Training Protocols for Multi-User MIMO Wireless LANs

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

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

***Abstract*—In this paper, we investigate the training problem of wireless local area networks (WLANs) with downlink multi-user multiple input multiple output (DL MU MIMO) capability. We extend the 802.11 MAC protocol and propose a few training protocols at the MAC layer to support DL MU MIMO. We provide a capacity analysis based on measurement results from an 802.11n systems, evaluate the overhead of these training protocols, and compare the performance of DL MU MIMO with that of a beam-forming (BF) based approach. Through simulation studies, we find that at high SNR and with implicit training, the DL MU MIMO mechanism provides significant performance gain over the BF based approach. Furthermore, our capacity analysis and simulation study also show that it is critical to define appropriate training intervals based on antenna configuration, channel mobility, and CSI feedback overhead.**

I. Introduction

Multiple-input multiple-output (MIMO) communication techniques have been extensively studied for next generation cellular networks and has been deployed in wireless local area networks (WLANs) using the IEEE 802.11n technology. A MIMO system takes advantage of two types of gains, spatial diversity gain and spatial multiplexing gain [1]. Spatial diversity is used to combat severe fading and can improve reliability of wireless links by carrying duplicate copies of the same information along multiple antennas. This is particularly useful for compensating against the effect of node mobility. Spatial multiplexing creates an extra dimension in spatial domain, which can carry independent information in multiple data streams. It has been shown that in a MIMO system with N transmit and M receive antennas, capacity grows linearly with $\min\{N, M\}$ [2].

Recent results show that similar capacity scaling applies when an N-antenna access point (AP) communicates with M users simultaneously [1]. Such 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. This 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 [3]–[5]. However, the important problems of training protocol design an training overhead for DL MU MIMO systems have not been addressed in these works. In [6], a training mechanism is adopted with channel state information (CSI) feedback. Because the feedback is transmitted over a cable instead of over an air interface, the training overhead is not taken into consideration in this work. In [7], the authors provide a comprehensive analysis of a few DL MU MIMO precoding techniques and characterize the degradation in the performance of transmitter optimization schemes with respect to delayed CSI feedback. However, this work does not consider exploiting MAC support for DL MU MIMO training.

In this paper, we investigate the training problem in WLANs with DL MU MIMO capability. The main contributions of the paper are as follows. First, we extend the 802.11 CSMA/CA based medium access protocol and propose multiple training mechanisms at the MAC layer to support DL MU MIMO. Second, based on measurement results taken from an 802.11n network, we study the capacity degradation of DL MU MIMO with respect to channel state information (CSI) feedback delay. Third, through OPNET simulations, we evaluate the overhead of training protocols, and compare the performance of DL MU MIMO with that of transmit beamforming (TxBF).

Through simulation studies, we find at high SNR and with implicit training, the DL MU MIMO mechanism provides significant performance gain over the BF based approach. Furthermore, our capacity analysis and simulation study also show that it is critical to define appropriate training intervals based on antenna configuration, channel mobility, and CSI feedback overhead.

The remainder of this paper is organized as follows. In Section II, we introduce the system model. The extended MAC protocol and the proposed training mechanisms are discussed in Section III. We present capacity analysis of DL MU MIMO versus CSI feedback delay in Section IV. Our simulation study is presented in Section V. Section VI concludes this paper and discuss future directions.

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 presented in (1), where Y_i is the received signal at the ith STA (with dimensions

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

$N_{Rx} \times 1$), 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 an 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 [8]. 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 [8]. 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,

$$Y = \sqrt{\frac{\rho}{M}} HWX + Z. \quad (2)$$

The MMSE precoding weights are then given as follows.

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

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

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

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 [9]. The system equation with single user TxBF is expressed as,

$$Y = \rho HVX + Z. \quad (5)$$

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.

