Cross-Layer Analysis of Rate Adaptation, DCF and TCP in Multi-Rate WLANs

Jaehyuk Choi*, Kihong Park† and Chong-kwon Kim*
*School of Computer Science and Engineering, Seoul National University, Seoul, Korea 151-744
Email: {jhchoi, ckim}@popeye.snu.ac.kr
†Dept. of Computer Sciences, Purdue University, West Lafayette, IN 47907, U.S.A
Email: park@cs.purdue.edu

***Abstract*—Wireless Internet access is facilitated by IEEE 802.11 WLANs that, in addition to realizing a specific form of CSMA/CA—distributed coordination function (DCF)—implement a range of performance enhancement features such as multi-rate adaptation that induce cross-layer protocol coupling. Recent works in empirical WLAN performance evaluation have shown that cross-layer interactions can be subtle, sometimes leading to unexpected outcomes. Two such instances are: significant throughput degradation (a bell-shaped throughput curve) resulting from automatic rate fallback (ARF) having difficulty distinguishing collision from channel noise, and scalable TCP performance over DCF that is able to curtail effective multiple access contention in the presence of many contending stations. The latter also mitigates the negative performance effect of ARF. In this paper, we present station-centric Markov chain models of WLAN cross-layer performance aimed at capturing complex interactions between ARF, DCF, and TCP. Our performance analyses may be viewed as multi-protocol extensions of Bianchi's IEEE 802.11 model that, despite significantly increased complexity, lead to tractable and accurate performance predictions due to modular coupling. Our results complement empirical and simulation-based findings, demonstrating the versatility and efficacy of station-centric Markov chain analysis for capturing cross-layer WLAN dynamics.**

I. Introduction

A. Motivation

Wireless Internet access is predominantly mediated by IEEE 802.11 WLANs that, in addition to realizing a specific form of CSMA/CA—i.e., DCF with RTS/CTS—implement several performance enhancement features including multi-rate adaptation that induce cross-layer protocol coupling. Whereas DCF's exponential backoff determines when a frame transmission is attempted, adaptive rate selection dictates at what data rate a frame is encoded. For example, IEEE 802.11b supports four data rates 1, 2, 5.5, and 11 Mbps whereas 802.11a/g support eight. Rate adaptation is left to vendor discretion (i.e., is not part of the IEEE 802.11 standard [1]) whose aim is to select a data rate commensurate with channel noise. When SNR is low, a low data rate yields greater resilience to noise (robust PHY modulation), and vice versa. A widely deployed method is automatic rate fallback (ARF) [14] which uses up/down thresholds to set the data rate. In the Enterasys RoamAbout 802.11 DS High Rate card with Orinoco chipset (similarly for Cisco Aironet 350 cards), two consecutive frame transmission failures—i.e., ACK frame is not received—result in rate downshift. Ten consecutive frame transmission successes trigger a rate upshift. Asymmetry in the threshold values injects a measure of conservativeness, reflecting the sensitive dependence of bit errors on SNR. Most chip firmware implement variants of the canonical ARF but are based on the same up/down counter mechanism [3], [15], [18]. This is not dissimilar to differences in TCP versions—e.g., Reno, NewReno—and implementations across operating systems that follow linear increase/exponential decrease as their core congestion control, the target of steady-state TCP analysis [17].

Although well-intentioned, ARF can do more harm than good. Since frame transmission failures—inferred from missing 802.11 ACKs—can result from channel noise or collision, ARF may respond to frame collision resulting in rate downshifts when channel noise is low. Empirical 802.11b WLAN measurements [7] show that even under moderate multiple access contention (4–15 wireless stations) WLAN throughput declines drastically, not because of network congestion but ARF confusing collision with channel noise. In the same physical environment, fixing the data rate at 11 Mbps (by default, ARF is enabled in WLAN cards) can give a 5-fold increase in system throughput. In [7] it is also shown that when TCP is active over ARF and DCF (the typical modus operandi for wireless Internet access), the detrimental influence of ARF is largely mitigated. Using experiments and simulation it is shown that this is due to the scalable performance of TCP over DCF which curtails effective multiple access contention. This, in turn, desensitizes ARF against misleading collision cues. The collective dynamics of ARF, DCF, and TCP can be subtle and complex.

B. New Contribution

Modeling cross-layer protocol interactions is, in general, a challenging task. Modeling the collective behavior of ARF, DCF, and TCP is no exception. In this paper, we tackle the problem of capturing cross-layer protocol interactions in 802.11 WLANs, with the aim of providing effective analysis tools for understanding and improving WLAN performance. First, we consider MAC layer interactions between ARF and DCF's exponential backoff which are joined by frame transmission events at the multi-rate PHY layer. We define an ARF Markov chain generated by the point process of transmission events. We derive a closed-form solution parameterized by the up/down thresholds and rate-dependent transmission

failure probabilities, which gives the probability of choosing a specific data rate in steady-state. The ARF chain is joined with a multi-rate extension of Bianchi's DCF chain [5]—parameterized by ARF's rate probabilities—which gives the rate-dependent transmission failure and attempt probabilities. The state-space explosion problem of the combined ARF-DCF system is handled by finding a globally consistent solution to the two parameterized subsystems—afforded by modular coupling—that lends itself to fixed-point methods. We validate the analysis by comparing the predicted results with *ns-2* simulations which show the qualitative fidelity and quantitative accuracy of the joint ARF-DCF model.

Second, we incorporate the influence of TCP running over IEEE 802.11 infrastructure WLANs with exponential backoff (w/ and w/o RTS/CTS) and ARF. Several solutions have been proposed for improving ARF performance [6], [15], [21], but the problem is more entangled because WLAN performance is affected by cross-layer interactions, in particular, TCP. The empirical findings in [7] show that TCP-over-DCF is able to regulate multiple access contention in a WLAN, even when the number of contending stations is large, such that ARF's miscues from frame collision—and consequent throughput degradation—is mitigated. Essential to this is the scalable throughput performance of TCP-over-DCF which drastically reduces the frequency of consecutive frame collisions. As in the ARF-DCF model, we adopt a station-centric approach where TCP-over-DCF dynamics is captured by a Markov chain over a station's backlog state, incorporating the symmetry of multiple access contention in DCF. We show that the combined model yields accurate predictions of both coarse-granular (throughput) and fine-granular (station dynamics) TCP-over-WLAN performance.

Remark. Our cross-layer WLAN models may be viewed as advancing the station-centric Markov chain analysis approach followed by Bianchi [5], whose surprisingly accurate performance predictions are only now being better understood [22]. Our results show that modular station-centric Markov modeling—even when the underlying Markov chains for ARF, DCF, and TCP are very different—accurately capture WLAN performance in the presence of cross-layer interaction spanning PHY, MAC and transport layers.

C. Related Work

Bianchi's Markov chain model of IEEE 802.11 DCF [5] has been extended in several directions including explicit consideration of carrier sense [30], maximum retransmission count [25], prioritized channel access [26], and capture effect [8]. They represent direct extensions of the DCF chain. A multi-rate generalization of Bianchi's model has been considered in [29], albeit with the severe restriction that stations are assigned fixed data rates. An interesting technical advance is a Markov chain model over the state-space counting the number of stations in a backoff stage [22] that trades Bianchi's independence assumption with stochastic approximation using average (i.e., deterministic) dynamics. The results are relevant because they help explain why the independence assumption leads to accurate quantitative predictions, a hallmark of Bianchi's Markov chain model.

