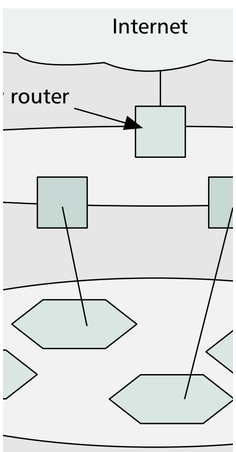
# A SURVEY OF QUALITY OF SERVICE IN IEEE 802.11 NETWORKS

HUA ZHU, MING LI, IMRICH CHLAMTAC, AND B. PRABHAKARAN
THE UNIVERSITY OF TEXAS AT DALLAS




IEEE 802.11 has gained popularity at an unprecedented rate. However, due to the lack of built-in quality of service support, IEEE 802.11 experiences serious challenges in meeting the demands of multimedia services and applications.

*This work was supported in part by the NSF grant IIS-0237954 on CAREER: Animation Databases.*

## ABSTRACT

Developed as a simple and cost-effective wireless technology for best effort services, IEEE 802.11 has gained popularity at an unprecedented rate. However, due to the lack of built-in quality of service support, IEEE 802.11 experiences serious challenges in meeting the demands of multimedia services and applications. This article surveys 802.11 QoS schemes, including service differentiation in the MAC layer, admission control and bandwidth reservation in MAC and higher layers, and link adaptation in the physical layer, designed to meet these challenges by providing the necessary enhancements for the required QoS. Furthermore, the article addresses issues that arise when end-to-end QoS has to be guaranteed in today's pervasive heterogeneous wired-cum-wireless networks. Among these challenges, protocol interoperability, multihop scheduling, full mobility support, and seamless vertical handoff among multiple mobile/wireless interfaces are specifically addressed.

## INTRODUCTION

Compared with a wired infrastructure, wireless LAN (WLAN) has unique advantages, such as broadband bandwidth capability and low deployment cost. Thanks to the technology provided by IEEE 802.11, the WLAN market is experiencing explosive growth in hot spots such as hotels, hospitals, and campuses, to mention just a few. With WLANs being deployed in an unlimited way as access points, wireless users can access real-time and Internet services virtually anytime, anywhere, while enjoying the flexibility of mobility and guaranteed connectivity.

IEEE 802.11 is designed for best effort services only. The lack of a built-in mechanism for support of real-time services makes it very difficult to provide quality of service (QoS) guarantees for throughput-sensitive and delay-sensitive multimedia applications. Therefore, modification of existing 802.11 standards is necessary. Although IEEE 802.11e is being proposed as the upcoming standard for the enhancement of service differentiation, QoS guarantee in 802.11 is still a very challenging problem and needs further study [1].

This article explores various 802.11 QoS schemes such as service differentiation in the medium access control (MAC) layer, admission control and bandwidth reservation in MAC as well as higher layers, and link adaptation in the physical layer. These schemes significantly enrich the enhancement of QoS from different aspects. Furthermore, many issues arise in smoothly providing end-to-end QoS guarantees while accessing wired networks from wireless and vice versa. Among these challenges, protocol interoperability, multihop scheduling, full mobility support, and seamless vertical handoff among multiple mobile/wireless interfaces are discussed.

This article is organized as follows. We provide an overview of the IEEE 802.11 standard. We discuss the existing QoS schemes for 802.11 networks. We introduce design challenges and future work. Finally, we conclude this article.

## AN OVERVIEW OF IEEE 802.11

IEEE 802.11 is the leading standard for wireless LAN [2]. It adopts the standard 802 logical link control (LLC) protocol but provides optimized physical layer (PHY) and medium access control (MAC) sublayers for wireless communications. 802.11 specifies two physical layers: direct sequence spread spectrum (DSSS) and frequency hopping spread spectrum (FHSS). Based on the transmission technologies and operating spectrum, the later revisions of 802.11 can be classified into three categories: 802.11a (orthogonal frequency-division multiplexing, OFDM, 5 GHz), 802.11b (high-rateDSSS, HR/DSSS, 2.4 GHz), and 802.11g (OFDM, 2.4 GHz). 802.11b is based on HR/DSSS and operates at the 2.4 GHz industrial, scientific, and medical (ISM) band with transmission rate from 1 to 11 Mb/s. 802.11a is based on OFDM and uses 5 GHz unlicensed national information infrastructure (U-NII) band in America with a transmission rate of 6–54 Mb/s. 802.11g is also based on OFDM but uses the 2.4 GHz ISM band and was formally ratified by the IEEE Standards Association's Standard Board in June 2003. This standard specifies a maximum transmission rate of 54 Mb/s, the same as 802.11a. However, since 802.11g uses the same spectrum between 2.4 and 2.4835 GHz and is inherently backward compati-

| Task Group | Responsibility |
|---|---|
| 802.11a — OFDM in 5 GHz Band | Specification enabling up to 54 Mb/s to be achieved in the 5 GHz unlicensed radio band by utilizing OFDM. |
| 802.11b — HR/DSSS in 2.4 GHz Band | Specification enabling up to 22 Mb/s to be achieved in the 2.4 GHz unlicensed radio band by utilizing HR/DSSS. |
| 802.11c — Bridge Operation Procedures | Provides required information to ensure proper bridge operations, which is required when developing access points. |
| 802.11d — Global Harmonization | Covers additional regulatory domains, which is especially important for operation in the 5 GHz bands because the use of these frequencies differ widely from one country to another. As with 802.11c, the 802.11d standard mostly applies to companies developing 802.11 products. |
| 802.11e — MAC Enhancements for QoS | Covers issues of MAC enhancements for quality of service, such as EDCF service differentiation and hybrid coordination function (HCF). |
| 802.11f — Inter Access Point Protocol (IAPP) | Provides interoperability for users roaming from one access point to another of different vendor. |
| 802.11g —OFDM in 2.4 GHz band | Specification enabling high data rates (36 or 54 Mb/s) to be achieved in the 2.4 GHz unlicensed radio band. |
| 802.11h — Dynamic Frequency Selection (DFS) | Dynamic channel selection and transmission power control. |
| 802.11i – Security | Specification for WLAN security to replace the weak Wired Equivalent Privacy (WEP). |

**Table 1.** *The family of IEEE 802.11 standards (OFDM: orthogonal frequency-division multiplexing).*

ble with 802.11b, it may attract more attention from industry than the earlier standardized 802.11a. Nevertheless, 802.11a possesses one noteworthy advantage: the unlicensed radio spectrum (5.15–5.35 and 5.725–5.825 GHz) it operates within is rarely used, while the 2.4 GHz spectrum for 802.11b and g has already been taken by many home electronic devices such as cordless phones, microwave ovens, and garage door openers. The family of IEEE 802.11 standards is shown in Table 1.