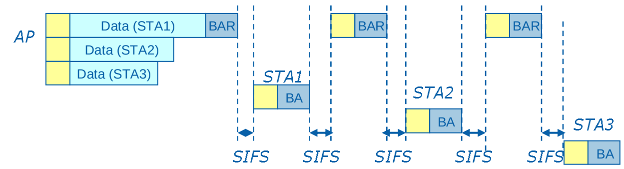

Fig. 1. CSMA/CA based DL MU MIMO protocol with polled response.

III. Training Protocols for DL MU MIMO

In this section, we first describe a CSMA/CA based protocol for DL MU MIMO WLANs. We then present the proposed training mechanisms for DL MU MIMO systems.

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). We extend the 802.11 protocol to support DL MU MIMO transmission. 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. Two response mechanisms can be used for the AP to collect acknowledgements from STAs. Fig. 1 illustrates a polled response mechanism, where the AP transmits block ACK request (BAR) frame to each destination STA in turn to solicit block ACKs (BAs).

Fig. 2 illustrates a scheduled response mechanism where the AP includes an offset in the frame header. The offset defines when a destination STA can transmit back a BA. After successfully receiving a data frame from the AP, 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 2 μs and SIFS is 16 μs, scheduled response with RIFS has smaller MAC overhead than scheduled response with SIFS.

The APs backoff procedure for a MU transmission is the following. 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 collision and initiates exponential backoff. The flow chart of the AP's backoff procedure is illustrated in Fig. 3.

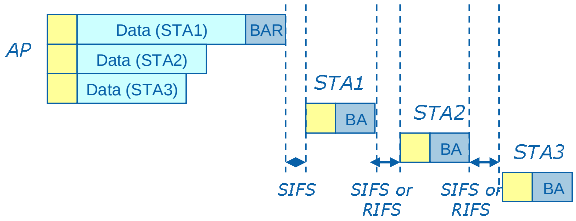

Fig. 2. CSMA/CA based DL MU MIMO protocol with scheduled response.

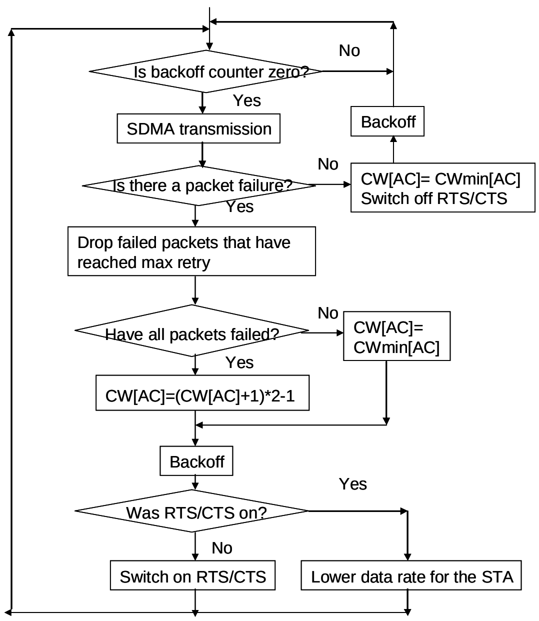

Fig. 3. Flow chart of the AP backoff procedure.

B. Training Protocols

Following the design philosophy of response mechanisms for DL MU MIMO, we propose two training protocols: polled training and scheduled training. As shown in Fig. 4, with the polled training mechanism, the AP transmits a Request frame to each STA. Upon receiving a Request frame, an STA replies with a Reply frame. As shown in Fig. 5, utilizing the scheduled training mechanism, the AP transmits a Request frame, containing a schedule that determines when each STA can transmit a response. Upon receiving the Request frame with a schedule, an STA transmits a Reply frame following the offset indicated in the Request frame.

Depending on whether the Channel State Information (CSI) is calculated by the transmitter or the receiver, there are two training mechanisms: i) training with implicit feedback or ii) training with explicit feedback. When training with implicit

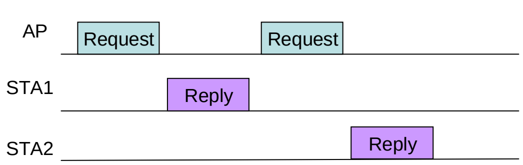

Fig. 4. Polled training mechanism.