Empirical performance evaluation of ARF on 802.11b WLANs performance were provided in [7], which showed drastic throughput degradation caused by ARF under moderate multiple access contention. ARF's unintended side effect has been a source of confusion in both academia and industry. For example, in [24] sharp WLAN throughput decline in an experimental WLAN was attributed to CSMA's multiple access congestion. In [13] it is noted that "as the number of contending stations increases, aggregate capacity drops precipitously (to less than 1 Mb/s with 10 contending stations)" which is blamed on CSMA. With ARF disabled, empirical findings in [7] show that WLAN throughput decreases gradually with increasing contention level. TCP-over-WLAN performance has been considered generally poor due to channel noise and multiple access contention [20], [28], given TCP's sensitive dependence on packet loss rate ($\propto p^{-1/2}$) [17]. This has prompted solutions aimed at reducing TCP's exposure to frame errors and collisions on the wireless channel [4], [9], [27]. In [7] it is shown through experiment and simulation that TCP-over-WLAN achieves high throughput, even when the number of contending stations is large. Our analysis provides an explanation for the surprising agility of TCP-over-WLAN performance.

Several methods have been proposed to deal with ARF's noise vs. collision differentiation problem [6], [15], [16], [18], [21]. These results focus on improving ARF performance through enhanced algorithms and protocol mechanisms. For example, in [15] a modified ARF that makes use of RTS/CTS is proposed which exploits the fact that RTS frames are small and always encoded at the lowest rate. Therefore a RTS frame transmission failure is likely the result of collision, whereas data frame transmission failures following a successful RTS/CTS handshake are likely due to channel error. An overview of existing methods can be found in [10].

II. Analytical Model of ARF

In this section, we present a performance analysis model of ARF, its closed-form solution, and performance validation.

A. Markov Chain Model of ARF

We assume a multi-rate IEEE 802.11 WLAN with L data rates $R_1 < R_2 < \cdots < R_L$. For example, in 802.11b $L = 4$ with rates 1, 2, 5.5, and 11 Mbps. In 802.11a/g $L = 8$ with data rates 6, 9, 12, 18, 24, 36, 48, 54 Mbps. Following Bianchi [5], we introduce the independence assumption that in equilibrium a frame transmission experiences collision with constant and independent probability p. By symmetry, each of the N stations is homogeneous, subject to the same channel conditions in steady-state. The frame error rates obey $e_1 \leq e_2 \leq \cdots \leq e_L$ due to the increased robustness of 802.11 PHY modulation at lower data rates. The conditional frame failure probability (from collision or noise) of a frame transmitted at rate R_i is given by $p_i = 1 - (1-p)(1-e_i)$.

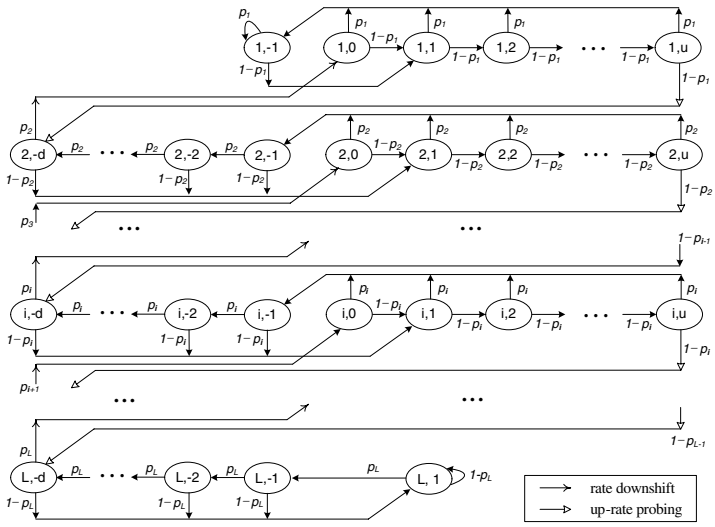

Fig. 1. Markov chain model of ARF with L data rates, up-threshold θ_u and down-threshold θ_d where $u = \theta_u - 1$ and $d = \theta_d - 1$.

Rate adaptation is a performance feature implemented in firmware that is left to vendor discretion. ARF is the most widely implemented method which uses two thresholds θ_d (down-threshold) and θ_u (up-threshold), where θ_d consecutive transmission failures result in a rate downshift and θ_u consecutive successes trigger a rate upshift (more precisely, up-rate probing). For example, in the Enterasys RoamAbout 802.11 DS High Rate card [3] (similarly for Cisco Aironet 350 cards), downshifting is triggered by θ_d=2 consecutive failures to receive an 802.11 ACK frame. The up-threshold is θ_u=10. A key weakness of current implementations of ARF is that channel noise and collision are not effectively distinguished. This can result in significant throughput degradation (a bell-shaped throughput curve) stemming from multiple access contention [7].

Our aim in this section is to capture the workings and impact of ARF on WLAN performance. Bianchi's DCF Markov chain [5] models the exponential backoff process by considering a Markov chain induced by the backoff stage and time counter. The chain is driven by the point process of frame transmission events under the aforementioned independence assumption. We define a Markov chain generated by the same point process that tracks the data rate selected by ARF. Let $r(t) \in \{1, 2, \ldots, L\}$ denote the data rate index. Let $c(t)$ ($-\theta_d + 1 \leq c(t) \leq \theta_u - 1$) denote the counter of consecutive failures ($c(t) < 0$) or successes ($c(t) > 0$) at rate $r(t)$. We adopt a discrete time model indexed by t which corresponds to the end time of the t-th transmission event of a tagged station. Assuming $e_1, \ldots, e_L$ and p are given—how p is determined is addressed in Section III—the ARF Markov chain $(r(t), c(t))$ is depicted in Fig. 1. For $p_i > 0$, the chain is irreducible and aperiodic, and we are interested in finding the unique equilibrium probability

$$\Pi_i = \sum_{k=-\theta_{\rm d}+1}^{\theta_{\rm u}-1} r_{i,k}, \quad 1 \leq i \leq L,$$

where $r_{i,k} = \lim_{t \to \infty} P\{r(t) = i, c(t) = k\}$. Π_i captures a station's probability of transmitting at data rate R_i.

Remark. Different rate adaptation methods result in different Π_i which provides a well-defined interface for integrating with DCF. Thus our modular cross-layer WLAN analysis approach is not limited to ARF, although one of the technical contributions of this paper is providing a rigorous performance analysis of ARF, to our best knowledge, a first.

