Quality of Service Provisioning in 802.11e Networks: Challenges, Approaches, and Future Directions

Naomi Ramos, Debashis Panigrahi, and Sujit Dey, University of California at San Diego

Abstract

In order to support diverse application requirements, a new standard called IEEE 802.11e is being proposed to improve quality of service in wireless LAN networks. There are, however, a few remaining challenges that need to be addressed in order to enable comprehensive QoS support using 802.11e. In this article we provide an overview of a few of these challenges, describe their possible impact on QoS, and provide a survey of techniques that potentially could be used to address the identified challenges. Specifically, we focus on three challenges: handling time-varying network conditions, adapting to varying application profiles, and managing link layer resources. Additionally, we present several potential future directions toward improved QoS in wireless networks.

Recently, there has been a growing interest in the use of wireless LAN (WLAN) technology. Examples of applications range from standard Internet services, such as Web access, to real-time services with strict latency/throughput requirements, such as multimedia video and voice over IP (VoIP). Future applications will be more demanding and may include high definition television (HDTV) and audiovisual support. With the high demands and varying requirements of these applications, there is a need to support quality of service (QoS) in WLAN.

As a step toward providing QoS support in WLAN, the 802.11 Working Group has been developing a new protocol, IEEE 802.11e [1], to provide differentiation mechanisms at the medium access control (MAC) layer. A number of studies have evaluated the current draft of the standard by both analytical evaluation [2], as well as simulation [3], and have demonstrated the usefulness of the proposed mechanisms in 802.11e.

Although 802.11e enhances the current access mechanisms, there are a number of challenges that must be addressed to enable comprehensive QoS support. In this article we identify a few remaining challenges in QoS provisioning with 802.11e. We specifically focus on three challenges: *handling time-varying network conditions*, *adapting to varying application profiles*, and *managing link layer resources*. We describe the impact of these challenges on QoS and provide a survey of the efforts being made in the research community to address these challenges. The techniques presented in the article improve QoS support, by either adapting the parameters available in the 802.11e mechanism or suggesting new modifications. Finally, we describe other open issues in QoS provisioning and possible future research directions.

IEEE 802.11e Background

We begin by providing an overview of the enhancements proposed in the 802.11e draft [1] for providing QoS. The 802.11e QoS framework defines a new coordination function called the hybrid coordination function (HCF), which multiplexes between two medium access modes: a distributed scheme called enhanced distributed channel access (EDCA), and a centralized scheme called HCF controlled channel access (HCCA). Both access schemes enhance or extend functionality of the original access methods, distributed coordination function (DCF) and point coordination function (PCF), specified in 802.11a/b/g.

Enhanced Distributed Channel Access (EDCA)

The first mode of channel access is EDCA, which is a parameterized version of the previous distributed channel access mechanism of 802.11b. To provide prioritized QoS, 802.11e enhances the original DCF by classifying traffic through the introduction of access categories (ACs). Each AC has its own transmission queue and its own set of channel access parameters. The prioritization in channel access through EDCA is shown in Fig. 1.

The differentiation in priority between each AC is realized by setting different values for the channel access parameters. The following are the most important additional parameters:

- *Arbitrary interframe space number (AIFSN)*: The minimum time interval for the medium to remain idle before starting backoff.
- *Contention window (CW_{min} and CW_{max})*: A random number is drawn from this interval for the backoff mechanism.
- *Transmission opportunity (TXOP) limit*: The maximum duration for which a node can transmit after obtaining access to the channel.

 0890-8044/05/$20.00 © 2005 IEEE

With these parameters, EDCA channel access prioritization works as follows. When data arrives from higher layers, the data is classified and placed in the appropriate AC queue. To determine the next packet to be sent by the node, an internal contention algorithm is used to calculate the total backoff time for each AC. This backoff time includes the time spent waiting for the medium to be idle, which is based on AIFSN, and the estimated backoff time within the contention window range, CW_{min} and CW_{max}. The AC with the smallest backoff wins the internal contention and uses this backoff value to contend externally for the wireless medium. Note that although this technique is similar to DCF, nodes with higher priority can access the channel earlier than other nodes using the parameterized values of the interframe space duration and CW range. Additionally, prioritized flows have the advantage of longer channel access with their TXOP.