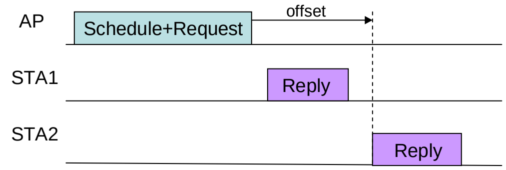

Fig. 5. Scheduled training mechanism.

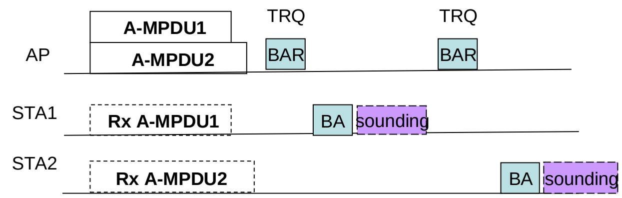

Fig. 6. Implicit feedback (polled).

feedback is utilized, the AP requests a sounding frame from an STA. The sounding frame contains a few pre-defined Long Training Fields (LTFs) that are known to both the transmitter and the receiver. Ideally enough LTFs are sent to sound the full dimensionality of the channel. The sounding frame is used by the AP to estimate CSI. Based on the estimated CSI, the AP can precode the DL MU MIMO data. Figs. 6 and 7 illustrate polled and scheduled training with implicit feedback. Note that to reduce training overhead, the training frames can be combined with data communication. In Fig. 6, the AP transmits one Aggregated MAC Protocol Data Unit (A-MPDU) intended for each destination STA. The AP can indicate training requested (TRQ) in each A-MPDU or in the block ack request (BAR) frame by setting the TRQ bit. Upon receiving the BAR frame that contains a TRQ bit, an STA replies with a block ack (BA) immediately followed by a sounding frame.

Instead of transmitting a separate sounding frame, which often is a Null Data Packet (NDP), STAs can also combine the sounding LTFs with a data frame. Such a frame is called a staggered sounding frame. Normally an LTF is transmitted per spatial stream to decode the data. With staggered sounding, additional LTFs are transmitted in the preamble in order to sound the additional channel dimensions.

When training with explicit feedback is utilized, the AP transmits the sounding frame and relies on the STAs to send back CSI feedback. The CSI feedback includes the signal-to-noise ratio (SNR) measured at each of the receiving chain and CSI matrix for each subcarrier. Therefore, the training

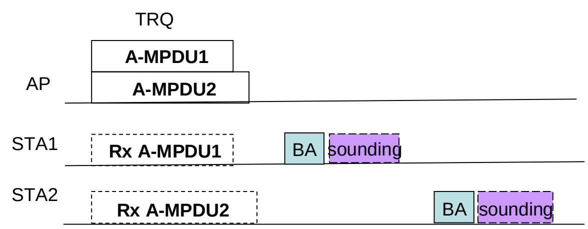

Fig. 7. Implicit feedback (scheduled).

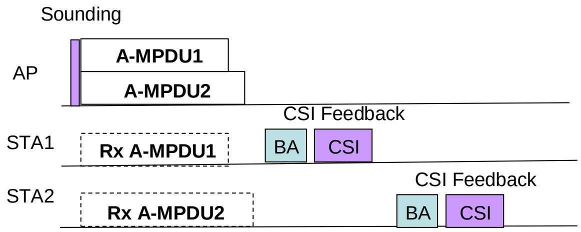

Fig. 8. Explicit feedback (scheduled, staggered).