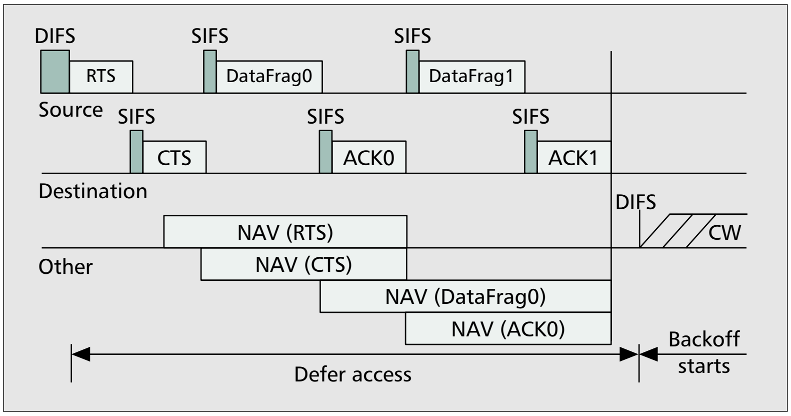
The 802.11 MAC supports two basic medium access protocols: contention-based distributed coordination function (DCF) and optional point coordination function (PCF). When PCF is enabled, the wireless channel is divided into superframes. Each superframe consists of a contention-free period (CFP) for PCF and a contention period (CP) for DCF. At the beginning of CFP, the point coordinator (usually the access point, AP) contends for access to the wireless channel. Once it acquires the channel, it cyclically polls high-priority stations and grants them the privilege of transmitting. Although the optional PCF is designed for delay-bounded services, it is centralized and can only be used in the network of infrastructure mode. In addition, the loose specification of PCF leaves many issues unsolved [3]:

- PCF experiences substantial delay at low load; stations must always wait for polling, even in an otherwise idle system;
- Since the AP needs to contend for the channel using DCF at the beginning of a CFP, the effective period of contention-free polling may vary.
- It is very difficult for the point coordinator to manage the polling of a large number of interactive streams without harming the applications using DCF contention.

In addition, PCF is a centralized approach that suffers from location-dependent errors. Therefore, PCF has not drawn much attention from either the research community or industry, and most existing schemes focus on the enhancement of DCF, which is a fully distributed protocol.

DCF is based on carrier sense multiple access with collision avoidance (CSMA/CA) instead of CSMA with collision detection (CSMA/CD) because stations cannot listen to the channel for collision while transmitting. In IEEE 802.11, carrier sensing (CS) is performed at both PHY and MAC layers: physical CS and MAC layer virtual CS. If the MAC frame length (including the payload and 34 bytes of MAC header) exceeds the *RTS_threshold*, request-to-send (RTS) and clear-to-send (CTS) are used by stations to solve the hidden terminal and capture effect problems. A MAC protocol data unit (MPDU) contains header information, payload, and a 32-bit cyclic redundancy check (CRC). The duration field indicates the amount of time after the end of the present frame the channel will be used to complete successful transmission of the data or management frame. Stations use the information in the duration field to adjust their network allocation vector (NAV), which indicates the amount of time that must elapse until the current transmission session is complete and the channel can be sensed again for idle status.

Before a station sends out a data frame, it senses the channel. If the channel is idle for at least a DCF interframe space (DIFS), the frame is transmitted. Otherwise, a backoff time slot is chosen randomly in the interval $[0, CW)$. The contention window (CW) is incremented exponentially with an increasing number of attempts to retransmit the frame. Upon receipt of a correct packet, the receiving stations waits a short interframe space (SIFS) interval and transmits a positive acknowledgment frame (ACK) back to




■ **Figure 1.** *The CSMA/CA-RTS/CT with fragmentation access scheme.*

the source station, indicating transmission success. During the backoff period, the backoff timer is decremented in terms of slot time as long as the channel is determined to be idle. When the backoff timer reaches zero, the data frame is sent out. If collision occurs, a new backoff time slot is chosen and the backoff procedure starts over until some time limit is exceeded. After successful transmission, the CW is reset to $CW_{min}$. For further increase of wireless channel utilization, payload length is divided into fragments of smaller size (if it exceeds the *Frag_threshold*) before a packet is transmitted within one CW. The advantage of this technique is that if an error occurs during its transmission of a specific fragment, a station does not have to wait to back off until the whole payload is transmitted. Also, it does not have to retransmit previous fragments that have been transmitted successfully. The timing diagram of CSMA/CA-RTS/CTS with fragmentation access is shown in Fig. 1. The range of *RTS_threshold* is 0–2347 bytes (default), while the range of *Frag_threshold* is 256–2312 bytes (default). However, vendors may choose different ranges for both thresholds.

Once an error occurs, a packet has to be retransmitted by the attempting station. Errors may be caused by many possible situations. For example, the corresponding CTS frame may not be returned after an RTS frame is transmitted. This may occur due to:
- Collision with the transmission of another station
- Interference in the channel during the transmission of other RTS/CTS frames
- The station receiving the RTS frame having an active virtual CS condition (indicating a busy medium time period)

Two retry counters, the *short retry count* and *long retry count*, are defined for use in packet retransmission. Packets shorter than *RTS_threshold* are associated with the short retry count; others are associated with the long retry count. The retry counters begin at 0 and are incremented whenever a frame (or fragment) transmission fails. A frame is dropped if the retry count exceeds the maximum retry limit. The short count is reset to 0 when:
- A CTS is received in response to a transmitted RTS.
- An ACK is received after a non-fragmented transmission.
- A broadcast or multicast frame is received.

The long retry count is reset to 0 when:
- An ACK is received for a frame longer than RTS threshold.
- A broadcast or multicast frame is received.

