

Research Article

Bandwidth Optimization in Centralized WLANs for Different Traffic Types

R. J. Haines, N. Fanning, T. Lewis, and J. Coon

Telecommunications Research Laboratory, Toshiba Research Europe Ltd., 32 Queen Square, Bristol BS1 4ND, UK

Received 31 May 2006; Revised 24 November 2006; Accepted 10 January 2007

Recommended by Wei Li

Allocating bandwidth between different forms of coexisting traffic (such as web-browsing, streaming, and telephony) within a wireless LAN is a challenging and interesting problem. Centralized coordination functions in wireless LANs offer several advantages over distributed approaches, having the benefit of a system overview at the controller, but obtaining a stable configuration of bandwidth allocation for the system is nontrivial. We present, review, and compare different mechanisms to achieve this end, and a number of different means of obtaining the configurations themselves. We describe an analytical model of the system under consideration and present two mathematical approaches to derive solutions for any system configuration and deployment, along with an adaptive feedback-based solution. We also describe a comprehensive simulation-based model for the problem, and a prototype that allows comparison of these approaches. Our investigations demonstrate that a self-adaptive dynamic approach far outperforms any static scheme, and that using a mathematical model to produce the configurations themselves confers several advantages.

Copyright © 2007 R. J. Haines et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

The IEEE802.11 protocols [1] have become the dominant standard for wireless local area networks (WLANs). These protocols have evolved to support a variety of traffic types, which benefit from different scheduling and control mechanisms. There are two common forms of traffic encountered by WLANs. The first is sporadic, bursty data traffic, which is most efficiently served by highly distributed contention-based access schemes. The second is traffic with stringent quality-of-service (QoS) requirements, such as bandwidth, delay and/or jitter, which needs a more structured approach to provide guarantees of access.

There are two complementary approaches to serving QoS-sensitive traffic: distributed and centralized access [2]. Distributed approaches have largely focused on differentiated access that prioritizes different traffic types but then relies on statistical guarantees of access for each priority level, although more recent developments also incorporate a distributed reservation mechanism [3]. Centralized approaches, where a central controller allocates resources, benefit from having a global view of the entire system, and from being able to concentrate complexity in a single (more feature-rich, more expensive, higher-powered) device.

The IEEE802.11 standards offer both centralized and distributed controls. In this work, we concentrate on the centralized point coordination function (PCF), which can be seen as a specialized case of the more flexible and complex hybrid coordination function (HCF) of IEEE802.11e [4]. These centralized approaches are most attractive in single-access-point scenarios such as commonly found in the home, as there are scheduling complexities that arise with multiple-access-point scenarios. The PCF allows the coexistence of both QoS-sensitive traffic and bursty data traffic through the polling of the former and the direct contention of the latter. This is achieved by overlaying a repeating time-division superframe onto the medium, with distinct phases for polled and contending traffic.

The configuration of this superframe directly affects the system's ability to support the two types of traffic effectively. If the configuration is badly wrong, then the QoS requirements may be missed, or the data traffic starved of access. Balancing these two competing classes of traffic in an optimal way is the fundamental subject of this work.

Published work in this specific area of configuring the superframe has, to date, relied on empirical, simulation-based studies of different scenarios to derive lookup tables

of superframe configurations [5]. This work improves upon these studies with a more comprehensive and accurate simulation model, and then goes on to propose novel solutions to this problem that have sound mathematical foundations and offer a more dynamic approach. This more flexible and adaptable approach allows a continuous optimized set of superframe parameters to be derived and the more theoretical basis permits greater confidence in the optimal nature of the values being employed than is possible with purely experimental results.

This paper is structured as follows: in Section 2 the IEEE802.11 PCF is explained to give a background to this problem area. In Section 3 we examine related work in this area and highlight how this contribution differs from, and improves upon, what has gone before. Section 4 presents our simulation model that improves upon that in the literature, whilst Sections 5 and 6 describe our mathematical approaches to this problem. In Section 7 we describe a simulation prototype that allows direct comparison of all of these approaches. Finally, in Section 8, we conclude this paper.

2. IEEE802.11 CENTRALIZED CONTROL

The IEEE802.11 standard [1] was created as a wireless alternative to wired local area networks (LANs), which at that time were predominately deployed in office environments to carry internet data traffic. Nonetheless, even at that time, it was recognized that support for QoS-sensitive traffic would be required. To achieve this, two complementary access schemes were specified, the best-effort contention-based distributed coordination function (DCF) for delay-insensitive traffic, and the optional centralized polling-based point coordination function (PCF) for time-bounded traffic, such as audio/video streams and voice over internet protocol (VoIP) traffic.

DCF is the mandatory access mechanism in IEEE802.11. For sporadic bursty data traffic, this offers a very efficient means of access: devices (stations, STA, in IEEE802.11 parlance) can compete for access to the medium as soon as they have a packet to transmit. The underlying access scheme is carrier-sense multiple access with collision avoidance (CSMA/CA). Multiple access and collision avoidance are achieved with a combination of prerequisite quiet periods on the medium (hence the carrier sense) followed by random backoffs to avoid collisions. The durations of the quiet periods (termed interframe spaces) prioritize access onto the medium. For example, the shortest interframe space (short interframe space, SIFS) is used between the transmission of a packet and the transmission by the receiving station of its acknowledgment. Transmission of this acknowledgement has the highest priority of any packet (as it is the only means by which the transmitting station can be aware of successful delivery, and therefore not retransmit the original packet), so it is allowed onto the medium with the shortest possible interframe space following the end of the original packet transmission. Stations newly contending for access must wait for a much longer interframe space (the DCF interframe space,

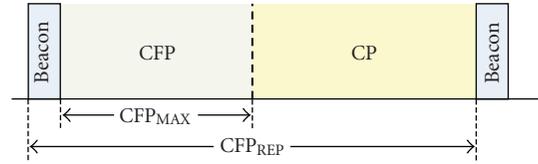


FIGURE 1: Superframe structure.

DIFS) before even being able to contend for access with the random backoff procedure.

However, for QoS-sensitive traffic where a packet must be sent at a guaranteed time, contending for access (and potentially losing) with every packet quickly becomes impossible under all but the lightest of network loads. To guarantee packet transmission, reservation and polling schemes must be considered. In these cases, the additional overhead of reserving a transmission in advance becomes acceptable. The centralized PCF of the original IEEE802.11 standard, and its progeny, the hybrid coordination function (HCF) in the IEEE802.11e standard [4], both introduce centralized coordination of resources to allow this QoS-sensitive traffic to coexist alongside contention-based data exchanges. The difference between the two is that HCF allows a more flexible allocation of transmission opportunities compared to PCF, although this is at the cost of increased complexity.

Centralized coordination imposes a time-based repeating superframe onto the medium (as illustrated in Figure 1), characterized by the transmission of a broadcast beacon, followed by a contention-free (polled) period (CFP) and then a contention-based access period (CP). The process of overlaying this structure onto the otherwise anarchic access mechanism of DCF is possible through the aforementioned interframe spaces: the central controller is able to use the PCF interframe space (PIFS), of shorter duration than the DIFS, to preempt contending stations and seize the medium to begin the superframe.

The structure of the superframe is determined by two parameters, its duration and the proportion of time spent in the contention-free phase. This duration (i.e., the beacon and CFP repetition rate) and the relative size of the CFP to the rest of the superframe, typically termed CFP_{REP} and CFP_{MAX} , respectively, are both configurable by the point controller (PC) entity located at the access point (AP). These two values are broadcast in the beacon to all stations.

These parameters determine the success of a given WLAN deployment from the perspective of the polled traffic, the contention-based traffic, or both. A badly configured system will fail to deliver the performance that the end user has the right to expect, irrespective of the headline data rate of the product.

3. RELATED WORK

The distributed approach to serving QoS-sensitive traffic has been closely studied in recent years, both in the guise of the enhanced distributed channel access (EDCA) subset of the

IEEE802.11e HCF [4] and in the WiMedia MAC [3] (formed from one of the survivors of the now-defunct IEEE802.15.3a standard). The latter offers extensions to the IEEE802.11e EDCA subset including a fully distributed solution including both hard and soft reservations of slots (soft reservation being the ability for a station to tentatively reserve a slot, and for it to be made available for other stations if unused). The performance of the WiMedia MAC has been evaluated, and the soft-reservation scheme is found to be particularly efficient [6]. A number of extensions and enhancements to these distributed schemes have been proposed from a number of different perspectives, the sheer number of which suggesting that there are several shortcomings to this approach. These extensions have included the use of admission control [7] by the higher layers, the addition of hybrid automatic repeat request (ARQ) mechanisms [8] and variable backoffs (contention windows) [9, 10] to the MAC protocol, and cross-layer schemes linking the differentiated access categories to the modulation and coding schemes of the physical layer [11].

The centralized approach has been less well studied, often because a distributed solution is viewed as being inherently more scalable and less complex [12]. However, under heavy and asymmetric loads such as would result from streaming high-definition television and similar demanding applications, it has been observed that the distributed approach results in a severe impact on the coexisting traffic streams [13, 14]. The complexity of the 802.11e HCF scheme has been highlighted as an issue, and an enhanced PCF (EPCF) has been proposed [15] to address some issues with PCF that HCF also addresses, whilst not imposing all of the complexity of HCF.

