

Exploiting Medium Access Diversity in Rate Adaptive Wireless LANs *

Zhengrong Ji, Yi Yang, Junlan Zhou, Mineo Takai, and Rajive Bagrodia

Computer Science Department
University of California Los Angeles
Los Angeles, CA 90095

{jizr, yangyi, zjl, mineo, rajive}@cs.ucla.edu

ABSTRACT

Recent years have seen the growing popularity of multi-rate wireless network devices (e.g., 802.11a cards) that can exploit variations in channel conditions and improve overall network throughput. Concurrently, rate adaptation schemes have been developed that selectively increase data transmissions on a link when it offers good channel quality. In this paper, we propose a *Medium Access Diversity* (MAD) scheme that leverages the benefits of rate adaptation schemes by aggressively exploiting multiuser diversity. The basic mechanism of MAD is to obtain instantaneous channel condition information from multiple receivers and selectively transmit data to a receiver that improves the overall throughput of the network, while maintaining temporal fairness among multiple data flows. We identify and address the challenges in the design and implementation of MAD's three phases: channel probing, data transmission, and receiver scheduling. We also use analytical models to examine the tradeoff between network performance improvement and overhead of channel probing, and derive an asymptotic performance bound for the receiver scheduling algorithms used by MAD. Results from the analysis and the extensive simulations demonstrate that, on average, MAD can improve the overall throughput of IEEE 802.11 wireless LANs by 50% as compared with the best existing rate adaptation scheme.

Categories and Subject Descriptors

C.2.1 [Computer-Communication Networks]: Network Architecture and Design—*Wireless Communication*; C.2.2 [Computer-Communication Networks]: Network Protocols

*This research is funded in part by the National Science Foundation under NRT grant ANI-0335302 *WHYNET: Scalable Testbed for Next Generation Mobile Wireless Networking Technologies*, and the DARPA NMS program under contract N66001-00-1-8937 *MAYA: Next Generation Performance Prediction Tools for Global Networks*.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MobiCom'04, Sept. 26-Oct. 1, 2004, Philadelphia, Pennsylvania, USA.
Copyright 2004 ACM 1-58113-868-7/04/0009 ...\$5.00.

General Terms

Design, Performance

Keywords

Medium Access, Multiuser Diversity, Scheduling, Wireless LAN

1. INTRODUCTION

The availability and low cost of 802.11 wireless networking products have encouraged a rapid growth in the deployment of wireless Local Area Networks (LANs). They can extend access to wired intranets such as campus networks, as well as support broadband access to the Internet — particularly at “hot spots”. As the use of powerful portable computing devices (laptops, PDAs) with wireless LAN connectivity becomes common, the ability of wireless LANs to cater to large number of users having applications with high bandwidth requirements is increasingly important. The increasing appetite for bandwidth in wireless LANs has in turn spurred extensive research efforts to provide high data rates at the physical (PHY) layer. However, the highest feasible rate is ultimately bounded by the channel signal to noise ratio, which is time-varying due to both slow and fast fading in mobile environments. This has encouraged the development of multi-rate adaptors, where the PHY layer data rate can effectively respond to wide variations in channel conditions; the difference between the lowest and highest data rates is expected to widen even further with emerging PHY layer technologies. This is evident by comparison of IEEE 802.11a [12] with legacy 802.11 standard [11]: the lowest data rate increased by a factor of 6 while the highest rate increased by a factor of 27. The number of distinct data rates supported by PHY is increasing as well and the trend in next generation PHY technologies such as MIMO and Adaptive-Bit-Loading OFDM [8] is to provide wider, almost continuous range of data rates that can be tailored to the given channel quality.

The multi-rate support at the PHY layer offers primitives for upper layers to respond in an agile manner to fluctuations in channel conditions and optimize channel utilization. Recent research efforts on rate adaptation schemes at the Medium Access Control (MAC) layer have focused on dynamic rate switching mechanisms by either sender's inference [15] or receiver's feedback [9] of the current channel conditions. Sadeghi et al. [23] shows that the success of rate adaptation depends not only on the ability to find

the highest feasible data rate for the given channel, but also on the ability to optimize the link utilization under time-varying channel conditions. Thus, the existing rate-adaptation schemes merely respond to the quality of a given channel by selectively increasing the transmission rates when the channel condition is favorable.

In contrast, this paper proposes a scheme called Medium Access Diversity (MAD), to *actively* exploit time and space varying channels at the MAC layer. In a wireless LAN environment, the instantaneous channel conditions from a sender to multiple receivers are time varying and are not correlated due to mobility, interference and other factors. This effect is referred to as *multiuser diversity* in the context of cellular networks [16] but is observable in mobile wireless environments in general. With multiuser diversity, a node can aggressively assess the instantaneous channel condition to each of the multiple receivers and select the receiver whose channel condition is near its peak. By thus leveraging the space-time heterogeneity of channels, this scheme can further exploit the benefits of any rate-adaptation scheme. To the best of our knowledge, this is the first attempt to explore the benefits of as well as identify and address the issues concerning the multiuser diversity in the wireless LAN environment whose PHY and MAC characteristics are substantially different from cellular networks. Further, we show that MAD can be implemented with relatively minor enhancements to the IEEE 802.11 standard.