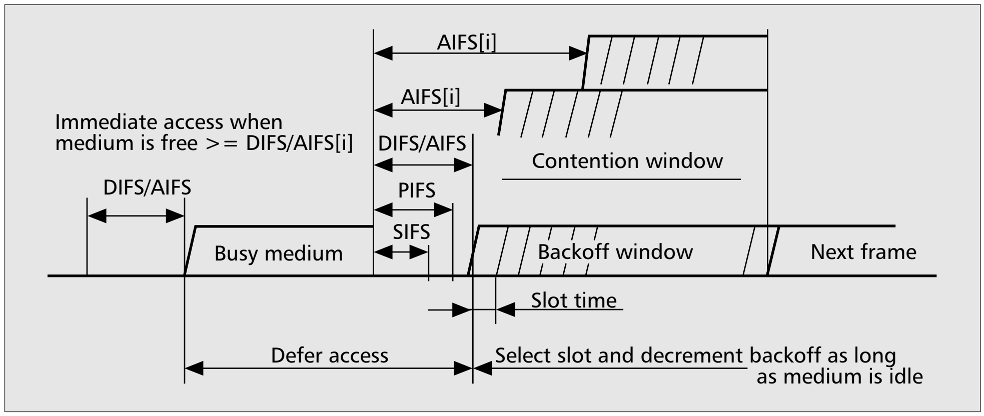

Figure 1. *EDCA channel access prioritization.*

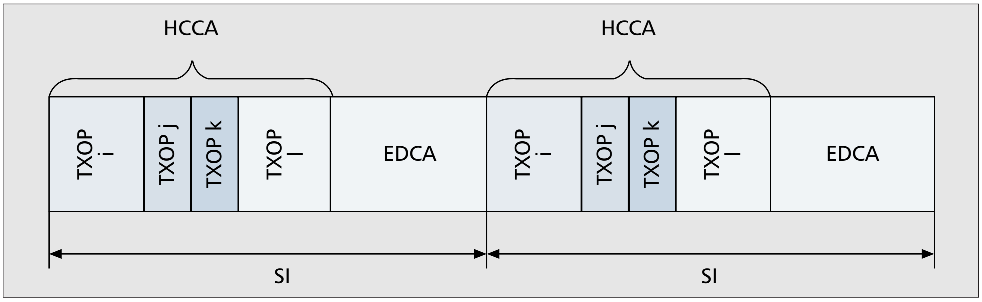

Figure 2. *HCCA channel access prioritization.*

HCF Controlled Channel Access

HCCA provides a centralized polling scheme to allocate guaranteed channel access to traffic flows based on their QoS requirements. Figure 2 illustrates the HCCA channel access scheme. In this period, the AP polls nodes for a TXOP duration, which is calculated from reservation requests sent by the nodes. The TXOP is initiated by a poll request from the AP and during this duration, transmissions can occur in both the uplink and downlink directions. The TXOP allows for multiple contention-free transmissions and ends if one of the following conditions occurs: neither the AP nor the node have any packets left to transmit, the channel idle time has exceeded the timeout period, or the TXOP duration has expired.

An important part of the HCCA framework is the scheduling mechanism used to generate TXOPs for each node based on its traffic flow specifications. The 802.11e Working Group provides a reference scheduler [4] for the HCCA period. Flows with strict QoS requirements send reservation requests to the AP containing flow information, such as mean data rate, mean packet size, MAC service data unit size, and maximum tolerable delay. Using this information, the AP determines the minimum service interval (SI) to be used for all of the nodes, where the SI is the time duration between successive polls for the node. The selected SI satisfies the delay requirements of each flow and is a submultiple of the 802.11e beacon interval duration. Next, the AP determines the TXOP duration of each node based on the mean application data rates of its requested flows. The multiplexing of HCCA and EDCA periods is shown in Fig. 2. The maximum time spent in HCCA for each SI is limited by the *dot11CAPMax* variable, and the total controlled access time in a beacon interval is limited by *dot11CAPRate*. The duration of the controlled access period can be limited using these parameters, and the effect of controlled access mode on traffic flows in contention access mode can be bounded.

Challenges and Approaches to Ensure Quality of Service Support in 802.11e Networks

Having given some background on 802.11e, we now present open challenges in supporting QoS in 802.11e networks. Mainly, we focus on the following set of challenges:

- *Handling time-varying network conditions*
- *Adapting to varying application profiles*
- *Managing link layer resources*

Figure 3 briefly summarizes the issues related to these three challenges and the negative impact they can have on QoS. In this section we first describe each of these challenges in detail and motivate the need to address the challenges by discussing their impact on QoS provisioning. We next provide a survey of the approaches proposed in literature that address aspects of these challenges. For each approach we describe the overall goal, discuss the effectiveness of the techniques, and compare the methods when appropriate. Finally, we present possible future directions in addressing each challenge.

Handling Time-Varying Network Conditions

One of the limitations of the current 802.11e is that it does not consider the impact of varying network conditions. We describe two different factors related to network conditions that impact the experienced QoS: channel conditions and network load. Although the impact of channel conditions and network load in WLAN has been evaluated in previous work, the effect of these factors in the context of QoS networks such as 802.11e has been a topic of recent study.

