , which is a set of packets in the present context, enforces a probability distribution $p_{\mathfrak{A}}(\cdot)$ on the types T_θ or on the parameter θ . $p_{\mathfrak{A}}(\theta)$ represents the

normalized frequency with which a packet of type T_θ is observed in a dataset \mathfrak{A} . Our approach is based on choosing a representative type $T_{\hat{\theta}}$ for a set \mathfrak{A} so that the average cost of misrepresenting members of \mathfrak{A} by type $T_{\hat{\theta}}$ is minimized. Thus from a given set we determine our model parameter as

$$\hat{\theta}_{\mathfrak{A}} = \arg \min_{\theta^*} \left(E_{p_{\mathfrak{A}}} \left[f(\theta, \theta^*) \right] \right), \quad (3)$$

where $E_{p_{\mathfrak{A}}} \left[f(\theta, \theta^*) \right]$ is the Expected KL Divergence (EKLD). Since $f(\theta, \theta^*)$ represents the loss in Shannon’s information (capacity), $E_{p_{\mathfrak{A}}} \left[f(\theta, \theta^*) \right]$ can be interpreted as the loss in the ergodic information (capacity). Note that $E_{p_{\mathfrak{A}}} \left[f(\theta, \theta^*) \right]$ measures the concentration of the types in set \mathfrak{A} around the type $T_{\hat{\theta}}$. Since in this case $\bar{x} \in \mathfrak{A}$ is defined on a binary symbol space the estimate of the representative Type (i.e. the model parameter $\hat{\theta}_{\mathfrak{A}}$), in accordance to (3), can be achieved rather simplistically. The following lemma shows that the parameter can be estimated by just taking an expectation over the training dataset.

Lemma 1: If $f(\theta, \theta^*) = D(\theta \parallel \theta^*)$, \mathfrak{A} s.t $\forall \bar{x} \in \mathfrak{A} \theta(\bar{x}) \neq 0$ and $\hat{\theta}$ is given by (2) then $\hat{\theta} = E_{p_{\mathfrak{A}}} [\theta]$.

Proof: The proof follows by solving $\frac{\partial}{\partial \theta^*} E_{p_{\mathfrak{A}}} \left[f(\theta, \theta^*) \right] = 0$.

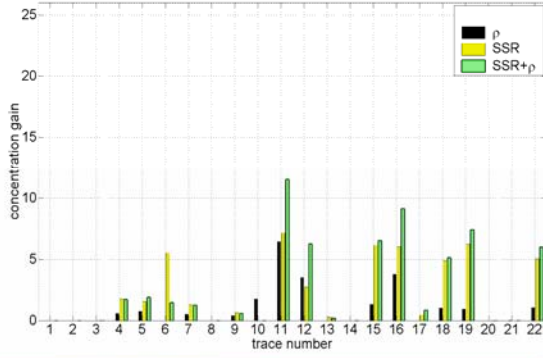
Until this point, our model was completely defined by a single parameter $\hat{\theta}$; however, the dataset we consider in this work is labeled. In particular, each \bar{x} is labeled by SSR and ρ . Thus we can obtain a set-decomposition of the training set \mathfrak{A} utilizing one or both of the labels. Let us denote these subsets by \mathfrak{A}_{SSR} , or \mathfrak{A}_ρ , or $\mathfrak{A}_{(SSR, \rho)}$, depending on the particular label(s) being used. For each of the subsets, we can determine a distinct parameter $\hat{\theta}$, which with a slight abuse of notation can be represented as $\hat{\theta}(SSR)$, or $\hat{\theta}(\rho)$ or $\hat{\theta}(SSR, \rho)$. These individual parameters are obtained by just taking the expectation in the subset. Thus the following equations define the correlation models that enable us to utilize BT and SSR as side-information

