A CSMA/CA MAC Protocol for Multi-User MIMO Wireless LANs

Michelle X. Gong, Eldad Perahia, Robert Stacey, Roy Want
Intel Corporation
Santa Clara, CA 95054-1549
Email: {michelle.x.gong, eldad.perahia, robert.j.stacey, roy.want}@intel.com

Shiwen Mao
Dept. ECE, Auburn University
Auburn, AL 36849-5201
Email: smao@ieee.org

***Abstract*—Multiple-input multiple-output (MIMO) is one form of the smart antenna technology that uses multiple antennas at both the transmitter and receiver to improve communication performance. In this paper, we investigate the problem of medium access control in wireless local area networks (WLANs) with downlink multi-user MIMO (DL MU MIMO) capability. We propose a CSMA/CA MAC protocol with three response mechanisms for DL MU MIMO and compare the performance of DL MU MIMO with the beam-forming (BF) based approach. A novel per-station weighted queuing mechanism is proposed to achieve fairness in the network. Performance analysis and simulation study both show that the proposed DL MU MIMO mechanism incurs low overhead and provides significant throughput performance gain over BF based approach in high SNR scenarios.**

I. Introduction

Multiple-input multiple-output (MIMO) is one form of the smart antenna technology that uses multiple antennas at both the transmitter and receiver to improve communication performance. MIMO communications have been extensively studied for next generation cellular networks and have been adopted for wireless local area networks (WLANs) as specified in the IEEE 802.11n standard [1].

A MIMO system takes advantage of two types of gains, namely, spatial diversity gain and spatial multiplexing gain [2]. Spatial diversity can combat severe fading and improve the reliability of the wireless link by duplicating information across multiple antennas. Spatial multiplexing takes advantage of the multiple physical paths between the transmit and receive antennas to carry multiple data streams. It has been shown that in a MIMO system with N transmit and M receive antennas, the channel capacity grows linearly with $\min\{N, M\}$ [3]. Recent results show that similar capacity scaling applies when an N-antenna access point (AP) communicates with M users simultaneously [2]. A multi-user (MU) MIMO system has the potential to combine the high capacity achievable with MIMO processing with the benefits of multi-user space-division multiple access. Such technology is being considered for the next generation of 802.11 (802.11ac). Particularly, we're interested in downlink (DL) MU MIMO systems, where an AP can transmit to multiple users simultaneously.

There has been prior work that studied the benefit of DL MU MIMO techniques in WLANs [4]–[6]. Applying an Earliest Deadline First (EDF) scheduling algorithm, Choi, Lee, and Bahk [4] demonstrated the performance benefit of DL MU MIMO over the single-user mechanism. This work focused on the performance analysis of DL MU MIMO, but did not consider MAC protocol design and MAC overhead in its analysis and simulations.

A MIMO distributed coordination function (DCF) protocol was presented in [7], using modified request-to-send/clear-to-send (RTS/CTS) frames to exchange antenna selection information and exploiting diversity and multiplexing gains. A modified acknowledgement (ACK) frame was also introduced to indicate whether a packet is received successfully on per spatial stream basis. In contrast, our proposed protocol does not modify RTS/CTS/ACK frames. A distributed MIMO-aware MAC was proposed in [8], assuming a three element antenna array based MIMO system that allows two simultaneous transmissions in a single collision domain. As will be discussed in Section III, our proposed solution can work for any antenna configuration.

In [5], the authors proposed a distributed DL MU MIMO MAC protocol that is based on the IEEE 802.11 MAC and provided an analysis of the proposed MU MAC protocol in terms of the maximum number of supported users and network throughput. The MAC protocol proposed in [5] is similar to one of our proposed protocols, i.e. the scheduled response mechanism (see Section III). However, we propose and evaluate multiple different response mechanisms for DL MU MIMO in this paper. We also propose enhancement mechanisms that work with DL MU MIMO, such as dynamic MAC protection and per-STA weighted queuing.