Varying channel conditions occur in wireless networks because of propagation loss, multipath effects, and interference. Poor channel variations can lead to retransmissions and dropped packets, and thereby increase latency while degrading throughput. The second factor on which we focus is the network load (i.e., the number of contending nodes in the net-

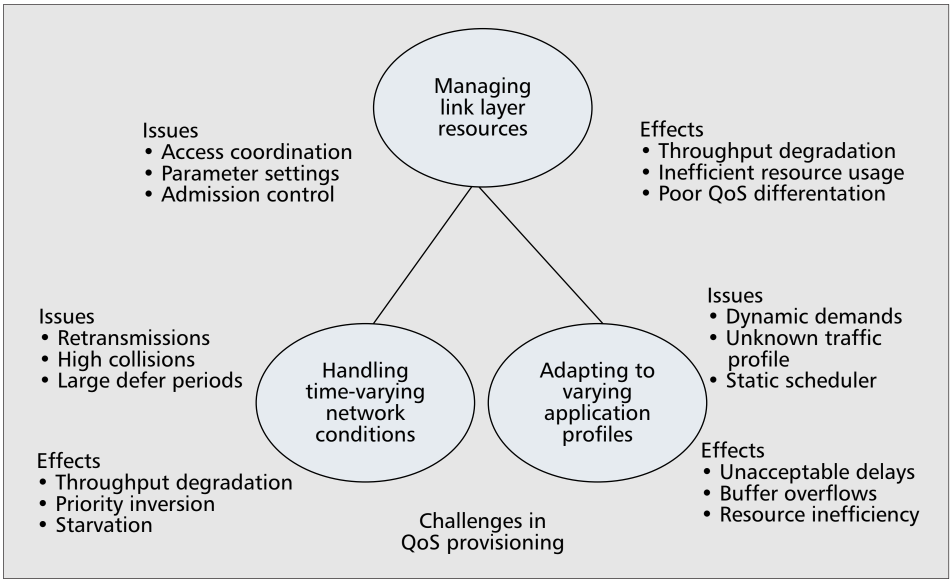

Figure 3. *Challenges in providing QoS in IEEE 802.11e.*

work). Since WLAN uses a shared channel access mechanism, the load of the network directly affects each node's performance.

In terms of providing QoS differentiation, channel condition is an important factor to consider because it can potentially weaken the service differentiation mechanisms of 802.11e. IEEE 802.11e provides a means of differentiation when nodes experience the same channel conditions. However, when the experienced channel conditions are different due to location, interference, mobility, and so on, the 802.11e mechanisms may be insufficient to support service differentiation and prevent degraded performance. In the worst case, poor channel conditions can lead to a situation called *priority inversion*, where higher-priority nodes suffer from significant throughput degradation and achieve lower performance than lower-priority nodes in the network. Such an example is described in [5], where a high-priority node experiencing poor channel conditions achieves a throughput of 618 kb/s, whereas a medium-priority node achieves throughput of 813 kb/s.

Similar to the impact of channel conditions, service differentiation can be negatively impacted by the number of contending nodes. With experienced retransmissions and deferrals, even a node with prioritized access may be unable to achieve its minimal required performance. For example, in the simulations presented in [6], a video flow using the EDCA prioritization mechanism still suffers from a goodput decrease of 33 percent when the number of nodes increases from 4 to 16.

Approaches — Having briefly described the impact of channel conditions and network load, we now describe some of the proposed approaches to improve node performance and provide robustness in these varying network conditions. While some of the channel conditions can be addressed by physical layer techniques, such as modulation scaling and forward error correction (FEC), in this article we focus on the effect of residual channel conditions observed by the MAC layer even with the physical layer enhancements in place. We highlight three pieces of work that attempt to address these challenges, and briefly describe their goals and approaches: channel-dependent packet-level tuning (ChaPLeT), adaptive EDCF, and adaptive fair EDCF. Each of these techniques adapts link-layer parameters to improve service differentiation by trying to improve the performance of nodes in a fair manner.

The first work, ChaPLeT, focuses on improving service differentiation of flows by configuring three selected parameters based on monitored channel conditions and ACs. Each of the adapted parameters, *fragmentation threshold*, *persistence factor*, and *defer countdown*, has an impact on different phases of transmission.

To minimize channel access delay, the first parameter ChaPLeT adapts is the persistence factor (PF). After a retransmission, the CW is doubled in an effort to reduce potential collisions. However, retransmissions can be caused by reasons other than collisions, such as poor channel conditions. In these cases expansion of the CW leads to unnecessary channel access delay and degraded throughput. First, ChaPLeT attempts to identify the cause of the retransmission by estimating the collision probability. Next, it varies the magnification factor, called PF, of the CW and adapts the PF value to reduce the delay prior to channel access. This technique improves the throughput experienced under poor channel conditions.

ChaPLeT's second parameter adaptation is the defer countdown. With distributed access, when a node overhears another node access the medium while it is waiting to transmit a packet, it begins a deferral process. During a deferral, a node freezes its backoff countdown and waits for the channel to be idle before resuming the residual backoff duration. The defer countdown is the countdown rate used for this residual backoff period. In order to avoid excessive delays in deferrals, this parameter is used by ChaPLeT to enable nodes that have experienced a high wait time to speed up their countdown and achieve higher throughput. The final parameter adaptation is changing the maximum allowable packet size, which is determined by the fragmentation threshold. By adapting this parameter appropriately, the probability of successful transmission is improved under poor channel conditions.