A self-adaptive scheme to configure the PCF superframe has been proposed [5]. This proposed scheme selects parameters from predefined lookup tables indexed by a quantized number of active polled stations and stepped values for the maximum allowable delay of the applications. The values populating the lookup tables are derived through experimental simulation results, which result in values of an almost random nature, as depicted in Figure 2.

These results do not take into account the minimum CFP and CP sizes mandated by the standard [1, 16], and crucially, there is no means of generating values outside of the simulation scenarios considered. Nonetheless, these values provide a valuable benchmark for the approaches considered herein. The traffic considered in this benchmark study is a combination of data and VoIP flows, an important area for investigation as internet telephony applications continue to gain popularity.

4. IMPROVED EMPIRICAL RESULTS

An improved (standard-compliant) simulation model, using the configuration proposed in [5], has been developed in OPNET.

The network model is constrained to 16 STAs and an AP throughout the study presented herein, with all stations located within a 300 m diameter. All 16 STAs produce voice

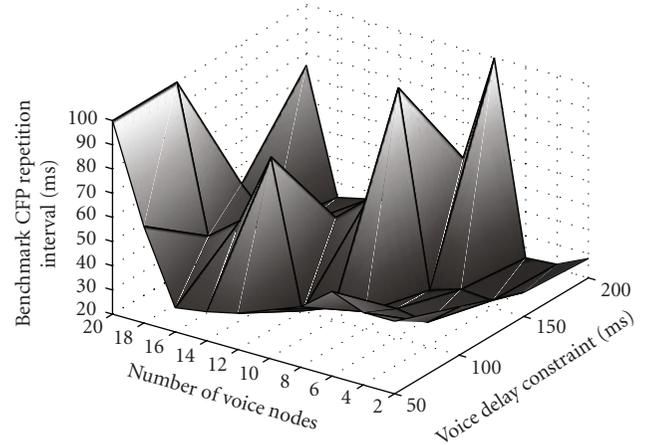


FIGURE 2: Benchmark CFP_{REP} values.

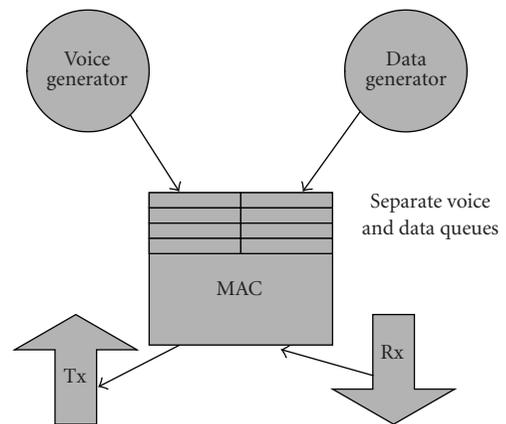


FIGURE 3: Node model.

traffic but only 6 of them produce data traffic. The PC function is performed in the AP which is the destination for all transmissions. The AP transmits only MAC control and management frames, such as ACKs, polls, and beacons.

An STA based on a generic node model (Figure 3) generates voice and potentially data application traffic along with the necessary MAC control frames. The different traffic streams are buffered in individual queues until the frames are transmitted. The data queue is served during the CP and the voice queue is served during the CFP. The interval between successive data MAC service data unit (MSDU) generations varies exponentially with a mean of 7.5 frames per second (fps). The data MSDUs vary exponentially in size with a mean of 1000 bytes. Brady's model [17] is employed for the voice traffic generator, which produces 200-byte MSDUs. To prevent idle CFPs and sudden traffic surges, the start times of the voice generators are random over the first two seconds of the simulation.

The AP model, which is based on the generic node model, controls the CFP with the transmission of beacons, polls, and CFP end (CF-END) frames using the PC function.

TABLE 1: Summary of model parameters.

Parameter	Value	Parameter	Value
Slot	20 μ s	Mean data MSDU	1 kbyte
SIFS	10 μ s	Mean data rate	7.5 fps
PIFS	30 μ s	Voice MSDU	200 bytes
DIFS	50 μ s	Voice mean on : off	1 s : 1.35 s
CW _{MIN}	31 slots	Voice on rate	64 kbps
PLCP time	192 μ s	Beacon	160 bytes
MAC header	28 bytes	ACK	14 bytes
Data rate	2 Mbps	Poll\CF _{end}	20 bytes
Control rate	1 Mbps	Queue sizes	250 Kbits

The AP responds to received data MPDUs with acknowledgements (ACKs) during the CPs. The QoS performance is also measured in the AP model as it provides sinks for the two types of traffic. The polling list, which consists of all 16 STAs, is cycled through continuously during the CFP. When a voice MPDU has been received in response to a poll frame, the AP acknowledges its reception in the proceeding poll frame by setting the frame type field to be a combined poll and acknowledgment. If a node does not have any voice packets queued when polled, it responds with a null data frame. At the beginning of a CFP, the polling is resumed where the previous CFP ended. If sufficient time remains in a CFP after all nodes have been polled, the polling cycle begins again. Intelligent polling schemes, such as biasing the polling to nodes that did not previously respond [18–20], are not utilized in this study. A check is made to ensure that sufficient time remains in the CFP to accommodate a polled voice frame exchange (i.e., poll + voice MSDU + 2SIFS + CF-END) prior to every poll transmission. An early CF-END is transmitted if insufficient time remains.

No check is made during the CP to ensure that the DCF access mechanism frame exchange sequence (DIFS + CW + data + SIFS + ACK) will be complete before the next expected beacon transmission. This will occasionally result in CP stretching which will shorten the duration of the proceeding CFP.

An IEEE802.11b physical layer (PHY) is assumed as this provides a fair comparison with the referenced work in this area. The fundamental behavior of a MAC is largely independent of the PHY technology, and when performing comparisons between different MAC solutions, the specifics of the PHY are not particularly relevant. The physical layer is modeled so that packet losses due to link errors do not occur. Packet losses occur due to collisions only, and so observations on the performance can be described purely in terms of MAC behavior. It is also assumed that there are no hidden stations, the capture effect does not occur, and none of the stations are in power-saving mode. The model parameters are summarized in Table 1.

TABLE 2: Simulated CFP_{MAX} and CFP_{REP} values.

Parameter	Values
CFP _{MAX} (%)	5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95
CFP _{REP} (ms)	50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250

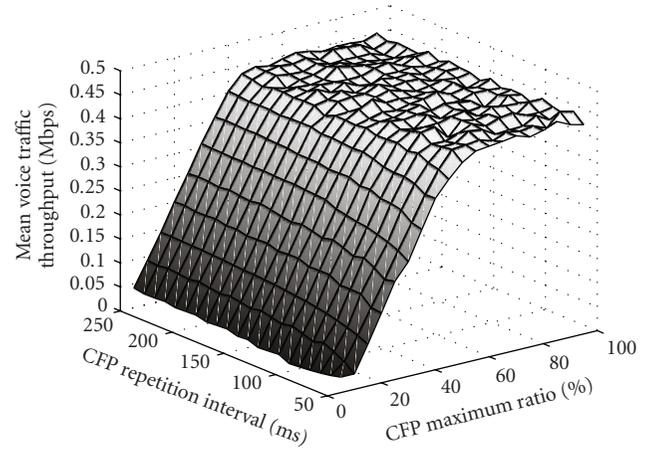


FIGURE 4: Mean voice throughputs.

Generally, the voice traffic has the more stringent performance requirements of the two traffic types. Therefore, the performance of the CFP, and that of the associated voice traffic, is focused on in the presentation of the results. Failure to satisfy these requirements results in wasted transmissions as packets received outside of the QoS constraints will probably be dropped at the transport or application layer. The approach taken is to determine how to configure the system so that the time-dependent voice traffic is satisfied whilst ensuring that the maximum possible amount of medium time remains for data traffic.

Simulations have been performed for all permutations of the CFP_{MAX} and CFP_{REP} settings contained in Table 2. This provided 399 simulations each covering 5 minutes of simulated time. However, some of the CFP_{MAX} and CFP_{REP} combinations will result in CP and CFP durations that are less than the minimum mandated by the standard. These invalid permutations can be discounted at a later stage.

The first set of simulation results is the mean voice traffic throughputs, which are illustrated in Figure 4. Sixteen STAs produce approximately 435 kbps of voice traffic within the network. The voice traffic throughput results show that the CFP_{MAX} value has to be around 45% and above so that all of the voice traffic generated can be accommodated.