The paper addresses three primary design issues with MAD that have a significant impact on the overall wireless LAN performance with contention based MAC protocols: channel probing, link goodput optimization, and receiver scheduling. First, we analyze the tradeoff of throughput gain and channel probing overheads in MAD to deduce optimal channel probing parameters that account for space and time varying channel conditions in mobile environments. Second, we propose Packet Concatenation (PAC) as a new rate adaptation scheme that efficiently exploits given channel quality to maximize individual link goodputs. Compared to existing rate adaptation techniques, PAC itself achieves up to 50% more goodput with IEEE 802.11a PHY, under favorable channel conditions, with relatively minor performance degradation under poor channel conditions. Third, the study derives an asymptotic performance bound of MAD by two scheduling algorithms that maintain temporal fair share among multiple data flows. Finally, we present the results from a comprehensive simulation study on MAD with existing rate adaptation schemes as well as PAC. The simulation results in various wireless LAN scenarios show that MAD achieves 120%~452% throughput improvement over ARF, the rate adaptation scheme used in Lucent's WaveLAN II networking devices, and on average more than 50% throughput improvement over the best known rate adaptation scheme, OAR.

The rest of the paper is organized as follows. The next section summarizes related work. Section 3 presents an overview of MAD, followed by three sections each of which respectively addresses the channel probing, data transmission, and receiver scheduling issues associated with MAD. Section 7 uses analytical models to examine the tradeoff between network performance improvement and overhead of channel probing, and derives an expected network throughput achievable by MAD. Section 8 presents extensive simulation results to evaluate the effects of MAD on wireless LAN throughput. Section 9 concludes the paper.

2. RELATED WORK

The multiuser diversity has its roots in the seminal work of Knopp and Humblet [16]. It underlies much of the recent work [4, 14, 19, 27] on the scheduler design for time-division downlink optimization in Code Division Multiple Access (CDMA) cellular networks, such as 1xEV-DO High Data Rate (HDR) [2] and High Speed Downlink Packet Access (HSDPA) [10] systems. The aforementioned work has found that as a consequence of the multiuser diversity effect, the more users to choose from, the larger the gain from the scheduler. In that context, a large channel variation is not a drawback, but is rather preferred [27]. In such systems, the channel conditions are fed back from all receivers via the CDMA uplink. One challenge to our work is that such a closed-loop feedback mechanism is not readily available in existing 802.11 wireless LANs. MAD differs from the above work in that it uses explicit multiuser channel probing to obtain the present channel condition information. The increasing overhead of channel probing may prohibit MAD from querying all receivers for each data transmission, hence posing a challenge to the scheduler design.

A channel-aware ALOHA protocol is proposed in [21] to exploit multiuser diversity gains. In this work, each user bases their transmission probability on their channel gain, assuming each user knows their own channel gain as well as the distribution of other users' channel gains. In contrast to previously mentioned work, it assumes that closed-loop channel condition feedback (in the context of ALOHA like protocol) is not available, and sender has no centralized control. Our approach differs from the above two as follows. The MAD mechanism is built on top of decentralized 802.11 MAC for channel acquisition. However, MAD gives each sender centralized control over its own multiuser diversity but with only partial feedback from its receivers.

A number of rate adaptation schemes have been proposed in wireless LANs. The Auto Rate Fallback (ARF) [15] was the first implementation of dynamic rate switching, in which the sender gradually increases data rates after consecutive successful transmissions and reverts to lower rates after a failure. The Receiver Based Auto Rate (RBAR) [9] improves throughput by letting the receiver measure the channel quality based on request-to-send (RTS) reception and feed back the transmission rate to the sender. Based on the observation that the coherence intervals are on the order of multiple packet transmission times, the Opportunistic Auto Rate (OAR) [23] transmits multiple packets (by treating them as fragments) when channel condition permits higher data rates, thus achieves the highest throughput among the three. The efficiency of MAD is highly dependent on the underlying link goodput optimization techniques. OAR is one such technique that can be employed by MAD directly. We devote section 5 to further discuss this issue, and comparative studies on MAD using OAR and our alternative rate adaptation scheme PAC is performed via analysis in section 7 and simulation in section 8.

3. MEDIUM ACCESS DIVERSITY

Consider an ideal wireless network where the instantaneous channel conditions of all receivers¹ are known by the

¹If the context is not clear, a sender (receiver) refers to the sender (receiver) of a data packet, not a control packet in this paper.

