Medium Access Control With Physical-Layer-Assisted Link Differentiation

Fanglei Sun, *Student Member, IEEE*, Victor O. K. Li, *Fellow, IEEE*, Zhifeng Diao, and Zhengyuan Xu, *Senior Member, IEEE*

Abstract—In this paper, we develop medium access control (MAC) schemes for both contention and contention-free accesses over wireless local area networks and give performance analysis of these MAC protocols. User detection and multirate adaptation (MRA) modules are proposed in the physical layer (PHY) to assist link differentiation. With these two modules, for contention accesses, a new distributed queuing MAC protocol (PALD-DQMP) is proposed. Based on different users' channel states, PALD-DQMP makes use of a distributed queuing system to schedule transmissions. To support multimedia transmissions, an enhanced PALD-DQMP (E-PALD-DQMP) is designed by providing two-level optimized transmission scheduling for four access categories, thus eliminating both external and internal collisions among mobile stations. For contention-free accesses, based on the same PHY-assisted link differentiation provided by the two modules, a new multipolling MAC protocol (PALD-MPMP) is proposed, which not only reduces the polling overhead but also prioritizes transmissions according to their delay requirements. Performance analysis and simulation results show that our proposed protocols outperform the standard MAC protocols for both delay-sensitive and best-effort traffics. All these improvements are mainly attributed to the awareness of cross-layer channel state information and the consequent MRA scheme.

Index Terms—Cross-layer design, medium access control (MAC), physical-layer-assisted link differentation (PALD), user detetion, 802.11 wireless local area network (WLAN).

I. INTRODUCTION

THE IEEE 802.11 wireless local area network (WLAN) [1] has emerged as a prevailing technology for broadband wireless access in offices, airports, and even onboard vehicles [2]. In this paper, we develop medium access control (MAC) schemes for both contention and contention-free accesses over WLANs. For contention accesses, in IEEE Standard 802.11 MAC, the distributed coordination function (DCF) is based on the carrier sense multiple access with collision avoidance (CSMA/CA) protocol. Retransmissions of collided packets are managed according to the binary exponential backoff (BEB) algorithm. To support multimedia applications such as voice and video applications over the wireless medium, the enhanced distributed channel access (EDCA) is being developed in IEEE Standard 802.11e [3]. EDCA works with four access categories (ACs), which are virtual DCFs, and each AC achieves a differentiated channel access. This differentiation is achieved by varying the amount of time a station senses the channel to be idle and the length of the contention window (CW) for a backoff.

However, DCF is originally designed to support best effort data services, which are delay tolerable, but require loss-free transmissions. The BEB algorithm embedded in DCF suffers from a fairness problem and yields low throughput under heavy load [4]. This is because it assumes that all losses are due to collisions. If a frame is lost because of a random link error instead of a collision, doubling the CW is inappropriate and can seriously degrade performance. For a detected collision, the BEB algorithm only generates a period of random backoff time within an increasing CW size and then blindly retransmits it until the retransmitting time reaches the maximum. In addition, when different users experience different channel variations, fairness is also a serious problem for the BEB algorithm. For multimedia services, EDCA also incurs bandwidth wastage caused by collisions and backoff. The problem is exacerbated in EDCA because collisions are caused not only among different mobile stations but also among different ACs in the same station [5]. Therefore, the elimination of such backoffs should improve the network performance for both DCF and EDCA.

For 802.11 MAC, Pang *et al.* [6] proposed a method to distinguish between collisions and link errors in the MAC layer. For a frame loss, if the sender receives a clear-to-send (CTS) but no ACK, it deduces that the loss is caused by wireless errors; if the sender receives neither a CTS nor an ACK, it deduces that the loss is caused by collisions. As in [6], we also try to distinguish between collisions and link errors, but we focus on the physical layer (PHY) solution. In this paper, effective user detection and multirate adaptation (MRA) modules in the PHY are designed to feed back the error-type information, the detailed channel gain information, and the corresponding supported maximum transmission rate. In [7], a MAC protocol based on the distributed queuing random access protocol (DQRAP) is originally proposed for a code-division multiple-access (CDMA) mobile communication system. Subsequently, Alonso *et al.* [8] transplanted the same method to WLAN networks. This mechanism changes the frame structure and the contention mechanism. It is not easy to implement in

current WLAN systems, and the reserved m contention periods (CPs) in each frame cause additional transmission overhead. However, the queuing method inspires us to develop a new transmission scheduling scheme. With the PHY-assisted link differentiation, we proposed the distributed queuing MAC protocol (PALD-DQMP) for best effort data transmission, which has the following features: 1) the user detection module can provide effective loss differentiation information and users' channel state information; 2) the MRA module can dynamically choose the maximum transmission rate supported by the current channel condition; 3) combined with the aforementioned cross-layer information, PALD-DQMP utilizes a queuing system to handle the random access and data transmissions, thus eliminating the bandwidth wastage caused by BEB; and 4) PALD-DQMP can effectively enhance the network performance and fairness, particularly in WLANs with heavy load. Based on PALD-DQMP, E-PALD-DQMP is developed to support multimedia transmissions. In this enhanced version, we find the following: 1) With the help of service priority differentiation and PHY link differentiation, the bandwidth wastage caused by BEB is eliminated; 2) the QoS requirements of delay-sensitive services are guaranteed; and 3) the throughput and fairness problem in best-effort data services can be resolved.

For contention-free access schemes, the point coordination function (PCF) is an optional centrally controlled channel access function to support time-bounded multimedia services. The IEEE 802.11 standard specifies a round-robin polling mechanism whereby the access point (AP) polls the stations in ascending order of their association IDs (AIDs). This algorithm is easy to implement. However, when the number of stations in the polling list is large, stations will suffer from long access delays. Bandwidth wastage is caused by sending CF-Polls and Null packets from the stations that have no pending packets to send. In addition, PCF only defines a single class scheduling algorithm, which cannot handle various types of traffics with different QoS requirements. To design an efficient polling scheme, Kim *et al.* [9] and Lo *et al.* [10] proposed two types of multipolling protocols. Both of them introduced a rather long status collection period at the beginning of a contention-free period (CFP), which increases the complexity and causes more overhead. In this paper, based on our PHY-assisted user detection and MRA modules, a new multipolling MAC protocol (PALD-MPMP) is presented, which is characterized by the following: 1) an effective polling list with the priorities based on delay requirements or levels; 2) the polling sequence sorted by the detected channel state information; and 3) increased bandwidth usage due to the multipolling and MRA schemes.

Combining the contention and contention-free access schemes, in the IEEE standard MAC, each super frame consists of a CP, where contention-based access schemes are used, and a CFP, where contention-free access schemes are used. The CFP will be repeated after each CFP repetition (CFPR) interval. Fig. 1 shows the superframe structure and an example of a foreshortened CFP. A foreshortened CFP is caused by the extension of the previous CP.

This paper is organized as follows: Section II presents the PHY user detection and MRA modules. Section III describes PALD-DQMP and its performance via simulations. Its en-
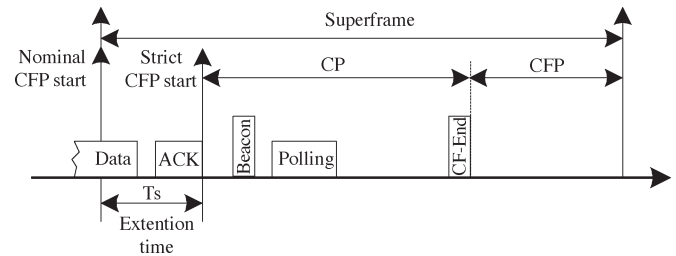

Fig. 1.   Superframe structure.

hanced version, i.e., E-PALD-DQMP, and the corresponding simulation results are presented in Section IV. Section V discusses the contention-free access control of PALD-MPMP. Section VI concludes this paper.

II. USER DETECTION AND MRA MODULES

A. User Detection Module

In IEEE 802.11, there are two types of WLAN topologies: 1) the independent basic service set (IBSS) as an *ad hoc* network and 2) an infrastructure network. In an infrastructure network, an AP is fixed and acts as a hub and connects the basic service set (BSS) network to an extended service set network. In contention-based access control schemes, we assume that any node can be chosen as a coordinating node (CN). Even with this CN, the IBSS still works in a distributed manner because CN only acts as a detector of the access requirement/ acknowledgment, and it behaves simply as a repeater to broadcast its information to other nodes. If the moving speed of the current CN exceeds the speed supported by IEEE 802.11 WLANs, another mobile node can be chosen as a CN. The information exchange can be effected by the reassociation-related services. In contention-free access control schemes, an AP is chosen as a CN. According to the statement of logical service interfaces in [1], the necessary information is exchanged between the AP/CN and other nodes by association-related services and authentication-related services at network initialization. Among such information, we are most interested in the transmitter addresses (TAs) of the active nodes, which are easily gathered by APs/CNs.

Multiuser detection is an effective technique originally used in CDMA systems. Recently, there has been an increased interest in applying multiuser detection in the cross-layer design of other types of network systems [12]. Multiuser detection can successfully separate users at the receiver based on their specific spreading codes [14]. In WLANs, there exists similar user-dependent information—TA. DCF describes two techniques in frame transmission. One is a two-way handshaking technique, which is called the basic access mechanism. The other is a four-way handshaking technique, which is known as the request-to-send/CTS (RTS/CTS) mechanism. The RTS/CTS mechanism improves the system performance by deducing the duration of a collision when long messages are transmitted. With the capability to cope with hidden terminals, the RTS/CTS access scheme is widely used. In addition, as retransmissions of data frames due to collisions may more frequently occur, the retransmission method becomes less cost-effective than the

use of the RTS/CTS frame exchanges. For these reasons, this paper mainly focuses on the RTS/CTS mechanism. In both contention-based and contention-free access control schemes, assuming users can synchronously or asynchronously send RTS frames, some multiuser detection techniques can be adapted to WLAN systems and provide further channel condition information. Based on this information, it is ensured that the MAC protocol can schedule the different users' transmissions, accounting for the users' QoS or fairness requirements. In terms of the throughput, by eliminating the collision errors and the consequent bandwidth wastage, user detection can effectively improve the network performance. Therefore, the study of user detection is important to improve the performance of MAC protocols. In this section, a synchronous user detection technique for WLANs will be presented. If RTSs cannot be synchronously transmitted, there are some asynchronous multiuser detection techniques used in CDMA systems, which can also support the PHY link differentiation. However, the study of asynchronous user detection will be discussed in our future work.

