

Contention-based Medium Access Control with Physical Layer Assisted Link Differentiation

Fanglei Sun, Victor O. K. Li

Department of Electrical and Electronic Engineering
University of Hong Kong
Hong Kong, China
Email: {flsun, vli}@eee.hku.hk

Zhifeng Diao, Zhengyuan Xu

Department of Electrical Engineering
University of California
Riverside, CA 92521 USA
Email: {zdiao, dxu}@ee.ucr.edu

Abstract—In this paper, we develop contention-based medium access control (MAC) schemes for both best-effort data transmissions and delay-sensitive multimedia transmissions over WLANs. A user detection module and a multi-rate adaptation module are proposed in the physical layer to assist in link differentiation. With these two modules, for best-effort data transmissions, 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 the transmissions. To support delay-sensitive 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. 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 contributed by the availability of cross-layer channel state information, and the consequent multi-rate adaptation scheme.

I. INTRODUCTION

For best-effort data transmissions, in standard IEEE 802.11 MAC [1], the binary exponential backoff (BEB) algorithm embedded in the distributed coordination function (DCF) suffers from a fairness problem and yields low throughput under heavy load [2]. This is because it assumes that all losses are due to collisions. To support multimedia transmissions over the wireless medium, enhanced distributed channel access (EDCA) is being developed in IEEE 802.11e [3]. EDCA works with four access categories (ACs), which are virtual DCFs, and each AC achieves differentiated channel access. However, EDCA also suffers from bandwidth wastage caused by collisions and backoff. The problem is exacerbated in EDCA since collisions are caused not only among different mobile stations, but also among different ACs in the same station [4].

For DCF, [5] proposed a method to distinguish collisions and link errors in the MAC layer. In our paper, an effective user detection module and a multi-rate adaptation module in the physical layer are designed to feed back the error type information, detailed channel gains and the corresponding supported maximum transmission rate. In [6], a MAC protocol based on the Distributed Queuing Random Access Protocol

(DQRAP) is proposed for WLANs. This mechanism changes the data frame structure and the contention mechanism. In addition, the reserved m contention periods in each frame cause additional transmission overhead. However, the queuing method inspires us to develop a new transmission scheduling scheme. In this paper, we propose PALD-DQMP, which has the following features: 1) the user detection module can provide effective loss differentiation information and users' channel state information; 2) the multi-rate adaption module can dynamically choose the maximum transmission rate supported by the current channel condition; 3) combined with the above cross-layer information, PALD-DQMP utilizes a distributed queuing system to handle the random access and data transmissions, thus eliminating the bandwidth wastage caused by the BEB algorithm; 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 that: 1) with the help of service priority differentiation and physical layer link differentiation, the bandwidth wastage caused by BEB is eliminated; 2) the QoS requirements of delay-sensitive services are guaranteed; 3) the throughput and fairness problem in best-effort data services can be resolved.

II. USER DETECTION AND MULTI-RATE ADAPTATION MODULES

A. User Detection Module

In contention-based access schemes, we assume that any node can be chosen as a coordinating node (CN). Even with this CN, the independent basic service set (BSS) 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 in order to broadcast its information to other nodes. According to the statement of logical service interfaces in [1], the necessary information can be exchanged between the AP/CN and other nodes by association-related services at network initialization. Among such information, we are most interested in the transmission addresses (TAs) of the active nodes, and such information is easily gathered by APs/CNs.

Multiuser detection is an effective technique in CDMA systems. It can successfully separate users at the receiver based

This research is supported in part by the Research Grants Council of the Hong Kong Special Administrative Region, China, under Grant No. HKU 7152/05E.

on their unique spreading codes [7]. In WLANs, there exists similar user dependent information - transmitter address (TA), which is indicated in the RTS frame. In this paper, 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. In this section, a synchronous user detection technique for WLANs will be introduced. The synchronization of RTSs will be presented at the end of Section III-A.2. If RTSs can not be transmitted synchronously, there are some asynchronous multiuser detection techniques used in CDMA systems, which can also be adopted in the physical layer link differentiation.

