IEEE 802.11 over multi-hop wireless networks: problems and new perspectives

Karthikeyan Sundaresan, Hung-Yun Hsieh, Raghupathy Sivakumar *

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, USA

Received 16 June 2003; accepted 10 July 2003

Abstract

The distributed coordination function (DCF) mode of the IEEE 802.11 MAC standard, though proposed for medium access in wireless local area networks, is seen as the de-facto medium access standard in multi-hop wireless networks. In this paper we contend that the unique characteristics that differentiate multi-hop wireless ad-hoc networks from local area wireless networks render the IEEE 802.11 MAC protocol inefficient in ad-hoc networks. Specifically, we focus on the band of contention and the fairness model employed by the IEEE 802.11 MAC protocol in our study. We substantiate our arguments through simulations of idealized (centralized) protocols, and consider the key changes required to adapt the IEEE 802.11 MAC protocol for multi-hop wireless networks. We then propose a simple medium access scheme within the IEEE 802.11 MAC framework, called flow based medium access (FBMA) that achieves significantly better fairness properties while adhering to the purely distributed operations of the basic IEEE 802.11 MAC scheme. We demonstrate the performance of the proposed MAC protocol through simulations.

1. Introduction

Over the last few years, there has been a proliferation of a variety of wireless technologies in the Internet. With the widely anticipated 3G effort into its test deployment phase, future wireless networks are expected to provide users with considerably higher data rates than those offered today. In parallel, the use of wireless technologies in the local area environment has garnered equally significant attention with the rising popularity of the IEEE 802.11a and 802.11b wireless standards [1] that offer users upwards of 54 Mbps and 11 Mbps of data rate respectively. The charm of tetherless networking, growing mobile user population, and the relatively higher bandwidth (when compared to existing 2.5G WWANs) have propelled the popularity of the IEEE 802.11 based wireless networks. Unlike most other technologies in their initial stages, not only are 802.11 based networks being deployed in enterprise environments, but such networks have found use even in public sites and private residences.

The popularity of the IEEE 802.11 MAC protocol, along with its support for infrastructure-less mode of operation (DCF mode), has made it conducive for operation in ad-hoc networks. Briefly, ad-hoc networks are stand-alone wireless networks that lack the services of a backbone...
infrastructure. They consist only of a collection of mobile stations, where the mobile stations double up as forwarders or routers for other mobile stations in the network. Such networks were initially designed for use in military and emergency-relief applications. Lately, the ad-hoc network model has also been proposed and used in other applications such as sensor networks, personal area networks, and regular wireless network applications by virtue of their better spatial reuse characteristics in comparison to the conventional cellular wireless network model [2]. Future wireless network standards including fourth generation wireless standards are expected to incorporate the ad-hoc model in some form [3]. In the meantime, CSMA/CA (carrier sense multiple access with collision avoidance), the MAC scheme used in the IEEE 802.11 DCF mode, has been assumed as the de-facto standard in ad-hoc networks [4–8]. Several of the research works on routing and transport layer protocols in ad-hoc networks assume the use of CSMA/CA as the medium access control protocol in their protocol stack. The key contributing factors for this assumption are (i) the infrastructure-less mode of operation that the DCF mode supports, and (ii) the popularity of the standard itself that requires no additional hardware for the ad-hoc mode of operation.

Given such de-facto acceptance of CSMA/CA as the MAC protocol in ad hoc networks, understanding its performance in such environments has gained significance. The focus of this work is to consider the consequences of using the CSMA/CA protocol in a multi-hop wireless environment. Unfortunately, CSMA/CA being primarily designed for a single-hop wireless network turns out to be inappropriate in a multi-hop wireless network. Although wireless LANs and ad-hoc networks share a few similar characteristics, they differ in the following respects: (i) All the nodes in a WLAN talk to the access point. Since only one node can communicate with the access point at any instant, the contention region encompasses the entire network, eliminating the possibility of any spatial reuse. However, in the case of ad-hoc networks, where any node could serve as a source or a destination, a reduction in the region of contention would potentially increase the degree of spatial reuse in the network and consequently the throughput. Also, the paths in ad-hoc networks typically consist of multiple hops. Hence, routing is an important factor that affects network performance, and the efficiency of the routing protocol used can indirectly depend on the underlying MAC protocol. (ii) In a wireless LAN, since the traffic generated by each node is typically its own, providing per-node fairness is quite reasonable. However, in an ad-hoc network where nodes cooperatively act as relays for other flows, per-node fairness is potentially unfair to heavily loaded nodes and flows traversing such nodes.

We contend that the above differences necessitate changes in the CSMA/CA protocol that are specific to the ad-hoc network environment. Specifically, we study two properties of the CSMA/CA MAC protocol:

- **Band of contention:** The area of the network inhibited by each per-hop transmission such that no other transmissions or receptions can occur within that area. While this property will have a direct impact on the throughput utilization in the network, we demonstrate that its impact has a wider scope including network fairness, and amount of performance gains achieved through better routing protocols.

- **Fairness model:** We show that the per-node fairness model supported by the CSMA/CA protocol significantly lowers both the network throughput and fairness performance. We consider an alternative fairness model and study the performance improvements gained through the new model.

We identify the limitations of the CSMA/CA protocol in ad-hoc networks through performance evaluation. We use both throughput and fairness as metrics in our evaluation. In addition to the basic performance of the CSMA/CA protocol, we also study the impact of the MAC protocol on the performance of the routing layer. Through comprehensive simulation results we substantiate our argument that the CSMA/CA medium access control protocol does not perform well in multi-hop wireless environments. We use the insights gained from the study therein to suggest ap-
proaches to improve its performance over a multi-hop network environment and provide bounds in the form of idealized protocol performance. Finally, we present a variation of the CSMA/CA protocol called flow based medium access (FBMA) that is suitable for multi-hop wireless networks.

The rest of the paper is organized as follows: Section 2 discusses background material on the CSMA/CA protocol and related work. In Section 3, we present the different algorithms that we use in the paper for both the objective evaluation of the CSMA/CA protocol and studying alternative approaches to improve network performance. In Section 4, we study and compare the performance of the CSMA/CA protocol with other alternative, but centralized, approaches. In Section 5, we identify the challenges and present an overview of the distributed version of an alternative approach that is equipped with a better fairness model than CSMA/CA. In Section 6, we discuss in detail the mechanisms and algorithm of the FBMA protocol. We evaluate the performance of the proposed distributed scheme and compare it with that of CSMA/CA for a variety of network parameters in Section 7. Finally, in Section 8 we identify some issues and conclude the paper.

2. Background and related work

2.1. Background

Since the focus of the paper is the evaluation of the CSMA/CA medium access control protocol over multi-hop wireless networks, we provide a brief overview of the CSMA/CA protocol.

2.1.1. The CSMA/CA medium access protocol

The IEEE 802.11 standard specifies two modes for medium access, namely the point coordination function (PCF) and the distributed coordinated function (DCF) modes. While the PCF is a centralized scheme, the DCF is a distributed one. Since the focus of this work is on DCF, we provide an overview of the DCF mode of operation below.