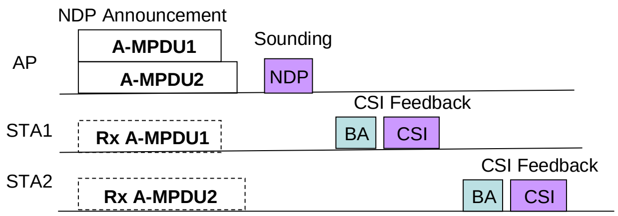

Fig. 9. Explicit feedback (scheduled).

overhead of explicit feedback is proportional to the total number of Tx antennas, the total number of Rx antennas, and the number of subcarriers.

Fig. 8 shows an example where the AP transmits A-MPDUs with staggered LTFs as sounding frames. Upon receiving the sounding frame, each STA replies with its measured CSI feedback. The AP can also transmit an NDP sounding frame after the A-MPDU transmission, as illustrated in Fig. 9. Note that the polled training mechanism can also be combined with explicit feedback. They are omitted here for simplicity.

In some implementations, it may take some time for an STA to compile the CSI feedback frame. Our proposed training mechanism is flexible enough to accommodate delayed feedback as shown in Fig.10.

IV. Capacity vs. Feedback Delay

A key parameter in the overall throughput of a closed-loop feedback system is the feedback rate and the overhead incurred by such feedback. The motion or Doppler in the environment will cause the channel to change during the interval between measurement of the CSI and application of weights on transmission. Different transmit weighting schemes will have different degrees of sensitivity to these environmental changes, which will dictate the feedback rate.

In [10], measurement results were captured in an 802.11n test bed deployed on one floor of an indoor office environment. Detailed information regarding the office environment, the test bed setup, and the conducted experiments can be found in [10].

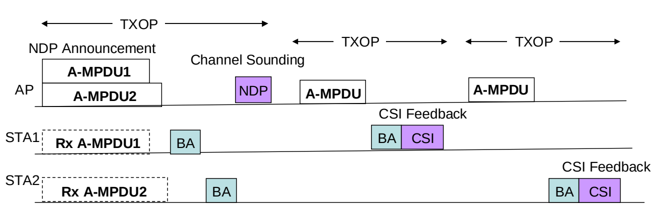

Fig. 10. Explicit feedback (delayed).

Using these measurements, TxBF capacity gain was computed as a function of feedback delay. Analysis demonstrated that the majority of the gain was maintained even with 100 msec of feedback delay. Therefore, we use 100 msec in the simulations for the training interval for TxBF.

With DL MU MIMO, aging of the CSI causes a mismatch between the precoding weights and channel resulting in multiple access interference. Due to this affect, DL MU MIMO will be more sensitive to changes in the channel and will require more frequent updates of the CSI.

To determine the appropriate training interval for the DL MU MIMO simulations, we process a subset of the measurements collected in [10]. A semi-analytic capacity formulation is used to compute capacity, as described in [9]. The expression for the mean square error (MSE) at the receiver is

$$J_{M\times M} = \frac{\rho}{M} G H_{eff} H_{eff}^{*} G^{*} + G\Phi_Z G^{*} - 2\sqrt{\frac{\rho}{M}}\mathrm{Re}\{G H_{eff}\} + I, \quad (6)$$

where H_{eff} is the effective channel matrix including the delayed MMSE precoding weights from (3), H_{eff}^{*} is the conjugate transpose of H_{eff}, Φ_Z is the noise covariance matrix, I is the identity matrix from the signal covariance, G is the MMSE interference cancellation equalizer from (4), and M is the number of data streams across all the STAs.

The output SNR for the ith data stream for MMSE is given by

$$\mathrm{SNR}_i = (1 - J_i)/J_i, \quad (7)$$

where J_i is the ith diagonal element of the MSE matrix given in (6). The formula for capacity based on output SNR is

$$C = \sum_{i=1}^{M} \log_2(1 + \mathrm{SNR}_i). \quad (8)$$

The DL MU MIMO capacity results for a system with an AP with three transmit antennas and two STAs are given in Fig. 11. In the figure there are two curves, one modeling capacity when the STAs have a single receive antenna and the other curve when the STAs two receive antennas. In both cases a single stream is transmitted to each STA and the SNR at the receiver is about 39dB. Fig. 11 illustrates that if the number of Rx antenna at an STA is more than the number of spatial streams destined for the STA and the number of transmit antennas is more than the total number of spatial streams, the capacity of DL MU MIMO does not degrade much with an increase in CSI feedback delay. However, if the above conditions are not completely satisfied, the capacity may decrease significantly when the CSI feedback is delayed.

