A Hybrid Distributed Coordination Function for Scalability and Inter-operability in Large-scale WLANs *

Nakjung Choi[†], Seongil Han, Yongho Seok, Yanghee Choi and Taekyoung Kwon
School of Computer Science and Engineering
Seoul National University
Shillimdong, Gwanak-gu, Seoul, Korea
{fomula, sihan, yhseok, yhchoi, tk}@mmlab.snu.ac.kr

Abstract

The distributed coordination function (DCF) of IEEE 802.11 MAC protocol is not scalable as the number of users on a network increases. We propose two new contention resolution schemes to reduce the number of collisions between a large number of contending mobile stations. The hybrid DCF (H-DCF) scheme splits contention resolution into two phases to improve scalability. Enhanced hybrid DCF (EH-DCF) adds fairness with expect to mobile stations using the current IEEE 802.11 MAC and further improves the efficiency of contention resolution. Inter-operability with the IEEE 802.11 DCF is the major advance over previous work. We use an ns-2 simulator to simulate the performance of these new schemes with the IEEE 802.11 DCF. The results show that an improvement in network performance by up to 10∼30%.

Keywords - IEEE 802.11; scalability; inter-operability; two phase contentions; null frame; channel occupancy regulation

1 Introduction

Users of wireless local area networks (WLANs) [1] are always seeking additional bandwidth, and WLAN service providers also desire increased bandwidth to accommodate more subscribers in hotspot areas. High-bandwidth WLAN technologies such as IEEE 802.11a/b/g [2-4] are becoming widespread. For example, in IEEE 802.11a the set of possible data rates is 6, 9, 12, 18, 24, 36, 48 and 54Mbps; IEEE 802.11g also allows transmission at up to 54Mbps; while the IEEE 802.11n [5] working group is defining modifications to the Physical (PHY) layer and Medium Access Control (MAC) layer that will deliver at least 100Mbps in next-generation WLANs. The MAC protocols of these standards are mostly based either on the Distributed Coordination Function (DCF) or on Enhanced Distributed Channel Access (EDCA) in IEEE 802.11e [6] supporting Quality of Service. Although there are centralized and controlled channel access schemes such as the Point Coordination Function (PCF) and HCF Controlled Channel Access (HCCA) in IEEE 802.11e, a distributed channel access schemes are more desirable in terms of complexity and cost, and such schemes are now widely used. However, the wireless channel utilization of the DCF and of the EDCA decreases as the number of competing mobile stations grows. In Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA), on which DCF and EDCA are based, each mobile station chooses a backoff time within a contention window (CW) and defers channel access to avoid a collision. The optimal CW value that maximizes network throughput depends on the number of mobile stations. Currently, the DCF primarily used in IEEE 802.11 standards adjusts its CW to the wireless channel conditions with an exponential backoff mechanism; but the resulting collision probability increases with the number of mobile stations.

Attempts to solve this scalability problem [7-11], have assumed the necessity of maintaining compatibility with a legacy 802.11-based device, which increases the difficulty of achieving a satisfactory solution. When a legacy 802.11-based device competes with an enhanced channel access-based device in a congested WLAN, the former has a higher channel access priority. To achieve the maximum network throughput in IEEE 802.11 WLANs, each mobile station must adjust the size of its contention window at each contention to reflect the number of active mobile stations. When a large number of mobile stations are competing, those enhanced mechanisms are especially effective: they decrease the probability of transmission occurring, and thus reduce the number of collisions, increasing the proba-

*This work was supported in part by the Brain Korea 21 project of Ministry of Education, 2006, Korea.
†The correspondent author

bility of a successful transmission by the DCF. But these enhanced mechanisms eventually reduce the transmission probability so much that stations start to experience starvation of wireless channel access. While continuing to support the compatibility with legacy 802.11-based devices, we propose a novel channel access mechanism to fulfill the following requirements: 1) *scalability*, 2) *low contention overhead*, and 3) *inter-operability* with DCF.

The rest of this paper is organized as follows. We begin in Section 2 by the IEEE 802.11 DCF and related research. In Section 3, we present our new distributed channel access mechanism, which is called hybrid DCF (H-DCF). In Section 4 we go on to present an enhanced version of H-DCF (EH-DCF) that is more efficient and provides fairness with DCF stations. The performance of H-DCF and EH-DCF is then evaluated by means of simulations, the results of which are presented in Section 5. We conclude this paper in Section 6.

