H-RCA: 802.11 Collision-aware Rate Control

K. D. Huang, K. R. Duffy and D. Malone

***Abstract*— Rate control methodologies that are currently available in 802.11 network cards seriously under-utilize network resources and, in addition, per-second throughputs suffer from high variability. In this article we introduce an algorithm, H-RCA, that overcomes these shortcomings, giving substantially higher, and less variable, throughput. The approach solely uses information already available at the driver-level to function and can be implemented on 802.11e commodity hardware.**

H-RCA's design objective is to minimize the average time each packet spends on the medium (including retries) in order to maximize total network throughput. It uses a development of a recently proposed estimation scheme to distinguish transmission failures due to collisions from those caused by channel noise. It employs an estimate of the packet loss ratio due to noise in assessing whether it is appropriate to change rate. We demonstrate experimentally that packet loss ratio is not necessarily a monotonic increasing function of rate; this is accounted for in H-RCA's design.

As H-RCA statistically separates noise losses from those caused by collisions, ns-2 simulations show that it is robust to changing environments. H-RCA does not require specific hardware support nor any change to the IEEE 802.11 protocol. This point is substantiated with results from an experimental implementation.

***Index Terms*— Rate Control, Collision-aware, WLAN, IEEE 802.11, TXOP.**

I. INTRODUCTION

IEEE 802.11 is the world's most commonly deployed WLAN technology. It supports several physical layer transmission rates, with 802.11b having four ($1, 2, 5.5$ and 11 Mb/s), 802.11a having eight ($6, 9, 12, 18, 24, 36, 48$ and 54 Mb/s), while 802.11g has all twelve of the 802.11b and 802.11a rates, and 802.11n has eight when using two streams at 20 Mhz and a guard interval of 800ns ($13, 26, 39, 52, 78, 104, 117, 130$ Mb/s). A range of rates are available as their modulation and coding schemes give them distinct robustness characteristics to noise on the medium. To maximize network performance, each station needs to select an appropriate rate for its current channel conditions. Rate Control (RC) algorithms that choose modulation and coding scheme pairs are designed for this purpose.

The 802.11 protocol has been the subject of extensive research since the mid 1990s. Despite this, the RC algorithms that are currently implemented in hardware can be wasteful of air-time resources, particularly in the presence of collision-based losses, leading to poor network performance [1]. For example, using the experimental apparatus described in Appendix I, Fig. 1 provides a representative illustration of these shortcomings. In a network of 5 stations that always have packets to send and are availing of the 802.11a rate-set, it plots per-second total network throughput with five minute mean reported in the legend. In each experiment, all stations are utilizing one of Minstrel [2], SampleRate, [3], AMRR [4] or Onoe [5] as implemented in the MadWiFi driver for the Atheros chipset or RRAA [6] as implemented by us. Also plotted is the same scenario but with the stations using the paradigm introduced in the present article, H-RCA, demonstrating the gains in utilization that are possible with it.

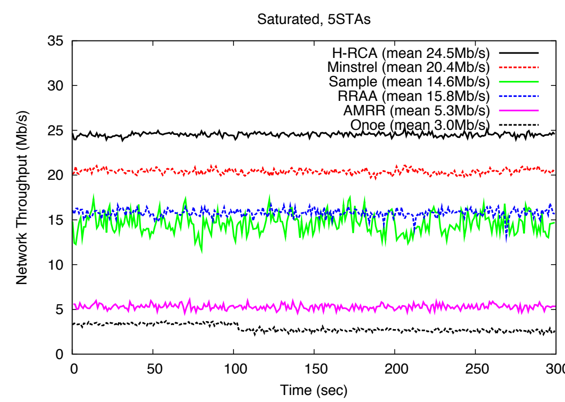

Fig. 1. WLAN consisting of 5 stations that always have 1,000 B packets to send using the 802.11a rate-set and a minimum contention window of 16. Throughputs for five existing rate control algorithms, Minstrel, SampleRate, RRAA, AMRR and Onoe, as well as the methodology proposed in this article, H-RCA. Experimental data

In this article we propose a principled design for a RC algorithm, H-RCA, that is applicable to all 802.11 rate-sets and is implementable on commodity hardware that supports 802.11e functionality. H-RCA's objective is to minimize the average time each packet spends on the medium, including MAC layer retries, in a fully decentralized fashion with no message passing. It employs a development of a recently proposed censored data technique based on the IEEE 802.11e TXOP[1] feature to distinguish transmission failures caused by collisions from those caused by noise [7]. H-RCA makes transmission rate choices based on the Packet Loss Ratio (PLR) due to noise alone[2], with Bayesian analysis used to determine rate-decrease decisions and an opportunity-cost metric used to determine the frequency at which rate-increase decisions are made. H-RCA does not alter the 802.11 MAC and can be implemented on existing network cards that possess 802.11e functionality. We give a concrete guide to the approach through detailed consideration of the 802.11a rate-set, including performance

Hamilton Institute, NUI Maynooth. SFI grant RFP-07-ENEF530 and The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7-ICT-2009-5) under grant agreement n. 257263 (FLAVIA project).

[1]If 802.11e is not supported, this technique can use fragmentation in lieu of TXOP, but at the expense of increased overhead.

[2]We reserve PLR for failures due to noise, not collision.

evaluation in simulation as well as initial results from an experimental implementation.

The rest of this paper is organized as follows. In Section II we describe related work. The H-RCA paradigm is defined in Section III and the reasoning behind its settings explained. During our worked example in Section III-A, through the use of experiments, we demonstrate that for a fixed signal to noise ratio (SNR) the robustness of 802.11a rates to channel noise can be, surprisingly, not a monotonic decreasing function of increasing rate. These experiments support the theoretical prediction [8][9] that the 802.11a 9 Mb/s rate is redundant. As it is not feasible to create an experimental setup with controllable losses due to noise, Section IV presents ns-2 simulation results illustrating H-RCA's performance for an 802.11a WLAN. These simulations reveal the merits of the approach in terms of throughput consistency in a fixed environment, adaptivity to changing channel conditions and robustness to collision-based transmission failures. As experimental systems can expose difficulties not captured by theory or simulations, in Section V we report on an experimental implementation of H-RCA for the 802.11a rate-set where it shows significant throughput gains over RC algorithms that are available with current hardware. The paper concludes with a discussion in Section VI.

II. Related Work

The IEEE 802.11 standard does not specify details of the RC algorithm to be used, so that 802.11 card vendors and researchers have proposed and implemented a variety of algorithms. There are two distinct strategies for RC algorithms. The first is the explorative type. In this approach the entire rate space is explored periodically to empirically identify the optimal rate. Examples of algorithms of this type include SampleRate [3], RBAR [10], OAR [11], WOOF [12], CHARM [13] and Smart-rate [14]. The second strategy is the incremental type where algorithms record statistics regarding their current rate and its neighboring rates, and make incremental changes. Example RC algorithms of this type include ARF [15], AARF [4], CARA [16], RRAA [6], SGRA [17], COLLIE [18], SoftRate [19] and AccuRate [20]. A number of these, including SampleRate, Minstrel [2], Onoe [5] and AMRR [4], are implemented, for example, on the Atheros chipset.

The choice of rate should be based on current channel conditions, so that a good estimate of channel quality is key to all RC algorithms. There are two dominant paradigms to measure channel conditions: determine SNR directly from physical layer estimates; or estimate channel conditions indirectly through packet loss information.

A. Physical Layer Based Estimates

The ideal information on which to base the choice of the transmission rate is SNR at the receiver. RBAR [10] uses an RTS/CTS exchange immediately prior to packet transmission to estimate SNR at the receiver and picks its rate accordingly. OAR [11] builds on RBAR and opportunistically transmits back-to-back frames using a fragmentation scheme when channel quality is good. CHARM [13] leverages reciprocity of the wireless channel to estimate average SNR at the receiver using packets overheard from the receiver to avoid the RTS/CTS overhead. It is not trivial, however, for the transmitter to accurately estimate the SNR at the receiver because signal strength exhibits significant variations on a per-packet basis [17].

B. Packet-Loss Based Estimates

Using packet loss information to infer channel conditions is the second option. Automatic Rate Fallback (ARF) is a scheme that uses patterns of packet losses as a trigger to change the transmission rate [15]. Adaptive ARF (AARF) continuously changes the threshold that decides when to try a higher rate to better reflect current channel conditions [4]. Adaptive Multi Rate Retry (AMRR) is AARF's practical realization. Its key idea is to use binary-exponential-backoff to control the probing period to sample other rates. The algorithm initially switches to a higher rate when ten consecutive packets have been transmitted without any failure and moves to a lower rate if two consecutive packets are not acknowledged. If packet failure occurs when a higher data rate is sampled, the interval for the next higher data rate sampling is exponentially increased. For these approaches to function correctly in a network with more than one active transmitter, the algorithm would need a mechanism to distinguish between transmission failures caused by collision and those caused by noise on the channel.

Some algorithms use the RTS/CTS scheme to identify failed transmissions due to collisions. If the first attempted transmission of a packet fails, CARA [16] uses RTS/CTS to test whether failure is caused by collision or noise. Since RTS/CTS costs substantial time on the medium, RRAA [6] reduces the frequency of using RTS/CTS. RRAA uses frame loss information gathered over tens of frames to adapt the rate and compares frame loss statistics both with and without RTS/CTS in order to decide if a loss is caused by collision or noise. It adaptively enables RTS/CTS more frequently as its estimate of the rate of failures due to collisions increases.

To avoid using RTS/CTS, WOOF [12] uses Channel Busy Time (CBT) as an indicator of network load. Higher CBT means heavier traffic in the network, so that a transmission failure is more likely to be caused by collision rather than noise. Running at the sender, COLLIE [18] analyzes the patterns of bit errors in the received packet in order to infer whether an error was due to a collision or the channel noise. Its rate adaptation protocol then solely depends upon channel noise failures. To detect bit errors, however, the COLLIE receiver must echo the entire received frame to the sender, incurring significant overhead. Running at the receiver, SoftRate [19] uses hints exported by the physical layer to compute the average Bit Error Rate (BER) for each received frame. To exclude failures due to collisions, it uses the ansatz that a sudden spike in bit errors is likely to have been caused by collision. The receiver sends its BER estimate to the sender where it picks the best rate for the next frame. In order to observe the impact of collisions more clearly, it also adds a 'post-amble' to every frame to enable the receiver to detect with high probability the portion of the sender's frame that lasts after the interference has ended. Thus SoftRate uses more time per frame on the medium and to implement it would require changes to hardware.