It has been previously mentioned that the TAs of active users are the information gathered at an AP/CN after network initialization. From the definition of the RTS frame format, a TA has a fixed 48-bit field embedded in an RTS frame; thus, at the PHY layer, this 48-bit information can be accordingly extracted. Similar to spreading codes in CDMA systems, in our detection scheme, TAs will be used to identify active users. In this section, we first assume that the RTSs are synchronously transmitted, and the realization of RTS synchronization will be presented in Section III-B. Consider a BSS system where K mobile nodes are associated with an AP/CN, and the AP/CN keeps a transmission address list of the active nodes in its BSS. The TA of mobile node k is $\mathbf{c}_k$, which is an N-bit ($N = 48$) vector; thus, the list of TAs can be recorded as follows:

$$\mathbf{C} = [\,\mathbf{c}_1, \quad \ldots, \quad \mathbf{c}_k, \quad \ldots, \quad \mathbf{c}_K\,] \tag{1}$$

where $\mathbf{c}_k = [c_{k,1}, \ldots, c_{k,n}, \ldots, c_{k,N}]^T$, and $c_{k,i}$ is the ith bit of the N-bit address $\mathbf{c}_k$. As in a synchronous CDMA system, if TAs are orthogonal to each other, then the best detection performance will be achieved. Therefore, in WLAN systems, to improve the detection performance, orthogonal 48-bit vectors are reassigned to users as TAs in different cells. Each orthogonal vector corresponds to each TA in a cell, and for different cells, these orthogonal vectors can be reused. For K users, the vector $\mathbf{h}$ represents the channel gains experienced in their respective wireless channels, i.e.,

$$\mathbf{h} = [\,h_1, \quad \ldots, \quad h_k, \quad \ldots, \quad h_K\,]^T. \tag{2}$$

At time t, we assume that l ($0 \leq l \leq K$) users have data to simultaneously transmit and send RTSs to contend the channel. For users who have not sent RTSs, their corresponding channel gains in vector $\mathbf{h}$ will be zero. After the RTS contention procedure, the received signal is given by

$$\mathbf{r} = \mathbf{C}\mathbf{h} + \mathbf{n} \tag{3}$$

where $\mathbf{n}$ is the noise. Consequently, according to the least squares criterion, the users' channel gain vector is estimated

at the receiver by the following formula:

$$\hat{\mathbf{h}} = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{r} = \mathbf{\Pi}\mathbf{r} \tag{4}$$

where $\mathbf{\Pi} = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T$. Basically, each calculated element $\hat{h}_k$ represents the detected channel gain of each active user. It is necessary to distinguish these results into three cases.

1) Silence state: The active node has no RTS requirement in the current contention round.
2) Link error state: The active user takes part in the channel contention; however, its channel quality is deduced to be inadequate for transmission.
3) Transmission state: The active user attends the current contention and obtains a good channel to transmit its pending frame.

To implement the aforementioned differentiation function, our user detection module utilizes the power value of h_k, namely $|\hat{h}_k|^2$, as the decision criterion and then constructs two thresholds to discriminate the aforementioned three cases. For those users who have not participated in the RTS contention, their $|\hat{h}_k|^2$ values should be zero. However, with noise present, the value of $|\hat{h}_k|^2$ could be nonzero. Thus, it is required to set the threshold $H_{\text{silence−thresh}}$ to distinguish the first case from the other two cases. By substituting (3) in (4), we obtain

$$\hat{\mathbf{h}} = \mathbf{h} + \mathbf{\Pi}\mathbf{n}. \tag{5}$$

Clearly, the power of the noise signal in the vector $\mathbf{\Pi}\mathbf{n}$ should be the lower bound of the silence threshold. We assume that $\mathbf{n}$ is an additive white Gaussian noise vector whose elements have a zero mean and a variance of σ^2. Denote the calculated inverse matrix by $\mathbf{\Pi} = [a_{i,j}]_{K \times N}$. Then, we obtain the lower bound of the silence threshold $H_{\text{lower−bound}}$, i.e.,

$$H_{\text{lower−bound}} = E\left\{|\mathbf{\Pi}\mathbf{n}|^2\right\}$$
$$= \sigma^2 \times \left[\sum_{j=1}^{N}(a_{1,j})^2, \ldots, \right.$$
$$\left. \sum_{j=1}^{N}(a_{k,j})^2, \ldots, \sum_{j=1}^{N}(a_{K,j})^2\right]_{1 \times K}^T \tag{6}$$

where $|\cdot|$ is an operator that takes the absolute value of each element in a vector. We denote the second threshold as $H_{\text{transmission−thresh}}$ to distinguish the second case from the third case. This parameter is related to the specific system performance, the user's QoS requirements, etc. In this paper, we consider the users' packet error rate (PER) requirements. With the relationship between the signal-to-noise ratio (SNR) and packet loss rate (PLR) [13], given a PER requirement, we can easily obtain the corresponding SNR value and furthermore obtain $H_{\text{transmission−thresh}}$ for a good channel with the detected noise power mentioned in the next section. In practice, the value of $H_{\text{silence−thresh}}$ will be chosen between the aforementioned lower bound and $H_{\text{transmission−thresh}}$. An improper $H_{\text{silence−thresh}}$ will cause two types of errors. One type of error is experienced by users whose channel gain power
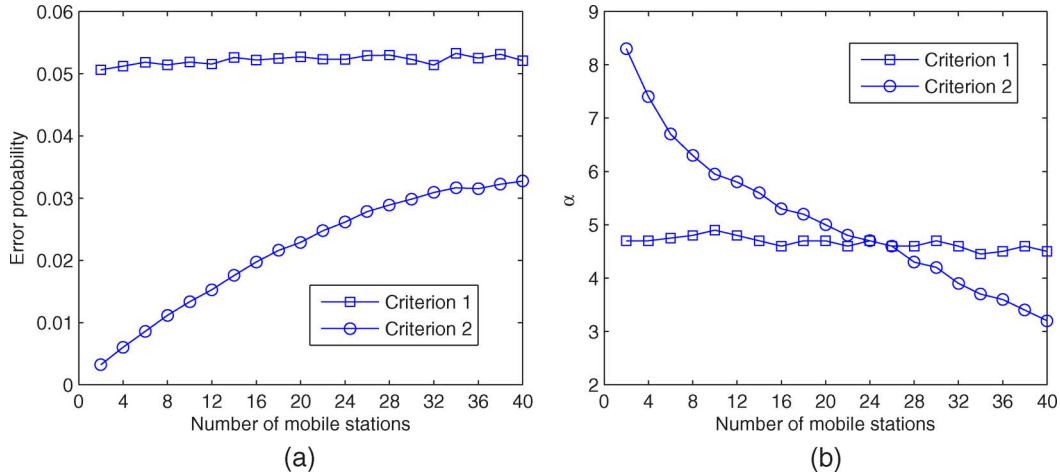
Fig. 2. Comparisons of the two criteria. (a) Error probabilities. (b) α values corresponding to the minimized error probabilities.

is higher than $H_{\text{silence-thresh}}$, but the detected values are lower than $H_{\text{silence-thresh}}$. The other type of error is experienced by users who have not sent RTSs, namely their channel gain power should be lower than $H_{\text{silence-thresh}}$, whereas the detected values should be higher than $H_{\text{silence-thresh}}$. Thus, in our multiuser detection module, we introduce a parameter α, upper bounded by $\alpha_{\max}$, which can minimize the detection error based on the following silence threshold:

$$H_{\text{silence-thresh}} = \alpha \times H_{\text{lower-bound}}. \qquad (7)$$

Two criteria can be adopted to determine a proper α value.

1) As previously introduced, we can conclude that one type of error happens to the set of active users, whereas the other type of error happens to the set of silent users. Through thousands of simulations, we average the error probability of active users as $p_{\text{active-error}} = N_{\text{error-active-users}}/N_{\text{active-users}}$ and the error probability of silent users as $p_{\text{silent-error}} = N_{\text{error-silent-users}}/N_{\text{silent-users}}$. Therefore, we obtain the first criterion, i.e.,

$$\alpha = \{\alpha \mid \min(p_{\text{active-error}} + p_{\text{silent-error}})\}$$
$$0 \leq \alpha \leq \alpha_{\max} \quad (8)$$

i.e., find α that minimizes the sum of $p_{\text{active-error}}$ and $p_{\text{silent-error}}$.

2) In the second case, we jointly consider the two types of errors; thus, the total error probability is $p_{\text{total-error}} = (N_{\text{error-active-users}} + N_{\text{error-silent-users}})/N_{\text{total-users}}$. Then, we have

$$\alpha = \{\alpha \mid \min(p_{\text{total-error}})\}, \qquad 0 \leq \alpha \leq \alpha_{\max}. \quad (9)$$

Fig. 2 shows the simulation results of the aforementioned two criteria. The Rayleigh fading model is used to generate channels in thousands of independent channel realizations. Channel noise is assumed to be white Gaussian. SNR is set to be 20 dB, $\alpha_{\max} = 10$, and $N = 48$. The number of mobile stations varies from 2 to 40. With different numbers of mobile stations, Fig. 2(b) shows the α values that are targeted to minimize the error probabilities according to the two criteria,

whereas Fig. 2(a) shows the corresponding minimized error probabilities. Here, the error probability considers both of the aforementioned two types of errors, either from active users or silent users. The simulation results show that the error probabilities based on the first criterion are higher than the corresponding values based on the second criterion, whereas the values of α for the first criterion are more stable than those values for the second criterion. This is a tradeoff. With different network scenarios, the first criterion can decrease the complexity of threshold detection but at the cost of a higher error probability. The second criterion can provide a lower error probability, but it needs to test different values of α. In fact, the aforementioned errors will not directly induce transmission failures. Section III-C will show that the performance variation caused by the detection error is much lower.

It should be mentioned that since $N = 48$, the maximum number of detectable users is 48. If there are more than 48 active users, we can consider adding additional address information in the RTS frame or assign new orthogonal addresses for them, but this will cause additional overhead. Alternatively, we can distribute additional users to other APs/CNs.

B. MRA Module

By using the concept of adaptive modulation [15], stations in a multirate WLAN system are assigned the modulation scheme and transmission rate according to the detected SNR and the required transmission quality. Each modulation scheme could be further mapped to a range of SNR at a given transmission power. To achieve a high transmission efficiency in WLANs, stations shall select the highest available rate modulation scheme according to the detected SNR. Based on the aforementioned user detection module, we assume that the noise power is detected at silent periods before the start of data transmissions; thus, APs/CNs can easily estimate the noise power by $\hat{\sigma}^2 = \|\hat{\mathbf{n}}\|^2/N$. Furthermore, by the definition of SNR, for any active user, we can easily detect its SNR value, i.e.,

$$\hat{\text{SNR}}_k = \frac{|\hat{h}_k|^2}{\hat{\sigma}^2} \quad (k = 1, 2, \ldots, K). \qquad (10)$$

In this paper, we are mainly concerned with the 802.11b WLAN, which provides four types of supported rates. In IEEE 802.11b, the PHY rates of 1 and 2 Mb/s are based on binary phase-shift keying (BPSK) and quadrature phase-shift keying (QPSK) modulations, respectively. To provide the higher PHY rates of 5.5 and 11 Mb/s, a complementary code keying (CCK) modulation scheme is also defined. CCK is a variation on m-ary orthogonal keying modulation, which uses an in-phase quadrature modulation architecture with complex symbol structures. Without MRA, we assume that the channel state is either good or bad, as differentiated by $H_{\text{transmission}-\text{thresh}}$. With MRA, a set of SNR thresholds for different modulations should be defined to select the appropriate data rates in the PHY. In the simulations, these thresholds have been selected based on the results presented in [16]. Thus

$$R = \begin{cases} 1 \text{ Mb/s (BPSK)}, & \text{SNR} < \gamma_1 \\ 2 \text{ Mb/s (QPSK)}, & \gamma_1 \leq \text{SNR} < \gamma_2 \\ 5.5 \text{ Mb/s (CCK)}, & \gamma_2 \leq \text{SNR} < \gamma_3 \\ 11 \text{ Mb/s (CCK)}, & \gamma_3 \leq \text{SNR}. \end{cases}$$

III. PALD-DQMP

A. Protocol Description

Based on the realization of the PHY-assisted link differentiation scheme, the proposed MAC protocol should address two issues: 1) how to synchronize the required RTSs but keep the random access characteristic for each node and 2) how to make use of such information to effectively schedule the data transmissions if the coexisting synchronized RTS signals are separated and the channel state information is detected. In this section, PALD-DQMP utilizes a queuing mechanism to solve these two problems.

A logical distributed queue is constructed in PALD-DQMP. This queue is simply represented by two integer numbers at each node, which are denoted as pDQ and DQ. DQ represents the total frame number of the data transmission queue. pDQ is the position of a given node in the data transmission queue. Its value ranges from 0 to DQ. The node does not have any position in the queue if it is 0. DQ has the same value for all active nodes, whereas pDQ has a specific value for each node. We assume that the queue is first-in first-out. Both values are initially set to zero and will be updated by broadcasted CTS/ACK frames from the AP/CN. A set of rules is required for updating the two queue parameters and scheduling the consequent frame transmissions. With this procedure, collision-free data transmissions can be achieved.

