A Distributed Fair Auto Rate Medium Access Control for Wireless LANs

Yanfeng Zhu and Zhisheng Niu

State Key Laboratory on Microwave and Digital Communication
Department of Electronic Engineering, Tsinghua University
Beijing 100084, P. R. China

Abstract—The IEEE 802.11 protocol provides a physical multi-rate capability. Existing auto rate schemes concentrate on passively tuning the rate according to the channel strength, but cannot exploit the time-variation of fading channel. In this paper, by means of an analytical model, we derive the rule maximizing the throughput of RTS/CTS based multi-rate WLANs. Then, based on the obtained rule we propose a distributed fair auto rate medium access control scheme. With the proposed scheme, after receiving a RTS frame, the receiver selectively returns the CTS frame according to both the signal-to-noise ratio of the RTS frame and current network status. Moreover, the returned CTS carries a piggyback information indicating the maximum feasible rate. The key feature of the proposed scheme is that it is capable of positively exploiting time-varying fading channel and maintaining fairness in asymmetric situation where the distribution of SNR varies with stations. Extensive simulation results show that the proposed scheme outperforms the existing fairness schemes in time-varying channel conditions.

I. INTRODUCTION

In multi-user scenario, multi-rate WLANs have potential capability to exploit the time-variation of the wireless channel strengths caused by the effect of Doppler shifts as well as larger scale effects such as path loss and shadowing. Therefore, they can benefit from carefully designed auto rate medium access control.

For cellular networks, an “opportunistic” downlink scheduling [3], which preferentially permits the stations with good channel gain to transmit, can benefit from the time-variation of channel by exploiting multi-user diversity. In the proposed scheme, a centralized control is required for collecting the fading level of all stations and scheduling transmissions. WLANs are based on distributed medium access control mechanism, and it is very expensive to employ a centralized control for uplink especially for independent structure where no access points are available. Existing WLANs based auto rate schemes (ARF in [1], RBAR in [2], and OAR in [5]) concentrate on passively tuning the data rate according to the channel strength, but few schemes can positively exploit the time-variation of channel.

Among the few, the authors in [6] proposed a Medium Access Diversity (MAD) based auto rate scheme, which employs a group RTS to measure the channel conditions of multiple candidate receivers and select the receiver with the best signal strength. However, if each sender has only one candidate receiver, which is the main scenario in practical WLANs, MAD cannot benefit from the time-variation of channel again. Furthermore, many modifications should be made to current IEEE 802.11 protocol when applying MAD in multi-rate WLANs. In our previous works [7] and [8], based on the assumption that each station knows the channel distributions of all stations, it is demonstrated that the overall throughput can be improved by permitting the transmitter to send data frames only when the signal-to-noise ratio (SNR) of channel exceeds a threshold. However, the proposed scheme starves the stations with bad channel distributions, and cannot provide fairness in asymmetric channel scenarios.

In this paper, we relax the assumption made in [7] [8] and develop an analytical model to investigate the rule maximizing the throughput of RTS/CTS based multi-rate WLANs. Based on the analytical results, we propose a Fair Auto Rate Medium access control scheme (FARM) to dynamically tune the threshold according to the network status sensed by CSMA mechanism. In FARM, after receiving a RTS frame, the receiver selectively return the CTS frame only when the channel strength measured by the RTS frame exceeds a threshold, which is dynamically tuned according to network status sensed by CSMA. Herein, the network status refers to the overall throughput of the network and the probability that receivers return CTS frames after receiving RTS frames. The main feature of FARM is that it is capable of positively exploiting the time-variation of channel and maintaining fairness in asymmetric time-varying fading channels.

The rest of this paper is organized as follows. In Section II, we develop an analytic model to investigate the rule maximizing the throughput. Then, FARM is described in Section III. The performance of the proposed scheme is evaluated by simulations in Section IV. Finally, the conclusions are drawn in Section V.