B. Steady-state Solution of ARF Markov Chain

Although different in structure from Bianchi's DCF chain, the ARF Markov chain possesses regularities that admit to a closed-form solution for Π_i as a function of the system parameters p_i, θ_d and θ_u. A key observation to finding the solution is that *the ARF chain can be transformed into a coarsified birth-death chain* by aggregating the states corresponding to $r_{i,k}$ for different counter values k into a single macro state i ($i \in \{1, \ldots, L\}$). Since the equilibrium distribution of a L-state birth-death chain with birth rates λ_i ($i \in \{1, 2, \ldots, L\text{-}1\}$) and death rates μ_i ($i \in \{2, \ldots, L\}$) is given by

$$\Pi_1 = \frac{1}{1 + \sum_{j=1}^{L-1} (\prod_{k=1}^{j} \frac{\lambda_k}{\mu_{k+1}})} \text{ and } \Pi_i = \frac{\lambda_{i-1}}{\mu_i} \Pi_{i-1}, \quad (1)$$

for $i \in \{2, \ldots, L\}$, it suffices to find closed-form solutions for λ_i and μ_i in terms of p_i, θ_d and θ_u. In the following, we set $u = \theta_u - 1$ and $d = \theta_d - 1$ for notational simplicity.

λ_i and μ_i denote the state transition rates of increasing the current rate i to $i+1$ and decreasing the current rate i to $i-1$, respectively. They can be written as

$$\lambda_i = \frac{r_{i,u} P\{i+1, -d \,|\, i, u\}}{P\{\text{current rate} = R_i\}} = \frac{r_{i,u}}{\Pi_i}(1 - p_i), \quad (2)$$

$$\mu_i = \frac{r_{i,-d} P\{i-1, 0 \,|\, i, -d\}}{P\{\text{current rate} = R_i\}} = \frac{r_{i,-d}}{\Pi_i} p_i. \quad (3)$$

Note that for λ_i the index is defined for $i \in \{1, \ldots, L-1\}$ whereas for μ_i the index ranges over $i \in \{2, \ldots, L\}$. The ARF

Markov chain (Fig. 1) obeys the balance equations

$$r_{i,k} = (1-p_i) r_{i,k-1}, \qquad 1 \le i < L,\ 2 \le k \le u, \quad (4)$$
$$r_{i,k} = p_i r_{i,k+1}, \qquad 1 < i \le L, -d < k \le -2, \quad (5)$$

which yield

$$r_{i,k} = (1-p_i)^{k-1} r_{i,1}, \qquad 1 \le i < L,\ 1 \le k \le u, \quad (6)$$
$$r_{i,k} = p_i^{-(k+1)} r_{i,-1}, \qquad 1 < i \le L, -d < k \le -1. \quad (7)$$

We also get the balance equations

$$r_{i,1} = (1-p_i) \sum_{k=-d}^{0} r_{i,k}, \qquad 1 \le i < L, \quad (8)$$
$$r_{i,-1} = p_i \sum_{k=0}^{u} r_{i,k} \qquad 1 < i \le L, \quad (9)$$
$$r_{i,0} = p_{i+1} r_{i+1,-d}, \qquad 1 \le i < L, \quad (10)$$
$$r_{i,-d} = p_i r_{i,-d+1} + (1-p_{i-1}) r_{i-1,u}, \qquad 1 < i \le L. \quad (11)$$

First, we consider λ_i. Π_i can be split into two parts

$$\Pi_i = \sum_{k=-d}^{u} r_{i,k} = \sum_{k=-d}^{0} r_{i,k} + \sum_{k=1}^{u} r_{i,k}.$$

Using Eqs. (8) and (6), we obtain

$$\begin{aligned}\Pi_i &= \frac{r_{i,1}}{1-p_i} + r_{i,1} \sum_{k=1}^{u} (1-p_i)^{k-1} \\ &= \frac{1-(1-p_i)^{u+1}}{p_i(1-p_i)} r_{i,1} = \frac{1-(1-p_i)^{u+1}}{p_i(1-p_i)^u} r_{i,u}.\end{aligned} \quad (12)$$

Applying Eq. (12) to Eq. (2),

$$\lambda_i = \frac{r_{i,u}}{\Pi_i}(1-p_i) = \frac{p_i(1-p_i)^{u+1}}{1-(1-p_i)^{u+1}} = \frac{p_i(1-p_i)^{\theta_u}}{1-(1-p_i)^{\theta_u}}. \quad (13)$$

Next, we consider μ_i. Π_i may also be written as

$$\Pi_i = \sum_{k=-d}^{u} r_{i,k} = r_{i,-d} + \sum_{k=-d+1}^{-1} r_{i,k} + \sum_{k=0}^{u} r_{i,k}.$$

Using Eqs. (7), (9) and (5), we get

$$\begin{aligned}\Pi_i &= r_{i,-d} + \frac{1-p_i^{d-1}}{1-p_i} r_{i,-1} + \frac{r_{i,-1}}{p_i} \\ &= r_{i,-d} + \frac{1-p_i^d}{(1-p_i)p_i} r_{i,-1} = r_{i,-d} + \frac{1-p_i^d}{(1-p_i)p_i^{d-1}} r_{i,-d+1}.\end{aligned} \quad (14)$$

From the definition of λ_i, $(1-p_{i-1}) r_{i-1,u} = \lambda_{i-1}\Pi_{i-1}$. The detailed balance equation, $\lambda_{i-1}\Pi_{i-1} = \mu_i \Pi_i$, holds for the birth-death chain. Applying these to Eq. (11), we have

$$r_{i,-d} = p_i r_{i,-d+1} + \mu_i \Pi_i = p_i r_{i,-d+1} + p_i r_{i,-d}$$

which yields

$$r_{i,-d+1} = \frac{1-p_i}{p_i} r_{i,-d}. \quad (15)$$

Substituting Eq. (15) into Eq. (14), we get

$$\Pi_i = r_{i,-d} + \frac{1-p_i^d}{(1-p_i)p_i^{d-1}} \frac{1-p_i}{p_i} r_{i,-d} = \frac{1}{p_i^d} r_{i,-d}.$$

Finally, we obtain the departure rate

$$\mu_i = \frac{r_{i,-d}}{\Pi_i} p_i = p_i^{d+1} = p_i^{\theta_d}. \quad (16)$$

Substituting λ_i and μ_i into Eq. (1), we arrive at an expression for Π_i as a function of p_k $(k = 1, \ldots, L)$, θ_u and θ_d.

C. ARF Performance Validation

We evaluate the accuracy of the ARF model by comparing the analytical results with those of *ns-2* simulations with CMU's wireless extension. We simulated an IEEE 802.11g WLAN in the saturated throughput regime (i.e., all stations are infinite source) where $L = 8$ with data rates 6, 9, 12, 18, 24, 36, 48, and 54 Mbps. The ARF thresholds are θ_d=2 and θ_u=10, and payload size is 1000 bytes. Channel conditions were set to generate frame error rates (FERs) based on empirical PHY measurements [2]. p_i is set to the value observed in the simulations. The case where p_i is predicted as part of the analysis is treated in the next section.

Fig. 2 compares the steady-state rate distribution probability Π_i obtained from the ARF Markov chain model and *ns-2* simulations. Irrespective of the number of stations $N = 2, 5, 10$, and 25, we find a close match between the analytical results and simulation. In addition to quantitative accuracy, the ARF model predicts the qualitative trend of current ARF implementations that dwell most of the time at low data rates even for moderate contention levels [7]. Due to space constraints we omit the IEEE 802.11b results which yield similar predictive accuracy.