The ChaPLeT framework improves service differentiation by selecting the parameters to adapt and changing their values based on the priority of the flows in the network. Using these adaptation policies, ChaPLeT shows an improvement in average throughput of up to two times under diverse network scenarios and can avoid problems such as priority inversion and starvation for high- and low-priority flows.

Another mechanism to address time-varying network conditions is adaptive EDCF (AEDCF) [7]. This work focuses on reducing the impact of collisions caused by high network load. The goal of this scheme is to adapt the CW appropriately to reduce the number of collisions, thereby improving goodput under high network load conditions. In the basic EDCA scheme, after each transmission the value of the CW is reset to the CW_{min} value. AEDCF, on the other hand, uses the past value of the CW as an indicator of the current network load rather than starting with the default value of CW_{min} under any network condition. After successful transmission, AEDCF sets the CW using both the previous value of the CW and the average collision ratio experienced. The extent of CW adaptation depends on the priority of the flow. With the AEDCF scheme of determining the CW, the achieved throughput and service differentiation are improved by using values reflective of the current network load and the priority of the traffic. It is shown that the AEDCF policy increases goodput up to 27 percent and improves medium utilization by 20 percent.

The final work we describe is adaptive fair EDCF [6]. This technique is focused on improving node performance by taking into account the network load and removing network inefficiency. Adaptive fair EDCF achieves this goal by adapting two parameters, the CW and the backoff mechanism. In order to avoid collisions when the network load is high, the proposed adaptation takes the precautionary action of expanding the CW whenever a deferral occurs. The proposed CW magnification is similar to the case when a retransmission occurs. Additionally, to improve performance, the scheme proposes to use an *adaptive fast backoff* mechanism, which speeds up the backoff countdown when the channel is idle. Since the criteria for the adaptive fast backoff mechanism and the upper limit of the CW are dependent on the AC, this adaptation policy improves service differentiation as well as the performance of the nodes. Performance evaluations show that the above schemes achieves 33 percent gain in goodput in a fully loaded network and achieves fairer allocation of resources to different priority traffic (i.e., audio, video, voice) than EDCA and AEDCF.

Summary and Future Directions — In summary, the approaches described above adapt various link layer parameters to improve the performance of nodes dealing with channel variance and high network load. Some of the methods proposed complement each other. For example, the fragmentation technique of ChaPLeT can be combined with both AEDCF and adaptive fair EDCF to adapt to poor channel conditions. Similarly, the adaptive fast backoff mechanism of adaptive fair EDCF and defer countdown adaptation of ChaPLeT can be combined to speed up the backoff counter when the node senses the channel is idle or has waited excessively.

Additionally, the effectiveness of the adaptation techniques differs depending on the present network condition. For example, the persistence factor adaptation in ChaPLeT is more suitable when retransmissions are caused by poor channel conditions. On the other hand, adaptive fair EDCF's technique of expanding the CW based on deferrals makes it more appropriate when the network load varies. AEDCF fares better when the network load does not vary significantly. Since the adaptation uses the past CW value as an indicator of network load with AEDCF, the CW value does not have to be recalculated with every deferral, as in the case of adaptive fair EDCF.

In the future, in order to address time-varying network conditions comprehensively there is a need to investigate adaptation of other 802.11e specific parameters in EDCA, such as TXOP, CW range, and AIFSN. Additionally, future work is necessary to understand if these techniques can be combined with changing retransmission limits and modulation schemes to improve differentiated service. Finally, there is a need to focus on the impact of time varying network conditions in the scheduling of nodes in HCCA. With a fixed set of nodes scheduled in HCCA and awareness of their experienced channel rate, appropriate time allocation can be given to each flow.

Adapting to Varying Application Profiles

The second set of challenges in quality-enhanced networks is adapting to varying application profiles. The traffic load generated by the application and the associated QoS requirements define an application profile. The QoS requirements of a flow, such as throughput, delay, and jitter, can vary significantly depending on application type and can also change over time. If the application traffic profile can be estimated correctly, the medium access mechanisms can be tuned accordingly to provide the desired QoS without degrading the utilization of the medium and affecting fairness. While variations in application profiles can affect flows in both EDCA and HCCA, the effect has been studied primarily in the context of HCCA flows with real-time requirements.

In the centralized scheme of HCCA, the scheduling scheme is designed to provide QoS guarantees. The reference scheduler of HCCA uses a round-robin method and allocates time based on the reservation information sent by the nodes. The reservation information used in the reference scheduler only includes averaged values of the traffic characteristics of the flows, such as mean packet size and mean required throughput. Hence, HCF can only allocate a fixed polling schedule suitable for constant bit rate (CBR) traffic.

