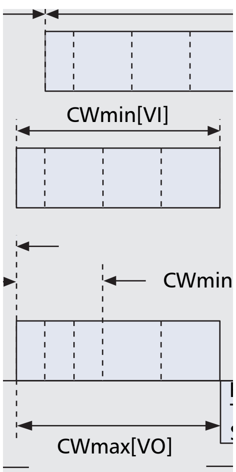
PROVIDING QOS SUPPORT AT THE DISTRIBUTED WIRELESS MAC LAYER: A COMPREHENSIVE STUDY

ORESTIS TSIGKAS AND FOTINI–NIOVI PAVLIDOU, ARISTOTLE UNIVERSITY OF THESSALONIKI



This article aims to provide a comprehensive study of the limitations and merits of mechanisms that have been proposed toward embedding QoS support to distributed wireless MAC protocols.

ABSTRACT

Holding the promise of making ubiquitous mobile access to IP-based applications and services a reality, wireless local area networks have been deployed in an unlimited way over the last few years. Due to their robust characteristics, distributed MAC protocols are the most widely used mechanisms to arbitrate access to the wireless channel. However, their ability to achieve high medium usage efficiency while providing services with meaningful performance assurances is being challenged by a wide range of existing and emerging applications that have lately migrated from other telecommunication networks to wireless environments. This article aims to provide a comprehensive study of the limitations and merits of mechanisms that have been proposed toward embedding QoS support to distributed wireless MAC protocols. A hybrid scheme that incorporates signaling and information sharing is proposed, and extensive simulation experiments are run to assess the efficiency of the access schemes in maximizing utilization of the wireless bandwidth while providing QoS support for heterogeneous applications.

INTRODUCTION

Since their standardization and development as a commercial product, wireless LANs (WLANs) have known increasing popularity, which is not spurred merely by the vision of optimal mobile connectivity. The intense drive to deliver high-speed Internet service to home users has seemingly tapped every conceivable pathway: copper and fiber optic cable, the air, power lines, even natural gas pipes. Among these emerging technologies, wireless access has prevailed as the most cost-effective and reliable way to confront the "last mile" access problem. Moreover, recent advances achieved in the physical layer have led to bit rates that were until recently achieved only in wireline networks, resulting in the deployment of WLANs to build high-speed backbone networks, along with their operation as last mile access networks. Last but not least, as the advent of satellite communications is still a forthcoming hope, WLANs comprise the most reliable solution to provide hassle-free and high-performance networking to areas where retrofitting is impossible.

WLANs own their popularity mainly to the powerful characteristics of distributed medium access control (MAC) protocols, which allowed the proliferation of low-cost broadband air interfaces and simplified the installation of WLANs while minimizing management and maintenance costs. However, their efficiency is being challenged by the emergence of applications with diverse throughput, loss rate, delay, or delay jitter requirements. Real-time video conferencing, Internet telephony, streaming stored video and audio, Web browsing of media-rich Web pages, and e-commerce find their way through the Internet, in addition to traditional data applications (e.g., telnet, file transfer, email). Moreover, as new applications are expected to thrive, the arguably limited quality of service (QoS) capabilities of distributed MAC schemes will be a serious impediment in deploying Internet-based applications in wireless networks.

Over the last few years, there has been vast research activity to define mechanisms and algorithms at the medium access layer for service differentiation, and continuous QoS provision to Internet traffic over WLANs. The underlying idea was that the mere extension of QoS control mechanisms, based on the hard QoS approach, to the wireless environment would be insufficient; therefore, soft QoS guarantees would have to be provided instead. Unfortunately, nothing could be further from the truth. Most studies provide the means just for service differentiation, which is of little significance when applications with diverse and stringent QoS are considered and populated wireless networks are examined. Providing soft QoS guarantees has proven to be far more difficult and complex than initially envisioned, since few proposed medium access schemes manage to provide meaningful QoS assurances.

This article aims to identify the progress toward the provision of QoS support to WLANs at the distributed access medium layer. Although mobility in wireless networks, hidden terminal occasions, and the intrinsically unstable nature of wireless channels pose a number of unique challenges of their own when QoS provisioning is considered, we confine our study to a single

1536-1284/08/$25.00 © 2008 IEEE

broadcast region and an ideal channel to focus our attention on the challenges that arise from the distributed nature of medium access schemes. We then discuss the approaches that have been proposed for providing different types of QoS and bring out the issues unique to wireless MAC protocols. A number of different medium access protocols embedded with QoS capabilities are examined, and a hybrid scheme is proposed. Finally, the ability of each mechanism to provide QoS support to current and future wireless networks is evaluated.

What Raised the Need for QoS Support?

The Internet architecture was designed with one major goal in mind: survivability. In this context the fundamental service model of the Internet, best effort delivery of IP packets, was concerned almost exclusively with reliable delivery of data content, as average performance guarantees were sufficient for the first data applications. Moreover, the Internet was founded on the concept that flow control would be performed in the end systems as a response to congestion signals. However, the Internet quickly burgeoned into a publicly accessible network, and some users intentionally misbehaved in order to capture more bandwidth. Moreover, real-time applications that invoke User Datagram Protocol (UDP) rather than Transmission Control Protocol (TCP) as their transport protocol captured more than their fair share of bandwidth.

The unfair sharing of network resources, along with the growing exigency for effective support of the diverse requirements of emerging multimedia applications, prompted the development of mechanisms, algorithms, and QoS frameworks for continuous provision of QoS to Internet traffic over wired and wireless networks.

Fairness

By arbitrating access to the shared channel, medium access algorithms have a direct impact on the fair sharing of the available raw bandwidth among the contending stations. Therefore, a MAC protocol should not exhibit preference to any single node when multiple nodes contend for channel access. As early as 1994, the authors in [1] noticed that MACAW, the access scheme of DCF, may result in the channel being captured by only one station, and stated that the streams of packets belonging to separate flows would have to be treated independently by the medium access scheme if fairness were to be provided.