To avoid hardware modifications, in [21] the authors use the number of idle slots between two consecutive busy periods to estimate the number of active stations in the network. Having obtained an estimate of number of the active stations, using a Bianchi-like model [22], the collision probability conditioned on transmission can be estimated. Armed with this estimate one can estimate the packet loss probability due to noise. The information on the number of idle slots is not available, however, in general commodity cards. Therefore, in [23] the authors overcome this difficulty by using retry information in 802.11 MAC headers as an indicator of the channel condition, as each station can monitor the medium and get access to the MAC header of every packet transmitted in the medium. If a large number of retransmitted packets are observed, it is inferred that the network is handling a high traffic load.

SampleRate [3] adopts a different strategy. It focuses on minimizing the service time required for successful transmission of a packet. SampleRate uses frequent probing of different transmission rates to calculate the Expected Transmission Count (ETC) for each rate. The ETC represents the average number of transmission attempts required for successful reception of a packet. The expected transmission time (ETT) is calculated using ETC information at a given transmission rate and accounts for the back-off time when the ETC metric predicts that a retransmission is required. SampleRate then decides to transmit data packets using the rate with the lowest expected transmission time. Since SampleRate does not distinguish between transmission failures caused by collisions and those caused by noise, it may make erroneous rate selection choices in the presence of collisions.

III. H-RCA

An outline of the H-RCA methodology is as follows.

A. Given a rate-set $\{r_1, \ldots, r_K\}$ Mb/s sorted in increasing order, $r_i < r_{i+1}$ for all $i \in \{1, \ldots, K-1\}$ (e.g. for 802.11a $\{6, 9, \ldots, 54\}$), use theory and experiment to identify rates r_i such that the PLR in given channel conditions at rate r_i is higher than the PLR for a higher rate r_j. These rates are excluded from H-RCA's rate-set.
B. To estimate PLR, H-RCA uses a technique based on TXOP to gain observations of packets solely susceptible to loss through channel noise.
C. Use theory or experiment to determine for each rate a critical PLR value, the rate lowering threshold, above which a lower rate would give higher throughput.
D. Use Bayesian inference to determine if the PLR of the current rate is above a rate lowering threshold.
E. Set rate increase frequency so that the opportunity-cost of sampling a higher rate is, in the worst-case, less than, say, 5%.

The H-RCA approach is to first evaluate the rate-set with theory and experiment to determine if increasing rate necessarily leads to a deterioration in PLR at each fixed level of channel noise. This process identifies problematic rates for incremental RC schemes. Theoretical predictions, for example, suggest that the 9 Mb/s rate is redundant in 802.11a as the higher 12 Mb/s rate always possesses greater robustness to noise [8][9] and we provide experimental evidence in support of this finding. As we consistently find this redundancy, including in experiments beyond those reported here, our approach is to exclude 9 Mb/s. At a minimum, these results clearly necessitate that any incremental RC algorithm must sample both the 9 Mb/s and 12 Mb/s rates when investigating rate increase decisions from 6 Mb/s or risk under-performance.

RC algorithms need to make two decisions: when to increase the rate and when to decrease it. Rate-increase decisions are necessarily exploratory as channel performance at the higher rate must be determined from new observations. Assuming lower rates are more robust, rate-decrease decisions can be made based on channel observations at the current rate. Thus these two decision making processes are distinct in nature. We use an opportunity-cost paradigm to dictate the frequency of rate-increase decisions. For rate-decrease decisions, we employ a recently proposed censored data technique [7] to separate losses due to collisions from those due to noise, followed by Bayesian decision making based on the resulting statistics. The resulting algorithm makes rate change decisions based on pre-calculated thresholds that can be stored in a lookup table. These comparisons are executed by the host CPU and not by the wireless card's firmware, so the algorithm's computational burden is, effectively, insignificant.

Use of the censored data technique based on TXOP developed in [7] enables us to overcome a serious challenge common to all algorithms: the base hardware cannot distinguish transmission failures that occur due to collisions from those that occur due to noise on the medium. This is important as if the rate of transmission failure increases there are two potential explanations, each of which would dictate distinct corrective action. If the channel is experiencing increased noise, transmission failures will result and the station should change to a lower, more robust rate. If, however, more stations become active, there will be an increase in transmission failure due to collisions. In this case, the station should not select a lower rate, as to do so would increase the time its packets spend on the medium, leading to increased congestion.

The primary goal of H-RCA is to maximize throughput of the whole network in a decentralized way by minimizing the average time each packet spends on the medium. Each station aims to choose a rate that minimizes the air time that its packets spend on the medium, including retries. Alternative objectives, such as each station selfishly maximizing its own throughput, are discussed in Section VI.

To fully illustrate the H-RCA methodology, throughout the exposition we use the 802.11a rate-set as a working example. Namely, after the explanation of each H-RCA design step, we have a brief section specifically related to 802.11a. The 802.11a parameters are summarized in Table I. Formulae are presented for H-RCA's parameterization. In practice these values could be determined dynamically or statically. For the purposes of this paper we use the latter, simpler scheme.

TABLE I
802.11A PARAMETERS

Parameters	Default values
Minimum Contention Window	16
Maximum Contention Window	1024
Long Retry Limit	4
Short Retry Limit	7
Slot Time	9 μs
SIFS Time	16 μs
DIFS Time	34 μs
Header Time	20 μs

A. Rate-Set Characteristics

When employing an incremental algorithm RC approach, such as with H-RCA, it is necessary to determine *a priori* the relative robustness of rates in the available rate-set. For the 802.11a and 802.11n rate-sets, it is possible to use theory, simulation and experiments in this investigation. The 802.11b/g 5.5 and 11 Mb/s rates employ a Complementary Code Keying modulation scheme. This scheme has proven resistent to analytic study so that no theoretical framework for their performance analysis currently exists. Using pseudo-theory with manufacturer-fit curves and detailed experiments, elsewhere we have shown that the 11 Mb/s rate in 802.11g is more robust that the 6 Mb/s rate [24].

1) Theoretical Prediction (802.11a): For this part of the methodology, it is, perhaps, easiest to explain the procedure by example. We first theoretically determine PLR as a function of SNR. Table II summarizes the modulation and coding information for each rate supported by 802.11a. For RC algorithms, the most significant feature of these rates is their robustness to noise at a given SNR.

With a fixed transmission rate and a fixed SNR, the PLR is the probability a transmission fails in a communication between one transmitter and one receiver in the absence of interference from other stations. For a theoretical calculation of PLR as a function of SNR, we begin by using the well-known relationships between BER and SNR that have been derived for different modulation schemes (BPSK, QPSK and QAM) [25][26]. The 802.11a rates also employ a convolutional code to provide Forward Error Correction (FEC) so that decoding errors also need to be taken into account. As recommended in the IEEE 802.11 standard, we assume a maximum likelihood hard-decision decoding scheme is used.

Our theoretical calculation of PLR, which depends on the channel model, follows from established reasoning [26] so we provide a guide rather than extensive details. We report on three channel models: Rayleigh fading; Rician fading; and Additive White Gaussian Noise (AWGN). The Rayleigh fading channel mirrors the situation where there is no significant propagation along a line of sight between transmitter and receiver [27]. Rician fading mimics urban environments where there is partial cancellation of the radio signal by itself due to multi-path interference [28]. The AWGN model is considered appropriate when there is line of sight between the transmitter and receiver, but no multi-path, no terrain-blocking and no interference [26]. With packets containing 1,000 B payloads, Fig. 2 shows the theoretical prediction of PLR vs. SNR for different transmission rates in the Rayleigh fading channel. Fig. 3 shows the equivalent plot for the Rician fading channel. Predictions using the Rayleigh fading channel model reveal two redundant rates, 9 Mb/s and 18 Mb/s, as the former has a higher PLR than the 12 Mb/s rate at every SNR while the latter has higher PLR than 24 Mb/s at every SNR. For the Rician channel, only the 9 Mb/s rate appears redundant. Plots for the AWGN channel can be found elsewhere [8][9][24] and mimic those shown here for Rician fading. Consequently, all of these channel models suggest redundancy of the 9 Mb/s rate. The 18 Mb/s rate is only redundant in the Rayleigh channel model. This suggests that adaptive RC algorithms should take care at rate increase/decrease decisions if 18 Mb/s performs poorly, as it is possible, but not certain, that 24 Mb/s will perform better.