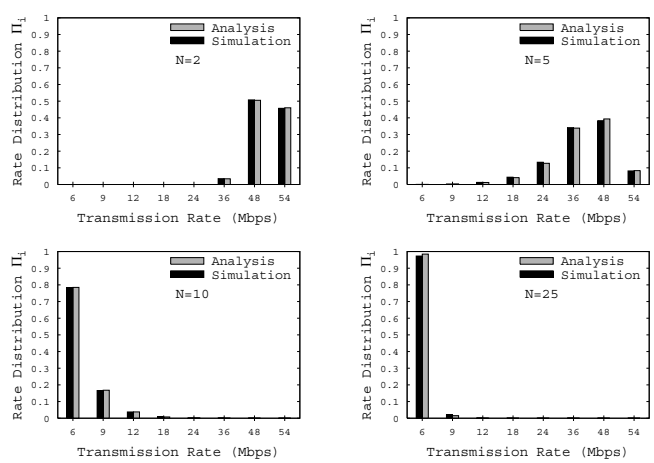

Fig. 2. Rate distribution (Π_i): Analysis vs. simulation ($\theta_u = 10$, $\theta_d = 2$) for $N = 2, 5, 10$, and 25 in IEEE 802.11g.

III. Performance Analysis of Combined ARF-DCF

The previous section modeled the behavior of ARF assuming $p_1, \ldots, p_L$ are given. In this section, we tackle the problem of analyzing multi-rate IEEE 802.11 WLANs with DCF and ARF when their influence extends in both directions.

A. Integrated ARF-DCF Multi-rate Model

We consider a multi-rate IEEE 802.11 WLAN with L data rates $R_1 < R_2 < \cdots < R_L$ accessed by N stations. Bianchi's DCF Markov chain $(s(t), b(t))$ is driven by the point process of frame transmission events under the independence assumption that a frame experiences the same channel condition in equilibrium [5]. $s(t)$ and $b(t)$ denote the backoff stage and backoff time counter, respectively. Our ARF chain $(r(t), c(t))$ is generated by the same point process and determines what data rate is used at backoff stage $s(t)$ when $b(t) = 0$. The outcome of the transmission attempt prescribes the next state. The combined system $(s(t), b(t), r(t), c(t))$ can be modeled as a Markov chain over their product space where the frame transmission failure probability p_i depends on the current rate $r(t) = i$. Not all state combinations, however, are reachable. For example, starting from initial state $s(0) = b(0) = 0$, $r(0) = L$ and $c(0) = 1$, the system cannot reach (x, y, L, z) for $x > \theta_d$ due to the rate downshift operation of ARF. Thus θ_d determines the state-space boundary that envelops the irreducible core. The latter is also aperiodic, hence ergodic.

B. Fixed-point Solution under Modular Coupling

A static multi-rate 802.11 model without rate adaptation is considered in [29] where each station in group i is assumed to use a *fixed* rate R_i, partitioned into L groups $n_1 + \cdots + n_L = N$. Consequently, n_i is fixed as well. Under the independence assumption, a station in group $i \in \{1, \ldots, L\}$ obeys Bianchi's DCF chain with homogenous transition rate p_i. Since frame collisions are assumed to occur independently—different rates only lead to variable collision slot durations—Bianchi's frame transmission attempt rate

$$\tau_i = \frac{2(1-2p_i)}{(1-2p_i)(W_0+1) + p_i W_0 (1-(2p_i)^m)} \tag{17}$$

holds with the added dependence on $i \in \{1, \ldots, L\}$. Here m is the maximum backoff stage and W_0 is the minimum backoff window size. When ARF is present, a station will adjust its data rate over time as frame transmissions succeed or fail. This implies that $n_1, \ldots, n_L$ are not given but *determined* by the dynamics of the integrated multi-rate WLAN with ARF and DCF. That is, $n_i(t)$ is a function of time.

The state-space explosion problem makes direct analysis difficult. We approximate the steady-state solution of the combined ARF-DCF Markov chain using *modular coupling*. First, we find parameterized solutions of the ARF chain $(r(t), c(t))$ and DCF chain $(s(t), b(t))$ separately, then resolve the parameters to find a globally consistent solution for the combined system. In the ARF chain, the closed-form solution (cf. Section II) has the form

$$\Pi_i = f(p_1, \ldots, p_L, \theta_d, \theta_u), \quad i \in \{1, \ldots, L\}, \tag{18}$$

parameterized by p_i which is determined by the DCF chain. In the DCF chain, a solution for (p_i, τ_i) is sought that satisfies Eq. (17) and

$$p_i = 1 - (1-\bar{\tau})^{N-1}(1-e_i) \tag{19}$$

where $\bar{\tau} = \sum_{i=1}^{L} \Pi_i \tau_i$ is the mean transmission rate of a station (i.e., averaged over the L data rates that it may employ). The latter is the fixed-point formulation of Bianchi's DCF Markov chain.

In the combined ARF-DCF model, we make use of the fact that Eq. (19) is parameterized by Π_i, $i \in \{1, \ldots, L\}$, since $\sum_{i=1}^{L} \Pi_i (1-\tau_i) = (1-\bar{\tau})$ which is the probability that a station does not transmit in a random slot. The Π_i, in turn, are determined by the ARF chain. We arrive at a globally consistent solution by finding (p_i, τ_i, Π_i), $i \in \{1, \ldots, L\}$, that satisfy Eqs. (17), (19), and (18) using fixed-point techniques. Thus the combined ARF-DCF model can be viewed as a multi-protocol extension of Bianchi's model whose total dimension has increased by one from (p_i, τ_i) to (p_i, τ_i, Π_i).

C. Combined ARF-DCF Throughput Computation

With (p_i, τ_i, Π_i) at hand, the main issue involved in computing the combined ARF-DCF throughput is estimating the slot duration when frames collide. Π_i plays an important role in this regard. Let $P_{S(i)}$ $(i = 1, 2, \ldots, L)$ denote the probability that a successful transmission at rate i occurs at a slot in steady-state. We have $P_{S(i)} = N\Pi_i\tau_i(1-\bar{\tau})^{N-1}(1-e_i)$. The probability that a frame transmitted at rate i does not collide but experiences a frame error is $P_{Err(i)} = N\Pi_i\tau_i(1-\bar{\tau})^{N-1}e_i$. The probability that a slot is idle is given by $P_I = (1-\bar{\tau})^N$. Hence the probability that a frame transmission collides is

$$P_C = 1 - P_I - \sum_{i=1}^{L} P_{S(i)} - \sum_{i=1}^{L} P_{Err(i)}.$$

The normalized system throughput contributed by frames transmitted at rate i is given by

$$X_i = \frac{P_{S(i)} S_l}{P_I \sigma + T_S + T_{Err} + T_C} \tag{20}$$

where S_l is the payload size, σ is the slot time, T_S, T_{Err}, and T_C are the average durations of successful, erroneous, and collided transmissions, respectively. Let $X = \sum_{i=1}^{L} X_i$ be the aggregate throughput. T_S and T_{Err} are given by

$$T_S = \sum_{k=1}^{L} P_{S(k)} T_{S(k)}, \quad T_{Err} = \sum_{k=1}^{L} P_{Err(k)} T_{Err(k)},$$

where for the basic access method in DCF (i.e., RTS/CTS is disabled) we have

$$T_{S(k)} = T_{\text{PHY}} + (S_{\text{h}} + S_{\text{l}})/R_k + \text{SIFS} + \text{ACK} + \text{DIFS},$$
$$T_{Err(k)} = T_{\text{PHY}} + (S_{\text{h}} + S_{\text{l}})/R_k + \text{EIFS}.$$

T_{PHY} is the duration of a PHY header and S_h represents the length of a MAC header. What remains is to estimate T_C.