It is not sufficient to concentrate solely on providing the necessary resource to accommodate all of the voice traffic to produce a successful system. The delay that is experienced is arguably more important for time-dependent voice services. The mean delays experienced by the voice traffic during the

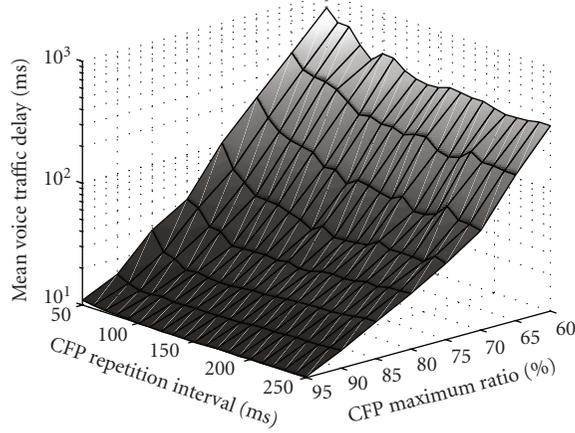


FIGURE 5: Mean voice delays.

simulations are illustrated in Figure 5. CFP_{MAX} values below 60% incur significant delays so only a subset of the CFP_{MAX} results is included. The CFP_{MAX} value of 45% suggested by the mean voice throughput results will result in voice delays in excess of three seconds, which is unacceptable for telephony services. Voice transmission requires delays below 25 milliseconds if echo cancellation is not available, 150 milliseconds for high quality with echo cancellation, and 400 milliseconds for acceptable quality with echo cancellation [21]. The results show that CFP_{MAX} values in the region of 70% and above are required to achieve mean delays below 150 milliseconds.

The mean voice delay results can generate a lookup table to select CFP_{MAX} and CFP_{REP} values that result in a given delay. They can also predict the performance of a particular superframe configuration generated by an optimization algorithm. This allows different optimization techniques to be compared. The most interesting observation of Figure 5 is the apparent immunity to CFP_{REP} variations that the near horizontal contours suggest.

Despite having similar mean delays, the probability density functions (PDFs) of instantaneous voice packet delays for given CFP_{REP} values are quite different. Figure 6 illustrates the distribution of delays that were experienced for a subset of the CFP_{REP} values with a constant CFP_{MAX} of 70%. This value of CFP_{MAX} provides mean delays in the region of 150 milliseconds. The distributions contain two peaks, the first occurring at $(nodes/2) \times \text{polled-exchange duration}$ and the second occurring at $CFP_{REP} \times (1 - CFP_{MAX})$. The former occurs due to the average wait experienced during a polling period, equal to half the time to poll all twelve stations and the latter due to packets having to wait for a CP to pass.

Figure 7 presents the cumulative distribution functions (CDFs) of the voice packet delays, and shows the percentage that satisfies a given delay constraint. A CDF is required if the maximum instantaneous delay is the important performance parameter. The CDF can predict the percentage of frames that may be dropped due to the delay constraints not being met. For a 400-millisecond instantaneous delay thresh-

old, a CFP_{MAX} setting of 70% requires a CFP_{REP} in the region of 170 milliseconds and above. This will provide a voice service of acceptable quality only if echo cancellation is included [21, 22]. The CDFs illustrate that delay distributions can be highly CFP_{REP} sensitive in certain regions. Figure 7 shows that the percentage of packets within the constraint of 100-millisecond maximum delay varies from 65 to 85 depending on the CFP_{REP} setting.

Focusing on the CFP and its associated voice traffic prevents valuable medium time from being wasted. However, it is also important to understand the effect of superframe configuration on the CP and the associated data traffic. Biassing resource allocation to the voice traffic is only sensible to the point where the voice services have their QoS constraints satisfied. Further biasing in the direction of voice traffic provides no noticeable improvements in the performance of voice services but it results in a noticeable degradation of the data services.

The data traffic throughput results, illustrated in Figure 8, show that values of CFP_{MAX} below approximately 80% are required to support all of the data traffic (360 Kbps) generated in the given scenario. The CDF of instantaneous voice traffic delays, Figure 7, has demonstrated that for 70% CFP_{MAX} , a minimum CFP_{REP} of 170 milliseconds is required for acceptable voice transmission. This superframe configuration provides sufficient CP capacity to fully accommodate the generated data traffic. Higher-quality voice transmissions demanding delays in the region of 150 milliseconds will require the superframe configuration to be biased further in favor of the CFP. CFP_{MAX} values in excess of 80% will reduce the amount of data traffic that can be supported. Reducing the proportion of medium time available for the CP increases the likelihood of CP stretching as there is a greater probability that data packets will be awaiting transmission at the end of the CP. This CP stretching will have a negative impact on the CFP albeit smaller than the positive impact of increasing the amount of resource allocated to the CFP.

5. NONLINEAR OPTIMIZATION

The first mathematical technique we propose as a candidate solution as a verifiable theoretical model is that of nonlinear optimization of an abstracted model of the data exchanges on the superframe [23]. Nonlinear optimization theory provides a number of means to optimize a number of variable parameters to provide a stable system solution. These techniques have been applied to a number of areas within communications, including wireless sensor network access [24] and deriving training sequences for orthogonal-frequency division multiplexing (OFDM) systems [25]. We use the barrier method [26] in this work.

No matter how robust the mathematical analysis technique adopted, its success is, of course, dependent on how closely the model being analyzed resembles reality. In the case of nonlinear optimization, this means that the formation of the objective and constraint functions is crucial. Our approach is to maximize the utilization of the contention-free and contending phases simultaneously within a number of

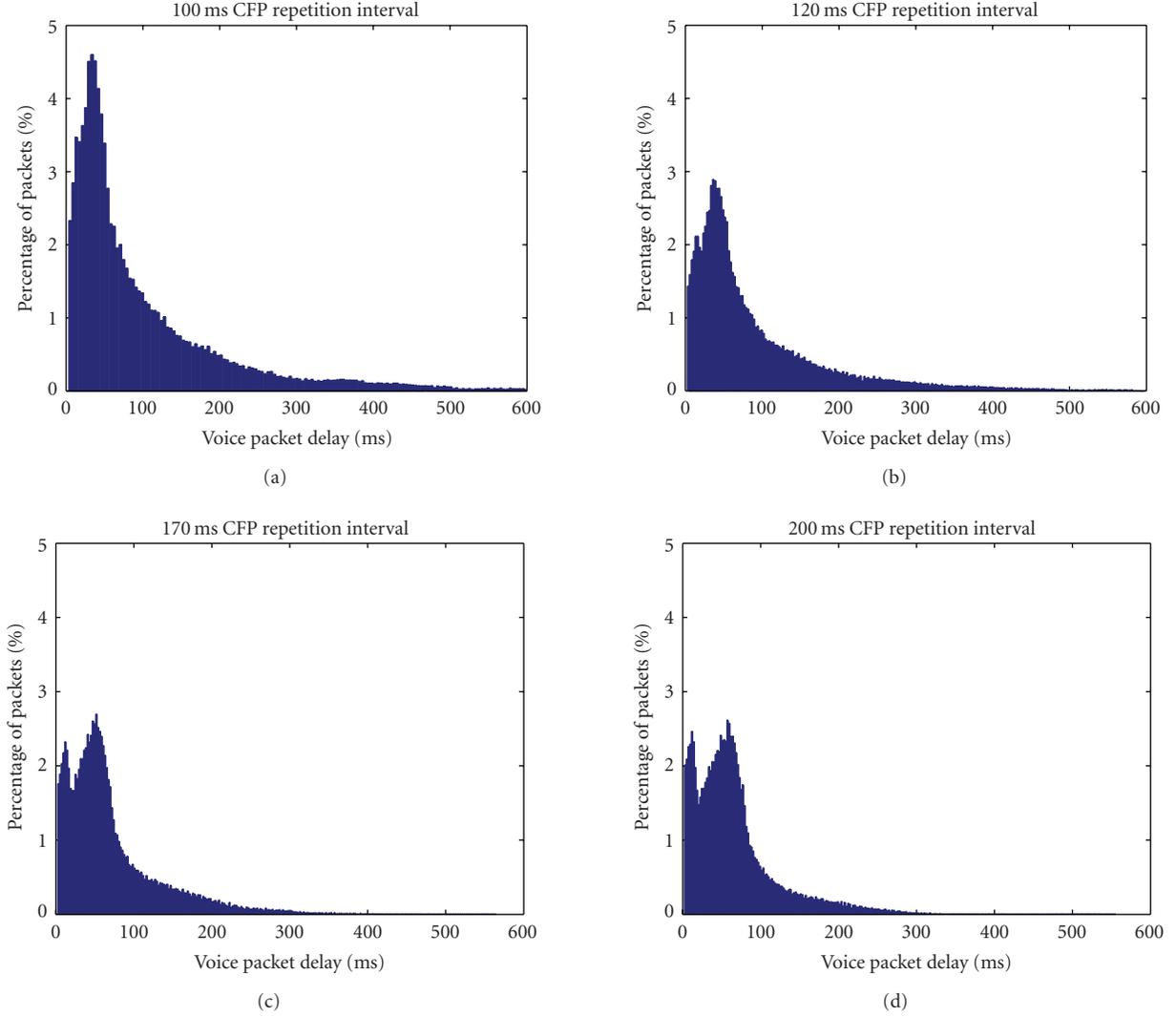


FIGURE 6: PDFs of voice packet delays at 70% CFP_{MAX}.

constraints, such that the two phases' utilizations are traded-off against each other. Therefore, expressions for these two phases must be carefully developed to represent the efficiency of the resource allocation in each phase, such that the resulting objective function can determine how far from the ideal each component is.