2 Background

2.1 IEEE 802.11 DCF

We will now briefly review random binary exponential backoff in the IEEE 802.11 DCF (Fig. 1). A transmitting mobile station must first sense the idle channel for a time period equal to the DCF Interframe Space (DIFS), after which it generates a random backoff time chosen uniformly from the range [0, CW], where CW is the *contention window*. At the first transmission attempt, CW is set to CW_{min}. When the backoff timer reaches 0, the mobile station transmits a frame. If that transmission is successful, the CW of the mobile station is reset to CW_{min}. Otherwise, its CW is doubled. A mobile station that overhears another's frame reads the duration field to update its Network Allocation Vector (NAV) which is a time period during which it assumes that the channel will remain busy, and defers its transmission for that length of time so as to avoid frame collisions. After a successful frame transmission sequence ends, mobile stations wait until they sense that the channel has been idle during a DIFS. But if a frame collision occurs or an erroneous frame transmission is detected, mobile stations wait until they sense that the channel has been idle for the longer period of an Extended Interframe Space (EIFS). When the DIFS or EIFS has expired, contention for the next frame transmission starts.

2.2 Enhancement of the IEEE 802.11 DCF

F. Cali et al. [7] showed that the IEEE 802.11 DCF achieves network performance which is far less than the theoretical limit. To improve the capacity of IEEE 802.11, they

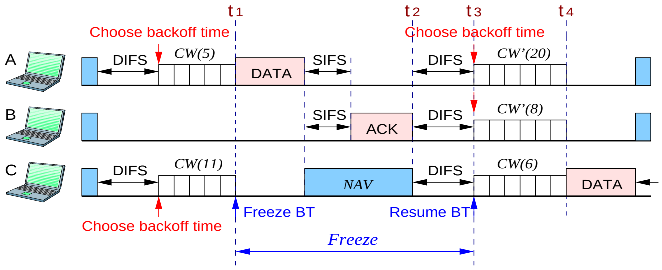


Figure 1. Random binary exponential backoff in IEEE 802.11 DCF.

proposed a mechanism to tune IEEE 802.11 MAC parameters dynamically, depending on the estimated number of active mobile stations. Y. Yang et al. [8] presented a theoretical analysis of a distributed CW control algorithm that achieves arbitrary bandwidth allocation policies and efficient channel utilization. By modelling bandwidth allocation as an optimal CW assignment problem, they proposed a general and fully distributed CW control algorithm. I. Aad et al. [9] have pointed out that it is inappropriate to change the contention level in IEEE 802.11 WLANs by resetting the CW to its minimum value after a successful transmission because that results in more collisions and retransmissions before the CW reaches its optimal value again. Therefore, they suggested the CW is slowly reduced after a successful transmission instead of being reset. Y. Kwon et al. [11] proposed an efficient way of resolving contention in WLANs, which they called the fast collision resolution (FCR) algorithm. It attempts to resolve a collision quickly by increasing the CW of not only the colliding stations but also of the deferring stations in the contention procedure. To reduce the number of idle slots, FCR decrements the backoff timer exponentially, instead of the linear decrement specified in the IEEE 802.11 DCF. H. Kim et al. [10] developed a model-based frame scheduling scheme (MFS) to enhance the capacity of IEEE 802.11 WLANs. This computes the current utilization of the wireless network and then determines the scheduling delay that needs to be introduced before a mobile station attempts to transmit its pending frames. The objective of MFS is to decrease the number of collisions by reducing the number of mobile stations competing for the wireless channel.

3 Hybrid-DCF (H-DCF)

Hybrid DCF (H-DCF) introduces two-phase contention by means of a *null* frame to reduce frame collisions. The first contention phase determines a set of eligible stations which are the only ones allowed to enter the second contention phase in which they contend for frame transmissions. The second contention phase is finished when all eligible stations have transmitted their frames regardless of success or failure.

As more mobile stations compete, the probability of a collision at a particular slot becomes higher in the IEEE 802.11 DCF, which uses single-phase contention for frame transmission. However, with H-DCF, mobile stations that choose the same time slot in the first contention phase do not experience a frame collision but go on to contend a second time. Therefore, the number of mobile stations that are actually contending for frame transmissions is significantly reduced and the probability of a frame collision is reduced.