More sophisticated distributed algorithms based on MACAW were later proposed to provide fair access and achieve rate-based differentiation [2–4]. However, the tight coupling between rate and delay under these schemes renders them inappropriate for providing delay guarantees, especially in the case of low-rate traffic requiring low delay bounds. Indeed, none of the schemes proposed in the above studies addresses the problem of delay differentiation.

Hard QoS Guarantees

Medium access schemes that aim to provide hard QoS guarantees follow the paradigm of algorithms and scheduling techniques developed within the framework of integrated services (IntServ) [5].

The IntServ QoS framework proposed extensions and modifications to the Internet architecture, protocols, and infrastructure to control bandwidth sharing among different traffic classes and provide deterministic QoS guarantees; that is, hard bounds on end-to-end packet delays, packet losses, and variation in queuing delays. Central to the conceptual foundation of the IntServ framework is the notion of *flow*. The fundamental assumption is that if users specify their QoS requirements by efficiently mapping user perceptual parameters into system QoS parameters, and characterize the nature of the traffic they inject into the network, routers will be able to meet the QoS requirements of each single flow by committing to provide a certain amount of service.

Four components are defined as necessary extensions to best effort networks so that deterministic or statistical end-to-end QoS assurances can be provided: the packet scheduler, admission control routine, classifier, and reservation setup protocol. The capstone of the research effort to provide end-to-end deterministic QoS guarantees in packet-switched networks is the theory of network calculus [6], which provides powerful tools to derive buffer requirements for loss-free operation and bounds on end-to-end delay for any network topology that implements the IntServ model.

The exact causes that prevented IntServ from becoming a widespread technology in wired networks, in spite of its appealing features, are beyond our ken. However, three limitations of IntServ are discussed with respect to distributed wireless networks.

The per-flow approach of IntServ constituted a fundamental change to the best effort network, imposing modifications to various network elements. The market potential at the time that the IntServ model was proposed did not justify the cost of upgrading or substituting a vast number of network elements in order to support end-to-end QoS. Indeed, most users seemed to accept the often inferior quality of perceived performance in exchange for the low cost. However, since the inception of IntServ, there have been significant changes in the telecommunications arena, and the numerous users of wireless networks seeking QoS support, the balanced costs of WLANs, and multiple vendors are enough to pave the way to the proliferation of QoS-aware wireless architectures.

Second, the main concern with hard QoS provisioning is scalability. Per-flow state has to be maintained in routers, and this information has to be retrieved to process each incoming packet. The high cost of per-flow management in processing resources is expected to increase in wireless networks, given the limited processing power of mobile devices. However, recent advances in core-stateless architectures [7] could

The Internet architecture was designed with one major goal in mind: survivability. In this context the fundamental service model of the Internet, best effort delivery of IP packets, was concerned almost exclusively with the reliable delivery of data content.

Although many medium access schemes that adopt the Service Differentiation approach have thrived, the benefits stemming from their deployment in a network domain seem to be far fewer than the issues raised.

efficiently be applied to distributed wireless networks to allow for the provision of delay and rate guarantees, and provide end-to-end assurance levels similar to those that can be provided with per-flow mechanisms.

The third and most serious concern when considering hard QoS provisioning at the distributed medium access layer is the time-varying MAC capacity. Hard QoS can only be provided if the capacity assigned to flows is deterministic. In distributed wireless networks, packet transmissions take place in a completely stochastic way, occasionally resulting in collisions. Therefore, the deterministic capacity assigned to flows is zero. Two alternatives to hard QoS that hold considerable appeal in distributed WLANs are the provision of service differentiation and soft QoS guarantees.

Service Differentiation

Due to its potential to cater for real-time services and its low complexity, the service differentiation approach has gathered considerable momentum in the wireless community. Rather than providing quantitative performance guarantees of any type, many studies propose to offer only better than best effort services, by classifying data packets in distinct traffic classes and designing the medium access mechanism to use different contention parameters for each traffic class. Flows with strict QoS requirements (e.g., voice, video) are given preferential treatment by ensuring that their prioritized contention parameters will cause packets from all other flows to defer access.

Although many medium access schemes that adopt service differentiation have thrived, the benefits stemming from their deployment in a network domain seem to be far fewer than the issues raised.

While handling the classification of applications into appropriate service classes, the MAC layer has to specify a service class that best matches the QoS required by an application, preventing overallocation of network resources to it. However, most of the distributed MAC protocols achieve service differentiation by overprovisioning high-priority classes. In this way the performance of high-priority classes is protected; yet the usage efficiency of the scarce and still expensive wireless resources is dramatically reduced. Packets belonging to low-priority classes experience frequent QoS violations, as they are required to defer access to packets that get an excess amount of service. Eventually, network utilization is reduced, since valuable resources that can be used to accommodate new users are wasted for applications that are already adequately serviced.

The problem is exacerbated if wireless access is considered just another hop in the communication path. By applying the tools of network calculus, it is easily shown that the overprovisioning of a flow at a single point in the network does not enhance the overall performance experienced by the end user, and also results in excess resource requirements in downstream nodes, causing packets to be dropped. Moreover, when heterogeneous telecommunication platforms are considered, the mapping of traffic classes from one telecommunication platform to another is more easily said than done. Last but not least, the service differentiation approach does not provide a means of QoS provisioning within a traffic class. When a traffic class becomes populated, the medium access scheme fails to meet the QoS requirements of the flows sharing the same class.

