

Supporting Real-Time Applications with Better QoS Guarantees in 802.11

Kiran Diwakar Sridhar Iyer
[kiran,sri]@it.iitb.ac.in
School of Information Technology
IIT Bombay

Abstract—The proliferation of IEEE 802.11 based wireless LANs has led to a lot of interest in providing Quality of Service(QoS) guarantees in such networks. 802.11e is the QoS variant of the 802.11 family of protocols. It suggests a flow-based solution based on priority queues at the node, for providing QoS guarantees. We propose a different approach, using a variant of TDMA, called Dynamic Time-division Multiple Access(DTMA). It is based on the observation that the ratio of data transmission time to control packet(poll or acknowledgement) transmission time is typically 6:1, and drops down almost to 2:1 when the data packet size becomes smaller than 600 bytes. DTMA focusses on reducing control information and thereby increasing time available for data transmission. More time available for data transmission implies increased throughput.

We expunge the overhead of control packets in HCF by using cumulative acknowledgements and piggybacking. Taking advantage of 802.11's inherent limited range, DTMA, a modified TDMA is used for data transfer in HCF, without encountering the problem of distributed time synchronization. In this paper, we do the analysis of 802.11e and DTMA using probability models and also present simulation results to support these models. Thus, aided by simulations and analytical methods, we prove that DTMA has stricter and lesser delay bounds than 802.11e, for real-time and QoS sensitive applications. We also show that DTMA enhances the overall throughput by almost 20%, thereby implying better QoS guarantees in the 802.11 domain.

I. INTRODUCTION

Wireless local area networks(WLANs) are now expected to support real time multimedia applications like videoconferencing, streaming servers, which require finite delay guarantees. The 802.11e Group, formed to enhance and improve the current 802.11 MAC, has suggested modifications to the legacy 802.11 to provide the required QoS guarantees[1].

The 802.11e Enhanced Distributed Co-ordination Function(EDCF) is based on Inter Frame Space(IFS) and Contention Window(CW) priorities, providing as many as 8 service classes. Each service class has a different IFS and also different CW. These two parameters help to distinguish between different service classes. The Hybrid Co-ordination Function(HCF) is an extension to the PCF mode. The Hybrid Coordinator(HC) controls the channel access and ensures QoS guarantees for prioritized flows by granting explicit access to the high priority nodes in the Contention Period (CP).

In DTMA we keep all the explicit control messages to the minimum. We use the dynamically slotted TDMA approach wherein, the privileged nodes do not require to be polled or be sent acknowledgements immediately after data transmission.

We use piggybacking and deferred cumulative acknowledgements to overcome the transmission of explicit control messages. The privileged nodes again contend for channel access in the CP, with a lower CW and also a lower IFS, ensuring better chance for transmission.

This paper is organized as follows. Section II describes the 802.11e extension to support QoS in WLANs. Sections III gives the motivation for a new scheme. Section IV gives the description of DTMA for providing better QoS guarantees followed by section V consisting of the analysis of 802.11e. Section VI provides the analysis of DTMA, followed by simulation results in the next section.

II. 802.11E - THE QOS EXTENSION

Legacy 802.11 is not suitable for QoS guarantees because there is no prioritization among nodes or flows. The MAC lacks the ability to differentiate between traffic streams. Hence the need of 802.11e to support QoS in the wireless networks. 802.11e [1] has the following modes of operation :

- 1) Enhanced Distributed Coordination Function
- 2) Hybrid Coordination Function

These modes make use of the concept of priority classes called Traffic Categories(TCs).

A. Enhanced Distributed Coordination Function

Enhanced Distributed Coordination Function(EDCF) is DCF with some elements of the MAC modified. EDCF is a flow based approach that classifies flows into different TCs. There is a queue for each of the TCs. The minimum Contention Window(CW) for each TC is different and this is the basis for giving priorities. Also an added IFS is used for service differentiation, called *Arbitration Inter Frame Space(AIFS)*. AIFS is the time the node has to check the channel to be idle before transmission. The AIFS value is dependent on the particular TC.

EDCF overcomes the flat nature of 802.11 DCF, allowing prioritization for the required QoS. The queues for each TC act as virtual stations within that node. The different backoff intervals and AIFS decide the time for scheduling a particular queue and thus decide the flow for transmission.