3.1 Basic Algorithm

3.1.1 The first contention phase

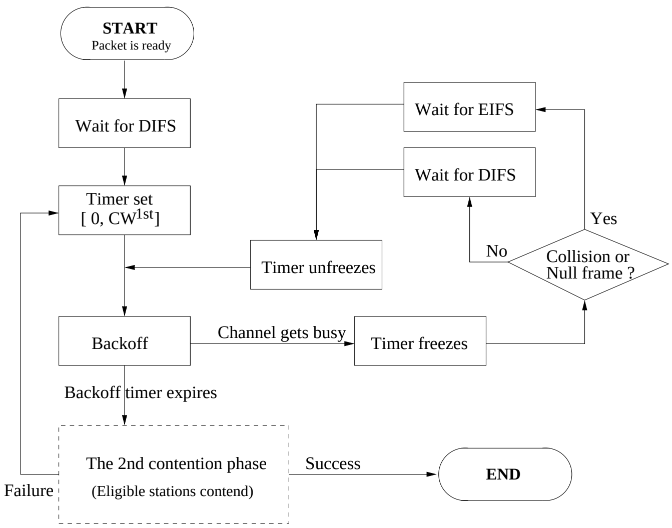


Figure 2. Flow chart of the first contention phase.

The basic principle underlying both contention phases is similar to that of the IEEE 802.11 DCF. In the first contention phase, a mobile station performs a random exponential backoff with its CW[1]. The mobile station chooses a backoff time between 0 and CW^{1st}. When the backoff timer reaches zero, it enters the second contention phase. After it transmits its frame in the second contention phase (to be explained below), it goes back to the first contention phase. If the channel is busy during backoff, the mobile station freezes its timer. When an erroneous frame or a collision is detected, the mobile station waits for the time period of an EIFS before resuming contention, which is what also happens with the DCF. When a successful transmission is detected, the mobile station waits for the length of a DIFS. Fig. 2 is the state diagram for the first contention phase.

The range of CW^{1st} is CW_{min} to CW_{max}. The value of CW_{max} is the same as it is in the IEEE 802.11 DCF. But CW_{min} has a different value. Otherwise, H-DCF would have more overhead than the IEEE 802.11 DCF due to the longer backoff time resulting from two contention phases. To reduce this overhead, we configure the CW_{min} of H-DCF to a half the value that it has in the IEEE 802.11 DCF[2].

3.1.2 The second contention phase

Eligible stations whose backoff timers reached zero in the first contention phase start the second contention phase by transmitting null frames. A null frame is a relatively short frame whose transmission time is the length of one slot. The null frame plays an important role in demarcating the first and the second contention phases of H-DCF. Meanwhile, as we have already seen, stations whose backoff timers are still running will now have to wait until the second contention phase is over.

Each eligible station chooses a backoff time between 0 and CW^{2nd}. When its backoff timer reaches zero, that mobile station transmits its frame. After that transmission ends, the mobile station goes back to the first contention phase. If the frame transmission fails, the mobile station doubles its CW^{1st} and attempts transmission again. Otherwise, its CW^{1st} is reset to CW_{min}. Meanwhile, the remaining eligible stations whose backoff timers are still running wait for a DIFS after the transmission of the winning station ends and then transmit null frames again. Mobile stations that are still in the first contention phase cannot join the second contention phase, which now continues with the eligible stations choosing a new backoff time between 0 and CW^{2nd}. This process allows all eligible stations to transmit their frames eventually, and when this has occurred the second contention phase concludes. Fig. 3 is the state diagram for the second contention phase.

In the second contention phase, CW^{2nd} is fixed to 7. Thus, the maximum backoff time is 7 slots. When mobile stations taking part in the first contention phase detect a null frame, they do not decrease their backoff timer for the length of an EIFS (e.g. 18.2 slots in IEEE 802.11b). During the EIFS, eligible stations start transmissions, which effectively freezes the backoff timers of the mobile stations that remain in the first contention phase.

[1]To distinguish the CWs in each contention phase, we will call the CW of the first phase CW^{1st} and the CW of the second phase CW^{2nd}.

[2]In IEEE 802.11b, CW_{min} of the DCF is 31, and so we set CW_{min} of H-DCF to 15.