In order to optimize the performance of DCF, a number of parameters [4] are tunable in both the PHY and MAC layers of 802.11. A few are selected and shown in Table 2. However, these parameters are basically station-based and therefore cannot effectively differentiate multiple flows within a station. Furthermore, the effects of tuning these parameters are limited in terms of increasing/decreasing MAC throughput/delay, respectively. Therefore, additional resolutions are demanded to guarantee QoS in 802.11.

## QoS Mechanisms

Most existing QoS mechanisms for 802.11 can be classified into three categories:
- Service differentiation
- Admission control and bandwidth reservation
- Link adaptation

While discussing different QoS schemes, we do not consider those based on higher layers since 802.11 specifies access technologies at the MAC and PHY layers.

| Parameter | Meaning and units | Tuning effect | |
|---|---|---|---|
| | | If increased | If decreased |
| Beacon interval | Number of Tus between transmission of beacon frames | Better throughput and longer battery life | Mobile stations can move faster and still maintain the network connectivity |
| RTS threshold | Frames longer than the threshold use RTS/CTS access method | Increasing the maximum theoretical throughput if no hidden terminal or interference | Higher throughput if there are a large number of hidden terminals |
| Fragmentation threshold | Frames longer than the threshold are fragmented | Increasing throughout in error-free channel | Increasing throughput in error-prone channel |
| Long/short retry limits | The maximum number of retransmission allowed for frames shorter/longer than RTS threshold | Lower frames drop rate, but it may incur longer backoff and throttle the throughput | Higher frames drop rate, but smaller buffer required |

■ **Table 2.** *Common tunable parameters in 802.11.*

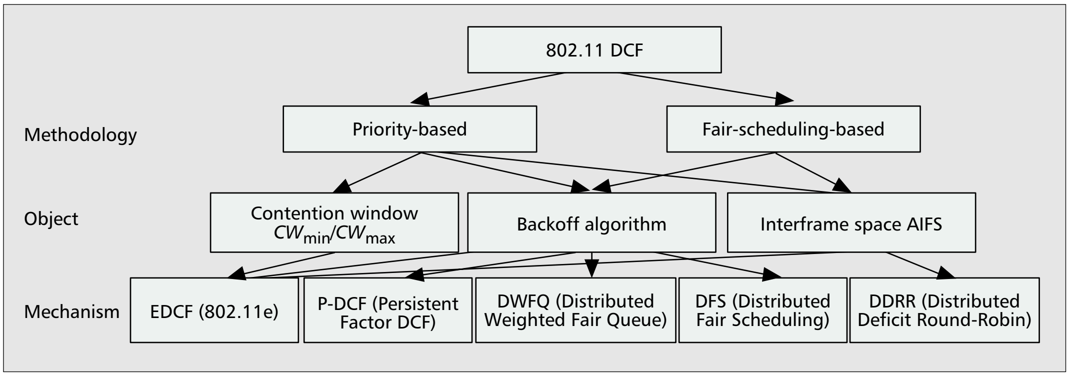



**Figure 2.** *IEEE 802.11 service differentiation mechanisms.*

Research has been done to provide certain DCF-based QoS enhancements, which mainly address the effective support of service differentiation. In fact, these mechanisms do not provide any QoS guarantee, but only better than best-effort services.

## Better than Best Effort Schemes: Service Differentiation (MAC)

Research has been done to provide certain DCF-based QoS enhancements that mainly address effective support of service differentiation. In fact, these mechanisms do not provide any QoS guarantee, only better than best effort services. Figure 2 shows a hierarchical taxonomy of QoS mechanisms that enable service differentiation in 802.11. Basically, service differentiation is achieved by two main methods: priority and fair scheduling [5]. While the former binds channel access to different traffic classes by prioritized contention parameters, the latter partitions the channel bandwidth fairly by regulating wait times of traffic classes in proportion according to given weights. The tunable parameters (or object in Fig. 2) for both approaches are CW size, backoff algorithm, and interframe space.

The specific service differentiation mechanisms are:

**Enhanced DCF (EDCF) [6]**: EDCF is a main part of the upcoming 802.11e standard for service differentiation. It prioritizes traffic categories by different contention parameters, including arbitrary interframe space (AIFS), maximum and minimum backoff window size ($CW_{max/min}$), and a multiplication factor for expanding the backoff window. Although all traffic categories keep using the same DCF access method, they have different probabilities of winning the channel contention by differentiating contention parameters. EDCF allows any combinations of these contention parameters according to the service provider's needs.

**Persistent Factor DCF (P-DCF) [7]**: In this algorithm, each traffic class is associated with a persistent factor $P$ (high-priority classes have smaller $P$). In a backoff stage, a uniformly distributed random number $r$ is generated in every slot time. Each flow stops the backoff and starts transmission only if ($r > P$) in the current slot time, given no transmission occurs in previous slot times. Therefore, the backoff interval is a geometrically distributed random variable with parameter $P$.

**Distributed Weighted Fair Queue (DWFQ) [8, 9]**: Two different algorithms using this strategy have been proposed. In the first, the backoff window size $CW$ of any traffic flow is adjusted based on the difference between the actual and expected throughputs. If the actual throughput is lower than the expected throughput, $CW$ will be decreased in order to increase the flow's priority, and vice versa. In the second algorithm, a ratio ($L_i' = R_i/W_i$) is calculated, where $R_i$ is the actual throughput and $W_i$ the corresponding weight of the $i$th station. By comparing its own $L_i$ with those of others, a station can adjust its $CW$; for example, if its $L_i$ is smaller than those of others, it will decrease its $CW$.

**Distributed Fair Scheduling (DFS) [8, 10]**: The main idea of DFS is to differentiate the backoff interval (BI) based on the packet length and traffic class, and the station with smaller BI transmits first. For the $i$th flow, $BI_i = \rho_i \times scaling \times factor \times L_i/\varphi_i$, where $BI_i$ is the backoff interval, $L_i$ is the packet length, $\varphi_i$ is the weight (the weights of high throughput classes are larger than that of low classes), and $\rho_i$ is a random variable uniformly distributed in [0.9,1.1]. $\rho_i$ is introduced to minimize the collision caused by multiple stations with the same backoff interval.