B. Hybrid Coordination Function

Hybrid Coordination Function is an extension to the PCF idea of the 802.11. It comprises of alternating Contention

Free Period(CFP) and the Contention Period(CP). The Point Coordinator in the legacy 802.11 PCF is replaced by a Hybrid Coordinator (HC). The HC governs the decision of channel access. In a given CFP, the HC polls the privileged stations. A station responds to a poll in SIFS amount of time, failing to do so, the HC takes control of the channel again. The HC then either polls the next station or sends *CFEnd* if there is no station left to be polled. The CP is now divided into two parts, the conventional CP, where all the nodes compete for channel access using the random backoff technique and the *Controlled Access Period(CAP)* where the HC gives channel access to the flow that has QoS guarantees unfulfilled. After the CAP, the HC again leaves the channel for contention based access till it needs to grant explicit access to a particular flow for transmission of data.

There is a new QoS field added to the MAC frame through which the stations convey the queue status, delay information to the HC. Based on this information the HC takes the scheduling decisions, giving *Transmission Opportunities(TXOP)* to the stations in need. This may be by polling in *CFP* or by stealing a time slot from the *CP* and allowing the priority node to transmit.

III. MOTIVATION FOR A NEW SCHEME

We observe that in the HCF mode the polling is similar to that of legacy 802.11. The data transmission time for a packet size of 512 bytes for a station is approximately $100\mu\text{secs}$ and the transmission time for the poll or acknowledgement is approximately $44\mu\text{secs}$ [6]. Thus, in the worst case the system spends almost half of the time sending just the control information. Even for a packet size of 2000 bytes, the ratio of data to control frame transmission time is almost 6 : 1, which becomes a substantial overhead when the number of priority flows is high. Some drawbacks in 802.11e are :

- 1) Controlled Contention Scheme : The stations having priority traffic send their requests to the HC, requesting for TxOPs in the *CAP*. The problem here is that a station can modify its channel requirement information only when it gets an opportunity during CP, which is probabilistic.
- 2) EDCAF does not give any guarantees for prioritized traffic. At high load there are high number of collisions even for flows with high priority.

We propose to use a dynamic slot allocation version of *TDMA approach* for transmission of high priority data. We eliminate the process of polling each station by having each node transmit in a defined time slot. Individual acknowledgements are removed by using one cumulative acknowledgement. With the beacon, the time slot value is communicated to each privileged node. Each node knows its position in the ordered list. Using this available information, the node calculates the time at which it has to transmit, without incurring the extra overhead of control information transmissions.

IV. DTMA : THE PROPOSED SCHEME

DTMA is Dynamic Time-division Multiple Access. As discussed in the earlier section, using explicit polls and acknowledgements can have a huge overhead when the number of pollable stations is high. So, DTMA assigns time slots to the privileged nodes for data transmission. The modified beacon broadcasts the time slot value. Each node is aware of its position in the ordered list. Now, based on this time slot value and position, the node calculates its time for transmission of data. So, after the beacon transmission, the privileged nodes start transmission without waiting for poll messages. The AP upon receipt of the data packet, 'marks' that particular node in the cumulative acknowledgement. At the end of *CFP*, AP sends cumulative acknowledgement.

We consider parameters like load on the network, average delay and throughput, for making the decision of slot size. A time slot value that optimizes these weighted parameters is chosen. The time slot value adapts to the system so that delay and throughput are optimized and the overall performance of the system is improved. Currently we assign slot-size statically, but a more complex function can be used. In addition to the value of time slot, the number of time slots also vary dynamically. If the AP finds that a privileged node does not have data to send, it does not allot the node any slot for transmission. Thus we remove the inherent drawback of the TDMA approach of wasting time slots, when there is no data to be sent.

Figure 2 shows the new super frame for DTMA. Along with the conventional fields in the beacon, there are a few additional fields to incorporate the dynamic slot approach. The new additions to the beacon are :

- *Cummulative Acknowledgement* : It is the combined acknowledgement for all the nodes that transmitted in the last superframe.
- *Ordered Privileged List* : It is the list of nodes, order of which decides the transmission order in *CFP* .
- *Time Value of Slot* : It is the time-slot value.