As described in [8, 9], this type of static scheduler may be unsuitable for the dynamic traffic requirements associated with some applications. For example, some multimedia traffic flows, such as quality-controlled MPEG4 or videoconferencing, do not have the CBR profile and instead use variable bit rate (VBR) encoding. Service providers commonly use VBR encoding of multimedia content to increase the capacity of the network by multiplexing between different flows. With non-CBR traffic, HCCA scheduling can be inefficient and unsuitable during the polling-based medium access. Although one possible solution would be to increase the TXOP by a fixed amount, this solution would lead to a decrease in the total capacity of the network.

In addition, depending on the application, accurate information about requirements of real-time flows may not be available at the beginning of the flow's transmission. For example, with the use of application layer adaptation techniques, the bit rate requirements of a video stream change based on the congestion in the end-to-end network; hence, it is not feasible to get accurate information at initialization.

Approaches — To address the issues of traffic variability and varying application requirements, there has been some work suggesting changes in the scheduling algorithm used in HCCA. To address this challenge, we describe three proposed techniques that adapt scheduling parameters, such as TXOP, service interval, and polling schedule.

The goal of the first scheme [10] is to improve the performance of the scheduler and enhance its flexibility. The work proposes using the mean TXOP as a guideline for allocating time and uses a token bucket scheme to allow nodes to vary their TXOP over time according to their needs. The authors also propose to change the service interval for each node based on the traffic profile, as opposed to using a constant SI for all nodes, and use earliest deadline first to determine the polling order. The above techniques provide flexibility for scheduling flows with varying application profiles. It has been shown that the proposed flexibility in the scheduler for voice and video traffic leads to significant reduction in average transmission delay (up to 33 percent) and packet loss ratio (up to 50 percent).

The next proposed adaptation we present is flexible HCF (FHCF) [8], which tries to address dynamic variability in traffic flows by adjusting the TXOP of each flow using queue length estimations. The AP uses the queue length information to calculate the current demands of the flows. If the current demand is more than the reserved time, the TXOP is increased. Similarly, the TXOP is reduced when the demand is less. When there is insufficient time to accommodate the increased TXOPs, the adaptation allocates the additional time between the nodes in a fair manner. Using this adaptation, the mean delay can be improved significantly: up to 13× reduction.

	Proposed adaptations		
	TXOP	**SI**	**Other**
Grilo *et al.* [10]	Maintain average TXOP value but allow flexibility	Allow each flow to have varying SI values	Ordering of the flows is by earliest deadline first (EDF)
Ansel *et al.* [8]	Determine based on queue length	Keep SI fixed	When no time left in HCCA, proportionally decrease or increase TXOP
Ramos *et al.* [9]	Keep TXOP value fixed but poll needing flows again after original polling schedule	Keep SI fixed	When no time left in HCCA, allow EDCA access if EDCA load is low

■ Table 1. *Comparison of proposed adaptations to address traffic variability in HCCA.*

The final work we describe addresses traffic variability by dynamically allocating additional time to VBR flows and is described in [9]. Rather than adjusting the TXOP, the adaptation proposes additional polling in order to avoid delays for the nodes scheduled later in the polling list. The decision to perform additional polling is made by the AP and is based on queue length information received from the nodes and the timestamp of when the node was given extra time in the HCCA. Similar to the adjustment of the TXOP, the additional polling is limited to the amount of time remaining in the HCCA period. Fairness is ensured by incorporating the weight of each flow, which factors in the last time the flow was adapted. The proposed adaptation was able to show a significant improvement for HCCA VBR flows, up to 40 times improvement in experienced delay.

Summary and Future Directions — Each of the described techniques improves the HCCA scheduler's ability to handle variable traffic flows. Note that the percentage improvement from the three techniques cannot be compared directly, since the evaluation framework used was different in each work. However, the major differences between the techniques can be summarized in Table 1. In terms of relative effectiveness, the technique presented in [10] would be the most effective, but has the overhead of supporting variable service intervals. Our evaluation shows that the approach presented in [8] provides better QoS when there are a number of VBR flows present, whereas the technique in [9] works better in a mixed traffic scenario.

Future directions in handling variability in application profiles include a better understanding of the traffic profiles of different real-time applications, accurate modeling, and appropriate scheduling mechanisms and adaptation based on the classified traffic profile. Additionally, the effect of application variability in EDCA flows should also be investigated. Beyond the link layer, mechanisms using cross-layer coordination (i.e., application-level adaptation and link layer scheduling) need to be developed to address application variability.

Managing Link Layer Resources

Although 802.11e provisions for service differentiation, there remains a need to manage and coordinate link layer resources and QoS parameters. In addition to handling network conditions and adapting to varying application requirements, the AP must also take into account system parameters and overall network goals in its management scheme. The management of the network in 802.11e must include techniques to coordinate between the distributed and centralized periods of 802.11e and an admission control scheme to limit the number of flows to provide stronger guarantees of QoS. Additionally, appropriate configuration of network and access parameters is necessary to ensure system differentiation.