Before delving into the details of our work, we first illuminate some terminology used in this paper. Successful transmission refers to the event that a receiver receives a RTS frame successfully and then returns the CTS frame, which implies the upcoming data frame will be transmitted successfully. Physical collision refers to the event that more than one stations send RTS frames simultaneously. While virtual collision refers to the event that the receiver receives a RTS frame but decides not to return the CTS frame.
II. MAXIMIZING OVERALL THROUGHPUT OF RTS/CTS
BASED MULTI-RATE WLANs

A. System model

We consider single cell RTS/CTS based IEEE 802.11 WLAN with \( N \) stations. Each station is always backlogged, which implies that each station always has packets available for transmission. All stations are within radio range of each other, and thus each station can sense all packets transmitted in the network. Each station randomly select a station as its receiver, and the channel between each station and the receiver is modeled as a time-varying, memoryless, and block-fading channel. The time-varying Signal-to-Noise Ratio (SNR), \( \gamma_i \), between station \( i \) and its receiver lies in \( 0 < \gamma_i < \infty \), and its probability density function of \( \gamma_i \) is \( f_i(\gamma_i) \). Clearly, it is feasible for the receiver obtains \( f_i(\gamma_i) \) by statistical measurements. To enhance the reliability of control information, all control information like RTS, CTS, and ACK frames are transmitted at a low base rate denoted by \( R_c \), which is assumed to be error free in any channel conditions.

The upper part of Fig.1 shows a typical transmission process in traditional RTS/CTS based WLANs. As shown in the figure, the exchange of RTS and CTS frames is prior to the actual data frame. From the standards, both RTS and CTS frames contain a Duration/ID field that defines the period of time that the channel is to be reserved for the upcoming data frame and ACK frame. After hearing RTS and CTS frames, other stations shall learn of the channel reservation and suspend their backoff counters. If a RTS frame is collided (see the third transmission), all stations restart their backoff counters after a DIFS period.

For auto rate schemes, the RTS frame play a new role that detecting channel conditions. Correspondingly, the CTS frame is ended with an additional functionality informing transmitters the feasible rate for the upcoming data frame, and is returned upon receiving the RTS frame regardless of the channel conditions. Alternatively, we make the receiver return the CTS frame only when the SNR of the RTS frame exceeds a threshold, which depends on the current network status. If the receiver decides not to return the CTS frame, other stations heard the corresponding RTS frame can detect a virtual collision because there are no frames transmitted after a RTS frame. In contrast, if the receiver returns the CTS frame, other stations will suspend their backoff counters until the transmitter finishes the data transmission.

B. Maximizing Overall Throughput

From the upper part of Fig.1, the time interval between the epoch that the receiver receives the RTS frame from station \( i \) and the epoch that station \( i \) completes its data transmission is given by

\[
T(\gamma_i) = \frac{L + H}{R(\gamma_i)} + \frac{L_{cts} + L_{ack}}{R_c} + 3t_{phy} + 3t_{sifs} \tag{1}
\]

where \( L \) and \( H \) are respectively the length of payload and MAC frame header, \( L_{cts} \) and \( L_{ack} \) are respectively the sizes of a CTS frame and a ACK frame, \( t_{phy} \) is the transmission time for a physical overhead, \( t_{sifs} \) is a SIFS period, \( R(\gamma_i) \) is the maximum feasible rate when the SNR is \( \gamma_i \). As mentioned above, the receiver returns the CTS frame only when \( \gamma_i \) exceeds a threshold denoted by \( \Gamma_i \). Clearly, if the receiver returns the CTS frame, the time interval \( T(\gamma_i) \) will be occupied exclusively by station \( i \). Then, the instant throughput in this time interval is given by \( \frac{L}{T(\gamma_i)} \). Herein, the throughput is defined as the payload information transmitted in a unit time. Alternatively, if the receiver decides not to return the CTS frame, this time interval will be shared by all stations for new contention. Then, the potential throughput in this time interval should be the throughput averaged in a long time interval. Therefore, the choice of \( \Gamma_i \) directly influences the overall throughput. The following proposition states the rule for choosing \( \Gamma_i \) to maximize the overall throughput.

