A Throughput Balancing Problem between Uplink and Downlink in Multi-user MIMO-based WLAN Systems

Hu Jin
School of EECS., KAIST
Daejeon, Korea
Email: jinhu@cnr.kaist.ac.kr

Bang Chul Jung
KAIST Institute for IT
Coinference, Daejeon, Korea
Email: bcjung@kaist.ac.kr

Ho Young Hwang
Department of ECE.,
University of Waterloo, Canada
Email: hyhwang@bbcr.uwaterloo.ca

Dan Keun Sung
School of EECS., KAIST
Daejeon, Korea
Email: dksung@ee.kaist.ac.kr

Abstract—Collision mitigation is one of classical research issues for wireless local area networks (WLANs). Recently, multiple-input multiple-output (MIMO) transmission techniques have been widely deployed in wireless systems, while a multi-user MIMO-based collision mitigation scheme in uplink WLANs was proposed by authors, and we showed the scheme is very efficient for the uplink performance. However, for an infrastructure-based WLAN, we observe a significant performance unbalance problem between uplink and downlink, compared to the conventional WLANs. Moreover, access point (AP) yields lower throughput performance than each contending station(STA). In order to solve this unbalance problem between uplink and downlink, we adopt a modified minimum contention window (CWmin) adjustment scheme and a random piggyback scheme to the multi-user MIMO-based WLANs. We also develop an analytical model to evaluate the performance of multi-user MIMO-based WLANs in a saturated traffic environment. The result shows that the random piggyback scheme performs more efficiently for the multi-user MIMO-based WLANs.

I. INTRODUCTION

Distributed coordination function (DCF) has been widely deployed in IEEE 802.11 wireless local area networks (WLANs) as a medium access control (MAC) protocol due to simplicity, low cost and efficiency[4]. The DCF is based on carrier sense multiple access with collision avoidance (CSMA/CA) where a binary exponential backoff scheme is used as a collision avoidance mechanism. Although this DCF can avoid some collisions in WLANs, collisions may occur more frequently as the number of contending stations(STAs) increases in a basic service set (BSS). This collision problem is considered as a key factor which degrades the system performance[5]. Thus, there have been several studies on enhancing the collision mitigation performance at MAC layer [6]-[8].

Multiple-input multiple-output (MIMO) transmission techniques have been widely deployed in wireless systems. Zheng et al.[1] adopted the concept of multipacket reception to the WLANs. Jin et al.[2][3] also proposed a collision mitigation scheme in uplink WLANs using multi-user MIMO antennas at an access point (AP). In this cross-layered design, the AP utilized a MIMO technique to simultaneously decode the multiple transmitted frames from multiple STAs. As a consequence, this scheme significantly enhances the uplink throughput performance. However, in this work, the effect of downlink traffic on the system performance was not considered. Since the multi-user MIMO-based scheme is only applied to the uplink, the AP has less channel access opportunities in downlink compared to the conventional WLANs and it may result in a severe performance degradation in downlink. Thus, an unbalance problem between uplink and downlink becomes more critical in these multi-user MIMO-based WLANs.

A weighted fairness problem between uplink and downlink has been discussed in conventional WLANs. Even if AP and STAs have the same channel access opportunities based on DCF in infrastructure-based WLANs, this DCF protocol causes a fairness problem between downlink and uplink when there are many STAs contending in uplink. To investigate the fairness and priority issues among STAs, Qiao and Shin[9] modified the DCF and presented a priority-based fair MAC (P-MAC) protocol which selects the minimum contention window (CWmin) size for each wireless station to reflect the relative weights among data traffic flows. Banchs and Perez[10] proposed a distributed weighted fair queuing (DWFQ) as an extension of the DCF to provide weighted fair queuing in WLANs. Vaidya et al.[11][12] presented a distributed fair scheduling (DFS) approach by modifying the DCF to allocate bandwidth in proportion to the weights of the flows sharing the channel. Gao et al. [13] improved the VoIP capacity of WLANs by adjusting the EDCA parameters of the AP. Jeong et al. [14] proposed a scheme that uses different means of backoff distribution to achieve weighted fairness between uplink and downlink. A token based piggyback scheme and a dynamic ratio adjustment based piggyback scheme were also proposed in [14] and [15], respectively.

In this paper, we first develop an analytical model to investigate the uplink and downlink throughput unbalance problem of the multi-user MIMO-based WLANs. Second, in order to solve the unbalance problem, we modify and adopt a CWmin adjustment scheme and a random piggyback scheme, which were typical throughput balancing solutions in conventional WLANs, to the multi-user MIMO-based WLANs, and discuss the suitableness of each scheme for the multi-user MIMO-
A multi-user MIMO-based collision mitigation scheme in uplink WLANs.