Similar results are provided in Fig. 12 for a system with three STAs. In this scenario, the AP transmits a total of three streams, one to each STA. The SNR at the receiver is about 34dB. Fig. 12 shows that with three different antenna-client configurations, two configurations show significant performance degradation with delayed CSI feedback, whereas one configuration, i.e. three Rx-antennas three clients, performs well even when the CSI feedback is delayed by 30 to 40ms.

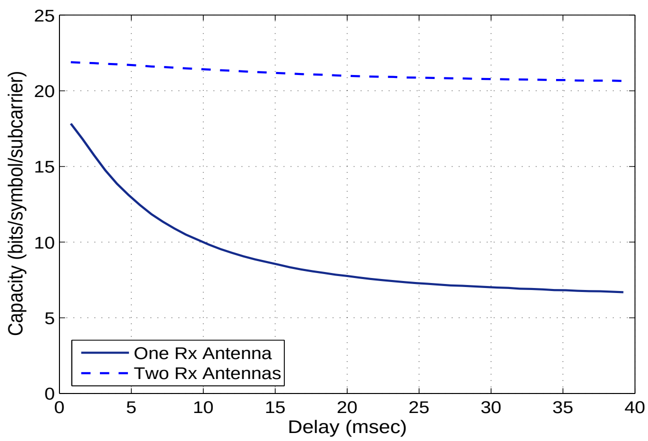

Fig. 11. DL MU MIMO capacity with regard to feedback delay (with 2 STAs).

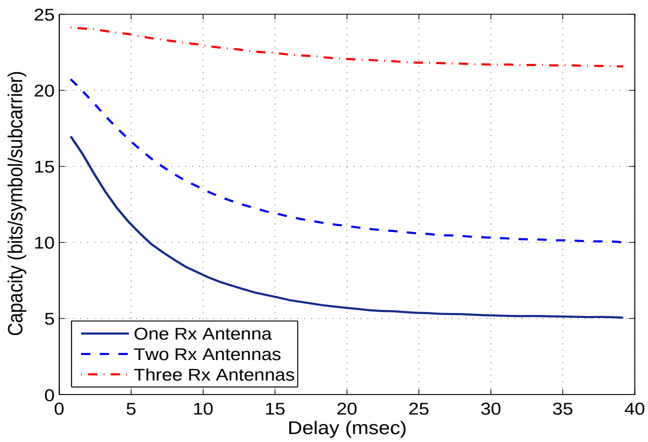

Fig. 12. DL MU MIMO capacity with regard to feedback delay (with 3 STAs).

In both sets of results, the sensitivity to delay and aging of the CSI is highly dependent on the number of transmitting antennas at the AP, the number of receiving antennas of the STAs and the interference cancellation ability of the STAs. With minimal interference cancellation, the capacity significantly degrades within 5 msec. However, with extra receive antennas at the STA, only a little degradation is experienced after 40 msec. Therefore in the DL MU MIMO simulations presented in Section V, we model the system performance with a range of trainings intervals, varying from 5ms to 40ms.

V. Simulation Study

Through OPNET simulations, we evaluate the saturation throughput of a DL MU MIMO system with different training protocols over a 20MHz channel. We compare the performance of DL MU MIMO with that of TxBF. A 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. Our simulation utilizes a typical wireless LAN topology, where there is one AP equipped with four antennas and three STAs each of which is equipped with two antennas. In the simulation, STAs use MAC frame aggregation schemes, such as A-MPDU, and multiple transmissions in one transmit opportunity (TXOP). Simulation parameters are defined in Table 1.