Before the model is developed, as with the preceding simulation study, we make assumptions of a reliable physical layer channel (no link errors, no collisions), and exclude hidden terminals, the capture-effect, and the power-saving mechanism, and assume that all stations are fully backlogged (i.e., they always have data to send).

Each phase is affected by two inefficiency components. The first is the efficiency of an individual exchange (which scales linearly with the number of exchanges) and the second is the efficiency of the whole phase, taking into account any unused airtime at the end of the phase.

Firstly, consider the QoS-sensitive polled traffic in the CFP (as illustrated in Figure 9). In the case of the first com-

ponent, due to the assumption that all stations are fully backlogged, no poll is wasted, so each packet polled from station incurs an overhead comprising just the interframe spaces between contention-free packets (SIFS):

$$C_a = 2 * \text{SIFS}. \quad (1)$$

The second component is the wastage at the end of the CFP if it is configured to any size not divisible exactly by the frame exchange duration (although note that, in practice, the central controller can terminate the CFP early and make this "wasted" period available to the CP). The overall efficiency for the CFP can be calculated as

$$V(N_p) = \left(1 - \frac{N_p(C_b - C_a)}{xy}\right), \quad (2)$$

where C_b is the entire polled exchange duration (ms) and C_a is the polled exchange overhead from (1), and x and y are CFP_{MAX} and CFP_{REP}, respectively. These parameters are tabulated for convenience in Table 3.

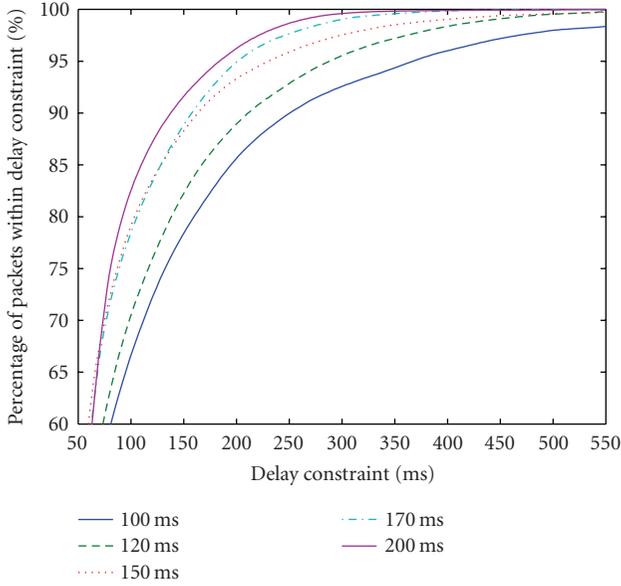


FIGURE 7: CDFs of voice packet delays at 70% CFP_{MAX} .

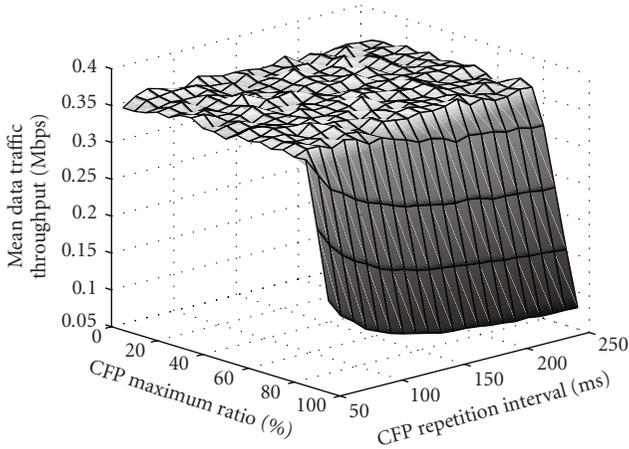
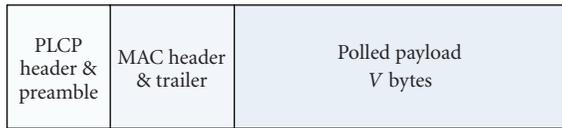
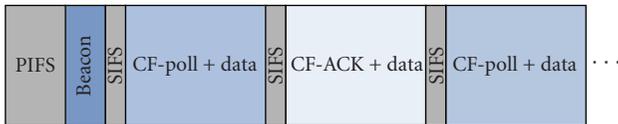


FIGURE 8: Mean data throughputs for various CFP_{MAX} .



(a) Polled frame



(b) Medium occupancy

FIGURE 9: Polling frame model.

TABLE 3: Model parameter definitions and values.

Parameter	Definition (value)
M_s	Standardized data exchange overhead (0.674 ms)
H_s	Standardized data exchange (4.978 ms)
C_a	Polled exchange overhead (0.02 ms)
C_b	Polled exchange duration (2.228 ms)
N_c	Number of data stations (11)
N_p	Numbers of polling stations considered (2, 4, ..., 20)
D	Polling rates under consideration (75, 87.5, 100, 112.5, ..., 200 ms)
P_r	Contending traffic packet generation rate (0.0075 s^{-1})
CFP_{MIN}	Minimal CFP size (39.922 ms)
CP_{MIN}	Minimal CP size (21.404 ms)



(a) Contention frame

(b) ACK Frame



(c) Medium occupancy

FIGURE 10: Contending frame model.

The number of polled terminals, N_p , is a parameter that the AP can reasonably be expected to know as all stations must associate with the AP if polling service is required.

For the CP (as illustrated in Figure 10), recall the operation of the DCF. Stations must wait for the DIFS period of silence on the medium (with the 802.11b physical layer, this is $50 \mu\text{s}$). If this period has elapsed without any activity on the medium, the station then performs a random back-off for a random number of slots (each of 20-microsecond duration in 11b) drawn from the range $[0, CW]$, where CW (contention window) begins at 31 (11b again) and can increase as a binary exponential up to the limit 1023.

If the station detects a transmission during the contention window before its backoff has finished, then the station has lost this particular contention to another station (which happened to choose a smaller backoff this time around), and it must suspend the countdown, and resume it on a later attempt. If a station gains access but experiences a collision on transmission, it will increase the size of CW for the next attempt. However, the “no collisions” assumption can be used to simplify this mechanism by freezing CW at its smallest value of 31, and taking the mean CW value of 15.5 for every contention. If every contention is assumed to win without any other terminal transmitting during the CW phase (although in reality the probability of seeing another terminal transmit is going to increase with the number of terminals present), then a single DIFS per contention can be assumed.

This gives the first efficiency component of the CP as

$$M_S = \text{DIFS} + \text{backoff} + \text{SIFS} + \text{ACK frame}. \quad (3)$$

The second efficiency component (wastage at the end of the phase) can be determined from the effective number of contending stations. This in turn depends on the traffic level and the total number of contending stations, N_c . If we know the approximate packet rate of this traffic, P_r , the effective number of concurrently sending stations will be $y \times P_r \times N_c$. Hence, the overall efficiency of the CP simplifies to

$$L(N_c) = \frac{P_r \times N_c (M_s - H_s)}{(1 - x)}, \quad (4)$$

where H_s is the entire standardized contended exchange duration (ms), M_s is the standardized contended exchange overhead (ms) from (3), N_c is the number of contending stations, y is CFP_{REP} , and x is CFP_{MAX} . We must further constrain this expression by the frame-generation rate of the traffic, otherwise this becomes almost a “self-optimizing” model that will always fill the CP to capacity. We can use the utilization functions L and V in the following objective function:

$$f_0(x, y) = (1 - L(N_c))^2 + (V(N_p))^2. \quad (5)$$

We use the $1 - L(N_c)$ term since higher values of L correspond to good performance (in contrast to high values of V , which indicate poorer performance), and square both terms to ensure that both are positive and continuously differentiable over the whole domain of interest. Substituting the expressions for V and L given in (2) and (4), respectively, and simplifying gives

$$f_0(x, y) = \left(1 - \frac{P_r N_c (M_s - H_s)}{1 - x}\right)^2 + \left(1 - \frac{N_p (C_b - C_a)}{xy}\right)^2. \quad (6)$$

A number of constraints on this solution can be identified. CFP_{MAX} is a ratio of two time periods, so it must be positive and less than one. CFP_{REP} is bounded by the worst-case polling frequency (“delay,” D) specified by the application. Additionally, both the CFP and CP are subject to minimum duration constraints (“ CFP_{MIN} ” and “ CP_{MIN} ,” resp.) according to the standard [1]. The CFP has to be at least big enough to contain one polled exchange comprising the largest payload possible in each direction, plus a beacon and a CF-end. The CP has to be large enough to contain an acknowledged exchange of the largest payload possible.