The DCF mode belongs to the carrier sense multiple access with collision avoidance (CSMA/CA) class of protocols. In this scheme, every DATA communication is preceded by an exchange of control packets when the data packet size exceeds a particular threshold. When a source S wants to transmit to a destination D (see Fig. 1), it senses its local channel (physical carrier sensing). If the channel is busy, it backs-off after exponentially increasing its back-off timer. Otherwise, the source transmits a request-to-send (RTS) control message to the destination. If the local channel around D is free, D replies with a clear-to-send (CTS) message, which is then followed by the data packet transmission from S to D, and an acknowledgment (ACK) packet transmission from D to S. If the channel around D is busy, S times out waiting for the CTS message, exponentially backs-off its timeout value and retransmits the RTS packet. We elaborate on the back-off mechanism later in this section. Both RTS and CTS packets contain the proposed duration of the upcoming data transmission. Nodes

![Fig. 1. Operation of CSMA/CA.](image-url)
located in the vicinity of communicating nodes, that overhear either (or both) of these control packets, must defer transmission for this proposed duration. This is called virtual carrier sensing which is performed in addition to the physical carrier sensing mentioned earlier. It is implemented by means of a variable called the network allocation vector (NAV). A node updates the value of its NAV with the duration field specified in the RTS or CTS. Thus the nodes lying within the transmission range of the transmitter or the receiver do not initiate any transmission while the communication is in progress. The RTS and CTS packets thereby reserve the local channel for the upcoming DATA transmission by silencing the nodes in the vicinity of the transmitter and the receiver. This in turn addresses the hidden terminal problem. Since the data packet transmission inhibits both neighbors around the source and the destination, we refer to the band of contention in CSMA/CA as being two.

The CSMA/CA MAC protocol uses a back-off interval to resolve channel contention. A source node $S$, before initiating a transmission chooses a random back-off interval in the range of $[0, cw]$ where $cw$ represents the contention window. The node $S$ then decrements its back-off counter by one after every idle slot time. When the back-off counter reaches 0, node $S$ transmits its packet. If the transmission from $S$ collides with some other transmission, $S$ doubles its $cw$, and chooses a new random back-off interval from the new range and then attempts retransmission. Note that collision of an RTS packet can be detected by the absence of a CTS within a timeout value. The contention window is doubled for every collision until it reaches a maximum threshold called the $cw_{max}$. While in the back-off stage, if a node senses the channel to be busy, then it freezes its back-off counter. When the channel becomes idle once again for a duration DIFS (DCF interframe spacing), the back-off counter is resumed to count down from its frozen value. Thus the back-off procedure is invoked only when the channel has been sensed to be idle for DIFS duration. Resuming the back-off counter from the frozen value ensures that the nodes that have deferred access to the channel for long have a higher probability to access the channel in the current slot. This in turn ensures that over a longer time span all the nodes have equal opportunity to access the channel, thereby achieving per-node fairness. A shorter interframe space, SIFS is used to separate transmissions pertaining to the same data packet (i.e., every node performs physical carrier sensing for SIFS before it actually transmits CTS, DATA or ACK frames). The packet transmissions and spacing values are illustrated in Fig. 1.

2.2. Related work

There has been a significant amount of research in the context of the IEEE 802.11 protocol. However most of them deal with WLANs that are characterized by single-hop flows. In [9,10], the authors evaluate the performance of the IEEE 802.11 protocol over wireless local area networks and identify its unfair performance characteristics. However, the scope of the evaluation is confined to last-hop wireless LAN environments, and does not include multi-hop wireless networks. The focus of [11,12] is to provide service differentiation to the flows based on priorities by varying either the contention window, the interframe spacing or the maximum frame length. The flows considered are one hop away from the access point. Hence flow fairness corresponds to node fairness in this work. The flows are all assumed to have different QoS requirements and consequently different priorities. Since the focus is only on single-hop flows, a central coordinating entity like the access point performs the task of service differentiation. The task group E of the IEEE 802.11 working group is currently working on an extension to the IEEE 802.11 standard called IEEE 802.11e. The proposed access mechanism, enhanced distributed coordinated function (EDCF) [13] combines two measures to provide service differentiation. The minimum contention window and the interframe spacing can be set differently for different priority and traffic classes. As before, the EDCF mechanism is aimed to provide service differentiation in WLANs where the communicating nodes are one hop from the access point, which in turn is different from the multi-hop environment that we consider in this work.
In [14], the authors identify the unfair nature of the IEEE 802.11 MAC protocol over wireless ad-hoc networks. The authors propose a better scheme to provide fairness. However, the scheme proposed is targeted toward achieving better node fairness and does not support flow based fairness addressed in this paper. Hence the performance inefficiency due to the per-node fairness model still exists. In [15], the authors investigate the performance of IEEE 802.11 over multi-hop wireless networks. Although the key conclusion drawn in that work is the same as in this work—that IEEE 802.11 is inappropriate for multi-hop wireless networks, the study is closely tied to evaluating IEEE 802.11's performance using TCP as the transport protocol. Moreover it does not provide insights into attainable performance improvement when the inefficiency with IEEE 802.11 over multi-hop wireless networks is resolved. In contrast, our work is not limited to any particular transport or routing protocol. In addition, we also propose a scheme that addresses the inefficiencies in IEEE 802.11's operation over ad-hoc networks. In [16], the authors advocate a per-flow fairness model. They consider distributed implementations of local and global fairness models, both of which require the construction of the conflict-free minimum spanning tree for the node graph. This tree is required to propagate flow information to the entire network or to the neighboring nodes depending on the fairness model used. The construction of the tree incurs overhead and also increases the complexity of the scheme. In contrast, the distributed local fairness model that we propose later in the paper, does away with the tree construction by using local coordination between the nodes and simple piggy-backing mechanisms.

3. Idealized protocols

In the case of WLANs, the environments for which CSMA/CA was initially designed, all the nodes talk to the access point that is one hop away. Since only one of the nodes can talk to the access point at any instant, the performance is not affected by the band of contention. Moreover all the flows are single-hop flows. As a result, providing per-node fairness is quite reasonable. However, in a multi-hop wireless environment, this is not the case. A band of contention of one as opposed to two would potentially increase the degree of spatial reuse and hence the network performance. Further, since flows traverse multiple hops, it is possible that some nodes are traversed by more flows than the others. In such a case, providing per-node fairness would not amount to providing per-flow fairness. Hence from the perspective of end-to-end flows the degree of fairness would be severely impacted. To corroborate the above issues, we proceed to evaluate the performance of CSMA/CA with respect to the two properties of band of contention and the fairness model. However, to hide the inefficiencies due to the distributed operation of the protocol, we take the approach of evaluating the performance using idealized MAC and routing protocols.

3.1. Idealized MAC protocols

In order to focus on the impact of band of contention and the fairness model, and mask the overheads of the CSMA/CA MAC implementation and inefficiencies due to the sub-optimal distributed operation (e.g. unnecessary idle or collided slots due to back-offs), we implement, and use a centralized version of the CSMA/CA MAC protocol during its evaluation and comparison with other approaches.

1. Ideal node scheduling—band 2 (INS-2): In the idealized MAC protocol, we add a transmission scheduler object to the network simulator. When the simulation begins, the MAC protocol at every node in the network registers with the centralized scheduler if it has a packet to transmit. The scheduler, for every transmission slot, chooses the node that has received the minimum service thus far. When more than one node with the minimum service counter exists, the node with the minimum two hop degree is chosen. Since the band of contention is two, a transmitting node will preclude any other node within its two hop neighborhood from transmitting simultaneously. Hence by choosing the node with the minimum two hop degree to transmit, we make it possible for more
simultaneous transmissions in the network, thereby increasing spatial reuse and consequently network utilization. Based on the first choice, it finds the second node that has received the minimum service among the other nodes and can transmit without interfering with the first transmission. The process continues until no more node transmissions can be accommodated for that transmission slot. We refer to the band of contention as being two since like the CSMA/CA protocol, the centralized scheduler does not allow any transmissions or receptions to occur in the vicinity of any transmitter or receiver. The protocol is referred to as the INS-2 protocol in the rest of the paper.