TABLE II
802.11A TRANSMISSION RATES

Rate (Mb/s)	Modulation Scheme	FEC Rate
6	BPSK	1/2
9	BPSK	3/4
12	QPSK	1/2
18	QPSK	3/4
24	16QAM	1/2
36	16QAM	3/4
48	64QAM	2/3
54	64QAM	3/4

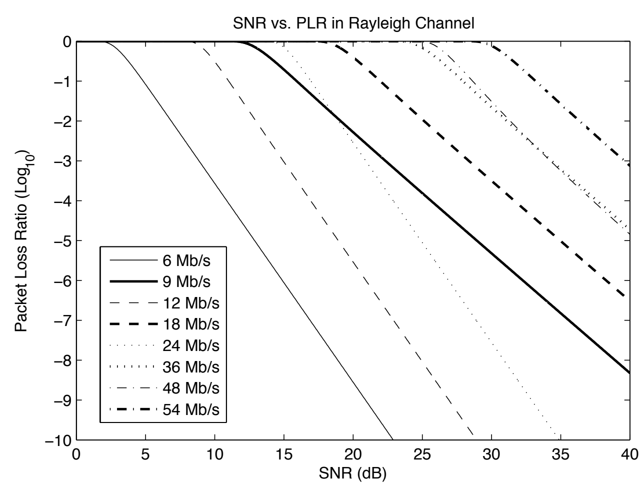

Fig. 2. PLR vs. SNR in a Rayleigh Channel. Theoretical prediction

2) Experimental Validation for the 802.11a rate-set: The theoretical prediction of PLR vs. SNR is based on the analysis of three diverse theoretical channel models of the environment. As these channel models are idealized, it is not clear if they can be used with confidence to draw accurate deductions for real WLANs. It is, therefore, essential to experimentally validate the predictions of possibly redundant rates. Our experimental

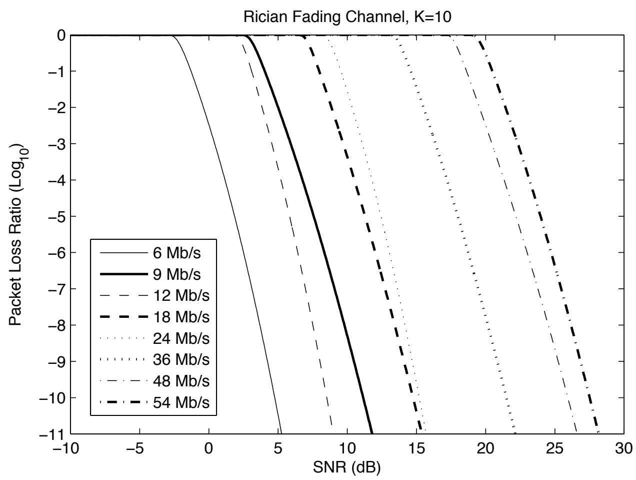

Fig. 3. PLR vs. SNR in a Rician Channel. Theoretical prediction

apparatus, which has been subject to substantial quantitative validation, is described in Appendix I.

We performed extensive measurements in three distinct environments: outdoor experiments on an open pitch and indoor experiments in an office environment both at night and during the day. For each rate, 20,000 packets with payload of 1,000 B were sent. The first 80 bytes of each payload were used to record experiment sequence number and transmission rate. The remaining payload bits were chosen randomly by a Bernoulli(1/2) process. The sequence number of correctly decoded packets at the receiver was collected. A binary sequence, which we call the loss sequence, was created with a 0 recorded for each packet that experiences a transmission failure and a 1 for each that is correctly received. Measurements were repeated with the laptops separated by increasing distances to vary the path SNR. In all experimental environments, we first investigated the auto-covariance of the loss sequence. Fig. 4 is a representative plot, the auto-covariance for the loss sequence at 12 Mb/s in the night-time indoor environment. The vertical range is extremely small and suggests that packet losses occur stochastically pairwise independently.

For the 6, 9, and 12 Mb/s rates Fig. 5 shows the complementary cumulative distribution functions of their PLR per second for an outdoor experiment at a separation of 160 meters. From this graph it can be clearly seen that at 9 Mb/s nearly every packet is lost, while the 12 Mb/s rate experiences packet loss of approximately 50%. Thus, while one intuitively expects lower rates to be more robust, this is not always the case. All higher rates experience 100% loss.

One set of indoor experiments was carried out at midnight to ensure that there was no human motion, which could cause variations in channel conditions. The transmitter and receiver laptops were placed in separate offices approximately 10 meters apart, with several partition walls between them. Fig. 6 reports complementary cumulative distribution function (ccdf) for PLR per second for this experiment. The 6 Mb/s has the lowest PLR most frequently. The ccdf for the 9 Mb/s rate, however, which is located at the extreme right of the plot, completely dominates the ccdf of the 12 Mb/s rate. Also shown are the 18 and 24 Mb/s rates. In these indoor experiments one might expect the Rayleigh fading model to be appropriate, but they do not confirm the second non-monotonic prediction of the Rayleigh fading channel of 18 Mb/s compared with 24 Mb/s.

A second collection of indoor experiments were performed at mid-day during a working week to investigate the impact of channel conditions driven by human motion as well as the switching on and off of computers with wireless cards. Fig. 7 is the autocovariance function for the loss sequence corresponding to an indoor daytime experiment at 12 Mb/s and, again, the vertical range is small suggesting little pairwise dependency. Fig. 8 reports ccdf for PLR that are typical of multiple experiments we performed. Although the absolute level of loss changes based on the environmental conditions, the redundancy of 9 Mb/s and the non-monotonicity feature of PLR do not change. From our experimental observations, we consistently find that the 802.11a 9 Mb/s rate is redundant. Consequently, we choose to eliminate it from the set of possible rates for RC. A more conservative scheme would be to sample both the 9 Mb/s and 12 Mb/s rates from 6 Mb/s, but the key point is that the monotonicity of robustness to noise of these rates cannot be taken for granted.

The question of the redundancy of the 802.11a 18 Mb/s rate is more subtle, as this has not been supported by any of our experiments. Adopting a risk-averse approach we suggest that care needs to be taken by adaptive RC algorithms when the 18 Mb/s rate appears to function poorly. In order to ensure that in this case H-RCA doesn't get stuck at the 12 Mb/s rate, on a rate increase decision from 12 Mb/s it samples the 18 Mb/s and 24 Mb/s rates in a round-robin fashion. Alternative sampling schemes to overcome these difficulties are discussed in Section VI-B.

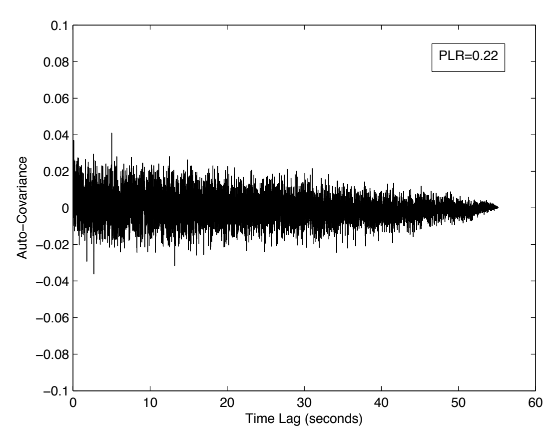

Fig. 4. Auto-Covariance of the loss sequence of 12 Mb/s in the night-time indoor environment at 10m separation. Experimental data

B. PLR Estimation

To distinguish failures due to noise from those caused by collisions, H-RCA uses a PLR estimation method based

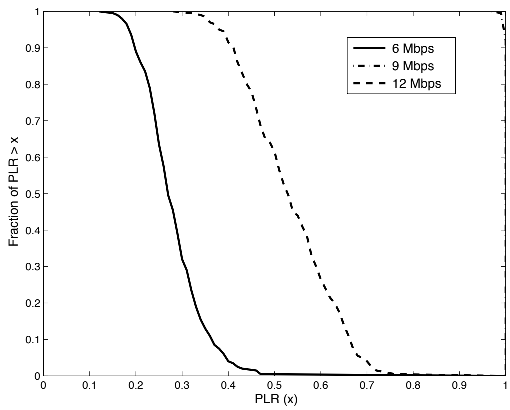

Fig. 5. Complementary cumulative distribution functions for PLR in the outdoor environment at 160m separation; 9 Mb/s curve barely visible in top right hand corner. Average PLR: 6 Mb/s 0.27; 9 Mb/s 0.99; 12 Mb/s 0.53; 18 Mb/s 1.00; 24 Mb/s 1.00. Experimental data

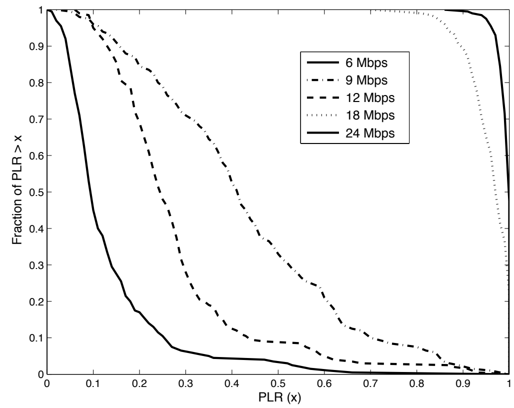

Fig. 8. Complementary cumulative distribution functions for PLR in daytime indoor environment at 10m separation. Average PLR: 6 Mb/s 0.13; 9 Mb/s 0.42; 12 Mb/s 0.27; 18 Mb/s 0.95; 24 Mb/s 0.99. Experimental data

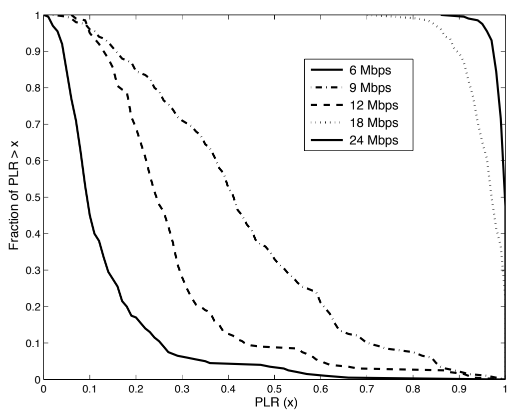

Fig. 6. Complementary cumulative distribution functions for PLR in the night-time indoor environment at 10m separation. Average PLR: 6 Mb/s 0.21; 9 Mb/s 0.28; 12 Mb/s 0.24; 18 Mb/s 0.73; 24 Mb/s 0.99. Experimental data

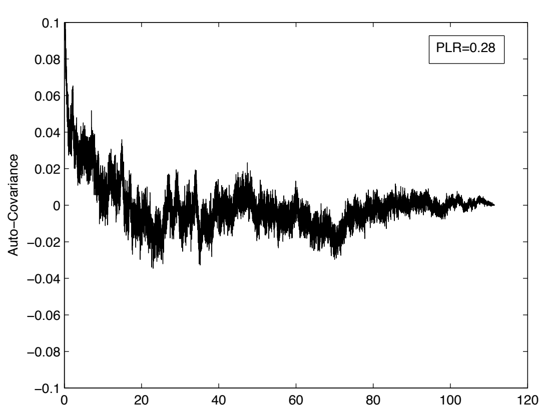

Fig. 7. Auto-Covariance of the loss sequence of 12 Mb/s in the daytime indoor environment at 10m separation. Experimental data

on the functionality of 802.11e's TXOP [7]. As defined in IEEE 802.11e, when a station gains access to the medium and successfully transmits a packet, if the remaining TXOP time is long enough for another packet transmission, the station can transmit the next packet after a short inter-frame space (SIFS, see Appendix II) without an additional back-off period. If any packet in the TXOP burst results in an unacknowledged transmission, no further packets are sent. At the time the second or later packets in the TXOP burst are transmitted, all other stations in the network see the medium as continuously busy so there can be no collision. In other words, if transmission of the second or later packets in the TXOP burst fails, it can only have been caused by noise[3]. Transmission failure of the first packet, however, can be due to both collision and noise. Thus it is necessary to record transmission statistics for these two classes of packets.