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. A transmission address is an N -bit ($N = 48$) vector, so the list of TAs can be recorded as follows:

$$\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_k, \dots, \mathbf{c}_K],$$

where $\mathbf{c}_k = [c_{k,1}, \dots, c_{k,n}, \dots, c_{k,N}]^T$. 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 can be re-assigned to users as TAs in BSSs. For K users, the vector \mathbf{h} represents the channel gains experienced in their respective wireless channels:

$$\mathbf{h} = [h_1, \dots, h_k, \dots, h_K]^T.$$

At time t , we assume that l ($0 \leq l \leq K$) users have data to transmit and send RTSs simultaneously 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

$$\mathbf{r} = \mathbf{C}\mathbf{h} + \mathbf{n}, \quad (1)$$

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:

$$\hat{\mathbf{h}} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{r} = \mathbf{\Pi} \mathbf{r}, \quad (2)$$

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) Collision state: the active user attends the contention and obtains a good channel to transmit its pending frame.

In order to implement the above 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 above three cases. With noise present, the value of $|\hat{h}_k|^2$ could be nonzero. So, it is required

to set a few thresholds for different users to distinguish the first case from the other two cases. By substituting (1) in (2), we obtain:

$$\hat{\mathbf{h}} = \mathbf{h} + \mathbf{\Pi} \mathbf{n}. \quad (3)$$

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 \mathbf{n} is an additive white Gaussian noise vector whose elements have zero mean and variance σ^2 . Denote the calculated inverse matrix by $\mathbf{\Pi} = [a_{i,j}]_{K \times N}$, then we obtain the silence threshold:

$$\begin{aligned} H_{silence-thresh} &\geq E\{|\mathbf{\Pi} \mathbf{n}|^2\} \\ &= \sigma^2 \times \left[\sum_{j=1}^N (a_{1,j})^2, \dots, \right. \\ &\quad \left. \sum_{j=1}^N (a_{k,j})^2, \dots, \sum_{j=1}^N (a_{K,j})^2 \right]_{1 \times K}^T, \end{aligned}$$

where $|\cdot|$ is an operator to take the absolute value of each element in a vector. We denote the second threshold as $H_{collision-thresh}$ 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, given the packet loss rate (PLR) and transmission power constrains, we can deduce the required channel gain power. In practice, the value of $H_{silence-thresh}$ will be chosen between the above lower bound and $H_{collision-thresh}$. For a particular network scenario, we can obtain an effective value which can minimize the error detection through simulations.

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 longer orthogonal addresses for them, but this will cause more overhead. Alternatively, we can distribute additional users to other APs/CNs.

B. Multi-rate Adaptation (MRA) Module

By the concept of adaptive modulation [8], based on the above user detection module, we assume that the noise power is detected at silent periods before the start of data transmissions, so 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:

$$SNR_k = \frac{|\hat{h}_k|^2}{\hat{\sigma}^2} \quad (k = 1, 2, \dots, K).$$

In this paper, we are mainly concerned with the 802.11b WLAN which provides 4 types of supported rates. Without multi-rate adaptation (MRA), we assume the channel state is either good or bad, differentiated by the $H_{collision-thresh}$. With MRA, a set of SNR thresholds should be defined in order to select the appropriate data rates in the physical layer (PHY). In the simulations, these thresholds have been selected based on the results presented in [9].

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

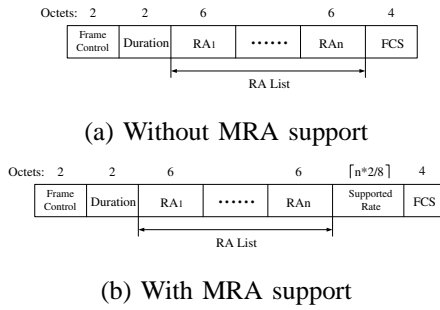


Fig. 1. Frame format of CTS.