**Distributed Deficit Round Robin (DDRR) [11]**: In this algorithm, the $i$th throughput class at the $j$th station is assigned with a service *quantum rate* ($Q_{i,j}$) equal to the throughput it requires, and a deficit counter ($DC_{i,j}$) that accumulates at the rate of $Q_{i,j}$ and is decreased by the packet length whenever a packet is transmitted. $DC_{i,j}$ is used to calculate the interframe space ($IFS_{i,j}$), which is the wait time before transmission or backoff starts, depending on whether backoff is used. A larger $DC_{i,j}$ results in a smaller $IFS_{i,j}$. In order to minimize the collision between stations with the same deficit counter, randomization of $IFS_{i,j}$ will be further adopted if a backoff scheme is eliminated.

Potentially, fair-scheduling-based schemes are advantageous to fairly allocate bandwidth among traffic classes and prevent starvation of a specific class. However, they often require substantial modification of existing 802.11 standards. Compared to fair-scheduling-based schemes, priority-based mechanisms require less modifications of the existing DCF access method and provide better QoS support for real-time applications. With the imperative demand to transmit video/voice data over WLANs, IEEE 802.11e has gained the interest of the majority of the community.

The AACA protocol was designed to effectively solve the hidden terminal and exposed terminal problems in multi-hop networks. Nevertheless, it can also be used to achieve bandwidth reservation.

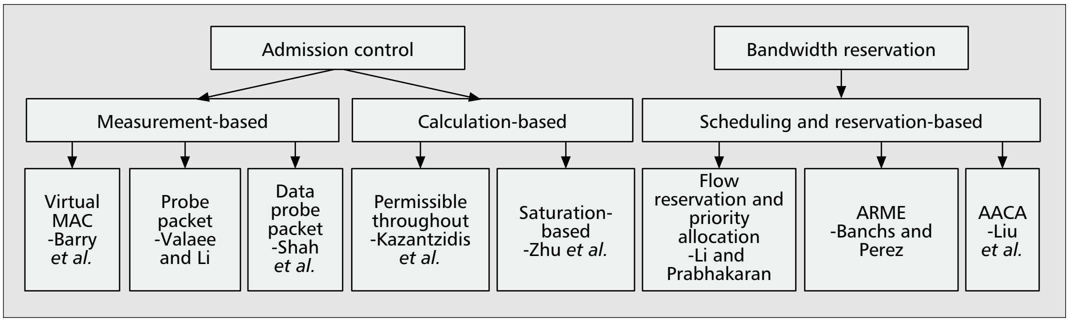



**Figure 3.** *Admission control and bandwidth reservation mechanisms.*

## QoS Mechanisms for Admission Control and Bandwidth Reservation (MAC)

Service differentiation is helpful in providing better QoS for multimedia data traffic under low to medium traffic load conditions. However, due to the inefficiency of IEEE 802.11 MAC, service differentiation does not perform well under high traffic load conditions [3]. In this case, admission control and bandwidth reservation become necessary in order to guarantee QoS of existing traffic. Otherwise, the extremely large saturation delay may lead to failure to support multimedia applications. However, unlike wired networks, in IEEE 802.11 wireless networks, a wireless node has no knowledge of the exact condition of the network and thus cannot make an accurate decision on whether or not to admit a new flow. In addition, with the contention-based CSMA/CA channel access mechanism, bandwidth provisioning is almost impossible, leading to only soft QoS guarantee. Because of these two major differences, admission control and bandwidth reservation in an IEEE 802.11 wireless network is quite difficult. Figure 3 shows a hierarchical taxonomy of QoS mechanisms that enable admission control and bandwidth reservation in 802.11. In general, admission control schemes require less modification to the 802.11 standards than bandwidth reservation schemes. Basically, admission control schemes can be broadly classified into measurement-based and calculation-based methods. In measurement-based schemes, admission control decisions are made based on the measurements of existing network status, such as throughput and delay. On the other hand, calculation-based schemes construct certain performance metrics or criteria for evaluating the status of the network [12].

**Measurement-based approaches**: *Virtual MAC* [13]: Barry *et al.* propose a virtual MAC algorithm that passively monitors the channel by virtual MAC frames and estimates local service level (i.e., throughput and delay) by the measurement of virtual frames. In addition, a virtual source (VS) algorithm allows application parameters to be tuned according to dynamic channel conditions by utilizing virtual MAC. Valaee and Li [14] proposed a measurement-based admission procedure using a sequence of probe packets for ad hoc networks. Instead of using probe packets, Shah *et al.* [15] proposed a measurement-based admission control scheme using data packets to measure the network load.

**Calculation-based approaches**: Kazantzidis *et al.* [16] present a heuristic solution using a simple parameter, *permissible throughput*, as the admission decision criterion. Implementation issues in multihop ad hoc networks, such as propagation of permissible throughput using the Ad Hoc On Demand Distance Vector (AODV) routing protocol and discovering the bottleneck along the path, have been investigated. Based on the proposed saturation coefficient, Zhu *et al.* [10] offer a calculation-based distributed admission control protocol. Here, stations can effectively determine whether they are approaching saturation condition based on piggybacked information, including the number of active stations, their corresponding traffic bit rates, and average packet lengths. In this protocol the admission control decisions are made dynamically at both source and destination stations in a fully distributed way. It can predict saturation and prevent it from happening.

**Scheduling and reservation-based approaches**: Li and Prabhakaran [1] introduce a novel framework for admission control with priority reservation and allocation, which is mainly focused on optimizing the usage of priority resources. Banchs and Pérez [8] present ARME as an extension of DCF. ARME uses a token-bucket-based algorithm to detect whether the network is in overloading condition, and improve the performance of the system by adjusting the CW appropriately. Liu *et al.* [17] define a class of reservation-based MAC access protocols with multichannel supports, including AACA-SDT, ACA-MDT, and AACA-RDT. The AACA adopted the RTS/CTS access method on a common channel solely for reservation purposes. After successful reservation, the station pair transmits packets uninterruptedly in the reserved channel. The AACA protocol was designed to effectively solve the hidden terminal and exposed terminal problems in multihop networks. Nevertheless, it can also be used to achieve bandwidth reservation.