The following describes our algorithm executed at each node. We separate it into two phases: channel contention phase and data transmission phase.

1) Channel Contention Phase:

1) In the initial state, both DQ and pDQ are zero. For users with pending data to transmit, their RTSs are easily transmitted synchronously after a period of distributed interframe spacing (DIFS). Along with this procedure, the network allocation vectors (NAVs) of the other active users are set to the longest time noted in the "Duration/ ID" field. This 16-bit field lies in the MAC header and
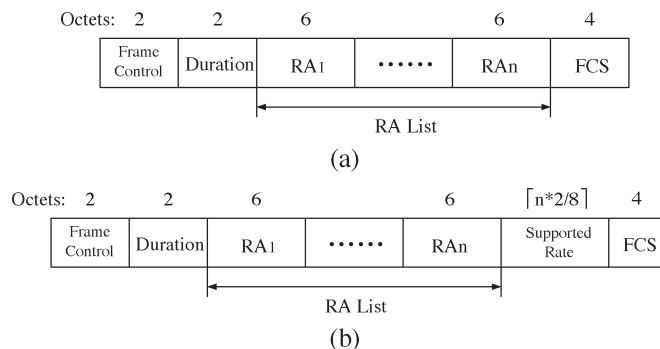

Fig. 3.   Frame format of CTS. (a) Without MRA support. (b) With MRA support.

contains a duration value for each frame type defined in the IEEE standard.

2) After the receiver obtains the coexisting signals, the PHY-assisted link differentiation will be triggered to detect users' channel conditions and distinguish multiple supported maximum rates. The traditional CTS format only includes one receiver address (RA) to which the CTS is destined. In our mechanism, we extend this field to an RA list, which includes multiple RAs to which CTS plans to feed back. To demonstrate the impact of the MRA module on the performance, our PALD-DQMP has two ways to broadcast CTS, which are different in the content of the RA list and supported rate field, as shown in Fig. 3. One is for PHY without MRA support. The other one is for PHY with MRA support.

a) The RA list only includes the MAC addresses of the users whose channels are in transmission states. In this step, PALD-DQMP sorts the RA list by the priority of their corresponding users' deduced power values of the channel gains. For the AP/CN, the DQ value can be easily obtained. For other active nodes, their DQ values are updated by counting the number of RAs in the RA list. To update pDQ, a user compares RAs in the CTS to its own address. If a certain RA value in the RA list is identical with its own address, its pDQ value is updated to the sequence of this RA in the RA list; if not, the pDQ value remains zero. By receiving this broadcast CTS with multiple RA information, all active users know that the frame transmission phase will start after a period of short interframe spacing (SIFS).

b) Let us consider users with both transmission and link error channel states. With MRA, different supported data rates can be dynamically chosen according to the users' current detected SNRs. This can eliminate the packet losses due to lower SNRs and ensure high-efficiency code modulation for the users with high SNRs. The RA list of the CTS should include not only the MAC addresses of the collided nodes but also the MAC addresses of the link error nodes. In addition, their supported maximum data rates shall also be reported back to the required users. Since the IEEE 802.11b WLAN system supports four types of data rates, two bits are assigned for each user to indicate
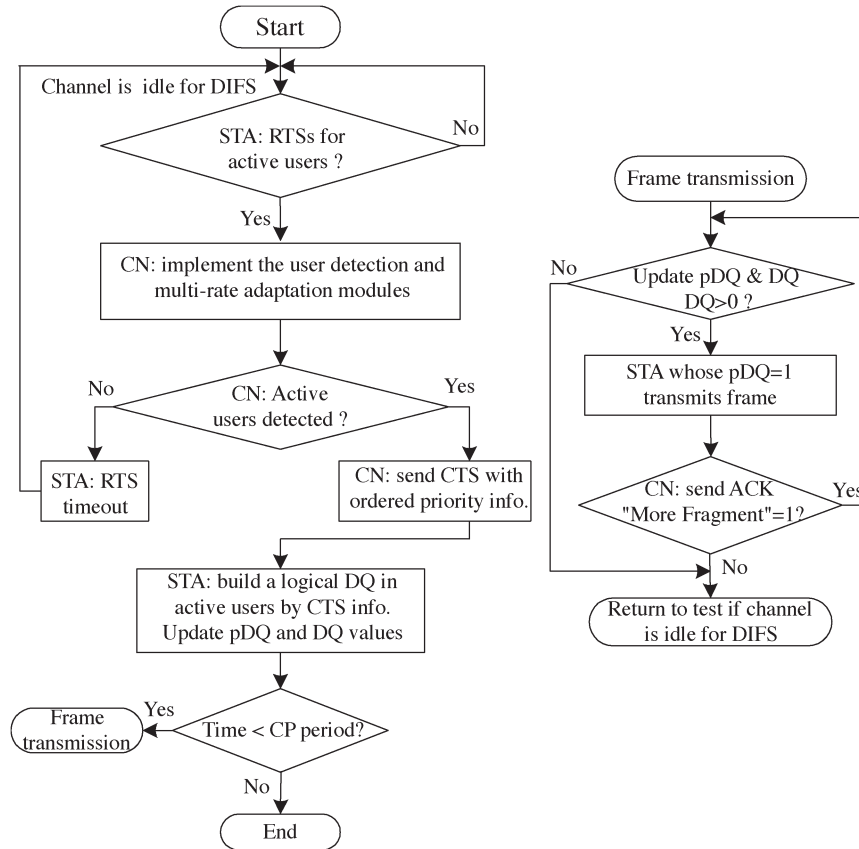
Fig. 4.   Flowchart of PALD-DQMP.

the currently supported data rate. Assuming that the current detected active user is n, the rate information in a CTS frame needs $\lceil n \times 2/8 \rceil$ bytes. The updating process of pDQ in mobile users is the same as that described in (a).

3) In general, during the channel contention or data transmission phases, if a node has data to transmit, it will check both the values of DQ and NAV. If one or both of them are nonzero, the RTS will be deferred to the next CP.

2) Data Transmission Phase: At the end of the channel contention phase, each node obtains the updated DQ and pDQ values and their corresponding supported maximum data rate. After a period of SIFS, data transmissions will be scheduled by two rules.

1) For the user whose pDQ value equals 1, it has the highest priority to seize the channel, and a frame is transmitted at the supported rate. After receiving the ACK of this frame broadcasted from the AP/CN, all active nodes reduce their DQ and pDQ values by 1. Consequently, the node whose pDQ value changes to 1 wins the next data transmission chance.

2) Once the last ACK that causes the values of DQ, pDQ, and NAV in all mobile nodes to change to zero is received, the current data transmission phase is finished, and PALD-DQMP repeats from the next channel contention phase. Fig. 4 shows the flowchart of PALD-DQMP.

B. RTS Synchronization

In the frame control field of MAC headers, there is a 1-bit "More Fragment" field, which indicates whether there is another fragment of the current MAC service data unit (MSDU). Our protocol utilizes this bit in a series of ACKs to inform all active users whether this ACK is an acknowledgment to the last scheduled frame. Since the AP/CN keeps the DQ information, if it deduces that the received data frame is not from the last scheduled user, it will broadcast an ACK with its "More Fragment" bit set to 1. On the contrary, if it deduces that the data frame is from the last scheduled user, it will broadcast the last ACK with the "More Fragment" set to 0. Upon receiving the last ACK, all users, either actively participating in the current-round transmission or silent, know that a new round channel contention phase will start. Thus, after a DIFS period, which can be exactly measured at each mobile station, the users who have data frames to transmit will synchronously send an RTS frame.

Through this distributed queuing system, the last ACK can be regarded as the AP/CN-issued beacon to synchronize the RTS transmissions from different users. We consider WLAN systems where only one channel exists. When considering the interference, for the standard MAC, the users' frame transmissions will be deferred by backoff procedures. However, for our designed protocols, all active nodes sense the channel to be idle by receiving the broadcasted AP/CN-issued beacon; then, after DIFS, they synchronously send RTSs. Therefore,

the interference has no effect on the sensing of the channel idle state in our protocols. In the PHY in WLAN systems [1], the short preamble header, including synchronization and start frame delimiter fields, is used to extract the received clock for carrier and bit synchronization. In this paper, we assume that the propagation delays for different users in the WLAN transmission range to reach the AP/CN are the same. This assumption is reasonable since the WLAN transmission range is relatively small. Thus, the multiuser detection module introduced in Section II-A effectively works. In addition, when considering asynchronous systems in which different propagation delays are experienced by different users, there are mature asynchronous multiuser detection techniques available for CDMA systems. When propagation delays are known or have been estimated, then the data model follows [17], which doubles the number of columns of matrix $\mathbf{C}$ to incorporate asynchrony. Either the subspace channel estimation method in [17] or our proposed detection scheme in Section II-A is applicable to estimate channel gains. In the absence of delay information, joint channel and delay estimation can be performed by adopting the subspace method for asynchronous CDMA systems detailed in [18]. Therefore, the multiuser detection techniques, whether in synchronous or asynchronous systems, can be used in the cross-layer design in WLAN networks.

In our protocols, for mobile users, if some of their RTSs are lost or some of them fail to receive the corresponding CTS, their data transmissions will be deferred to the next-round channel contention phase. In terms of data frames, because the channels of the scheduled users have already been checked by the previous RTSs, the frame loss probability is really small. The following situations only happen with some unexpected changes on the wireless channels. In case that some scheduled users fail to receive ACKs, to smoothly carry on the scheduled transmissions, the system will not give them more chances to retransmit, and the loss recovery will be deferred to the next-round transmissions. Since ACKs are broadcasted information, if active users fail to receive the next ACK within a timeout period, they will automatically reduce their DQ and pDQ values by 1. For the receiver AP/CN, if data frames are not received from the scheduled users for a long time, the AP/CN will timeout to broadcast an ACK, which may be the last ACK, to inform the mobile users to update their related information and to continue the following frame transmissions. It is important to point out that during the channel busy period, the last ACK is used to synchronize the RTSs in the next CP, and the time interval between these last ACKs dynamically changes with the variation of the number of scheduled users in a transmission round. However, during the channel idle period, a control frame, such as the last ACK with "More Fragment" set to 0, is broadcasted in a constant time interval to synchronize the new randomly generated RTS frames. This method is also useful when CTS is not successfully received by all active users.

C. Performance Analysis

In this section, we use IEEE 802.11b as an example. The notations and parameter values used in this paper are given in Table I.

TABLE I
NOTATIONS AND PARAMETER VALUES

Notations	Descriptions	Values
MAC	MAC header overhead	224 bits
DIFS	The period of DIFS	$50\,\mu s$
SIFS	The period of SIFS	$10\,\mu s$
PIFS	The period of PIFS	$30\,\mu s$
RTS	RTS frame size	160 bits
CTS	CTS frame size	112 bits
Payload	Packet size	1000 bytes
ACK	ACK frame size	112 bits
CWmin	Minimum value of CW	31
CWmax	Maximum value of CW	1023
SlotTime	Backoff slot time	20
PHY	PHY header overhead	192 bits
DataRate	Physical rate for data frame	5.5 M
BasicRate	Physical rate for control frame	1 M
t_p	Propagation delay	$1\,\mu s$

1) Throughput Bound Analysis: Throughput bounds of the standard DCF and of our PALD-DQMP are calculated for the WLANs where the wireless channels are perfect, i.e., without fading and noise. We assume that each MSDU only includes one frame.