In this paper, we investigate the problem of medium access control in WLANs with DL MU MIMO capability. The main contributions of this paper are summarized as follows: First, we propose a CSMA/CA based medium access protocol with multiple response options for DL MU MIMO WLANs. A dynamic MAC protection scheme is proposed to reduce the overhead of MAC protection. Secondly, we propose a novel per-STA weighted queuing mechanism to achieve fairness in the network. We derive the optimal saturation throughput with respect to the number of simultaneous contending devices. The proposed MAC protocol can fully exploit spatial multiplexing and maximally reduce overhead associate with the MAC mechanism. It can achieve better performance than the 802.11n transmit beamforming (TxBF) mechanisms [1], as demonstrated in our simulation studies.

The remainder of this paper is organized as follows. In Section II, we introduce the system model. The proposed MAC protocol is discussed in Section III. We present an analysis of the proposed MAC protocol in Section IV. Our simulation evaluation of the proposed protocol is presented in Section V. Section VI concludes this paper.

II. System Model

We consider an enhancement to an IEEE 802.11n system where the AP has N transmit and N receive antennas. Assume the AP transmits simultaneously to different stations (STAs) in the same basic service set (BSS). With N transmit antennas, the AP can transmit a total of N spatial streams. These N streams can be distributed across a maximum of N STAs.

When the AP transmits different streams to multiple STAs, interference from streams intended for one STA will cause interference to the other STAs. This is represented by the following equation.

$$\begin{aligned} Y_i &= \sqrt{\frac{\rho}{M}} H_i W_1 X_1 + \cdots + \sqrt{\frac{\rho}{M}} H_i W_i X_i + \\ &\quad \cdots + \sqrt{\frac{\rho}{M}} H_i W_M X_M + Z_i \\ &= \sqrt{\frac{\rho}{M}} [W_1, \cdots, W_M] \begin{bmatrix} X_1 \\ \vdots \\ X_M \end{bmatrix} + z_i, \end{aligned} \tag{1}$$

where Y_i is the received signal at the ith STA (with dimensions $N_{Rx} \times 1$), X_i is the transmitted streams to the ith STA (with dimensions $N_{ss} \times 1$), N_{ss} is the number of spatial stream for each STA, H_i is the channel between the AP and the ith STA (with dimensions $N_{Rx} \times N_{Tx}$), W_i's are weights applied at the transmitter (with dimensions $N_{Tx} \times N_{ss}$), ρ is the received power, M is the number of STAs, Z_i is addition white Gaussian noise at the ith STA (with dimensions $N_{Rx} \times 1$), N_{Rx} is the number of receiving antennas at a STA, and N_{Tx} is the number of transmitting antennas at the AP.

The signal $H_i W_j X_j$ received by Y_i causes interference when decoding its streams X_i when $i \neq j$. The AP can mitigate this interference with intelligent beamforming techniques [9]. For example, if we select weights such that $H_i W_j = 0$ when $i \neq j$, then the interference from other STAs is canceled out.

A simple linear processing approach is to precode the data with the pseudo-inverse of the channel matrix [9]. To avoid the noise enhancement that accompanies zero forcing techniques, the minimum mean square error (MMSE) precoding can be used instead. To describe this approach, we first present the entire system model including all STAs as follows.

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_M \end{bmatrix} = \sqrt{\frac{\rho}{M}} \begin{bmatrix} H_1 \\ \vdots \\ H_M \end{bmatrix} \begin{bmatrix} W_1 \\ \vdots \\ W_M \end{bmatrix}^T \begin{bmatrix} X_1 \\ \vdots \\ X_M \end{bmatrix} + \begin{bmatrix} Z_1 \\ \vdots \\ X_M \end{bmatrix}.$$

That is,

$$Y = \sqrt{\frac{\rho}{M}} HWX + Z. \tag{2}$$

The MMSE precoding weights are then given as follows.

$$W = \sqrt{\frac{\rho}{M}} H^\dagger \left(\frac{\rho}{M} H H^\dagger + \Phi_z \right)^{-1}, \tag{3}$$

where Φ_z is the noise covariance matrix and $H^\dagger$ is the Hermitian of H.