$$\hat{\theta}(SSR) = E_{p_{\mathfrak{A}_{SSR}}} [\theta] \quad (4)$$

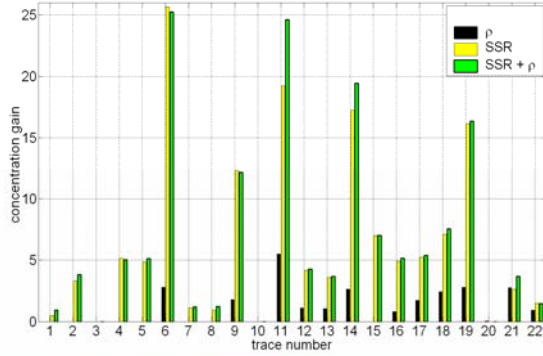
$$\hat{\theta}(\rho) = E_{p_{\mathfrak{A}_\rho}} [\theta] \quad (5)$$

$$\hat{\theta}(SSR, \rho) = E_{p_{\mathfrak{A}_{(SSR, \rho)}}} [\theta] \quad (6)$$

Thus in this section, we consider four models, which are



(a) 5.5Mbps



(b) 11Mbps

Figure 4 Concentration gain due to side-information .

referred to as *Average*, *SSR_aware*, *BT_aware*, *SSR+BT_aware* when the defining equation is (3), (4), (5) or (6) respectively.

Model Testing for Channel State Inference:

We test the models obtained in the above manner on various traces. Each trace Λ is represented by a vector time series $\{\theta_i, Z_i, SSR_i, \rho_i\}_{i=1:N}$ where $\theta_i, Z_i, SSR_i, \rho_i$ are obtained from the i^{th} packet in the trace. We use $Z=0$ if the packet is error free and $Z=1$ otherwise. As has been explained before, we focus our analysis entirely on packets for which $Z=1$. N simply represents the total number of packets in each trace. We use the models defined by (3)-(6) to obtain an estimated series $\{\hat{\theta}_i\}_{i=1:N}$. The utility of the model, for a particular trace, in estimating the type of a packet is obtained as

$$EKLD[\Lambda] = \frac{1}{N_1} \cdot \left(\sum_{i=1, Z_i=1}^N D(\theta_i \| \hat{\theta}_i) \right) \quad (7)$$

where N_1 represents the number of packets with $Z=1$ in trace Λ . Depending on the model being employed we can utilize SSR_i, ρ_i as side-information to obtain the estimate $\hat{\theta}_i$. In the absence of any side-information¹ $\hat{\theta}_i = \hat{\theta}$ where $\hat{\theta}$

is the average statistic obtained by employing (3), but when side-information is utilized we can write the estimate as $\hat{\theta}_i = \hat{\theta}(SSR_i)$, $\hat{\theta}_i = \hat{\theta}(\rho_i)$, or $\hat{\theta}_i(i) = \hat{\theta}(SSR_i, \rho_i)$ depending on whether (4), (5) or (6) is being used for model definition.

C. Results and Analysis

In this sub-section, we evaluate the performance of the above described models. The model parameters are obtained in accordance to (3), (4), (5) and (6). For each PHY data rate, we consider a distinct training set and this set is composed of the traces 1 to 11 at the data rate. We test the estimation accuracies of the trained models on each of the traces. Depending on the model being deployed, we get 4 measures for each traces which can be denoted by $EKLD_{Ave.}[\Lambda]$, $EKLD_{SSR}[\Lambda]$, $EKLD_{\rho}[\Lambda]$ and $EKLD_{SSR+\rho}[\Lambda]$. We are interested in quantifying the CSI improvements provided by the present side-information. For this purpose we define a measure, called concentration gain, as:

$$concentration\ gain = 10 \cdot \log_{10} \left(\frac{EKLD_{Ave.}[\Lambda]}{EKLD_M[\Lambda]} \right) \quad (8)$$

where $M \in \{SSR, \rho, SSR+\rho\}$. Thus the concentration gain represents the reduction in EKLD, on account of utilizing side-information, as compared to that obtained when only an Average statistic is employed. A gain greater than zero would imply that side-information leads to improving the accuracy of our estimate. The value of EKLD reduces when the values that the true parameter θ takes are closely concentrated around the estimate $\hat{\theta}$. Thus concentration gain is so named as it precisely represents the improvement in the concentration about the model parameters.

