Empirical validation of the modeling assumptions about Signal Strength and Error Rates under a Fading Wireless Channel

Vinay Kolar*, Saquib Razak*, Petri Mähönen† and Nael B. Abu-Ghazaleh‡
* Department of Computer Science,
Carnegie Mellon University, Qatar.
{vkolar,srazak}@cmu.edu
† Department of Wireless Networks,
RWTH Aachen University, Germany.
pma@mobnets.rwth-aachen.de
‡ Department of Computer Science,
State University of New York, Binghamton.
nael@cs.binghamton.edu

***Abstract*—The uncertainty of the wireless channel in realistic wireless networks inhibits the effectiveness of the of the higher level protocols and models to predict the effect of channel. Measurement based models and protocols overcome this hurdle by measuring the essential lower level parameters and use these empirical data as an input for decision making at the higher layers. Received signal strength and error rates are two such measured parameters that are extensively used for capturing the effect of randomness in the wireless channel. Statistical characterization of the empirical data not only helps to understand the randomness of their behavior but also benefit the measurement based applications that base their judgment on the observed values. In this paper, we analyze the empirical values of received signal strength and error rates from IEEE 802.11 indoor wireless network. We methodically classify the links based on their behavior, analyze their statistical distribution and independence properties, and study the effect of fading for varying packet sizes and standard modulation schemes. While classifying links into discrete categories has been proposed before, the statistical properties of each category and the mechanism to recognize the category of the link remains uncertain. In this paper we show that while all links do not exhibit a general statistical behavior, the statistical behavior of each category can be generalized. Each category has different behavior in terms of statistical distribution, memory and the temporal variation of the parameters. We observe that the error rates of strong and weak links are more stable and predictable, while the effect of fading introduces significant volatility in an average quality link.**

I. Introduction

Recent years have seen a tremendous growth in the applications and the deployment of wireless data networks due to its ability to enable pervasive applications. With such a wide-spread demand for wireless access and sparse capacity of wireless networks, predicting and optimizing the performance of the applications operating in a wireless data networks is a high-impact and a hard research problem. A perfect characterization of wireless data networks is unrealistic since the wireless links typically experience high error rate, large channel variations and complex interference patterns. Existing experimental studies [1], [2] have shown that simplistic assumptions, that are often employed by simulators and models to judge the performance of the network, deviate significantly from the observed phenomenon.

In such an volatile environment, measurement based models [3]–[5] and protocols [6], [7] are being extensively used to empirically measure the unpredictable parameters and use these measured values for predicting or optimizing the applications at different layers of the network stack. Such *Measurement Based Applications (MBAs)* employ a measurement phase where measurement traffic is injected into the network for measuring the necessary parameters. Then educated heuristics are employed to predict or optimize the performance of the protocol by using these measured parameters. They are more susceptible to changes in network topology and traffic and have shown outperform the non-measurement based counterparts by a large margin [6], [7].

One of the first steps in realizing the capability of MBAs in realistic networks is to capture the important low-level parameters by considering the practical limitations of the wireless cards. A vast majority of the current wireless data networks use IEEE 802.11 based cards which restrict the altering and fetching all the low-level parameters. For example, off-the-shelf Atheros based 802.11 cards provide allows to fetch the Received Signal Strength Indicator (RSSI) and Noise at the granularity of a packet. While it allows us to choose transmission rate from a set of standard rates, it dictates the modulation used for each rate. Received Signal Strength (RSS), Signal-to-noise ratio (SNR) and the packet/bit error rates (PER/BER) of a link are some of the primary measurable parameters that are used by many higher layer protocols. For example, a mixture of these parameters are used in several higher layer protocols like: (1) AP association in WLANs [8]; (2) Transmission rate-control in WLAN and Multi-hop Wireless Networks (MHWNs) [6], [9]; (3) Access the quality of the links for routing in MHWNs [7]. We focus on the RSS and the error rate parameters in this paper since SNR can be derived directly from the measured RSS and noise values.

Based on the measurement period, MBAs can be graded from *static MBAs* where values are measured once and used forever to *real-time MBAs* where measurement is performed periodically. While frequent measurement and dissemination of these values to the neighboring nodes will enable better performance of the MBAs, it drastically increases the measurement overhead. For example, many routing protocols like OLSR and ETX requires transmission of broadcast packets to infer the link quality. Constant transmission of these control packets results in humongous measurement overhead while infrequent measurement is vulnerable to stale and unrepresentative values. It is necessary to study the temporal variation of these primary parameters in order to optimize these measurement window. In this paper, we first study the this trade-off by analyzing the decay of the measured values with respect to time for different variety of links. Specifically, our analysis answers the following questions: (1) Do these parameters vary over different quality of links and different modulation schemes? If so, how? (2) How often should the RSS be disseminated in the neighborhood such that they are representative? Can this dissemination overhead be justified? From these experiments, we conclude that the time-variation of the parameters is strongly dependent upon the quality of the link and dissemination of RSS values once in 2-5 seconds is reasonable for a majority of the links.

Another important aspect for MBAs is the approximating the variation of of the measured values within the measurement period. A wide variety of assumptions are employed in characterizing the variations of the primary parameters within the measurement period which have not been thoroughly validated by experiments. These parameters have been represented in different forms in the literature. While some research work assume these parameters as constants [5], [10], [11], others approximate them as i.i.d. random variables from some known statistical distribution [3], [12]. Many studies have even pointed out that the RSS varies rapidly and randomly [2] and it is hard to deterministically or statistically characterize them. With such differing abstractions and conclusions about these primary parameters, empirical study and characterization not only provides insight into their behavior, but also aids in the design of practical real-time models, network planning/provisioning tools and design of efficient dynamic protocols. The second contribution of the paper is to validate the generality of these assumptions by analyzing the statistical properties of RSS and error rates for different link qualities and transmission rates. We find, as previously stated, that the distribution of the RSS and PER values are strongly dependent on the link quality. Strong links and weak links are separated by a layer of transitional grey zone where fading of the channel leads to unexpected and interesting distribution of the error rates. We analyze the variation of error rates with respect to different transmission rates. Finally, we use our conclusions from the empirical study of these parameters to develop a throughput model and show our model the results of the model with the experimental values.

II. Experimental observation

In this section, we analyze the statistical properties and the behavior of RSS and error rates.

A. Measurement methodology

The experiments were conducted using laptops with NEC Aterm WL54AG(S) wireless cards with MadWiFi driver [13]. These cards have Atheros chip-sets and a port for attaching an external antenna. We have modified the driver to collect the information of received packets. It also provides the information about erroneous packets where only header was received successfully and the rest of the packet failed the CRC test. This allows us to log the information about the RSS, transmission rate and time about the all the packets where only header correctly received, thus providing a finer granularity RSS measurement.

Unless mentioned, IEEE 802.11a with basic mode is used in the experiments. IEEE 802.11a is used to avoid the external interference in the 2.4GHz band. Spectrum analyzers are used to ascertain the absence of external interference while the experiment is in progress. Transmission rate-control modules are disabled and fixed-rate is used to deterministically characterize the behavior of the parameters under a known rate and modulation scheme.

A CBR client and server application is used to generate the broadcast and unicast traffic. Since the objective of the paper is to measure and analyze physical layer parameters like RSS and error rates, and not to study the effect of MAC protocols like exponential back-off, we employ saturated broadcast traffic. Atheros cards exhibit a transmission strategy where alternate packets are transmitted at two different power-levels if the source does not receive any form of acknowledgement (e.g. in broadcast transmissions). The effect of this design anomaly in Section IV-B. But, for a majority of the paper, we analyze the RSS by forcing the card to transmit all the packets at a single power (as suggested by Giustiniano *et al.* [14]) since we are interested in a more general case of capturing the effects of wireless propagation without this anomaly.

B. Analysis of Received Signal Strength

In this section, we study the variation of RSS time-series to conclude if we can approximate the RSS as a random distribution over large time scale. This is also of critical importance in many protocols and models that assume that RSS values from a node is either known or communicated to the neighboring nodes [11], [15]. If the RSS varies to a large extent over fraction of seconds, then measurements have to be taken in very small durations of time and thus adding to a lot of measurement and control overhead. The analysis also answers the question about how often to exchange the RSS measurements. For the distributed protocols and models which assume such a knowledge of RSS, it is important to answer the important system designing questions like: (1) How often should RSS values be disseminated to the neighbors such that the protocols and analytical models will have a fairly consistent view of the neighbors' RSS?; (2) Does the variation of the RSS values depend upon the quality of the link? In this section, we empirically analyze the RSS values to answer such important questions.

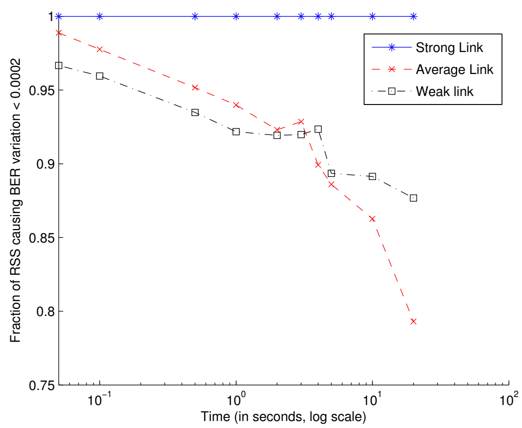


Fig. 1. Percentage of RSS that cause a BER variation of 0.0002 or less

1) How often should RSS information be disseminated: We analyzed the RSS time-series for a number of links, both having a line-of-sight in component (LoS links) and non-line-of-sight (nLoS links), in an indoor office environment. Similar to several existing studies [16] and for the clarity of presentation we classify the links into three zones: (1) A *Low-loss zone* where the links observe a very few packet losses; (2) A *High-loss zone* where the link experiences very high rate of packet errors; and (3) An intermediate *Gray-zone* where the link quality fluctuates between low and high losses. We use the term *strong, average and weak link* to refer to the link in low-loss, gray and high-loss zone, respectively. We observed that links can be classified into these zones with fair accuracy. Due to space limitations, we present the results by choosing one representative link in each zone. However, the conclusions we derive are based on the analysis of all the links that were measured.

