A Scheduling Scheme to Enhance Throughput In Infrastructure-mode 802.11 WLANs

Osama Abu-Sharkh and Ahmed H. Tewfik

Department of Electrical and Computer Engineering.
University of Minnesota, Minneapolis, MN 55455
{osama, tewfik} @ece.umn.edu

***Abstract* — In IEEE 802.11 Wireless LANs, stations can transmit data at different rates. Stations far away from an access point in infrastructure mode will start transmitting at lower data rates depending on how much this station is subject to signal fading and interference. It was observed in the literature and our studies that degradation in the throughput occurs when stations operate at different speeds. The throughput of all wireless stations is decreased to the order of magnitude of the lowest. We proposed a solution in [3] that avoid this reduction. We suggested that the frame size should be a function of the transmission speed. In this paper, we introduce an alternative solution based on scheduling. Stations should send data on the channel at certain time slots according to their data rates. Simulation is performed using NS-2. An improvement in the system is achieved.**

***Index Terms* — 802.11, Infrastructure, Scheduling, Throughput, Wireless LANs**

I. Introduction

Recently, IEEE 802.11 standard has become the predominant technology for connectivity in WLANs, especially 802.11b. One of the most used configurations for 802.11 network consists of an access point surrounded by many wireless stations. In many situations, a station far away from an access point will transmit at different data rates depending on signal fading and interference. It was observed in the literature [2], and our studies that degradation in the throughput occurs when stations transmit at different speeds. The stations which operate at low data rates reduce the throughput of all other stations transmitting at higher rates to a value in the order of the throughput of the lowest. For example, a station operates at 1Mbps reduces the throughput of all other stations that operate at 11Mbps to a value below 1Mbps. This is due to the CSMA/CA channel access method used in 802.11 which guarantees that the long term channel access probability is equal for all wireless stations.

In [3], we proposed a solution to stop slow stations from degrading the throughput of fast stations in the network. The solution is based on adjusting the frame size of a station according to its speed. Wireless stations that transmit at high data rates would send larger frames than the ones that transmit at lower rates. It was found that in a 802.11b network if a slow station transmits at the lowest speed; i.e. 1Mbps, the payload should be set to a considerably low value in order to avoid the degradation in the system. This is impractical as the overhead would be relatively large compared to the payload size. Thus, we recommended that this slow station should stop sending on the channel until it is able to transmit at a higher rate.

In this paper, we introduce a scheduling scheme that also solves the problem. Any station, even the ones that runs at 1Mbps, can access the medium using this scheme. Our scheduling is based on dividing the time on the medium into two slots. One slot is allocated to stations that operate at full data rate, and another one is allocated to stations that operate at slower rates. By doing this, we avoid slow stations from penalizing fast stations as they have different time slots. Slow stations send their data and capture the channel only in their allocated time slot.

The rest of the paper is organized as follows. In section II, we provide a review of the IEEE 802.11 DCF access mechanism, and then we explain our scheduling scheme in Section III. In Section IV-A the network model used in the simulation is introduced. Simulation results are presented in Section IV-B, and we conclude the paper in Section V.

II. IEEE 802.11 DCF Access Mechanism

The DCF is the basic access mechanism of the IEEE 802.11 MAC (Medium Access Control). It achieves automatic medium sharing between competing stations through the use of CSMA/CA (Carrier-Sense Multiple Access with Collision Avoidance). Before a station starts transmission, it senses the channel to determine if it is idle. If the channel is idle, the station starts sending on the channel, otherwise the station waits until the end of the in progress transmission. The station should ensure that the medium has been idle for an interframe interval before attempting to transmit. The interframe space is a fixed amount of time, independent of the transmission speed. The Distributed Interframe Space (DIFS) is the minimum medium idle time for contention-based services. It is used by stations operating under the DCF access mechanism. A station may have immediate access to the medium if it has been free for a time period that is longer than DIFS. Moreover, to reduce the collisions between the stations that try to access the channel after a DIFS, the stations generate and defer an additional backoff period of time, which is determined by randomly choosing an interval within their contention window. The backoff counter is stopped when a transmission is detected on the channel, and reactivated when the channel is sensed idle again. The size of the contention window is increased to double the previous size for every unsuccessful transmission until it reaches its maximum value.