Mathematically, the problem reduces to an optimization problem over two variables, x and y : minimize $f_0(x, y)$ from (6), subject to the set of constraints:

$$\begin{aligned} \text{CFP}_{\text{MIN}} - xy &\leq 0, \\ \text{CP}_{\text{MIN}} - (1 - x)y &\leq 0, \\ 0 &\leq x \leq 1, \\ 0 &\leq y \leq D. \end{aligned} \quad (7)$$

5.1. Nonlinear vector optimization of model

Before standard optimization techniques can be unleashed on the problem, the objective function must be first reformulated in vector form with a single variable. Let $\mathbf{z} = (x, y)^T$, and define the two unit vectors $e_1 = (1, 0)^T$ and $e_2 = (0, 1)^T$. We can then rewrite the objective function as

$$f_0(\mathbf{z}) = \left(1 - \frac{\alpha}{1 - e_1^T \mathbf{z}}\right)^2 + \left(1 - \frac{\beta}{z^T E \mathbf{z}}\right)^2. \quad (8)$$

Here $\alpha = P_r N_c (M_s - H_s)$, $\beta = N_p (C_b - C_a)$, and $E = e_1 e_2^T$. Other parameters are defined in Table 3, along with the values used in the application of this model. The constants are determined by the physical layer under consideration and the characteristics of the traffic flows.

In vector notation, the constraints can be restated as follows:

- (i) $\text{CFP}_{\text{MIN}} - z^T E \mathbf{z} \leq 0$: first constraint;
- (ii) $\text{CP}_{\text{MIN}} - e_2^T \mathbf{z} + z^T E \mathbf{z} \leq 0$: second constraint;
- (iii) $e_1^T \mathbf{z} - 1 \leq 0$: third constraint, upper bound;
- (iv) $-e_1^T \mathbf{z} \leq 0$: third constraint, lower bound;
- (v) $e_2^T \mathbf{z} - D \leq 0$: fourth constraint, upper bound;
- (vi) $-e_2^T \mathbf{z} \leq 0$: fourth constraint, lower bound.

Before the barrier method [26] can be used to solve this problem, there is one more hurdle to overcome. This objective function is not convex, and furthermore may have multiple solutions (local minima). Two of these minima may occur at the extreme values of the feasible set, with a third local minimum from the objective function. Feasible starting points must be determined to guide the solution in the right direction. By examining the inequality constraints of the original problem, it is possible to find feasible starting points x_0 and y_0 that can be used to initialize the barrier method. Consider the following two inequalities:

$$\begin{aligned} \text{CFP}_{\text{MIN}} &\leq xy, \\ \text{CP}_{\text{MIN}} &\leq (1 - x)y = y - xy. \end{aligned} \quad (9)$$

These are obtained by rearranging the first two inequalities of the original problem statement. Solving the second inequality for xy enables the composite inequality to be written as $\text{CFP}_{\text{MIN}} \leq xy \leq y - \text{CP}_{\text{MIN}}$.

Thus, for a given $y = y_0$, a feasible $x = x_0$ can be taken from the interval

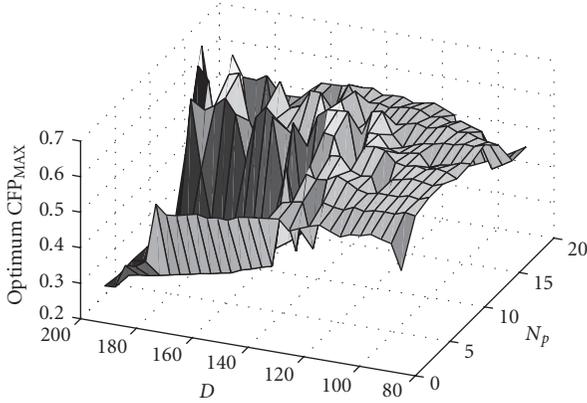
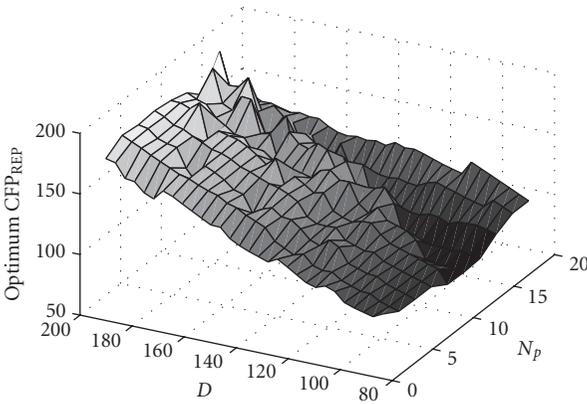
$$x_0 \in \left(\frac{\text{CFP}_{\text{MIN}}}{y}, 1 - \frac{\text{CP}_{\text{MIN}}}{y}\right), \quad (10)$$

and the following feasible starting point constraint must be met:

$$\text{CFP}_{\text{MIN}} > y_0 - \text{CP}_{\text{MIN}}. \quad (11)$$

5.2. Application of model

The assumptions and parameters used in [5] and the simulation model in Section 4 can be adopted by this model to give

FIGURE 11: CFP_{MAX} optimization results.FIGURE 12: CFP_{REP} optimization results.

some concrete values. These parameters are given in Tables 1, and 3 gives the resulting concrete values for the constants in the model. The starting point constraint in (11) can be met for these values when, for example, $CFP_{MIN} = 39.922$, $CP_{MIN} = 21.404$, and $y_0 = 48$.

Three local minima were discovered using the following set of initial x values:

- (1) $1.2 * (CFP_{MIN}) / y$;
- (2) $0.5 * (1 - CP_{MIN} - CFP_{MIN}) / y$;
- (3) $0.8 * (1 - (CP_{MIN} - CFP_{MIN})) / y$.

The first of these is a point near the lower end of the feasible set, the second a point in the middle, and the third a point towards the top end of the feasible set for x . For many values of D and N_p , all of these local minima were found to be identical, indicating that the local minimum is a global minimum. In the case where a number of local minima were found, the objective function was evaluated at each one and the true minimum chosen. The minimum values obtained are illustrated in Figures 11 and 12—compare the relatively smooth surface of Figure 12 with that of the benchmark results shown in Figure 2. These configurations have been verified by comparison with comprehensive simulation [23].

The optimum values of CFP_{MAX} are fairly variable, especially for larger values of D and the smaller values of N_p . This variability seems to occur mainly when the objective is most flat: in that it does not vary much over a wide range of CFP_{MAX} values. This means that the instability happens in exactly the situations where choosing a precise value of CFP_{MAX} is least important. The CFP_{REP} optima tend to be close to the maximum D , especially for smaller D where the constraints do not permit much variation anyway. For larger D , the optimum values are significantly smaller than D , this is in line with the fact that there is much more potential to fit the polled and contention periods within a smaller repetition time.

6. QUEUING THEORY APPROACH

Queuing theory models can be used to analyze the performance of many aspects of wireless networks. Here we apply this approach to the polling phase of the PCF procedure. In these models, the system is thought of as a queue which is filled with packets by an *arrival* process and is emptied by a *servicing* process. In this application, the arrival process is the voice packet generation system, and the servicing process is the polling mechanism as implemented by the AP. Queuing models aim to provide information about the distributions of the time spent in the queue (the *waiting* time) and queue length distributions. The waiting time depends on the mixture of arrival time distribution and service time distribution. The arrival time model for this application is a simple Poisson process when the voice stream is in “on” mode; we assume here that the switch from “on” to “off” occurs sufficiently infrequently to not influence the waiting time distribution. The service time distribution is dependent on the exact polling process used by the AP.

A specific use of this technique to packet delay of polled protocols can be found in [27]. The technique of Laplace-Stieltjes transforms (LST) allows the treatment of the service time distributions to be as general as possible and provides more detailed information about the full distribution of the waiting times. We present the analysis in this form here primarily for the first reason, since we do not use information beyond the mean waiting time explicitly in this paper. The service time distribution is given either as a cumulative distribution function (CDF), or its derivative, the probability density function (PDF). The LST of a CDF of a random variable $F(t)$ is given by

$$\phi(s) = \int_0^{\infty} e^{-st} dF(t). \quad (12)$$

These CDFs (and corresponding LSTs) are used to capture the distributions of service times and waiting times. A central result [28] in queuing theory analysis for a queue with exponential arrival times (mean rate λ) and general service time distribution (with LST $\eta(s)$ and mean τ) is that the LST of the waiting time is given by

$$w(s) = \frac{s(1 - \lambda\tau)}{s - \lambda(1 - \eta(s))}. \quad (13)$$

This is known as the Pollaczek-Khintchine (PK) formula. Inverting the corresponding LST to get back to the more useful PDF of the waiting times is often intractable. However, we can readily extract the set of *moments* (M_n) of the PDF distribution using the following formula:

$$M_n(F) = (-1)^n \left(\frac{d^n}{ds^n} \phi(s) \right)_{s=0}. \quad (14)$$

All the properties of a distribution can be deduced from its full set of moments, but this may require computation of a large number of them. The mean (μ) and variance (σ^2) can be calculated directly from just the first and second moments:

$$\begin{aligned} \mu &= M_1(F), \\ \sigma^2 &= M_2(F) - \mu^2. \end{aligned} \quad (15)$$

6.1. Application to PCF delay model

This theory can be applied to analyze the delay times of the polling procedure in 802.11 PCF. The polling procedure that the AP runs flips between two states, polling and contending. We make two assumptions in this model.

- (1) Service times of the polling mechanism are independent.
- (2) The time to poll and receive responses from the complete set of stations is constant.