2. Ideal node scheduling—band 1 (INS-1): The INS-1 protocol is similar to the INS-2 protocol in that a centralized scheduler is used to achieve the scheduling. However, unlike INS-2 where the band of contention is two, and hence no transmitters or receivers are allowed in the vicinity of both the transmitter and receiver, INS-1 has a band of inhibition of only one: transmissions are allowed subject to the condition there can be no other transmissions in the vicinity of a receiver, or no other receptions in the vicinity of a transmitter. The choice of the nodes for transmission is based on the service enjoyed by the nodes until that point. Ties are broken based on the two hop node degree as before.

3. Ideal flow scheduling (IFS): The INS-1 and INS-2 protocols are scheduling protocols where the fairness model is node based. In other words, service counters are maintained purely on a node by node basis, and nodes are chosen for transmissions. In IFS, the centralized scheduler is responsible for scheduling flows instead of nodes. Service counters are maintained per flow and not per node. When a flow is scheduled for transmission, all hop-by-hop transmissions for that flow are scheduled sequentially. If after a flow is scheduled but before its first hop transmission commences, another flow with a lower service counter arrives, the former flow may be re-scheduled to accommodate the latter flow to ensure short-term fairness. However, if the latter flow arrives after the one or more hop-transmissions of the former flow, no such pre-emption is performed. The band of contention is equal to one in the IFS implementation. IFS being the crux of this work, we continue to provide a more detailed explanation of the protocol.

Essentially, the centralized scheduler periodically draws up a schedule for multi-hop transmissions within the network based on the information provided by nodes. The transmission schedule maximizes throughput subject to a fair per-flow service. The scheduler then broadcasts the schedule to the nodes. The algorithm used by the centralized scheduler to perform the scheduling and the variables used in the algorithm are shown in Fig. 2.

Periodically, the node \( n \) updates the centralized scheduler with the location information \( L(n) \) such as neighbor list or GPS location (line 1). It also informs the scheduler the service backlog \( B(f) \) of the flow \( f \) for which it acts as the source node (line 2). Upon receipt of the updated information, the scheduler computes a connection matrix \( CM \) of the network and the optimal route \( R(f) \) for each flow \( f \) (lines 4–6). For each scheduling-period (every \( sp \) time slots), the scheduler iterates through the list of flows with backlogged services (line 10), and for each flow schedules the hops along the path that the flow traverses (lines 21–31). Once all the flows are accommodated within the schedule, the scheduler iterates once again through the schedule and attempts to fill in more end-to-end transmissions for the flows within the schedule (lines 11–20). The process is repeated until the schedule cannot be filled in with any flow. A transmission schedule \( SR \) of length \( sp \) is broadcast to nodes every \( sp \) time slots (line 3). The centralized scheduler always tries to provide fair service before trying to enhance throughput. In other words, when the “refilling” process is done, flows with less service are provided priority over flows with more service. Flows that have schedules beyond the current scheduling-period have slots reserved during the next scheduling-period (\( SO \)) irrespective of the newly contending flows during the next scheduling operation (i.e. flows once scheduled are not preempted).

At each node, a single output queue is maintained for all packets to be forwarded. When the MAC layer requests for a packet from a specific flow (according to the schedule drawn by the
3.2. Idealized routing protocols

1. Shortest path routing (SPR): Similar to our centralized implementation of the CSMA/CA MAC protocol, we also implement a centralized and ideal version of a shortest path routing protocol. Once the network is initialized, the centralized routing protocol computes the shortest paths between every source destination pair in the network and updates the routing tables in the network accordingly. We still use the dynamic source routing (DSR) [4] as the routing layer. However, the routes for DSR are furnished through the centralized routing module instead of through its route re-computation module.

2. Widest-shortest path routing (WSR): In order to demonstrate the effect of the MAC protocol on the performance gains achieved through a better routing layer, we use a load-balanced routing algorithm called WSR. The advantages gained by employing a reduced band of contention and a per-flow fairness model with the SPR routing protocol are further enhanced by the use of WSR. The reduction in the band of contention to one aids WSR by reducing the amount of physical coupling between the routes. This in turn increases the probability of finding decoupled routes and consequently helps improve the network utilization. Unlike SPR which is based on a single metric—the hop-count—for route selection, WSR uses a 3-tuple—(interference along a path, interference caused by the flow, hop-count)—to choose from the available set of paths. Each link is associated with a single weight. When a flow is assigned a particular path, the weights of all links that will contend with the flow are incremented. The first parameter of the 3-tuple is the maximum of the link-weights of the path being considered. The second parameter is the aggregate increase in weights of the links due the considered path if the flow were to be assigned to it. The third parameter is a simple hop-count. The path that has the minimum lexicographic value for the 3-tuple is chosen by WSR. WSR is also implemented as an

---

**Fig. 2. Algorithm for the IFS protocol.**

---

There exists another subtle benefit which results from performing flow scheduling. We highlight this benefit in our discussion of results.
independent centralized routing module that furnishes routes to DSR.

4. Performance study of CSMA/CA

We present details of the simulation environment used for our evaluations of CSMA/CA in this section. Specifically, we describe the physical layer, topology, traffic model, and metrics used in the simulations. We use the ns-2 network simulator for all the simulations presented in the paper [17].

- **Physical layer**: A combination of the free space propagation and two-ray ground reflection model is used to model the signal propagation in the simulations. The signal strength falls as $1/r^2$ ($r$ is the distance) within a constant cross-over distance. Above the cross-over distance, the signal strength falls as $1/r^4$. The cross-over distance used for all our simulations is 100 m.

- **Topology**: We use a 1500 m $\times$ 1500 m network grid with 100 nodes randomly distributed within the grid for all the simulations. The random seed for the topology creation is varied for the different scenarios used. Twenty different scenarios are used for each data point in the simulation results. A constant transmission range of 250 m is used. We do not consider mobile scenarios in this evaluation since the focus is on evaluating the performance of the MAC protocol alone.

- **Traffic model**: We primarily use 100 TCP flows as the traffic content in the network where the source and destination pair for each TCP flow is randomly chosen from the set of 100 nodes. CBR sources are used to feed the TCP flows. Different loads from 16 to 256 Kbps are used in the simulations. In order to depict the validity of our results for other traffic scenarios, we also use UDP flows for one set of results with a different traffic model wherein sources can be clustered instead of being randomly distributed.

- **Metrics**: The average throughput and normalized throughput deviation are used as measures of the throughput and fairness performance. The throughput deviation is normalized to the average throughput for the fairness measure. Each simulation is run for 60 s, and results of 20 different scenarios are averaged for every data point shown in the results.