The rest of this paper is organized as follows: In Section II, a multi-user MIMO-based WLAN system is described and the corresponding performance analysis model is proposed. In Section III, in order to solve the unbalance problem, using a desired downlink and uplink throughput ratio, a CWmin adjustment scheme and a random piggyback scheme are adopted and their throughput performance is evaluated. Finally, conclusions are presented in Section IV.

II. MULTI-USER MIMO-BASED WLANS

A. System Model

Various MIMO transmission techniques have recently been adopted in wireless systems in order to achieve high spectral efficiency. For WLANs, the IEEE 802.11n specification was standardized to adopt a single-user MIMO scheme[16]. Although the single-user MIMO scheme can support high transmission data rate, it cannot solve a collision problem among STAs during frame transmissions in WLANs and the performance enhancement is limited. In order to mitigate the collision problem, Jin et al.[2] proposed a multi-user MIMO-based collision mitigation scheme which adopts a multi-user MIMO scheme to resolve the collisions at physical layer. The performance between this collision mitigation scheme and the legacy CSMA/CA protocol needs to be slightly modified. As shown in Fig. 2, the AP can correctly recover the two transmitted data and it has to send two ACK frames consequently to each STA within the SIFS time. Since the data transmission time for each STA is different, the ACK frames consequently have to be sent after the complete transmission of both data frames. If the receiver can recover the data streams, it has to send $N$ ACK frames to each STA. In this paper, each STA is assumed to have a unique preamble chosen from an orthogonal preamble sequence set. Hence, the AP can estimate the channel coefficients of each STA.

B. Performance Analysis

Bianchi [5] proposed a simple discrete-time Markov chain (DTMC) model to compute the throughput in a saturation traffic environment for conventional IEEE 802.11 WLANs. Based on this DTMC model, we extended it and analyzed the performance of the multi-user MIMO-based collision mitigation scheme in uplink WLANs [2][3]. In this section, we further extend the analysis by considering uplink and downlink traffic. We consider a network environment where AP is located at the center of a basic service set (BSS) and $n$-contending STAs communicate with the AP. In order to evaluate the performance in MAC layer, we simply assume that there is no transmission error except collisions, and the payload size and its transmission time are identical for each STA and AP.

For each STA, let $\tau$ be the transmission probability and $p$ be the backoff stage transition probability that the STA retransmits the previously transmitted frame. Since there are $M$ receiver antennas for simplicity. This model can be easily extended to the case of multiple transmit antennas at each STA. If the AP can estimate the channel coefficients, it can recover the transmitted data streams from different STAs by using MIMO decoding techniques, such as zero forcing, minimum mean square error and maximum likelihood. Fig. 1(a) shows an example of a multi-user MIMO-based collision mitigation scheme in case of $M = 2$ and $N = 2$.

On the other hand, Fig. 1(b) shows a collision between AP and STA1. Since the AP is in transmit state, it cannot recover the data stream from the STA1. Furthermore, the destination, STA2, of the data stream from the AP, also cannot recover the data stream because it only has one receiver antennas. Thus, in the case of the simultaneous transmission between the STA and the AP, it should be considered as a collision. This collision between uplink and downlink traffic is not considered in the previous work [1]-[3] and we will analyze it in this paper.

In order to support simultaneous transmissions of STAs, the modified CSMA/CA protocol needs to be slightly modified. As shown in Fig. 2, the AP can correctly recover the two transmitted data and it has to send two ACK frames consequently to each STA within the SIFS time. Since the data transmission time for each STA is different, the ACK frames consequently have to be sent after the complete transmission of both data frames. If the receiver can recover $N$ simultaneous data streams, it has to send $N$ ACK frames to each STA. In this paper, each STA is assumed to have a unique preamble chosen from an orthogonal preamble sequence set. Hence, the AP can estimate the channel coefficients of each STA.
no channel errors, the backoff stage transition probability is the
same as the collision probability. If an AP can receive up to maximum \( N \) uplink simultaneous transmissions, a transmis-
sion failure occurs when there are more than \( N \) simultaneous transmissions. From the viewpoint of a given
backoff procedure with retry limit \( R \), we can obtain the relationship between \( \tau \) and \( p \), based on the DTMC model in [5],

\[
\tau = \frac{2(1 - p^{R+1})}{W(1 - 2^{L}p^{R+1}) + Wp[\sum_{i=0}^{L-1} (2p)^{i}] + (1 - p^{R+1})},
\]

(1)

where \( W \) represents the minimum contention window size \( CW_{\text{min}} \). The term \( L \) is the maximum number of doublings of the \( CW \), which is identical to \( \log_{2}((CW_{\text{max}} + 1)/(CW_{\text{min}} + 1)) \).