The sequence of events taking place in DTMA are :

- 1) As shown in the figure 2, the frame starts with the reception of the beacon. All the nodes get the beacon frame. All nodes now know the *Time Value of Slot* for the dynamic slot scheme. Also the *Ordered Privileged List* field in the beacon is checked by all the privileged nodes, that gives each the position in the sequence of transmissions. From its position in the ordered list and the time slot value, each node calculates the exact time at which it has to transmit data. Thus no polling or acknowledgements are required.
- 2) After this the nodes start transmitting at their respective time instants. The AP keeps receiving the data, but does not send the acknowledgement to the sender node. Instead it marks the acknowledgement in a *Cummulative Acknowledgement*. If data is received without any errors, 1 is inserted in the position corresponding to the node in the list, else a 0 is inserted.

- 3) After the last node in the *Ordered Privileged List* has finished with its transmission, the AP sends the cumulative acknowledgement for all the nodes. The privileged nodes check for the bit in their corresponding position in the list. If it is '1' the node removes the sent packet from its buffer and prepares to send the next one, else it prepares to retransmit the old packet that was not received properly.
- 4) Following the combined acknowledgement is the EDCF part as per the 802.11 standard. All the nodes compete as per backoff and transmit rules.
- 5) After the EDCF, the superframe ends and the beacon is transmitted indicating the start of a new superframe. The *Cumulative Acknowledgement* in the beacon though redundant, is kept for reliability. If the original acknowledgement is erroneous, the Cumulative Acknowledgement in the beacon can be used to avoid retransmissions.

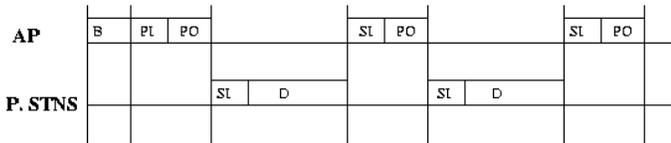


Fig. 1. HCF Transmission Sequence

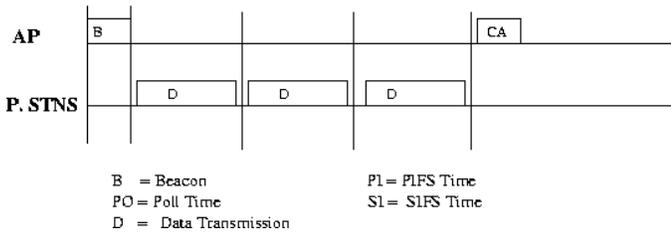


Fig. 1. DTMA Transmission Sequence

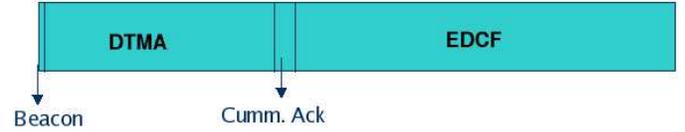
Figures 1a and 1b show working of HCF vis-a-vis DTMA.

V. ANALYSIS OF 802.11E

Yunli Chen et al [2] have done an analysis of 802.11e EDCF. The average delay for any real time flow in EDCF is given. This analysis considers only EDCF for calculating the average delay. For the complete analysis of 802.11e we need the average delay over both HCF and EDCF combined together. Our handling of the analysis is thus different. We calculate the delay for the privileged node in HCF and EDCF individually and then calculate the average delay over the entire superframe.

Let the number of privileged nodes be x and the unprivileged be y . The backoff interval is assumed to be $[1, N]$. For ease of analysis, we assume that each node requires the same amount of time for data transmission, m . The number of transmission time slots in *CFP* and *CP* are $t1$ and $t2$ respectively. The superframe time is T units.

For tractability in the analysis, we divide the *CFP* and *CP* time into a number of slots in which data transmissions can



$$\text{Beacon} = \text{Beacon} + \text{Cumm. Ack} + \text{Ordered Privileged List} + \text{Time value for each slot}$$

Fig. 2. Proposed Super Frame

take place. So,

$$t1 = \frac{CFPTime}{m}$$

$$t2 = \frac{CPTime}{m}$$

We first start by calculating the average delay a privileged high priority node suffers in the *CP*.