Interference cancellation techniques can be implemented in the receiver to further reduce degradation from multiple access interference. When the receiving STA has more receive antennas than the number of spatial streams it intends to received, the extra antennas can be used to cancel out the spatial streams intended for other STAs. If channel state information (CSI) is known for the channel dimensions of the interference streams (i.e., $H_i W_j$), the CSI can be used to null interference in an MMSE receiver. This type of equalizer structure is given by $G_i Y_i$, where

$$G_i = \sqrt{\frac{\rho}{M}} W_i^H H_i^H \left(\sum_{k=1}^{M} \frac{\rho}{M} H_i W_k W_k^H H_i^H + \Phi_z \right)^{-1}. \tag{4}$$

To compare DL MU MIMO to single user 802.11n TxBF, we assume that the transmitter weights are generated using the eigenvectors from singular value decomposition (SVD). Though a specific weighting scheme is not defined in 802.11n, SVD yields maximum likelihood performance with a simple linear receiver [10]. The system equation with single user TxBF is expressed as,

$$Y = \rho HVX + Z. \tag{5}$$

where the SVD of H is $U\Sigma V$. When the AP has more antennas than transmitted spatial streams, the TxBF gain can be substantial even when the receiver has the same number of receive antennas as spatial streams.

III. CSMA/CA Based DL MU MIMO Protocol

In this section, we describe a DL MU MIMO MAC protocol based on CSMA/CA. Three different response mechanisms are proposed, as well as a novel weighted queuing mechanism to mitigate the fairness problem.

A. CSMA/CA Based DL MU MIMO MAC Protocol

The IEEE 802.11 MAC protocol is based on carrier-sense multiple access with collision avoidance (CSMA/CA) [11]. In this section, we propose to extend the 802.11 MAC to support DL MU MIMO transmission. With the proposed extension, an AP contends for the medium using the normal 802.11 enhanced distributed channel access (EDCA) procedure. Once an STA wins the channel, the AP transmits multiple packets that are destined for different STAs simultaneously.

We describe three response mechanisms that can be used for the AP to collect acknowledgments from STAs. The first response mechanism is illustrated in Fig. 1. This is a polled response mechanism, where the AP transmits block ACK

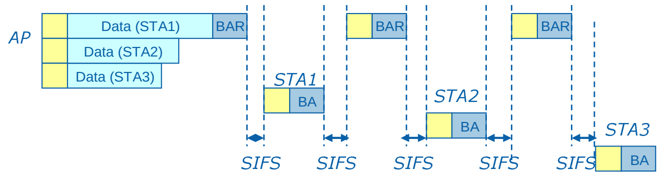

Fig. 1. CSMA/CA based DL MU MIMO protocol with polled response.

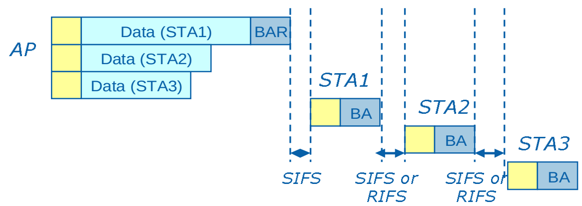

Fig. 2. CSMA/CA based DL MU MIMO protocol with scheduled response.

request (BAR) frame to each destination STA in turn to solicit block ACKs (BAs).

The remaining two scheduled response mechanisms are illustrated in Fig. 2. With these approaches, the AP includes an offset in the frame header. The offset defines when a destination STA can return a BA. Each STA transmits a BA, following the offset defined in the header of the received frame. In one option, BAs from different STAs are separated by short inter-frame space (SIFS); in another option, BAs are separated by reduced inter-frame space (RIFS). Because RIFS is 2us and SIFS is 16us, scheduled response with RIFS has smaller MAC overhead than scheduled response with SIFS.

In the case when one of the packets in the DL MU MIMO burst is not successfully received, the corresponding receiver would not reply a BA. Using the polled response mechanism, if an AP senses the medium idle PIFS after transmitting a BAR frame, it immediately transmits a BAR frame towards the next destination STA, as illustrated in Fig. 3. This error-recovery mechanism serves two purposes:

- to avoid gaps between responses and keep the medium busy so that other STAs do not attempt channel access and collide with the remaining BAs, and
- to reduce the response overhead by not waiting for the duration of the BA.