We present simulation results that demonstrate the performance of the CSMA/CA MAC protocol in a multi-hop wireless network. Also, we present simulation results for the other medium access approaches briefly described in Section 3.1. We organize the rest of the section into four parts: (i) CSMA/CA vs. INS-2: We present the performance difference between the centralized and distributed versions of the CSMA/CA protocol, (ii) INS-2 vs. INS-1: We then identify the throughput performance enhancement provided by reducing the band of contention of the CSMA/CA protocol from two to one. We further show the impact of the same on the network fairness and performance gains achieved through employing the load balancing WSR routing algorithm described in Section 3.2. (iii) INS-1 vs. IFS: We present performance enhancements in terms of both throughput and fairness when the fairness model of CSMA/CA is enhanced from a per-node fairness model to a per-flow fairness model. We also show the impact of such an upgrade in the fairness model when a transport protocol other than TCP is employed. (iv) Furthermore, in order to substantiate that our arguments remain valid even when operating over different traffic and network models, we use a clustered traffic distribution model with UDP traffic and compare the throughput and fairness performance of the four flavors of the MAC protocol considered.

4.1. Protocol inefficiency (CSMA/CA vs. INS-2)

Fig. 3(a) and (b) show the performance difference between CSMA/CA and INS-2 in terms of throughput and fairness. In INS-2, when the centralized scheduler permits a particular station to transmit, the station still performs the RTS–CTS–DATA–ACK exchange as in CSMA/CA. Hence, the throughput performance improvement shown for INS-2 in the figures is a direct measure of the fact that INS-2 is a centralized approach and thus does not suffer from the inefficiencies of packet collisions and unnecessary contention based back-
offs as in the distributed CSMA/CA protocol. Similarly, the fairness performance improvement is a measure of the unfairness properties of CSMA/CA due to its distributed nature. Although the CSMA/CA fairness model is based on node fairness, the unfair nature of CSMA/CA has also been profiled in related works [9,10,14]. However, note that the focus of this paper is not to study the performance enhancements achieved through a centralized operation of CSMA/CA. Rather, the goal of this section is to consider a centralized implementation such that subsequent comparisons presented in the rest of the section are fair.

4.2. Effect of reduced band of contention (INS-2 vs. INS-1)

Fig. 4(a) and (b) show the throughput and fairness performance improvement achieved over the performance of INS-2 when the band of contention is reduced from two to one. Details on how the reduction is performed were furnished in Section 3.1. The throughput enhancement is obvious since reducing the band of contention results in an immediate increase in the amount of spatial reuse in the network. However, it is interesting to note that the fairness also improves when the band of contention is reduced. This is a result of the reduced band of contention allowing less privileged flows to catch up to the more privileged flows in terms of throughput. To observe the impact of reducing band of contention on the routing protocols, Fig. 5 shows the performance enhancement achieved when WSR is used in place of SPR. The performance enhancement due to different routing algorithms can be seen when INS-2 is used and when INS-1 is used. The absolute values of the INS-1 results are greater than the INS-2 results as expected, yet from the spacing between the curves, it can be made out that the gains when using WSR is greater than using SPR when the underlying medium access control layer is INS-1. This can be intuitively explained as follows: WSR attempts to distribute flows in the network, whereby it reduces the degree of coupling between the routes such that they do not contend with each other. However, such a distribution will be beneficial to overall network utilization only until there are unused resources in the network. When INS-1 is used, because of the smaller band of contention, such a saturation point (in terms of resource usage) is reached much later than in the case of INS-2 resulting in better overall network utilization.

4.3. Effect of fairness model (INS-1 vs. IFS)

Fig. 6(a) and (b) present the throughput and fairness results for INS-1 and IFS respectively.
Recall that the difference between the two protocols is purely in the fairness model supported. As expected, Fig. 6(b) demonstrates a significant improvement (close to 40% for a load of 128 Kbps) in terms of fairness when IFS is used. However, it is interesting to note an improvement, albeit a small one, in terms of throughput also. The reason for the throughput improvement can be explained by the absence of losses in the IFS model (since one hop transmission of a packet means that the other hops have also been scheduled) resulting in a better utilization of network resources. However, that the improvement in the throughput performance looks marginal can be explained by the nature of the transport protocol used. TCP is an adaptive transport protocol that reacts to losses. Hence, the use of TCP ensures that not many such losses in the network occur. In order to substantiate this observation, in Fig. 7, we present the throughput improvement seen for IFS when UDP is used as the transport protocol. It can be observed that the throughput improvement is significantly larger than in the case of using TCP as the transport layer.

### 4.4. Effect of clustered traffic distribution

For all simulation results presented thus far, a randomly distributed traffic model is used in that sources and destinations are randomly selected from the 100 node population. In this section we present representative results that demonstrate the observations made earlier in the section about the in-efficacy of the CSMA/CA protocol or the efficacy of the other schemes still holds good when other traffic models are considered. We consider a clustered traffic distribution scheme wherein the 100 flows in the network have \( k \) sources, and \( k \) is varied from 1 to 100. When \( k \) is 100, the scenario is
the same as the ones presented earlier in the section. Such a traffic distribution is also realistic when considered in the typical Internet client–server realm where multiple clients can access the same server, or in wireless sensor networks where traffic in the network might be targeted toward few sinks. Fig. 8(a) and (b) show the throughput and fairness performance of all flavors of the MAC protocol considered thus far when the source cluster size is varied (results for $k > 25$ are not shown in the figures as the trend remains the same beyond $k = 25$). While it is evident that throughput performance remains consistent with our earlier discussions, fairness can become a serious issue when the sources are clustered and IFS remains the only protocol that can effectively address the issue of fairness even in such a heavily shared environment.

4.5. Summary

We have demonstrated in this section that the CSMA/CA MAC protocol can be significantly improved both in terms of its band of contention and in terms of its fairness model for wireless multi-hop networks. We have shown that the performance of the MAC protocol not only has a direct impact on the performance of the network, but also indirectly impacts the performance gains achieved through using smarter higher layer protocols. Finally, we have also shown that our observations remain valid under different conditions of transport protocols, and traffic distribution scenarios. These studies help us realize that a band of contention of one and a per-flow fairness model are essential for the effective and efficient functioning of CSMA/CA in multi-hop wireless networks.
5. Flow based medium access—challenges and overview

Inspite of its significant performance improvement, IFS (with a band of contention of one) still cannot be used in ad-hoc networks owing to its requirement of central coordination. Hence we focus on a distributed version of the centralized IFS algorithm called flow based medium access (FBMA) in this section. While a straight-forward approach is to design a MAC protocol from scratch that addresses the inefficiencies in CSMA/CA’s operation over ad-hoc networks, we have taken an alternative approach. Since 802.11’s DCF mode of operation (CSMA/CA) has become the de-facto standard for ad-hoc networks and owing to its gaining momentum and popularity, we have implemented our distributed scheme, FBMA within the IEEE 802.11 MAC framework by effecting changes to its mechanism of operation. FBMA retains the functionalities of CSMA/CA as is, and supplements it with additional mechanisms in achieving its objectives. We elaborate on this in the later sections. Since FBMA is implemented within the IEEE 802.11 MAC framework, it has a band of contention of two as opposed to IFS, that has a band of contention of one. Note that this is inevitable for any MAC scheme making use of two-way control packet handshakes (RTS–CTS, DATA–ACK) as in CSMA/CA. Hence the goal of FBMA is directed toward improving only the fairness model in a distributed fashion while employing a band of contention of two. However, when we discuss the impact of better routing protocols, we show that FBMA is also able to achieve an improvement in throughput when load-balanced routing protocols such as the WSR are employed.

In this section, we introduce the FBMA scheme. We first identify the challenges in the realization of the FBMA scheme and then provide an overview of FBMA.