Soft QoS Guarantees

In between the hard QoS approach, which fits badly into the framework of wireless networks, and the service differentiation approach, adopted by the vast majority of QoS-aware MAC protocols, is a class of medium access schemes that provides soft QoS guarantees. Soft QoS provision of a session is defined as the graceful acceptance of QoS specification violation over transient periods of time, provided that the session QoS requirements are honored over the total connection time.

By allowing the QoS commitments of sessions to be violated over short time periods, MAC schemes can significantly increase a wireless network's multiplexing gain. Hard QoS provisioning schemes are conservative in admitting new flows into the network, assuming that flows cooperate to simultaneously yield their worst case traffic behavior. However, the performance experienced by applications is typically much better than conservatively computed worst case bounds.

With respect to end users, strict QoS guarantees have higher cost, since network resources assigned to a single session are not available to other flows. Many users would be willing to loosen their stringent QoS requirements and accept an imperceptible degradation of service in return for a reduction in cost. Furthermore, continuous media applications (video and audio) are resilient to infrequent packet losses without any quality degradation being perceived by end-users. Therefore, short-timescale QoS commitment violation will not harm the performance experienced by most multimedia applications.

In order to honor the specific needs of traffic classes and achieve high efficiency in terms of throughput, MAC protocols that offer soft QoS assurances require support from both medium access and scheduling algorithms. Indeed, in the literature there are a number of studies that aim to provide soft QoS guarantees at the distributed wireless MAC layer by applying the ideas behind scheduling algorithms proposed for wired networks. However, deploying existing scheduling algorithms proposed for wired networks to control channel access in a distributed wireless environment comprises a number of challenges. Since there is no management entity that can obtain all the information (e.g., number of active sessions, link states, status of session queues) needed to make a scheduling decision, it is the medium access algorithm that determines how close the transmission of packets is to the idealized schedule.

QoS-Aware Distributed MAC Protocols

The medium access algorithm of MAC protocols, as well as the QoS mechanism that they incorporate, have a direct impact on the type of QoS assurances that they can provide [8].

The approaches that have been proposed for infusing QoS capabilities in the distributed wireless MAC layer can be classified into three categories [9]. The first one uses prioritized contention and backoff parameters to allow faster access to the channel to traffic classes with higher priority. Based on the locally computed values of the parameters, stations determine individually when to access the medium. While this approach is the simplest to implement, it merely provides differentiated levels of performance. In the second approach nodes exchange information about the packets stored in their buffers, in order to assess their relative priority. Since nodes have incomplete knowledge of the priority indices of all other packets in the broadcast region, the probability that a high-priority packet is preempted by a lower-priority one is nonzero. In the third approach highest-priority nodes signal first to preempt lower-priority ones from gaining access to the channel. However, a higher probability of granting access to the highest priority packet comes at the expense of increased overhead.

Below we review four distributed MAC protocols. Their medium access algorithms are widely used to arbitrate access in distributed wireless networks, while the QoS mechanism they incorporate corresponds to one of the aforementioned approaches. Moreover, we propose a hybrid scheme that uses both signaling and information sharing to achieve higher efficiency. Examples of their channel access cycle and characteristics are presented in Fig. 1 and Table 1, respectively.

Priority Broadcast for DCF

Distributed coordination function (DCF) is the fundamental access method of the IEEE 802.11 standard [10]. According to DCF, if the medium is idle for a period of time equal to a distributed interframe space (DIFS), nodes may transmit their packet. Otherwise, they defer access until the end of the current transmission. Nodes generate a backoff counter, which is a random number of slots uniformly distributed in the range of 0 to *CW*. The Contention Window (CW) parameter takes an initial value of *aCWmin*, and is doubled at every unsuccessful attempt to transmit a packet, until it reaches the value of *aCWmax*. The *CW* is reset to *aCWmin* after every successful transmission. Each time the medium becomes idle for a period longer than *DIFS*, the backoff timer is decremented once every slot time; when it is zero, nodes are allowed to transmit.

To provide soft QoS guarantees using DCF, a distributed priority scheme is proposed in [11] that aims to approximate an ideal Earliest Deadline First (EDF) scheduler. The priority index of the transmitting nodes' head-of-line packet is piggybacked onto existing handshake messages of the 802.11 DCF. Neighbors monitor these transmissions and keep a table of the times in order to assess the relative priority of their own head-of-line packet. Upon the successful transmission of a packet in the table, each node removes the current entry from its scheduling table. A node defers from contention, as long as a time on its table precedes the arrival time of its own head-of-line packet. Specifically, given a node's j local scheduling table S_j, its rank r_j in its local scheduling table, and the packets' retransmission attempt l, the following equation is used to calculate the CW:

$$CW(S_j) = \begin{cases} Uniform[0, 2^l CW_{\min} - 1] & r_j = 1, l < m \\ CW_{\min} + Uniform[0, CW_{\min} - 1] & r_j > 1, l = 0 \\ Uniform[0, 2^{l+1} CW_{\min} - 1], & r_j > 1, l \geq 1 \end{cases}$$

EDCF

EDCF stands for enhanced DCF and is part of the 802.11e standard for service differentiation [10]. EDCF did not add any wisdom to the wireless community, as it is just the realization of many research efforts over the last few years to embed service differentiation capabilities in DCF. The 802.11e standard defines four access categories (ACs) labeled according to their target application: *AC_VO* (voice), *AC_VI* (video), *AC_BE* (best effort), and *AC_BK* (background). Differentiated behavior to different traffic classes is provided by introducing two modifications to the DCF. First, there are no global *CWmin* and *CWmax* values; rather, each traffic class AC has its own CW limits, *CWmin[AC]* and *CWmax[AC]*. Higher-priority traffic employs lower values for *CWmin* and *CWmax* than low-priority traffic. Second, the arbitrary interframe space (AIFS) is used, so high-priority traffic has a higher probability of making a transmission attempt. Instead of using DIFS for each traffic class, the backoff counter of traffic class AC may begin decrementing after *AIFS[AC]* time has passed from the end of the last transmitted frame. Differentiated control of access to the medium is provided by letting high-priority traffic have shorter AIFSs.