The AP's backoff procedure for an MU transmission is as follows. If a response is received from at least one of the STAs address in the DL MU MIMO burst, the AP assumes there is no collision. If a response is not received from any of the STAs address in the burst, then the AP assumes a

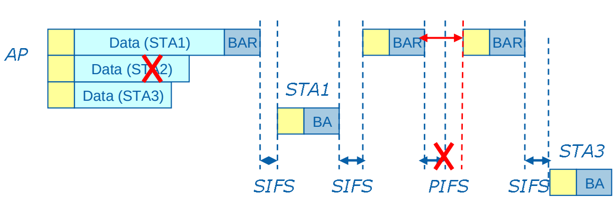

Fig. 3. Error recovery for polled response mechanism.

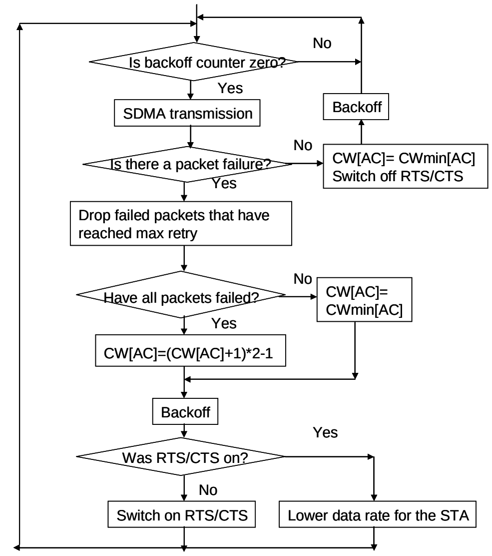

Fig. 4. Flow chart of the AP backoff procedure.

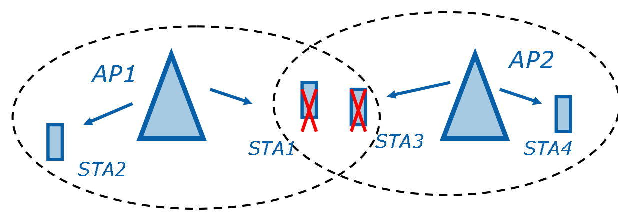

Fig. 5. Illustration of a hidden node scenario.

collision and initiates exponential backoff. A dynamic MAC protection scheme is combined with the AP backoff procedure, in which the AP does not turn on MAC protection until a failure occurs. This is because RTS/CTS exchange introduces a fixed overhead. If there is no collision in the network, MAC protection incurs overhead rather than providing benefits. The flow chart of the AP backoff procedure is illustrated in Fig. 4.

B. *Per-STA Weighted Queuing Mechanism*

After a DL MU MIMO transmission, the AP does not initiate exponential backoff when only one STA does not respond with a BA. However, in some cases, the AP may choose not to transmit to the STA that fails to transmit a BA. Fig. 5 illustrates a scenario where AP1 and AP2 are hidden nodes with respect to each other and thus cannot detect each other's transmission. Because AP2's transmission to STA3 can interfere with AP1's transmission to STA1 and *vice versa*, packets destined for STA1 and STA3 would collide with each other. If STA1 is part of a DL MU MIMO group but STA3 is not, AP1 would keep on transmitting to STA1 without exponential backoff because other STAs in the DL MU MIMO group have successfully received their packets while AP2 would keep on backing off exponentially. As a result, STAs in a DL MU MIMO group gain an unfair advantage on channel access.

To mitigate this fairness problem, we propose a per-STA

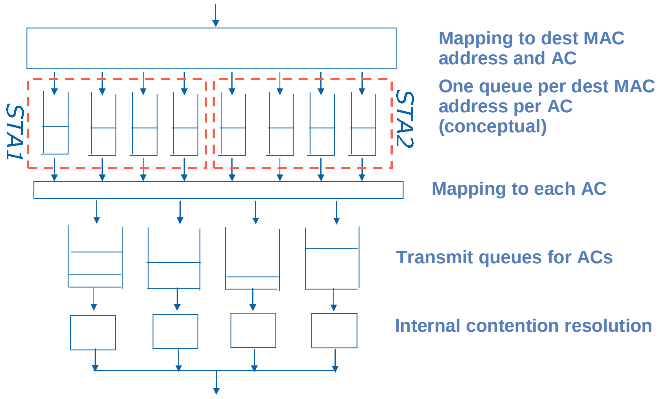

Fig. 6. Illustration of Weighted Queuing at the AP.