5.1. Challenges

1. Fair contention: Since CSMA/CA provides per-node fairness, there is no distinction between nodes belonging to the same contention region as far as the contention resolution algorithm is concerned. Hence the nodes belonging to the same contention region will have the same probability of accessing the channel. However, due to the possibility of some nodes servicing more flows than the others, a per-node fairness model would fail to provide per-flow fairness. Hence it becomes necessary to distinguish nodes that service more flows from those that service lesser number of flows. This requires that at a high-level some form of
priority be incorporated in the contention resolution algorithm such that nodes that service more flows than the others have a higher probability of access to the channel. However, note that a mere priority assignment to the nodes based on the number of flows serviced by them would not achieve a per-flow fairness model. To illustrate this point, consider a multi-hop flow that traverses several hops. Let the node upstream (closer to the source) be a high priority node, servicing more flows than other nodes in its local neighborhood. Hence this node will have a higher probability of access to the channel. Now let the node downstream (closer to the destination) be a low priority node, servicing fewer flows than other nodes in its local neighborhood. Hence this node will have a lower probability of access to the channel. From the flow’s perspective, though the upstream node is able to forward the packet belonging to the flow, the downstream node is not able to do so owing to its lower channel access probability. Hence the high priority assignment to the upstream node does not directly translate to a higher service for the flow. Instead a node that services a backlogged flow with the lowest service counter in its local neighborhood should obtain a higher probability of access to the channel than any other node in its neighborhood.

2. **Fair queuing**: Once the nodes have been prioritized based on the service obtained by the flows traversing them, it becomes essential that this priority in channel access (assigned to node) is experienced by the packet that belongs to the flow with the lowest service amongst all other packets in the queue at the node. This needs to be ensured in order to provide fairness on a per-flow basis. Further, this is achieved by modifying the queuing operations within the node (intra-node fairness). If we consider fairness amongst the flows with respect to the throughput obtained by them, then at any instant \(t\), the distribution of the service obtained by the various flows (throughput distribution) is an indication of the degree of fairness in the network over the time window \([0,t]\). Although the nodes have a notion of priority in accessing the channel based on the service (throughput) obtained by the flows traversing them, this comes into effect only after a packet has been dequeued from the queue. Hence, it is equally important that the queue also implements some notion of priority in dequeuing packets that belong to a flow having a lower service, before those that belong to flows having a higher service.

We elaborate on each of these issues with the appropriate mechanisms addressing them in Section 6 where we present our FBMA scheme.

### 5.2. Overview

In the CSMA/CA protocol, the “contention window” \(cw\) parameter plays a vital role in determining a node’s access to the channel. It represents the time in number of slots (each slot being 20 \(\mu\)s) that a node backs off before transmitting. By varying this parameter we can vary the time for which a node waits before transmitting when it finds the channel to be idle. The contention window specifies a maximum value \(cw\) and the node chooses a random value between 0 and \(cw\). This represents the number of slots that the node will wait before it tries to access the channel again. The contention window parameter is adapted in FBMA to achieve per flow fairness. The intuition behind adapting (scaling) the contention window parameter is to make packets belonging to flows with a lower service counter contend with a smaller \(cw\) value than those belonging to flows with a higher service counter. This would in turn ensure that flows with a lower service are provided with a higher probability to access the channel than those with a higher service, thereby helping them catch up on the service difference.

The MAC layer first dequeues the packet belonging to the flow with the lowest service among all packets in its queue. The contention window at the node is then scaled during the transmission of the packet depending on the service obtained by the flow to which the packet belongs. By service, we refer to the total number of packets that the destination of a flow has obtained at any instant. The service obtained by the flow is normalized with the minimum service in the network to obtain the scaling factor for the contention window. Since only the flows in the two hop neighborhood are involved in channel contention, every node needs information only about the service counters of the
flows that the nodes in its two hop neighborhood are servicing in determining the minimum service. This flow information in the local neighborhood is made available at the nodes by means of exchange (propagation) through piggy-backing onto control packets. Each of these mechanisms is explained in detail in Section 6. On determining the scaling factor, the node scales its contention window with this factor and contends for channel access with its new contention window parameter. This would ensure that the higher the service for a flow, the higher will be the scaling factor and hence the larger the contention window for the transmission of the packet. This would give the flows with a lower service a high probability to access the channel. Over a longer time span we would expect all the flows to achieve their fair share of throughput.

6. FBMA—mechanisms and algorithm

6.1. Mechanisms

We now describe the mechanisms that help FBMA achieve a per-flow fairness model. The mechanisms can essentially be grouped under the following four key components, namely (i) queuing operations, (ii) scaling of contention window, (iii) obtaining local service information, and (iv) propagation and update of state.

1. Queuing operations: Varying the contention window alone with respect to the service obtained will not be able to provide the flows with a fair share of the channel capacity. The operation of the queue at each node must also change. Specifically, every node must give priority to the packet belonging to a flow with a lower service when it is being dequeued. This assignment of priority can be incorporated in two ways: either during the enqueue process or during the dequeue process. In the enqueue process the packets could be placed such that the head of the queue always contains the packet with the lowest service and the packets in the queue are sorted in terms of the service obtained by their respective flows. During the dequeue process, the head of line packet is removed and serviced from the queue. On the other hand, all the packets could be enqueued without any intelligence like a drop tail queue. In this case, it is not necessary to sort the packets in the queue based on their flows’ service, but rather enqueue them at the tail as and when they arrive. However, the dequeue process has to be made more intelligent in order to dequeue the packet whose flow has obtained the lowest service thus far. While intelligence in the enqueue process (dequeue process unchanged) would require \(O(\log(n))\) operations to parse the queue and place the packet in the ascending order of service, intelligence in the dequeue process (enqueue process unchanged) would require \(O(n)\) operations to parse the entire queue to determine the packet belonging to the flow with the lowest service. Moreover, the routing layer packets are given the highest priority and are always enqueued at the head of the queue. Therefore, FBMA incorporates intelligence in the enqueue process, and hence the packets are inserted in the queue in the ascending order of their flows’ service. The dequeue process is left unchanged, where the packet at the head of the queue is always dequeued.

2. Scaling of contention window: Once the MAC layer has dequeued the packet belonging to the flow with the lowest service from its queue, it needs to scale its contention window before contending for channel access to transmit this packet. This scaling of contention window should depend on the service obtained by the flow to which the packet belongs. However, the problem with using the service directly as the scaling factor is the resulting inefficiency. As time progresses, the service of the flows would increase and hence the scaled contention window for the flows would keep increasing indefinitely. Thus it could cause the flows to wait longer unnecessarily before accessing the channel thereby leading to a potential under-utilization of the channel capacity. On the other hand, the service obtained for a flow at any instant can be normalized with the minimum service in the network. The normalized value can then be used as the scaling factor for the base contention window (\(cw\)). This would help reduce the inefficiency pointed out earlier. However, normalizing the service with the minimum service in the network also has two issues associated with it: (i) In a distributed network, it is not possible to easily obtain
the minimum service in the entire network, and (ii) normalizing by the minimum service in the entire network will reduce the degree of spatial reuse and could consequently result in under-utilization. Hence in FBMA, the service of any flow is normalized with the minimum service value obtained by the flows in its two hop neighborhood. Since only the flows in the two hop neighborhood are involved in channel contention, it is sufficient to consider the service counters of these flows alone in determining the minimum service. Furthermore, this requirement of local (two hop) information alone, also helps keep the scheme distributed. However, to be able to obtain the minimum service in the two hop neighborhood it is necessary for a node to determine the service counters of the flows that are serviced by it and also the service counters of the flows in its two hop neighborhood. The mechanisms that help a node obtain these information are outlined in the following.