TABLE III

802.11A TXOP PARAMETERIZATION

Rate (Mb/s)	TXOP (for 1,000 B packets)	STh
6	0.0030s	361
12	0.0016s	589
18	0.0011s	779
24	0.0009s	893
36	0.0007s	1140
48	0.0006s	1349
54	0.0005s	N.A.

In H-RCA, all packets are sent in TXOP bursts[4]. The TXOP value is rate and packet-size dependent. It is set to allow two packets transmitted in each TXOP burst (for the 802.11a example, see Table III). For the sequence of first packets in the bursts, we define $F(k) := 1$ if the k^{th} packet is successfully received and $F(k) := 0$ if it is not acknowledged by its intended receiver. For the sequence of second packets

[3]Loses due to hidden nodes are not directly considered in this paper, see Section VI-D.

[4]The question of lightly loaded stations is addressed in Section VI-E.

in the TXOP bursts, which only exist if $F(k) = 1$, we define $S(k) := 1$ if the k^{th} packet is successfully received and $S(k) := 0$ otherwise. During time periods that rate change decisions are made, which will be shown to be short, we assume that $\{S(k)\}$ forms an i.i.d. sequence. Based on, for example, Fig. 4 and Fig. 7, this is a reasonable hypothesis. Define

$$P(S(k) = 0) := p_n,$$

where p_n is the probability of failure due to noise. Again, during the short time intervals during which rate change decisions are made, we assume that $\{F(k)\}$ is i.i.d. and that collisions are independent of noise so that

$$P(F(k) = 0) = 1 - (1 - p_n)(1 - p_c) =: p_l,$$

where p_c is the probability of failure due to collision and p_l is the probability of failure due to either collision or noise.

Ideally, H-RCA would only make its rate change decisions based on the sequence $\{S(k)\}$, but it is possible that this sequence will be completely censored by transmission failures of the first packets in each TXOP burst (when $P(F(k) = 1) = 0$). Thus a principled strategy is required to make decisions based on the statistics of the first packets too.

C. Rate Reduction Decision

The fundamental goal of H-RCA is to minimize the average time that packets spend occupying the medium, a quantity that we now determine as a function of MAC parameters, average packet size, p_n and p_c. Define $T_{tx}(r)$ to be the time on the medium that a first packet in a TXOP burst during a transmission with a physical layer (PHY) rate r Mb/s. Then

$$T_{tx}(r) = \text{DIFS+Header+(Payload)/r+SIFS+Header+ACK/}r_{ack},$$

where DIFS is the DCF inter-frame spacing, SIFS is the short inter-frame spacing and r_{ack} is the rate at which the ACK is sent. In ns-2 r_{ack} is set to 6 Mb/s. In our 802.11a experimental apparatus, r_{ack} is 6 Mb/s for the 6 and 9 Mb/s rates, 12 Mb/s for the 12 and 18 Mb/s rates, and 18 Mb/s for all higher rates. In H-RCA, all the re-transmissions are proceeded with the current rate r with Multi-Rate Retry mechanism disabled. From the above assumptions, using analysis along the lines found in [29][30], if a packet is the first packet in a TXOP burst and the PHY rate is r Mb/s, its expected time on the medium is[5]

$$\sum_{i=0}^{M} p_l^i T_{tx}(r) = \frac{1 - p_l^{M+1}}{1 - p_l} T_{tx}(r), \tag{1}$$

where M is the 802.11 retry limit (see Appendix II). If a packet is the second in a TXOP burst, its first transmission is delayed by a SIFS rather than a DIFS, where SIFS is two idle slots (σ) shorter than DIFS, SIFS = DIFS-$2\sigma\mu s$. Should it experience a collision, which can only be due to noise, it becomes the first packet in the next TXOP burst, but can at most experience $M-1$ more collisions before being discarded. Thus at rate r Mb/s the expected time on the medium is

$$\begin{aligned} T_{tx}(r) - 2\sigma + p_n \sum_{i=0}^{M-1} p_l^i T_{tx}(r) \\ = T_{tx}(r) - 2\sigma + p_n \frac{1 - p_l^M}{1 - p_l} T_{tx}(r). \end{aligned} \tag{2}$$

These two expected waiting times have to be weighted based on the likelihood that when a packet is first transmitted it is the first or second packet in a TXOP burst. These two events are not equally likely as if a second packet in a TXOP burst experiences transmission failure, it becomes the first packet in the next burst.

Under the above assumptions, the stochastic process that determines whether a packet is initially a first or second packet in a TXOP burst forms a Markov chain on two states. The first state corresponds to a packet initially being a first packet and the second corresponds to it initially being a second packet. The Markov chain's transmission matrix is

$$\Pi = \begin{pmatrix} p_l^{M+1} & 1 - p_l^{M+1} \\ 1 - p_n(1 - p_l^M) & p_n(1 - p_l^M) \end{pmatrix}.$$

The entries of Π can be understood as follows: if a packet is initially the first packet in a TXOP burst, the next one will also be if it is discarded, which happens with probability p_l^{M+1}. If a packet is initially a second packet in a TXOP burst and it experiences a failed transmission due to noise, becomes a first packet and is then not discarded, which happens with probability $p_n(1 - p_l^M)$, then the next packet will be a second packet too. The stationary distribution, λ where $\lambda\Pi = \lambda$, of this Markov chain gives the likelihood that a packet is initially a first packet in a TXOP burst or a second:

$$\left(\frac{1 - p_n(1 - p_l^M)}{2 - p_l^{M+1} - p_n(1 - p_l^M)}, \frac{1 - p_l^{M+1}}{2 - p_l^{M+1} - p_n(1 - p_l^M)} \right).$$

Thus, with PHY rate r Mb/s, the average time that a packet spends being transmitted is

$$\begin{aligned} T_s(r) = \frac{1 - p_n(1 - p_l^M)}{2 - p_l^{M+1} - p_n(1 - p_l^M)} \left(\frac{1 - p_l^{M+1}}{1 - p_l} T_{tx}(r) \right) + \\ \frac{1 - p_l^{M+1}}{2 - p_l^{M+1} - p_n(1 - p_l^M)} \left(T_{tx}(r) - 2\sigma + p_n \frac{1 - p_l^M}{1 - p_l} T_{tx}(r) \right). \end{aligned} \tag{3}$$

For a given rate and channel model, the probability of loss due to noise p_n can be determined as a function of SNR.

1) Rate-reduction (802.11a): In Fig. 9 the channel is AWGN and the probability of transmission failure due to collision is $p_c = 0.3$, corresponding to a station competing with 6 stations that always have packets to send [31]. In Fig. 10 the channel is Rayleigh Fading and $p_c = 0.01$, corresponding to a station competing for access in a lightly loaded network [29]. These figures are representative of graphs for AWGN and Rayleigh fading channels and a wide range of conditional collision probabilities. In both figures, the rate with the lowest average transmission time is the optimal rate for H-RCA at

[5] If successful and failed transmissions take distinct times, this can readily be taken into account.

that SNR value. The cross-points between these lines are the critical points where H-RCA should decrease its rate.

Corresponding to Figs 9 and 10, in terms of p_n ($\log_{10}$ scale) vs. SNR, Figs 11 and 12 plot these crossing-points. In both figures it is clear that those cross-points are distributed around the value $p_n = 0.1$. This pattern is the same for a wide range of conditional collision probabilities irrespective of the channel model and, therefore, for convenience in our simulations and experiments with the 802.11a rate-set H-RCA uses $p_n > 0.1 = p_{\text{thresh}}$ as the threshold to trigger rate reduction at all rates. H-RCA could, of course, be set up with a distinct p_n threshold value for each rate. In selecting 0.1 we are typically erring on the conservative side and demonstrate that this does not come at a significant performance cost. For each rate, the average transmission time at the SNR corresponding to the threshold value p_{thresh} is indicated in Figs 9 and 10.

D. Rate-reduction: Bayesian inference

Based on observations of $\{F(k)\}$ and $\{S(k)\}$, we adopt a Bayesian paradigm to the rate lowering decision. First note that the statistics of $\{F(k)\}$ depend on p_c as well as p_n, while $\{S(k)\}$ only depends on p_n. Given a sufficient number of observations of $\{S(k)\}$ to enable a good estimate of p_n, it is possible to estimate p_c from $\{F(k)\}$. However, this is not an approach we use as if p_n is large, it is possible that we get no $\{S(k)\}$ observations.

Instead we will take a rate lowering decision based on the experience of either $\{F(k)\}$ or $\{S(k)\}$, but use a worst-case *a priori* upper bound on p_c for $\{F(k)\}$ based on Bianchi's well-known model [31]. Assuming that stations always have packets to send, we use Bianchi's relationship between the conditional collision probability, p_c, and the number of stations in the WLAN. In practice, a WLAN is not likely to have a network with over 40 stations that always have packets to send. For such a situation $p_c \approx 0.6$ and so H-RCA finds the transmission failure probability, p_l, of the first packet in the TXOP burst over $1-(1-p_c)(1-p_{\text{thresh}}) = 0.6+0.4\,p_{\text{thresh}}$ then $p_n > p_{\text{thresh}}$ and it should choose a lower rate. For example, for the 802.11a rate-set with $p_{\text{thresh}} = 0.1$, this value is $p_l = 0.64$.