EY-NPMA

Elimination-Yield Non-Preemptive Priority Multiple Access is the medium access scheme of the HIPERLAN standard [13]. EY-NPMA provides QoS support by granting access to packets with lower residual lifetimes. Depending on their residual lifetime, packets are assigned one of the five priorities from 0 to 4, with 0 being highest. The synchronized channel access cycle comprises three phases: the prioritization, contention, and transmission phases.

The prioritization phase ensures that only the data transmission attempts with the highest channel access priority survive this phase. The contention resolution algorithm of EY-NPMA comprises two subphases: the elimination and yield phases. During the elimination phase, contending nodes transmit a channel access burst, whose length in slots is random between 0 and a

The service differentiation approach does not provide a means of QoS provisioning within a traffic class. When a traffic class becomes populated, the medium access scheme fails to meet the QoS requirements of the flows sharing the same class.

We used event-driven stochastic simulations to assess the efficiency of each medium access scheme in maximizing the effective use of the wireless bandwidth while providing QoS support for heterogeneous applications.

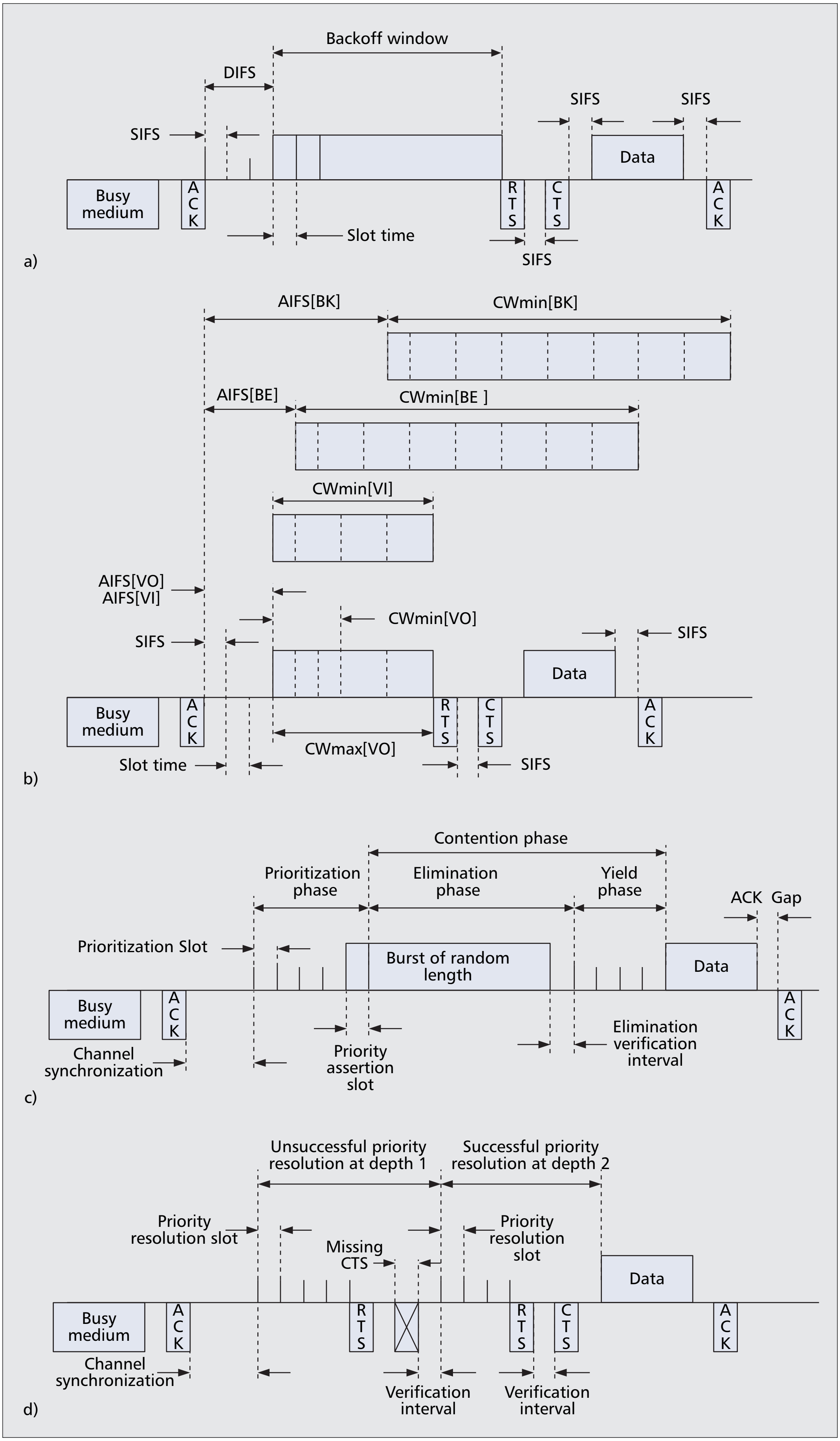


Figure 1. *Channel access cycles: a) PBDCF; b) EDCF; c) EY-NPMA; d) ATPB.*

Medium access scheme	QoS mechanism	Type of QoS support	Fairness	Collision rate	Overhead
PBDCF	Sharing information	Soft QoS	Poor	High	Moderate
EDCF	Prioritized parameters	Service differentiation	Poor	High	Moderate
EY-NPMA	Signaling	Soft QoS	Good	Low	High
ATPB	Signaling, sharing information	Soft QoS	Very good	Collision-free	Moderate

Table 1. *Properties of the four channel access schemes.*

predefined maximum according to a truncated geometric distribution, and then listen to the channel. If the channel is sensed as idle, nodes proceed to the yield phase, where they back off for a random number of slots, and if the channel is idle, they immediately enter the transmission phase by transmitting the packet stored in their buffer. All other stations sense the beginning of the transmission and refrain from transmitting.