An acknowledgment (ACK) frame is sent by the receiver upon successful reception of a data frame. The transmitter assumes successful delivery of the corresponding data frame only after correctly receiving an ACK frame. A specified time interval between the reception of a data frame and the transmission of its ACK frame is defined and denoted by the Short Interframe Space (SIFS). It is smaller than DIFS. Using this small time space between transmissions within the frame exchange sequence prevents other stations that are waiting for the medium to be idle for a time longer than DIFS from attempting to use the medium, thus giving priority to completion of the in progress frame exchange sequence. On the other hand, if an ACK frame is received in error, i.e., received with an incorrect frame check sequence, the transmitter attempts to retransmit the unsuccessful frame on the medium after an Extended Interframe Space (EIFS) interval. Otherwise, if no ACK frame is received within an SIFS interval, due to an erroneous reception of the preceding data frame, the transmitter contends again to retransmit the frame after a period of time ACKTimeout.

In order to reduce collisions, 802.11 also employs an optional mechanism; RTS/CTS protocol, which exchanges smaller packets RTS (Request to Send) and CTS (Clear to Send) between the transmitter and the receiver before a data frame transmission. The IEEE 802.11 standard requires that a data frame is discarded by the transmitter's MAC after a certain number of unsuccessful transmission attempts. If the length of a data frame is less than or equal to a threshold (RTSThreshold), the number of transmission attempts is limited by the value of ShortRetryLimit, otherwise the maximum number of transmission attempts is set to LongRetryLimit.

III. Scheduling Scheme

Our scheduling scheme uses a centralized access control method. Access to the medium is controlled by the access point. A specialized function is implemented in the access point to add this feature. Associated stations can transmit data only when they are allowed to do so.

We divide the time on the medium into two time slots. One slot is allocated to stations that operate at full data rate, and another one is allocated to stations that operate at slower rates. At the beginning of each period, the access point transmits an introduced control frame. This frame includes the maximum duration of the upcoming period T_M. It also includes whether this time period is allocated to fast stations or slow ones. After receiving the control frame, stations that are not allowed to access the medium set their Network Allocation Vector (NAV) to the announced maximum duration. Other stations can send data on the channel and operate using the DCF access mechanism. These stations are not allowed to send any amount of data that takes a time period which is more than T_M or the remaining fraction of T_M. To achieve this, a counter is added in each station. This counter takes the value of the announced T_M at the beginning of each period, and is decremented by time. A station checks on its speed and amount of data it is intending to send before transmitting on the channel. If the transmission requires duration of time more than the value of the counter, it is optional to either send a fraction of its data or not to send anything.

To ensure that there are transmissions occur during any of the allocated time slots; i.e. stations have packets to send, we define a time value T_I. The access point interrupts a time slot by sending the control frame with new announcement if the medium is idle more than T_I. A proper value to set T_I is the value of the maximum period of time it is required for a station to send the largest frame size to the medium.

Note that we are proposing this scheduling scheme to be applied on IEEE 802.11 networks whose access mechanism do not account for time sensitive data and do not give any priorities. In order to avoid the dropping of packets when the stations wait long periods of time before their allocated slot starts, larger queue size should be used.

Slow stations stop penalizing fast stations as they have different time slots. Slow stations send their data and capture the channel only in their allocated time slot.

As our object is to make the stations get the same throughput they were supposed to get assuming that the slow stations in the network are sending at full data rate instead of lower ones, it can be easily found that the time slot that is allocated to fast stations should be more than the one allocated to the slower ones by a factor equals to

$$F = \frac{Throughput_F}{Throughput_A - Throughput_F}$$

where,

$Throughput_F$: The throughput that every fast station is supposed to get assuming that all stations in the network including slow ones are sending data at full rate.

$Throughput_A$: The throughput that every fast station gains assuming that only actual number of fast stations are in the network and no slow stations exist.

The values of the throughput (shown in Fig 2) should be stored in the access point. It uses them to set the value of T_M.

We divide the time on the medium only into two time slots, in spite the amount of degradation changes when different values of data rates are in the system, because stations do not lower their transmission speed unless they face signal fading or interference and this unlikely to happen except for only few stations which are located far away from the access point. Thus, if we divided the medium into more than two time slots that every slot is allocated to stations that operate at same speed, then the packets of the stations will face a big delay waiting their allocated time slot to start.