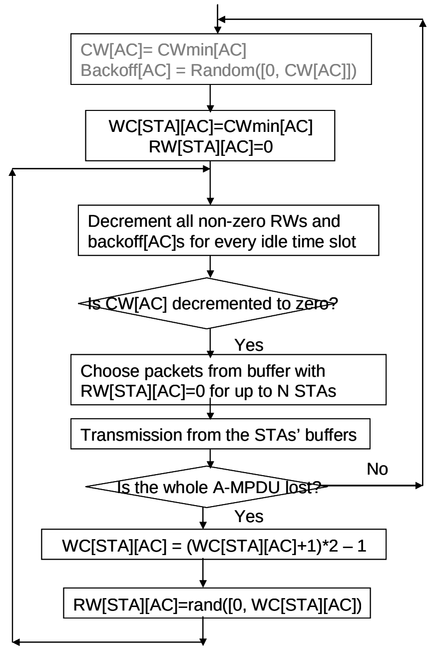

Fig. 7. Flow chart of weighted queuing.

weighted queuing mechanism at the AP [12], as illustrated in Fig. 6. When downlink traffic arrives at the AP, it is buffered according to its destination MAC address and its access category (AC). For each queue, there is one associated weight counter, i.e. WC[STA][AC], and one random weight, i.e. RW[STA][AC].

In Fig. 7, we show the flow chart of the per-STA weighted queuing mechanism. Initially, WC[STA][AC] is set to CWmin[AC], and RWs are set to zero, where CWmin[AC] is the minimum contention window defined for an AC. For every idle time slot, all non-zero RWs are decremented by one. When the AP is ready to transmit packets from a particular AC, it only chooses packets from the queues where RW[STA][AC] is zero. If the transmission for a particular STA receives a response, the corresponding RW[STA][AC] is set to zero and WC[STA][AC] is set to CWmin[AC]. If the transmission for a particular STA does not receive a response, the corresponding WC[STA][AC] is incremented as follows:

$$\mathrm{WC[STA][AC]} = (\mathrm{WC[STA][AC]} + 1) \times 2 - 1, \tag{6}$$

and RW[STA][AC] is drawn as a random integer from a uniform distribution over an interval [0, WC], which is RW[STA][AC]= Random([0, WC]).

This weighted queuing mechanism is equivalent to implementing an internal per-STA backoff procedure at the AP such that all STAs that are involved in a collision initiate exponential backoff.

IV. Performance Analysis

In this section, we derive the saturation throughput of WLAN system using the proposed protocol. The system's saturation throughput is defined as the combined throughput achieved at the top of the MAC layer when all nodes in the systems are fully loaded at all times.

It is assumed that the devices use MAC frame aggregation schemes, such as aggregated-MAC Protocol Data Unit (A-MPDU), and multiple transmissions in one transmit opportunity (TXOP). We follow the assumptions made in [13] and the same 2-D Markov chain model. In the Markov chain mode, each state is represented by $\{s(t), b(t)\}$, where $s(t)$ is defined to be the stochastic process representing the backoff stage $[0, 1, \cdots, m]$ of the station at time t and $b(t)$ is the stochastic process representing the backoff time counter for a given station. The maximum backoff stage, i.e., m, takes the value such that $CW_{max} = 2^m CW_{min}$, where CW_{max} is the maximum contention window and CW_{min} is the minimum contention window.