The first aspect of MAC resource management is coordination between the two channel access periods and deciding the appropriate channel access mode. The current approach to mapping between the two periods is to restrict real-time flows to the HCCA period and map other traffic flows to the EDCA period. However, as shown in [9], this mapping may lead to high experienced delay for nodes with real-time variable traffic in HCCA. Additionally, it can lead to network inefficiency. One example of this inefficiency is when a node that has high-priority traffic in EDCA experiences low throughput despite available time in the HCCA period.

A related challenge to coordination is the management task of admission control. Without a form of admission control, the network, despite provisions for service differentiation, will be unable to support QoS adequately. Although the 802.11e draft includes an admission control policy to be used with the reference scheduler, the policy determines the number of nodes to admit in HCCA, assuming all nodes have a CBR profile, and does not describe an admission control policy to limit nodes in the EDCA period.

Approaches — Work in the area of managing link layer resources with 802.11e has been limited, and has mainly been in the area of admission control improvements. We describe three techniques that have been proposed, one focused on coordination between the HCCA/EDCA techniques and two efforts providing EDCA admission control.

In terms of coordination between the HCCA and EDCA period, we revisit the adaptation presented in [9]. In addition to enhancing the scheduling algorithm, the work also describes an algorithm to dynamically associate traffic flows to the two medium access modes. The proposed algorithm mainly monitors the load in the two modes of access and decides the appropriate access mode based on queue size and estimated load. With this proposed method, when HCCA flows require additional time, the AP encourages flows to send during the EDCA period only if the network load in EDCA is low. Rather than reduce the time allocation of other flows already scheduled in HCCA, the algorithm attempts to give variable traffic flows an additional opportunity to send without negatively affecting the other HCCA nodes. Additionally, depending on the time available in the HCCA period, the algorithm chooses EDCA nodes based on their weights to decide which EDCA flow can be polled during the HCCA period. As shown in [9], the percentage of improvement gained using the transition of nodes from one period to another ranged from 15 to 36 percent for EDCA flows and was approximately 2× for HCCA flows assuming HCCA occupies a third of the channel access time within a beacon interval.

In the field of admission control, in [11], the AP attempts to limit the transmission duration of each AC. Mainly, the proposed work introduces a transmission budget for each AC, and this value is announced every beacon interval. Whenever a packet is sent during that beacon interval, the transmission budget for the appropriate AC is decreased, and all overhearing stations update their internal variables. After the allocated time for the AC has been used, transmissions must wait until the next beacon. The results indicate that with the use of the proposed admission control, the flows in EDCA can achieve the desired throughput, and the total throughput can be improved due to less contention.

Another piece of work in admission control is [12], where the AP limits the number of nodes allowed in the EDCA period. Whenever a new flow request is received, the algorithm estimates the effect of the admittance of the new flow on the achievable throughput of flows already in the network. For each flow, the achievable throughput is computed using monitored collision statistics and is compared with its requested bandwidth. If a node is unable to meet its desired throughput, the AP attempts to change one of two parameters, CW or TXOP. By reducing the CW or increasing the TXOP, the algorithm's desired effect is to increase the probability of accessing the network. If the CW or TXOP cannot be changed, the flow is rejected and the parameters are reset. The results indicate that the proposed admission control policy is able to prevent overloading the EDCA.