We note that, for the privileged node to transmit successfully in EDCF, it must select the lowest random backoff. That is, if it has a random backoff value i , all the remaining nodes must select a value that is greater than i . Thus, the probability that the privileged node chooses a random backoff value i is uniform :

$$P1 = \frac{1}{N}$$

Now, the probability that all other privileged nodes ($x-1$) choose a random backoff that is greater than i is :

$$P2 = \left(\frac{N-i}{N}\right)^{x-1}$$

Also the constraint that needs to be satisfied is that no other non-privileged node of the given y selects a random backoff less than or equal to i . Thus, the probability of a non-privileged node selecting a backoff greater than i becomes :

$$P3 = \left(\frac{N-i}{N}\right)^y$$

If EDCF is assumed,

$$P3 = \left(\frac{N-i}{N+k}\right)^y$$

where, k is a positive integer that ensures that the non-privileged nodes have a larger interval to select a backoff $[0, (N+k)]$.

Since the events are independent, the total probability becomes,

$$P = P1 * P2 * P3$$

Summing over all possible values of i in the given interval,

$$P = \sum_{i=1}^N \frac{1}{N} \left(\frac{N-i}{N}\right)^{x-1} \left(\frac{N-i}{N+k}\right)^y$$

Now, the probability that a privileged node is able to transmit successfully in the k th attempt is :

$$P_k = (1 - P)^{k-1} * P$$

Now the probability is calculated such that this k th attempt happens before the end of the superframe and in the *CP*.

$$P(k \leq T) = \sum_{k=t1+1}^{t1+t2} (1 - P)^{k-1} * P$$

In the given scenario, the worst case would be when all the nodes, both privileged as well as the non-privileged have packets outstanding in the queue. The expectation that this privileged node transmits in the *CP* after k attempts is:

$$E(k) = \sum_{k=t1+1}^{t1+t2} P_k * k$$

The average delay for a privileged node is:

$$D = E(k) * m$$

D is the average delay in the worst case that a high priority privileged node suffers in the *CP*. The closed form expression for the total average delay over the entire superframe is now calculated. The privileged node gets polled in the *CFP* for data transmission followed by an acknowledgement for the same by the AP.

The probability of transmission in the *CFP* is :

$$P_{CFP} = P_{Poll} * P_{Data} * P_{Ack}$$

where, P_{Poll} and P_{Ack} are the probabilities of successful receipts of poll and acknowledgement frames respectively. P_{Data} is the probability of successful transmission of data packet.

We assume that once a poll is received the data probability of data transmission is 1. So, $P_{Data} = 1$. Hence,

$$P_{CFP} = P_{Poll} * P_{Ack}$$

The delay encountered due to the overhead of polls and acknowledgements is :

$$P_{poll} * P_{ack} * (T_{overhead})$$

where,

$$T_{overhead} = T_{Poll} + T_{Ack} + 2 * T_{SIFS}$$

$$T_{Poll} = \text{Time for Poll}$$

$$T_{Ack} = \text{Time for Acknowledgment}$$

$$T_{SIFS} = \text{Short Interframe Space}$$

$$\text{Let } D1 = P_{poll} * P_{ack} * (T_{overhead})$$

As mentioned above, the average delay for the privileged node over the entire superframe is total number of data transmissions per superframe time.

$$AvgDelay = \frac{T}{\frac{t2*m}{D} + \frac{t1*m}{D1}} \quad (1)$$

Thus the average delay depends on the number of nodes of the privileged and unprivileged type, the value of backoff interval and the delays due to the control information.

VI. ANALYSIS OF DTMA

There is a bounded delay requirement for the high priority nodes that generate real time, delay sensitive traffic. We derive a closed form expression for this bounded delay. We also prove the following claims :

- The maximum delay that any high priority node suffers is less than or equal to *superframe time* (T).
- The delay for real time traffic is strictly bounded and we give the expression for this bounded delay.
- The bounded delay given by DTMA is stricter and lesser than that given by 802.11e.

In HCF, all the high priority privileged nodes transmit their data in the assigned time slot in the *CFP*. Now as the *CFP* is repeated after every superframe, it is trivially proved that the *maximum delay* that is suffered by a privileged node is T . The first claim, (a) is proved here.