A distinctive feature of multi-rate frame collision is that *the collision time is determined by a frame encoded with the lowest data rate* [5],[29]. Collisions involving $2, 3, \ldots, N$ stations can be approximated by pairwise collisions [5] as they dominate the higher order terms. In a multi-rate system, an additional source of combinatorial explosion arises due to the

different ways N stations using data rates $R_1, \ldots, R_L$ can collide. Let $(n_1, n_2, \ldots, n_L)$ denote the average number of stations using rates $R_1, R_2, \ldots, R_L$ in steady-state. The number of combinations to consider is $(N+L-1)!/N!(L-1)!$ which is unwieldy. We employ a "mean field" approximation as follows. T_C is bounded by

$$P_{\mathrm{C}}\Big(T_{\mathrm{PHY}}+\frac{S_{\mathrm{h}}+S_{\mathrm{l}}}{R_{\mathrm{L}}}+\mathrm{EIFS}\Big) < T_{\mathrm{C}} < P_C\big(T_{\mathrm{PHY}}+\frac{S_{\mathrm{h}}+S_{\mathrm{l}}}{R_{\mathrm{l}}}+\mathrm{EIFS}\big).$$

We split the difference contributed by the two extremes $1/R_L$ (highest rate among colliding frames is R_L) and $1/R_1$ (lowest rate is R_1) by taking the mean

$$T_{\mathrm{C}} = P_{\mathrm{C}}\Big(T_{\mathrm{PHY}} + (S_{\mathrm{h}}+S_{\mathrm{l}})\big(\frac{a_1}{R_1}+\frac{a_2}{R_2}+\cdots+\frac{a_{\mathrm{L}}}{R_{\mathrm{L}}}\big)+\mathrm{EIFS}\Big) \quad (21)$$

where a_i represents the ratio that rate R_i is slowest among the colliding frames. We have $\sum_{k=1}^{L} a_k = 1$.

The steady-state probabilities Π_i are crucial to obtaining a_i. Since $n_i = \Pi_i N$ in steady-state, the mean attempt rate of stations using rate R_i is $\tau_i \Pi_i N$. The total attempt rate of the system is $\sum_{i=1}^{L} \tau_i \Pi_i N$. Hence the ratio of transmission attempts of stations using R_i is

$$c_i = \frac{\tau_i \Pi_i N}{\sum_{k=1}^{L} \tau_k \Pi_k N} = \frac{\tau_i \Pi_i}{\sum_{k=1}^{L} \tau_k \Pi_k}.$$

Ignoring the contribution of three or more stations colliding simultaneously [5] (their contribution is marginal unless N is very large), pairwise collisions are given by c_i^2 (between the same rate R_i) and $2c_ic_j$ for $i \neq j$. Thus

$$a_i = c_i \sum_{j=i}^{L} c_j + (\sum_{j=i+1}^{L} c_j)c_i = c_i^2 + 2c_i \sum_{j=i+1}^{L} c_j \quad (22)$$

for $1 \leq i < L$ and $a_L = c_L^2$. Substituting into (21), we obtain an estimate of T_C. In Section III-E, we show that the combined ARF-DCF model yields accurate prediction of multi-rate 802.11 WLAN performance.

D. Interaction of ARF with RTS/CTS

It has been noted [12], [15] that RTS/CTS (by default disabled) can be useful in helping ARF distinguish noise from collision, both of which result in frame transmission failure. Since an RTS frame is transmitted at the lowest data rate and its size is small, a failure in RTS frame transmission is likely to stem from collision. On the other hand, unsuccessful data frame transmission following a successful RTS/CTS handshake is likely the result of channel noise. A modified ARF that makes use of this information, assuming RTS/CTS is enabled, can help discriminate channel noise from collision. The influence of RTS/CTS on ARF performance is readily captured in our ARF-DCF model by recalculating the expressions of slot durations. We omit the RTS/CTS derivation due to space constraints.

E. Combined ARF-DCF Performance Validation

We evaluate the predictive accuracy of the combined ARF-DCF IEEE 802.11 WLAN model by comparing the analytical results with *ns-2* simulations.

1) ARF Rate Distribution: Section II-C showed performance validation of the ARF Markov chain when p_i, $i \in \{1, \ldots, L\}$, were *given*. Fig. 3 compares *predicted* steady-state rate distribution Π_i in the combined ARF-DCF IEEE

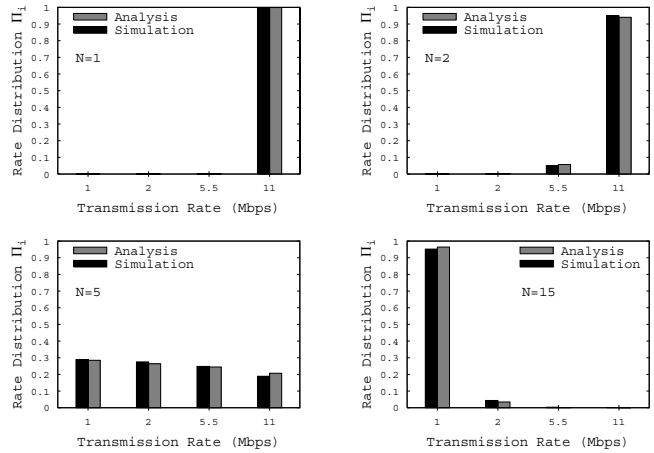

Fig. 3. Rate distribution (Π_i): Combined ARF-DCF analysis vs. simulation (θ_u=10, θ_d=2) for $N = 1, 2, 5$, and 15 in IEEE 802.11b (moderate error).

802.11b model with simulation results. Simulation duration is 1000 seconds. The results show that the combined solution obtained from parameterized coupling of the ARF and DCF chains gives accurate performance predictions.