1) DCF: With the increase in the number of mobile nodes, in the BEB algorithm, the high collision probability causes frequent start-ups of the backoff procedure. Obviously, this will waste bandwidth and cause fairness problems. Therefore, the throughput bound will be reached when only one pair of mobile nodes is transmitting data. In this case, there are no backoffs and no bandwidth wastage. Thus, we have the throughput bound of the standard MAC protocol in (11), shown at the bottom of the next page.

2) PALD-DQMP: However, with perfect channels, the throughput bound for DCF corresponds to the worst case for the PALD-DQMP protocol. The reason is that multiple users share the same RTS/CTS (with RA list) transmitting time, and their transmissions are scheduled by a logical queue, without bandwidth wastage. Thus, the overhead of the RTS/CTS frame will be really small if a great number of nodes are contending for the channel; then, we have the new throughput bound of our protocol in (12), shown at the bottom of the next page, where K is the number of users. In addition, if the MRA scheme is used, higher data rates will be supported; thus, (MAC + Payload)/DataRate will have a smaller value, and a better throughput performance is obtained.

For fairness consideration, the BEB scheme resets the CW of a successful sender to the minimum value, whereas other nodes continue to maintain larger CWs, which reduces their chances of seizing the channel and results in channel domination by successful nodes. However, in PALD-DQMP, users are already fairly scheduled according to their channel conditions, as will be demonstrated in our simulation results.

2) Throughput Analysis With Fading Channels: In a practical scenario, a certain amount of packet loss is unavoidable due to the bad quality of wireless channels. In the standard MAC protocol, failed users will first generate a period of backoff time, which may increase with the increase of the retry time,

and then retransmit an RTS frame. When its retry time reaches the maximum, this frame will be dropped. Therefore, more backoffs will cause more bandwidth wastage.

For PALD-DQMP, collided users with good channels can be effectively separated, and then, their transmissions are scheduled. Although some users' current channels are weak, the user detection model can still detect their data transmission requirements. Without the MRA scheme, users with bad (error-prone) channels will be deferred to the next CP, thus only wasting its shared RTS/CTS overhead. With the MRA scheme, their transmissions will be served with lower data rates. Therefore, the throughput will be close to the aforementioned bound.

3) Throughput Analysis With Noisy Channels: For a channel with noise, in PALD-DQMP, the user detection module uses $H_{\text{silence-thresh}}$ to identify the active users who have RTS requirements. If the effect of noise is small, the deduced threshold is effective. When the noise is big, we need to consider two types of errors caused by the improper threshold. One type of error is experienced by the user who has sent an RTS but may be on a weak channel. The value of $|\hat{h}_k|^2$ should be between $H_{\text{transmission-thresh}}$ and $H_{\text{silence-thresh}}$, but the detected result should be lower than the corresponding $H_{\text{silence-thresh}}$. In this case, a user will wait until timeout and then attend the next channel contention round. Compared with DCF, this type of error will not affect the throughput. The other type of error is experienced by the user who has not sent an RTS, but the detection module mistakenly includes its MAC address in the RA list. In this case, the user only sends a short frame without data payload to inform other nodes to update their DQ and pDQ values.

D. Simulation Results

This section validates the performance of the standard DCF, PALD-DQMP with a constant data rate, and PALD-DQMP with the MRA scheme. All simulations in this paper are conducted using Matlab V7.1. We simulate WLANs where K nodes contend to transmit to an AP/CN. Each node is assumed to always have enough data to send. Simulation parameter values are shown in Table I. In the PHY, a Rayleigh fading model with a channel gain equal to 1 is chosen. In the simulations, we did not consider the specific channel variations caused by different locations of mobile nodes or their position changes. We consider the more general conditions where users' channels fluctuate under the Rayleigh fading channel model. Channel noise is assumed to be white Gaussian with SNR equal to 20 dB.
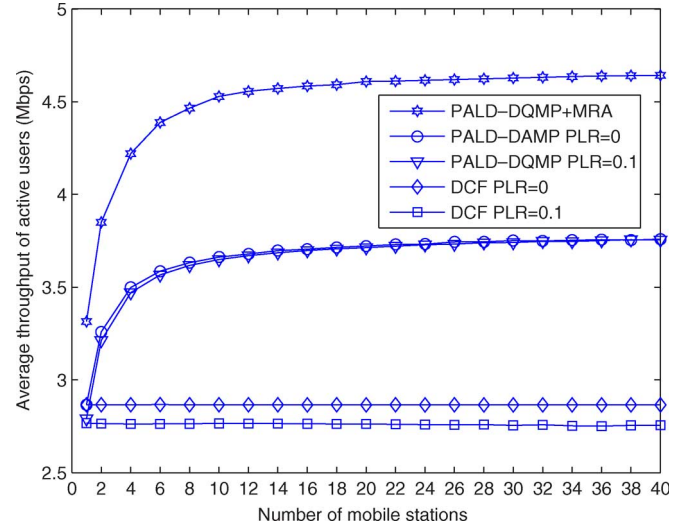

Fig. 5. Throughput comparisons.

In addition, we use $H_{\text{transmission-thresh}}$ to control the network PLR. Under perfect (PLR = 0) and lossy (PLR = 0.1) channel conditions, we test the network performance when the number of nodes changes from 1 to 40. All simulations in this paper are done in 100-s periods.

1) Throughput Comparisons: As shown in Fig. 5, the throughputs of the protocols are similar to each other when the number of nodes is 1. When the number of mobile nodes increases, the performance of DCF becomes slightly worse. It is because when the number of mobile nodes increases, the collisions among them will also increase, which may lower the network throughput. However, the chances for random accesses controlled by the BEB algorithm will increase. Combining the effects of these two factors will slightly decrease the throughput. For the performance of PALD-DQMP, as analyzed in Section III-C1, the RTS/CTS overhead is shared by users that have pending data to transmit. Thus, with the increase in the number of these users, the throughput should increase and be close to the theoretical bound under perfect channels. With MRA, higher data rate support causes another distinct improvement of our protocol.

2) Fairness Comparisons: The fairness index f is defined as $f = (\sum_{i=1}^{K} x_i)^2 / K(\sum_{i=1}^{K} x_i^2)$ [19], where K is the number of active users in the network, and x_k is the throughput achieved by user k, $1 \le k \le K$. Fig. 6 shows the fairness index comparison. It can be found that our proposed PALD-DQMP always

$$T_{\text{std}} = \frac{\text{Payload}}{\frac{\text{RTS} + \text{CTS} + \text{ACK}}{\text{BasicRate}} + \frac{\text{MAC} + \text{Payload}}{\text{DataRate}} + \text{DIFS} + 3\text{SIFS} + 4\text{PHY} + 4t_p} \tag{11}$$

$$T_{\text{new}} = \frac{\text{Payload}}{\left(\frac{\text{RTS} + \text{CTS}}{\text{BasicRate}} + \text{DIFS} + \text{SIFS} + 2\text{PHY} + 2t_p\right)/K + \frac{\text{ACK}}{\text{BasicRate}} + \frac{\text{MAC} + \text{Payload}}{\text{DataRate}} + 2\text{SIFS} + 2\text{PHY} + 2t_p} \tag{12}$$
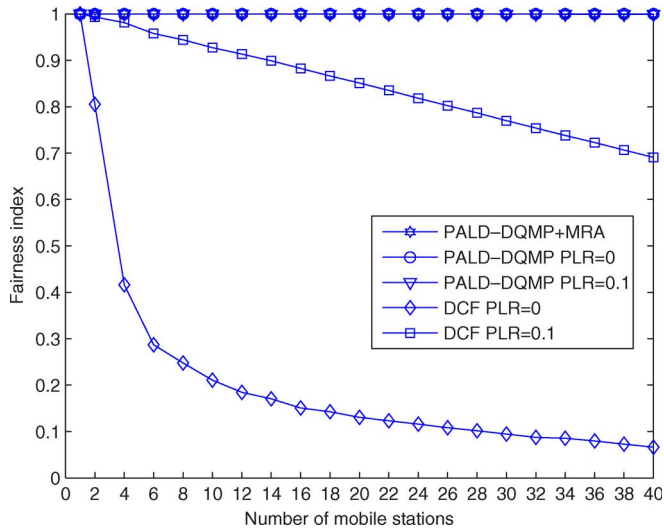
Fig. 6. Fairness comparisons.

keeps the fairness index close to 1. For DCF, the fairness index under perfect channels sharply decreases with the increase in the number of mobile users, whereas under lossy channels, the fairness in DCF is improved, but its performance is still much worse than that of PALD-DQMP.

IV. E-PALD-DQMP

In multimedia transmission systems, time-bounded multimedia applications (e.g., voice over Internet Protocol and video conference) require certain bandwidth, delay, and jitter guarantees. With DCF, all stations compete for the channel with the same priority. To support the QoS requirements for multimedia services, EDCA is developed with four ACs or virtual DCFs, and each AC achieves a differentiated channel access priority by assigning different arbitration interframe space (AIFS) and maximum and minimum CW size (CWmin and CWmax, respectively) values [20]. However, EDCA is only an extension of DCF, and the BEB algorithm adopted in DCF meets similar, or even worse, problems in EDCA. To solve these problems, based on our improved PALD-DQMP over DCF, E-PALD-DQMP is proposed to support the QoS requirements for multimedia services.

A. Protocol Description

EDCA supports eight types of traffics with different priorities, which are further mapped into four ACs. E-PALD-DQMP also works with four ACs, which are virtual PALD-DQMPs, as shown in Fig. 7. In the traditional IEEE 802.11e MAC, EDCA is executed in every backoff entity with different parameter values for the EDCA parameter set. The same EDCA parameter set is used by the backoff entities of the same AC in different stations. It is essential that the same values for the parameters are used by all backoff entities, whereas in E-PALD-DQMP, each frame unit is regarded as the basic contention entity. Their different access priorities are achieved by the following two-level contention mechanism.

1) Service-Level Priority: As shown in Fig. 8, each type of service also starts the timer after detecting the medium as
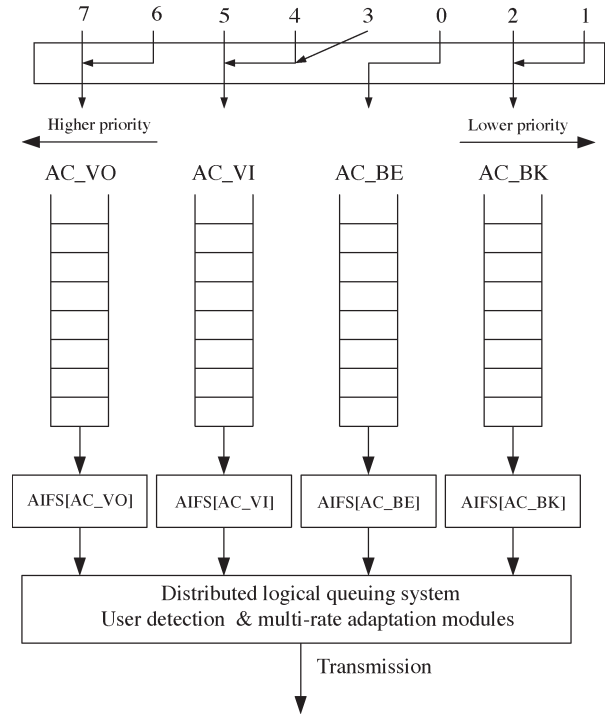

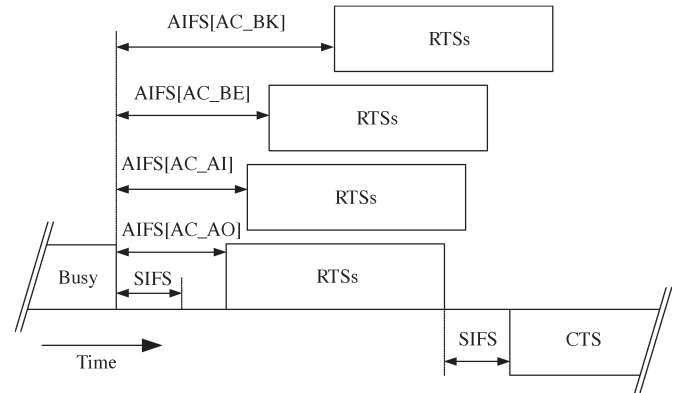Fig. 7. Station structure with E-PALD-DQMP.