III. PALD-DQMP

A. Protocol Description

A logical distributed queue is constructed in PALD-DQMP. This queue is simply represented by two integer numbers at each node, 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, while pDQ has a specific value for each node. Both values are initially set to zero and will be updated by broadcasted CTS/ACK frames from the AP/CN. We separate this queuing algorithm 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 data to transmit, their RTSs are easily transmitted synchronously after a period of DIFS.

(2) After the receiver obtains the co-existing signals, the physical layer 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 the CTS plans to feed back. In order to demonstrate the impact of the multi-rate adaptation module on the performance, our PALD-DQMP broadcasts two CTS frame formats, which are different in the content of the RA list and supported rate field, as shown in Fig. 1.

i) Without MRA, the RA list only includes the MAC addresses of the users whose channels are in collision states. PALD-DQMP sorts the RA list by the priority of the 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 the frame transmission

phase will start after a period of SIFS.

ii) Let us consider users with both collision 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 not only include 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, 2 bits are assigned for each user to indicate the currently supported data rate. Assuming the current detected active user is n , the rate information in a CTS frame needs $\lceil n * 2/8 \rceil$ bytes. The updating process of pDQ in mobile users is the same as that described in i).

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

2) *Data Transmission Phase:* At the end of the channel contention phase, each node gets the updated DQ, pDQ values and their corresponding supported maximum data rate. After a period of SIFS, data transmissions will be scheduled by the following 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 which causes the values of DQ, pDQ, and NAV in all mobile nodes to change to zero is received, a new contention phase will be triggered. In the frame control field of MAC headers, there is a one-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 users whether this ACK is an acknowledgment to the last scheduled frame. If it is, after a period of DIFS, nodes which have RTS requirements will contend the channel again, and the synchronization of RTSs can be solved. Then the PALD-DQMP repeats from the channel contention phase. Fig. 2 shows the flow chart of PALD-DQMP.

Thus, without collisions, the throughput bound of DCF will be reached when there is only a pair of mobile nodes in a BSS. For PALD-DQMP, the transmission overhead of RTS/CTS frames is shared by all active users, so the larger the number of active users, the higher the theoretical throughput bound.

B. 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. We simulate WLANs where K nodes contend to transmit with an AP/CN. Each node is assumed

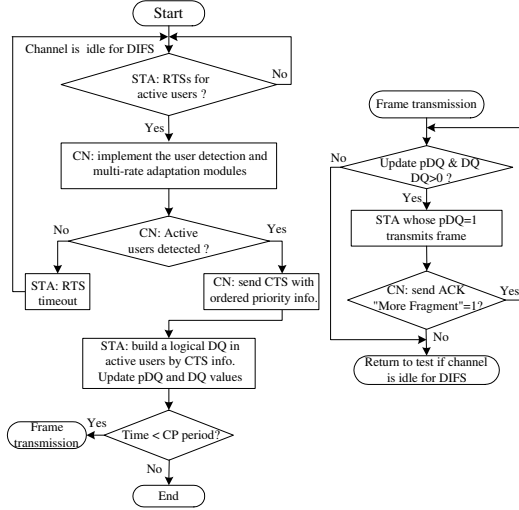


Fig. 2. Flow chart of PALD-DQMP.

TABLE I
NOTATIONS AND PARAMETER VALUES

Notations	Descriptions	Values
MAC	MAC header overhead	224 bits
DIFS	The period of DIFS	50 μs
SIFS	The period of SIFS	10 μs
PIFS	The period of PIFS	30 μs
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 μs
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 μs

to always have data to send. Simulation parameter values are shown in Table I. In the physical layer, a Rayleigh fading model with channel gain power equal to 1 is chosen. Channel noise is assumed to be white Gaussian noise model with average SNR equal to 20 dB. In addition, we use $H_{collision-thresh}$ to control the network packet loss rate (PLR). Under perfect channel (PLR=0) and lossy channel (PLR=0.1) conditions, we test the network performance when the number of nodes changes from 1 to 40. All simulations in this paper are done in 100s.

1) *Throughput Comparisons*: As shown in Fig. 3, the throughputs of the protocols are similar to each other when the node number is 1. When the number of the mobile nodes increases, the performance of the standard protocol gets slightly worse, but the performance of PALD-DQMP improves. With MRA, our protocol has another distinct improvement.