Many protocols and models assume accurate value of the RSS at the nodes. In reality, the correlation between the true RSS and the measured RSS value decays over time. We study the decay of the correlation between expected RSS and observed RSS with respect to time. One approach of studying the decay is to study the decrease in the auto-correlation function (ACF) of the RSS time-series. However, this has two drawbacks: (1) The Bit Error Rate (BER) does not vary linearly with the RSS. Hence, a small difference of RSS for a link in the low-loss zone has lesser effect on the performance than the same difference for a gray-zone link; (2) As we will see later, the distribution of the RSS varies widely depending upon the quality of the link.

Our results indicate that comparing the decay of ACFs of different distributions will not quantify the variance of RSS over time. Hence, we use a more direct approach: We consider a series of measured RSS values in the time interval T and find its mean. We then calculate the empirical BER (by our measurement data) for the mean RSS and predict the RSS interval which causes a change of δ in the BER. Then, we find out the fraction of the RSS population that would have resulted in the BER within δ tolerance. Intuitively, this gives the fraction of the RSS in an interval T that does not significantly alter the BER. We conduct this experiment for different values of T to study how the fraction of the population deviates from the expected BER.

Figure 1 shows the population fraction that is within $\mathrm{BER} \pm 0.0002$. The strong link does not have any variations that cause the BER shift of more than 0.002. The RSS of the average and weak link decays faster initially (upto around 2 seconds), but the the average link is prone to have greater variation in the BER at time scales of 10's of seconds. This is intuitive since the average link is in the gray-zone where slight alteration of RSS will lead to large variation in BER[1]. Hence, BER is almost constant for strong links, stabilizes for the weak link at around 2-3 seconds and the decay increases drastically after 2-3 seconds for the average link. Hence, a choice of 1s to 3s will be a reasonable period to disseminate RSS values to the neighbors. Updating RSS values by overhearing the beacon packets in single-hop wireless networks or HELLO messages sent by certain routing protocols in multi-hop wireless networks[2] gives a fairly good estimation of RSS from the neighbors. The measurement models and other protocols can benefit from recording the RSS values of such control packets.

2) Distribution of RSS: The approximation of RSS as a constant or a i.i.d. random variable from some distribution has been extensively used in the literature [3], [10]. However, the validity of these assumptions and the effect of such approximations is not extensively tested, even though the assumption of i.i.d. random variables has a large effect on the effectiveness of the model. In this section, we analyze the empirical RSS values to infer about the independence and distribution of the RSS values.

The RSS time-line for each link is first divided into segments of duration 1.5 seconds each (approximately 750 RSS values for 1460 byte sized packets) and the independence and distribution-fitting tests are applied for each of these segments.

The measured RSS values for each packet is verified for independence by plotting the *Auto-correlation function (ACF)*. ACF is a statistical metric for verifying if there is a repeating pattern in the measured values, thus indicating the non-randomness in the data. ACF is defined at various positive integer points n (called as *lags*) and is a value between $[0, 1]$ at each lag. ACF at lag n denotes the correlation between

[1] BER variation of $\delta = 0.0002$ will approximately change the PER by 0.1 for a 512 byte sized packet under operating RSS values. Different reasonable δ values were experimented and it was concluded that the trend of the curves remains the same.

[2] The default interval of HELLO messages in widely used routing protocols like OLSR is around 2s.

the measurement point at time t and $t+n$. A perfect random variable should have a ACF of 1 for lag 0 and ACF of 0 at all the other lags.

For the distribution-fitting tests, we have to extrapolate the histogram of discrete RSS values into a continuous variable pdf by using an appropriate *Kernel Density Estimation (KDE)* technique. This is because the actual RSS is a continuous variable and the cards report only discrete values of RSS (integer values of RSSI). Without KDE, the distribution fitting tests fail to match any of the continuous distributions even when there is a very high visual similarity in the PDF of the measured RSS values. We then sample 100 values from this extrapolated population and conduct distribution-fitting tests using Kolmogorov-Smirnoff tests.

In order to summarize these tests for a large number of samples, we use the box-plot notation. The box-plot summarizes groups of data (e.g. ACF for each packet transmission in Figure 6(a)) by a box (that bounds the upper and lower quartiles of the data), a median (a horizontal line) and the outliers (denoted by '+' marks).

a) Histograms of RSS: Figure 2 shows the overall histograms of the RSS for the strong, average and the weak links. We can infer that strong links have very little variation of RSS and as the link gets weaker, the variation of RSS increases. To analyze the variation of RSS in a much shorter duration of time, we analyzed the variation of RSS in 1.5 second interval. The results indicate that the largest standard deviation was found for the weak links (a value of approximately 2 dB) and strong links exhibited almost no variation in such time interval.

b) Distribution fitting tests: Kolmogorov-Smirnov (KS) test with significance level $\alpha = 0.05$ was used to test the distribution of the extrapolated RSS with standard distributions and the results were randomly verified by visual matching of the PDF, CDF and the Q-Q plots. Log-likelihood test were also conducted to verify the conclusions. Certain tests like Lillifor test for normality was conducted if KS test for normal distribution resulted in a positive match. For succinctness, we demonstrate the results of KS test in this paper.

Figures 3, 4 and 5 shows the outcome of the KS tests. It can be inferred that distribution of the weak links can be coarsely approximated as log-normal distribution. This was true in a majority of the weak links which suggests that the approach used by Qiu *et al.* [3] is valid for weak links. Strong links do not follow any of the distribution. As the histogram in 2(a), it was observed that RSS for the strong link is very well approximated as a constant rather than a random variable from a specific distribution. The average link in the gray zone match normal distribution in approximately 50% of the cases (e.g. Figure 4). However, the approximation is not very accurate. .

c) Independence of RSS: Figure 6 shows that most of the RSS values in the time-period of 1.5 seconds are independent since the ACF at lag 1 is approximately zero. However, strong auto-correlation was observed for weak and average links as the period was increased. This suggests that slow variation of the channel over longer period.

In summary, from the histograms, distribution-fitting tests and the independence tests, we conclude that: (1) Strong links are best approximated as constant and the disseminated RSS values can be used for a longer duration of time (approxmiately 10s of seconds); (2) Weak links can be coarsely approximated as i.i.d random variables from a log-normal distribution with appropriate parameters; (3) Average links are independent random variables but their approximation as a random variable from a specific distribution is not consistently true. However, the variation of RSS is very small for a time-frame of approximately 1.5 seconds. This also suggests that a constant RSS approximation is reasonable for this time-frame.

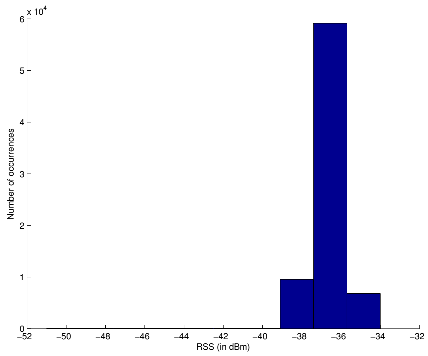
3) Variance of RSS in different periods of time: The above section specifies the average decay of the correlation of the RSS values over time with a rough estimate of the tolerated variance in RSS. However, it does not explicitly quantify the amount of variance of RSS for different time intervals. It also does not capture the short-term variation of RSS during different time phases of the link where the RSS is inherently low or high due to environmental effects. We now measure the variation of the RSS in different time-intervals T and scenarios to quantify the fluctuations of RSS over T and different time phases. This experiment is conducted for two reasons:

- To examine the time intervals during which the RSS can be assumed to be constant, thus aiding measurement based models and protocols to set a precise intervals to update the RSS data among the nodes.
- Establish empirical relationships between the short-term and long-term RSS fluctuations. This helps to identify the phases where RSS can be used as a stable link quality estimator.

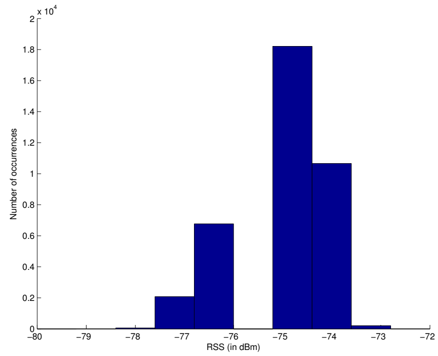
Similar to the experiments in the previous section, the effect of variance in indoor scenario with LoS and nLoS scenarios are measured. However, a saturated CBR traffic with a smaller packet (64 bytes including the MAC, IP and CBR headers) is selected. This enables us to receive an RSSI value for every 132 microseconds and hence measure the fluctuation with a finer granularity.