Fig. 8. E-PALD-DQMP timing diagram.

idle for a duration defined by AIFS (AIFS$[i]$, $i = 0, \ldots, 3$), instead of DIFS, which is used in PALD-DQMP. AIFS$[i]$ can be enlarged per AC with the help of the AIFS number (AIFSN$[i]$, $i = 0, \ldots, 3$). AIFSN$[i]$ also denotes the duration of AIFS[AC] according to

$$AIFS[i] = SIFS + AIFSN[i] \times aSlotTime \qquad (13)$$

where AIFSN$[i] \geq 1$. The parameter aSlotTime defines the duration of a slot. The smaller the AIFSN$[i]$, the higher the medium access priority. In this paper, we assign the values 1, 2, 3, and 4 to the services of voice, video, best effort, and background data, respectively. In E-PALD-DQMP, by simply setting the same AIFSN$[i]$ for the same service for different users, the service-level priority is guaranteed.

2) Frame-Level Priority: Since E-PALD-DQMP also works with four ACs, four logical distributed queues are constructed in E-PALD-DQMP, which are responsible for the transmission
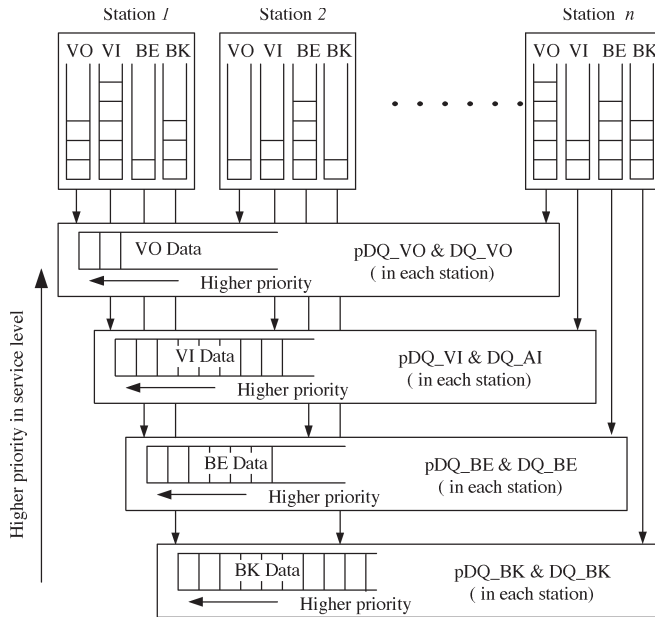
Fig. 9.   Queuing system of E-PALD-DQMP.

scheduling for voice, video, best effort, and back ground data. Similar to the employment of distributed queue in PALD-DQMP, each of these four queues is represented by two integers at each node. Thus, eight integers need to be recorded at each node. DQ_AO, DQ_AI, DQ_BE, and DQ_BK represent the total frame numbers of these four queues, with the same value for each node, whereas pDQ_AO, pDQ_AI, pDQ_BE, and pDQ_BK record the position of a given node in the corresponding queue. These four variables range from 0 to DQ_AO, DQ_AI, DQ_BE, and DQ_BK. If the node does not have a position in these queues, the corresponding position value is equal to 0. All these variable values are initially set to zero and updated by broadcasting CTS/ACK frames from APs/CNs. Fig. 9 illustrates the queuing system of E-PALD-DQMP in 802.11e MAC. ACs with the same priority in different nodes work as nodes with PALD-DQMP. They have the same AIFS$[i]$ values. Fig. 10 is the flowchart of E-PALD-DQMP.

In EDCA, for independent backoff entities, the smaller the CW, the higher the priority in medium access. However, the collision probability increases with a smaller CW if there are more than one backoff entity. Furthermore, it is noted that there are also retry counters that limit the number of retransmissions. Recently, some researchers have focused on adaptively adjusting the CW size to solve the problem. Romdhani *et al.* [21] adjusted the size of the CW of each AC according to both the traffic characteristics and the network conditions. Wong and Donaldson [22] changed the size of the CW of each AC considering the real-time data delay bound. However, based on [23], the CW size range of the high priority traffic is already very narrow. Thus, the algorithm could not have obvious improvements. As previously described, during contention, when a virtual collision occurs, the backoff entity with the higher priority will transmit, whereas all other backoff entities will act as if a collision occurred on the medium. The transmission of the backoff entity with a higher probability may still collide with another transmission initiated by other stations. Therefore,

both the internal and external collisions will cause bandwidth waste, delay, and fairness problems. However, in E-PALD-DQMQ, without backoff processing for collisions, the two-level optimized scheme can solve the problems faced by EDCA, and the QoS requirements of multimedia transmissions in WLAN can be guaranteed.

B. Simulation Results

The performances of EDCA, E-PALD-DQMP without MRA, and E-PALD-DQMP with MRA are compared using computer simulations. Each protocol operates according to the network parameters shown in Table I. The wireless channel model is the same as that used in Section III-D, but PLR is set to 0.1. We assume that each station generates a single type of traffic. There are 12 voice, 12 video, and 12 data stations. The parameters of the voice and video models are summarized in Table II. The three traffic models are described in the list that follows.

1) Voice traffic model (CBR): The voice traffic is modeled as a two-state Markov process with talkspurt and silent-gap states. The durations of talkspurts and silent gaps follow exponential distributions, with the mean durations equal to 1 and 1.35 s, respectively.
2) Video traffic model (VBR): As shown in Fig. 11, video traffic is modeled as a multiple-state model, where a state generates a continuous bit stream for a certain holding duration [24]. The bit rate values of different states are obtained from a truncated exponential distribution with minimum and maximum bit rate values. The holding times of the states are assumed to be statistically independent and exponentially distributed. We assume that each state has the same mean holding time. In the simulations, the generated VBRs are quantized into 16 levels.
3) Data traffic model: In data traffic, the packet interarrival time is assumed to follow the exponential distribution, with a mean value of 8 ms. The packet size is set to 1000 bytes; thus, the data traffic generation rate is 1 Mb/s.

The performance metrics include throughput, delay, and dropping rate. Delay and dropping rate performances are only recorded for voice and video traffics. The corresponding simulation results will be analyzed later in this paper.

1) Throughput: The average throughput values of all users in the same AC$[i]$ $(i = 0, \ldots, 3)$ is

$$\text{Throughput}[i] = \frac{\text{SucTxNum}[i] \times \text{FrameSize}[i]}{\text{TxTime}[i] \times \text{UserNum}[i]} \quad (14)$$

where for AC$[i]$, SucTxNum$[i]$ is the total number of successfully transmitted frames, FrameSize$[i]$ is the frame size, TxTime$[i]$ is the transmission time, and UserNum$[i]$ is the number of users with the same services.

2) Delay: To calculate the delay at one user, we include the delays generated by both successful and dropped packets. The delay of a successful packet is calculated from the time when it is generated to the time when it is
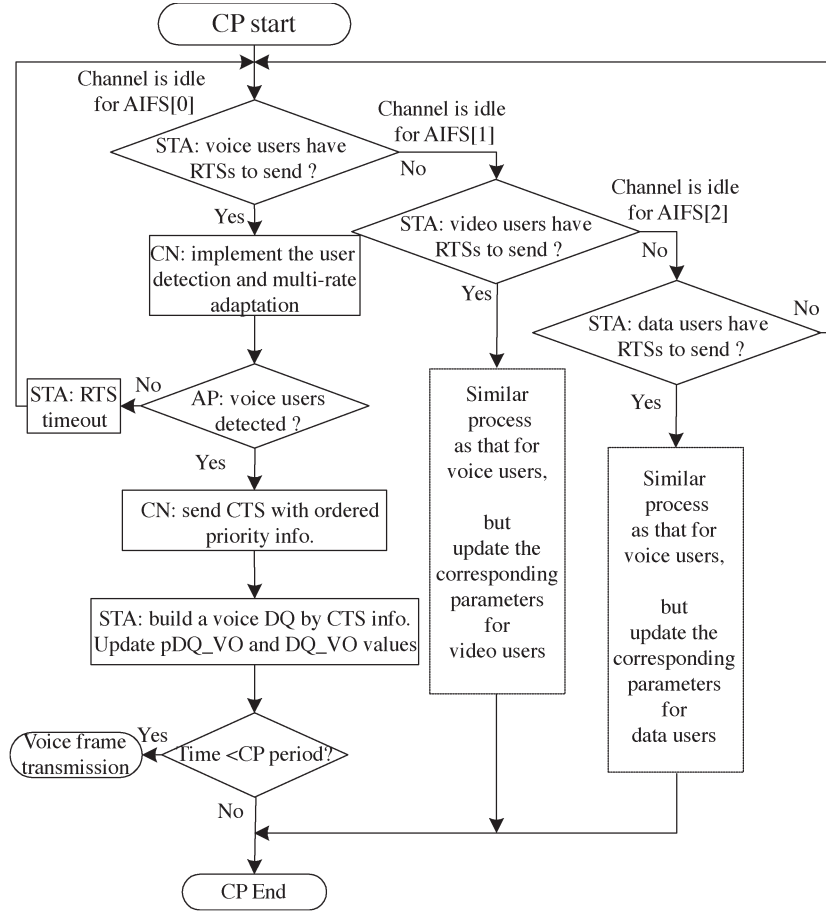
Fig. 10. Flowchart of E-PALD-DQMP.

TABLE II
VOICE AND VIDEO TRAFFIC MODELS

Voice Model	Values	Video Model	values
Talkspurt	1 s	Peak rate	420 Kbps
Silent gap	1.35 s	Minimum rate	120 Kbps
Data rate	64 Kbps	Average rate	240 Kbps
Delay bound	25 ms	Mean state time	160 ms
Packet size	200 Bytes	Delay bound	50 ms
		Packet size	800 Bytes

acknowledged. The delay of a dropped packet is regarded as the delay bound of its AC, i.e.,

$$\text{Delay}[i] = \frac{\sum \text{SucDelay}[i] + \sum \text{DropDelay}[i]}{\text{SucTxNum}[i] + \text{DropNum}[i]} \qquad (15)$$

where $\sum \text{SucDelay}[i]$ is the total delay accumulated by all successfully transmitted frames, whereas $\sum \text{DropDelay}[i]$ is the total delay accumulated by all dropped frames. $\text{SucTxNum}[i]$ and $\text{DropNum}[i]$ are the total numbers of the successfully transmitted frames and the dropped frames, respectively.

3) Dropping rate: It is the ratio of the number of dropped packets due to exceeding the delay bound to the total number of transmitted packets in the same AC, i.e.,

$$\text{Dropping\_rate}[i] = \frac{\text{DropNum}[i]}{\text{SucTxNum}[i] + \text{DropNum}[i]}. \qquad (16)$$
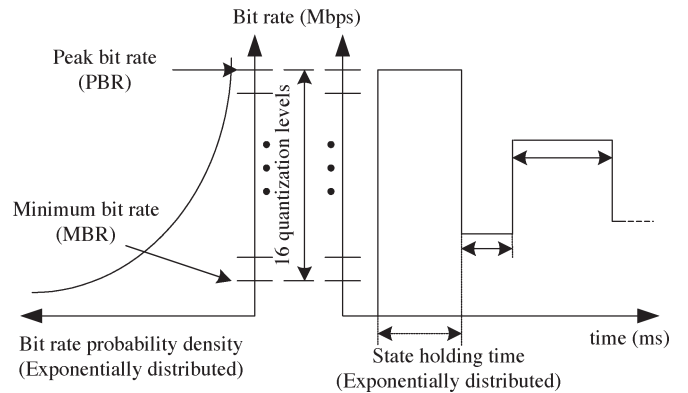

Fig. 11. VBR traffic model.