2) *Fairness Comparisons*: The fairness index f is defined as $f = \frac{(\sum_{i=1}^K x_i)^2}{K(\sum_{i=1}^K x_i^2)}$ [10], where K is the number of active users in the network and x_k is the throughput achieved by user k ,

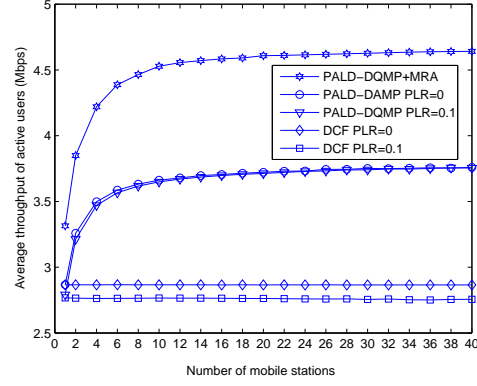


Fig. 3. Throughput Comparisons.

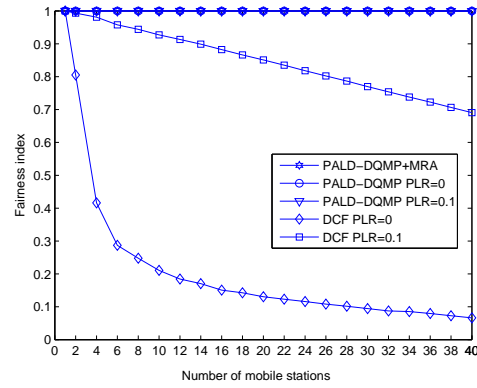


Fig. 4. Fairness Comparisons.

$1 \leq k \leq K$. Fig. 4 shows the fairness index comparison. It can be found that our proposed PALD-DQMP always keeps the fairness index nearly close to 1. For DCF, the fairness index under perfect channels is much worse compared to that with a PLR of 0.1. This is because under perfect channels, a successful user may easily hold on to the channel without any backoffs.

IV. E-PALD-DQMP

With DCF, all stations compete for the channel with the same priority. To support QoS requirements for multimedia services, EDCA is being developed, with four access categories (ACs) to implement differentiated channel access priorities. However, EDCA is just an extension of DCF, and BEB 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

E-PALD-DQMP also works with four ACs. In traditional IEEE 802.11e MAC, EDCA is executed in each backoff entity with different parameter values for the EDCA parameter set, whereas, in E-PALD-DQMP, each frame unit is regarded as

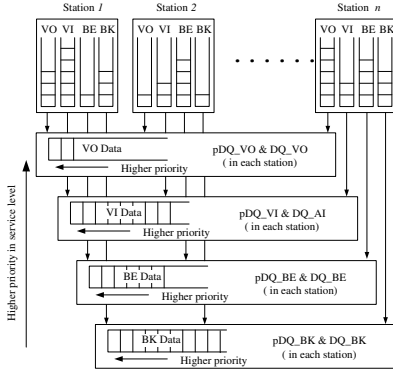


Fig. 5. Queuing system of E-PALD-DQMP.

the basic contention entity. Their different access priorities are achieved by the following two-level contention-based mechanism.

1) *At the Service Priority Level:* Each type of service also starts counting down the timer after detecting the medium being idle for a duration defined by the arbitration interframe space (AIFS[i], $i=0, \dots, 3$), instead of DIFS, which is used in PALD-DQMP. The AIFS[i] can be enlarged per AC with the help of the arbitration interframe space number (AIFSN[i], $i=0, \dots, 3$). The AIFSN[i] defines the duration of AIFS[AC] also according to:

$$AIFS[i] = SIFS + AIFSN[i] * aSlotTime,$$

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 in different users, priority in service level is guaranteed.