Now let us calculate the average delay for a high priority node.

For ease of comparison, the parameters used for analysis of DTMA are the same as those for 802.11e. On the lines of initial steps of derivation of (1), the probability of a privileged node selecting a random backoff i , that is the lowest among all the nodes for EDCF is,

$$P = \sum_{i=1}^N \frac{1}{N} \left(\frac{N-i}{N+k} \right)^{x-1} \left(\frac{N-i}{N+k} \right)^y$$

Now, the probability that a privileged node is able to transmit successfully in the k th attempt is :

$$P_k = (1-P)^{k-1} P$$

So, the probability of the privileged node transmitting in the *CP* is:

$$P(k \leq T) = \sum_{k=t1+1}^{t1+t2} (1-P)^{k-1} P$$

For the privileged node, average number of transmissions it waits before it transmits would therefore be the expectation, which is:

$$E(k) = \sum_{k=t1+1}^{t1+t2} P_k * k$$

The worst case average delay for the privileged node is:

$$D = E(k) * m$$

The closed form expression for the total average delay over the entire superframe for DTMA is:

$$AvgDelay = \frac{T}{\frac{t2*m}{D} + 1} \quad (2)$$

We see, when compared with (1), the delay bound given by DTMA is stricter as well as lesser than that given by 802.11e HCF and EDCF. We have removed the overhead of the polls and acknowledgements and this has helped us to grant more time for the actual data transmissions and thereby reduce the delay and improve performance and throughput. Equation (2) proves the claims (b) and (c).

For the throughput analysis, we replaced Crawford's model [4] by DTMA in their work. By removing the delay overheads of polling and acknowledgements, and subsequently adding the delay induced by the cumulative acknowledgement of DTMA, we recalculated the throughput for the privileged stream. Using DTMA, we had an enhancement of approximately 19% in throughput. The improvement would be more prominent for more number of real time flows. For numerous flows, the control information transfer becomes very high and so the reduction in that improves the overall performance.

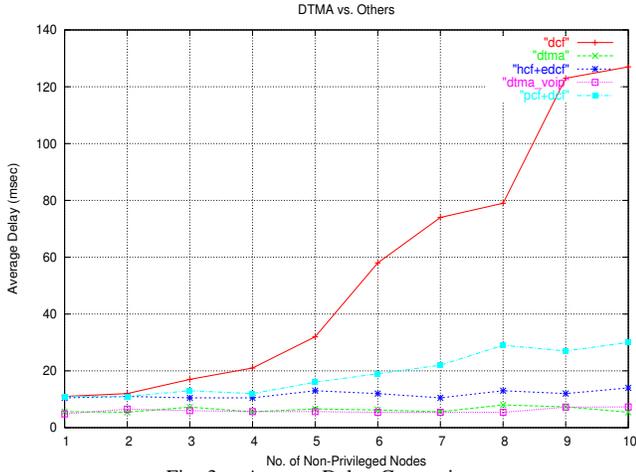


Fig. 3. Average Delay Comparison

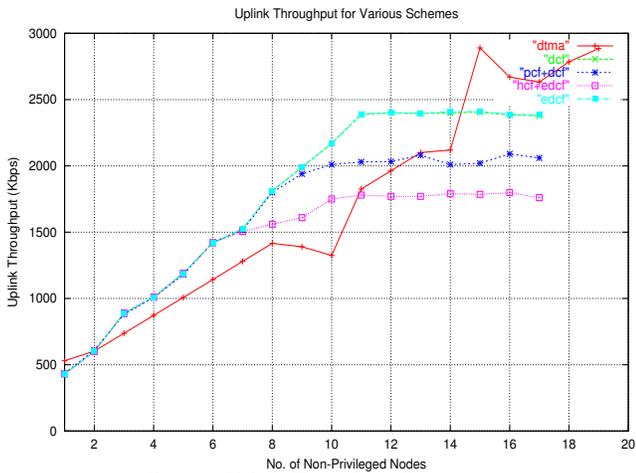


Fig. 4. Uplink Throughput Comparison

VII. SIMULATION RESULTS

We verified our theoretical model using computer simulations. Network Simulator was used for our experiments.