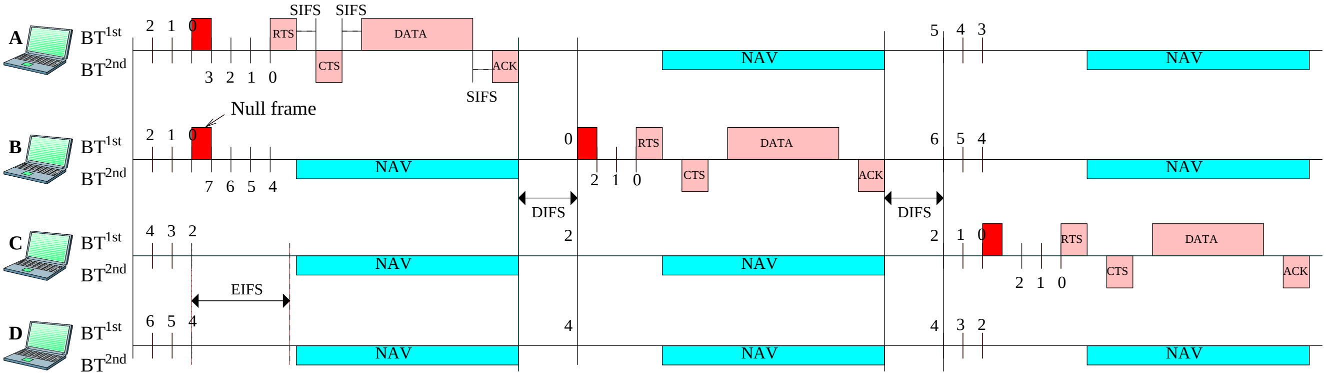


Figure 4. Timing diagram of H-DCF.

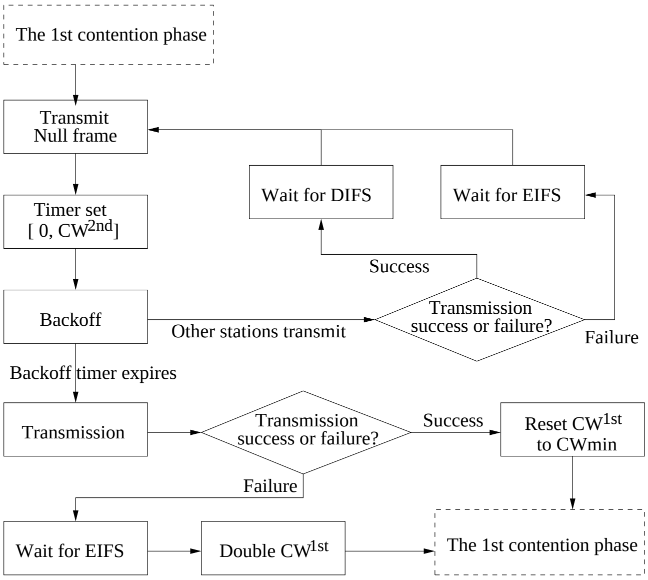


Figure 3. Flow chart of the second contention phase.

3.2 Illustration of the H-DCF Protocol

Fig. 4 shows the basic operation of H-DCF. BT^{1st} and BT^{2nd} are the backoff times chosen for the first and second contention phases. In the first contention phase, the backoff timers of mobile stations *A* and *B* reach zero at the same time while those of the other mobile stations *C* and *D* are still running. Mobile stations *A* and *B* now transmit null frames and then enter the second contention phase. Mobile stations *C* and *D* receive the null frames, which force them to wait for an EIFS before continuing to decrement their backoff timers. As mentioned before, mobile stations in the first contention phase cannot participate in the second contention phase due to transmissions of null frames sent by eligible stations, if any, in the second contention phase. The eligible stations in the second contention phase are *A* and *B*, and they now contend for the wireless channel until the EIFS expires. If the BT^{2nd} of mobile station *A* reaches zero before that of *B*, then *A* transmits its data frame. If the frame transmission is successful, mobile station *A* resets its CW^{1st} to CW_{min}, and then returns to the first contention phase. If its frame transmission fails, mobile station *A* doubles its CW^{1st} instead of resetting it to CW_{min}. The other eligible station in the second contention phase, which is *B*, waits for a DIFS after mobile station *A*'s transmission and then transmits another null frame. *B* resets its BT^{2nd} (to 2 in Fig. 4) and transmits its data frame when the timer expires. Until both the eligible stations in the second contention phase, namely *A* and *B*, have finished their frame transmissions, each of the other two mobile stations *C* and *D* must freeze their BT^{1st} and wait for the end of the second contention phase. When the second contention phase is finally completed, all the mobile stations can start contention for the wireless channel once more, by entering the first contention phase.

4 Enhanced Hybrid-DCF (EH-DCF)

When H-DCF-compliant mobile stations and IEEE 802.11 DCF-compliant mobile stations are contending together in the same cell, the DCF-compliant mobile stations will be starved since CW^{1st} is less than the CW of the DCF and a null frame will cause the DCF-compliant mobile stations to freeze during the second contention phase, at least once during every interval of length $(CW^{1st} + CW^{2nd})$. This means that DCF stations will not be able to access the channel if H-DCF stations have any frames to send. We have therefore introduced Enhanced Hybrid-DCF (EH-DCF) to improve fairness with the IEEE 802.11 DCF. To reserve a proportion of the wireless channel for DCF stations, we employ a channel occupancy regulation algorithm that stops H-DCF stations from transmitting when the ratio of H-DCF traffic is higher than a pre-determined threshold. By controlling the channel occupancy of H-DCF stations, appropriate utilization is guaranteed to DCF-compliant mobile stations. In addition to improving fairness, we make the length of a null frame variable in EH-DCF, so as to reduce the number of mobile stations that take part in the second contention phase.