2) Combined ARF-DCF Throughput: Fig. 4 compares combined ARF-DCF throughput predicted by analysis with simulation as the number of contending stations N is varied. We consider two different channel conditions with different bit error rates: (i) moderate noise where $\mathrm{BER}_{11\mathrm{Mbps}} = 10^{-6}$, and (ii) high noise at which $\mathrm{BER}_{11\mathrm{Mbps}} = 10^{-3}$ (the impact of noise is evaluated in the next section). The resultant FER (i.e., $e_1, e_2, \ldots, e_L$) of each PHY modulation is determined by empirical BER-vs-SNR curves from Intersil [2]. Under

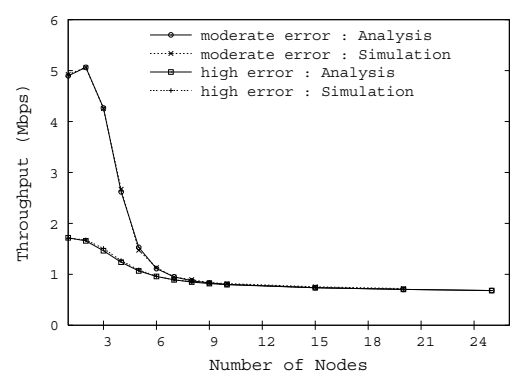

Fig. 4. IEEE 802.11 combined ARF-DCF throughput as a function of contention level: Analysis vs. simulation.

moderate channel noise conditions, we find a bell-shaped curve whose throughput drops precipitously at moderate contention, consistent with empirical performance results from real-world IEEE 802.11b WLANs [7]. The skewness in Π_i caused by

ARF's inability to effectively differentiate channel noise from collision translates to a steep decline in throughput. When channel noise is high, the throughput decline due to ARF's missteps is significantly dampened but still present (down to a factor of 2 from a factor of more than 5). Fig. 4 shows that the combined ARF-DCF model gives accurate quantitative predictions of IEEE 802.11 performance in the presence of cross-layer protocol interactions.

3) Effect of Channel Noise: We evaluate predictive accuracy of the combined ARF-DCF model over a wide range of stationary channel noise. Fig. 5 compares throughput from analysis and simulation as a function of SNR with and without RTS/CTS when $N = 1$ (to remove multiple access contention). The fixed-rate curves for 1, 2, 5.5, and 11 Mbps from physical measurements are shown for reference. We observe that predictive accuracy remains high over a wide SNR range. The influence of RTS/CTS is also easily accounted for in the combined ARF-DCF model.

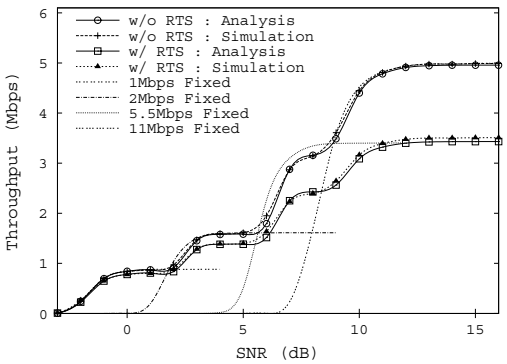

Fig. 5. IEEE 802.11b combined ARF-DCF throughput: Analysis vs. simulation ($\theta_u = 10$, $\theta_d = 2$) with and without RTS/CTS.

4) Model-based ARF Calibration: Fig. 6 shows ARF-DCF throughput for different combinations of the up/down thresholds under moderate channel noise as contention level N is varied. We observe that the default values $\theta_u = 10$ and $\theta_d = 2$ implemented in WLAN cards performs worst due to the excessive conservativeness. When asymmetry is instituted in the opposite direction, i.e., $\theta_u = 2$ and $\theta_d = 10$, throughput significantly improves. The optimal values of θ_u and θ_d depend on channel SNR. Since the IEEE 802.11 multi-rate WLAN is accurately modeled by our combined ARF-DCF chain, the optimal control problem may be tackled using a Markov decision process (MDP) formulation. Solution of the MDP is left for future work.

IV. Scalable Performance of TCP-over-WLAN

The previous section analyzed the dynamic interaction between ARF and DCF over a multi-rate PHY layer following a station-centric Markov chain approach that achieved accurate, tractable analysis through modular coupling. In this section, we extend the approach with the aim of understanding the surprisingly scalable performance of TCP-over-WLAN which also has a strong mitigating influence on ARF.

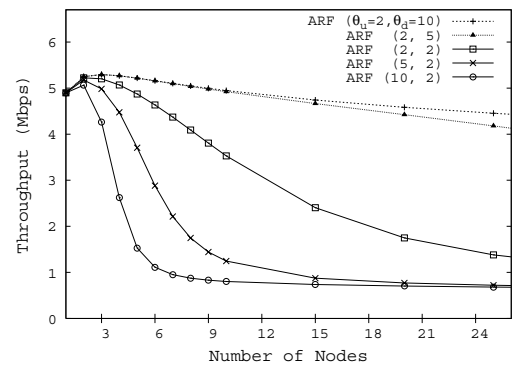

Fig. 6. ARF-DCF throughput for various θ_u and θ_d combinations.

A. Markov Chain Analysis of TCP-over-WLAN

IEEE 802.11 WLANs are predominantly used to facilitate wireless Internet access, and like their wireline brethren [19], the bulk of the workload is TCP-mediated client/server file traffic [23]. We consider an IEEE 802.11 infrastructure WLAN where N wireless stations access the Internet through a shared AP. We assume a typical client/server environment where each station downloads files mediated by TCP (mostly HTTP traffic), which means that the bulk of data traffic flows downstream—from server through AP to wireless client—and the bulk of upstream traffic is TCP ACK traffic. We ignore differences in wireline bandwidth and delay from AP to servers which can cause TCP unfairness issues on the wireline side (an orthogonal issue). We assume each station uses TCP NewReno without SACK and delayed ACK is disabled. The traffic rate (counted in packet unit) in steady-state contributed by the AP to multiple access contention vis-à-vis a wireless station is $N : 1$. Under heavy traffic conditions (e.g., N large), the AP's downstream buffer is constantly backlogged. A wireless station's egress buffer is empty until a data frame is received from the AP which triggers an upstream TCP ACK response. Enabling delayed ACK only changes the traffic ratio to $2N : 1$. Note that the traffic ratio formulation is robust with respect to details in congestion control.

Next, we analyze the single-rate environment which focuses on scalable TCP-over-DCF performance. This is followed by the multi-rate case where the mitigating influence of TCP-over-DCF on ARF is considered.

1) Analysis of Single-rate Environment: We define a station-centric TCP-over-DCF Markov chain $x(t)$ over the state-space of backlogged frames awaiting IEEE 802.11 MAC transmission. In the wireless Internet access context, $x(t) = k$ if there are k outstanding TCP ACK packets awaiting transmission at client side. Similarly for the AP due to symmetry—there is no distinction between AP and wireless station in DCF when engaging in CSMA competition—albeit for TCP data packets. Let λ_g, μ_g denote the birth and death rate of the AP's (g=ap) and a station's (g=sta) egress buffer, respectively. The key difference between AP and wireless station is the $N : 1$ traffic ratio which leads to different transition rates λ_{ap}, μ_{ap},

λ_{sta}, μ_{sta}. Assuming homogeneous clients, due to DCF's symmetry the probability that a station receives a TCP packet from the AP in a given slot is $\lambda_{sta} = \frac{1}{N}\tau_{ap}(1-p_{ap})$ where τ_{ap} is the AP's attempt rate and p_{ap} is its frame transmission failure rate. The departure rate is given by $\mu_{sta} = \tau'_{sta}(1-p_{sta})$ where τ'_{sta} is the conditional attempt rate of a station when it has a packet to transmit. A station's overall attempt rate is given by

$$\tau_{sta} = \sum_{k=1}^{\infty} \pi_{sta,k}\tau'_{sta} = \tau'_{sta}(1-\pi_{sta,0})$$

where $\pi_{sta,k}$ is the steady-state probability that a station has k backlogged packets. In an $M/M/1$ birth-death chain, $\pi_{sta,k} = (1-\rho)\rho^k$ where $\rho = \lambda_{sta}/\mu_{sta}$, which yields

$$\tau_{sta} = \frac{\tau_{ap}(1-p_{ap})}{N(1-p_{sta})}. \tag{23}$$

Eq. (23) can serve as a heavy traffic approximation (i.e., $N \to \infty$ and $\rho \to 0$) when the number of contending stations N is large (a light traffic approximation from a client's queue perspective).