Let \( \tau_{0} \) and \( p_{0} \) be the AP’s transmission probability and stage transition probability, respectively. With the same reason as for
the STA, we can obtain

\[
\tau_{0} = \frac{2(1 - p_{0}^{R+1})}{W(1 - 2^{L}p_{0}^{R+1}) + Wp_{0}[\sum_{i=0}^{L-1} (2p_{0})^{i}] + (1 - p_{0}^{R+1})},
\]

(2)

As shown in Fig. 1(b), since the transmission of AP is only successful when there is no transmission of the STAs, a stage transition occurs for the AP with the following probability if any STA transmits data to the AP:

\[
p_{0} = 1 - (1 - \tau)^{n},
\]

(3)

If the AP can receive up to maximum \( N \) simultaneous transmissions from \( N \) STAs together, for each STA, a backoff stage transition can occur when there are more than \((N - 1)\) STAs transmitting data at the same time among remaining \((n - 1)\) STAs. Furthermore, a simultaneous transmission from any STA also fails if AP is in transmit state, as shown in Fig. 1(b). Thus, the relationship between \( \tau, \tau_{0} \) and \( p \) from this viewpoint is,

\[
p = \tau_{0} + (1 - \tau_{0}) \left[ 1 - \sum_{m=0}^{N-1} \binom{n - 1}{m} \tau^{m}(1 - \tau)^{n-1-m} \right],
\]

(4)

where the probability that \( m \) STAs transmit frames simultaneously among the remaining \((n-1)\) STAs is expressed as \( \binom{n-1}{m} \tau^{m}(1 - \tau)^{n-1-m} \).

We can numerically solve Eqs. (1) - (4) for the transmission probabilities \( \tau \) and \( \tau_{0} \), and backoff stage transition probabilities \( p \) and \( p_{0} \). From these obtained values, \( \tau \) and \( \tau_{0} \), the system performance can be calculated. First, the probability \( P_{\tau_{0},\text{STA}} \) that there is transmission from either STA or AP in a time slot is expressed as

\[
P_{\tau_{0},\text{STA}}^{m} = 1 - (1 - \tau_{0})(1 - \tau)^{n}.
\]

(5)

The probability \( P_{\tau_{0},\text{STA}}^{m} \) that there are simultaneous transmissions from \( m \) STAs in a time slot is written as

\[
P_{\tau_{0},\text{STA}}^{m} = (1 - \tau_{0}) \binom{n}{m} \tau^{m}(1 - \tau)^{n-m}.
\]

(6)

The probability \( P_{\tau_{0},\text{AP}}^{m} \) that there is transmission only from AP in a time slot is expressed as

\[
P_{\tau_{0},\text{AP}}^{m} = \tau_{0}(1 - \tau)^{n}.
\]

(7)

The collision probability \( P_{c}^{m} \), in which there are simultaneous transmissions from more than \( N \) STAs in a time slot or there is a collision between AP and STAs, is written as

\[
P_{c}^{m} = (1 - \tau_{0}) \sum_{i=0}^{m} \frac{m!}{n!} \tau_{0}^{i} \left( 1 - \tau_{0} \right)^{n-i}.
\]

(8)

The average payload size for a uplink successful transmission in data transmission time is obtained as

\[
E_{\text{ul}[\text{payload}]} = \sum_{m=1}^{N} m \cdot N_{\text{tr,STA}}^{m} \cdot E[\text{Payload}],
\]

(9)

where \( E[\text{payload}] \) is the average payload size for each STA. The average payload size for a downlink successful transmis-
sion in the data transmission time is expressed as

\[
E_{\text{dl}[\text{payload}]} = P_{\tau_{0},\text{AP}}^{m} E[\text{Payload}].
\]

(10)

Then, the average uplink and downlink throughput can be obtained as

\[
\text{Throughput}_{ul} = \frac{E_{\text{ul}[\text{payload}]} \cdot E[\text{length of a time slot}]}{E[\text{length of a time slot}]},
\]

(11)

\[
\text{Throughput}_{dl} = \frac{E_{\text{dl}[\text{payload}]} \cdot E[\text{length of a time slot}]}{E[\text{length of a time slot}]},
\]