## QoS Mechanism for Link Adaptation (PHY)

802.11 specifies multiple transmission rates that are achieved by different modulation techniques in the PLCP header of the PHY layer. However, it intentionally leaves the rate adaptation and signaling mechanisms open. Since transmission rates differ with the channel conditions, an appropriate link adaptation mechanism is desirable to maximize the throughput under dynamically changing channel conditions. Figure 4 shows a hierarchical taxonomy of QoS mecha-

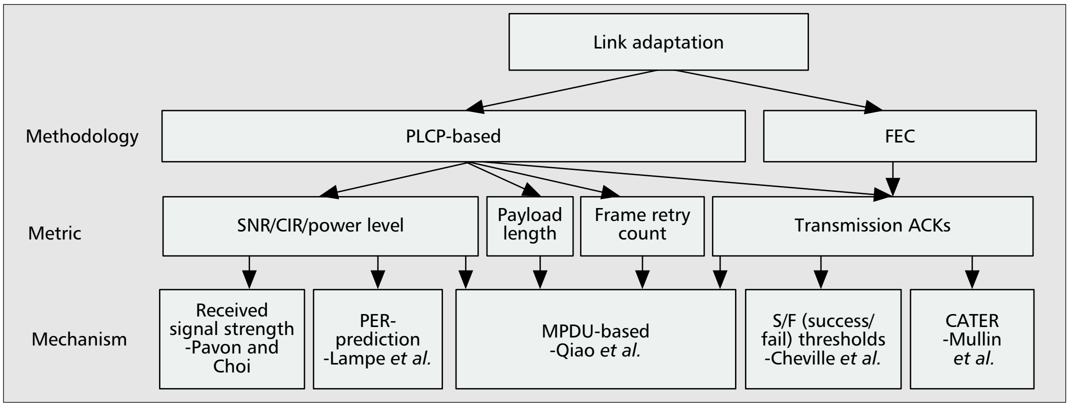



**Figure 4.** *Link adaptation mechanisms.*

With more and more real-time multimedia applications subscribed by mobile users, there is an urgent demand that end-to-end QoS guarantee should be provided in wired-cum-wireless heterogeneous networks.

nisms that enable link adaptation. Most link adaptation mechanisms focus on algorithms to switch among transmission rates specified in PLCP, without the need to modify existing standards. However, there is a novel idea to adjust the length of DSSS pseudo-noise (PN) code in 802.11b, with slight modifications of 802.11b DCF. Metrics used in existing link adaptation algorithms include channel signal-to-noise ratio/carrier-to-interference ratio (SNR/CIR), received power level, average payload length, transmission acknowledgments, or combinations.

**Received signal strength (RSS) [18]**: Pavon and Choi choose the RSS as the metric for the adaptation algorithm with the assumption that transmission power is fixed. The authors also assume a linear relationship between the average RSS and SNR. A rate adaptation algorithm at every station maintains its own 12 RSS thresholds and corresponding transmission rates. Based on the measured RSS, the station dynamically switches to an appropriate transmission rate.

**PER-prediction [19]**: Lampe *et al.* propose a link adaptation scheme in which decisions are made based on packet error rate (PER) prediction. A main contribution of this article is the prediction of PER with not only SNR/CIR but also the momentary channel transfer function.

**MPDU-based link adaptation [20]**: Observing the effects of multiple factors on the throughput, Qiao *et al.* proposed to use a combination of SNR, average payload length, and frame retry count as the metric for the link adaptation algorithm. The proposed algorithm pre-established a table of best transmission rate for decision making.

**Link adaptation with success/fail (S/F) thresholds [21]**: Chevillat *et al.* use the ACKs of transmitted frames as a measurement of channel condition. The basic idea of this algorithm is: if the number of consecutive successful transmissions exceeds *S*, the transmission rate is increased. On the other hand, if the number of consecutive failed transmission exceeds *F*, the transmission rate is decreased. The ACKs of frames are used to indicate whether a transmission succeeded or failed.

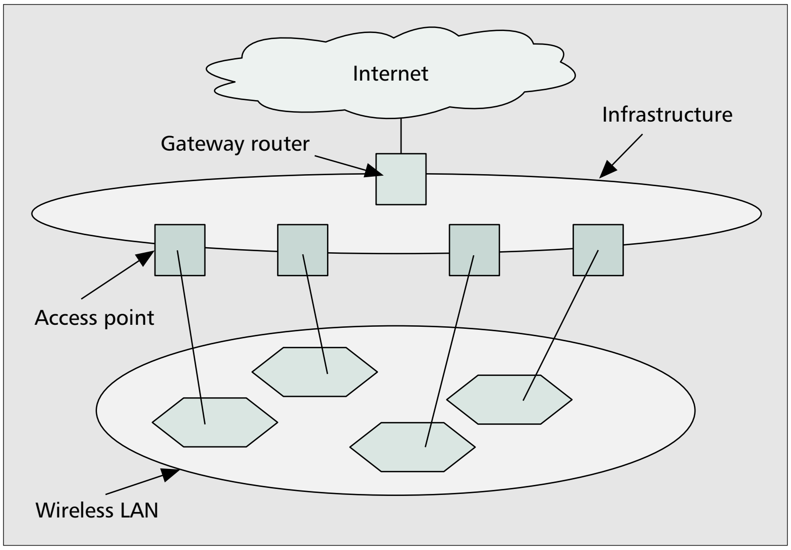
**Code Adapts To Enhance Reliability (CATER) [22]**: Mullins *et al.* describe an adaptive PN code algorithm for direct sequence spread spectrum (DSSS) used in 802.11b. This algorithm improved the throughput under high bit error rate (BER) channel conditions. However, due to the additional signaling overhead of the proposed scheme, the throughput under low BER channel conditions is lower than the standard 802.11b.

## Challenges and Future Work

### Wireless Internet and Interoperability

IEEE 802.11 WLANs have been successfully applied as the last mile technology in the increasingly pervasive computing environments where wireless/mobile users accesses Internet services via the access point (AP). With more and more real-time multimedia applications subscribed to by mobile users, there is an urgent demand for end-to-end QoS guarantee to be provided in wired-cum-wireless heterogeneous networks. The architecture of a heterogeneous wired-cum-wireless network is depicted in Fig. 5 (revised from the MIRAI architecture [23]). Currently, there are some noteworthy approaches that focus on the interoperability between IEEE 802.11 and differentiated services (DiffServ) or integrated services (IntServ). These approaches explore the possibility of protocol interoperability at the AP. Since MAC protocols are prevalent in IEEE 802.11 WLAN, cross-layer interaction is necessary for both network and MAC layers to share the flow QoS characteristics and network topology information without duplicative efforts.