1) Throughput Comparisons: Fig. 12 shows the throughput comparisons. For voice and video traffics, with or without MRA, our proposed protocols demonstrate great throughput improvements. For data traffics, without MRA, there is no throughput improvement. However, with MRA, the throughput improvement is substantial.

2) Delay Comparisons: Average delay results for voice and video traffics are presented in Fig. 13. Consistent with the throughput results, for voice traffics, the delay performance of our proposed protocol has a slight improvement. For video, as compared with EDCA, both E-PALD-DQMP and
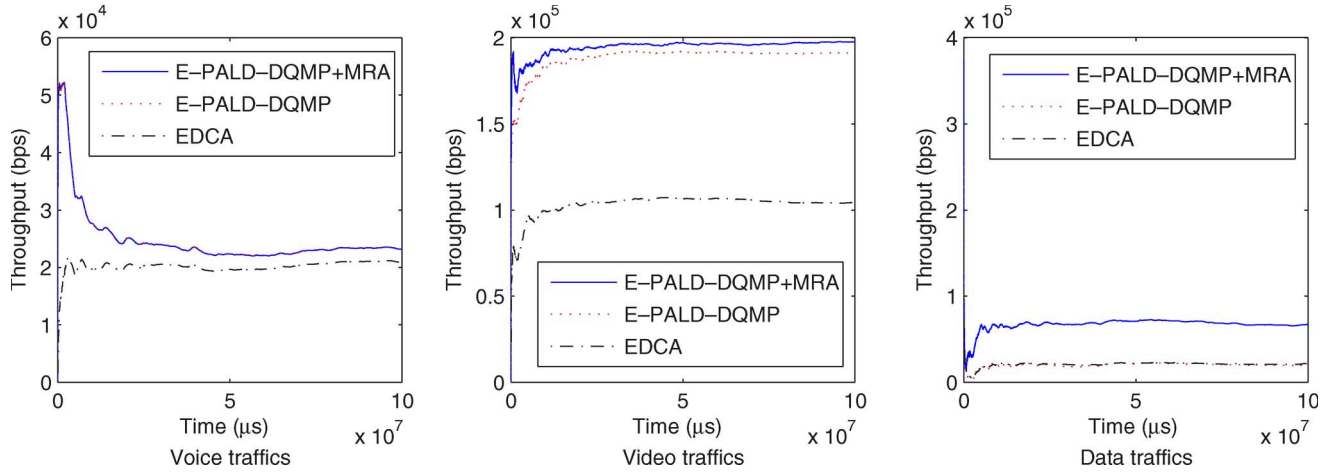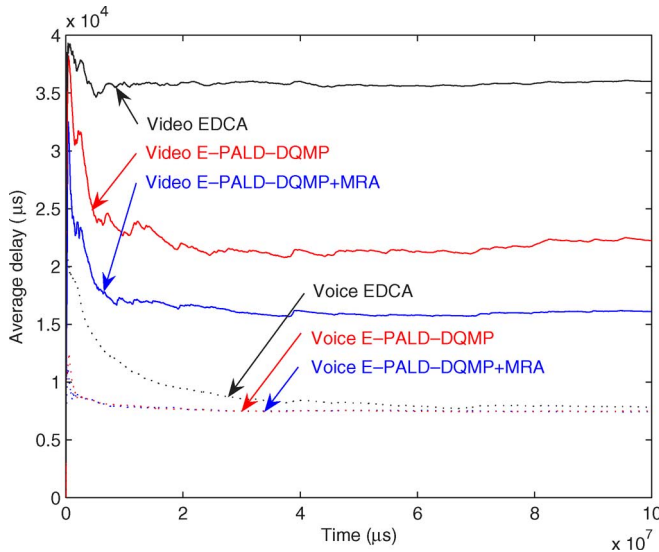
Fig. 12. Throughput comparisons.

Fig. 13. Delay comparisons.

Fig. 14. Dropping rate comparisons.

E-PALD-DQMP with MRA enjoy significantly lower transmission delays.

3) Dropping Rate Comparisons: According to our dropping rate definition, best effort data services are not delay sensitive; thus, no frames are dropped due to delay expiration. Fig. 14 shows the comparisons of voice and video services. For voice traffics, the dropping rate in EDCA is more than 0.1, whereas the corresponding values of our protocols are close to 0. For video traffics, the dropping rate in EDCA is up to 0.5, whereas E-PALD-DQMP with and without MRA have dropping rates less than 0.05.

V. PALD-MPMP

In contention-free access schemes, users in CFPs employ the polling strategy to provide time-bounded services. To combat the overhead caused by the polling frame, piggybacking is adopted in IEEE 802.11. However, this still cannot reduce bandwidth wastage caused by polling stations that have no packets to send. To improve the polling efficiency, several multipolling MAC protocols are proposed in [9]–[11]. The common
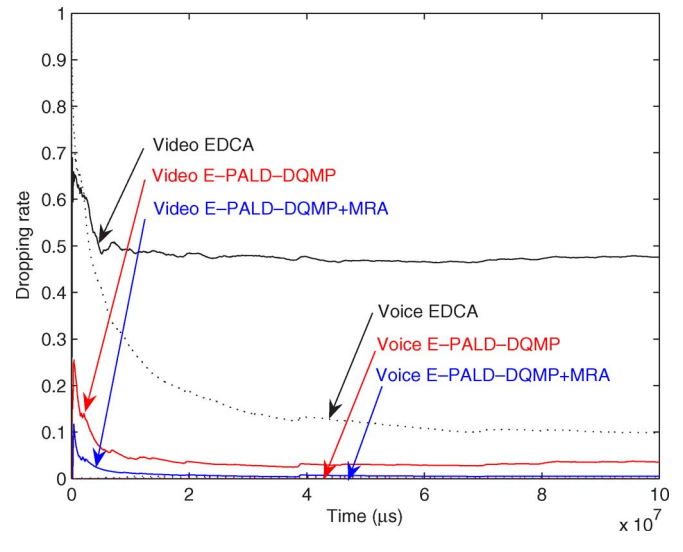
feature is that a rather complex status collection process is performed at the beginning of a CFP. These methods improve the polling efficiency; however, the status collection period causes new overhead, and the complexity is increased. Based on the PHY-assisted link differentiation introduced in Section II, a new multipolling MAC protocol, which is called PALD-MPMP, is presented in this section.

A. Protocol Description

Fig. 15 shows the timing diagram of PALD-MPMP. The CFP period consists of several subpolling periods (SPP), i.e.,

$$\sum_{i=1}^{N} \text{SPP}_i \leq \text{CFP} \tag{17}$$

where N is the maximum number that is supported in a CFP, and its value is dynamically changed for different CFPs. Each SPP is further divided into two subperiods: a station selecting period (SSP) and a data transmission period (DTP). To illustrate the implementation of PALD-MPMP, three issues will be discussed in this section: 1) management of the polling list;
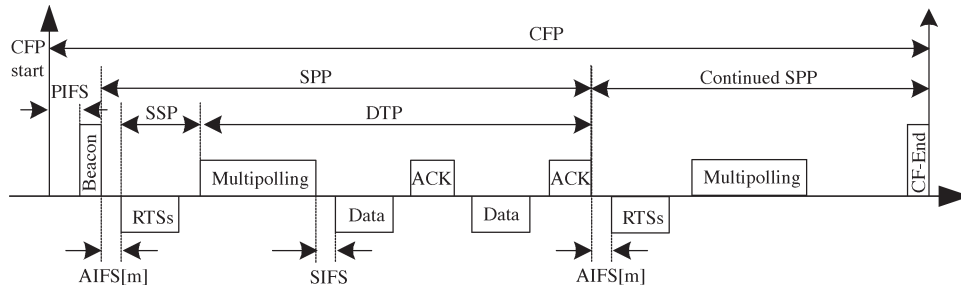
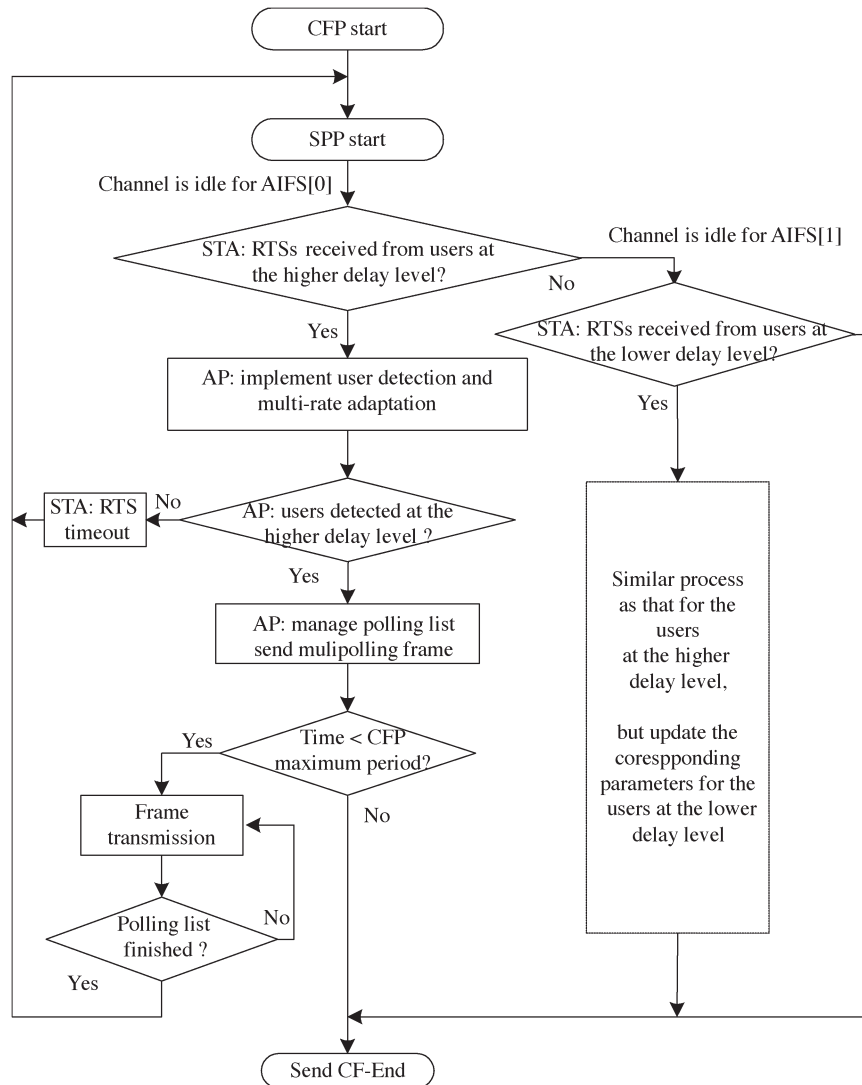Fig. 15.   Timing diagram of PALD-MPMP.

Fig. 16.   Flowchart of PALD-MPMP.

2) determination of the polling sequence; and 3) processing of the multipolling scheme. The first two functions are performed in SSP, whereas the third function is performed in DTP.