2) *At the Frame Priority Level:* Since E-PALD-DQMP also works with four ACs, four logical distributed queues are constructed in E-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_VO, DQ_VI, DQ_BE, and DQ_BK represent the total frame numbers of these four queues, with the same value for each node, whereas pDQ_VO, pDQ_VI, pDQ_BE, and pDQ_BK record the position of a given node in the corresponding queue. These four variables range from 0 to DQ_VO, DQ_VI, 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 variables are initially set to zero and updated by broadcasting CTS/ACK frames from APs/CNs. Fig. 5 illustrates the queuing system of E-PALD-DQMP. ACs with the same priority in different nodes work as nodes with PALD-DQMP. They have the same AIFS[i] values. Fig. 6 is the flow chart of E-PALD-DQMP.

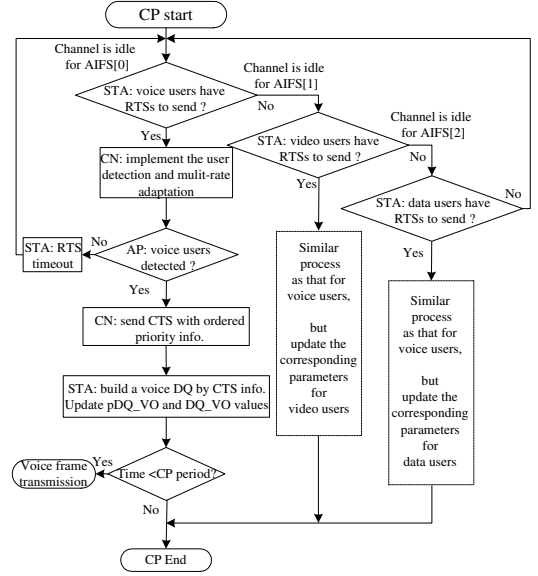


Fig. 6. Flow chart of E-PALD-DQMP.

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

B. Simulation Results

The simulation scenario is the same as that used in Section III-B, but PLR is set to 0.1. Each station generates a single type of traffic. There are 12 voice, 12 video, and 12 data stations. The parameters of voice (CBR) and video (VBR) models are summarized in Table II. For data traffic, the packet inter-arrival time is assumed to follow the exponential distribution and the data generation rate is 1 Mbps. Delay and dropping rate performances are only recorded for voice and video traffics.

Fig. 7 shows the throughput comparison. For voice and video traffics, with or without MRA, our proposed protocols have great throughput improvements. For data traffics, without MRA, constrained by the remaining bandwidth resource, our throughput has a slight improvement. However, with MRA, the throughput performance has distinct improvement. The average delay results for voice and video traffics are presented in Fig. 8. For voice traffics, the delay performance of our proposed protocols has a slight improvement. While, for video, compared with EDCA, both E-PALD-DQMP and E-PALD-DQMP with MRA decrease the transmission delay greatly. In Fig. 9, for voice traffics, the dropping rate in EDCA is over 0.1, while the corresponding values of our protocols are close to zero. For video traffic, the dropping rate in EDCA

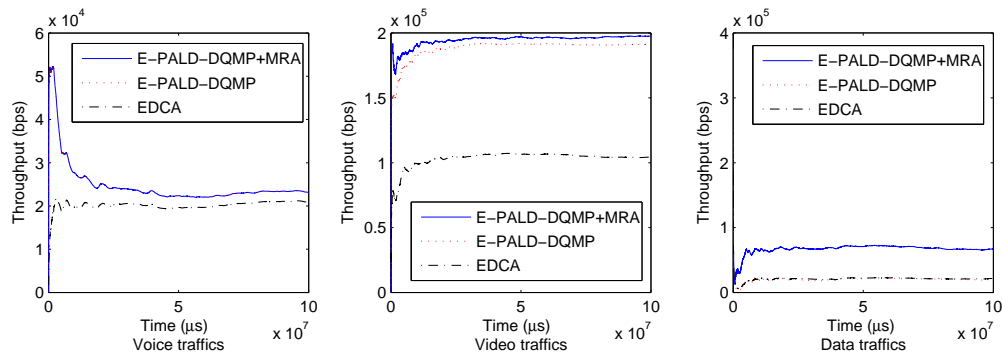


Fig. 7. Throughput comparisons.