The first is not strictly the case here since there is a deterministic switch between polling and CP modes. This means that the short delay that occurs in polling mode is very likely to be followed by an equally short delay, and similarly longer delays will tend to follow longer delays when the system is in CP mode. In practice, this assumption should only restrict the range of parameters over which the results are valid, since the deterministic process is likely to be more stable in the face of configurations that would otherwise cause the polling mechanism to break down with unacceptably large delays.

The second is an approximation since if a station has a packet, its response will take longer than if it is returning a null frame. Thus it will take longer to poll the full set of stations at the beginning of the CFP when most stations are waiting with a packet than it does at the end when most have empty queues. In the model, we approximate such a delay by looking at the expected number of stations that has packets and combining it with the with-packet and without-packet polling times, building a weighted average for the polling time, which we denote by r . This constant rate assumption will have greatest effect on large superframe configurations since the variation in total time to poll will be the largest across the whole frame in these configurations.

Next we construct a CDF for the service time for the polling traffic. Each station gets polled a total number $n_{\text{poll}} = \lfloor xy/r \rfloor$ of times each superframe. As in the previous section, we use x to denote CFP_{MAX} and y to denote CFP_{REP} . In each of these occasions, the service time is r . In the following time slot, the CFP ends and the service time is equal to the length of the CP, $y(1-x)$. So the service time has value r with probability $n_{\text{poll}}/(n_{\text{poll}}+1)$, and value $y(1-x)$ with probability

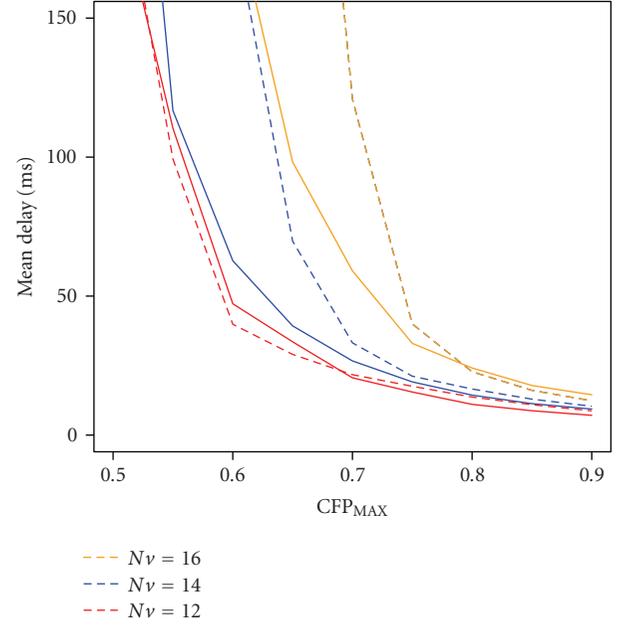


FIGURE 13: Mean packet delays for a range of voice stations. Solid lines show model predictions, dotted lines show simulated values.

$1/(n_{\text{poll}}+1)$. This translates to a PDF for the service times of

$$\text{ServPDF}(t) = \frac{\delta(t-r)n_{\text{poll}}}{n_{\text{poll}}+1} + \frac{\delta(t-y(1-x))}{n_{\text{poll}}+1}. \quad (16)$$

Here we use $\delta(t)$ the Dirac delta function to represent in the PDF what would be discontinuities in the CDF. The required CDF is given by the integral of this function. This service time has corresponding LST given by

$$\text{LSTServ}(s) = \frac{e^{-rs}n_{\text{poll}}}{n_{\text{poll}}+1} + \frac{e^{sy(x-1)}}{n_{\text{poll}}+1}. \quad (17)$$

We insert this in the PK formula (13), assuming that the voice source is in talk-spurt mode with a Poisson arrival rate of packets with mean λ . If we compute the first moment using (14), we obtain the following formula for the mean packet delay:

$$D_{(x,y,r,\lambda)} = \frac{\lambda(n_{\text{poll}}r^2(x-1)^2y^2)}{2(\lambda(x-1)y + (1-\lambda)r n_{\text{poll}} + 1)}. \quad (18)$$

Once suitable values of λ and r are set from the scenario parameters, the mean packet delay can be computed. Figure 13 shows the mean delay predicted by this method compared to the mean delay observed from OPNET simulation. Here we fix CFP_{REP} to be 120 milliseconds and show the delays for a range of CFP_{MAX} values from 50% to 90%. For 12 voice stations, there is very close agreement between the model and the simulations. For larger numbers of stations, there is more discrepancy for smaller values of CFP_{MAX} , but this is where both model and simulation tend to break down anyway due to high packet delays.

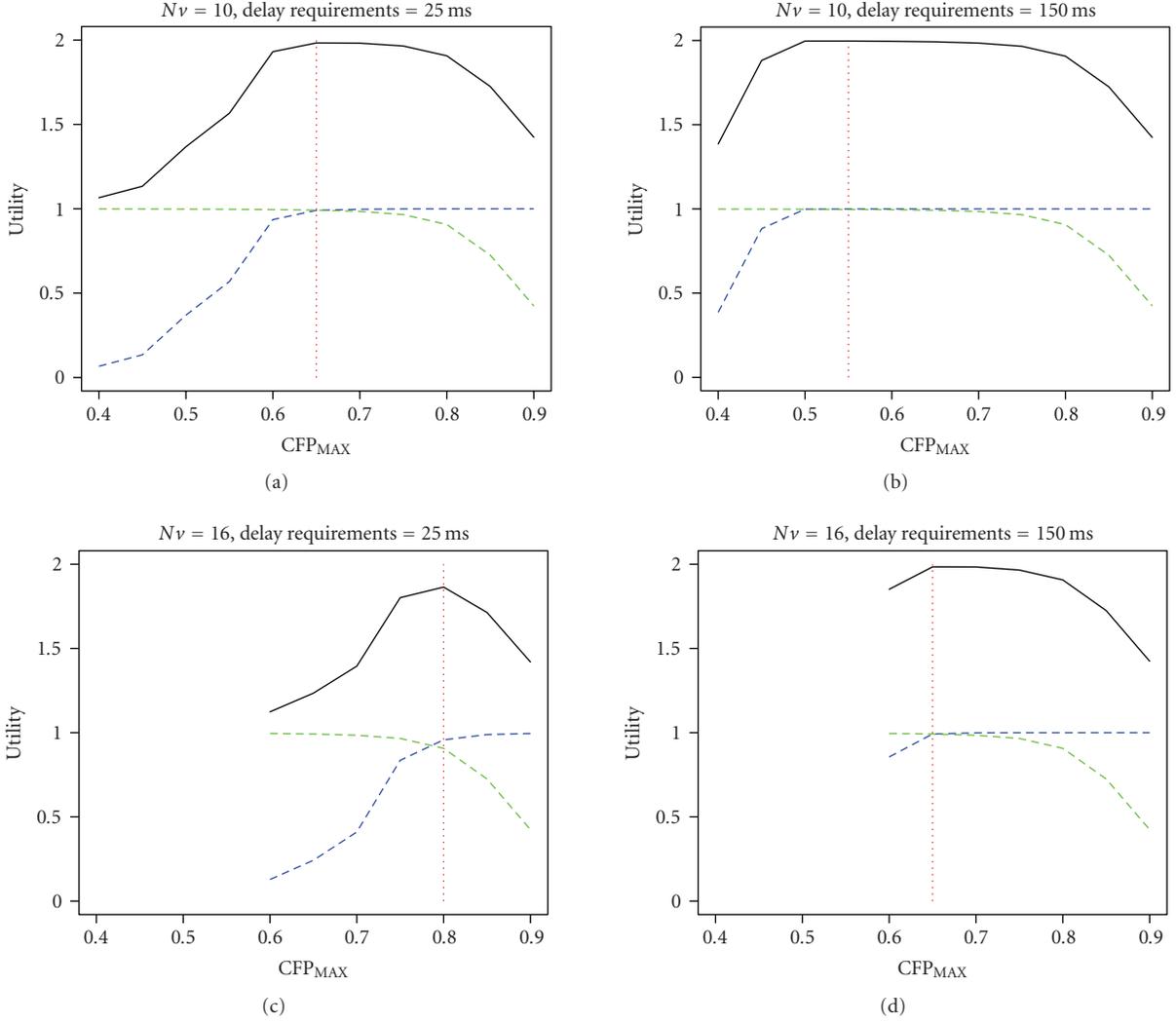


FIGURE 14: Objectives for a set of different network sizes and delay requirements. The blue dotted line shows the polled traffic utility, the green gives the data utility. The black line is the combined objective; red dotted line is the optimum CFP_{MAX} .

Generally though, the agreement is close enough to answer the kind of questions that are required to optimize the performance of this system. For example, the value of CFP_{MAX} that is required to bring the average delay below 50 milliseconds for 14 voice stations is approximately 0.67 for both model predictions and simulations.

6.2. Selection of optimal CFP_{MAX}

This approach does not yet provide enough information to optimize for the values of CFP_{REP} . Future work would look at the higher moments provided by the LST approach and uses these to more accurately predict the percentage of packets satisfying delay requirements. However, we can still use the mean packet delay to select CFP_{MAX} values.