1) Management of the Polling List: As introduced in contention-based access control schemes, after the synchronized RTSs from users who have transmission requirements are received in APs/CNs, the PHY-assisted link differentiation can effectively obtain the active user ID, channel gain, and supported maximum data rate information. In contention-free access control schemes, such a mechanism will still work due

to its simplicity and small overhead for the polling process. Therefore, at the beginning of each SSP, synchronized RTSs for the users who have transmissions are also used to send the stations' information to APs. When considering different classes of services, such as voice, video, and best effort data traffics, E-PALD-DQMP uses a series of AIFS$[i]$ $(i = 0, \ldots, 3)$ to distinguish the different access priorities of these traffics. However, since the polling strategy is targeted at the QoS for time-bounded services, we differentiate the transmission priorities in delay levels, and these delay levels can be set by

the delay bounds of different services or other delay thresholds. In this paper, we assume that only the delay-sensitive users are pollable and simply choose the voice delay bound and the video delay bound to distinguish the users' access priorities. Among the users who have data to send, those users whose packet expiration time is shorter than the voice delay bound have a higher access priority than those users whose packet expiration time is between the voice and video delay bounds. To control the access of various delay level requirements, AIFS[m] $(m = 1, \ldots, M)$ is also used, where AIFS[m] has the same definition as in Section IV-A2, and M is the number of levels. It is noticed that these AIFSs do not overlap in time with the PIFS used at the initiation of a CFP. They are only used between the CFP beacon and CFP end period. Since only voice and video delay bounds are considered here, the values 0 and 1 are assigned to AIFSN[m] $(m = 1, 2)$. In each SPP, stations that send RTSs at the same delay level will be added into the polling list.

2) Determination of the Polling Sequence: After the synchronized RTSs are received at an AP, the PHY user detection and MRA modules can detect their channel gains and further deduce their supported data rates. In PALD-MPMP, the user detection module tries to effectively identify the users that have pending data and adequate channels for transmissions among all pollable users. With this detected information, at the same delay level, users who have better channel gains or higher supported data rates will have higher transmission priorities.

3) Multipolling Processing: Instead of sending back CTS and building a distributed logical queues among the active users to schedule their data transmissions in PALD-DQPM/E-PALD-DQMP, APs use a multipolling frame to control the wireless medium in PALD-MPMP. The format of the multipolling frame is similar to the CTS frame format, as shown in Fig. 3. However, the 2-byte "Duration" field is changed to the 6-byte related base station information, namely the "BISSID" field. In addition, the 6-byte RA information is mapped to the 2-byte "AID." The differences of the multipolling frame format with and without MRA support are the same, as shown in Section III-A1. After receiving this frame, polled stations will transmit their data with supported maximum data rates one by one. As used in Section III-A2, once the last ACK with "More Fragment" field in the MAC header equal to 1 is broadcasted, after a period of SIFS, a new SPP starts. Fig. 16 is the flowchart of PALD-MPMP.

B. Simulation Results

1) Simulation Results of PALD-MPMP: This section presents the simulation results of the standard PCF with piggybacked scheme and our proposed multipolling protocols. The simulation scenario and evaluation criteria used in this section are the same as those used in Section IV-B, but the performance tests are only performed for delay-sensitive voice and video traffics.

As shown in Fig. 17, for both CBR (voice) and VBR (video) traffics, the throughput performances of our proposed PALD-MPMP protocols demonstrate obvious improvements. Delay results for voice and video traffics are illustrated in Fig. 18. It is noted that for both CBR and VBR transmissions, delay
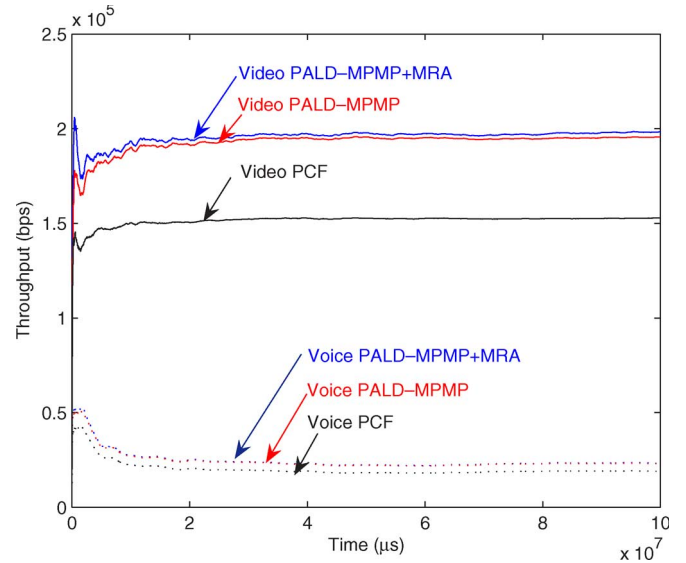

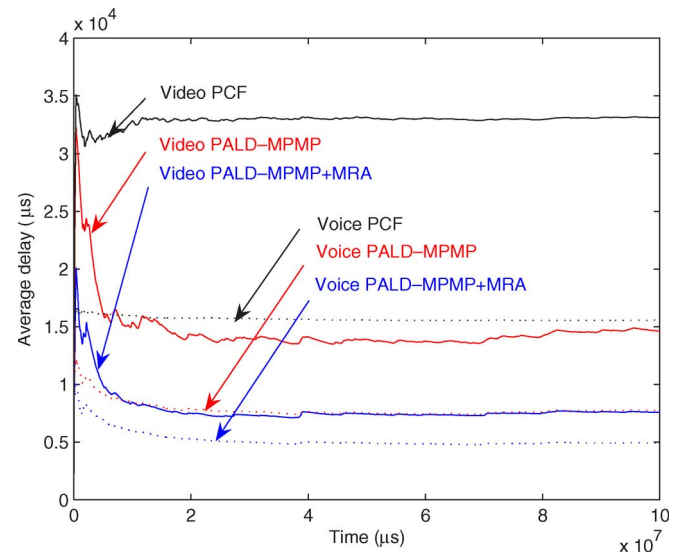Fig. 17.   Throughput comparisons.
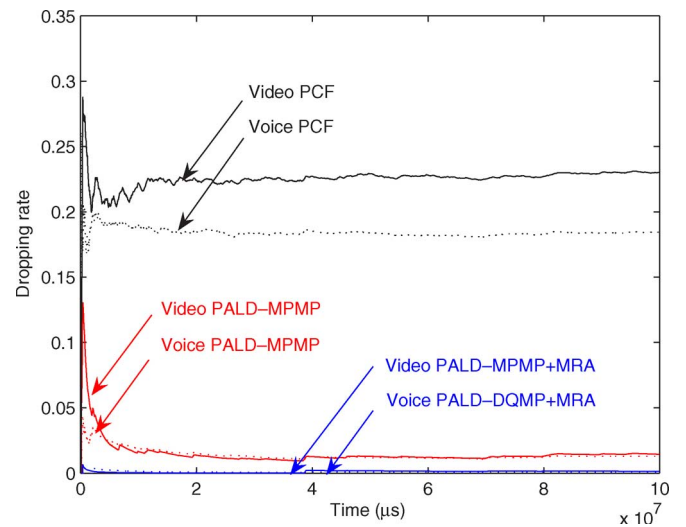

Fig. 18.   Delay comparisons.

Fig. 19.   Dropping rate comparisons.
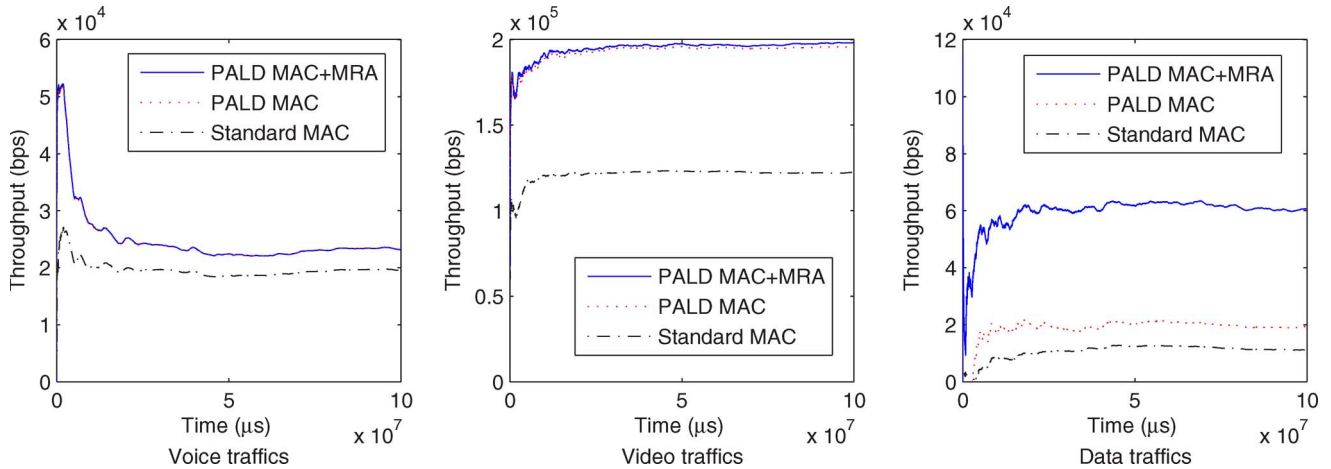
Fig. 20.   Throughput comparisons.
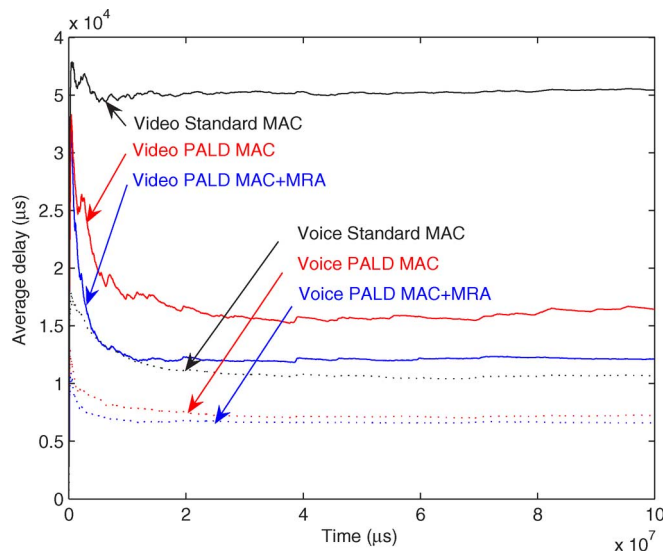

Fig. 21.   Delay comparisons.

Fig. 22.   Dropping rate comparisons.

can be effectively decreased with a priority classification in delay levels. With MRA, new improvements are obtained by reducing the frame transmission time with higher supported data rates. In Fig. 19, for both voice and video traffics, the dropping rate is around 0.2 in PCF. In PALD-MPMP without MRA, it is decreased to about 0.01. With MRA, the dropping rate of PALD-MPMP is nearly zero.