**Proposition 1** In the \( p \)-persistent IEEE 802.11 protocol based multi-rate WLANs with \( N \) transmitters, if the date rate increases with the SNR, the overall throughput \( S \) reaches to the maximum if and only if both

1. \( \Gamma_1 = \Gamma_2 = \ldots = \Gamma_N = \Gamma \), and
2. \( S = L/T(\Gamma) \),

hold.

**Proof:** Reduction to absurdity is employed. Firstly, it is assumed that the maximum throughput is \( S_m \), and condition 1 does not hold when the throughput reaches to \( S_m \). Without loss of generality, it is assumed that \( \Gamma_1 \leq \Gamma_2 \leq \ldots \leq \Gamma_N \), and then there exists at least one \('<\'\) that can be replaced by \('<\'\).

Therefore, we have \( \Gamma_1 < \Gamma_N \). Since \( T(\gamma_i) \) is a monotonically decreasing function with respect to \( \gamma_i \), we have

\[
L/T(\Gamma_1) < L/T(\Gamma_N)
\]

Then, at least one of the two inequalities below hold

\[
\begin{cases}
S_m > L/T(\Gamma_1) \\
S_m < L/T(\Gamma_N)
\end{cases}
\tag{2}
\]

If \( S_m > L/T(\Gamma_1) \) holds, there must exists a time interval which includes a successful transmission of station 1 with \( \Gamma_1 \leq \gamma_1 < T^{-1}(L/S_m) \). As shown in the upper part of Fig.1 (for Data 2), the receiver returns the CTS frame to station 1 when \( \Gamma_1 \leq \gamma_1 < T^{-1}(L/S_m) \), and station 1 then sends Data 2. If the receiver increases \( \Gamma_1 \) to \( T^{-1}(L/S_m) \), it will decide not to return the CTS frame. Then, a virtual collision occurs. Due to memoryless of the \( p \)-persistent model, this change does not influence the channel contention process of other stations. Clearly, if the receiver increases \( \Gamma_1 \) to \( T^{-1}(L/S_m) \), the new transmission process will change to the lower part of Fig.1. Since \( S_m > L/T(\gamma_i) \) with \( \Gamma_1 \leq \gamma_1 < T^{-1}(L/S_m) \), the throughput in lower part of Fig.1 is larger than that in the upper part of Fig.1, which is inconsistent with that the maximum throughput is \( S_m \).

\[^1\text{In [10][11], the authors demonstrated that the p-persistent protocol closely approximates the standard protocol with the same average backoff window size. Therefore, it is reasonable to infer the behavior of the standard protocol from the analytical results based on the p-persistent protocol.}\]
Similarly, if $S_m < L/T(\Gamma_N)$ holds, the throughput can be increased as well by permitting the receiver to return the CTS frame to station $N$ when $T^{-1}(L/S_m) \leq \gamma_N < \Gamma_N$. In a word, if the throughput reaches to the maximum, both conditions holds. The proof of the inverse proposition is similar to that above and is omitted.

C. Capacity Analysis

Herein, we take a simple scenario, where all channels are independently and identically distributed with the probability density function $f(\cdot)$, for example to investigate the capacity of RTTs/CTS multi-rate WLANs. From proposition 1, when the throughput reaches to the maximum, the thresholds for all stations are the same. Let $\Gamma$ denote the threshold. Then, the probability, $q$, that the receiver returns the CTS frame after receiving a RTS frame (called return probability in the remainder of this paper) is given by

$$q = \int_\Gamma^\infty f(\gamma)d\gamma$$

(3)

In the $p$-persistent model, the probability that a station sends a RTS frame in a time slot is denoted by $p$. Let $P_f$ denote the probability that a time slot is idle, and $P_S$ and $P_C$ denote the probabilities that a RTS transmission occurring on the channel is successful and suffers a physical collision, respectively. Then, we have

$$\begin{cases} P_f = (1 - p)^n \\ P_S = np(1 - p)^{n-1} \\ P_C = 1 - P_f - P_S \end{cases}$$

(4)