The Bayesian decision to change rate is based on the following question: using a uniform prior for p_n on $[0,1]$, conditional on the fact that the noise packet loss probability, p_n, is over p_{thresh}, in N packet transmissions, the Bayesian sampling window, how many failures should be observed before H-RCA has over 95% confidence that $p_n > p_{\text{thresh}}$? With $p_l(p_c,p_n) := 1-(1-p_c)(1-p_n)$ and p_c known, this corresponds to finding the minimal value of K that satisfies the following inequality:

$$\frac{\int_{p_{\text{thresh}}}^{1} \binom{N}{K} (1-p_l(p_c,p_n))^{N-K} p_l(p_c,p_n)^K \, dp_n}{\int_0^1 \binom{N}{K} (1-p_l(p_c,p_n))^{N-K} p_l(p_c,p_n)^K \, dp_n} \geq 0.95. \tag{4}$$

1) Bayesian inference (802.11a): Using (4), out of $N = 50$ transmission samples for each sequence, $\{F(k)\}$ and $\{S(k)\}$, and assuming $p_c = 0.6$ for $\{F(k)\}$ and $p_c = 0$ for $\{S(k)\}$, H-RCA should observe at least 39 failures out of 50 first packet transmissions or 9 failures out of 50 second packet transmissions to be over 95% confident that the $p_n > 0.1$ for the current rate and to decide to choose a lower rate. Note that these numbers, based on a principled design, are efficient. If the noise probability p_n is small enough that a rate reduction decision is not taken based on $F(1),\dots,F(N)$, then we will quickly get a sufficient sample of second packets in order to make an accurate decision based on noise-only failures. On the other hand, if the WLAN has less than 40 stations that always have packets to send and H-RCA sees more than 39 transmission failures for 50 first packets, it can confidently decide that $p_n > 0.1$ and lower the transmission rate.

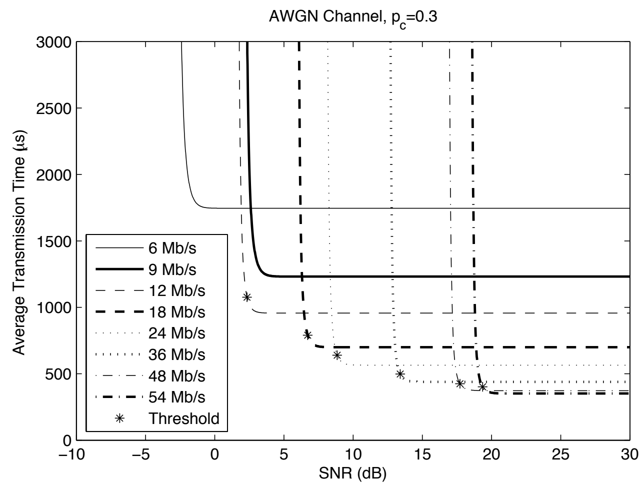

Fig. 9. Average Transmission Time T_s in AWGN channel with $p_c = 0.3$. Threshold points correspond to $p_n = 0.1$. Theoretical prediction

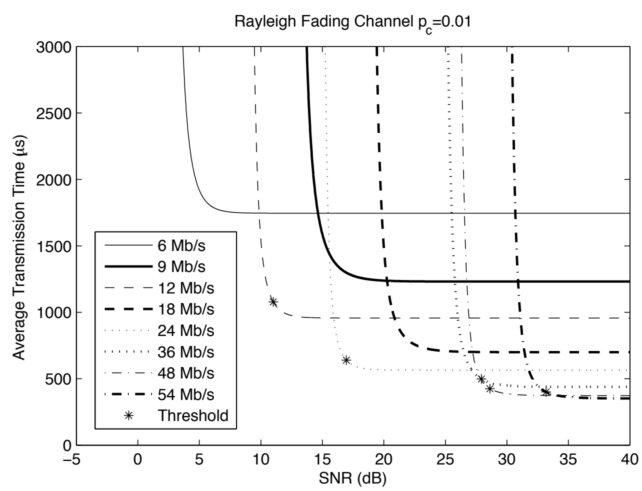

Fig. 10. Average Transmission Time T_s in Rayleigh Fading channel with $p_c = 0.01$. Threshold points correspond to $p_n = 0.1$. Theoretical prediction

E. Rate Increase Frequency

A commonly adopted process [15][4][16] for an adaptive algorithm to decide if it should try a higher rate is when it experiences a fixed number of successful transmissions.

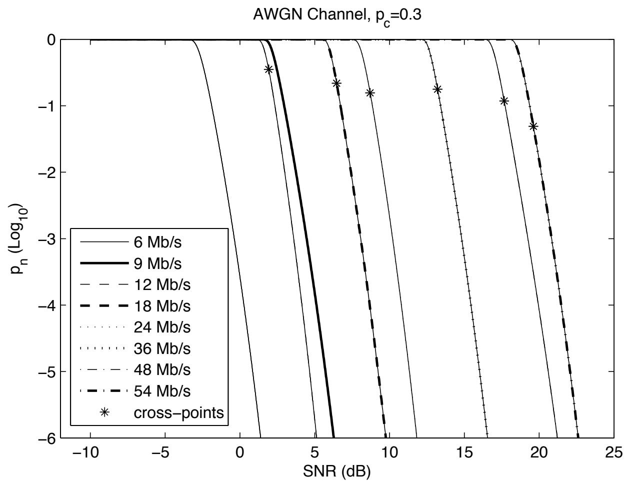

Fig. 11. Slow Down Points in AWGN Channel with $p_c = 0.3$. Theoretical prediction

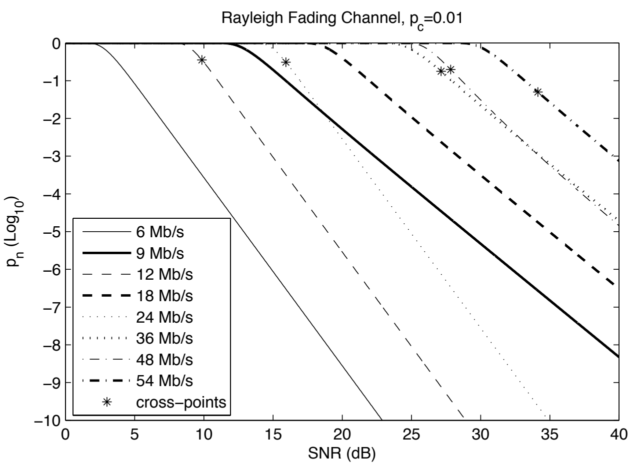

Fig. 12. Slow Down Points in Rayleigh Fading Channel with $p_c = 0.01$. Theoretical prediction

However, if the algorithm samples a higher rate that is not suitable for the current SNR too frequently, it will significantly decrease the station's throughput. On the other hand, if the algorithm samples higher rates rarely, it will be insensitive to changes in channel quality.

The solution that H-RCA employs to overcome this conundrum is to employ an opportunity-cost approach and have rate dependent successful transmission thresholds (STh). These thresholds are decided based on the following logic. The worst case scenario is if H-RCA is operating at a given rate, r Mb/s, with p_n=0 and attempts to transmit at a higher rate r' Mb/s whose p_n=1. Due to our Bayesian inference mechanism, H-RCA will not drop back to rate r Mb/s until it observes N transmission samples, which are all first packet transmissions due to $P(F(k) = 1) = 0$. Consequently, the time wasted on trying the high rate r' Mb/s will be N consecutive transmissions plus the back-off times between them,

$$H(r') := NT_{tx}(r') + \sigma \left(\frac{W \sum_{i=0}^{N-1} 2^{\min(i \bmod M, m)} - N}{2} \right).$$

where W is the minimum contention window, $2^m W$ is the maximum contention window and M is the discard limit (see Appendix II). Note that if a rate change decision can only be made after a packet's success or discard, as happens in Atheros hardware, then the upper limit in the sum should be set to $\lceil N/M \rceil M - 1$. Assuming no collisions, i.e. $p_c = 0$, instead of trying the high rate r', during this time $H(r')$, we could successfully transmit

$$\begin{aligned} X &:= \left\lfloor H(r')/\frac{1}{2}\left(T_{tx}(r) + T_{tx}(r) - 2\sigma + \sigma\frac{(W-1)}{2}\right)\right\rfloor \\ &= \left\lfloor H(r')/\left(T_{tx}(r) - \sigma + \sigma\frac{(W-1)}{4}\right)\right\rfloor \end{aligned}$$

packets at rate r Mb/s, half as first packets and half as second packets in TXOP bursts. Therefore, if the station can transmit DX packets at rate r Mb/s with trying rate r' Mb/s, this station could transmit $(D+1)X$ packets without trying rate r' Mb/s. To ensure that the penalty in lost transmission opportunities at the higher rate would result in achieving, at worst, we pick a target of 95% of the throughput of the current rate r Mb/s setting $DX/((D+1)X) = 95\%$ where $D = 19$. Thus, when currently at rate r Mb/s, the station changes to a higher rate r' Mb/s every time the station observes STh= $19X$ not necessarily consecutive successful transmissions (c.f. Table III).

1) Rate-increase frequency (802.11a): For packets of size 1,000 B, these values are given in Table III. Successful transmissions are counted over both first and second packets in each TXOP burst. To enable H-RCA to drop back quickly to its current rate if the higher rate proves to have $p_n > 0.1$, during the first instance of observation of a higher rate the algorithm uses Bayesian sampling window $N = 10$. Based on 95% confidence and the Bayesian analysis in equation (4), this means that the algorithm will drop back to its original rate if it observes 9 first packet transmission failures or 1 failed transmission for second packets. Should neither of these events occur, H-RCA stays at its current rate and resets N to be 50.

IV. 802.11a H-RCA Performance Evaluation

The following are three natural characteristics that can be used to evaluate the performance of a RC algorithm.

1) Accuracy: can it find the right rate for a given SNR?
2) Speed: how quickly does it converges to the right rate?
3) Noise vs. Collisions: is it robust to collision induced transmission failures?

As it is challenging to build an experimental wireless channel with controllable noise characteristics, here we present results from ns-2 simulations. Data from an experimental implementation is reported in Section V. In simulation we used both the AWGN and Rayleigh Fading channel models to determine p_n, the probability of failure due to noise, at each rate at a given SNR. As results for both channel models are similar and our experimental results suggest that the AWGN is the more appropriate of the two, we provide graphs only for the AWGN channel.