**802.11e and DiffServ**: Park and Kim [24] propose a collaborative end-to-end QoS architecture across wired WAN, wired LAN, and WLAN, based on DiffServ, IEEE 802.1D/Q, and IEEE 802.11e, respectively. The 802.3 MAC frames carry the user priority values via the 802.1Q virtual LAN (VLAN) tag. These user priorities are forwarded through 802.1D MAC bridge to 802.1e MAC and used by EDCF to differentiate flows. To realize this, it is necessary to establish mapping between DiffServ code point (DSCP) values, defined in the DiffServ (DS) field, and traffic category identification (TCID), defined in 802.1e. Two kinds of mapping have been defined in [24]: direct and hierarchical. In direct mapping, when the IP packets are encapsulated in MAC frames, they are placed in priority queues without preemption. In hierarchical mapping, IP packets are classified and shaped according to the priority of the DSCP values




■ **Figure 5.** *Architecture of the wireless Internet.*

before being forwarded to 802.1e priority queues. Essentially, hierarchical mapping enables more accurate end-to-end QoS traffic control required by user applications.

**802.11 and IntServ**: Based on a proposed MAC layer flow reservation and admission control protocol in IEEE 802.11 WLAN, called WRESV, Li and Zhu [25] suggest integrating RSVP and WRESV for the support of IntServ in heterogeneous wired-cum-wireless networks. In their approach, features of Resource Reservation Protocol (RSVP) and the characteristics of the wireless medium are carefully considered (e.g., multicast receivers in a WLAN can be handled by broadcasting at the AP). Message mappings at the AP are implemented by cross-layer interaction, and user priorities are mapped to 802.11 MAC priorities with 802.1p. Since WRESV can work with most of the existing MAC schedulers such as DCF, EDCF, and DFS, this integration scheme is more general and leaves space for further enhancement. One interesting feature is that this integration scheme also considers support of both node mobility and QoS in the situation of handoff.

## Support of Full Mobility

802.11 supports mobile stations (MSs) within an 802.11 extended service set (ESS) to roam among multiple APs. This roaming capability is achieved through MSs' beacon scanning in a channel sweep. When an MS enters a new basic service area, it first scans across all channels, remaining on each channel for a specified period of time to detect the signal radiated from the AP, and then acquires the channels from the AP.

Currently, 802.11 WLAN service is only available for low-mobility devices in isolated hot spots with coverage from dozens of meters up to a few hundred meters. However, recent efforts have been made to extend 802.11 WLANs into outdoor cellular networks to provide fully mobile broadband service with ubiquitous coverage and high-speed connectivity. Leung *et al.* [26] investigate the throughput of 802.11 MAC protocol with the service area (cell) size of 6 km and claim that without modifying the standard, the DCF access method with RTS/CTS is feasible for large outdoor cellular coverage. To extend the coverage of 802.11, smart antenna design at the radio link is critical. By transmitting packets in a beam instead of in all directions, Vivato Inc. extends the coverage of 802.11 to kilometers, while other antennae (e.g., Cisco's Aironet, Motorola Inc.'s Canopy Radio, and Proxim Inc.'s Tsunami QuickBridge) may reach up to 10 km. [27]. However, a cell with large outdoor coverage does not necessarily guarantee high-speed connectivity. Due to the unavoidable channel contention, throughput may degrade when a cell is overcrowded. Therefore, it is more desirable to have overlapped small cells with fast handoff mechanisms. So far, a few companies (e.g., MeshNetworks) have accomplished highway-speed connectivity through 802.11 compatible protocols. In their implementation, data packets from the Internet are transmitted through APs hanging on rooftops or streetlights.

## QoS and Mobility Management in Hybrid Wireless Networks

In addition to roaming and horizontal handoff among 802.11 WLANs, supporting QoS anytime, anywhere, and by any media requires seamless vertical handoffs between different wireless networks such as WLAN, mobile ad hoc network (MANET) [28], Bluetooth, Universal Mobile Telecommunications System (UMTS), and wideband code-division multiple access (WCDMA). Many new architectures/schemes have been proposed recently for seamless integration of WLAN and various wireless network interfaces, discussed below.

**Integration of WLAN and MANET**: Lamont and Wang [29] investigated the issue of maintaining session connectivity while mobiles continuously roam across multiple WLANs and MANETs. In the proposed network architecture, routing within MANETs is handled by the Optimized Link State Routing (OLSR) protocol, and handoff between WLANs and MANETs is supported through automatic mode detection and node switching capabilities of the mobiles. To achieve efficient mobility management, functionalities of OLSR are extended to support Mobile IPv6.

**Integration of WLAN and Bluetooth**: Conti and Dardari [30] proposed an integrated analytical model for evaluation of the interference between IEEE 802.11 and Bluetooth. The model takes both PHY and MAC layers into account and can easily be implemented. The performance is evaluated by packet error probability in terms of the relative distances between the two systems for different conditions.

**Integration of WLAN and 3G wireless networks**: An architecture for integrating UMTS and IEEE 802.11 WLANs was proposed by Jaseemuddin [31]. Since 802.11 is used primarily for high-speed best effort service, a mobile node can maintain two connections in parallel (i.e., data connection through WLAN and voice connection through UMTS). Park and Yoon [32] investigate vertical handoff between IEEE 802.11 WLANs and CDMA cellular networks. In their handoff strategy, traffic characteristics are considered in order to guarantee low handoff latency. Specifically, real-time traffic takes into account the handoff delay, and best effort traffic takes only throughput into account . Finally, Buddhikot and Chandranmenon [33] suggest combining the features of wide-coverage low-rate 3G networks and high-rate small-coverage WLANs to improve the QoS and flexibility of wireless data services. A loose integration approach is realized with an IOTA gateway and a new client software in order to support seamless mobility, QoS guarantees, and multiprovider roaming agreements.

With the decreasing size of cells in next-generation multimedia-enabled wireless networks, the number of handoffs during a call's lifetime increases. Thus, for integration of WLAN and 3/4-G wireless networks, an essential element of seamless end-to-end QoS guarantee is ensuring low call dropping probability in the 3/4-G networks. Lou and Li [34] propose an adaptive bandwidth allocation scheme, termed *measurement-based preassignment*, to prevent handoff failure in wireless cellular networks. With periodic measurement of traffic status within a local cell, the number of channels reserved for a handoff can be adjusted, thus eliminating the signaling overhead of status information exchange between involved cells.