Figure 4 (a) and (b) show the concentration gains for 5.5 Mbps and 11 Mbps respectively. It can be clearly seen that both BT and SSR can provide concentration gains. Note a gain of 3dB implies an improvement in concentration by a factor of 2. Thus for 14 out of 20 traces at 11Mbps and for 10 out of 20 traces at 5.5Mbps, utilizing SSR as side-information can improve the accuracy of estimating BER by atleast a factor of 2 over a scheme that just utilizes an average statistic. It can be clearly seen that on some traces the improvement is in excess of 24 dB, corresponding to an improvement in concentration by a factor greater than 250. Thus the improvement in CSI by just utilizing SSR, despite the presence of BT, can be significant in many practical scenarios. As compared to SSR, the gains provided by BT are modest. However, there still exist a few traces [e.g., traces 6, 11, 14, 19, 21 at 11Mbps and traces 11, 12, 16 at 5.5Mbps] where the gain is close to or above 3dB. In certain cases, it is possible to jointly use SSR and BT side-

¹ Since we are concentrating all our analysis on corrupted packets it is implicit that Z is always being used as side-information. Thus in the remainder when

we discuss about the presence or absence of side-information we are specifically referring to the use of SSR and BT with an implicit value of Z=1.

information to achieve further gains. For examples of such cases, see traces 11 and 14 for 11Mbps, and traces 11, 12 and 16 for 5.5Mbps.

IV. CHANNEL STATE PREDICTION

In this section, we focus on CSP using temporal correlation in the BER of adjacent packets. The problem of CSP specifically refers to predicting the state of packet that has not been received as yet and thus we have no information about the associated SSR and ρ values too. Hence, the CSP problem considered in this section is distinct from the CSI problem considered in the last section. Accurate CSP can be fed back to the transmitter, which can in turn adapt the rate of source and channel codes according to the predicted state of future packets.

A. Motivation

Many research studies (e.g., [10]-[13]) have observed that temporal correlations can be observed in the channel states in 802.11b wireless networks. Figure 3 shows the correlation coefficients, calculated on the basis of the BER process, for some of the error traces utilized in this work. The coefficients can be calculated on the basis of the BER process $\{\theta_i\}$ as

$$\Gamma = \frac{E[\theta_i \theta_{i+1}] - E[\theta_i]E[\theta_{i+1}]}{\sqrt{\text{var}[\theta_i] \cdot \text{var}[\theta_{i+1}]}} \quad (9)$$

where the mean and variance are just the sample mean and sample variances derived from the traces. Figure 5 clearly exhibits the existence of significant temporal correlation in a number of traces.

B. Channel Prediction using BER temporal correlations

In the above sub-section we exhibited the temporal correlation in the BER in two adjacent packets. This BER correlation can be leveraged to predict the BER in a future packet. We can do this in a simplistic manner by utilizing the BER in the current packet as an estimate of the BER in the next packet:

$$\hat{\theta}_{i+1} = \theta_i \quad (10)$$

However, improved predictions can be obtained by developing a link-specific model that takes advantage of the temporal variations specific to the receiver's wireless link. This can be achieved by empirically evaluating the conditional probability distribution $p_\Lambda(\theta_{i+1}|\theta_i)$ to count the frequency with which various Type pairs $(T_{\theta_i}, T_{\theta_{i+1}})$ occur in a trace Λ . On the basis of this distribution we get an improved prediction as:

$$\hat{\theta}_{i+1} = g_\Lambda(\theta_i) = E_{p_\Lambda(\theta_{i+1}|\theta_i)}[\theta_{i+1}]. \quad (11)$$

C. Utilizing CSI from observable variables for CSP

The predictors described by (10)-(11) in the previous sub-section cannot be employed in many practical situations