3. Obtaining local service information: Every node needs to obtain the service information corresponding to the packet that its MAC layer has dequeued from the queue. More generally, the node needs to know the service counters of each of the flows serviced by it. The mechanism by which the service counter information of a flow serviced by the node is updated (recorded) is as follows: The destination node keeps track of the number of packets belonging to a flow that it has received thus far. When the destination node receives a packet belonging to a flow, it piggy-backs the service that has been obtained by that flow on its MAC layer ACK packet. The downstream nodes propagate this service information to the upstream nodes of the flow in their respective MAC layer ACK packets for that flow. In addition, they also update their service counter information for the flow corresponding to the packet. In the start-up phase of the network, the nodes do not have service information for the flow that they service, we approximate the number of packets that a node has forwarded for a particular flow as the service counter for that flow.

4. Propagation and update of state: We now explain the detailed mechanism by which a node obtains the service information of the flows in its two hop neighborhood. To be able to use the collected information effectively in determining the scaling factor, every node maintains five parameters locally, namely $min_{hop0}$, $min_{hop1}$, $owner_{hop1}$, $min_{hop2}$ and $owner_{hop2}$. The significance of these parameters is explained below.

Since the MAC layer at every node knows the service counter of the flow it serves, it propagates this information to help other nodes find the minimum service counter in their two hop neighborhood. It piggy-backs this value on all the control packets that it transmits as its local service minimum ($min_{hop0}$). Whenever the MAC at any node receives a control packet destined to it or by overhearing, it obtains the $min_{hop0}$ value that the originating node had stamped on the packet. This would be compared against the one hop service minimum ($min_{hop1}$) maintained at the node. If the new value is lower than the already stored ($min_{hop1}$) value then it updates ($min_{hop1}$) to this new value obtained from the control packet. It also stores the ID of the node ($owner_{hop1}$) from which this information was obtained. Every node stamps only two values on its control packets, namely its $min_{hop0}$ and $min_{hop1}$. The $min_{hop0}$ value that is obtained from control packets from neighboring nodes is used to update the $min_{hop1}$ value at the node. The two hop service minimum ($min_{hop2}$) value and the corresponding owner ID ($owner_{hop2}$) are calculated as follows: every node in addition to its $min_{hop0}$ value also stamps the $min_{hop1}$ value that it maintains, onto the control packets. Hence any node receiving a control packet, in addition to using the $min_{hop0}$ value on the packet to update its $min_{hop1}$, also uses the $min_{hop1}$ value on the packet to update its $min_{hop2}$. The corresponding owner IDs are also updated. Note that the nodes always update their $min_{hop1}$ and $min_{hop2}$ values if their maintained owner IDs are the same as the node from whom the new packet is received. On the other hand if the transmitting node of the packet is different from the owners of the $min_{hop1}$ and $min_{hop2}$
values maintained at the node, then the values and corresponding owner IDs are updated only if the stamped values are lesser than the maintained minimum values.

In the case that a flow leaves the network, to avoid storing stale service information, every node runs a timer corresponding to the owners of the $min_{hop1}$ and $min_{hop2}$ values that it maintains. If the node does not get an update from the respective owner within the expiry of the timer, the state stored corresponding to that owner is reset. The timers are set for a duration that is dependent on the data sending rate of the application. The local, one hop and two hop minimum service values and their corresponding owner IDs at any node are updated in this manner through advertisements in control packets. To make the state at every node more accurate and consistent, the nodes send out the next minimum service packet in their queues as their local minimum values. This would help the neighboring nodes obtain information about the service of the next packet that the initiating node is going to contend for, in the upcoming time slot. By this mechanism every node in the local neighborhood has knowledge about the service counter of the packet for which the nodes in the neighborhood will be contending for in any slot. This helps the nodes contend appropriately, thereby achieving a significant degree of per-flow fairness in the network. However, since the nodes do not have global information, the two hop neighborhood information alone will not help them achieve optimal fairness, as we discuss in Section 6.2.

6.2. Algorithm

The detailed algorithm for FBMA (Fig. 9) is presented in this section. Once the MAC dequeues a packet based on the packet with lowest service in the queue, it finds the ratio of the service of this packet (obtained from the downstream nodes of the flow relaying the service information from the destination through MAC layer ACKs) to the minimum service in its two hop neighborhood.

```
Variables used in the algorithm
min serv --> minimum service of flow in 2 hop neighborhood
scaling factor --> factor that scales the contention window
 cw --> contention window parameter
hop1 --> flow with minimum service traversing the node
owner hop1 --> id of node with minimum service flow in 1 hop neighborhood
hop2 --> flow with minimum service in the node's 2 hop neighborhood
owner hop2 --> id of node with minimum service flow in 2 hop neighborhood
hop0 stamp --> hop0 value stamped by the node on its control packets
hop1 stamp --> hop1 value stamped by the node on its control packets
servflow stamp --> service of flow stamped by the destination node
destserv stamp --> id of the destination that stamped the flow's service

Basic operation
1 packet from higher layer to ifq
2 intelligent sorting in ifq
3 MAC obtains packet from ifq after every successful transmission
4 min serv -- getmin(hop0,hop1,hop2)
5 scaling factor -- mypkt serv/min serv
6 cw -- scaling factor+cw

Outgoing control packets
7 hop0 stamp -- get neaptxt serv()
8 hop1 stamp -- hop1
9 hop0 -- mypkt serv

Incoming control packets
10 Hop1 updating
  if (src == owner hop1)
  hop1 --> hop0 stamp
11 elseif (hop1 < hop0 stamp)
12 hop1 --> hop0 stamp
13 owner hop1 --> src

Hop2 updating
14 if (src == owner hop2)
15 hop2 -- hop1 stamp
16 elseif (hop2 < hop1 stamp)
17 hop2 -- hop1 stamp
18 owner hop2 --> src

Outgoing ACK packets
20 if (node id == dest flow)
21 ++ serv flow
22 servflow stamp -- serv flow
23 destserv stamp -- dest serv

Incoming ACK packets
24 if destserv stamp belongs to the list of destination nodes
25 serv flow|destserv stamp -- servflow stamp
26 if (my id != src of the flow)
27 propagate it to the next node upstream
```

Fig. 9. Algorithm for the FBMA protocol.
The minimum service in the two hop neighborhood is obtained as the MIN\(\left(\min_{\text{hop}0}, \min_{\text{hop}1}, \min_{\text{hop}2}\right)\). This ratio forms the scaling factor for its contention window. Hence if the minimum service packet in the queue has a large service when compared to the service of its two hop neighbors then this node will have a large scaling factor of its contention window and hence will defer its access to the channel for a longer time. Thus nodes having a packet with a lesser service in the neighborhood will get to access the channel with a higher probability. A brief explanation of the pseudo-code for the algorithm is provided below.