Adaptive *m*-ary Tree Algorithms with Priority Broadcast

In [14] we proposed an adaptive medium access scheme, which is based on *m*-ary tree algorithms to resolve the priorities of contending nodes. The proposed medium access scheme builds a tree that is assumed to be dynamic and adaptive in the sense that the structure of the tree and the degree k of the root node are allowed to vary from epoch to epoch depending on traffic conditions and the deadline of packets.

At the beginning of each access cycle, all stations that intend to transmit a packet wait for a time period equal to a channel synchronization (CS) interval, and priority resolution starts at depth 1. Based on the residual lifetime (RL) of its head-of-line packet, a node determines the subtree to which it belongs, and senses the channel for as many slots as the index of the subtree. If the channel is idle for the whole sensing interval, the station transmits a request to send (RTS) message; otherwise, the source exits contention. If the clear to send (CTS) packet is correctly received, the source transmits its data. Otherwise, all sources that transmitted an RTS message continue the priority resolution procedure at depth 2, with a new RTS/CTS handshake. In the case of a missing CTS, the priority resolution procedure is continued in the same manner until a source is granted clear transmission of its data packet. The receiving node acknowledges the successful reception of a packet with an acknowledgment (ACK) message.

To compute k, stations overhear the transmitted information during the channel access cycles and deduce an upper limit RL^+ on the residual lifetime of the packet that is being transmitted. Let

$$\overline{RL_r^+}$$

denote the average upper limit of the residual lifetime of packets that were transmitted in the preceding r access cycles. Then, the degree k_r of the root node is calculated as

$$k_r = \left\lfloor \frac{S}{\overline{RL_r^+}} \right\rfloor.$$

The number r of access cycles over which k_r is calculated determines how fast the protocol responds to traffic variations.

To enhance further the performance of our proposed medium access scheme, a mechanism similar to the PBDCF is infused in it. Transmitting nodes piggyback onto existing handshake messages the value of k they compute based on the residual lifetime time of their HOL packet,

$$k_{HOL} = \left\lfloor \frac{S}{RL_{HOL}} \right\rfloor.$$

Neighbors monitor these transmissions and keep a table of the values of k_{HOL}. Then at the beginning of each access cycle, they compute k as the maximum of k_r and k^1_{HOL}, where k^1_{HOL} is the first entry in their table.

Performance Evaluation

We used event-driven stochastic simulations to assess the efficiency of each medium access scheme in maximizing the effective use of the wireless bandwidth while providing QoS support for heterogeneous applications. The simulations aimed to compare the medium access algorithms and not the respective implementations, as expressed in the standards. Toward this end, we assumed a channel rate of 24 Mb/s, while the duration of any slot used by the four protocols during the channel access cycle was set to the same value. The only exception was the *channel synchronization interval* in the case of EY-NPMA and ATPB, which was set equal to DIFS and *AIFS[VO]* (3 slots) to allow a fair comparison.

The physical channel was considered to be ideal; that is, the only reason behind erroneous reception was the simultaneous transmission of more than one station (packet collision). Furthermore, all network stations were within one hop of each other, thus eliminating the appearance of hidden/exposed terminals. The performance metrics of interest were the achieved medium utilization, the probability of correct scheduling, and the admissible region of each scheme. These metrics were examined for different node populations (1–200 stations).

Each station initiated four sessions carrying four types of data: voice, video, streaming audio, and Web traffic. The traffic characteristics as well as the QoS requirements of each session are

Application	Traffic characteristics				QoS requirements	
	Max. bit rate	Mean bit rate	Max. packet size	Mean packet size	Delay	Packet loss rate
Voice	64 kb/s	32 kb/s	400 bytes	400 bytes	Uniformly distributed, 1–10 ms	< 2%
Video	400 kb/s	64 kb/s	2048 bytes	1684 bytes	Uniformly distributed, 5–20 ms	< 1%
Streaming audio	32 kb/s	32 kb/s	2048 bytes	2048 bytes	Uniformly distributed, 20–100 ms	< 1%
Web browsing	64 kb/s	22.4 kb/s	2048 bytes	598 bytes	Uniformly distributed, 100–500 ms	0

Table 2. *Traffic characteristics and QoS requirements of the examined applications.*

depicted in Table 2. Upon their arrival, packets were assigned a lifetime which was equal to the delay budget associated with the flow that they belonged to. Packets that could not be delivered within the allocated lifetime were discarded.

Simulation Parameters and Traffic Source Models

Voice traffic: The widely adopted two-state on-off voice activity model with exponentially distributed duration of voice spurts and gaps was applied to voice sources. A voice source generated a signal that followed a pattern of talk spurts and silent gaps, with equal mean times of 500 ms each. A G.711 codec was considered that produced one 8-bit value every 125 µs, resulting in a 64 kb/s bitstream during talk spurt intervals. The encoded voice stream was packetized into 400-byte packets, resulting in packetization delay of 50 ms.

Video traffic: To better approximate compressed video traffic, real frame sizes of H.263 encoded video were used. The video frame traces were taken from a videoconference session [14]. The bit rate of the encoder output was 64 kb/s, its peak to mean bit rate ratio was equal to 6.2, and the peak to mean frame size was equal to 8.48. The maximum packet size was set to 2048 bytes.

Streaming audio: Streaming audio was modeled as a constant bit rate source, periodically generating packets at a rate of 32 kb/s. The packet size was set to 2048 bytes.