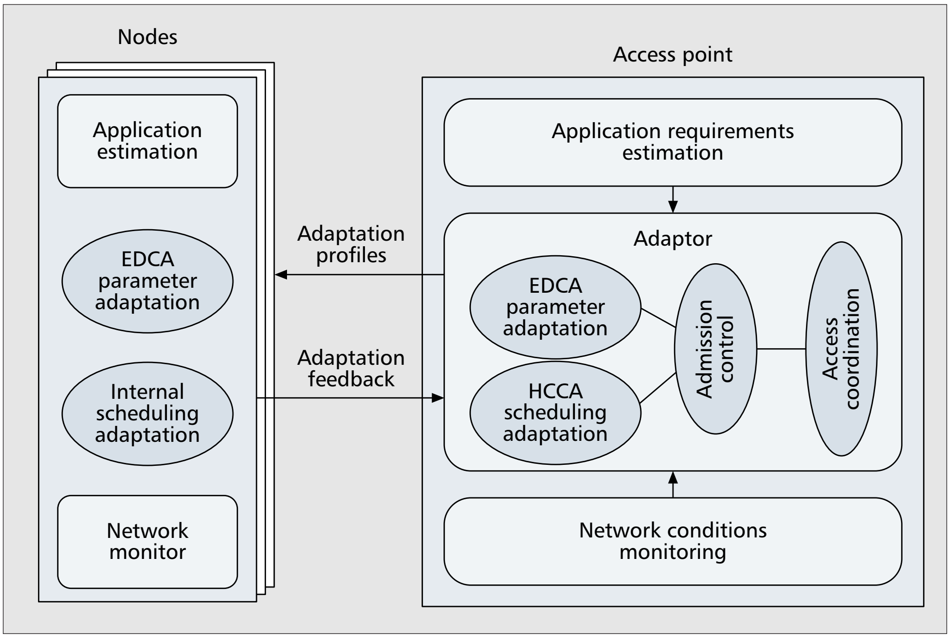

Figure 4. *Proposed HCF control framework.*

Summary and Future Directions — In summary, managing link layer resources in 802.11e networks can improve network efficiency. Although there has been some preliminary work in this area, further studies should look beyond queue-based adaptation for moving flows between channel access periods and incorporating other factors, such as inaccurate or misrepresented reservation information. Additionally, there is a need to determine appropriate parameters for different access categories taking into account node characteristics and conditions. Some example parameters include the ratio of centralized to distributed channel access, as well as the channel access parameters for the EDCA period. The values for each AC are important for determining the extent of service differentiation in the EDCA network. Finally, there is also a need to exploit the flexibility of the scheduler in designing an efficient admission control algorithm to maximize the capacity of the network.

The Proposed HCF Controller: An Integrated Approach to QoS Support

To enable comprehensive QoS support, we discuss the potential of creating a new HCF controller framework that can support the approaches presented and future work in QoS adaptation. Figure 4 illustrates the control framework, which takes into account channel conditions, application profiles, and some feedback on the effects of the adaptations.

The HCF framework must be aware of the possible management and coordination techniques available to adapt to the current conditions. Some of the key blocks of the framework on the AP side include access coordination, admission control, EDCA policy and parameter adaptation, and HCCA scheduling adaptation. The access coordination block consists of techniques that address both the HCCA and EDCA, such as the ratio between the periods. After coordinating between the two modes of access, the admission control policy is then used to limit the number of flows in both HCCA and EDCA. Once the flows are admitted to either the HCCA or EDCA period, the HCF controller must configure the appropriate parameters associated with each period to address the challenges identified. For EDCA, the appropriate EDCA adaptation policy for the network is selected. One example of such an adaptation policy is the mechanism used to determine the CW setting in the network. The EDCA block must also beacon out the parameter values for the various ACs. The final block in the AP is used to determine the HCCA scheduling adaptation. In this block the AP must decide the polling schedule, service interval, and TXOP durations of each flow. Finally, at the node level, each of the nodes has its own adaptations for its EDCA parameter settings and HCCA internal transmission scheduling.

By monitoring the network load and understanding the application requirements of the nodes scheduled in the network, the HCF controller can make decisions on these various aspects and can send the appropriate adaptation profile to the nodes. In order to enable and improve adaptation at the AP, each of the nodes provides feedback on the effect of the adaptation, as well as the monitored application requirements and network conditions experienced by the node.

Future Directions and Conclusions

In this article we describe three major challenges in improving QoS support in IEEE 802.11e networks. Next, we briefly discuss other related challenges that need to be addressed. An important problem is misbehaving nodes in the context of 802.11e networks. With service differentiation techniques, malicious nodes can use inaccurate reservation information in HCCA and change channel access parameters or AC classifications to gain an unfair advantage. Hence, techniques that identify these situations and reduce the impact of misbehaving nodes on other traffic flows are needed.

Another future direction is to understand the interaction of 802.11e techniques with applications and higher-layer QoS schemes. Although there is a suggested mapping between application type and access category, the method of mapping application requirements from higher to lower layers or reservation requests has not been investigated in depth. Studies on the impact of 802.11e combined with existing QoS schemes, such as IntServ and DiffServ, should be performed to understand the implications of 802.11e schemes on end-to-end performance.

Finally, there is a need to evaluate the trade-off between QoS techniques and energy consumption. Due to limited battery resources, energy consumption is a challenge in wireless networks and may be a larger problem in 802.11e networks. In EDCA, service differentiation can lead to longer wait times depending on traffic priority, thus affecting energy consumption. Additionally, depending on the scheduling scheme, nodes must remain awake to answer AP polls in the HCCA period. Techniques need to be developed that incorporate the battery resources and sleeping schedules of flows in the network.

In conclusion, in this article we provide an overview of challenges facing the 802.11e implementation of QoS and provided a survey of techniques associated with these efforts. We also discuss the potential of creating a new HCF controller framework that can support the approaches presented. By developing adaptation techniques to address the challenges presented in this article, we can provide improved service in future WLAN networks and enable comprehensive QoS.

Acknowledgments

The authors would like to acknowledge the support of the University of California, San Diego Center for Wireless Communications (CWC) and the U.C.-Discovery Grant (com02-10123).