The packet is passed on to the interface queue \((ifq)\) from the link layer (line 1). The queue then inserts the packet into the queue (line 2) such that the queue has all the packets sorted in the ascending order of their flows’ service. MAC layer dequeues the head of line packet from the queue whenever it has completed a transmission successfully. The head-of-line packet will always be the one with the minimum service in the queue (line 3). The minimum service at the node is obtained as the minimum of its local \((\text{hop}0)\), one \((\text{hop}1)\), and two hop \((\text{hop}2)\) service minimum values (line 4). The calculation of the scaling factor which is the ratio of the service (of the flow corresponding to the packet) to the minimum service calculated at this node for its two hop neighborhood is then performed (line 5). Using this scaling factor, the contention window is scaled (line 6). The local and one hop service information are stamped onto the control packets (lines 7–8). The local minimum represents the service of the packet that has been dequeued from the \(ifq\) which in turn is the minimum service packet at this node (line 9). This is followed by the update of one hop service minimum values and their owner IDs (lines 10–14). If the source of the packet \((\text{src})\) is the same as the owner of one hop minimum service value then the one hop minimum service value is updated, or else it is updated only if the new value is older than the already stored value. The same process is repeated for the update of the two hop minimum service values (lines 15–19). The service obtained by a flow is stamped by the destination node along with its ID onto the MAC layer ACK packet (lines 20–23). The nodes upstream of the destination, update their service counter values for each destination to whom they have been servicing packets on receiving such ACK packets (lines 24–25). Finally the nodes keep propagating such service information only as long as they are not the source of the flow (lines 26–27). Once this information reaches the source it is not propagated any further.

6.3. Practical considerations

FBMA retains the functionalities of CSMA/CA and supplements it with additional mechanisms to achieve a per-flow fairness model. This is illustrated in Fig. 10 where the FBMA module consists of the CSMA/CA (IEEE 802.11) module in conjunction with another module that lies between the CSMA/CA module and the physical layer. The new module, comprising of the fairness and the state maintenance modules, is responsible for the mechanisms that help ensure a per-flow fairness model. The interface queue is also modified to...
perform intelligent enqueuing such that the packets in the queue are always sorted in the increasing order of the service counters of the flows to whom the packets belong. The dequeue process is straight-forward and always dequeues the head of line packet that corresponds to the flow with the lowest service counter at the node. When CSMA/CA dequeues a packet, the packet header is passed to the state maintenance module. The state maintenance module maintains the service counters of the flows serviced by the node. The flow ID from the packet’s header is used to look-up the service obtained by the flow currently being serviced. In addition, the module also maintains the one and two hop minimum service values and also the owner IDs of the corresponding values. These values are obtained from the information piggy-backed onto the control packets by the neighboring nodes. The state maintenance module thus possesses all the information required to determine the scaling factor. This information is then used by the fairness module to compute the scaling factor for the base contention window. The fairness module feeds the computed scaling factor as input to the back-off mechanism employed by CSMA/CA, where CSMCA/CA contends for channel access with the newly scaled contention window parameter. Once the CSMA/CA protocol has gained access to the channel and is ready to transmit the packet, the packet is passed to the state maintenance module. Here the packet header is stamped with the local and one hop minimum service values and their respective owner IDs as maintained by the module, before passing the packet on to the physical layer. In the other direction, whenever the physical layer receives a packet, it is first passed onto the state maintenance module, from where it gets passed onto the CSMA/CA layer. The state maintenance module would update its one and two hop minimum service values and their respective owner ID’s based on the information piggy-backed on the (control) packet.

Thus, FBMA does not require any change in the functionalities of the CSMA/CA protocol. Instead, the desired objectives of FBMA are achieved by simply supplementing CSMA/CA with an extension module as explained above.

6.4. Properties of FBMA

In summary, the following are the key properties that help FBMA achieve a per-flow fairness model:

- It uses contention window as a tool to provide nodes with lower service, a higher probability (priority) to access the channel.
- It manipulates the queuing operations intelligently to aid in the prioritization process whilst also reducing the computational complexity.
- It obtains the information necessary to achieve prioritization (by scaling the contention window appropriately) in a purely distributed and localized manner by piggy-backing pertinent information onto the control messages. This in turn also helps keep the overhead minimal.

7. Performance evaluation of FBMA

The topology considered for performance evaluation is a 1500 m x 1500 m grid containing 100 nodes. We consider only static scenarios in our simulations. The simulations are all run for 100 s and each data point on the graph has been averaged over 10 random seeds for the above specified parameters. The 10 random seeds/scenarios have been generated using the setdest tool. UDP is used as the transport protocol with CBR as the traffic generator and there are 30 flows each having a load of 32 Kbps (512 byte packet at 8 packets per second) in all the scenarios unless otherwise specified. However, the load on the network will vary when we study the impact of load on the different MAC protocols. Further we use the shortest-path routing (SPR) for all the scenarios except the case where we study the impact of other routing protocols like widest-shortest path routing (WSR). In all the scenarios the performance of FBMA is compared with CSMA/CA.

We consider the following two metrics in our evaluation:

- **Throughput**: This represents the end-to-end data rate observed by a flow and is measured in bits per second. It is measured by the number
of data packets successfully delivered to the destination of a flow.

- **Relative unfairness index:** The common method to measure the degree of unfairness is to obtain the normalized standard deviation. But the normalized standard deviation would not capture the effectiveness of the fairness model. This is because a per-flow fairness model does not necessarily imply that all flows in the network obtain the same throughput. It is possible that some of the flows traverse through less congested regions and hence are capable of securing more throughput than the others. Since the model must also ensure efficiency (in terms of utilization), while the flows in a contention region will obtain their fair share of throughput, those in lesser congested regions will capitalize on their topological advantage to obtain more throughput. Hence what we are actually interested in, is the distribution of throughput amongst only the flows that are contending. In other words, the metric should ideally capture the deviation in the throughput distribution amongst flows in the different contention regions separately. However, the normalized standard deviation would include the contention-free (lesser contention) flows along with the contending flows in evaluating the fairness model. Though the throughput distribution amongst the flows may be fair based on the contention pattern in the network, it would not be highlighted by this metric. A reasonable way to capture this would be to compare the throughput distribution of the fairness model against that of IFS, which by virtue of being a centralized model provides optimal fairness for a given contention pattern. Hence we define another metric called the *relative unfairness index* as the deviation of the throughput distribution of FBMA (or CSMA/CA) from that of IFS. This metric is similar to the Kullback Leibler Fairness Index employed in [9].

In the rest of the section we evaluate the performance of *distributed IFS* over a variety of network parameters such as (i) impact of load, (ii) impact of node distribution, (iii) impact of traffic distribution, and (iv) impact of load-balanced routing protocols.

### 7.1. Impact of load

In order to study the impact of load we maintain the number of flows in the network to be a constant of 30, while we vary the load injected by each flow into the network, from 1 to 15 pkt/s which varies the total load from 120 Kbps to 1.8 Mbps.

Fig. 11(a) and (b) represent the relative unfairness index and throughput for the two protocols considered under different load conditions. The improvement in the degree of fairness is evident from Fig. 11(a) where FBMA shows a marked reduction in unfairness compared to CSMA/CA by over several folds. Further the unfairness index seems to be little disturbed by the increasing load unlike CSMA/CA whose degree of unfairness increases with the load. This is because relative unfairness index indirectly measures how far the fairness model deviates from the ideal per-flow fairness model (IFS). Since FBMA is an approximation of IFS, it must be able to closely track IFS and hence its deviation from IFS must remain almost a constant. On the other hand, since CSMA/CA follows a per-node fairness paradigm, its deviation from the per-flow model will tend to increase as the load on the network increases. This can be attributed to the fact that as the load on the network increases, the fraction of nodes servicing more flows than the others tends to increase. Hence a per-node fairness model in such cases will only contribute to more deviation from the IFS. It can be seen that FBMA achieves a throughput similar to that of CSMA/CA as the load on the network is increased in Fig. 11(b). It must be noted that any significant improvement in throughput cannot be expected by FBMA owing to its objective of fairness. FBMA tries to achieve per-flow fairness by increasing the probability of channel access for the flows that have obtained lesser service at the cost of flows that have obtained higher service. The flows with higher service give up their slot to access the channel with a probability proportional to their service obtained. Hence the goal
of FBMA is to achieve per-flow fairness without suffering from any degradation in throughput.