## Summary

Research in providing QoS guarantee in IEEE 802.11 WLANs and/or MANETs has been extensively studied in recent years, and many QoS schemes have been proposed. We classify these approaches into the following three categories:
- Link adaptation in the physical layer
- Channel access coordination in the MAC layer
- Admission control strategies in MAC and higher layers

Essentially, these approaches focus on different network layers and are tightly interrelated. Without admission control and resource allocation, providing QoS guarantees only by differentiating flows and coordinating the order of channel access cannot be effective under high traffic loads. Also, all MAC and higher-layer protocols are significantly affected by link adaptation mechanisms in the physical layer. Therefore, the current trend is to design frameworks that can share flow characteristics across multiple layers and cooperate to meet applications' QoS requirements.

Besides the aforementioned QoS mechanisms, this article also discussed challenges encountered in designing wireless Internet, multihop ad hoc networks, and heterogeneous wireless networks. We expect interesting work in the following areas:
- Interoperability between QoS mechanisms in 802.11 WLAN and existing QoS architectures for IP-based Internet, such as IntServ and DiffServ
- QoS guarantee in a self-organized multihop MANET, where link condition and resource are highly dynamic
- Support of full mobility in 802.11, which may require the extension of coverage of 802.11 by modifying the design of the smart antenna at the radio link
- QoS and mobility management in hybrid mobile/wireless networks where WLAN coexists with other networks such as MANET, Bluetooth, and 3G networks

These topics and issues combined extend the concepts of QoS guarantee in 802.11 and make it better suited for today's high-speed high-mobility wireless networks.

So far, a few companies (e.g., MeshNetworks) have accomplished highway-speed connectivity through 802.11-compatible protocols. In their implementation, data packets from the Internet are transmitted through access points hanging on rooftops or streetlights.

## References


[1] M. Li, B. Prabhakaran, and S. Sathyamurthy, "On Flow Reservation and Admission Control for Distributed Scheduling Strategies in IEEE802.11 Wireless LAN," *Proc. 6th ACM Int'l. Wksp. Modeling, Anal. and Sim. Wireless and Mobile Sys.*, Sept. 2003.

[2] IEEE 802.11, 1999 ed. (ISO/IEC 8802-11: 1999), "IEEE Standards for Information Technology — Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Network — Specific Requirements — Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications."

[3] A. Lindgren, A. Almquist, and O. Schelén, "Evaluation of Quality of Service Schemes for IEEE 802.11 Wireless LANs," *Proc. 26th Annual IEEE Conf. Local Comp. Nets.*, Tampa, FL, Nov. 2001.

[4] M. Gast, *802.11 Wireless Networks: the Definitive Guide*, O'Reilly, 2002.

[5] W. Pattara-Atikom, P. Krishnamurthy, and S. Banerjee, "Distributed Mechanisms for Quality of Service in Wireless LANs," *IEEE Wireless Commun.*, vol. 10, June 2003, pp. 26–34.

[6] IEEE Std 802.11e-D3.3, "Draft Supplement to IEEE Standard for Telecommunications and Information Exchange between Systems — LANMAN Specific Requirements,. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY)," Oct. 2002.

[7] Y. Ge and J. Hou, "An Analytical Model for Service Differentiation in IEEE 802.11," *IEEE ICC '03*, vol. 2, May 2003, pp. 1157–62.

[8] A. Banchs and X. Pérez, "Providing Throughput Guarantees in IEEE 802.11 Wireless LAN," *IEEE WCNC '02*, vol. 1, 2002, pp. 130–38.

[9] A. Banchs and X. Pérez, "Distributed Weighted Fair Queuing in 802.11 Wireless LAN," *IEEE ICC '02*, vol. 5, Apr. 2002, pp. 3121–27.

[10] N. H. Vaidya, P. Bahl, and S. Gupta, "Distributed Fair Scheduling in a Wireless LAN," *Proc. ACM MOBICOM 2000*, Aug. 2000, pp. 167–78.

[11] W. Pattara-Atikom, S. Banerjee, and P. Krishnamurthy, "Starvation Prevention and Quality of Service in Wireless LANs," *Proc. 5th Int'l. Symp. Wireless Pers. Multimedia Commun.*, HI, Oct. 2002.

[12] H. Zhu *et al.*, "Distributed Admission Control for IEEE 802.11 Ad Hoc Networks," submitted, *IEEE/ACM Trans. Net.*

[13] M. Barry, A. T. Campell, and A. Veres, "Distributed Control Algorithms for Service Differentiation in Wireless Packet Networks," *Proc. IEEE INFOCOM 2001*.

[14] S. Valaee and B. Li, "Distributed Call Admission Control in Wireless Ad Hoc Networks," *Proc. IEEE VTC 2002*, Vancouver, BC, Canada, Sept. 24–28, 2002.

[15] S. H. Shah, K. Chen, and K. Nahrstedt, "Dynamic Bandwidth Management for Single-Hop Ad Hoc Wireless Networks," *Proc. IEEE Int'l. Conf. Perv. Comp. and Commun.*. 2003.

[16] M. Kazantzidis, M. Gerla, and S.-J. Lee, "Permissible Throughput Network Feedback for Adaptive Multimedia in AODV MANETs," *IEEE ICC '01*, vol. 5, June 2001, pp. 1352–56.

[17] K. Liu *et al.*, "A Reservation-based Multiple Access Protocol with Collision Avoidance for Wireless Multihop Ad Hoc Networks," *IEEE ICC '03*, vol. 2, May 2003, pp. 1119–23.

[18] J. D. P Pavon and S. Choi, "Link Adaptation Strategy for IEEE 802.11 WLAN via Received Signal Strength Measurement," *IEEE ICC '03*, vol. 2, May 2003, pp. 1108–13.

[19] M. Lampe, H. Rohling, and J. Eichinger, "PER-Prediction for Link Adaptation in OFDM Systems," *OFDM Wksp.*, Hamburg, Germany, 2002.

[20] D. Qiao, S. Choi, and K. G. Shin, "Goodput Analysis and Link Adaptation for IEEE 802.11a Wireless LANs," *IEEE Trans. Mobile Comp.*, vol. 1, no. 4, 2002, pp. 278–92.