Web data traffic: The Web traffic model used was an on-off two-state model with exponentially distributed on and off times, with mean values of 0.7 s and 1.3 s, respectively. The on rate was set to 64 kb/s, and the size of the document page had a Pareto distribution,

$$p(x) = \frac{ak^a}{x^a},$$

where $\alpha = 0.9$ and $k = 150$ bytes.

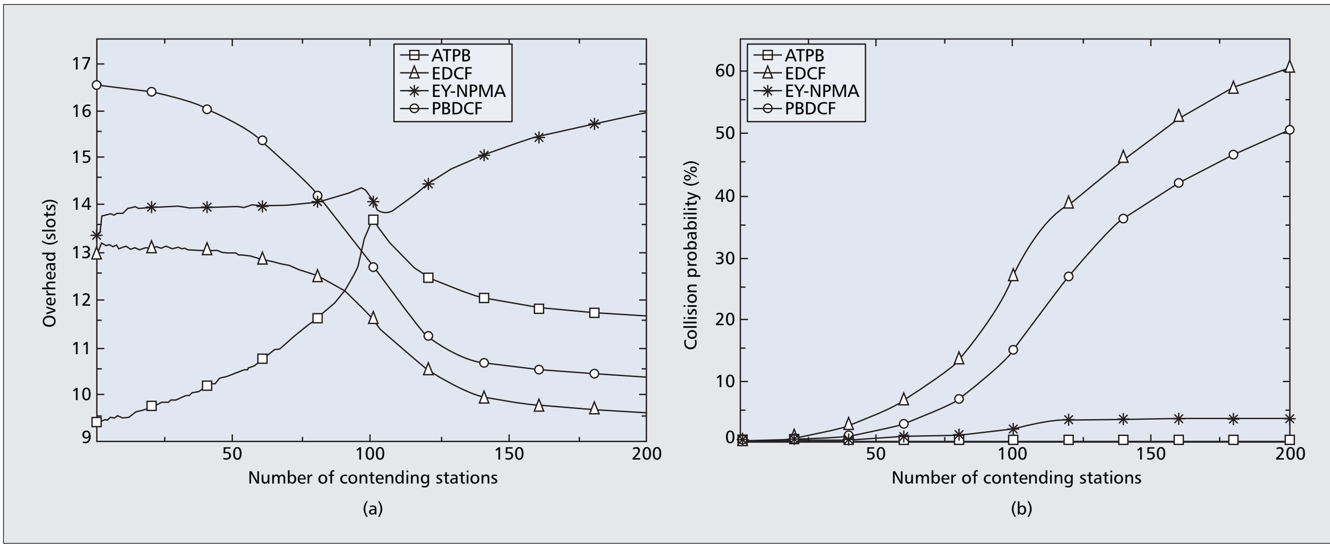
Simulation Study

The main performance metric of a distributed medium access scheme is its mean utilization of the available raw bandwidth, depicted in Fig. 3a. Its efficiency is governed to a great degree by the equilibrium of two conflicting qualities: collision rate and exhibited overhead. Arbitrarily low collision rates may be achieved at the cost of long access cycles and vice versa. Moreover, when considering QoS-aware protocols, medium utilization is tightly coupled with the ability to make the correct scheduling decision as well.

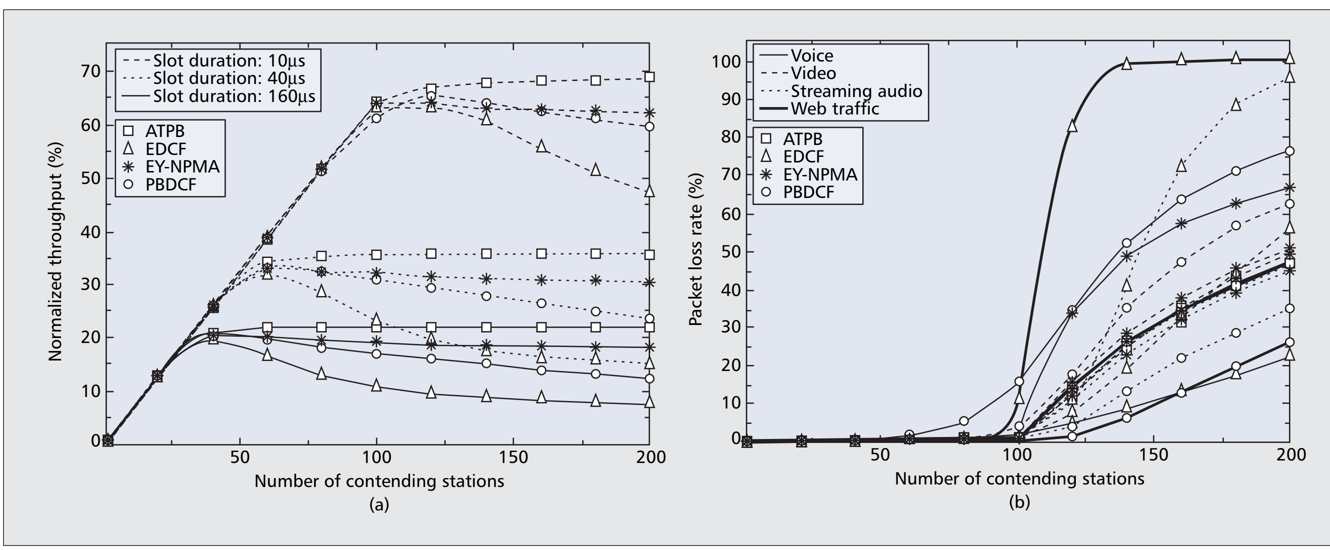
Figures 2a and 2b present the overhead exhibited by each scheme and the probability of a collision occurrence, respectively. While, the overhead of EDCF and PBDCF is rapidly reduced when the offered load increases, their high collision rate has an adverse impact on their achieved medium usage efficiency. Although collisions are limited just to the exchange of RTS/CTS messages, roughly 60 percent of the time is spent without any packet transmissions. The collision rate of EDCF is higher than that of PBDCF, since the low values of the contention parameters used for high-priority traffic result in frequent collision occurrences.