4.1 Channel Occupancy Regulation for Fairness between DCF and EH-DCF

To prevent H-DCF from monopolizing, we need to yield an appropriate portion of network capacity to DCF in the same cell. We will see how we can allocate a proportion of network capacity to DCF-compliant mobile stations. The proportion can be set manually by an administrator, or dynamically by the network management system.

Using EH-DCF, all frame transmissions happen in the second contention phase. If we can measure how much of the link capacity is being used by this phase, we can estimate the channel occupancy of mobile stations using EH-DCF. The transmission of a null frame means that a data frame will be transmitted during the CW^{2nd}. Hence, we can estimate the channel occupancy of EH-DCF from the number of null frames on the wireless channel. Each mobile station using EH-DCF counts the number of null frames and calculates the relative channel occupancy of EH-DCF by dividing the number of null frames by the duration of the count.

To limit the proportion of the channel used by EH-DCF stations, an EH-DCF compatible station calculates the relative channel occupancy whenever a data frame is transmitted by any stations (DCF or EH-DCF). If the occupancy (fr_O) of EH-DCF compatibles is greater than a pre-defined threshold, then that mobile station delays its transmission by setting its NAV to the excessive duration of the transmission. R is the target proportion of channel capacity to be used by EH-DCF compatible stations. If EH-DCF stations count n_{null} null frames during a time $t_{measure}$, then the proportion of channel capacity used by EH-DCF stations and the value of NAV can be determined by the following equations;

$$fr_O = \frac{n_{null} \times T_{2nd\_phase}}{t_{measure}}$$

$$T_{2nd\_phase} = T_{null} + T_{BO\_2nd} + T_{TX}$$

$$NAV = \frac{n_{null}}{R} \times T_{2nd\_phase} - t_{measure}.$$

T_{2nd_phase} is time required for a series of EH-DCF transmissions, consisting of a null frame, the average backoff time in the second contention phase, and the transmission time of DATA and ACK frames plus SIFS.

4.2 Improving Throughput

Although H-DCF improves the network performance of WLANs, more mobile stations will now enter the second contention phase if the number of competing mobile stations becomes large. The probability of a frame collision will also grow because CW^{2nd} is fixed at 7. To reduce the number of mobile stations entering the second contention phase, we insert an additional contention process between the first and the second contention phases by changing the length of a null frame so that it is either the length of one slot, or two. Eligible stations which transmit shorter null frames lose the contention, and then they have to wait for an EIFS, so that only stations that transmit longer null frames will contend. The eligible stations that lose in the second contention do not go back to the first contention phase and transmit null frames, whose length is one or two slots, after the winning station has transmitted its data frames. Intuitively, we can see that fewer frame collisions will happen as the possible length of a null frame becomes larger. However, if the maximum length of a null frame were to be 3 slots or greater, the overall performance of EH-DCF would be damaged by the increased contention overhead. That is why a null frame is eigher 1 or 2 slots in length.

4.3 Illustration of EH-DCF Protocol

Fig. 5 shows the operation of EH-DCF. A mobile station transmits a null frame for the duration of 1 or 2 slots when its BT^{1st} reaches zero, which starts the second contention phase. After eligible station *C* has transmitted a short null frame, it detects the longer null frames that have been sent by the other eligible stations, which are *A* and *B* in this case. Eligible station *C* now has to wait for EIFS, as though it were in the first contention phase. This forces eligible station *C* to give up the wireless channel access for an EIFS. Eligible stations *A* and *B*, which have transmitted long null frames, now enter the second contention phase and the winner, which is *A*, transmits its data frame. After a DIFS has elapsed, the remaining eligible stations *B* and *C* transmit null frames again which begins the next second contention phase.