Antonio Grilo et al evaluate the performance of 802.11e [3]. Three types of traffic sources are considered : bursty data(HTTP sessions), VoIP and a video source. The audio source generated messages of size 60 bytes at the bitrate of 24 Kbps. The video source model has an output rate of 256 Kbps. VoIP and video are the high-privileged sessions.

We conducted our experiments in a similar environment and traffic flows were replicated. The various schemes compared are DCF, PCF and DCF, EDCF and HCF, and DTMA.

Figure 4 compares the average delay suffered by a priority node running different schemes. DTMA shows better performance for both video as well as the VoIP flows. Enhanced performance of DTMA is attributed to the reduced MAC overhead of polling and individual acknowledgements. As we had predicted, the number of non-privileged nodes does not have an effect on the performance of DTMA. The average delay remains less than 7msec consistently as against 10msec for HCF-EDCF, for any number of flows. We can see that the VoIP traffic suffers even lesser delay due to the less data rate and also the considerable smaller packet size. Thus the

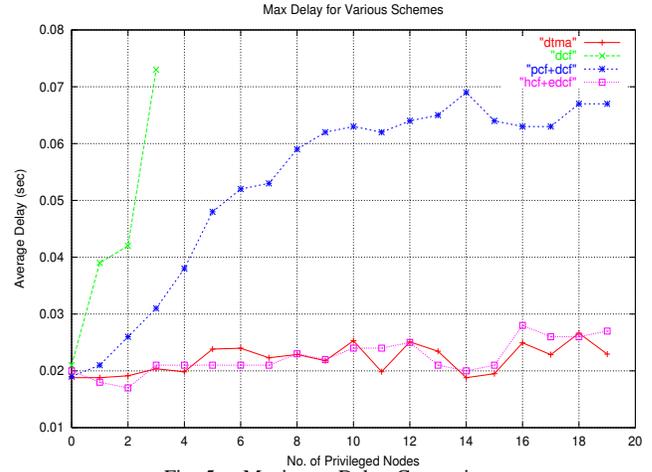


Fig. 5. Maximum Delay Comparison

average delay suffered using DTMA is lesser. The throughput comparison shown in fig 5 also conforms to our prediction. The performance enhancement is not much for lesser number of nodes, though the slope followed is similar compared to the other schemes. For the other schemes, the throughput stabilizes after the number of non-privileged nodes starts increasing over 8. But we can see that the throughput is increasing for DTMA till node 15. The stabilization occurs for DTMA when the number of nodes reach around 22. Figure 6 shows the comparison of the maximum delays suffered by a packet for various schemes. The maximum delay for DTMA is almost the same when compared to 802.11e.

VIII. CONCLUSION

The current schemes to provide QoS can be modified for better QoS guarantees to the end user. DTMA approach for HCF of 802.11e, in which the privileged nodes will transmit as per dynamically allocated time slots, avoids the overheads of explicit control messages and thus ensures more time for the actual data transfer and hence less delays for the privileged nodes. Overall throughput given by DTMA is better than the proposed standard 802.11e by almost 20%. The analytical work done shows that the delay bound for DTMA is stricter and lesser than 802.11e. Our simulation results also support the analytical model. Thus DTMA improves the overall performance of the system by evading the transmission of explicit control information, allowing more time for the actual data transfer.

REFERENCES

- [1] IEEE 802.11e Draft Standard for Wireless LANs with QoS, P802.11 July 2001.
- [2] Yunli Chen, Qing-An Zeng and Dharma P. Agarwal, "Performance analysis of IEEE 802.11e Enhanced Distributed Co-ordination Function", *The 11th IEEE International Conference on Networks*, October 2003.
- [3] Antonio Grilo, Mario Nunes, "Performance Evaluation of IEEE 802.11E" *Proceedings of PIMRC02*, Lisbon 2002.
- [4] James Crawford, "IEEE 802.11E/A Throughput Analysis", *Magis Document White Paper No. E10282*, 2003 .
- [5] Qiang Ni, Lamia Romdhani, Thierry Turetli, Imad Aad, "QoS Issues and Enhancements for IEEE 802.11 Wireless LAN", *RINRIA*, November 2002.