7.2. Impact of node distribution

We study the impact of node distribution on the protocols considered. We vary the density of node distribution from a dense to a more sparse network. The way we do it is as follows: Out of the 100 nodes, 50 of them are randomly distributed over the grid of 1500 m × 1500 m. The entire grid is divided into 16 partitions of equal areas. We start by placing the remaining 50 nodes in only one of the 16 partitions. This represents the most dense scenario. We then decrease the density by increasing the number of partitions over which the remaining 50 nodes are placed from 1 to 16. The case of 16 would reduce to the case of a randomly distributed scenario. The two extremes (50% of nodes in one grid and random distribution) are illustrated in Fig. 12(a) and (b) respectively.

If the network is dense, then the fraction of nodes that serve as intermediate routers for several flows will tend to be more. Hence if the per-node fairness model is employed then as the density of the network decreases the degree of unfairness should also decrease. This can be seen from the CSMA/CA curve in Fig. 13(a). FBMA closely tracks IFS in its operation and hence has a fairness index that is almost constant unlike CSMA/CA. Further it can be seen that FBMA is able to reduce the degree of unfairness by a factor of about 5. It can also be seen from Fig. 13(b) that FBMA closely tracks CSMA/CA in its throughput distribution without suffering from any degradation.

7.3. Impact of traffic distribution

We now study the impact of clustered traffic on the performance of the MAC protocols. This kind of a scenario is common when several nodes try to access a backbone server through a base station in which case the traffic arriving at the base station tends to be clustered. We vary the degree of clustering by varying the number of destinations keeping the total number of flows (30 in number) and load on the network a constant. Hence when only one destination is employed, all the 30 sources will try to reach the same destination. On the other hand, when the number of destinations is 30, the scenario reduces to that of a completely distributed scenario with unique source–destination pairs.

Fig. 14(a) and (b) represent the fairness and throughput results for the two protocols. When the number of destination is 1, we have a single hot spot. As the number of destinations increases we increase the number of hot spots. However, since
the total load on the network is the same the load directed to a single hot spot decreases and the traffic tends to become more distributed in the network. This in turn increases the fraction of nodes that serve more flows than the others. Hence as we keep distributing the traffic till we obtain a randomly distributed traffic pattern, (destinations = 30), the degree of unfairness should tend to increase. This can be observed in Fig. 14(a) where CSMA/CA shows an increasing trend in unfairness as the number of destinations is increased. However, FBMA tends to remain stable with little variation, showing an improvement of over a factor of 5 when the relative unfairness index is measured. As in the other cases, the throughput of FBMA closely tracks that of CSMA/CA without experiencing a significant degradation in throughput.

7.4. Impact of load-balanced routing protocols

In all the above scenarios we have considered the shortest path routing (SPR) protocol. Now we consider the widest-shortest path routing (WSR)
protocol to study its impact on the MAC protocols. We consider a load of 32 Kbps per flow with 30 flows in the network. The results are all averaged over 10 seeds.

From Fig. 15(a) it can be seen that FBMA reduces the relative unfairness index from as high as 2.56 in CSMA/CA to 0.56 or in other words, increases the degree of fairness by about 80%. Also a marginal improvement in throughput can be seen from Fig. 15(b). Fig. 15(a) and (b) also present results for the distributed approach using the load-balanced routing protocol WSR. No significant performance improvement can be seen in the case of CSMA/CA. In fact the relative unfairness index degrades for CSMA/CA when WSR is employed. On the other hand, WSR aids FBMA by reducing the degree of unfairness by about 50% in the case of relative unfairness index. Also there is an improvement in throughput by about 16%. This can be attributed to the reason that WSR being a load-balanced routing protocol, the contention amongst the flows is reduced, which in turn helps the mechanism (of reducing contention and increasing probability of channel access for low service flows) in the FBMA protocol. This result re-emphasizes our thesis in Section 4 that significant performance gains can be obtained when a smarter routing protocol is used with a MAC protocol that is based on the per-flow fairness paradigm.
8. Issues and conclusion

8.1. Issues

We present some of the issues pertaining to the distributed implementation of the per-flow fairness model in this subsection. (i) The dynamics in the network with flows arriving and leaving at different time instants can be handled by considering throughput instead of the number of packets reaching the destination as the service indicator. Throughput would incorporate an additional dimension of time into the fairness model thereby taking care of the dynamics in the network. (ii) The update of state in the local neighborhood when flows leave the network can be achieved by employing timers that help nodes keep track of state in their local neighborhood. This in turn helps nodes recover from stale information whenever a flow leaves the network. (iii) Finally, the requirement of different rates (QoS parameters) for different flows can be incorporated into the existing fairness model by modifying the function that scales the contention window appropriately. A possible solution would be to include the priority (measure of the QoS class) of the flows in the function that determines the scaling factor.

8.2. Conclusion

The CSMA/CA MAC protocol was designed for wireless local area networks to provide fair and efficient medium access control to stations sharing a wireless channel. It has been adopted as the de facto standard for medium access control in multi-hop wireless networks also. In this paper, we argue that certain improvements in terms of reducing its band of contention and adopting a per-flow fairness model are necessary in order to realize efficient and fair medium access. We demonstrate the performance gains that can be achieved both directly and indirectly by reducing the band of contention and by employing the per-flow fairness model. We also present a MAC scheme that aims to realize the ideal flow scheduling protocol in a distributed fashion. We show through simulation results that the proposed distributed version improves the fairness of CSMA/CA MAC protocol to a significant extent. A critical issue not considered in this paper is mobility. While the contribution of the paper remains and can be used in static multi-hop wireless networks (e.g. sensor networks and wireless broad-band access networks), ongoing work is investigating the impact of a mobile multi-hop wireless network.

References


Karthikeyan Sundaresan received his Bachelor’s degree in Electronics and Communication Engineering from Anna University in 2001. He then joined the Electrical and Computer Engineering discipline at Georgia Institute of Technology in 2001 where he received his Masters degree in 2003. He is currently working towards his Ph.D at Georgia Tech. His research interests are in the field of wireless network protocols, mostly at the medium access, routing and transport layers of the protocol stack and also in the area of wireless multi-carrier communications.

Hung-Yun Hsieh received the B.S. and M.S. degrees in Electrical Engineering from National Taiwan University, Taiwan, ROC. He is currently a Ph.D. candidate in the School of Electrical and Computer Engineering at Georgia Institute of Technology. His research interests include wireless systems, mobile computing, and network protocols.

Raghupathy Sivakumar received his Masters and Doctoral degrees in Computer Science from the University of Illinois at Urbana-Champaign in 1998 and 2000 respectively. He joined the School of Electrical and Computer Engineering at Georgia Institute of Technology as an Assistant Professor in August 2000. His research interests are in wireless network protocols, mobile computing, and network quality of service.