A time interval T is selected and the variation of the RSSI value over T is measured. During each interval the mean and variance of the RSS values are collected. Since the mean RSS over T varies over longer time frame, the mean RSS values are divided into different bins and the variance of the RSS in that bin is measured. This enables us to measure the RSS variance as a function of the fluctuation of across different time intervals T. The time interval T was selected from 5ms to 3s. Figures 7 and 8 show the variance of RSS for different LoS and nLoS scenarios. The error bar plots the 95% confidence intervals for RSS variance. As a general trend, it can be seen that:

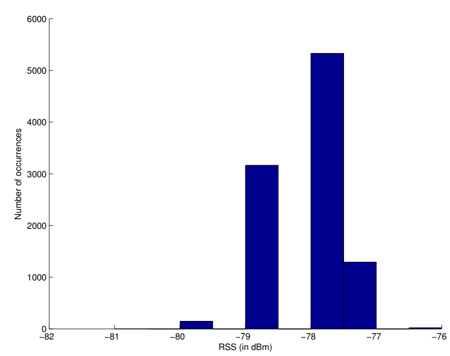
- The variance generally decreases as the RSS values increases, reaches minimum at a point (which is the long-term mean of the RSS), and slightly increases or remains the same. Hence, the variation of short-term RSS are highly unstable if their values are much lower than the long-term average.



(a) Low-loss link

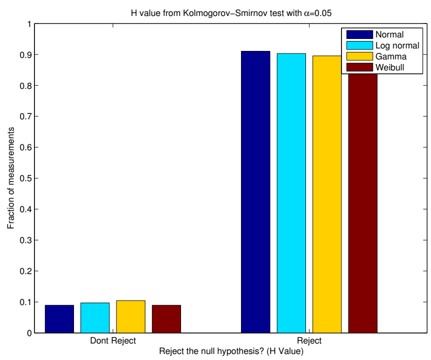


(b) Gray-zone link

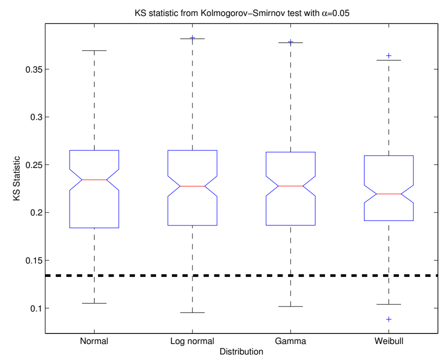


(c) High-loss

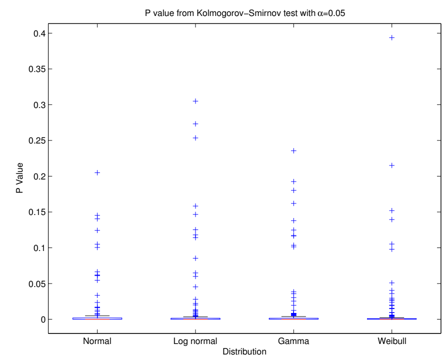
Fig. 2. Histograms of RSS



(a) K-S Test: H-value link



(b) K-S Test: KS Statistic



(c) K-S Test: P-value

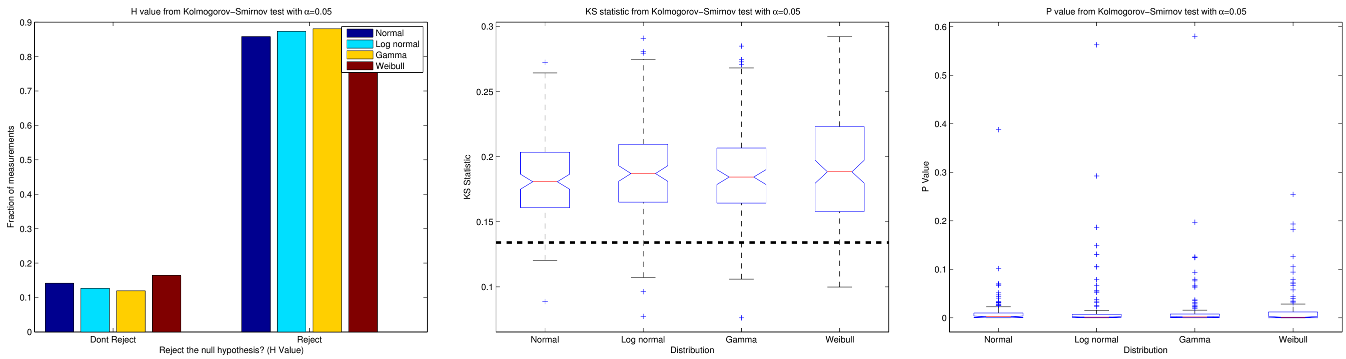
Fig. 3. Distribution of RSS: Strong link

- The variation of the RSS variance, as given by the error bars, indicates the fluctuation of *variance of RSS values*. A larger fluctuation indicates that RSS variance fluctuates significantly when the mean RSS is smaller. The error bars are very high when the mean RSS is low, reaches minimum at around the long-term mean of the RSS, and slightly increases thereafter. This suggests that the accuracy of using the short-term average RSS values will be ineffective in predicting link quality when the short-term average RSS is significantly lower than the long-term average of RSS.
- The variance of RSS for stronger links does not vary significantly over time scales. For example, the difference between the variance curves for the strongest LoS link (Figure 7(a)) is much lesser than that of the strong LoS link(Figure 7(c)). Hence, stronger links can be updated with lesser frequency while the weaker links should be updated with higher frequency.
- The variance of the RSS does not fluctuate significantly even at 3s intervals (around 1.2 dBm in the worst case). Hence, RSS dissemination by beacons and control messages of the routing protocols is a viable option to keep track of RSS. Very frequent updates of RSS values (order of fraction of seconds) which causes extreme overhead, especially in the multi-hop wireless networks, is not necessary.

C. Error rate analysis

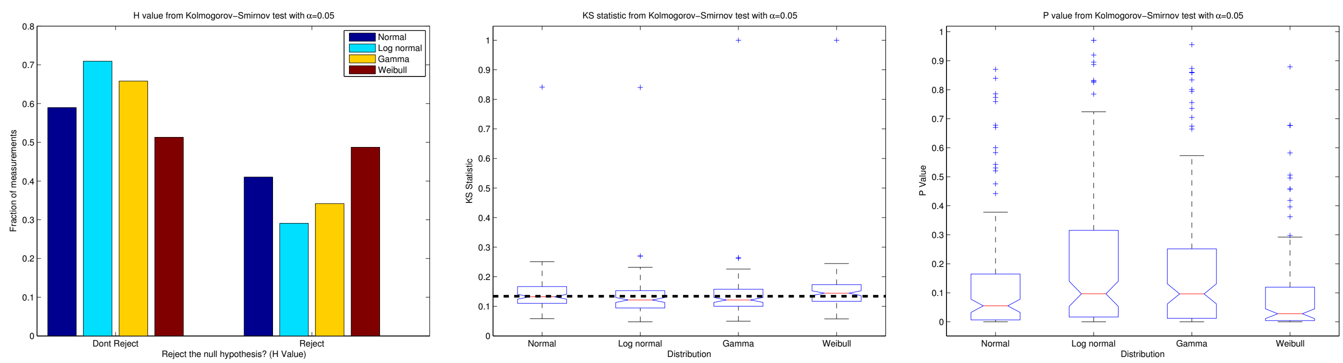
In this section, we study empirical Packet Error Rate (PER) and Bit Error Rate (BER), other important parameters that are important to predict the performance of the link. In this section, we first study the PER with RSS values which gives a rough estimate of the PER with respect to the observed signal strength. We then perform an in-depth analysis to isolate the effect of fading and state the mechanisms for identifying the gray-zone links where the link error rates are extremely volatile. We then test if PER can be assumed as an i.i.d. random variable and analyze the results of distribution fitting for the PER values. This is useful to validate the assumptions of many models which assume PER as either constant or as an i.i.d. random variable from some distribution. Finally, we observe the effect of different packet sizes and transmission-rates on PER.

1) Error rate as a function of RSS: Figure 9(a) shows the Packet Error Rate (PER) against the measured RSS values with the error bars capturing the $[12.5\%, 87.5\%]$ percentiles. It also plots the theoretical PER calculated for BPSK under AWGN and Rayeigh fading channels. As in the specification of 802.11a broadcast packets, we used the theoretical curves were



(a) K-S Test: H-value link (b) K-S Test: KS Statistic (c) K-S Test: P-value

Fig. 4. Distribution of RSS: Gray zone link



(a) K-S Test: H-value link (b) K-S Test: KS Statistic (c) K-S Test: P-value

Fig. 5. Distribution of RSS: High-loss link

calculated for BPSK modulation with convolutional codes. We used Rayleigh fading since we are more interested in indoor environments where line-of-sight path is not readily available. The PER is calculated assuming that the BER values observed during the packet transmission time are constant and

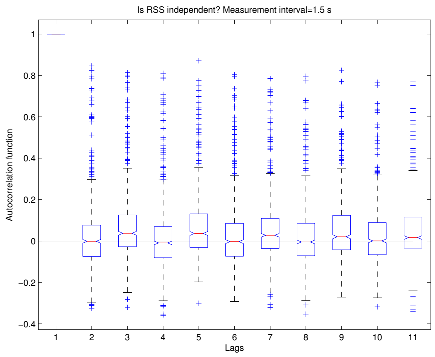
$$\mathrm{PER} = 1 - (1 - \mathrm{BER})^D \quad (1)$$

where D is the size of the complete packet (including PHY and MAC headers). This equation is often used while modeling the MAC layer throughput to abstract the PHY level details like modulation and coding. Hence, we computed the BER curves for the combination of the protocol (802.11a/b/g) and the allowed transmission rate and use it as a basis for MAC layer modeling.