2) Simulation Results of the Combined MAC: This section gives the simulation results of the combined contention-based and contention-free MAC protocols. CFPR is set to be 100 ms. CFP and CP are 40 and 60 ms, respectively. Fig. 20 shows the throughput comparisons of the three combined protocols. It can be found that for both delay-sensitive and best effort traffics, our proposed PALD MAC protocols show distinct improvements. Fig. 21 gives the delay results. For E-PALD-DQMP, the delay improvement for voice is limited because the next-round voice transmissions will start when the current-round video and data transmissions are finished. However, with the help of PALD-MPMP, priorities in delay levels are considered. Therefore, both voice and video delay performances are improved. Fig. 22 shows the dropping rate comparison results. It is noted that with
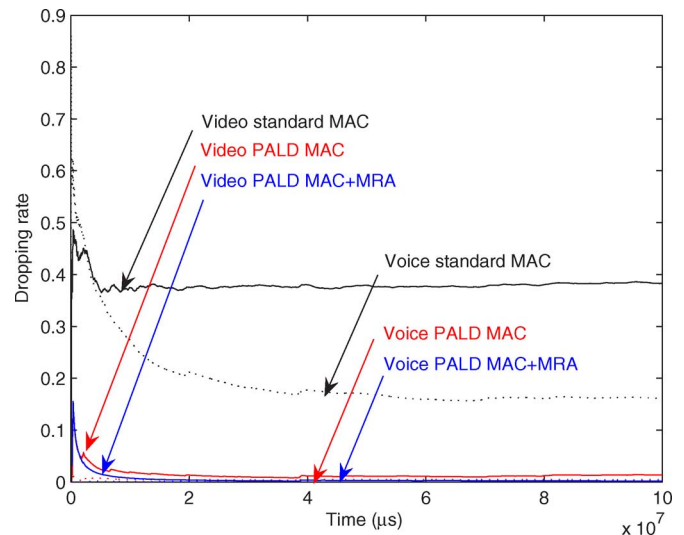
and without MRA, the dropping rates of our PALD MAC are always close to zero.

VI. CONCLUSION

In this paper, with the synchronization of RTS frames and the awareness of TAs in the MAC layer, user detection and MRA modules implemented in the PHY could detect detailed information, including the IDs of active users, their channel gains, and supported maximum data rates. Using such useful information, in PALD-DQMP, a logical distributed queuing system is built to fairly schedule the data transmission. In its enhanced version, i.e., E-PALD-DQMP, with the differentiation by AIFS among different ACs, prioritized services are guaranteed. For transmissions at the same service priority, users' frame transmissions are sorted by their detected channel gains. In these two proposed protocols, the bandwidth wastage and unfairness problems caused by the BEB algorithm are eliminated. Based on the same PHY-assisted mechanisms, PALD-MPMP is proposed with an effective multipolling scheme. It decreases the polling overhead and gives QoS guarantees for delay-sensitive services. By combining the proposed contention-based and

contention-free MAC protocols, simulation results show that a better performance can be obtained for both delay-sensitive and best effort services.

IEEE 802.11 WLAN systems are originally designed to support modest speeds and have already had many successful applications in both general wireless communication fields and vehicular communication systems. To support the wireless transmissions for users with high speeds, IEEE 802.16e Worldwide Interoperability for Microwave Access (WiMAX) is a promising technology. Compared with WLANs, WiMAX networks have different MAC schemes. The study of this topic is very interesting and will be considered in our future work.

ACKNOWLEDGMENT

REFERENCES

[1] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,* IEEE 802.11, Aug. 1999.
[2] B. Sardar, P. Chand, and D. Saha, "A novel version of wireless TCP for vehicular ObIPNs," in *Proc. IEEE 63rd Veh. Technol. Conf.*, Melbourne, Australia, May 2006, vol. 2, pp. 876–880.
[3] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS),* IEEE 802.11e/D13.0, Jan. 2005. draft supp. to IEEE Std. 802.11.
[4] W. Kuo and C. C. J. Kuo, "Enhanced backoff scheme in CSMA/CA for IEEE 802.11," in *Proc. IEEE 58th Veh. Technol. Conf.*, Orlando, FL, Oct. 2003, vol. 5, pp. 2809–2813.
[5] Q. Ni, "Performance analysis and enhancements for IEEE 802.11e wireless networks," *IEEE Netw.*, vol. 19, no. 4, pp. 21–27, Jul./Aug. 2005.
[6] Q. Pang, S. C. Liew, and V. C. M. Leung, "Design of an effective loss-distinguishable MAC protocol for 802.11 WLAN," *IEEE Commun. Lett.*, vol. 9, no. 9, pp. 781–783, Sep. 2005.
[7] L. Alonso, R. Agusti, and O. Sallent, "A near-optimum MAC protocol based on the distributed queuing random access protocol for a CDMA mobile communication system," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 9, pp. 1701–1718, Sep. 2000.
[8] L. Alonso, R. Agusti, and O. Sallent, "WLAN throughput improvement via distributed queuing MAC," *IEEE Commun. Lett.*, vol. 9, no. 4, pp. 310–312, Apr. 2005.
[9] B. Kim, S. W. Kim, Y. Fang, and T. F. Wong, "Two-step multipolling MAC protocol for wireless LANs," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 6, pp. 1276–1286, Jun. 2005.
[10] S. C. Lo, G. Lee, and W. T. Chen, "An efficient multipolling mechanism for IEEE 802.11 wireless LANs," *IEEE Trans. Comput.*, vol. 52, no. 6, pp. 764–778, Jun. 2003.
[11] Z. T. Chou and S. L. Wu, "A new QoS point coordination function for multimedia wireless LANs," in *Proc. Int. Conf. Distrib. Comput. Syst.*, Tokyo, Japan, Mar. 2004, pp. 40–47.
[12] C. Comaniciu, M. B. Mandayam, and H. V. Poor, *Wireless Networks: Multiuser Detection in Cross-Layer Design*. New York: Springer-Verlag, May 2005.
[13] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.
[14] Z. Xu, "Blind identification of co-existing synchronous and asynchronous users for CDMA system," *IEEE Signal Process. Lett.*, vol. 8, no. 7, pp. 212–214, Jul. 2001.
[15] N. Morinaga, M. Nakagawa, and R. Kohno, "New concepts and technologies for achieving highly reliable and high-capacity multimedia wireless communications systems," *IEEE Commun. Mag.*, vol. 37, no. 1, pp. 34–40, Jan. 1997.
[16] J. D. Prodo and S. Choi, "Link adaptation strategy for IEEE 802.11 WLAN via received signal strength measurement," in *Proc. IEEE Int. Conf. Commun.*, Anchorage, AK, May 2003, pp. 1108–1113.
[17] S. E. Bensley and B. Aazhang, "Subspace-based channel estimation for code division multiple access communication systems," *IEEE Trans. Commun.*, vol. 44, no. 8, pp. 1009–1020, Aug. 1996.
[18] M. Torlak and G. Xu, "Blind multiuser channel estimation in asynchronous CDMA systems," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 137–147, Jan. 1997.
[19] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," Digital Equipment Corp., Maynard, MA, DEC Res., Tech. Rep. TR-301, 1984.
[20] Y. Xiao and H. Z. Li, "Voice and video transmissions with global data parameter control for the IEEE 802.11e enhance distributed channel access," *IEEE Trans. Parallel Distrib. Syst.*, vol. 15, no. 11, pp. 1041–1053, Nov. 2004.
[21] L. Romdhani, Q. Ni, and T. Turletti, "Adaptive EDCF: Enhanced service differentiation for IEEE 802.11 wireless adhoc networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, New Orleans, LA, Mar. 2003, vol. 2, pp. 1373–1378.
[22] G. W. Wong and R. W. Donaldson, "Improving the QoS performance of EDCF in IEEE 802.11e wireless LANs," in *Proc. IEEE Pacific Rim Conf. Commun., Comput. Signal Process.*, Aug. 2003, vol. 1, pp. 392–396.
[23] *IEEE 802.11 TGe—Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: MAC Enhancements for Quality of Service*, IEEE 802.11e/D8.0, Mar. 2004.
[24] S. T. Sheu and T. F. Sheu, "A bandwidth allocation/sharing/extension protocol for multimedia over IEEE 802.11 *ad hoc* wireless LANs," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 10, pp. 2065–2080, Oct. 2001.

Fanglei Sun (S'06) received the B.S. and M.S. degrees from Xidian University, Xi'an, China, in 2001 and 2004, respectively. She is currently working toward the Ph.D. degree in the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong.

She was a Research Assistant from December 2002 to December 2003 with the Department of Electrical and Electronic Engineering, The University of Hong Kong. From November 2005 to August 2006, she was a Visiting Researcher with the Department of Electrical Engineering, University of California, Riverside. Her research interests include wireless protocol design, cross-layer algorithm design over WLAN/WiMAX networks, and OFDM/MIMO systems.

Victor O. K. Li (S'80–M'81–SM'86–F'92) was born in Hong Kong in 1954. He received the S.B., S.M., E.E., and Sc.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1977, 1979, 1980, and 1981, respectively.

In February 1981, he joined the University of Southern California (USC), Los Angeles, where he became a Professor of electrical engineering and the Director of the Communication Sciences Institute. Since September 1997, he has been with The University of Hong Kong, Hong Kong, where he is an Associate Dean (Research) of Engineering and a Chair Professor of information engineering. From September 1997 to June 2004, he also served as the Managing Director of Versitech Ltd., which is the technology transfer and commercial arm of The University of Hong Kong. He serves on various corporate boards. His research interest is in information technology, including all-optical networks, wireless networks, and Internet technologies and applications. Sought by government, industry, and academic organizations, he has extensively lectured and consulted around the world. He is very active in the research community and has chaired various international conferences and served on the editorial boards of various international journals. He has given distinguished lectures at various universities around the world, as well as keynote speeches at many international conferences.

Prof. Li has received numerous awards, including, most recently, the PRC Ministry of Education Changjiang Chair Professorship at Tsinghua University, Beijing, China, the U.K. Royal Academy of Engineering Senior Visiting Fellowship in Communications, the K. C. Wong Education Foundation Lectureship, the Croucher Foundation Senior Research Fellowship, and the Order of the Bronze Bauhinia Star, Government of the Hong Kong Special Administrative Region, China.

Zhifeng Diao received the B.S. and M.S. degrees from Xidian University, Xi'an, China, in 1999 and 2002, respectively, and the Ph.D. degree from The University of Hong Kong, Hong Kong, in 2005.

From September 2005 to September 2006, he was a Postdoctoral Fellow with the Department of Electrical Engineering, University of California, Riverside. He is currently a Postdoctoral Researcher with the Department of Electrical Engineering, Arizona State University, Tempe. His research interests include wireless scheduling algorithm design, smart antennas, cross-layer design, OFDM, MC-CDMA, and Wimax.
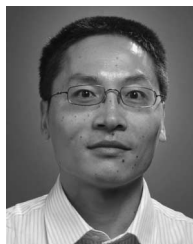
Zhengyuan Xu (S'97–M'99–SM'02) received the B.E. and M.E. degrees in electronic engineering from Tsinghua University, Beijing, China, in 1989 and 1991, respectively, and the Ph.D. degree in electrical engineering from Stevens Institute of Technology, Hoboken, NJ, in 1999.

From 1991 to 1996, he was a System Engineer and the Department Manager with the Tsinghua Unisplendour Group Corporation, Tsinghua University. Since 1999, he has been a Faculty Member with the Department of Electrical Engineering, University of California, Riverside, where he is currently an Associate Professor with tenure. He has held visiting positions at Stanford University, Stanford, CA, and the University of Science and Technology of China, Hefei, China. His research interests are in wireless communications and related signal processing, including multiuser spread spectrum, impulse radio, ultrawideband, wireless optics, hybrid radio frequency and optical communication systems, and sensor and *ad hoc* networks.

Dr. Xu is an Elected Member of the IEEE Signal Processing Society's Technical Committee on Signal Processing for Communications. He has served as a Session Chair and a Technical Program Committee Member for many international conferences. He is an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and a Guest Editor of the Special Issue on "Performance Limits of Ultra-Wideband Systems" of the IEEE JOURNAL OF SELECTED TOPICS ON SIGNAL PROCESSING. He also served as an Associate Editor of the IEEE COMMUNICATIONS LETTERS during 2002–2005 and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY during 2002–2006. He was a recipient of the Outstanding Student Award and the Motorola Scholarship from Tsinghua University, the Peskin Award from Stevens Institute of Technology, and the Academic Senate Research Award and the Regents' Faculty Award from the University of California, Riverside.