In existing commodity hardware, the physical layer performs automatic re-sends on transmission failure and the network card driver is only made aware of the transmission result at the MAC layer. To mirror this constraint, in ns-2 H-RCA works on information at the MAC-level, so H-RCA is only informed of the totality of a packet's transmission results after it has been successfully sent or discarded. That is, H-RCA receives new data only when the MAC layer finishes servicing each packet.

All stations transmit fixed 1,000 B UDP packets to an Access Point (AP) and always have packets to send. H-RCA's TXOP and STh values are set as in Table III. We have also implemented SampleRate [3] in ns-2. In order to provide a fair comparison we use the same simulation settings, including the same rate-set, TXOP values and the redundant 9 Mb/s rate is excluded from SampleRate's rate list.

We perform two sets of simulations to determine accuracy and speed. One set with a single station, so there are no collisions. The second set has five stations so that transmission failure can be caused by collisions.

We report on H-RCA's performance under two distinct, evolving SNR conditions: 1. step changes in channel quality; 2. gradual changes in channel quality. In the step-change case, SNR changes with the following discontinuous function:

$$\mathrm{SNR}(t) = \begin{cases} (15 + G(t))\,\mathrm{dB} & \text{if } 0s \le t \le 300s \\ (10 + G(t))\,\mathrm{dB} & \text{if } 300s < t \le 600s \\ (5 + G(t))\,\mathrm{dB} & \text{if } 600s < t \le 1200s \\ (10 + G(t))\,\mathrm{dB} & \text{if } 1200s < t \le 1500s \\ (15 + G(t))\,\mathrm{dB} & \text{if } 1500s < t \le 1800s \end{cases}$$

While in the gradient-change case, SNR varies as the following continuous V-shaped function:

$$\mathrm{SNR}(t) = \begin{cases} \left(15 - \dfrac{t}{90} + G(t)\right)\mathrm{dB} & \text{if } 0 \le t \le 900s \\ \left(-5 + \dfrac{t}{90} + G(t)\right)\mathrm{dB} & \text{if } 900s < t \le 1800s \end{cases}$$

In both cases, $\{G(t)\}$ is Gaussian Process with mean 0 and variance 1 and the resulting p_n is determined from the AWGN channel model.

A. Single Station, No Collisions

Our first simulation takes place in a WLAN with a single active client so that it experiences no collisions. In the discontinuous SNR scenario, Fig. 13 shows the WLAN's second-by-second throughput when using either H-RCA or SampleRate. The 30 minute average throughput for each RC algorithm and shows that H-RCA gets higher throughput than SampleRate. SampleRate looses throughput as it frequently samples the whole rate space. H-RCA's sampling frequency is restricted by design, so it achieves a consistently higher throughput than SampleRate. For this simulation we implemented an omniscient optimal algorithm that knows channel conditions in advance and can select the best rate at all times. Comparisons with this all-knowing algorithm shows that H-RCA is accurate, finding the correct rate and sampling the one above it. The 30 minute average throughput of the omniscient algorithm is 15.96 Mb/s. H-RCA gets 95% of this figure, as one would expect based on its rate-increase opportunity cost approach. The figure shows that H-RCA is responsive to a sudden change in SNR where it only takes seconds to adapt and stabilize rate in response to the dramatically different environmental conditions.

Fig. 14 shows simulation results in the case of SNR gradient demonstrating that H-RCA still delivers greater and less variable throughput than SampleRate. H-RCA sustains network throughput when SNR decreases slowly (from 200s to 600s), while the throughput of SampleRate drops continuously and is highly variable. Sizeable drops in SampleRate's throughput flag instances when SampleRate adapts its rate. H-RCA, however, makes better decisions more quickly and more accurately. In comparison to the omniscient algorithm, again H-RCA gets 95% of this maximum throughput.

H-RCA's decision making process is shown in Fig. 15. For the second half of the simulation, it plots indicates the instances at which rate change decisions were made in the example shown in Fig. 14. The rate at which decisions to increase rate are made reflect opportunity cost scheme where with one station sampling of higher rates can occur frequently without a performance detriment. These simulations demonstrate that even in the absence of collisions, H-RCA exhibits gains over SampleRate.

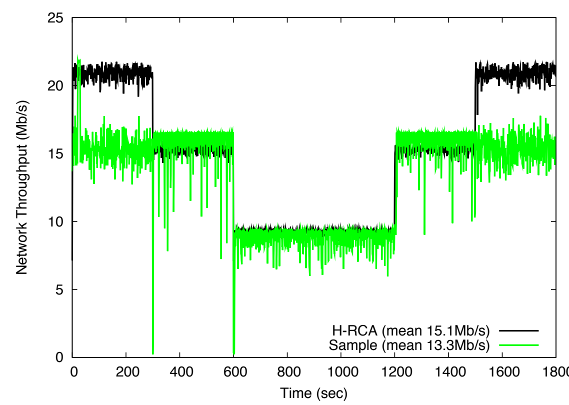

Fig. 13. Throughput in AWGN channel, SNR Step and 1 station. Simulation data

B. Five Stations With Collision

As a test of robustness to failures caused by collisions, the next simulation models a WLAN with five active stations, which is not an unrealistic practical scenario. Fig. 16 reports throughput in the SNR step case, while Fig. 17 is for the SNR gradient case. In comparison to the omniscient algorithm, for Fig. 11 H-RCA gets 98% of its throughput and for Fig. 17 it gets 97%. This demonstrates that H-RCA is accurate, always finding the best rate and sampling the one above it.

With the increased packet losses due to collisions in this network, SampleRate's decision making has been significantly affected. The influence of collision packet losses on H-RCA's

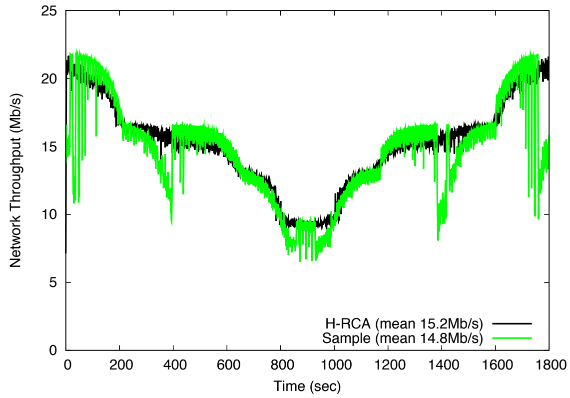

Fig. 14. Throughput in AWGN channel, SNR Gradient and 1 station. Simulation data

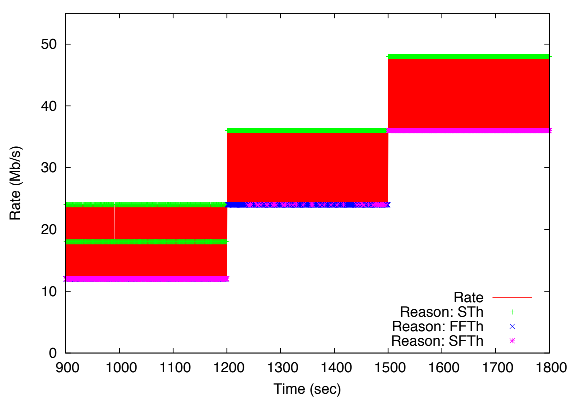

Fig. 15. Rate change decisions, AWGN channel, SNR Gradient and 1 station. UP indicates a rate increase decision, FFTh indicates a rate decrease decision based on first packets in a TXOP burst and SFTh indicates a rate decrease decision based on second packets in a TXOP burst. Simulation data

decision making is small. As H-RCA only makes rate change decisions after it has observed a certain number of packet transmissions, when the number of active clients in the network increases, this estimation time increases. Therefore, in this scenario H-RCA's reaction time is longer in comparison to the single-station network, but not unacceptably so.

The rate change decisions for a single station in the network are plotted in Fig. 18 in which the round-robin approach to sampling 24 and 36 Mb/s from 18 Mb/s between 900s and 1000s is clear. In contrast to Fig. 15 it can be seen that the presence of other active stations necessarily slows down the real-time adaptivity of the algorithm. This occurs as the rate at which RC algorithms gain channel information is a function of the rate at which they get to transmit packets.

Again these simulations demonstrate that H-RCA's reaction time is short. Most importantly, H-RCA is robust to transmission failures caused by collisions. Network performance, therefore, does not degrade in the presence of several active stations as is common in practice.

V. 802.11A EXPERIMENTS

It is not feasible to construct an experimental scenario with controllable noise characteristics, so in Section IV we inves-

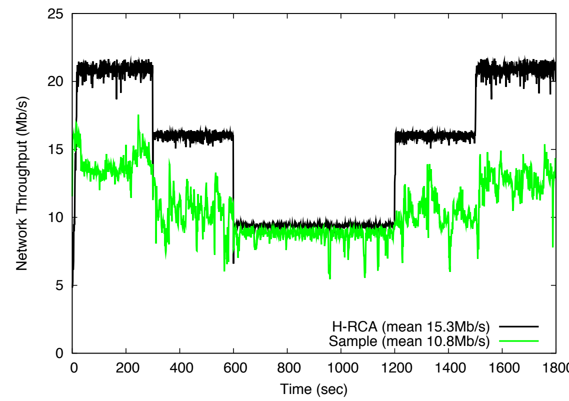

Fig. 16. Total Throughput in AWGN channel, SNR Step and 5 station. Simulation data

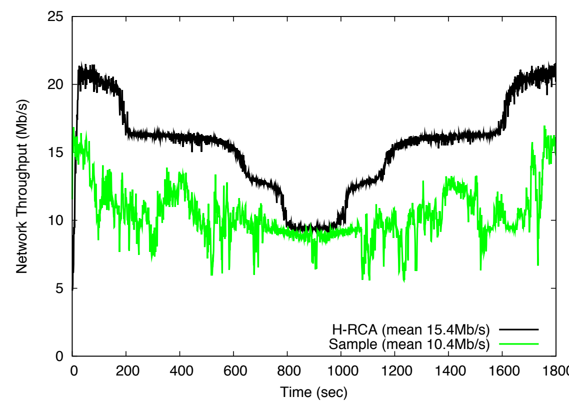

Fig. 17. Total Throughput in AWGN channel, SNR Gradient and 5 station. Simulation data