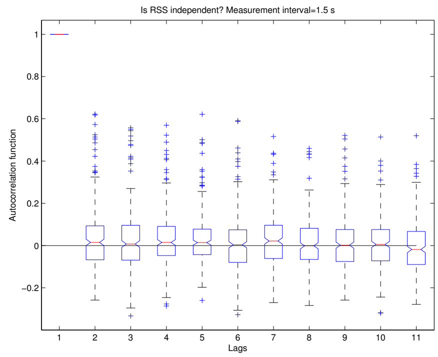
The division of PER into three piecewise zones can be clearly identified by observing the experimental PER curves: (1) A low-loss zone with low and constant PER with small variance; (2) A gray zone where PER varies widely (from 0.2 to 0.9); and (3) A high-loss zone where PER approaches 1 with acceptable variation. It can be observed that irrespective of the link type (strong LoS link or strong nLoS link), the RSS can be directly mapped to an almost constant value in the low-loss zone. In the gray zone, we observe that aggregate metrics of mean PER is not sufficient to capture the error rate due its huge variation. Section II-C2 investigate this region in detail.

Identifying the relation between RSS and BER (instead of the PER) is useful for modeling and protocol design of the wireless networks where varying packet sizes are often transmitted. Since the BER depends upon the type of the channel and the card specifications, accurate measurement of the BER is infeasible for dynamic applications. Hence, we derive the BER from the easily measured PER values. Figure 9(b) shows BER curve and compares it with the theoretical BER curve, with and without fading. From the Figure 9(b), we infer that: (i) In the low-loss zone, constant BER assumption is a better approximation than the fading models; (ii) In the gray-zone the BER is significantly deviates from the regularly observed trends while Rayleigh fading approximation is good in some areas; and (3) In high-loss zones, the measured BER approaches 1 more rapidly than predicted by theoretical models as the RSS value is lowered. The observed illustrates the need for using empirical BER values in protocols and models that predict higher layer performance in wireless networks with standard wireless cards.

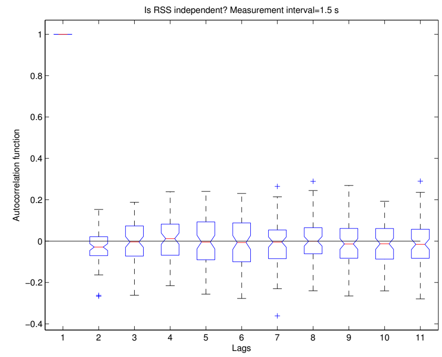
2) Determining the cross-over RSS values: Figure 9(a) discussed the variation of PER with RSS for three types of links. In this section, we answer another important question for



(a) Strong link

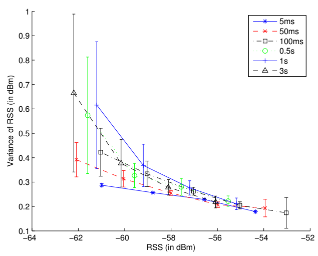


(b) Average link

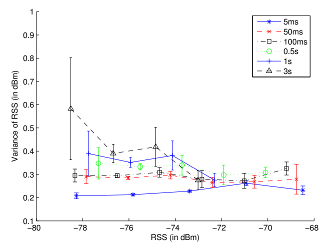


(c) Weak link

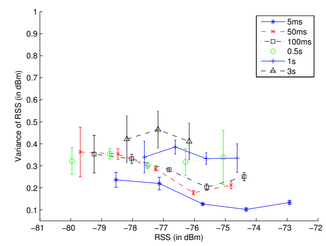
Fig. 6. Independence of RSS



(a) LoS 1: Strong

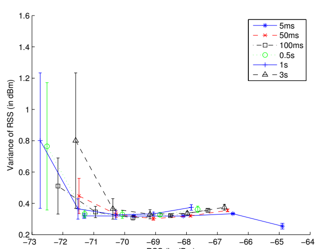


(b) LoS 2: Medium

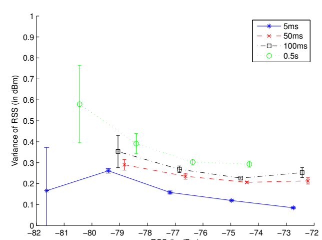


(c) LoS 3: Weak

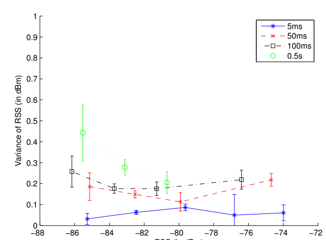
Fig. 7. Mean vs. Variance in LoS links



(a) nLoS 1: Strong



(b) nLoS 2: Medium



(c) nLoS 3: Weak

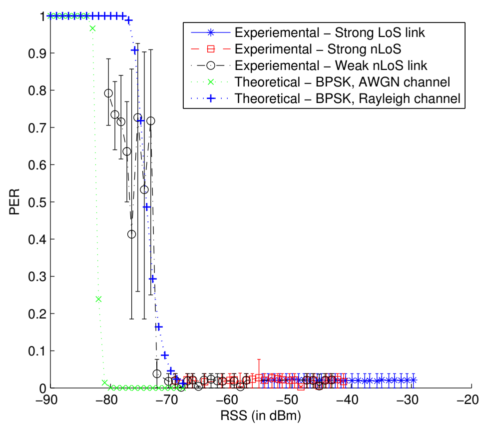
Fig. 8. Mean vs. Variance in nLoS links

measurement based models and protocols: "How to infer the zone in which a given link operates?". Based on the inference, protocols and the models can tune their BER/PER curves to estimate the link quality.

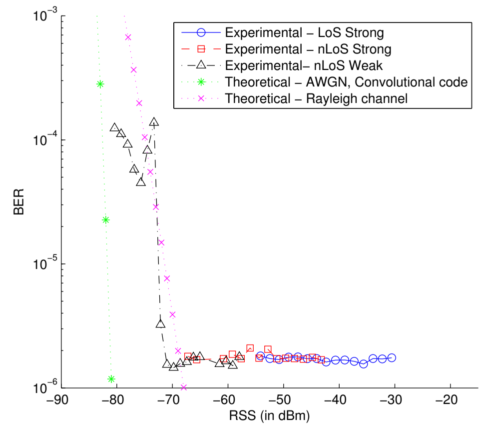
Different links cross-over from the low-loss zone to gray-zone at different ranges of RSS due to the variation of the channel between the nodes of different links. The RSS value for which a strong outdoor line-of-sight link crosses from the low-loss zone to the gray-zone is different from the cross-over point for an indoor non-line of sight link due to the varying effect of fading. Hence, absolute cross-over points cannot be dictated for estimating the quality of the link. Secondly, as shown in Section II-B2 (Figure 2), the RSS has an almost similar distribution for links in all the zones: a small variation and the histogram is unimodal for links in all the zones. Hence, sole measurement of RSS per link will not suffice to decide the cross-over points of PER.

We now analyze the gray-zone of the PER for inferring the cross-over ranges. This is a critical area of the PER curve where PER changes dramatically and follows unintuitive trends. For example, figure 9(a) shows that average PER dramatically increases instead of decreasing as the RSS is increased at around -74 dBm in our measurements. We study detailed PER measurements for analyzing and answering the trends in this region.

Figure 10 plots the histograms of the PER and the RSS values in an experimental setup where the links do not have a line of sight. We vary the transmit power at the sender