References

[1] IEEE-802.11WG, "IEEE 802.11e Standard Draft/D8.0: Draft Supplement to Standard for Telecommunications and Information Exchange between Systems — LAN/MAN Specific Requirements — Part 11: MAC enhancements for quality of service (QoS)," Feb. 2004.

[2] H. Zhu and I. Chlamtac, "An Analytical Model for IEEE 802.11e EDCF Differential Services," *Proc. 12th Int'l. Conf. Comp. Commun. and Networks*, Dallas, TX, Oct. 2003.

[3] S. Choi, J. Prado, and S. Shankar N, "IEEE 802.11e Contention-Based Channel Access (EDCF) Performance Evaluation," *Proc. IEEE ICC*, 2003.

[4] J. Prado, "Mandatory TSPEC Parameters and Reference Design of a Simple Scheduler," Nov. 2002.

[5] N. Ramos, D. Panigrahi, and S. Dey, "ChaPLeT: Channel-dependent Packet Level Tuning for Service Differentiation in IEEE 802.11e," *Proc. Int'l. Symp. Wireless Pers. Multimedia Commun.*, Yokoshuka, Japan, 2003.

[6] M. Malli et al., "Adaptive Fair Channel Allocation for QoS Enhancement in IEEE 802.11 Wireless LANs," *Proc. IEEE ICC*, Paris, France, 2004.

[7] L. Romdhani, Q. Ni, and T. Turletti, "Adaptive EDCF: Enhanced Service Differentiation for IEEE 802.11 Wireless Ad Hoc Networks," *Proc. IEEE Wireless Commun. and Net. Conf.*, New Orleans, LA, 2003.

[8] P. Ansel, Q. Ni, and T. Turletti, "An Efficient Scheduling Scheme for IEEE 802.11e," *Proc. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, 2004.

[9] N. Ramos, D. Panigrahi, and S. Dey, "Dynamic Adaptation Policies to Improve Quality of Service of Multimedia Applications in WLAN Networks," *Proc. Int'l. Wksp. Broadband Wireless Multimedia*, San Jose, CA, 2004.

[10] A. Grilo, M. Macedo, and M. Nunes, "A Scheduling Algorithm for QoS Support in IEEE 802.11E Networks," *IEEE Wireless Commun.*, June 2003, pp. 36–43.

[11] Y. Xiao and H. Li, "Evaluation of Distributed Admission Control for the IEEE 802.11e EDCA," *IEEE Radio Commun. Supp., IEEE Commun. Mag.*, Sept. 2004, pp. S20–S24.

[12] D. Pong and T. Moors, "Call Admission Control for IEEE 802.11 Contention Access Mechanism," *Proc. IEEE GLOBECOM*, 2003.

Biographies

NAOMI RAMOS [S'00] (naomir@ece.ucsd.edu) received a B.S. degree in computer engineering from the University of Washington. She received an M.S. degree in electrical and computer engineering from the University of California, San Diego, and is currently a Ph.D. candidate in electrical and computer engineering at the same university. From June to September 2002, she worked at TRW/Northrop Grumman. Her research interests are in the area of wireless local area networks, with an emphasis in data link layer adaptations for energy efficiency and quality of service support for multimedia applications.

DEBASHIS PANIGRAHI [S'99] (dpani@ece.ucsd.edu) received a B.Tech. degree in computer science and engineering from the Indian Institute of Technology, Kharagpur. He holds an M.S. degree in electrical and computer engineering from the University of California, San Diego, and is currently a Ph.D. candidate in the Electrical and Computer Engineering Department at the same university. Since November 2004 he has also been affiliated with Ortiva Wireless. His research interests include wireless network optimization, application adaptation for wireless networks, wireless sensor networks, and embedded system design.

SUJIT DEY [S'90, M'91, SM'03] (dey@ece.ucsd.edu) is a professor with the Department of Electrical and Computer Engineering, University of California, San Diego, where he heads the Mobile Systems Design and Test Laboratory, engaged in developing configurable platforms, consisting of adaptive wireless protocols and algorithms, and deep submicron adaptive systems-on-chip for next-generation wireless networks and appliances. He is affiliated with the California Institute of Telecommunications and Information Technology (Cal-IT2), and the UCSD Center for Wireless Communications. He is also the founder and CEO of Ortiva Wireless, a company developing technologies and products for wireless multimedia. Prior to starting Ortiva, he was an advisor to ST Microelectronics and NEC USA, and chaired the Advisory Board of Zyray Wireless until its acquisition by Broadcom in July 2004. Prior to joining UCSD in 1997, he was a senior research staff member at the NEC C & C Research Laboratories, Princeton, New Jersey. He received a Ph.D. degree in computer science from Duke University, Durham, North Carolina, in 1991. He has co-authored more than 140 publications, including journal and conference papers, a book on low-power design, and several book chapters. He is the co-inventor of 10 US patents, with several others pending. He received Best Paper awards at several conferences. He has been General Chair, Program Chair, and a member of organizing and program committees for several IEEE conferences.