IV. Simulation

Simulation is performed using NS-2 simulator. The network model and the simulation results are explained in the following.

A. Network Model

The network model used in the simulation is a wired-cum-wireless network as shown in Fig. 1, where an access point is connected to the wired domain through a fixed node F. The Link between the access point and F has a data rate R and propagation delay D_p. N wireless stations are allocated with the access point. R, D_p in the simulation are set to 100Mbps, 2ms respectively. The capacity of the wired connection between the access point and the wired node F is set to a large value in order to accept all the uplink traffic without any loss.

The wireless LAN operates using 802.11b. The basic Distributed Coordinated Function (DCF) is employed. The stations are allowed to transmit at four data rates 11, 5.5, 2, and 1Mbs.

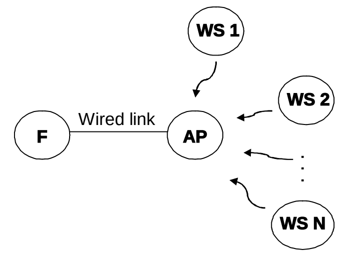

Fig. 1 wired-cum-wireless Network

Random movement is generated for the stations. For simulation simplicity, we assume that stations switch to lower data rates when they move away from the access point as the distance plays the major role. By doing this, we won't loose the generality of taking reasons other than the distance, as switching implementation and reasoning differs from one vendor to another.

TCP is the used transport layer protocol. The packet size is set to 1000 bytes. Greedy uplink traffic is generated between stations and wired domain through the access point. The stations are always backlogged.

B. Simulation Results

- *Scenario 1*

We first run the simulation without any modifications on the standard for a period of time equals to 200 seconds.

Average throughput is calculated in the simulation for different number of stations as shown in Fig 2. The throughput strongly depends on the number of active stations in the network where it decreases with increasing the number of stations. It is also seen from the graph that stations which operate at low data rates influence the overall performance of the system as expected and explained earlier.

We should note that the larger the number of stations that transmit at slow speeds in the network, the more reduction in the throughput occurs.

- *Scenario 2*

We divide the medium into two time slots and allocate one of the slots to stations that transmit at full speed and the other one to slower stations. The simulation time is also set to 200 seconds.

Fig 3 shows that F increases with the increasing number of stations in the network to get the appropriate throughput in a system that does not have any slow stations. Note that F is implicitly dependant on the number of slow stations in the network, but it is independent on the amount of throughput reduction they cause when using the unmodified version of the standard.

Top

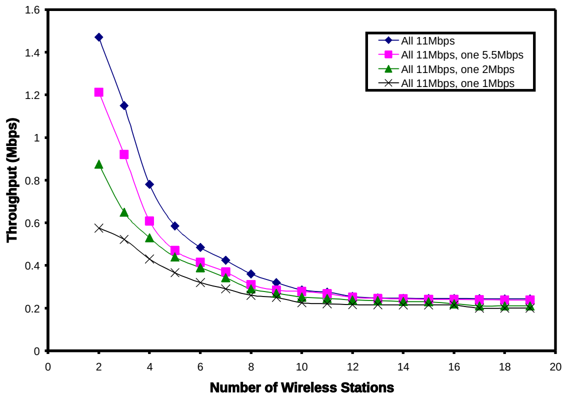

Fig. 2 Throughput of stations that run at 11Mbps when only one station runs at 1, 2 or 5.5 Mbps.

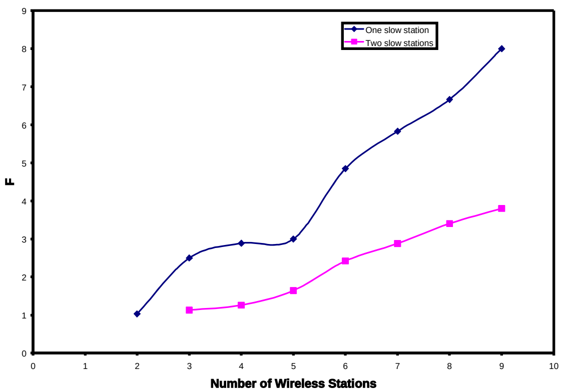

Fig. 3 The ratio of the time period that should be allocated to fast stations to the one that should be allocated to slow stations with the increasing number of stations in the network.

Thus, we conclude that no matter what the speed of slow stations in the network, the solution does not change for fast stations.