All MAC and higher-layer protocols are significantly affected by link adaptation mechanisms in the physical layer. Therefore, the current trend is to design frameworks that can share flow characteristics across multiple layers and cooperate to meet the applications' QoS requirements.

[21] P. Chevillat *et al.*, "A Dynamic Link Adaptation Algorithm for IEEE 802.11a Wireless LANs," *IEEE ICC '03*, vol. 2, May 2003, pp. 1141–45.
[22] B. E. Mullins, N. J. Davis IV, and S. F. Midkiff, "An Adaptive Wireless Local Area Network Protocol that Improves Throughput via Adaptive Control of Direct Sequence Spectrum Parameters," *ACM SIGMOBILE Mobile Comp. and Commun. Rev.*, vol. 1, no. 3, Sept. 1997, pp. 9–20.
[23] P. J. M. Havinga and G. Wu, "Wireless Internet on Heterogeneous Networks," *Proc. Wksp. Mobile Commun. in Perspective*, Enschede, the Netherlands, Feb. 2001.
[24] S.-Y. Park *et al.*, "Collaborative QoS Architecture between DiffServ and 802.11e Wireless LAN," *Proc. IEEE VTC '03-Spring*, Jeju, Korea, Apr. 2003.
[25] M. Li *et al.*, "End-to-end QoS Guarantee in Heterogeneous Wired-cum-Wireless Networks," Tech. rep. no. UTDCS-25-04, Univ. of TX at Dallas, July 2004.
[26] K. K. Leung *et al.*, "Outdoor IEEE 802.11 Cellular Networks: MAC Protocol Design and Performance," *IEEE ICC '02*, vol. 1, Apr. 2002, pp. 595–99.
[27] S. M. Cherry, "Broadband a Go-Go," *IEEE Spectrum*, June 2003, pp. 20–25.
[28] I. Chlamtac, M. Conti, and J. Liu, "Mobile Ad Hoc Networking: Imperatives and Challenges," *Ad Hoc Networks J.*, vol. 1, no. 1, 2003.
[29] L. Lamont *et al.*, "Integrating WLANs & MANETs to the IPv6 Based Internet," *IEEE ICC '03*, vol. 2, May 2003, pp. 1090–95.
[30] A. Conti *et al.*, "Bluetooth and IEEE 802.11b Coexistence: Analytical Performance Evaluation in Fading Channels," *IEEE JSAC*, vol. 21, Feb. 2003, pp. 259–69.
[31] M. Jaseemuddin, "An Architecture for Integrating UMTS and 802.11 WLAN Networks," *IEEE ISCC '03*, June 2003, pp. 716–23.
[32] H. Park *et al.*, "Vertical Handoff Procedure and Algorithm between IEEE 802.11 WLAN and CDMA Cellular Network," *Mobile Commun.: 7th CDMA Int'l. Conf.*, Seoul, Korea, 2002.
[33] M. Buddhikot *et al.*, "Integration of 802.11 and Third-Generation Wireless Data Networks," *Proc. IEEE INFOCOM 2003*.
[34] X. Lou *et al.*, "An Adaptive Measured-based Pre-Assignment Scheme with Connection-level QoS Support for Mobile Networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 3, 2002, pp. 521–30.

## Biographies

HUA ZHU (zhuhua@utdallas.edu) received a B.S. in electrical engineering from Beijing Polytechnic University, China, in 1998. He received an M.S. in telecommunication engineering in University of Texas at Dallas in 2001. He is now working toward a Ph.D. degree in electrical engineering at the University of Texas at Dallas, where his research interests include QoS mechanisms for various wireless access technologies and wireless multimedia communications.

MING LI (mingli@utdallas.edu) received his B.S. and M.S. in engineering from Shanghai Jiao Tong University, China, in 1995 and 1998, respectively. He is currently a Ph.D. candidate in the Department of Computer Science, University of Texas at Dallas, where he received an M.S. degree in computer science in December 2001. His research interests include QoS schemes for mobile ad hoc networks and multimedia over wireless networks.

IMRICH CHLAMTAC [M'86, SM'86, F'93] (chlamtac@utdallas.edu) received a Ph.D. degree in computer science from the University of Minnesota, Minneapolis. Since 1997 he has been the Distinguished Chair in Telecommunications, University of Texas at Dallas. He also holds the titles of the Sackler Professor at Tel Aviv University, Israel, the Bruno Kessler Honorary Professor at the University of Trente, Italy, where he is currently on sabbatical, and University Professor at the Technical University of Budapest, Hungary. He has published close to 300 papers in referreed journals and conferences, and is the co-author of *Local Area Networks* (Lexington Books, 1981, 1982, 1984) and *Mobile and Wireless Networks Protocols and Services* (Wiley, 2000), and *IEEE Network*'s 2000 Editor's Choice. He serves as the founding Editor-in-Chief of *ACM/URSI/Kluwer Wireless Networks*, *ACM/Kluwer Mobile Networks and Applications*, and *SPIE/Kluwer Optical Networks*. He is a Fellow of the ACM, a Fulbright Scholar, and an IEEE Distinguished Lecturer. He is the winner of the 2001 ACM Sigmobile annual award and the IEEE ComSoc TCPC 2002 Award for contributions to wireless and mobile networks and of multiple best paper awards in wireless and optical networks.

B. PRABHAKARAN (praba@utdallas.edu) is on the faculty of the Computer Science Department, University of Texas at Dallas. He has been working in the area of multimedia systems: animation and multimedia databases, authoring and presentation, resource management, and scalable Web-based multimedia presentation servers. He received the prestigious National Science Foundation CAREER Award in 2003 for his proposal on animation databases. He has published several research papers in various refereed conferences and journals in this area. He has served as Associate Chair of the ACM Multimedia Conferences in 2003, 2000, and 1999 He has served as guest editor (special issue on Multimedia Authoring and Presentation) for *ACM Multimedia Systems*. He is also serving on the editorial board of *Multimedia Tools and Applications*, Kluwer. He has also served as program committee member on several multimedia conferences and workshops. He has presented tutorials at ACM Multimedia and other multimedia conferences. He has served as a visiting research faculty with the Department of Computer Science, University of Maryland, College Park. He also served as a faculty in the Department of Computer Science, National University of Singapore, and at the Indian Institute of Technology, Madras.