On the other hand, EY-NPMA and ATPB aim at maximizing the mean medium utilization by reducing or even preventing the collision of transmitted packets. Their low collision probability comes at the expense of increased overhead, yet their throughput remains high. The overhead of both schemes shows an increasing trend as traffic load increases, as more slots are spent to resolve the priorities of contending packets. Even though the zero collision rate of ATPB would justify a larger overhead than EY-NPMA, its adaptive qualities allow it to quickly detect the highest-priority node without spending a large number of slots.

As depicted in Fig. 3a, ATPB achieves the highest throughput of all four protocols, followed by EY-NPMA, PBDCF, and EDCF. When considering a slot duration of 10 ms, which is typical in current wireless networks, 30 percent or more of the available bandwidth is spent as overhead. The portion of wasted bandwidth is expected to increase in future wireless networks. As advances in the physical layer lead to higher bit rates, the duration of each slot becomes a significant fraction of the time needed to transmit the actual data payload. Because of the wireless environment, but also for technical reasons, there is a lower limit on the slot duration. Turnaround times, propagation delay, and delay spread demand that the slot duration for both bursting and backing off exceeds a cer-



■ **Figure 2.** *a) Overhead; b) collision rate of the four access schemes.*



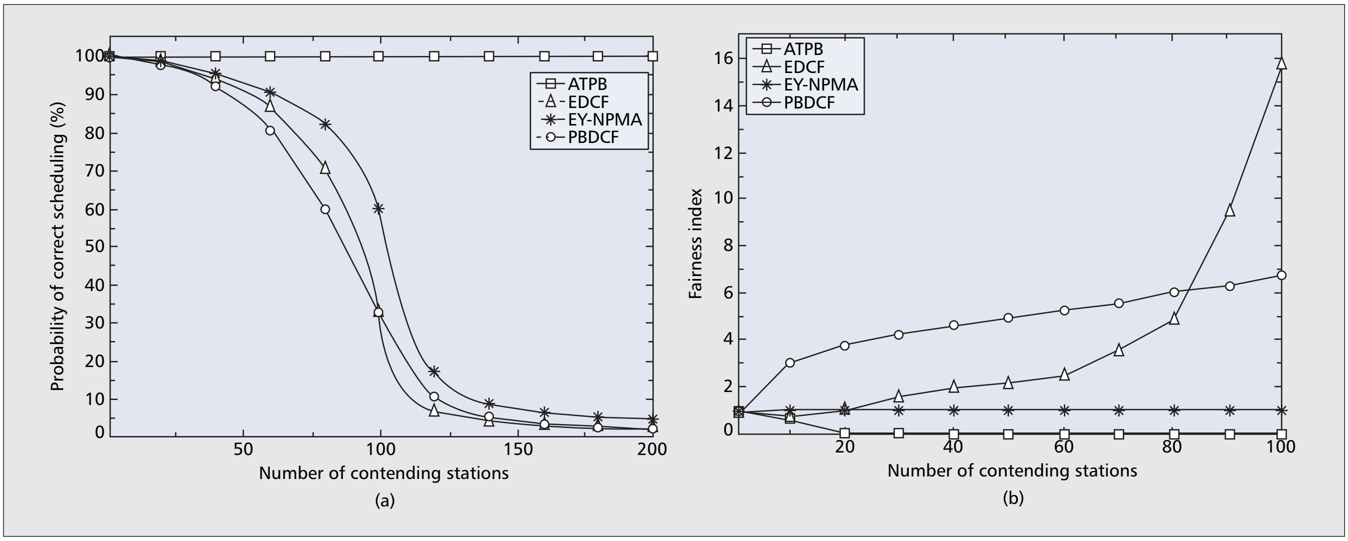
■ **Figure 3.** *a) Normalized throughput; b) packet loss rate of the four access schemes.*

tain threshold. The normalized throughput derived for slot duration of 40 µs and 160 µs is equal to the normalized throughput that would be achieved by the four access schemes if the bit rate of the physical medium was increased 4 and 16 times, respectively, and the slot duration remained 10 µs.

The increased medium utilization of ATPB in all environments is credited to its high probability of correct scheduling, presented in Fig. 4a. By granting access to the station that has the packet with the lowest residual lifetime, ATPB serves first those packets whose deadline is about to expire and would otherwise be lost. In contrast, all other schemes fail to make the correct scheduling decision, resulting in an increased portion of lost packets. Packets whose QoS requirements are about to be violated are preempted by packets that have enough residual lifetime to be transmitted in succeeding access cycles.

Figure 3b presents the packet loss ratio experienced by the four traffic types for each access scheme. It should be stressed that under ATPB all traffic types have the same packet loss probability. On the other hand, EY-NPMA and PBDCF exhibit an increased portion of lost voice packets. It turns out that their QoS mechanisms are insufficient when traffic with very stringent QoS requirements is to be supported. EDCF manages to protect voice packets by underprovisioning all other traffic classes. Under loaded traffic conditions, EDCF drops almost all of the packets belonging to Web sessions in order to ensure that voice sessions will get preferential treatment. Table 3 presents the maximum number of users that can be admitted into the network while ensuring that all four sessions meet their QoS requirements.