Fig. 4 shows the throughputs that slow stations gain with the increasing number of slow stations in the network. We fix the number of fast stations to five for illustrated purposes only. We found that the amount of reduction in throughput of slow stations by using the scheduling scheme is lower than the one when we make the frame sizes a function of station's speeds. For example, Table I shows the throughput of a slow station when there are only two stations; one transmits at full rate and the other one on a lower rate.

Table I

	1Mbps	2Mbps	5.5Mbps
Scheduling scheme	0.39	0.70	1.25
Frame size Variation	0.09	0.22	0.69

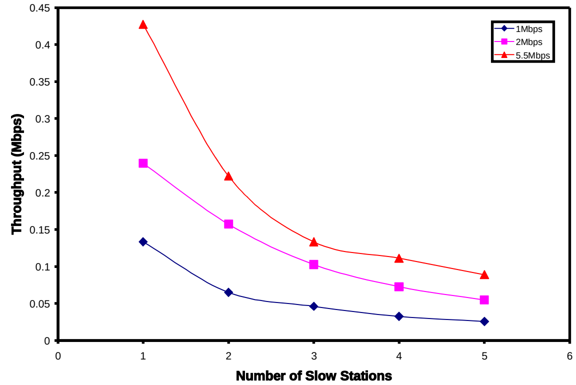

Fig. 4 Throughput gained by slow stations. Number of fast stations is five.

To use this scheduling scheme or the solution proposed in [3] is a matter of trading off.

Varying the frame size according to the speed of the stations is simple to implement and does not require any changes in the access mechanism. It can also be applied on either infrastructure or Ad-hoc modes. The major drawback of this solution is that slow stations would gain very low throughput as was discussed in [3]. It is also recommended that stations that run at 1Mbps should stop accessing the medium. This can be explained by the fact that a packet with a small payload size has a large overhead.

On the other hand, by applying the scheduling scheme, any slow station, no matter what the value of its speed, is allowed to send data on the channel. Slow stations would gain a high throughput relative to the ones they were supposed to get by changing their frame sizes. One of the drawbacks of adopting this approach is the use of a centralized access control mechanism and the need to make major changes on the original access mechanism of the standard. The time delay that faces the packets while waiting in the queues until their allocated time slot starts is a considerable drawback mainly for time sensitive traffic.

V. Conclusion

In this paper we introduced another solution to avoid the degradation in throughput that happens in 802.11 WLANs, when some stations run at lower speeds than the nominal rate. The solution is based on a TDMA scheduling scheme. Giving slow stations a separate time slot reduces their effect on the performance of the system.

Top

We discussed the two proposed solutions. We found that if the time delay and simplicity of the system are our main concerns, then the solution in [3] should be used. Otherwise, the scheduling scheme should be used especially it gives better results in terms of slow stations' throughput.

REFERENCES

[1] IEEE Computer society. 802.11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. June 1997.

[2] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance anomaly of 802.11b," IEEE Infocom'03, San Francisco, USA, Mar. 2003.

[3] O. Abu-Sharkh, A. Tewfik, "Throughput Evaluation and Enhancement in 802.11 WLANs with Access Point", IEEE 61st Vehicular Technology Conference, Stockholm, May 2005.

[4] C. Koksal et al., "An analysis of short-term fairness in wireless media accesss protocols," in Proceddings of ACM SIGMETRICS, 2000.

[5] S. Lu, V. Bharghavan, and R. Srikant, "Fair scheduling in wireless packet networks," ACM Sigcomm'97, Cannes, France, Sep. 1997.

[6] M. Shreedhar and George Varghese, "Efficient fair queuing using deficit round-robin." IEEE/ACM Transactions on Networking, no. 3, pp. 375-385, Jun 1996.

[7] M. Heusse, P. Starzetz, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Bandwidth Allocation for DiffServ based Quality of Service over 802.11," IEEE Globecom 2003, San Francisco, USA, 2003.

[8] S. Lu, T. Nandagopal, and V. Bharghavan, "A wireless fair service algorithm for packet cellular networks," IEEE/ACM MobiCom, 1998.

[9] D. Qiao, S. Choi, K. Shin, "Goodput Analysis and link Adaption for IEEE 802.11a Wireless LANs," IEEE Transactions on Mobile computing. Vol. 1, No.4, October-Demeber 2002.