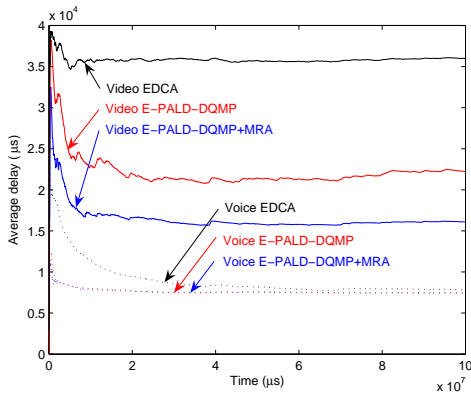


Fig. 8. Delay comparisons.

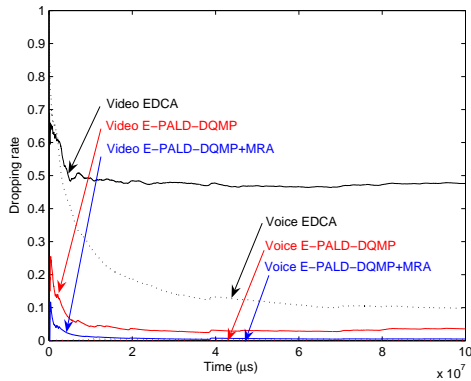


Fig. 9. Dropping rate comparisons.

is around 0.5, while E-PALD-DQMP with and without MRA have dropping rates under 0.05.

V. CONCLUSIONS

In this paper, with the synchronization of RTS frames and the awareness of TAs in the MAC layer, user detection and multi-rate adaptation modules implemented at the physical layer could detect detailed information including the IDs of

active users, their channel conditions and supported maximum data rates. Using such information, in PALD-DQMP, a logical distributed queueing system is built to schedule the best-effort data transmissions fairly. In its enhanced version E-PALD-DQMP, the two-level optimized MAC scheme can effectively support the QoS for multimedia transmissions. Finally, the simulation results show that our proposed contention-based MAC protocols outperform the standard MAC protocols for both delay-sensitive and best-effort traffics.

REFERENCES

- [1] IEEE 802.11, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," Aug. 1999.
- [2] W. Kuo and C. C. J. Kuo, "Enhanced backoff scheme in CSMA/CA for IEEE 802.11," *IEEE 58th VTC*, vol. 5, Oct. 2003.
- [3] IEEE 802.11e/D13.0, Part 11, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: medium access control (MAC) enhancements for quality of service (QoS)," draft supp. to IEEE 802.11 std., Jan. 2005.
- [4] Q. Ni, "Performance analysis and enhancements for IEEE 802.11e wireless networks," *IEEE Network*, vol. 19, no. 4, pp 21-27, July-Aug. 2005.
- [5] Q. Pang, S. C. Liew, V. C. M. Leung, "Design of an effective Loss-Distinguishable MAC Protocol for 802.11 WLAN," *IEEE Communications Letters*, vol. 9, no. 9, Sep. 2005.
- [6] L. Alonso, R. Agusti and O. Sallent, "WLAN throughput improvement via distributed queueing MAC," *IEEE Communication Letters*, vol. 9, no. 4, Apr. 2005.
- [7] Z. Xu, "Blind identification of co-existing synchronous and asynchronous users for CDMA system," *IEEE Signal Processing Letters*, vol. 8, no. 7, pp. 212-214, Jul. 2001.
- [8] 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.
- [9] J. D. Prodo and S. Choi, "Link adaptation strategy for IEEE 802.11 WLAN via received signal strength measurement," *IEEE ICC03*, pp. 1108-1113, May 2003.
- [10] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," DEC Research, Tec. Report, TR-301, 1984.
- [11] S. T. Sheu and T. F. Sheu, "A bandwidth allocation/sharing/extension protocol for multimedia over IEEE 802.11 Ad Hoc Wireless LANs," *IEEE JSAC*, vol. 19, no. 10, pp. 2065-2080, Oct. 2001.