Let S be the normalized system throughput, defined as the fraction of time when the channel is used to successfully transmit the payload bits. S can be expressed as the average payload bits transmitted in a TXOP divided by the average length of a TXOP. Based on the 2-D Markov chain mode, we extend the analysis in [13] and derive the system saturation throughput as:

$$\begin{aligned} S &= P_{AP} \frac{P_s P_{tr} \sum_{j=1}^{M} \sum_{i=1}^{N_j} \mathrm{E}[P_{ij}]}{(1-P_{ij})\sigma + P_{tr} P_s T_s + P_{tr}(1-P_s) T_c} + \\ & \quad P_{STA} \frac{P_s P_{tr} \sum_{j=1}^{n-1} \sum_{i=1} N_j \mathrm{E}[P_{ij}]}{(1-P_{ij})\sigma + P_{tr} P_s T_s + P_{tr}(1-P_s) T_c} && (7) \\ T_s &= TXOP_{dur} && (8) \\ T_c &= TRS + DIFS && (9) \\ P_t r &= 1 - (1-\tau)^n && (10) \\ P_s &= \frac{n\tau(1-\tau)^{n-1}}{1-(1-\tau)^n}, && (11) \end{aligned}$$

where P_{AP} is the probability that the AP wins the contention, P_{STA} is the probability that a STA wins the contention, M is the number of users to which an AP can transmit simultaneously, T_s is the average time consumed by a successful TXOP, T_c is the average medium time a collision consumes, σ is the duration of a time slot, RTS is the transmission duration of the RTS frame, n is the number of contending devices in the

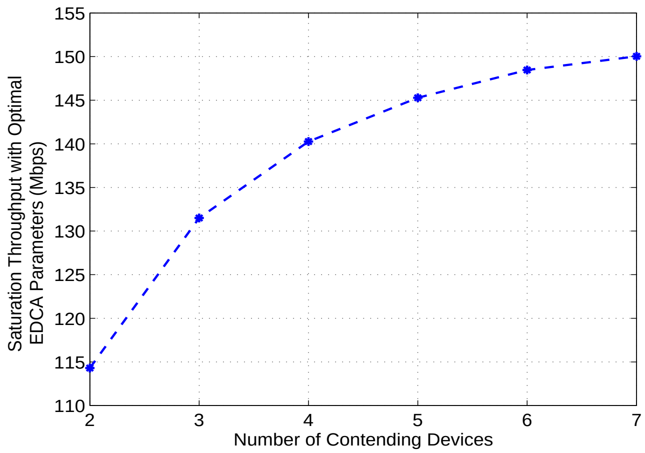

Fig. 8. Saturation throughput S vs. number of contending devices n (optimal EDCA parameters).

network, including the AP and the stations, τ is the probability that a device transmits in a randomly chosen time slot, P_s is the probability that a TXOP is successfully set up, P_{tr} is the probability that there is at least one transmission in the considered slot time, $\sum_{i=1}^{N_j} \mathrm{E}\left[P_i\right]$ is the combined average payload size of N_j A-MPDUs that are transmitted in the TXOP.

Equation (7) can be rearranged as follows:

$$S = \frac{\frac{1}{n}\left(\sum_{j=1}^{M}\sum_{i=1}^{N_j}\mathrm{E}[P_{ij}] + \sum_{j=1}^{n-1}\sum_{i=1}^{N_j}\mathrm{E}[P_{ij}]\right)}{T_s - T_c + \frac{T_c-(1-\tau)^n(T_c-\sigma)}{n\tau(1-\tau)^{n-1}}}. \tag{12}$$

Under condition $\tau \ll 1$, τ can be estimated as [13]

$$\tau \approx \left(n\sqrt{\frac{T_c}{2\sigma}}\right)^{-1}. \tag{13}$$

In Fig. 8, we plot the relationship between the optimal saturation throughput S and the number of contending devices n in the BSS. When all contending devices have equal transmission opportunities, the saturation throughput of the network increases with the number of contending devices due to spatial diversity gain achieved by DL MU MIMO.

V. Simulation Study

The proposed DL MU MIMO MAC protocol is implemented in OPNET Modeler [14]. Using OPNET simulations, we evaluate the performance of the proposed DL MU MIMO MAC protocol and compare its performance with that of the beam-forming protocol. Our simulations consider a typical WLAN topology, consisting of one AP, equipped with four antennas, and multiple STAs, each of which is equipped with two antennas. Other simulation parameters are presented in Table I.