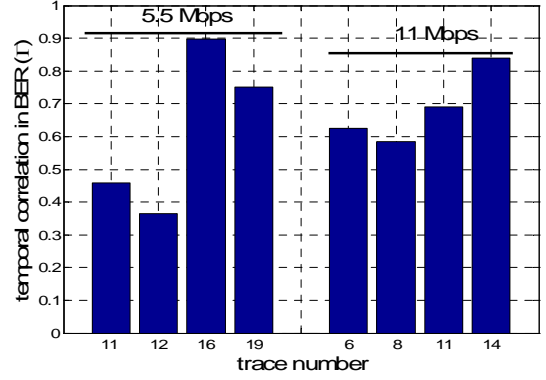


Figure 5 Temporal correlation in BER, PHY rate 11Mbps.

because θ_i is not observable whenever $Z_i=1$. In such circumstances we can utilize the side-information based CSI mechanism developed in the previous sub-section [refer to equations (4), (5), (6)] to estimate θ_i . For clarity and focus, we shall only utilize SSR as side-information. Thus the predictor described by (6) can be realized as

$$\hat{\theta}_{i+1} = \hat{\theta}(SSR_i). \quad (12)$$

We realize the predictor of equation (11) by estimating θ_i and then applying the temporal correlation model $g_\Lambda(\cdot)$ on the estimate. Thus the predictor can be realized as:

$$\hat{\theta}_{i+1} = g_\Lambda(\hat{\theta}(SSR_i)) = E_{p_\Lambda(\theta_{i+1}|\hat{\theta}(SSR_i))}[\theta_{i+1}]. \quad (13)$$

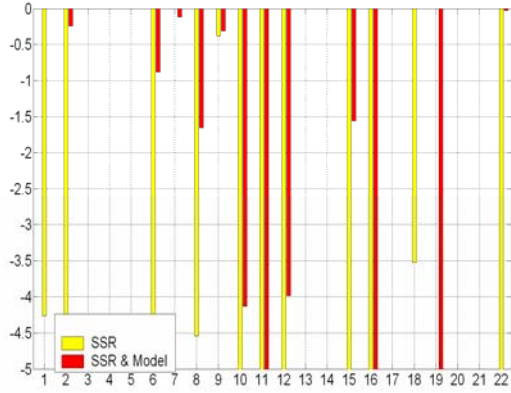
For the remainder of this section we refer to the predictor obtained from equation (11) as *Model based predictor*, the predictor obtained from equation (12) as *SSR based predictor* and the predictor obtained from (13) as *SSR + Model based predictor*.

D. Results and Analysis

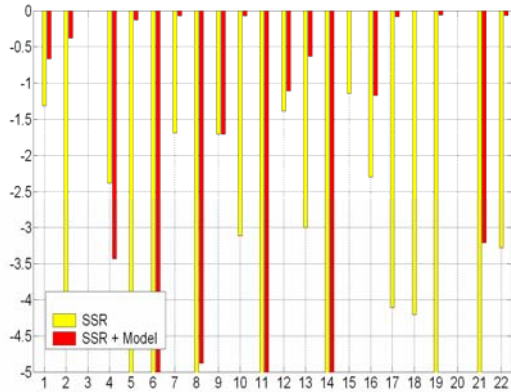
As in the previous section, we measure the accuracy of our various predictors by employing the EKLD measure as follows:

$$EKLD[\Lambda] = \frac{1}{N_{11}} \cdot \left(\sum_{i=1, Z_i=1, Z_{i+1}=1}^{N-1} D(\theta_{i+1} || \hat{\theta}_{i+1}) \right). \quad (14)$$

As expressed by equation (14), we limit the analysis for our predictors to only predicting errors in the corrupted packets. However, our methods can be easily generalized for packet level predictions, that is, predicting that a future packet will be corrupted. Due to brevity constraints and to maintain focus, we defer the development of such a model to a subsequent version of this work. Also note that in equation (14) N_{11} represents the number of times we see two consecutive corrupted packets in trace Λ .



(a) 5.5 Mbps



(b) 11 Mbps

Figure 6 Concentration loss due to imperfect CSI. In the above Figures, x-axis represents the trace number, while y-axis represents the concentration loss.