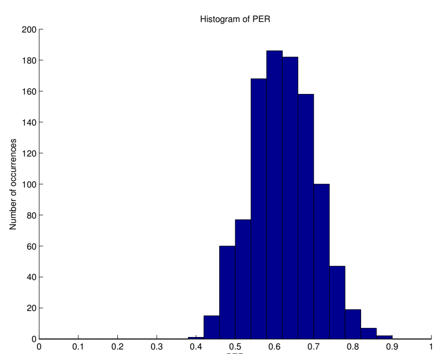


(a) PER

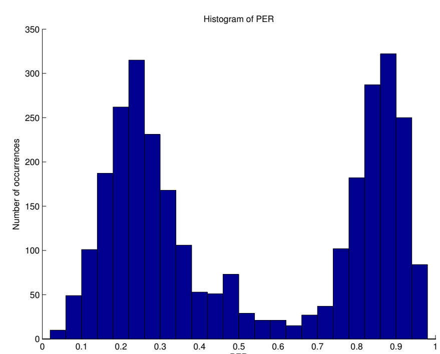


(b) BER

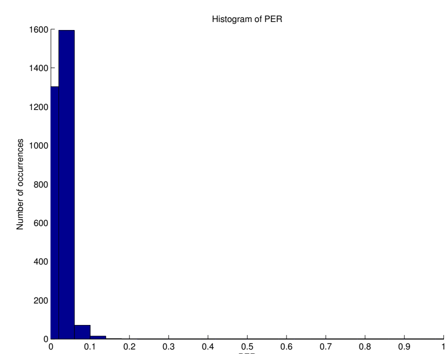
Fig. 9. Error rate vs. RSS



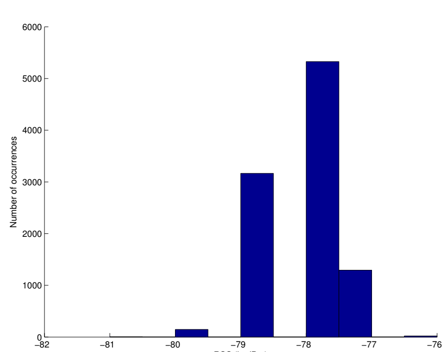
(a) PER at 0 dBm transmit power



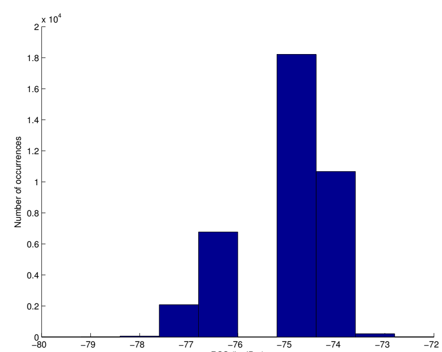
(b) PER at 3 dBm transmit power



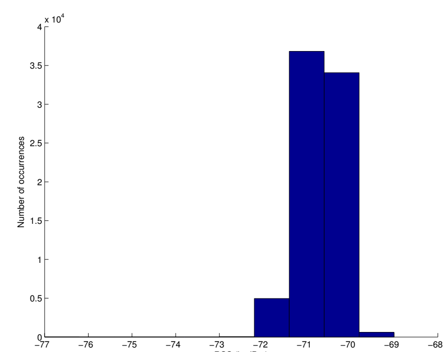
(c) PER at 6 dBm transmit power



(d) RSS at 0 dBm transmit power



(e) RSS at 3 dBm transmit power



(f) RSS at 6 dBm transmit power

Fig. 10. Effect of fading on PER

and observe its effect on PER and RSS. Figure 10(a) shows that the PER is high and has a nice unimodal distribution when the transmit power is low. At high transmit powers, a low PER with a similar unimodal distribution is observed (Figure 10(c)). Intuitively, we expect the PER histogram to shift from high to low (with probably decreased variance) as we increase the transmit power. However, Figure 10(b) shows that at a certain intermediate RSS value, the PER histogram dramatically differs from the expected one. As shown in Figure 10(b), the PER fluctuates dramatically between low and high values over a very small RSS range (even on a time scale of fraction of seconds). It becomes bimodal with a large separation between the two modes. This is clearly an effect of fading [17]. In addition, Figures 10(d), 10(e), 10(f) also see that there is no significant difference in the shape of the RSS histograms for these transmit powers. The histogram of the measured PER plots indicates the presence of the link in the gray-zone. The cross-over points can be computed by measuring PER for various transmission powers.

We observe that this tipping point varies depending upon the quality of the channel between the transmitter and the receiver. Hence, generalizing the tipping point based only on the observed RSS values leads to false estimates. This indicates that a link should not only be aware of the RSS values, but also the precise nature of the channel between them. Disseminating or measuring the channel information (e.g. the PER for a large range of RSS values) is valuable when there is a high probability that link operates in this gray-zone. However, as seen in Figure 9(b), RSS value can be directly mapped to the PER or BER values for links in other zones. Hence, disseminating the RSS information is sufficient for predicting the link behavior when the link is not operating in the gray-zone. A vast majority of the links operate in high-loss or low-loss zones because the limited range of RSS which causes such behavior. Moreover, strong and relatively stable PER links are preferred by wireless nodes in WLAN and multi-hop wireless networks.

3) Can PERs be approximated as i.i.d. random variables?: In existing MAC layer modeling, PER is computed from the measured RSS values. Some studies conclude PER as a random variable from a given distribution since RSS was assumed to be from a specific distribution (e.g. Qiu *et al* assume RSS and error rate to be log-normally distributed [3]). Other research studies derive it from a constant RSS value (or an average of many measurements) (e.g. [5]). We have shown in Section II-C2 that the PER has a peculiar bi-modal distribution in gray-zones and unimodal distributions elsewhere. In this paragraph, we statistically analyze the empirical PER values in all the zones to check if they can be assumed to be i.i.d. random variables.

For the sake of simplicity, we demonstrate the distribution of the PER by choosing three types of links based on their link quality: strong links in low-loss zone, average links in gray zone and weak links in high-loss zone. Each source broadcasts saturated UDP traffic with packet size of 1460 bytes (including MAC headers). We measure PER once in every 50ms (approximately 30 packets). Since we are interested in analyzing PER distribution in one measurement period, we group the PERs in one measurement period (which is approximately 1.5 seconds as analyzed in Section II-B1) and perform distribution fitting tests on the data. The histograms of the PER were already analyzed in Section II-C2. To summarize, (i) in the low-loss zone, the PER has a unimodal histogram with very small variance. (ii) In the high-loss zone, the histogram is still unimodal, but the variance is larger. (iii) Gray-zone links have a bi-modal distribution with widely varying PER values. Experiments on different type of links (strong/weak, LoS/nLoS) links also revealed similar shapes of the histograms. Hence, it is reasonable to assume that the PER is constant for strong links, and not for average and weak links, within a measurement period.

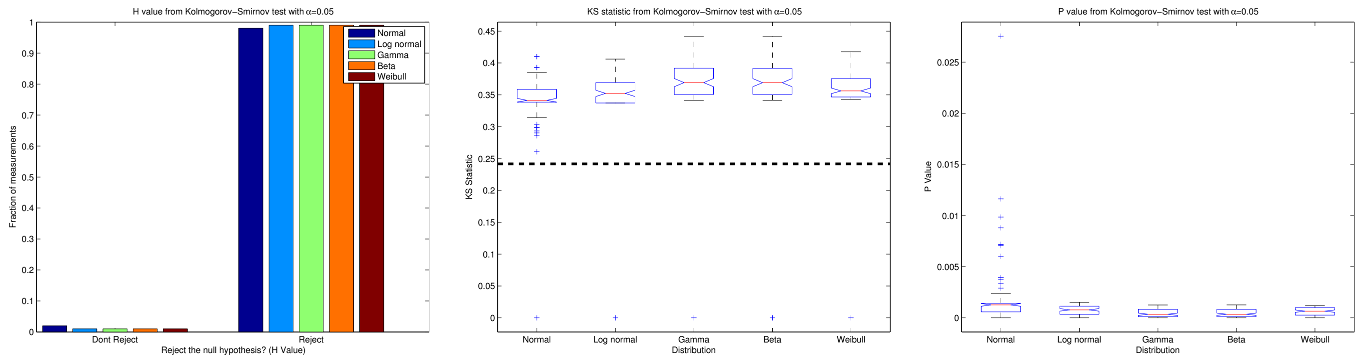
a) Distribution tests for PER: In this paragraph, we explain the results of distribution fitting tests to conclude the distribution of the PERs. Figures 11, 12 and 13 show the results of the Kolmogorov-Smirnov tests (K-S Test) for the PERs in each measurement period. The summary is represented by the H-value of the K-S Test. The results demonstrate that the strong link does not fit into any distribution (it is near constant) while the distribution of the high-loss links is well-approximated by Log-normal, Beta or Weibull distributions. As we have seen earlier, the gray-zone links are bi-modal.

b) Independence of PER: In this paragraph, we verify if the PER can be considered as independent random variables. Figure 14 shows the box-plot of the auto-correlation function of the links. It can be seen that the ACF of the low-loss link alternates with a very high probability. However, most of the PER values for strong links are near-constant with very little variance and they an be assumed as constant. The ACF of the gray-zone link, gradually decreases to droops indicating the effect of the fading of the channel over time. Hence, PER of the gray-zone links have memory and cannot be assumed as independent random variables. The ACF of the weak links drops to near zero values at the first-lag, thus showing that the PER an be assumed as independent variables.

In summary, observing the results of independence and the distribution of PER in the time-frame of seconds, we can infer that: (i) PER of strong links should be best approximated as a constant; (ii) PER of gray-zone links are unpredictable, have memory and are bi-modally distributed; and (iii) PER of the high-loss links can be approximated by i.i.d from a Log-normal, Beta or Weibull distributions.

4) Effect of packet size: Figure 15(a) shows the variation of the observed PER for different packet sizes for one representative link. It shows that the cross-over values are approximately the same for a given link with varying different packet sizes in the low-loss zone. The mean PERs are also similar in the grey zone, but the amount of variation differs for different packet sizes.

The effect of packet size on the observed error rate is one of the primary factors to access the measurement mechanism. If the observed BER (which is empirically calculated from Equation 1) for different packet sizes deviates significantly

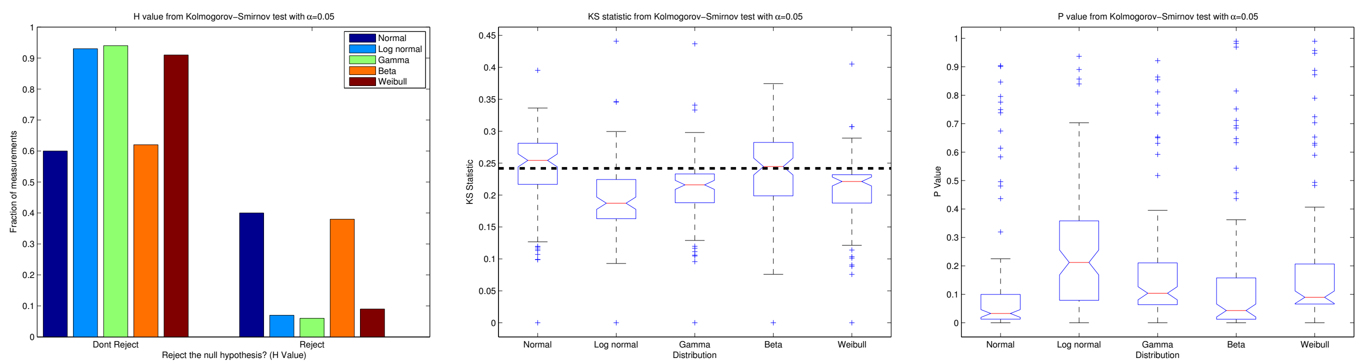


(a) K-S Test: H value

(b) K-S Test: K-S Statistic

(c) K-S Test: p value

Fig. 11. Distribution of PER: Low-loss zone



(a) K-S Test: H value

(b) K-S Test: K-S Statistic

(c) K-S Test: p value

Fig. 12. Distribution of PER: Gray-zone link **It does not make sense putting this here. We know that it is bi-modal**