Let $E[T_s]$ denote the average time spent for a successful transmission. Then, given $\Gamma$, the expected value of $T_s$ is given by

$$E[T_s] = \frac{\int_\Gamma^\infty (T(\gamma) + L_{rts}/R_c + t_{phy})f(\gamma)d\gamma}{q}$$

(5)

where $L_{rts}$ is the RTS frame size. On the other hand, Let $T_{cp}$ and $T_{cv}$ denote the time that the channel being captured due to a physical collision and the time for a virtual collision, respectively. Then, we have

$$\begin{align*}
T_{cp} &= \frac{L_{rts}}{R_c} + t_{phy} + t_{diff} \\
T_{cv} &= \frac{L_{rts} + L_{cts}}{R_c} + 2t_{phy} + t_{sifs} + t_{diff}\end{align*}$$

(6)

From the analysis above, the probability that a successful transmission occurs in a time slot is $P_Sq$. Similarly, the probability that a virtual collision occurs in a time slot is given by $P_S(1 - q)$. The throughput $S$ is thus given by

$$S = \frac{P_SqL}{P_SqE[T_s] + P_CT_{cp} + P_S(1 - q)T_{cv} + P_f\delta}$$

(7)

where $\delta$ denotes a back-off slot time. From proposition 1, the maximal throughput conforms to

$$S = \frac{L}{T(\Gamma)}$$

(8)

Given $p$, we can get $\Gamma$ which leads to the local maximal throughput from (7) and (8). Then, instituting it into (8), we derive the corresponding throughput. Furthermore, we can derive the throughput capacity through searching for the optimal $p$ which results in the global maximal throughput in the region $(0, 1]$. Herein, taking IEEE 802.11a MAC protocol for example, we give the numerical results over Rayleigh fading channels. We focus on the throughput capacity of MAC protocol, and thus the physical layer is assumed to be ideal. From Shannon’s law, in the ideal condition, the data rate follows $R(\gamma) = W \log_2(1 + \gamma)$ where $W=16.6$ MHz is the system bandwidth of IEEE 802.11a protocol. However, to enhance the reliability of control information, 9 Mbps is used for the transmission of all control information. Some parameters of IEEE 802.11a protocol are listed in Table I.

Fig.2 shows the capacity versus the number of stations for various payload sizes in Rayleigh fading channel with the average SNR as 20 dB. In particular, when there is a single station in the network, the throughput reaches the maximum...
with \( p = 1 \). In the meantime, the channel utilization is the highest due to no collision and idle. Generally, the throughput should markedly decreases with the number of stations due to the increment of collisions. However, as shown in the figure, a surprising result is that the capacity slowly decreases to a stable level as the number of stations increases. This is mainly due to two reasons. One is that the selection of optimal \( p \), which is dependent of the number of stations, can balance the costs for collisions and idles. The other is that from (7) the throughput can be improved by adapting \( q \) to the number of stations, which allows only the stations with high channel gain to transmit. Therefore, the dynamic tuning of \( q \) can bring multi-user diversity gain.

In IEEE 802.11 protocol, fairness is maintained by configuring the minimum contention window and the maximum contention window of all stations with the same values, which leads to the probabilities that stations transmit in a time slot are the same. Similarly, in FARM, the fairness is achieved by keeping the return probability of each station as the same value and providing same transmission time for each successful transmission. In this way, the time-share fairness can be guaranteed. Herein, the return probability, \( q \), refers to the probability that the receiver returns the CTS frame after receiving a RTS frame, i.e.

\[
q = \frac{\text{Number of CTS frames received in a unit time}}{\text{Number of RTS frames received in a unit time}}.
\]  

(9)

Due to the employment of carrier sense multiple access (CSMA) mechanism, it is very easy for the receiver to obtain the return probability averaged among all stations in the condition that all stations are within radio range of each other.

The primary step of FARM is to estimate the overall throughput \( S \) and the average return probability \( q \). For this issue, it is considered in two scenarios. If all stations stay awake all the time, \( S \) and \( q \) can be independently measured on-line by all stations with CSMA. On the other hand, if there exist some stations working in the power-saving mode, from the standard [13] an AP is strictly required to wake up the power-saving stations periodically. In such a scenario, the AP is responsible for estimating \( S \) and \( q \) and informing the power-saving stations when waking up the power-saving stations.