5 Performance Evaluation

We implemented and evaluated H-DCF and EH-DCF using the ns-2 simulator [12]. The performance of these schemes is compared with the IEEE 802.11 DCF in terms of scalability and fairness. We measured how the number of frame collisions and the aggregate throughput varies with the number of competing mobile stations. In our simulations, IEEE 802.11a and IEEE 802.11b PHY parameters are used (See Table 1), and the size of each data frame is 1000

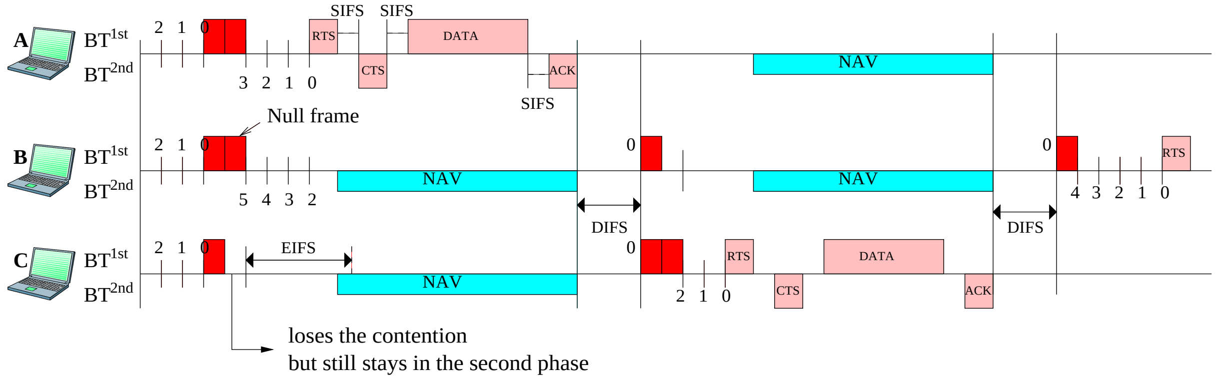


Figure 5. Timing diagram of EH-DCF.

bytes. All mobile stations are assumed to be located within range of an access point at the maximum transmission rate.

Table 1. IEEE 802.11a/b PHY parameters

Time	IEEE 802.11a	IEEE 802.11b
slot time	9 μs	20 μs
SIFS	16 μs	10 μs
DIFS	34 μs	50 μs
EIFS	89 μs	364 μs
CW	15 to 1023	31 to 1023

For H-DCF to operate correctly, the EIFS must be longer than 7 time slots. If this condition is not satisfied, the second contention phase may be interrupted by mobile stations which are still in the first contention phase. The parameters in Table 1 satisfy this requirement, achieving the IEEE 802.11a/b PHYs to be used for H-DCF. The following subsections show the results of our simulations.

5.1 Scalability of H-DCF and EH-DCF

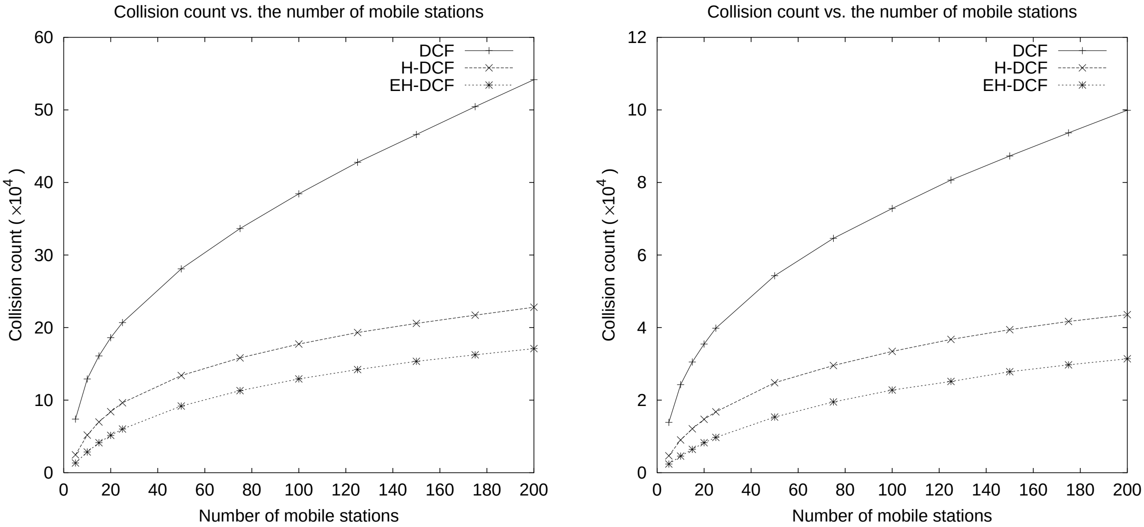
The main aim of H-DCF is to improve the scalability of WLANs. We therefore simulated between 1 and 200 competing mobile stations on a WLAN. Each mobile station attempts to transmit data frames to one access point, and always has backlogged data frames. As shown in Figs. 6 and 7, H-DCF and EH-DCF always achieves higher throughput than DCF. When the number of competing mobile stations is small, the difference between their throughputs is marginal because there is little contention for the wireless channel. But as the number of mobile stations increases, contention for the wireless channel becomes severe and the probability of frame collision increases. Under these circumstances, both H-DCF and EH-DCF reduce the number of frame collisions, resulting in a higher aggregate throughput. These two schemes are much less affected by the growth of the number of mobile stations than DCF. Even when the number of competing mobile stations is over 50, there is little loss of throughput with H-DCF or EH-DCF (Fig. 7). There seems to be a clear benefit from the extra contention phase, and the changing length of a null frame made by EH-DCF achieves a small additional gain in throughput.