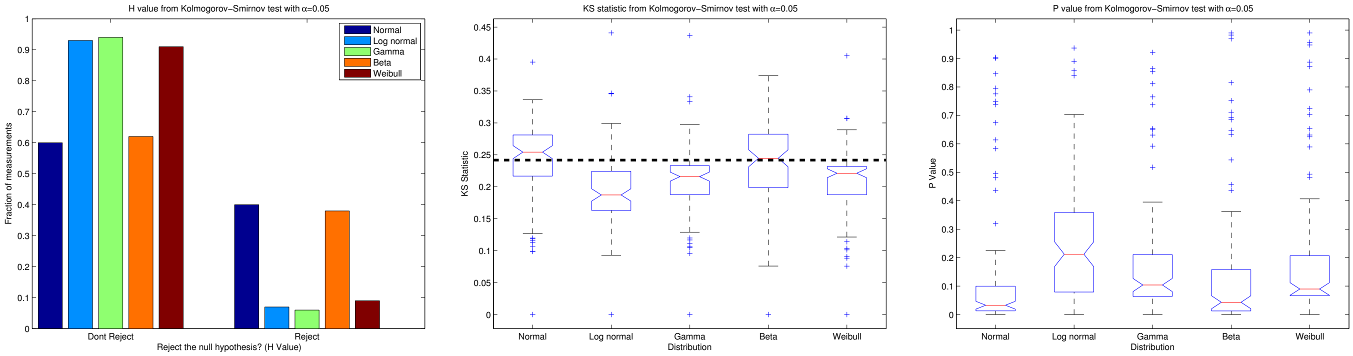
from each other, then the MBAs has to replicate the empirical PER measurement for different packet sizes to infer the BER curve for a dynamic and realistic traffic that consists of varying packet sizes. This adds an overwhelming measurement overhead, thus making the measurement procedure infeasible for realistic application. Figure 15(b) shows that a BER can be assumed as independent of packet sizes for the strong links. However, the BER curves deviate in the grey and high-loss zones.

5) Effect of modulation and rate control: In the previous sections, we analyzed the PER for different links with varying transmission power, but by fixing a standard transmission rate and modulation scheme. This was done in order to study the relationships between the RSS and the PER and the distribution of the PER. In this section, we fix the link and analyze the results the for different modulation schemes. This is helpful the analyze the effect of standard modulation and transmission rates, which are vital for the widely used rate-control modules. Each transmission rate uses a specific modulation scheme with a fixed set of parameters in 802.11. Hence, altering the transmission rate invariably alters the modulation. Measurements were carried out for all the specified transmission rates and a representative subset was chosen to illustrate the effect in a simple and uncluttered manner.

Figure 16(a) studies the effect of transmission rate (or modulation scheme) on the observed PER for a subset of 802.11 transmission rates. An interesting point to be noted is the general trend that as the transmission rate is increased (by using more complex modulation schemes), the cross-over point from low-loss zone to the gray-zone happens at lower SNR. This is counter-intuitive since stable modulation schemes like BPSK should yield lesser error rates when compared to the more advanced modulation schemes like 64-QAM. Consistent with the result observed in Figure 8, the traces indicated that there is no large variation in RSS [3] when the link is in transitional stage. Hence, we conjecture that this observation is due to the fact that higher modulation schemes transmit the packet in much smaller time than the lower modulation schemes since the transmission rates are much higher. Hence, modulation schemes like BPSK (which take approximately 2ms to transmit a 1460 byte size packet at 6Mbps) is more vulnerable to deep fading of the channel when compared to the 64-QAM (which takes approximately 0.2ms using 54Mbps).

Finally, we focus on the cross-over points that were ob-

[3]A large number of packets were received with correct headers but CRC with errors. In 802.11a, the PLCP header is always transmitted using a more stronger BPSK modulation at 6Mbps for all the values of data transmission rates. Hence, we observe a large number of erroneous packets with correct headers at higher transmission rates.

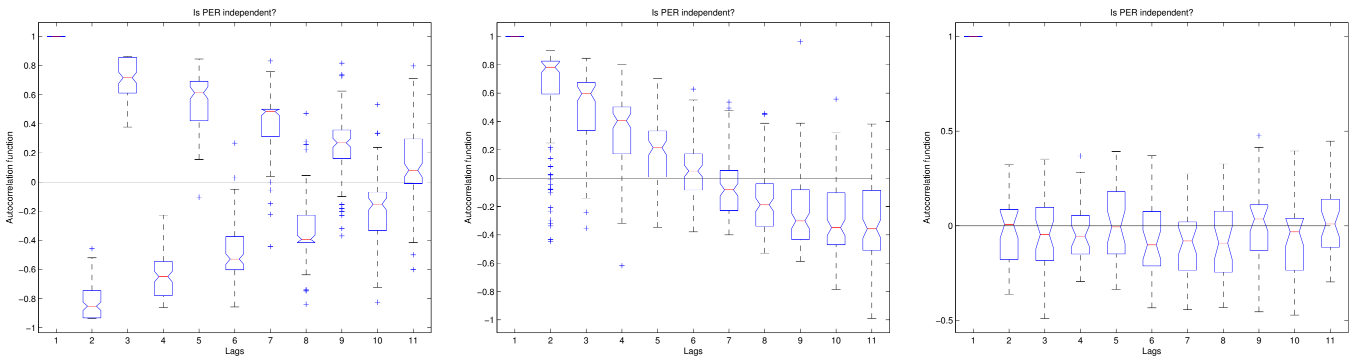


(a) K-S Test: H value

(b) K-S Test: K-S Statistic

(c) K-S Test: p value

Fig. 13. Distribution of PER: High-loss link

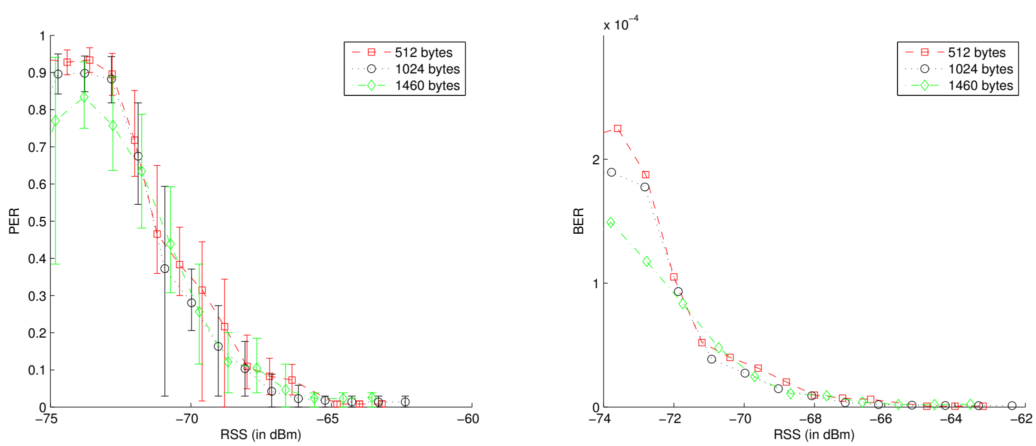


(a) Low-Loss link

(b) Gray-zone link

(c) High-loss link

Fig. 14. Independence of PER



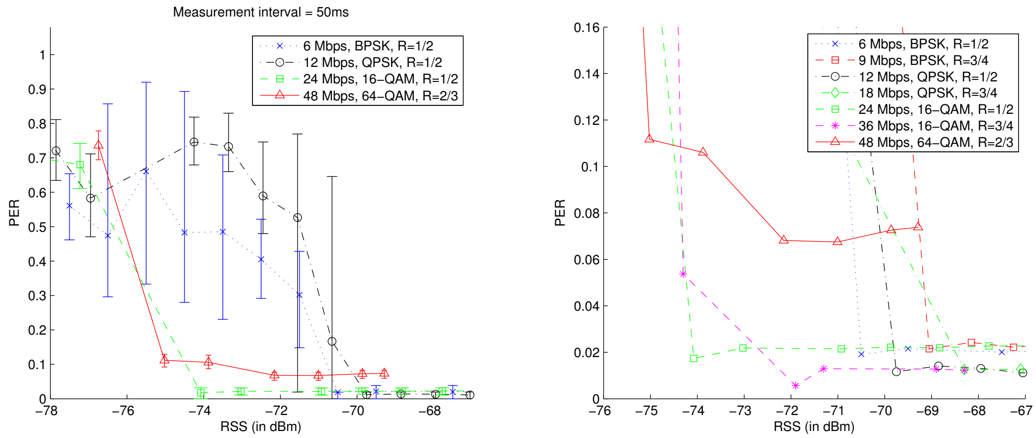
(a) PERs for different packet sizes

(b) BER study

Fig. 15. Effect of packet size on error rates

Tx-Rate (in Mbps)	6	9	12	18	24	36	48
Cross-over (in dBm)	-69.5	-69.1	-70.1	-68.2	-74.1	-73.6	-74.8

TABLE I
CROSS-OVER POINTS FOR DIFFERENT TRANSMISSION RATES



(a) PERs for different modulation schemes

(b) Cross-over values

Fig. 16. PER and rate-control

served on our simulations for different modulation schemes. Figure 16(b) focusses on the low-loss zone and cross-over points. This also points to the trend that higher transmission rates with complex modulation schemes generally have cross-over points at a lesser RSS than the lower transmission rates with simple and robust modulation. Table I shows the cross-over points that were measured for a link in our testbed. The cross-over points for 48Mbps and 6Mbps is separated by over 5 dB.