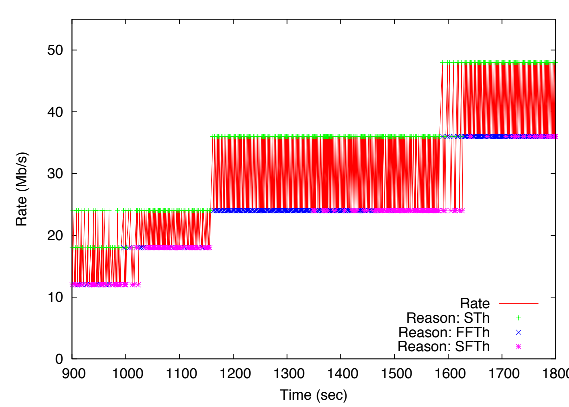

Fig. 18. Rate change decisions, AWGN channel, SNR Gradient and 5 stations. UP indicates a rate increase decision, FFTh indicates a rate decrease decision based on first packets in a TXOP burst and SFTh indicates a rate decrease decision based on second packets in a TXOP burst. Simulation data

tigated H-RCA's performance in the controlled environment of ns-2 simulations. As experiments can reveal difficulties not predicted by theory or simulation, in this section we report on experiments using the apparatus described in Appendix I. We compare H-RCA's throughput with that of Minstrel [2], SampleRate, [3], AMRR [4] and Onoe [5] as implemented in the MadWiFi driver for the Atheros chipset, and RRAA [6] as implemented by ourselves in the absence of a publicly available implementation. In all experiments, stations always have 1,000 B UDP packets to send using the 802.11a rate-set.

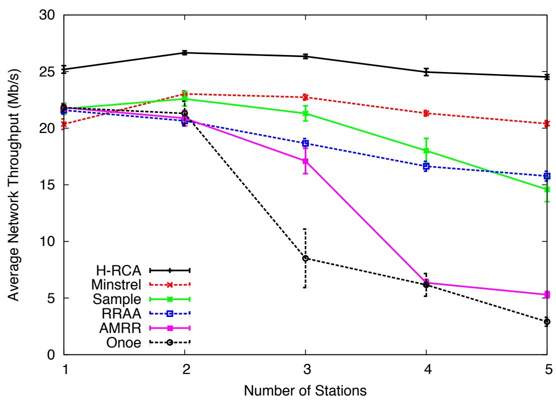

Fig. 19. Long run throughput for H-RCA, Minstrel, SampleRate, RRAA, AMRR and Onoe. Experimental data

For a WLAN consisting of between 1 and 5 stations, all of which are running the same RC-algorithm (H-RCA, Minstrel, SampleRate, RRAA, AMRR or Onoe), Fig. 19 reports the 5 minute average throughput in each WLAN. As the number of stations increases, the likelihood of failed transmission due to collisions, p_c, increases. SampleRate, AMRR and Onoe misinterpret collisions as being a consequence of bad channel conditions resulting in unnecessary increased rate sampling and, consequently, network resources are underutilized. RRAA distinguishes noise from collisions, but the overhead in using RTS/CTS costs throughput. Minstrel has an improved sampling technique over SampleRate and so does not over-react to increased collisions. The gain in throughput that is available in H-RCA by distinguishing collisions from noise by using the TXOP-based methodology is apparent. While stations are not likely to be constantly back-logged for long periods in practice, 5 back-logged stations would not be an uncommon scenario; indeed poor rate selection decisions make these periods more likely as they lead to increased congestion.

Fig. 20 and Fig. 1 (in Section I) report the dynamic throughput in the 2 and 5 station WLANs, respectively, on a second-by-second basis. These graphs demonstrate that H-RCA is consistent in its rate selection and its higher throughput does not come with any increased variability. The latter point is substantiated in Fig. 21 and Fig. 22 which provide histograms of these dynamic throughputs. As well as offering increased mean throughput, it is clear that H-RCA also offers decreased variance. This consistency is desirable for both real-time applications and TCP as its performance depends upon round-

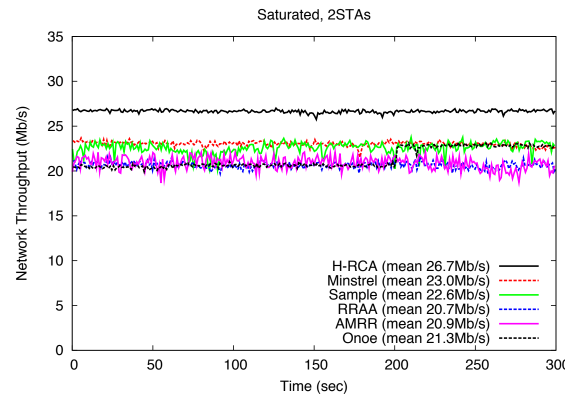

Fig. 20. Dynamic throughputs for H-RCA, Minstrel, SampleRate, RRAA, AMRR and Onoe in a 2-station WLAN. Experimental data

trip time statistics.

These experiments demonstrate the H-RCA methodology can deliver higher throughput with decreased variability.

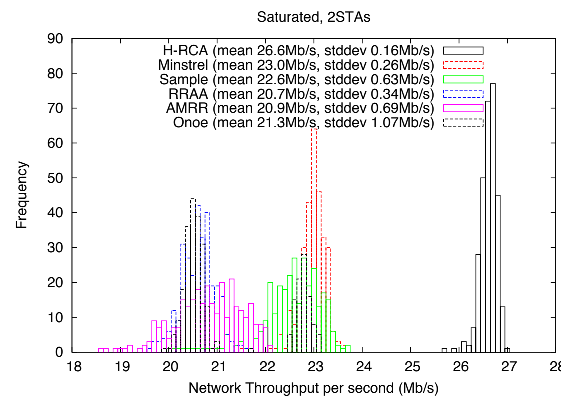

Fig. 21. Histogram of dynamic throughputs for H-RCA, Minstrel, SampleRate, RRAA, AMRR and Onoe in a 2-station WLAN. Experimental data

VI. Discussion and Conclusion

A. H-RCA's objective, possible alternatives

H-RCA is designed to maximize overall network throughput, but other objectives are possible. For example, if the station wished to selfishly optimize its own throughput, it could minimize the average service time of its packets. This would be achieved by changing (3) through appending to equations (1) and (2) a term corresponding to the mean MAC back-off time while the packet is at the head of the line awaiting transmission. This quantity can be calculated based on the model introduced in [31]. In order to implement this change, the station would need an estimate of the mean busy slot time on the medium, which would be practically challenging with existing hardware. Assuming this information is available, we have implemented this approach in ns-2 and the results (data not shown) display little difference from those based on equation (3).

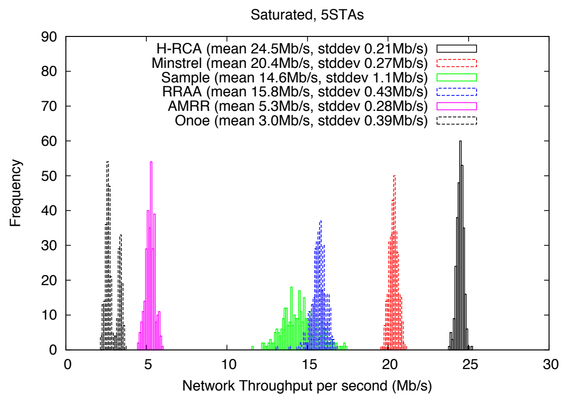

Fig. 22. Histogram of dynamic throughputs for H-RCA, Minstrel, SampleRate, RRAA, AMRR and Onoe in a 5-station WLAN. Experimental data

B. The 18 Mb/s 802.11a Rayleigh Fading Issue, other stratagems

To overcome the possible redundancy of the 802.11a 18 Mb/s rate, which is predicted by theory but not substantiated by experiments, we employ a round-robin strategy when a rate increase decision occurs from 12 Mb/s. Many other schemes could be proposed, including adaptive schemes, but the simplest one appears to function adequately in our tests. For example, we implemented a weighted scheme in which, on a rate increase decision the 18 Mb/s and 24 Mb/s rates were selected with frequencies related to the potential lost bandwidth if $p_n = 1$. In all tests, the results from this scheme were directly comparable to the round-robin methodology.

C. Redundancy of 9 Mb/s, other stratagems

All theoretical models predict that the 12 Mb/s rate is more robust to noise than the 9 Mb/s rate. Every experiment we performed was consistent with this prediction, so we elected to remove 9 Mb/s from the set of rates that H-RCA considered. Alternatively, one could employ a dual sampling strategy of investigating both 9 Mb/s and 12 Mb/s from 6 Mb/s, as we have advocated for the 18 Mb/s issue in Section VI-B above.

D. Hidden nodes

Rate-control in itself isn't a solution to the problem of hidden nodes, but can be used to mitigate their impact. The scheme we employ to separate collision and noise-based losses can be extended using standard 802.11e functionality to distinguish between hidden nodes based losses too [7]. By using the echoed NAV values in MAC ACKs, one can test if noise-based losses change in response to virtual carrier sense. If they do, these noise losses can be classified as losses as due to hidden nodes at the receiver. H-RCA can be readily modified to make use of this information in order to keep transmission rates high when noise is low, but hidden nodes are present. This is a only partial solution as completely mitigating for hidden nodes is within the remit of power and channel selection.

E. Non-saturated stations

We have focused on stations that have back-to-back packets to send so that they can be packaged in pairs in TXOP bursts. This is reasonable as it is when stations have a lot of traffic to send that efficient usage is of particular importance. If stations are extremely lightly loaded there are two alternative stratagems. If stations have large packets, the MAC can split them into two fragments the second of which is not subject to collisions. If stations have small packets, this is unnecessary and they can be sent individually as the dominant component of the transmission delay comes from fixed overheads in that case.