5.2 Fairness between DCF and EH-DCF

We will now see what happens when DCF stations contend with stations running EH-DCF, which is designed to achieve fairness in this situation. Fig. 8 shows the throughput of DCF and EH-DCF with IEEE 802.11a and IEEE 802.11b PHYs. EH-DCF guarantees a certain fraction of the channel capacity to DCF by limiting its own channel occupancy to a pre-defined threshold. Fig. 8 shows the results when that threshold is 0.5. Although the regulation algorithm is working and achieving party in channel occupancy, but the ratio between throughput of EH-DCF and DCF remains at 2 : 1 irrespective of the number of competing mobile stations. This is because the CW of EH-DCF is generally smaller than the CW of the DCF, allowing EH-DCF to utilize the wireless channel more aggressively than DCF. The low probability of frame collision with EH-DCF reinforces this effect. That is why the throughput ratio is not exactly 1 : 1 even though the channel is fairly allocated.

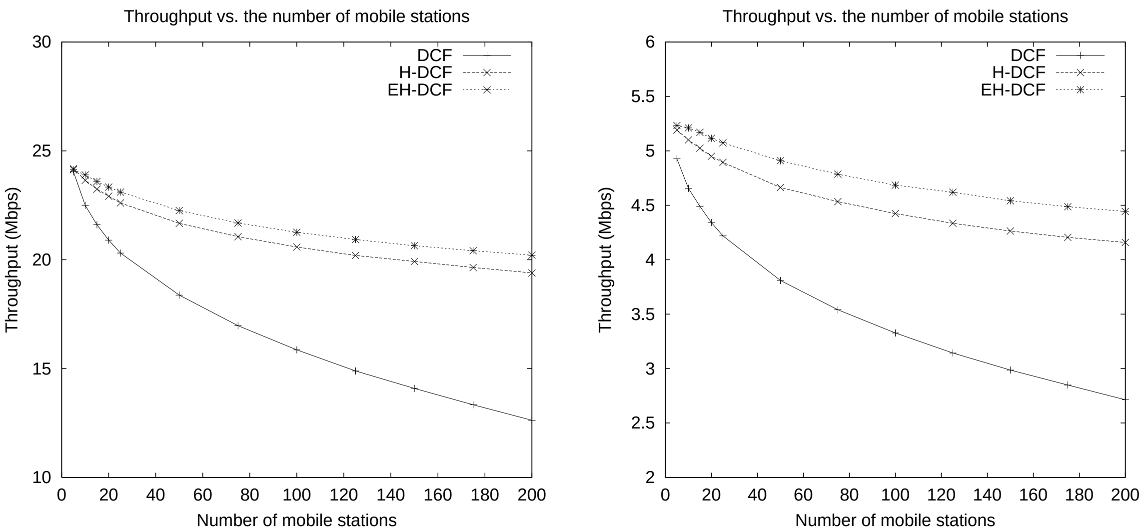
6 Conclusions

WLANs are required to accommodate large number of users in hotspots, and also to provide each user with high bandwidth. The IEEE 802.11 DCF is widely used in such an environment, but it has the disadvantage that network performance degrades due to the growth of frame collisions as the number of competing mobile stations becomes large. To solve this problem we propose a novel distributed channel access function called Hybrid DCF (H-DCF) to increase