We can combine this packet delay model with a DCF throughput model derived from the one described in [29] to produce a combined performance objective function. The

polled component of the objective is constructed from the required and predicted delay in the following way:

$$Obj_{polled} = U\left(\frac{delay_{req.}}{delay_{predicted}}\right). \quad (19)$$

Here $U(x)$ is a utility function designed to map the region $[0, \infty)$ onto $[0, 1)$ (here we use $(Ax^k + x)/(Ax^k + x^2 + 1)$, with $A = 18$, $k = 6$). Similarly for the data traffic, with predicted and required levels of throughput,

$$Obj_{data} = U\left(\frac{throughput_{predicted}}{throughput_{req.}}\right). \quad (20)$$

The total objective is simply the sum of these two components; maximizing the value of this function provides optimum CFP_{MAX} values.

Figure 14 illustrates the objectives for different numbers of voice stations and delay requirements. The curves for the

set of 16 voice stations are only partially given since the delay model is not applicable to values of CFP_{MAX} less than 0.6 with this number of stations. Note also that with the level of data traffic described in Section 4 (with 6 data nodes), the performance for these devices does not drop off significantly until CFP_{MAX} reaches beyond 0.8. This causes the optimum CFP_{MAX} values to be driven chiefly by the polled traffic. The optima do change in an intuitive fashion, since increasing the number of voice stations and tightening the delay restriction both force a higher value of CFP_{MAX} in order to give higher priority to the voice traffic.

7. APPLICATION

So far in this paper, we have presented a number of alternative means for configuring the PCF superframe. Clearly, what is needed is a means for fairly comparing these different approaches. To achieve this, a prototype has been developed in OPNET, embodying an adaptive mechanism that reconfigures the superframe as needed. This model can be configured to use a fixed superframe structure (and thereby, replicate the results of the aforementioned simulations), but can also be configured to react to changes in traffic conditions. Additionally, the model can provide benchmark results, it can be configured to use the results from Li et al. [5], or can carry all traffic via the contention-based DCF mechanism.

7.1. Adaptive model

The key to any adaptive model is the provision of a feedback path, in this case to feed back information on the degree to which QoS requirements are being met, such as the end-to-end delays are experienced. In this model, this feedback path is abstracted as a statistic wire in OPNET, although a real implementation would, of course, require additional signaling to transfer this information.

The adaptation mechanism comprises a check as to whether the QoS requirements are being met as the superframe is started. This check is performed periodically, the period dictated by the tolerance to jitter. If the requirements have drifted, then the reconfiguration process is triggered.

The reconfiguration process supports multiple data sets. These data sets can either be a fixed lookup table (as it is the case with the Li results), or can be generated dynamically by an online optimization algorithm. This approach models the range of possible implementations that could be considered, from simple products with a limited number of lookup tables to more complex adaptive solutions.

The entire adaptation algorithm can be disabled (thereby supporting a predefined fixed superframe configuration), and the traffic streams can all be diverted to DCF to show the effect of not having a polling mechanism at all.

7.2. Scenarios

In this paper, we concentrate on two specific scenarios of interest. To allow comparison with the benchmark results, the physical layer rates are those supported by 802.11b physical

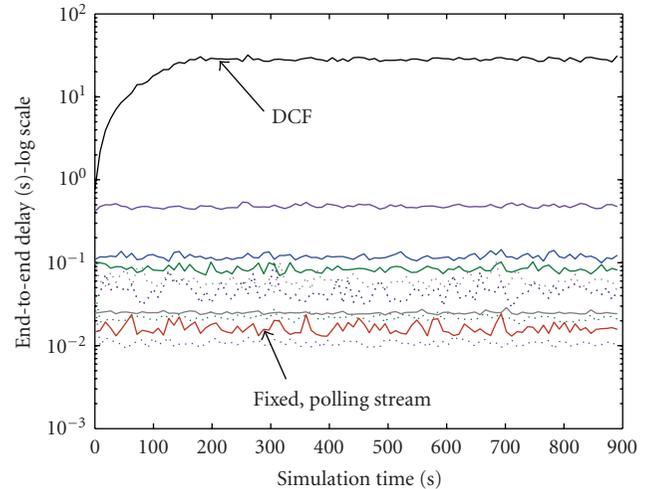


FIGURE 15: End-to-end delays for scenario 1.

layer (2 Mbps for data and 1 Mbps for control messages). The first scenario comprises an AP plus:

- (i) 6 STA: local file transfer (DCF, 1500-byte mean);
- (ii) 10 STA: carrying voice traffic (30-millisecond maximum delay, fixed MSDU 200 bytes, 64 kbps to correspond to a G.711-style voice CODEC, duty cycle 1 : 2.35).

The second scenario is an extension of the first. In this second scenario, an additional “glut” of QoS-sensitive traffic is initiated halfway through the simulation, in the form of an additional six voice traffic streams.

7.3. Results

The first scenario is a static scenario in terms of the traffic stream profile, as no new traffic streams begin and no existing traffic streams end in the course of the simulation. It would be expected that measurements of performance parameters will soon reach a fairly steady state, and this is borne out by results shown in Figures 15 and 16.

The results are for a DCF-only configuration (i.e., all traffic having to contend for access), a fixed superframe scheme and adaptive schemes using data from the benchmark (Li), and the results from Sections 4, 5, and 6. In the following traces, in all but the DCF case, there are separate traces for the polling and contending traffic flows, the polling results are marked by solid lines, the contending are marked by dashed lines. Specific traces of interest are highlighted in the figures.

The DCF benchmark configuration shows increasing instantaneous end-to-end delay, but offers the best receive rate of all the contention schemes as no time is spent polling. The other schemes all achieve the required end-to-end delay requirements of the voice traffic, including the fixed superframe configuration because it has an “ideal” configuration selected for this scenario (CFP_{MAX} of 85% and CFP_{REP} of 30 milliseconds). The received data rates (Figure 16) have all clustered in a similar way, with the polled traffic getting considerably greater throughput.

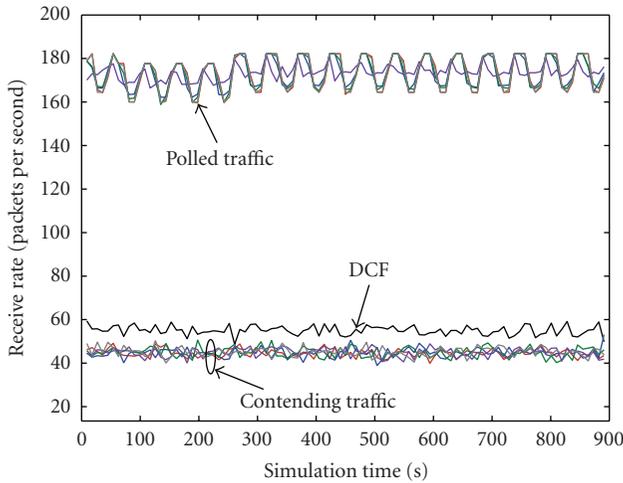


FIGURE 16: Receive rates for scenario 1.

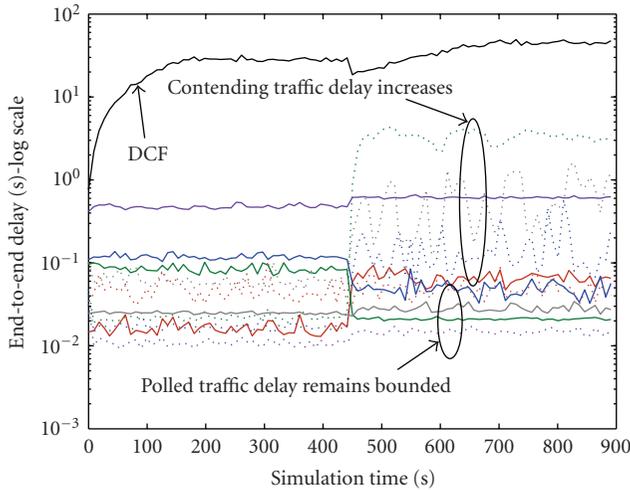


FIGURE 17: End-to-end delays for scenario 2.

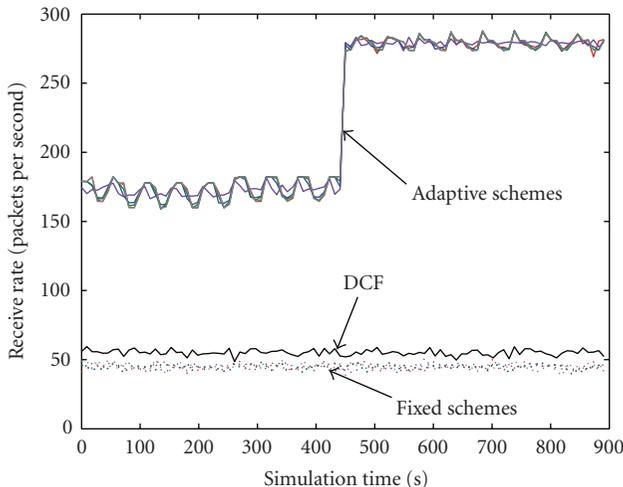


FIGURE 18: Receive rates for scenario 2.