An important inference from the result observed in this section is that rate-control modules have to be sensitive to the loss zone of the link. If the link is in the gray-zone region, which can be inferred from bi-modal PER due to the effect of fading, then it is reasonable to attempt using a higher transmission rate to reduce the probability of deep fade. This will not only result in lower and stable PER, but also lead to a drastic improvement in throughput because of the increase in transmission rate. This property makes it effective for both real-time applications with smaller packet size and data traffic with larger packet size. As far as our knowledge, none of the existing rate-control modules use this inference. Majority of the rate-control modules switch to a lower rate when the greater packet errors are observed.

III. Related Work

This paper presented an empirical analysis of the RSS and error rates by focusing on the requirements for the MBAs deployed in a IEEE 802.11 based networks that are susceptible to fading. Several studies have analyzed the data gathered from such networks. For the sake of clarity, we divide this section by studying the related work which analyzes distribution and the temporal stability of RSS and PER and the effect of fading on them.

A. Temporal stability

Fixing the order of time to analyze the temporal stability is obviously a vital step before capturing the temporal stability. In this aspect, very few research work has focussed on quantifying the time range that is reasonable for a MBA to assume that the measured value is representative. Wide-ranging conclusions has been drawn by studying the experimental traces of RSS under different environments. Some studies conclude that RSS fluctuates rapidly over fraction of seconds [18]. Other studies [4], [19], [20] concur with out observation that the RSS values remains representative in the order of seconds which confirms our observation. Research studies on the temporal stability of the error rates are also inconclusive. Unlike the above papers, we measure the stability of RSS in more detail: we show the variation of the RSS for different categories of links and map the stability of the RSS with respect to the empirical error rate and the zone of the link. We believe that this is of direct use in the design of MBAs.

B. Inference of statistical distribution

A discord about the distribution of the RSS and PER is prevalent seen in the literature. Shrivastava *et al.* [19] observe that the distribution of RSS has a very large variation. Qiu *et al.* [3] observe a log-normal distribution of the RSS and use it to model the link throughput. Reis *et al.* [4] and Srinivasan *et al.* [20] observe very low variation of the RSS. We conjecture that flat distribution [19] of RSS in static networks are due to external interference or unexpected design of Atheros chipset to enable transmit diversity [14] and not due to the propagation effects. In addition to comparing the RSS to several distributions and performing tests to check the independence, our contribution in this paper unifies the other conflicting conclusions by observing that the distribution of RSS can follow constant or log-normal behavior based on the operating zone of the link. We also study the statistical tests to infer the independence and distribution of the PER which is assumed as constant [10], as a log-normal or normal i.i.d. random variable [3], [12].

C. Effect of fading and grey zone

Existing studies [4], [16] have observed a specific pattern where errors rates jump between high, low and a transitional zone. However, detailed effect of fading and the bimodal distribution of PER in this region is observed by Zuniga *et al.* [16] and Awoniyi *et al.* [17]. However, both these studies do not show the effect of different modulation schemes on the observed PER which is vital for MBAs like rate-control modules. Moreover, the study in [17] is simulation based and Zuniga *et al.* [16] measure using sensor motes which does not use IEEE 802.11 protocol.

In summary, the paper contributes to the existing measurement based research by analyze the statistical properties of RSS and error rate for (1) different links where channel state might be different; (2) different transmission powers over a single link which isolates the effect of widely varying channel; and (3) Effect of transmission rate which is not.

IV. Discussion

A. Analysis of RSS at smaller time-scale

In this section, we verify the statistical properties of the received signal strength at a microsecond granularity. The spectrum analyze is configured to scan the channel at a very low time granularity (once every 10 microseconds). The observed RSS values are analyzed for two main properties: (1) Independence; and (2) Distribution.

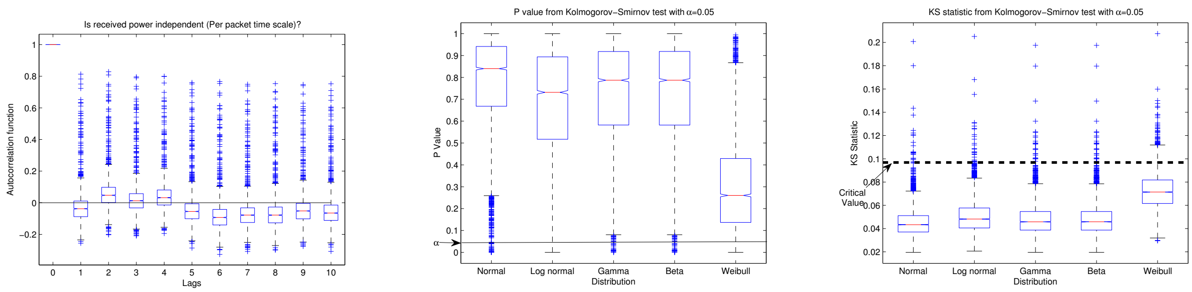
a) Independence of RSS values: The measured RSS values for each packet (around 200 consecutive measurement points) is first verified for independence by plotting the *Autocorrelation function (ACF)*. ACF is a statistical metric for checking if there is a repeating pattern in the measured values, thus indicating the non-randomness in the data. ACF is defined at various positive integer points n (called as lags) and is a value between $[0, 1]$ at each lag. ACF at lag n denotes the correlation between the measurement point at time t and $t+n$. A perfect random variable should have a ACF of 1 for lag 0 and ACF of 0 at all the other lags. Figure 17(a) shows the box-plot of the ACF for the RSS values measured for the packets when the source and the spectrum analyzer were separated by around 2m distance. The box-plot summarizes groups of data (e.g. ACF for each packet transmission in Figure 17(a)) by a box (that bounds the upper and lower quartiles of the data), a median (a horizontal line) and the outliers (denoted by '+' marks). It can be seen that a vast majority of the ACF drops to almost 0 at lag 1, thus indicating the randomness of the received signal strength. Similar ACF values are also seen when the distance between source and the spectrum analyzer is altered, thus indicating that RSS can be assumed to be a random variable at a microsecond time granularity.

b) Distribution estimation: After inferring that the RSS values can be assumed as independent random variables, we now estimate the distribution of the RSS at a small time scale. We first compared the observed RSS values with more than 40 classical distributions and inferred that the family of distributions that denote bell-shaped curves. The skewness was also observed to be very low. We then carried out an extensive comparison of measured RSS with four standard distributions that were indicated as good fits: Normal, Log-normal, Gamma and Weibull distribution. We use the Kolmogorov-Smirnov test (K-S test) to estimate the best-fitting distribution to the measured RSS values. K-S test performs a goodness-of-fit test for the empirical distribution of the measured value (RSS, in our case) with the cumulative distribution for a specified distribution. The K-S test is performed for a given significance level (α) and will output four metrics based on which we either reject or do not reject the hypothesis that the measured values belong to a given distribution. Figures **??** and **??** shows the box-plot of the *p-value* and *K-S statistic* of the measured RSS values. The hypothesis that the observed empirical data belongs to the compared distribution is rejected if the *p-value* is lesser than α and the K-S statistic value is greater than the *critical value* (that is taken from a standard table). From the figures, we can see that the observed RSS values are a good-fit for the chosen distributions and the Normal distribution best fits the RSS values. The Shapiro-Wilk test also provided strong positive result that ascertains that data fits Normal distribution. We do not describe those tests due to the lack of space.

As we show in Section **??**, another important quantity that is required for the derivation of *Bit-error rate* (BER) is the type of modulation used. Theoretically, BER is approximated as a closed form function of received signal strength and noise for different modulation scheme. However, these functions are complex and it is analytically intractable to derive the distribution of the BER from approximating the RSS as a random variable. Hence, we numerically estimate the distribution of the BER from the measured RSS.

c) RSS: Constant or Random Variable: The data from the spectrum analyzer shows that RSS can be approximated as a normally distributed independent random variable. However, the lowest granularity of the received signal strength that can be obtained from the realistic wireless cards is a single value of the *Received Signal Strength Indicator (RSSI)* per packet. This value is measured when the preamble of the packet is received. RSSI can be converted to an actual received signal strength (RSS) through standard conversion techniques [4]. We now evaluate if per-packet RSS information is sufficient to represent the fluctuations of RSS over packet transmission time. Specifically, we empirically analyze the answers to the following questions: (1) Does the RSS measured during the preamble be a good indicator of the mean RSS during the packet transmission?; (2) The effect of approximating the RSS as constant.

Previous paragraph argued that the RSS is well-approximated as a random normally distributed variable. While the per-packet RSS can be taken as the expected value of this random variable, the complete distribution cannot be inferred since other critical parameters like standard deviation are not reported. In order to have an accurate and realistic model that can be used in standard wireless nodes, we evaluated the dependency of the mean and standard deviations from the distribution fitting tests that we conducted.