To support DL MU MIMO, we assume that the STA implements interference cancellation techniques necessitating more receive antennas than received spatial streams. Therefore in the simulations, the AP only transmits one spatial stream

TABLE I
Simulation Parameters

Parameter (unit)	*Value*	*Parameter (unit)*	*Value*
DL MU MIMO data rate (Mbps)	65	aSlotTime (μs)	9
BF data Rate (Mbps)	130	aSIFSTime (μs)	16
Control rate (Mbps)	24	TXOP duration (ms)	3
RTS (byte)	20	A-MPDU size (byte)	1,500
CTS (byte)	14	CWmin	7
BA size (byte)	32	CWmax	63

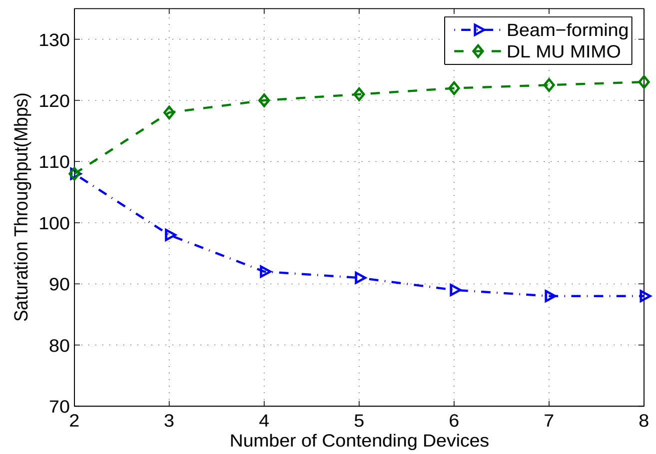

Fig. 9. Saturation throughput S vs number of contending devices n (bidirectional traffic).

to each STA, which has two antennas. However, when TxBF is used in the simulations, each STA can receive two spatial streams. Because STAs are placed close to the AP, on average the achievable signal-to-noise ratio (SNR) at each receiver is at least 30dB.

We first compare the saturation throughput of DL MU MIMO with that of TxBF with respect to the number of contending devices n. The simulation results are plotted in Fig. 9. It can be seen that saturation throughput achieved by beam-forming degrades. The reason is that DL MU MIMO can effectively take advantage of the spatial diversity gain, which is larger when the number of contending STAs increases, while the beam-forming scheme does not have this capability.

We next evaluate the performance of the three DL MU MIMO response mechanisms, assuming a WLAN with one AS and three STAs. As illustrated in Fig. 10 and Fig. 11, due to spatial multiplexing gain, when the AP can transmit simultaneously to three STAs, all DL MU MIMO techniques achieve higher saturation throughput than TxBF. When MAC protection is not enabled, the polled response mechanism performs better than the scheduled ACK mechanisms. This is because the polled response mechanism implements an error recovery mechanism and thus is more robust than scheduled ACK mechanisms when there is no MAC protection. When MAC protection is enabled, schedule ACK mechanisms perform better than the polled ACK mechanism due to the lower

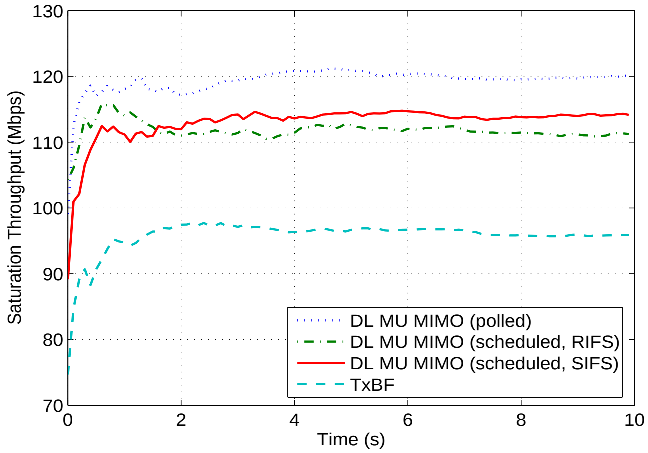

Fig. 10. Saturation throughput S (without MAC protection).