To better track the changes of \( S \) and \( q \), in FARM, the receiver updates the estimates at intervals. Extensive simulations show that 100 ms is a proper value for the interval, with which the channel variation can be tracked on-line. Let \( S(k) \) and \( q(k) \) denote the updated values of \( S \) and \( q \) after the \( k \)th interval, respectively. During the following interval, the receiver gets the statistics on the throughput and the return probability as \( \hat{S} \) and \( \hat{q} \), respectively. Then, we update \( S \) and \( q \) after the \( (k+1) \)th interval as follow

\[
\begin{align*}
S(k+1) &= \alpha S(k) + (1 - \alpha) \hat{S} \\
q(k+1) &= \alpha q(k) + (1 - \alpha) \hat{q}
\end{align*}
\]  

(10)

where \( \alpha \) is a smoothing factor, which is widely adopted in the network protocols to obtain reliable estimates. Extensive simulations show that \( \alpha = 0.9 \) is a good choice when the network conditions change because it potentially reduces the length of transient phases, and hereby we adopt \( \alpha = 0.9 \) as the default value in this paper.

Different from the analysis in the previous section, in practical standards [13][14] the feasible rates are just limited discrete values. Without loss of generality, we assume that there are \( M \) feasible rates \( (R_1 < R_2 < ... < R_M) \) and for \( R_i \) there will be error free when the SNR is larger than \( G_i \). To keep time-share fairness, each station is granted to transmit multiple back-to-back frames in proportion to the ratio of the achievable rate over the base rate. In this case, the time interval between the epoch that the receiver receives the RTS
frame from station $i$ and the epoch that station $i$ completes its transmission is given by
\[ T^\ast(\gamma_i) = \frac{R^\ast(\gamma_i)}{R_c} \left[ L + H + \frac{L_{\text{ack}}}{R_c} + 2t_{\text{phy}} + 2t_{\text{sifs}} \right] + \frac{L_{\text{ets}}}{R_c} + t_{\text{phy}} + t_{\text{sifs}} + t_{\text{disfs}} \]
\[ (11) \]

Note that the value of $\frac{R^\ast(\gamma_i)}{R_c}$, which indicates the number of data frames transmitted in a transmission opportunity, is usually not a integer. A feasible method to hold time-share fairness is randomly transmitting $\lceil \frac{R^\ast(\gamma_i)}{R_c} \rceil$ or $\lceil \frac{R^\ast(\gamma_i)}{R_c} \rceil - 1$ frames with respective probability $\frac{\frac{R^\ast(\gamma_i)}{R_c}}{\left[ \frac{R^\ast(\gamma_i)}{R_c} \right]}$ and $\frac{\left[ \frac{R^\ast(\gamma_i)}{R_c} \right] - 1}{\left[ \frac{R^\ast(\gamma_i)}{R_c} \right]}$, where $\lceil x \rceil$ and $\lfloor x \rfloor$ is respectively the largest integer no greater than $x$ and the smallest integer no less than $x$. For example, $\frac{R^\ast(\gamma_i)}{R_c} = 1.5$ can be approached by sending 2 frames with probability 0.5 and 1 frames with probability 0.5. Correspondingly, the decision-making mechanism of the receiver is shown in Fig.3. Herein, $\theta(\gamma) = \frac{R^\ast(\gamma)}{R_c}$ for simplicity,

![Fig. 3. The operation of the receiver after receiving a RTS frame in FARM](image)

and $\text{Rand}(0, 1)$ return a float number uniformly distributed in $[0, 1]$.