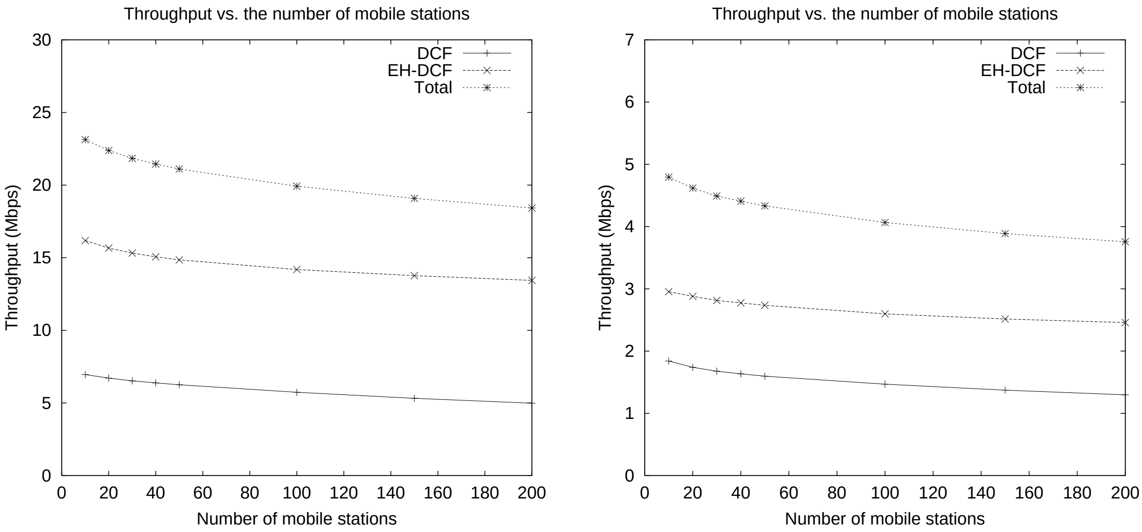
(a) IEEE 802.11a PHY (b) IEEE 802.11b PHY

Figure 6. Number of collisions against number of mobile stations for 3 MAC schemes.



(a) IEEE 802.11a PHY (b) IEEE 802.11b PHY

Figure 7. Throughput against number of mobile stations for 3 MAC schemes.



(a) IEEE 802.11a PHY (b) IEEE 802.11b PHY

Figure 8. Throughput against number of mobile stations for a mixed scenario of two MAC schemes.

channel utilization. H-DCF uses two contention phases demarcated by a null frame to decrease the probability of frame collision. However, if there are mobile stations using DCF and others using H-DCF in the same cell, the mobile stations using DCF may experience severe starvation of bandwidth compared to those using H-DCF. Therefore, we go on to introduce a more advanced scheme called Enhanced H-DCF (EH-DCF) that guarantees fairness between mobile stations running DCF and those running EH-DCF by restricting the channel occupancy of the EH-DCF stations. EH-DCF also uses a variable null frame, occupying one or two time slots, which reduces the probability of frame collision and improves throughput. By means of comprehensive ns-2 simulations, we have demonstrated that H-DCF and EH-DCF really do improve the aggregate throughput, and that EH-DCF will indeed reserve a fraction of the channel capacity for DCF stations.

References

[1] IEEE WG. *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. IEEE 802.11 Standard, 1999.

[2] IEEE WG. *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: High-speed Physical Layer in the 5 GHz Band*. IEEE 802.11 Standard, 1999.

[3] IEEE WG. *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-speed Physical Layer Extension in the 2.4 GHz Band*. IEEE 802.11 Standard, 1999.

[4] IEEE WG. *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Further Higher Data Rate Extension in the 2.4 GHz Band*. IEEE 802.11 Standard, 2003.

[5] IEEE WG. *TGn Sync Proposal Technical Specification*. IEEE 802.11-04/889r2, 2005.

[6] IEEE WG. *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Medium Access Control (MAC) Quality of Service Enhancements*. IEEE 802.11 Standard, 2005.

[7] F. Cali, M. Conti, and E. Gregori. Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit. *IEEE/ACM Transactions on Networking*, 8(6):785–799, December 2000.

[8] Y. Yang, J. Wang, and R. Kravets. Distributed optimal contention window control for elastic traffic in wireless LANs. *Proceedings of IEEE INFOCOM 2005, Miami, USA*, 2005.

[9] I. Aad, Q. Ni, C. Barakat, and T. Turletti. Enhancing IEEE 802.11 MAC in congested environments. *Elsevier Computer Communications Journal*, 28(14):1605–1617, September 2005.

[10] H. Kim and J. C. Hou. Improving protocol capacity for UDP/TCP traffic with model-based frame scheduling in IEEE 802.11-operated WLAN. *IEEE Journal on Selected Areas in Communications*, 22(10):1987–2003, December 2004.

[11] Y. Kwon, Y. Fang, and H. Latchman. Design of MAC protocols with fast collision resolution for wireless local area networks. *IEEE Transactions on Wireless Communications*, 3(3):793–807, May 2004.

[12] VINT group. Ucb/lbnl/vint network simulator - ns (version 2). *http://www.isi.edu/nsnam/ns/*.