The ability of each scheme to provide fair sharing of network resources is depicted in Fig. 4b. To assess the fairness of each scheme, we assumed an environment where all stations generate packets at a constant bit rate equal to the channel rate; thus, all stations were continuously



■ **Figure 4.** *a) Probability of correct scheduling; b) fairness index of the four access schemes.*

Medium access scheme	Schedulability region
ATPB	93
EY-NPMA	82
EDCF	71
PBDCF	50

■ **Table 3.** *Number of users meeting their QoS requirements.*

backlogged. The instantaneous domination in the channel sharing of each scheme is described by its fairness index (FI). The FI is calculated as the probability that the previous successful station becomes the next successful transmitter multiplied by the number of contending stations. Therefore, medium access schemes that distribute the bandwidth fairly among the competing stations should have an FI equal to one or less. The unfair characteristics of PBDCF and EDCF become evident, as these schemes tend to favor the last transmitting station. On the other hand, the FI values of EY-NPMA and ATPB stay at almost the same level as the number of contending stations increases from 1 to 100. The FI of EY-NPMA is equal to one, meaning that all stations have equal opportunities to gain access to the channel. The zero FI of ATPB is ascribed to the fact that it always grants access to the packet with the lowest residual lifetime. Since we considered constant bit rate sources, the station with the lowest residual lifetime packet was other than the last transmitting one.

Conclusions

In this article we discuss the potential of providing QoS support at the distributed MAC layer. MAC protocols generally have a dominant effect on the ability of a wireless system to deliver on a QoS contract. Different medium access schemes, as well as approaches toward QoS provisioning, are spelled out, and a hybrid combination of them is proposed. We evaluate the ability of each scheme and its corresponding QoS mechanism to provide QoS support to heterogeneous applications while maximizing effective use of the wireless bandwidth. Extensive simulation results show that medium access schemes that use signaling and effectively adapt to traffic conditions can achieve high efficiency. However, further research has to be conducted to ameliorate the detrimental effects of node mobility, unstable channel conditions, and hidden terminal occurrences on providing QoS support to a distributed wireless network.

References

[1] V. Bharghavan *et al.*, "MACAW: A Medium Access Protocol for Wireless LAN's," *Proc. ACM SIGCOMM '94*, Aug. 1994, pp. 212–25.

[2] T. Nandagopal *et al.*, "Achieving MAC Layer Fairness in Wireless Packet Networks," *Proc. MOBICOM*, Boston, MA, 2000.

[3] N. H. Vaidya, P. Bahl, and S. Gupta, "Distributed Fair Scheduling in a Wireless LAN," *Proc. MOBICOM*, Boston, MA, 2000.

[4] A. Banchs and A. X. Perez, "Distributed Weighted Fair Queuing in IEEE 802.11 Wireless LAN," *Proc. IEEE ICC*, 2002.

[5] R. Braden, D. Clark, and S. Shenker, "Integrated Services in the Internet Architecture: an Overview," IETF RFC 1633, June 1994.

[6] J.-Y. Le Boudec and P. Thiran, "Network Calculus," *LNCS*, vol. 2050.

[7] I. Stoica and H. Zhang, "Providing Guaranteed Services Without Per Flow Management," *Proc. ACM SIGCOMM*, Boston, MA, Sept. 1999.

[8] A. Chandra, V. Gummalla, and J. Limb, "Wireless Medium Access Control Protocols," *IEEE Commun. Surveys & Tutorials*, vol. 3, no. 2, 2000.

[9] J. A. Stine and G. de Veciana, "A Paradigm for Quality-of-Service in Wireless Ad Hoc Networks Using Synchronous Signaling and Node States," *IEEE JSAC*, vol. 22, no. 7, Sept. 2004, pp. 1301–21.

[10] IEEE 802.11, "Wireless LAN Medium Access Control (MAC) and Physical Layer Specification," 1997.

[11]V. Kanodia *et al.*, "Distributed Priority Scheduling and Medium Access in Ad Hoc Networks," *Kluwer Wireless Networks*, 2002, pp. 455–66.

[12] IEEE 802.11e, "Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS)," 2003.

[13] ETSI, "Broadband Radio Access Networks: HIgh Performance Local Area Network (HIPERLAN) Type I: Functional Specification," EN 300 652 v1.2.1, 1998.
[14] O. Tsigkas and F.-N. Pavlidou, "An Adaptive Medium Access Control Protocol Using m-ary Tree Algorithms for Quality-of-Service Support in Single-Cell Ad Hoc Networks," to appear, *Ad Hoc Networks*.
[15] F. H. P. Fitzek and M. Reisslein, "MPEG-4 and H.263 Video Traces for Network Performance Evaluation," *IEEE Network*, vol. 15, no. 6, Dec. 2001, pp. 40-54.

Biographies

FOTINI-NIOVI PAVLIDOU received a Ph.D. degree in electrical engineering from Aristotle University of Thessaloniki, Greece, in 1988 and a diploma in mechanical-electrical engineering in 1979 from the same institution. She is currently a full professor at the Department of Electrical and Computer Engineering at Aristotle University, teaching in the undergraduate and post-graduate program in the areas of mobile communications and telecommunications networks. Her research interests are in the field of mobile and personal communications, satellite and HAP communications, multiple access systems, routing and traffic flow in networks, and QoS studies for multimedia applications over the Internet. She is Involved in many national and international projects in these areas, and chaired the European COST262 Action on Spread Spectrum Techniques. She has served as a member of the Technical Program Committees of many IEEE/IEE conferences. She is a permanent reviewer for many international journals. She has published about 80 papers in refereed journals and conferences. She is currently chairing the joint IEEE VTS & AESS Chapter in Greece.

ORESTIS TSIGKAS (torestis@auth.gr) received his Diploma in electrical and computer engineering from Aristotle University in 2002. He is currently working toward his Ph.D. degree in the same department. His research interests include medium access as well as quality of service provisioning.