In this section, we are interested in determining the loss in accuracy incurred by employing *SSR based predictor* or *SSR + Model based predictor* instead of the *Model based predictor*. Thus we measure the concentration loss as:

$$\text{concentration loss} = 10 \cdot \log_{10} \left(\frac{EKLD_{Model}[\Lambda]}{EKLD_M[\Lambda]} \right) \quad (15)$$

where the term *Model* specifically refers to the predictor (11) obtained by training on a trace Λ , while $M \in \{SSR, SSR + Model\}$ is the predictor given by (12) or (13). It is important to note that the SSR models are obtained by training on the traces 1 to 11, just like Section II. The concentration loss thus represents the loss in accuracy due to an inability to observe the actual BER in the packet.

Figure 6 (a) and (b) show the concentration loss at 5.5Mbps and 11Mbps respectively. Ideally we would like the concentration loss to be never greater than 0dB, however such a performance is not consistently feasible. In practice slight loss of accuracy does not affect the

performance of an eventual error control mechanism significantly. Never the less, the exact impact of loss of accuracy is application dependent. For the purpose of discussion we consider two thresholds 0.15dB and 3dB which correspond to a loss in accuracy by a factor of 1.035 and 2 respectively. It can be seen that when we use the SSR based predictor the concentration loss, for 9/20 traces at 11Mbps and for 6/20 traces at 5.5Mbps, is less than 3dB. These numbers drop to 0 and 6 when we limit the concentration loss to approx. 0.15dB. Thus even though the SSR based predictor very often provides satisfactory prediction, a more consistently robust prediction mechanism is necessary. Hence, to reduce the loss, a link-specific model can be utilized in conjunction with SSR. It can be observed in Figure 6 that for *SSR + Model based predictor* the concentration loss is less than 3dB for 15/20 traces at both 5.5 Mbps and 11 Mbps. In fact it can be seen that for 10 traces at 5.5 Mbps and for 8 traces at 11 Mbps the concentration loss is within 0.15dB. Thus the proposed mechanism of combining a global and observable side-information-based model with a link-specific temporal correlation model can lead to significant performance benefits.

V. RELATED WORK

Measurement based studies are often important pre-cursors to design of efficient communication protocols. The past few years have seen some excellent comprehensive studies of 802.11b networks, see for instance [10]. However almost all of these studies limit their measurements to packet losses and are void of any measurement/analysis/modeling of bit errors. The authors' previous work is among the few studies that investigate the channel impairments in 802.11b networks at bit-level granularity [2], [4], [11]-[14]. Our initial efforts ([11]-[13]) in analyzing/modeling bit-errors were limited to developing channel models that tried to capture the temporal correlation in the bit-error process. Even though such modeling techniques are capable of capturing temporal correlations of a specific link, they are not very useful in CSI/CSP in an actual multimedia system primarily because the states of these models are not observable.

In [14] we presented a new modeling framework which utilized observable variables – in particular SSR – that have a link-invariant relationship with the bit error process and therefore can be used as meaningful side-information. The analysis in [14] is primarily applicable for developing simulation models. Even though the analysis in [14] provided us with insight that SSR can be used for CSI, the measure employed in [14] does not directly quantify the utility of SSR in CSI. In [14] we measure the KL-divergence $D(p_\Lambda(\theta) \| p_M(\theta))$, where p_Λ is the probability distribution on types obtained from the traces and p_M is the distribution obtained from a model, to measure the accuracy of the model. Thus, even though the analysis in

[14] motivated the utilization of SSR as a side-information in an actual multimedia system in [4], the measures employed in either [4] or [14] did not directly quantify the improvement in BER estimates. The present work fills this important gap in our previous contributions. Moreover, the present work makes many additional contributions; including the impact of BT on CSI and utilizing side-information for CSP, which to the best of the authors' knowledge have not been studied in prior literature

VI. CONCLUSION

In this work we investigated the feasibility of utilizing observable variables in channel state inference (CSI) and prediction (CSP). It was observed that both BT and SSR can be used to improve the accuracy of CSI. The gains provided by BT when used by itself or when used along with SSR are evident but not always significant. Thus there is enough promise in utilizing BT as side-information, but for improved performance finer measurements of BT are required. On the other hand, the gains provided by SSR in CSI are more significant and consistent. These gains were observed on a variety of traces which contained varying amounts of BT. Our experiments showed that the suggested CSI mechanism can be efficiently combined with the temporal correlations in the BER to facilitate channel prediction. We suggested a mechanism to combine a global SSR-based CSI mechanism with a BER-based link-specific temporal correlation model to facilitate channel prediction. This mechanism provided performance comparable to what would be achieved if the BER in a packet was observable.