(12)

where \( E[\text{length of a time slot}] \) is the average length of a
time slot which can be backoff slot time, successful data transmission time, or collision time. It can be expressed as

\[
E[\text{length of a time slot}] = (1 - P_{c}^{m}) \sigma + P_{\tau_{0},\text{AP}} \cdot T_{\tau_{0},\text{AP}} + \sum_{m=1}^{N} P_{\tau_{0},\text{STA}}^{m} T_{\tau_{0},\text{STA}}^{m} + P_{c}^{m} \cdot T_{c},
\]

(13)

where \( \sigma \) is the backoff slot time. \( T_{\tau_{0},\text{AP}} \) is the data frame transmission time at the AP and it is expressed as

\[
T_{\tau_{0},\text{AP}} = \text{DataTime} + \text{SIFS} + \text{ACKtime} + \text{DIFS}.
\]

(14)

\( T_{\tau_{0},\text{STA}}^{m} \) is the time duration used to simultaneously transmit \( m \) frames including overhead and it is expressed as

\[
T_{\tau_{0},\text{STA}}^{m} = \text{DataTime} + m \cdot (\text{SIFS} + \text{ACKtime}) + \text{DIFS}.
\]

(15)

\( T_{c} \) is the time duration used due to a collision through simultaneous transmissions from more than \( N \) STAs or the simultaneous transmission between AP and STAs, and it is expressed as

\[
T_{c} = \text{DataTime} + \text{ACKtimeout} + \text{DIFS}.
\]
Table I lists the MAC layer parameters based on the IEEE 802.11a[17] specifications. We evaluate the performance of the multi-user MIMO-based WLAN system in the case of saturated traffic environment. The data rate is fixed to a single rate of 24Mbps and the frame size is 1000bytes.

Fig. 3 shows the uplink and downlink throughput performance obtained from both analysis and simulation. The number of receiver antennas at the AP varies from 1 to 3. The analytical results agree well with simulation results. In uplink, the multi-user MIMO-based collision mitigation scheme yields much better performance than the conventional WLAN(N = 1). The throughput enhancement is more significant as the number of receiver antennas at the AP increases. However, in downlink, the scheme shows lower throughput performance than the conventional WLAN as the number of receiver antennas at the AP increases. Since, the simultaneous transmission from both the AP and STA is considered as a collision in downlink, as shown in Fig. 1(b), there are more collisions at the AP and, consequently, AP has less successful transmission opportunities compared to the STAs.

Fig. 4 shows the transmission probability and collision probability for the AP and STA when there are two receiver antennas at the AP. The AP has lower transmission probabilities and higher collision probabilities than STA.

Fig. 5 shows the throughput performance ratio of downlink to uplink. In the case of the conventional WLANs(N = 1), since AP has the same channel access probability as each STA, the ratio of downlink to uplink is 1/n. However, when it comes to the multi-user MIMO-based collision mitigation scheme, it becomes significantly low when there are more than 10 STAs. Thus, there is a severe unbalance problem between the uplink and downlink for the multi-user MIMO-based collision mitigation scheme and, consequently, balancing the throughput of the uplink and downlink is more important for the multi-user MIMO-based collision mitigation scheme.

III. Throughput Balancing Schemes for Multi-user MIMO-based WLANs

A. CWmin Adjustment Scheme

Since the AP and STAs have the same channel access opportunities based on DCF in infrastructure-based conventional WLANs, AP has a reduced channel access ratio of 1/n compared with the n STAs when there are n STAs in a BSS. This unbalance problem have gotten attentions [9]-[14]. Most of the previous work were based on adjusting the DCF parameters. In this section, based on the analysis model in Section II-B, we attempt to find the appropriate CWmin size.
calculate the CWmin size of each STA, values into Eq. (15), we can
obtain from Eq. (1), we can
AP also can access the channel just after the transmissions
B. Random Piggyback Scheme
addresses from the transmitting frames at the BSS or by other
the number of active STAs by observing the unique MAC
conventional WLANs.
user MIMO-based collision mitigation scheme, the slope of
the CWmin size over the number of STAs is steeper than
in the multi-user MIMO-based WLANs.
In order to adjust the CWmin size at STAs, we can estimate
from STAs in SIFS intervals. If there are simultaneous
transmissions from m STAs in uplink, the AP has mq random
transmission opportunities in which the AP can transmit its
first ⌊mq⌋ frames and one additional frame with probability
mq − ⌊mq⌋. Here ⌊x⌋ denotes the largest integer value that
is equal to or smaller than x. The uplink and downlink
throughput performance ratio can be controlled by adjusting
the parameter q. An example of frame exchange between
uplink and downlink is shown in Fig. 7. There are two
STAs’ simultaneous transmissions, and, correspondingly, the
AP has 2q random transmission opportunities after the STAs’
transmissions. In Fig. 7, since ⌊2q⌋ = 1 and the randomly
generated variable is smaller than 2q − ⌊2q⌋, the AP transmits
two frames.