Due to constant backlog, the AP's attempt rate follows the solution of the DCF chain (Eq. (17) in Section III-B),

$$\tau_{ap} = \frac{2(1-2p_{ap})}{(1-2p_{ap})(W_0+1) + p_{ap}W_0(1-(2p_{ap})^m)}. \tag{24}$$

The transmission failure probabilities are given by

$$\begin{aligned} p_{ap} &= 1-(1-\tau_{sta})^N(1-e_{ap}), \\ p_{sta} &= 1-(1-\tau_{ap})(1-\tau_{sta})^{N-1}(1-e_{sta}), \end{aligned} \tag{25}$$

where e_{ap} and e_{sta} are the channel error probabilities. (23)-(25) are four equations in four unknowns (τ_{ap}, τ_{ap}, τ_{sta}, and τ_{sta}) which can be solved using fixed-point techniques.

2) Analysis of Multi-rate Environment: ARF enters into the picture through $\Pi_{g,i}$, $g \in \{sta, ap\}$, $i \in \{1,\ldots,L\}$, which affects average transmission rate

$$\bar{\tau}_g = \sum_{i=1}^{L} \Pi_{g,i}\tau_{g,i}, \qquad g \in \{sta, ap\}.$$

Since the transmission failure probabilities $p_{g,i}$ are given by (Eq. (19) in Section III-B),

$$\begin{aligned} p_{ap,i} &= 1-(1-\bar{\tau}_{sta})^N(1-e_{ap,i}), \\ p_{sta,i} &= 1-(1-\bar{\tau}_{ap})(1-\bar{\tau}_{sta})^{N-1}(1-e_{sta,i}), \end{aligned}$$

we use modular coupling from ARF-DCF analysis to solve the fixed-point problem with increased dimension introduced by $\Pi_{g,i}$. Throughput calculation for TCP-over-DCF with ARF follows the steps described in Section III-C, accounting for differences in TCP data and ACK frame sizes.

Another important quantity in TCP-over-WLAN dynamics is the average number of active (i.e., backlogged) stations. Its small value—resulting from the $O(1/N)$ negative drift in the backlog chain—is the primary reason for scalable TCP-over-DCF performance, which also mitigates the detrimental effect of ARF. The active station count in equilibrium is given by $N_{active} = N_{ap} + N_{sta} = 1 + N(1-\pi_{sta,0})$.

B. TCP-over-WLAN Performance Validation

1) Scalable TCP-over-DCF Throughput: Fig. 7 shows TCP-over-DCF throughput (both data and ACK traffic) as the number of contending wireless stations N is increased from 1 to 50. The results from TCP-over-DCF Markov chain analysis accurately predict throughput obtained from *ns-2* simulation. Even with 50 contending stations, there is hardly any throughput decline. The key reason lies in the $O(1/N)$-factor difference between λ_{sta} and μ_{sta} in a station's TCP-over-DCF birth-death chain, which imparts a strong negative drift for large N that inhibits the growth of a station's backlog queue in the presence of heavy multiple access contention.

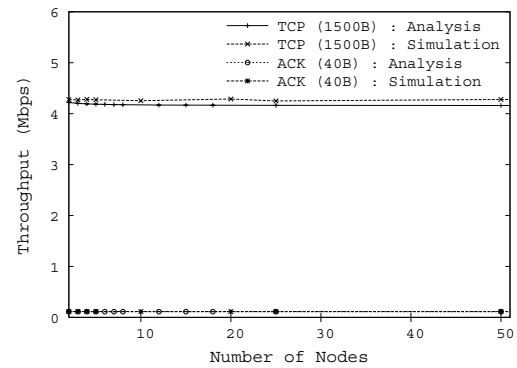

Fig. 7. TCP-over-DCF throughput as a function of contending stations N.

From a "station versus AP perspective," the $O(1/N)$-factor difference between τ_{sta} and τ_{ap} captures the fact that an AP—because it competes equally with stations under DCF—is unlikely to win a multiple access competition when N is large. The larger the number of backlogged stations, the smaller the probability that the AP will succeed in transmitting a data packet to a station, thereby increasing the backlog queue of an already backlogged station or increasing the number of backlogged stations. This effect can be seen in Table I which shows the average number of backlogged stations N_{active} as a function of N. The results from TCP-over-DCF analysis, which accurately agree with simulation results, show that the average number of backlogged stations in equilibrium is a little over 2 even for large N.

TABLE I
BACKLOGGED STATIONS AS A FUNCTION OF CONTENDING STATIONS N

N	2	5	10	20	50	100
Analysis	2.07	2.11	2.12	2.13	2.14	2.14
Simulation	2.14	2.16	2.17	2.20	2.23	2.26

This implies that the *effective multiple access contention* experienced by TCP-over-DCF remains invariant at 2–3 stations resulting in low collision and high system throughput.

2) TCP's Mitigating Effect on ARF: Fig. 8 shows the mitigating effect of TCP-over-DCF on ARF which, unlike the bell-shaped ARF-DCF throughput curves in Figs. 4 and 6 without TCP, remain flat as contention level N is increased.

The main reason for ARF not getting miscued by collision is the low collision rate afforded by scalable TCP-over-DCF throughput which makes consecutive collision unlikely. Even when they occur, recovery is fast. This allows ARF to focus on frame errors stemming from channel noise in accordance with its intended design. TCP over DCF and ARF performance analysis also shows that the mitigating influence of TCP is insensitive to details in the underlying ARF parameters.

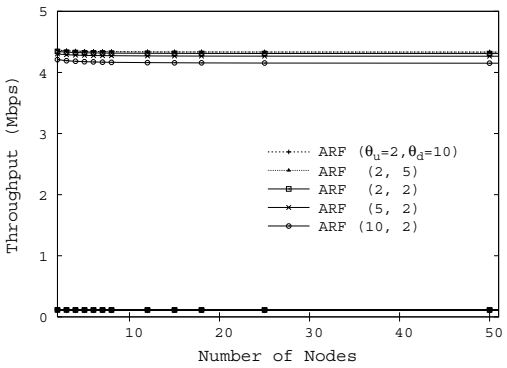

Fig. 8. Scalable TCP-over-WLAN throughput in the presence of ARF.