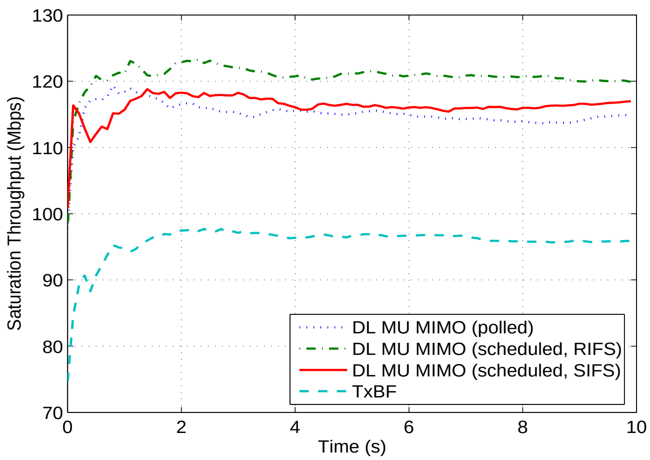

Fig. 11. Saturation Throughput S (with MAC protection).

MAC protocol overhead.

VI. Conclusion

In this paper, we proposed and evaluated a CSMA/CA based DL MU MIMO protocol with three response mechanisms. Furthermore, we propose a novel weighted-queuing mechanism to achieve fairness in the WLAN where DL MU MIMO is utilized. Analysis and simulation study both show that when the number of contending STAs increases, the saturation throughput achieved by DL MU MIMO also increases. Furthermore, our simulation results show that DL MU MIMO can achieve better performance than BF when there are more than two STAs in the network.

Acknowledgment

Shiwen Mao's research is supported in part by the National Science Foundation under Grants ECCS-0802113 and CNS-0855251, and through the Wireless Internet Center for Advanced Technology at Auburn University (NSF Grant IIP-0738088).

References

[1] IEEE, "Enhancements for higher throughput," 2009, IEEE standard amendment 802.11n.

[2] D. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1859–1874, Sept. 2004.

[3] A. Goldsmith, S. Jafar, N. Hindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

[4] Y. Choi, N. Lee, and S. Bahk, "Exploiting multi-user MIMO in the IEEE 802.11 wireless LAN systems," *Wireless Personal Commun.*, May 2009, DOI: 10.1007/s11277-009-9731-2.

[5] L. Cai, H. Shan, W. Zhuang, X. Shen, J. Mark, and Z. Wang, "A distributed multi-user MIMO MAC protocol for wireless local area networks," in *Proc. IEEE GLOBECOM'08*, New Orleans, LA, Nov./Dec. 2008.

[6] M. Zhao, M. Ma, and Y. Yang, "Applying opportunistic medium access and multiuser MIMO techniques in multi-channel multi-radio WLANs," *J. Mobile Netw. Appl*, vol. 14, no. 4, pp. 486–507, Aug. 2009.

[7] J. Mirkovic, G. Orfanos, H.-J. Reumerman, and D. Denteneer, "A MAC protocol for MIMO based IEEE 802.11 wireless local area networks," in *Proc. IEEE WCNC'07*, Hong Kong, China, Mar. 2007, pp. 2131–2136.

[8] D. Dechene, K. Meerja, A. Shami, and S. Primak, "A novel MIMO aware distributed media access control scheme for IEEE 802.11 wireless local area networks," in *Proc. IEEE LCN'07*, Dublin, Ireland, . 2007, pp. 125–132.

[9] Q. Spencer, C. Peel, A. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 60–67, Oct. 2004.

[10] E. Perahia and R. Stacey, *Next Generation Wireless LANs: Throughput, Robustness, and Reliability in 802.11n*. Cambridge, UK: Cambridge University Press, 2008.

[11] IEEE, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specification," 2007, IEEE standard 802.11.

[12] C. Oottamakorn, S. Mao, and S. Panwar, "On generalized processor sharing with regulated multimedia traffic flows," *IEEE Trans. Multimedia*, vol. 8, no. 6, pp. 1209–1218, Dec. 2006.

[13] G. Bianchi, "Performance analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[14] OPNET, "OPNET Modeler," [online] Avalable: http://www.opnet.com/.