F. Summary

We have presented H-RCA, an adaptive collision-aware wireless rate control methodology. As H-RCA does not require specific hardware support nor any change in IEEE 802.11 standard, we have implemented it on commodity hardware.

Due to its TXOP technique used to distinguish the collision loss, H-RCA adapts appropriately to collision induced losses. Its rate decrease decision making process employs Bayesian analysis to ensure reasonable outcomes. Its increase-rate decision frequency is chosen in a way that guarantees near optimal performance in an unchanging environment.

As well as offering increased mean throughput over existing algorithms, experiments demonstrate that H-RCA also offers decreased variance in throughputs. This consistency is desirable for real-time applications that rely on high throughput and low jitter. It is also desirable for non-real-time traffic as the performance of TCP is dependent upon stable round-trip time statistics.

Appendix I
Experimental Apparatus

For the experiments described in Section III-A, two laptops, one used as a transmitter and one as a receiver, were equipped with version 0.9.4 of the MadWifi driver that was modified to enable the selection of a fixed transmission rate for multicast packets as well as disabling power control. Multicast packets were used to circumvent potential additional complexities caused by MAC level retries. Care was taken to avoid problems of station disassociation, lease expiry and so forth. In addition, it is known that human motion heavily influences measurement outcomes [32], so this was excluded from the clean-environment experiments. A spectrum analyzer was used to check that the channel chosen for the experiments was clear of interference from other signals.

The experimental apparatus used for the results presented in Sections I and V employs a PC acting as an Access Point (AP) and another 5 PCs acting as client stations. The WLAN is set up in infrastructure mode. All systems are equipped with an Atheros AR5215 802.11 PCI card with an external antenna. These cards do not suffer the serious difficulties reported in [33] and [34]. All stations, including the AP, use a version of the MadWiFi wireless driver that supported H-RCA, SampleRate, AMRR, Onoe and Minstral with RTS/CTS disabled as well as RRAA with RTS/CTS enabled. All stations are equipped with a 100 Mbps wired Ethernet port that is solely used for control of the test bed from a distinct PC. In the experiments, UDP traffic is generated by the Naval Research Laboratory's MGEN in periodic mode. All packets are generated in client stations before being transmitted to the AP.

APPENDIX II
A BRIEF OVERVIEW OF 802.11'S BEB ALGORITHM

On detecting the wireless medium being idle for a period DIFS, each station initializes a counter to a random number selected uniformly in the range $\{0, 1, \ldots, W-1\}$. Time is slotted and this counter is decremented once during each slot that the medium is observed idle. The count-down halts when the medium becomes busy and resumes after the medium is idle again for a period DIFS. Once the counter reaches zero the station attempts transmission and if a collision does not occur it can transmit for a duration up to a maximum period TXOP (defined to be one packet except in the Quality of Service MAC extension 802.11e). If two or more stations attempt to transmit simultaneously, a collision occurs. Colliding stations double their Contention Window (CW), up to a maximum value $2^m W$, selects a new back-off counter uniformly and the process repeats. If a packet experiences more collisions than the retry limit, M, where $M = 7$ in 802.11a, the packet is discarded. After the successful transmission of a packet or after a packet discard, CW is reset to its minimal value W and a new count-down starts regardless of the presence of a packet at the MAC. If a packet arrives at the MAC after the count-down is completed, the station senses the medium. If the medium is idle, the station attempts transmission immediately; if it is busy, another back-off counter is chosen from the minimum interval. This bandwidth saving feature is called post-back-off. The revised 802.11e MAC enables the values of DIFS (called the Arbitration Inter-Frame Spacing, AIFS, in 802.11e), CW and TXOP to be set on a per-class basis for each station. That is, traffic is directed to up to four different queues at each station, with each queue assigned different MAC parameter values.

REFERENCES

[1] K. Ramach, H. Kremo, M. Gruteser, and P. Spasojevi. Scalability analysis of rate adaptation techniques in congested ieee 802.11 networks: An orbit testbed comparative study. In *in Proc. of IEEE WoWMoM*, 2007.

[2] Minstrel rate adaptation algorithm. http://madwifi-project.org/browser/madwifi/trunk/ath_rate/minstrel, Jan 2010.

[3] J. Bicket. Bit-Rate Selection in Wireless Networks. Master's thesis, Massachusetts Institute of Technology, 2005.

[4] M. Lacage, M. Manshaei, and T. Turletti. IEEE 802.11 Rate Adaptation: A Practical Approach. In *ACM MSWiM*, 2004.

[5] Onoe rate adaptation algorithm. http://madwifi-project.org/browser/madwifi/trunk/ath_rate/onoe, Jan 2010.

[6] S. Wong, H. Yang, S. Lu, and V. Bharghavan. Robust Rate Adaptation for 802.11 Wireless Networks. In *ACM MobiCom Conference*, 2008.

[7] D. Giustiniano, D. Malone, D. J. Leith, and K. Papagiannaki. Measuring transmission opportunities in 802.11 links. *IEEE/ACM Transactions on Networking*, 18(5):1516–1583, 2010.

[8] D. Qiao, S. Choi, and K. G. Shin. Goodput Analysis and Link Adaptation for IEEE 802.11a Wireless LANs. *Mobile Computing, IEEE Transactions*, 1(4):278–292, 2002.

[9] S. Choudhury and J. D. Gibson. Payload Length and Rate Adaptation for Throughput Optimization in Wireless LANs. In *Vehicular Technology Conference*, 2006.

[10] G. Holland, N. Vaidya, and P. Bahl. A Rate-Adaptive MAC Protocol for Multi-Hop Wireless Networks. In *ACM MobiCom Conference*, 2001.

[11] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly. Opportunistic Media Access for Multirate Ad Hoc Networks. In *ACM MobiCom Conference*, 2002.

[12] P. A. K. Acharya, A. Sharma, E. M. Belding, K. C. Almeroth, and D. Papagiannaki. Congestion-Aware Rate Adaptation in Wireless Networks: A Measurement-Driven Approach. In *IEEE SECON Conference*, 2008.

[13] G. Judd, X. Wang, and P. Steenkiste. Efficient Channel-aware Rate Adaptation in Dynamic Environments. In *ACM MobiSys Conference*, 2008.

[14] M. A. Y. Khan and D. Veitch. Isolating physical PER for smart rate selection in 802.11. In *IEEE INFOCOM*, 2009.

[15] A. Kamerman and L. Monteban. Wavelan II: A High-Performance Wireless LAN for the Unlicensed Band. *Bell Labs Technical Journal*, 2(3):118–133, 1997.

[16] J. Kim, S. Kim, S. Choi, and D. Qiao. CARA: Collision-aware Rate Adapataion for 802.11 Wireless Networks. In *IEEE INFOCOM*, 2006.

[17] J. Zhang, K. Tan, J. Zhao, H. Wu, and Y. Zhang. A Practical SNR-Guided Rate Adaptation. In *IEEE INFOCOM Conference*, 2008.

[18] S. Rayanchu, A. Mishra, D. Agrawal, S. Saha, and S. Banerjee. Diagnosing Wireless Packet Losses in 802.11: Separating Collision from Weak Signal. In *IEEE INFOCOM Conference*, 2008.

[19] M. Vutukuru, H. Balakrishnan, and K. Jamieson. Cross-Layer Wireless Bit Rate Adaptation. In *ACM SIGCOMM*, 2009.

[20] S. Sen, N. Santhapuri, R. R. Choudhury, and S. Nelakuditi. AccuRate: Constellation based rate estimation in wireless networks. In *In USENIX NSDI*, 2010.

[21] J. Choi, J. Na, K. Park, and C. Kim. Adaptive optimization of rate adaptation algorithms in multi-rate WLANs. In *ICNP*, 2007.

[22] J. Choi, K. Park, and C. Kim. Cross-layer analysis of rate adaptation, DCF and TCP in multi-rate WLANs. In *IEEE INFOCOM*, 2007.

[23] J. Choi, J. Na, Y. s. Lim, K. Park, and C. k. Kim. Collision-aware design of rate adaptation for multi-rate 802.11 WLANs. *Selected Areas in Communications, IEEE Journal on*, 26(8):1366 – 1375, 2008.

[24] K. D. Huang, D. Malone, and K. R. Duffy. The 802.11g 11 Mb/s rate is more robust than 6 Mb/s. *IEEE Transactions on Wireless Communications*, 10(4):1015–1020, 2011.

[25] A. Goldsmith. *Wireless Communications*. CUP, 2005.

[26] M. K. Simon and M. S. Alouini. *Digital Communication over Fading Channels*. Wiley-IEEE Press, 2005.

[27] M. B. Pursley and D. J. Taipale. Error Probabilities for spreadspectrum packet radio with convolutional codes and viterbi decoding. *IEEE Trans. Commun.*, 35(1):1–12, 1987.

[28] W C Lindsey. Error probabilities for rician fading multichannel reception of binary and n-ary signals. *IEEE Transactions on Information Theory*, 10(4):339–350, 1964.

[29] D. Malone, K. Duffy, and D. J. Leith. Modeling the 802.11 Distributed Coordination Function in non-saturated heterogeneous conditions. *IEEE/ACM Transactions on Networking*, 15(1):159–172, 2007.

[30] K. Duffy and A. J. Ganesh. Modeling the impact of buffering on 802.11. *IEEE Comm. Lett.*, 11(2):219–221, 2007.

[31] G. Bianchi. Performance Analysis of IEEE 802.11 Distributed Coordination Function. *IEEE JSAC*, 18(3):535–547, March 2000.

[32] P. Barsocchi, G. Oligeri, and F. Potortia. Validation for 802.11b Wireless Channel Measurements. Technical report, Istituto di Scienza e Tecnologie dell'Informazione 'Ales sandro Faedo', 2006.

[33] G. Bianchi, A. Di Stefano, C. Gianconia, and L. Scalia. Experimental assessment of the backoff behavior of commercial IEEE 802.11b network. In *IEEE INFOCOM*, 2007.

[34] D. Giustiniano, G. Bianchi, L. Scalia, and I. Tinnirello. An explanation for unexpected 802.11 outdoor link-level measurement results. In *IEEE INFOCOM*, 2008.