The benefits of PCF and of having an adaptive scheme soon become apparent when the second scenario is considered. Firstly, let us consider the disadvantages of a DCF-only system. As can be seen in Figure 18, the received data rate for DCF does not change when the system is further loaded with additional traffic, and, in Figure 17, the end-to-end delay increases.

The adaptive schemes are able to respond to the change in traffic stream demand and reconfigure to provide a nearly constant end-to-end delay for the polling traffic, sacrificing some of the end-to-end delay performance of the contending traffic, which is an acceptable and even sensible tradeoff.

A more detailed examination of the adaptive schemes reveals that the nonlinear optimization approach offers the most stable configurations, but the queuing-theory-based approach offers comparable results and has the benefit of having more potential for distributed solutions in this area. The nonlinear optimization approach does well on the polled delays, but that is at the expense of the contention traffic, which incurs a greater penalty than with the other approaches. There is the clear benefit with models that cater for all of the constraints of the IEEE802.11 specification, making any solution based on those results fully compliant with the standard.

8. CONCLUSIONS

This paper has described the IEEE802.11 centralized control schemes, concentrating on the PCF. There has been a considerable amount of research into the support of QoS-sensitive traffic in more distributed aspects of IEEE802.11, but much less investigation into centralized solutions. An existing superframe configuration solution has been described and opportunities for improvement have been identified.

A number of solutions for configuring the PCF superframe have been presented. Firstly, an improved simulation model has been used to provide an accurate set of results for any lookup-table-oriented solution. This model confers the advantage over the literature available to date of being fully compliant with the standard. This approach demonstrates the need to focus on the time-dependent services and shows the importance of considering several performance measurements.

Secondly, two mathematical models have been developed, resulting in optimized sets of values for a given configuration, and, critically, general purpose algorithms that provide optimal results for any set of model constraints.

Finally, an adaptive prototype has been presented that can show each approach in active use, highlighting the effects of changes in traffic requirements. This prototype has highlighted the consistency of the more mathematically based approaches, as well as demonstrating the benefits of both centralized control and adaptive solutions.

In terms of future work, we hope to extend this solution to the more general case of the IEEE802.11e HCF, as well as investigating the benefits (and disadvantages) of distributed methods of handling mixed traffic networks such as the distributed reservation protocol [3].

REFERENCES

- [1] IEEE (Institute of Electrical Electronics Engineers), "IEEE Standard 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," 1999.
- [2] Y. Xiao, "QoS guarantee and provisioning at the contention-based wireless MAC layer in the IEEE 802.11e wireless LANs," *IEEE Wireless Communications*, vol. 13, no. 1, pp. 14–21, 2006.
- [3] WiMedia-Alliance (Ecma International (Ecma)), "Standard ECMA-368 High Rate Ultra Wideband PHY and MAC Standard," 2005, <http://www.wimedia.org/en/resources/eis.asp?id=res>.
- [4] IEEE (Institute of Electrical Electronics Engineers), "IEEE Standard 802.11e—Part 11 Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements," 2005.
- [5] C. Li, J. Li, and X. Cai, "A study of self-adaptive transmission for integrated voice and data services over an IEEE 802.11 WLAN," in *Proceedings of 15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '04)*, vol. 3, pp. 1922–1926, Barcelona, Spain, September 2004.
- [6] Y. Zang, G. Hiertz, J. Habetha, B. Otal, H. Sirin, and H.-J. Reumerman, "Towards high speed wireless personal area network - efficiency analysis of MBOA MAC," in *Proceedings of International Workshop on Wireless Ad-Hoc Networks (IWWAN '05)*, pp. 10–20, London, UK, May 2005.
- [7] Y. Xiao and H. Li, "Evaluation of distributed admission control for the IEEE 802.11 e EDCA," *IEEE Communications Magazine*, vol. 42, no. 9, pp. S20–S24, 2004.
- [8] J. Wall and J. Y. Khan, "An adaptive ARQ enhancement to support multimedia traffic using 802.11 wireless LANs," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '04)*, vol. 5, pp. 3037–3041, Dallas, Tex, USA, November 2004.
- [9] P. Chatzimisios, A. C. Boucouvalas, and V. Vitsas, "IEEE 802.11 wireless LANs: performance analysis and protocol refinement," *EURASIP Journal on Wireless Communications and Networking*, vol. 2005, no. 1, pp. 67–78, 2005.
- [10] L. Gannoune and S. Robert, "Dynamic tuning of the contention window minimum (CWmin) for enhanced service differentiation in IEEE 802.11 wireless ad-hoc networks," in *Proceedings of 15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '04)*, vol. 1, pp. 311–317, Barcelona, Spain, September 2004.
- [11] M. Bandinelli, F. Chifi, R. Fantacci, D. Tarchi, and G. Vannuccini, "A link adaptation strategy for QoS support in IEEE 802.11e-based WLANs," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC '05)*, vol. 1, pp. 120–125, New Orleans, La, USA, March 2005.
- [12] A. Iera, G. Ruggeri, and D. Tripodi, "Providing throughput guarantees in 802.11e WLAN through a dynamic priority assignment mechanism," *Wireless Personal Communications*, vol. 34, no. 1-2, pp. 109–125, 2005.
- [13] D. Chen, D. Gu, and J. Zhang, "Supporting real-time traffic with QoS in IEEE 802.11e based home networks," in *Proceedings of 1st IEEE Consumer Communications and Networking Conference (CCNC '04)*, pp. 205–209, Las Vegas, Nev, USA, January 2004.
- [14] G. Smith and D. Dillon, "QoS over IEEE 802.11e: the need for HCCA for video applications," *Bermai*, pp. 1–13, 2004.
- [15] J. N. Al-Karaki and J. M. Chang, "Quality of service support in IEEE 802.11 wireless ad hoc networks," *Ad Hoc Networks*, vol. 2, no. 3, pp. 265–281, 2004.
- [16] S. Mangold, S. Choi, G. R. Hiertz, O. Klein, and B. Walke, "Analysis of IEEE 802.11 e for QoS support in wireless LANs," *IEEE Wireless Communications*, vol. 10, no. 6, pp. 40–50, 2003.
- [17] P. Brady, "A model for generating on-off speech patterns in two-way conversation," *Bell System Technical Journal*, vol. 48, pp. 2445–2472, 1969.
- [18] J. Zheng and E. Regentova, "An improved polling scheme for voice support in IEEE 802.11 wireless network," in *Proceedings of International Conference on Information Technology: Coding and Computing (ITCC '05)*, vol. 2, pp. 603–608, Las Vegas, Nev, USA, April 2005.
- [19] X. Ma, C. Du, and Z. Niu, "Adaptive polling list arrangement scheme for voice transmission with PCF in wireless LANs," in *Proceedings of Joint Conference of the 10th Asia-Pacific Conference on Communications and the 5th International Symposium on Multi-Dimensional Mobile Communications Proceedings (APCC/MDMC '04)*, vol. 1, pp. 21–25, Beijing, China, August–September 2004.
- [20] R. Y. W. Lam, V. C. M. Leung, and H. C. B. Chan, "Polling-based protocols for packet voice transport over IEEE 802.11 wireless local area networks," *IEEE Wireless Communications*, vol. 13, no. 1, pp. 22–29, 2006.
- [21] ITU-T (International Telecommunications Union - Telecommunication Standardisation Sector), "Transmission systems and media: general characteristics of international telephone connections and international telephone circuits: one-way transmission time," 1996.
- [22] X. Ma, Y. Wu, Z. Niu, and T. Sato, "Performance analysis of the packetized voice transmission with PCF in an IEEE802.11 infrastructure wireless LAN," in *Proceedings of 9th Asia-Pacific Conference on Communications (APCC '03)*, vol. 2, pp. 571–575, Penang, Malaysia, September 2003.
- [23] R. J. Haines, T. Lewis, J. Coon, and N. Fanning, "Non-linear optimization of IEEE 802.11e super-frame configuration," in *Proceedings of 63rd IEEE Vehicular Technology Conference (VTC '06)*, vol. 3, pp. 1211–1215, Melbourne, Australia, May 2006.
- [24] B. Krishnamachari and F. Ordóñez, "Analysis of energy-efficient, fair routing in wireless sensor networks through non-linear optimization," in *Proceedings of 58th IEEE Vehicular Technology Conference (VTC '03)*, vol. 5, pp. 2844–2848, Orlando, Fla, USA, October 2003.
- [25] M. Sandell and J. Coon, "Near-optimal training sequences for MIMO OFDM systems with nulled subcarriers," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '05)*, pp. 2244–2249, St. Louis, Mo, USA, November–December 2005.
- [26] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, Mass, USA, 2004.
- [27] L. Wang, M. Hamdi, R. Manivasakan, and D. H. K. Tsang, "Multimedia-MAC protocol: its performance analysis and applications for WDM networks," *IEEE Transactions on Communications*, vol. 54, no. 3, pp. 518–531, 2006.
- [28] R. B. Cooper, *Introduction to Queueing Theory*, Elsevier/North Holland, Amsterdam, The Netherlands, 1981.
- [29] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, 2000.