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
Training data rate (Mbps)	39	BA size (byte)	32
Control rate (Mbps)	24	CTS (byte)	14
Max A-MPDU size (byte)	64,000	RTS (byte)	20
MPDU size (byte)	1,500	CWmin	7
TXOP duration (ms)	3	CWmax	63
Training interval for DL MU MIMO (ms)	5, 10, 20, 40	Training interval for TxBF (ms)	100

Our measurement results show that to achieve reasonable DL MU MIMO capacity, an STA needs to implement interference cancellation techniques necessitating more receive antennas than received spatial streams. Therefore in the simulation, the AP only transmits one spatial stream (SS) to each STA with two antennas. However when TxBF is utilized, each STA can receive two SSes. Because STAs are placed close to the AP, on average the achievable signal-to-noise ratio (SNR) at each receiver is at least 30dB. The simulation assumes that the channel aging effect is negligible within the same interval. For implicit feedback, the number of LTFs for the sounding frame is the same as the sounded spatial streams. For explicit feedback, the CSI feedback sent back from STAs is proportional to the total number of Tx antennas, the total number of Rx antennas, and the number of subcarriers.

We first compare the saturation throughput of DL MU MIMO with that of TxBF when implicit training mechanism is utilized. Fig. 13 shows that DL MU MIMO with polled training protocol achieves 45% higher saturation throughput than the TxBF protocol whereas DL MU MIMO with scheduled training protocol achieves 50% higher saturation throughput than the TxBF protocol.

We then evaluate the training overhead of explicit CSI feedback on the performance of DL MU MIMO. Note that the training interval of TxBF scheme is fixed to 100ms because TxBF is not very sensitive to CSI feedback delay [10]. Fig. 14 illustrates a case when the AP has four Tx antennas, each STA has two Rx antennas, and the AP can transmit a DL MU MIMO burst simultaneously to three STAs. In this case, when the training interval increases from 5ms to 40ms, the saturation throughput of DL MU MIMO with polled training or scheduled training mechanism improves by 15%.

Fig. 15 illustrates a case when the AP has eight Tx antennas, each STA has two Rx antennas, and the AP can transmit a DL MU MIMO burst simultaneously to four STAs. In this case, due to the high training overhead, the performance of DL MU MIMO is much more sensitive to the training interval. When the training interval increases from 5ms to 40ms, the

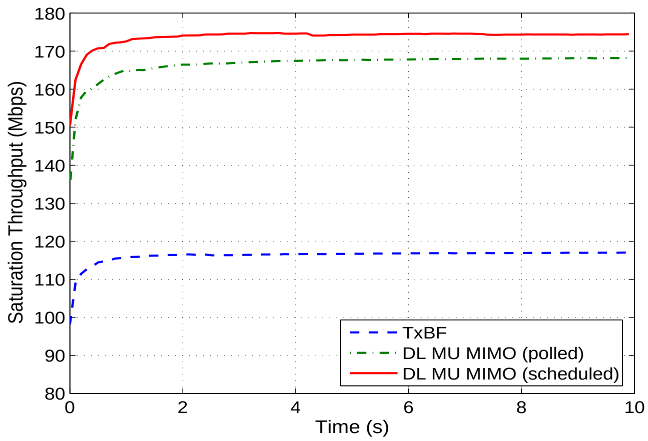

Fig. 13. Performance Comparison of DL MU MIMO and TxBF with implicit feedback.