IV. PERFORMANCE EVALUATION

In this section, we take IEEE 802.11a based WLANs for example to evaluate the performance of the proposed FARM as compared to the existing throughput fairness schemes in various channel conditions with NS-2 [16] simulator. The physical layer characteristics used in the simulations follow the standard, where the SNR required for each data rate is shown in Table II [15]. Note that the required SNR’s for 54 Mbps and 48 Mbps are very close, which implies that the data rate of 54 Mbps is more efficient than that of 48 Mbps. Therefore, only the data rate of 54 Mbps is considered in the simulations. Similarly, we choose 36 Mbps from the pair of 36 and 24 Mbps, 18 Mbps from the pair of 18 and 12 Mbps, and 9 Mbps from the pair of 9 and 6 Mbps. In addition, 9 Mbps is also used for the transmissions of control information. The minimum contention window and the maximum contention window are respectively set to 31 and 1023, and other parameters about IEEE 802.11a protocol can be found in [14].

In the simulations over time-varying fading channels, the time-varying SNR associated with the Rayleigh fading, i.e. the probability density function of the SNR is given by
\[ f(\gamma) = \frac{1}{\gamma} e^{-\frac{1}{\gamma}}, \quad 0 < \gamma < \infty \]
\[ (12) \]
where $\bar{\gamma}$ is the average SNR.

To analyze the convergence of the proposed FARM, we first explore the adaption of FARM to the time-variation of Rayleigh fading channels. We simulate a scenario where 20 stations suffer from independent and identical Rayleigh fading channel with various average SNR’s. Fig.4 show the convergence of a single station’s throughput. From the figures, it is observed that for various average SNR FARM can effectively bound the throughput in a stable oscillation range. Therefore, the proposed scheme has good convergence.

The overall throughput versus the average SNR in Rayleigh fading channel is first simulated. For the purpose of comparison, we also simulate OAR [5] which is a typical time-share fairness scheme. A scenario with 20 stations is considered. Herein, all channels are statistically independent and identical Rayleigh fading. The results are shown in Fig.5. From the

<table>
<thead>
<tr>
<th>Rate (Mbps)</th>
<th>SNR (dB)</th>
<th>Modulation</th>
<th>Coding Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>54</td>
<td>24.56</td>
<td>64 QAM</td>
<td>3/4</td>
</tr>
<tr>
<td>48</td>
<td>24.05</td>
<td>64 QAM</td>
<td>2/3</td>
</tr>
<tr>
<td>36</td>
<td>18.80</td>
<td>16 QAM</td>
<td>3/4</td>
</tr>
<tr>
<td>24</td>
<td>17.04</td>
<td>16 QAM</td>
<td>1/2</td>
</tr>
<tr>
<td>18</td>
<td>10.79</td>
<td>QPSK</td>
<td>3/4</td>
</tr>
<tr>
<td>12</td>
<td>9.03</td>
<td>QPSK</td>
<td>1/2</td>
</tr>
<tr>
<td>9</td>
<td>7.78</td>
<td>BPSK</td>
<td>3/4</td>
</tr>
<tr>
<td>6</td>
<td>6.02</td>
<td>BPSK</td>
<td>1/2</td>
</tr>
</tbody>
</table>

Table II: Simulation results for IEEE 802.11a protocol with various average SNR's.
figure, it is observed that FARM significantly outperforms OAR. In particular, when $\bar{\gamma} = 25$ dB, FARM can improve the overall throughput by 34% as compared with OAR; when $\bar{\gamma} = 35$ dB, the improvement is just 15%; and when $\bar{\gamma} = 10$ dB, the improvement is very little. Similar to the analysis above, the return probability approaches to 1 when the average SNR is very large or very small. In these cases, the difference between FARM and OAR decreases. Apparently, if only a single rate available, FARM will degrades to OAR.

From the figure, it is observed that FARM can maintain time-share fairness among all stations, but the throughput increases with the average SNR.

V. CONCLUSIONS

In this paper, we started from analyzing the rule maximizing the throughput of multi-rate WLANs, and then based on the analytical results we proposed a RTS/CTS based distributed fair auto rate medium access control. Extensive simulations showed that the proposed scheme outperforms the existing auto rate schemes and is capable of maintaining time-share fairness among asymmetric stations in time-varying fading channels.

ACKNOWLEDGMENT

This paper has been partially supported by the National Nature Science Foundation of China (60272021).

REFERENCES