REFERENCES

- [1] L. Larzon, M. Degermark, and S. Pink, "UDP Lite for Real Time Multimedia Applications," IEEE International Conference of Communications (ICC), Vancouver, June 1999.
- [2] S. A. Khayam, S. S. Karande, H. Radha and D. Loguinov, "Performance Analysis and Modeling of Errors and Losses over 802.11b LANS for High-Bitrate Real-Time Multimedia," Signal Processing: Image Communication, vol. 18, Aug 2003.
- [3] S. Karande and H. Radha, "Hybrid Erasure-Error Protocols for Wireless Video," to appear in IEEE Transactions on Multimedia.
- [4] S. Karande, U. Parrikar, K. Misra, and H. Radha, "Utilizing Signal to Silence Ratio indications for improved Video Communication in presence of 802.11b Residue Errors", IEEE International Conference on Multimedia & Expo (ICME), July 2006.
- [5] A. Servetti, J.C. De Martin, "Link-Level Unequal Error Detection for Speech Transmission over 802.11b Networks," Proc. Special Workshop in Maui (SWIM), Maui, Hawaii, January 2004.
- [6] E. Masala, A. Servetti, J.C. De Martin, "Standard Compatible Error Correction for Multimedia Transmissions over 802.11 WLAN", Proc. of IEEE International Conference on Multimedia & Expo (ICME), July 2005.
- [7] R. Riemann and K. Winstein, "Improving 802.11 Range with Forward Error Correction", MIT-CSAIL-TR-2005-011
- [8] F. Pauchet, C. Guillemot, M. Kerdranvat, S. Pateux, P. Siohan, "Conditions for SVC bit error resilience testing" Joint Video Team (JVT) 17th Meeting: Nice, FR, 14-21 October, 2005
- [9] S. Yi, Y. Shan, S. Kalyanaraman and B. Azimi-Sadjadi, "Header Error Protection for Multimedia Data Transmission in Wireless Ad Hoc Networks", ICIP 2006.
- [10] D. Aguayo, J. Bicket, S. Biswas, G. Judd, R. Morris, "Link-level Measurements from an 802.11b mesh network", Sigcomm 2004.
- [11] S. A. Khayam, H. Radha, S. Aviyente, and J. R. Deller, Jr., "Markov and Multifractal Wavelet Models for Wireless MAC-to-MAC Channels," to appear in Elsevier Performance Evaluation Journal.
- [12] S. A. Khayam and H. Radha, "Linear-Complexity Models for Wireless MAC-to-MAC Channels," ACM/Kluwer Wireless Networks (WINET) - Special Issue on Selected Papers from MSWiM'03, vol. 11, no. 5, pp. 543-555, September 2005.
- [13] S. A. Khayam and H. Radha, "Constant-Complexity Models for Wireless Channels," IEEE Infocom, April 2006.
- [14] S. Karande, U. Parrikar, K. Misra and H. Radha, "On Modeling of 802.11b Residue Errors," Conference on Information Sciences & Systems (CISS), March 2006.
- [15] T. M. Cover and J. A. Thomas, "Elements of Information Theory," Wiley Series in Telecommunications.
- [16] M. Schaar and S. Shankar, "Cross-Layer Wireless Multimedia Transmission: Challenges, principles, and new paradigms," IEEE Wireless Communications Magazine, vol. 12, no. 4, pp. 50-58, August 2005
- [17] T.S. Rappaport, "Wireless Communications – Principles and Practice, 2nd ed," Pearson Education, Singapore, 2002.