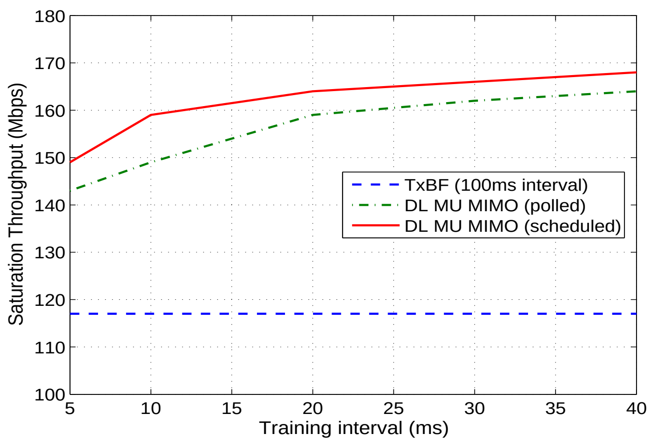

Fig. 14. Saturation throughput vs. training interval, with explicit feedback. The AP has four Tx antennas and there are three STAs, each with two Rx antennas.

saturation throughput of DL MU MIMO with polled training or scheduled training mechanism improves by 40%. Note that the simulation assumes that channel aging effect is negligible within the same training interval. In reality, depending on the antenna configuration and channel mobility, the performance of DL MU MIMO degrades when there is CSI feedback delay as shown in Figs. 11 and 12.

VI. Conclusion and Future Work

We have proposed and evaluated multiple training protocols for supporting DL MU MIMO in WLANs. Capacity analysis based on measurement data shows that the configuration of Tx/Rx antennas and the number of spatial streams have a profound impact on DL MU MIMO capacity and channel aging. Simulation results show that at high SNR and when AP has four Tx antennas, DL MU MIMO achieves 50% higher saturation throughput than TxBF. Furthermore, the training overhead with explicit feedback is non-negligible for DL MU MIMO systems, esp. when the AP has eight antennas. Therefore, a training interval needs to be carefully chosen based on channel mobility, antenna configurations, and CSI

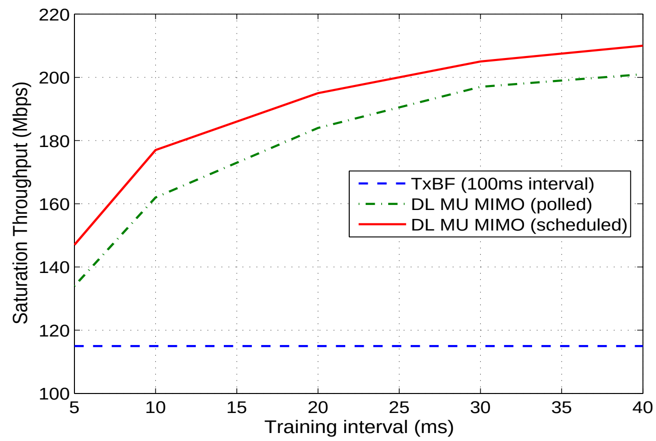

Fig. 15. Saturation throughput vs. training interval, with explicit feedback. The AP has eight Tx antennas and there are four STAs, each with two Rx antennas.

feedback overhead. For future work, we plan to develop a joint analysis of antenna-client configurations, channel aging, and CSI feedback overhead and provide recommendations for choosing appropriate training intervals.

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 (under NSF Grant IIP-0738088).

References

[1] 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.

[2] 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.

[3] Y. Choi, N. Lee, and S. Bahk, "Exploiting multi-user MIMO in the IEEE 802.11 wireless LAN systems," *Wireless Personal Commun.*, May 2009.

[4] 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.

[5] 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.

[6] D. Samardzija, H. Huang, R. Valenzuela, and T. Sizer, "An experimental downlink multi-user system with distributed and coherently coordinated transmit antennas," in *Proc. IEEE ICC'07*, Glasgow, Scotland, June 2007, pp. 5365–5370.

[7] D. Samardzija, N. Mandayam, and D. Chizhik, "Adaptive transmitter optimization in multi-antenna multi-user systems: Theoretical limits, effect of delays, and performance enhancements," *EURASIP J. Wireless Commun. Netw.*, no. 3, pp. 298–307, Aug. 2005.

[8] 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.

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

[10] E. Perahia, A. Sheth, T. Kenney, R. Stacey, and D. Halperin, "Investigating into the doppler component of the IEEE 802.11n channel model," submitted to Globecom 2010.