V. Conclusion

In this paper, we have advanced station-centric Markov chain models of protocol interaction between ARF, DCF, and TCP in multi-rate IEEE 802.11 WLANs. We have shown that the performance analyses accurately predict cross-layer WLAN performance, including fine-granular station dynamics and coarse-granular throughput, capturing subtle system traits such as bell-shaped throughput curve under ARF-DCF and scalable throughput performance of TCP-over-WLAN. This is achieved through modular coupling which facilitates tractable analysis by curbing the state-space explosion of direct analysis without sacrificing accuracy. Our modular station-centric performance analysis spanning transport, MAC, and PHY layers may be viewed as a multi-protocol extension of Bianchi's IEEE 802.11 DCF model, which demonstrates the versatility and efficacy of station-centric Markov chain modeling for capturing cross-layer WLAN dynamics. Several avenues for future work remain, including solution of an optimal ARF design problem using a MDP approach enabled by the ARF-DCF chain, and multi-WLANs using a hierarchical extension of modular station-centric Markov chain analysis.

VI. Acknowledgments

The authors would like to thank Sunwoong Choi of Samsung Electronics for his valuable discussion. This research was supported, in part, by NSF grant ANI-0082861, and a BK-21 grant. Part of the research was carried out while Jaehyuk Choi was visiting the Network Systems Lab at Purdue University.

References

[1] IEEE, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Std 802.11-1999, 1999.

[2] Intersil, "HFA3861B; Direct Sequence Spread Spectrum Baseband Processor," 2000.

[3] http://www.enterasys.com/, online link.

[4] H. Balakrishnan, V. Padmanabhan, S. Seshan, and R. Katz, "A comparison of mechanisms for improving TCP performance over wireless links," *IEEE/ACM Trans. Networking*, 5(6):756–769, 1997.

[5] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordinated Function," *IEEE J. Selected Areas in Commun.*, vol. 18, no. 3, pp. 535-547, 2000.

[6] P. Chevillat, J. Jelitto, A. N. Barreto, and H. Truong, "A Dynamic Link Adaptation Algorithm for IEEE 802.11a Wireless LANs," in *Proc. IEEE ICC03*, pp. 1141-1145, 2003.

[7] S. Choi, K. Park, and C. Kim, "On the Performance Characteristics of WLANs: Revisited," in *Proc. ACM SIGMETRICS'05*, pp. 97-108, 2005.

[8] Z. Hadzi-Velkov and B. Spasenovski, "An analysis of CSMA/CA protocol with capture in Wireless LANs," in *Proc. IEEE WCNC'03*, vol.2 pp. 16-20, 2003.

[9] W. Haitao, P. Yong, L. Keping, C. Shiduan, and M. Jian, "Performance of reliable transport protocol over IEEE 802.11 wireless LAN: Analysis and enhancement," in *Proc. IEEE INFOCOM '02*, pages 599–607, 2002.

[10] I. Hatatcherev, K. Langendoen, R. Lagendijk, H. Sips, "Hybrid rate control for IEEE 802.11," in *Proc. ACM MobiWac'04*, pp.10-18, 2004.

[11] M. Heusse, F. Rousseu, G. Berger-Sabbatel, and A. Duda, "Performance Anomaly of 802.11b," in *Proc. IEEE INFOCOM'03*, pp. 836-843, 2003.

[12] C. Hoffmann, M. H. Manshaei, T. Turletti, "CLARA: Closed-Loop Adaptive Rate Allocation for IEEE 802.11 WirelessLANs," in *Proc. IEEE WIRELESSCOM'05*, 2005.

[13] V. Joel. "Exploding the myth of WLAN performance," Telephony Online, 2004.

[14] A. Kamerman and L. Monteban, "WaveLAN 2: A High-performance Wireless LAN for the Unlicensed Band," *Bell Labs Tech. Journal*, 1997.

[15] J. Kim, S. Kim, S. Choi, and D. Qiao,"CARA: Collision-Aware Rate Adaptation for IEEE 802.11 WLANs,"in*Proc. IEEE INFOCOM'06*, 2006.

[16] M. Lacage, M. H. Manshaei, T. Turletti, "IEEE 802.11 rate adaptation: a practical approach," in *Proc. ACM MSWiM'04*, pp. 126-134, 2004.

[17] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP throughput: A simple model and its empirical validation," in *Proc. ACM SIGCOMM '98*, pp. 303–314, 1998.

[18] Q. Pang, Leung, V.C.M. Liew, S.C. Liew, "A rate adaptation algorithm for IEEE 802.11 WLANs based on MAC-layer loss differentiation," in *Proc. IEEE Broadband Networks'05*, pp. 709-717, 2005.

[19] K. Park and W. Willinger, "Self-similar network traffic: An overview," in K. Park and W. Willinger, editors, *Self-Similar Network Traffic and Performance Evaluation*. Wiley-Interscience, 2000.

[20] K. Pentikousis, "TCP in wired-cum-wireless environments," *IEEE Communications Surveys*, Fourth Quarter(7):2–14, 2000.

[21] D. Qiao and S. Choi, "Fast-Responsive Link Adaptation for IEEE 802.11 WLANs," in *Proc. IEEE ICC'05*, pp. 3583-3588, 2005.

[22] G. Sharma, A. Ganesh and P. Key,"Performance analysis of Contention Based Medium Access Control Protocols,"in *Proc. IEEE INFOCOM'06*.

[23] D. Tang, M. Baker, "Analysis of a Local-Area Wireless Network," in *Proc. ACM MobiCom'00*, pp. 1-10, 2000.

[24] A. Vasan and U. Shankar. "An empirical characterization of instantaneous throughput in 802.11b WLANs," *Technical Report CS-TR-4389, UMIACS-TR-2002-69, Department of Computer Science and UMIACS, University of Maryland College Park*, 2002.

[25] H. Wu, Y. Peng, K. Long and J. Ma, "Performance of Reliable Transport Protocol over IEEE 802.11 Wireless LAN: Analysis and Enhancement," in *Proc. IEEE INFOCOM'02*, vol.2, pp. 599-607, 2005.

[26] Y. Xiao, "An Analysis for Differentiated Services in IEEE 802.11 and IEEE 802.11e Wireless LANs," in *Proc. ICDCS'04*, 2004.

[27] G. Xylomenos and G. Polyzos, "TCP and UDP performance over a wireless LAN," in *Proc. IEEE INFOCOM '99*, pp. 439–446, 1999.

[28] G. Xylomenos, G. Polyzos, P. Mähönen, and M. Saaranen, "TCP performance issues over wireless links," *IEEE Communications Magazine*, vol. 39(4), pp. 52–58, 2001.

[29] D. Yang, T. Lee, K. Jang, J. Chang and S. Choi,"Performance Enhancement of Multi-Rate IEEE 802.11 WLANs with Geographically-Scattered Stations," *IEEE Trans. Mobile Computing*,vol. 5, no. 7, pp. 906-919, 2006.

[30] E. Ziouva and T. Antonakopoulos, "CSMA/CA performance under high traffic conditions: throughput and delay analysis," *Computer Communications*, vol. 25 , pp.313-321, 2002.