(a) Auto-correlation of RSS (b) K-S test for RSS: p-value (c) K-S test for RSS: K-S Statistic

Fig. 17. Broadcast traffic analyzed at smaller time scale

Thus, we have used the detailed measurement using the spectrum analyzer to derive a realistic model based on the data obtained from the commercial wireless cards.

However, spectrum analyzer results cannot be used for long-term RSS analysis since:

- Analysis of long-term data using spectrum analyzer produces inconsistent data since the spectrum analyzer will pause for a long duration (around 200 ms) after each *sweep cycle* (approximately 20 ms in our setting) to log the collected data. Increasing the sweep cycle time is not feasible since it reduces the accuracy of the signal measurement.
- Protocols and analytical models that run on the node will have direct access to the RSS obtained from the Atheros cards.

B. Anomalies on Atheros chipsets

In this paragraph, we measure the RSS of a saturated broadcast transmission. The measurements reveals an unexpected and interesting effect. As shown in Figure 18(a), the broadcasted packets toggle between two significantly differing power-levels. The effect of such alternating power-levels has a drastic influence on the performance of the network since the difference in alternating RSS is very high (with a difference of around 6 dBm). Since the Atheros based card provides the Linux kernel with per-packet RSSI values, the receiving node is able to monitor the alternating effect of RSS (as observed in Figure **??**).

This effect of alternating RSS was also recently reported by Giustiniano *et al.* [14] and is attributed to the antenna transmit diversity. The authors concluded that the firmware switches transmitting each packet using a different antenna when it is not able to since it is unable to infer any feedback on the transmitted broadcast packets. success, the card switches to transmit each packet using a different antenna. This paper extends this observation by: (1) producing a more detailed spectrum-analyzer traces; (2) statistical analysis of the effect of such alternation; and (3) throughput modeling and analysis of its effects on network performance under such default scheme.

Transmit diversity anomaly in Atheros cards

Mean RSS (MRSS) values are computed for each packet from the data obtained from the spectrum analyzer and the variation of these are observed. We first infer if the MRSS values are independent. Figure 19(a) shows the ACF of the MRSS values. It is clearly seen that the consecutive MRSS values are negatively correlated, thus indicating that MRSS are not independent and it alternates (as observed in Figure 18(a)). Figure 19(b) shows the histogram of the observed RSS values. The bimodal distribution of RSS with two peaks separated by around 8 dBm can be clearly seen; a sharper and better MRSS that corresponds to the "stronger" antenna diversity transmission scheme and a more flat and weaker MRSS that corresponds to the packet reception from the "weaker" scheme.

References

[1] Y. Gao, D.-M. Chiu, and J. C. S. Lui, "The fundamental role of hop distance in ieee802.11 multi-hop ad hoc networks," in *ICNP*, 2005.

[2] D. Aguayo, J. Bicket, S. Biswas, G. Judd, and R. Morris, "Link-level measurements from an 802.11b mesh network," in *SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*. New York, NY, USA: ACM Press, 2004, pp. 121–132.

[3] L. Qiu, Y. Zhang, F. Wang, M. K. Han, and R. Mahajan, "A general model of wireless interference," in *MobiCom '07: Proceedings of the 13th annual ACM international conference on Mobile computing and networking*. New York, NY, USA: ACM, 2007, pp. 171–182.

[4] C. Reis, R. Mahajan, M. Rodrig, D. Wetherall, and J. Zahorjan, "Measurement-based models of delivery and interference in static wireless networks," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 4, pp. 51–62, 2006.

[5] A. Kashyap, S. Ganguly, and S. R. Das, "A measurement-based approach to modeling link capacity in 802.11-based wireless networks," in *MobiCom '07: Proceedings of the 13th annual ACM international conference on Mobile computing and networking*. New York, NY, USA: ACM, 2007, pp. 242–253.

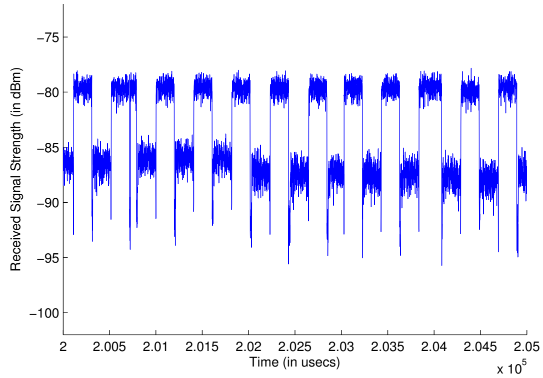
[6] S. H. Y. Wong, H. Yang, S. Lu, and V. Bharghavan, "Robust rate adaptation for 802.11 wireless networks," in *MobiCom '06: Proceedings of the 12th annual international conference on Mobile computing and networking*. New York, NY, USA: ACM, 2006, pp. 146–157.

[7] R. Draves, J. Padhye, and B. Zill, "Comparison of routing metrics for static multi-hop wireless networks," in *SIGCOMM*, 2004.

[8] G. Pollini, "Trends in handover design," *Communications Magazine, IEEE*, vol. 34, no. 3, pp. 82–90, Mar 1996.

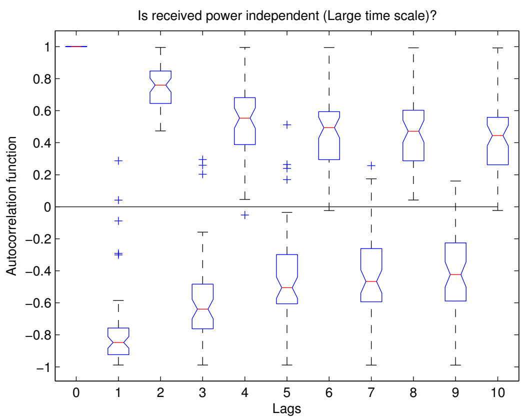
[9] J. Pavon and S. Choi, "Link adaptation strategy for ieee 802.11 wlan via received signal strength measurement," *Communications, 2003. ICC '03. IEEE International Conference on*, vol. 2, pp. 1108–1113 vol.2, 2003.

[10] D. Lal, A. Manjeshwar, F. Herrmann, E. Uysal-Biyikoglu, and A. Keshavarzian, "Measurement and characterization of link quality metrics in energy constrained wireless sensor networks," *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, vol. 1, pp. 446–452 Vol.1, Dec. 2003.

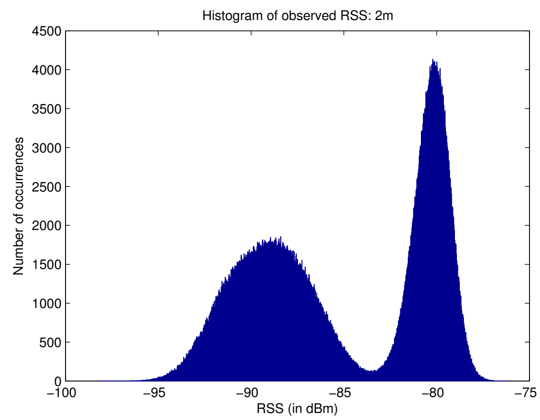


(a) RSS (2m apart)

Fig. 18. RSS graph for broadcast traffic **Put the transmitter RSS graph here**



(a) Auto-correlation of RSS



(b) Histogram of RSS

Fig. 19. Broadcast traffic analyzed at larger time scale

[11] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, "Impact of interference on multi-hop wireless network performance," in *MobiCom*, 2003.

[12] A. Woo, T. Tong, and D. Culler, "Taming the underlying challenges of reliable multihop routing in sensor networks," in *SenSys '03: Proceedings of the 1st international conference on Embedded networked sensor systems*. New York, NY, USA: ACM Press, 2003, pp. 14–27.

[13] "Madwifi: Linux driver for atheros cards," http://madwifi.org/.

[14] D. Giustiniano, I. Tinnirello, L. Scalia, and A. Levanti, "Revealing transmit diversity mechanisms and their side-effects in commercial ieee 802.11 cards," *Telecommunication Networking Workshop on QoS in Multiservice IP Networks, 2008. IT-NEWS 2008. 4th International*, pp. 135–141, Feb. 2008.

[15] V. Kolar and N. B. Abu-Ghazaleh, "A Multi-Commodity Flow Approach to Globally Aware Routing in Multi-Hop Wireless Networks," in *IEEE Pervasive Computing and Communications (PerCom)*, 2006.

[16] M. Zuniga and B. Krishnamachari, "Analyzing the transitional region in low power wireless links," *Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004. 2004 First Annual IEEE Communications Society Conference on*, pp. 517–526, Oct. 2004.

[17] O. Awoniyi and F. Tobagi, "Effect of fading on the performance of voip in ieee 802.11 a wlans," *Communications, 2004 IEEE International Conference on*, vol. 6, pp. 3712–3717 Vol.6, June 2004.

[18] A. Keshavarzian, E. Uysal-Biyikoglu, D. Lal, and K. Chintalapudi, "From experience with indoor wireless networks: A link quality metric that captures channel memory," *Communications Letters, IEEE*, vol. 11, no. 9, pp. 729–731, September 2007.

[19] V. Shrivastava, D. Agrawal, A. Mishra, S. Banerjee, and T. Nadeem, "Understanding the limitations of transmit power control for indoor wlans," in *IMC '07: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*. New York, NY, USA: ACM, 2007, pp. 351–364.

[20] K. Srinivasan and P. Levis, "RSSI is Under Appreciated," *Third Workshop on Embedded Networked Sensors (EmNets 2006)*, May 2006.