The analysis model of the random piggyback scheme for
the multi-user MIMO-based WLANs has to be modified from
the model in Section II-B. The AP has more transmission
opportunities due to the random piggyback scheme and the
average payload size in a data transmission time in downlink is

\[ E_d' \cdot \text{payload} = \left( P_{tr,AP} + \sum_{m=1}^{N} q \cdot m \cdot P_{tr,STA}^m \right) \cdot E[\text{payload}] . \]

(16)

On the other hand, there is no change in uplink average
payload size. Thus, the downlink and uplink throughput
performance ratio is

\[ \psi = \frac{1}{n} \frac{\tau_0 (1 - p_0)}{\tau (1 - p)} + q . \]

(17)

The AP can adjust the piggyback probability q to achieve the
desired ratio \( \psi \). In case of \( N = 1 \), since \( \tau_0 = \tau \) and \( p_0 = p \),
we have \( q = \psi - \frac{1}{n} \). In case of \( N \geq 2 \), since \( \frac{1}{n} \frac{\tau_0 (1 - p_0)}{\tau (1 - p)} \to 0 \)
as shown in Fig. 5, the AP can approximate \( q \approx \psi \). Thus,
in the conventional WLANs, in order to achieve the desired
ratio, the AP has to be aware of the number of active STAs in
the BSS. However, in the multi-user MIMO-based WLANs,
the ratio also can be achieved without the knowledge about
the number of active STAs.

Moreover, the random piggyback scheme modifies the
duration of the channel access time when there are m
simultaneous transmissions in the uplink.

\[ T_{tr}^m = \text{DateTime} + m \cdot (\text{SIFS} + \text{ACKtime}) + \text{DIFS} \\
+ q \cdot m \cdot (\text{SIFS} + \text{DateTime} + \text{SIFS} + \text{ACKtime}) . \]

With above modifications, the throughput performance can be
obtained with Eqs. (11) and (12).

C. Numerical Result

The system parameters are set to the same as in Section II-
C. Fig. 8 shows the total throughput performance of both the
CWmin adjustment scheme and the random piggyback scheme
when the desired throughput ratio between downlink and up-
link is one. For the random piggyback scheme, the throughput
enhancement is significant with increasing the number of
receiver antennas at the AP, compared to the conventional
transmitting property of the multi-user MIMO-based collision random piggyback scheme does not change the simultaneous STAs adjust their CWmin sizes to a large value as shown in Fig. 6, and, thus, there is a small number of simultaneous achieve the throughput ratio between downlink and uplink, the scheme, the throughput enhancement is very small. In order to WLANs. However, when it comes to the CWmin adjustment this unbalance problem, we adopted the CWmin adjustment W/L WLANs, it works more efficient than the CWmin inherits the collision mitigation property of the multi-user MIMO-based WLANs. Since the random piggyback scheme adjustment scheme. In this paper, we considered the saturated traffic case and we leave the unbalance problem in a non-saturated traffic case for further studies.

IV. CONCLUSIONS

In this paper, we investigated the uplink and downlink throughput performance of a multi-user MIMO-based WLAN through both analysis and simulation. With the numerical result, we explained a significant throughput unbalance problem between uplink and downlink in the multi-user MIMO-based WLAN. Since the multi-user MIMO-based collision mitigation scheme is applied for the uplink, the AP has less channel access opportunities, compared to each STA. In order to solve this unbalance problem, we adopted the CWmin adjustment scheme and the random piggyback scheme for the multi-user MIMO-based WLANs. Since the random piggyback scheme inherits the collision mitigation property of the multi-user MIMO-based WLANs, it works more efficient than the CWmin adjustment scheme. In this paper, we considered the saturated traffic case and we leave the unbalance problem in a non-saturated traffic case for further studies.

ACKNOWLEDGMENT

This research was supported in part by the BroMA IT Research Center supervised by the Institute of Information Technology Advancement. (IITA-2008-C1090-0801-0037)

REFERENCES