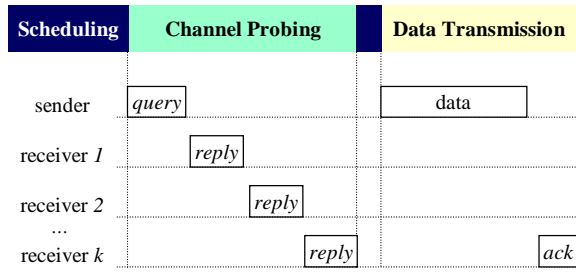


Figure 1: Medium Access Diversity: sender decides to send data to receiver k after channel probing with query/reply exchange.

sender. It is possible for the sender to exploit multiuser diversity [16] by adopting scheduling algorithms that serve a flow when its channel condition is near its peak within a short latency. Such intelligence provides at least two potential benefits. First, the head-of-line (HOL) blocking problem [5] can be alleviated. Instead of continuing retransmission to a receiver when it experiences a poor channel condition with a burst of packet retransmission failures, the sender could try scheduling transmission of data packets to another receiver. Second, with complete knowledge of channel conditions at all receivers, a sender can not only select a receiver with the best channel condition for data transmission, but also choose the highest feasible rate. This approach could improve the overall network throughput, and such improvement peaks when the channel utilization is optimized, particularly with the presence of favorable channel conditions. However, existing contention based MAC protocols such as IEEE 802.11 DCF cannot be applied with this concept without modification, as the time taken to probe the channel conditions of all receivers is unpredictable, and the information will be obsolete, particularly when expected coherence time in a wireless network is short.

To realize the above idea in a practical manner for existing IEEE 802.11 wireless LANs, we propose a medium access mechanism to exploit multiuser diversity, henceforth referred to as Medium Access Diversity (MAD). MAD consists of three phases: channel probing, data transmission and scheduling, where scheduling decisions are required before each of the former two phases. The basic mechanism of MAD works as follows (see Figure 1). The sender broadcasts a *query* message including the list of potential receivers for channel condition probing. Following the order indicated in the list, each receiver sends its own *reply* message back to the sender in the corresponding time slot. The *reply* message contains channel condition information obtained from evaluating reception of the preceding *query* message. After collecting the information from all *reply* messages, the sender chooses the receiver that maximizes potential throughput gain and begins its data transmission. The chosen receiver will reply with an *ack* message if it receives the data packet(s) correctly. This process repeats after the current data transmission is completed.

Although conceptually straight-forward, the MAD mechanism will be effective only if the following requirements are met in its design and implementation:

- MAD must limit the overhead incurred by channel probing, implying that only a subset of receivers can

be queried for their channel conditions. MAD must be able to improve overall network performance even with such limited information.

- MAD has to maximize the effective throughput for each data transmission, especially when the current channel condition at a receiver supports high data rates in order to improve the throughput gains. This could also compensate for the overhead of channel probing.
- The MAD scheduling algorithm must balance the need to improve network throughput with fairness expectations among multiple traffic flows. The two objectives often conflict with each other.

The next three sections elaborate the details of MAD for IEEE 802.11 wireless LANs with the above requirements. We start with a generic realization of the channel probing and data transmission phases of MAD for IEEE 802.11 DCF, assuming readers are familiar with the details of IEEE 802.11 wireless LANs [12, 11]. We then present a theoretical scheduling model with two approximation algorithms for MAD. In section 7, an analysis is carried out to shed insights about optimal MAD design choices and asymptotic performance gain by MAD.

4. CHANNEL PROBING

The use of request-to-send (RTS) and clear-to-send (CTS) control packets as the handshake before actual transmission of data packet is a well-known distributed channel acquisition mechanism employed by the IEEE 802.11 wireless networks. MAD enhances this handshake mechanism to facilitate channel probing as follows.

To allow a sender to query multiple receivers for channel condition feedbacks, we introduce an enhanced RTS control packet, referred to as the Group RTS (GRTS). The formats of regular RTS and our GRTS are shown in Figure 2(a) and 2(b) respectively. The only difference between GRTS and regular RTS is that a GRTS control packet contains a list of addresses (RAs) of receivers polled by the GRTS sender. The rank of an RA in the RA list indicates the order by which the receiver should respond to the GRTS. When the number of RAs in the list is one, GRTS is identical to the original RTS.

We also modify the CTS control packet in the standard by including an additional field (two octets) named *Feedback*. This field consists of two subfields, *Rate* and *Gain* as shown in Figure 2(c). The *Rate* (4 bits) subfield stores the encoded data rate selected by a receiver. In MAD, a sender strictly follows a receiver's choice of data rate if the sender decides to transmit data to that receiver in the subsequent data transmission phase. The receiver uses a simple rate selection algorithm to suggest a *Rate* to the sender. The algorithm will select the modulation scheme that has the highest data rate among those with bit error rates less than 10^{-5} for the measured SNR of previous (G)RTS packet. A similar approach was used in [9].

The *Gain* (12 bits) subfield of *Feedback* in our CTS control packet stores the relative gain of a receiver's current channel condition over its average. In practice, a receiver can use the measured signal-to-noise ratio (SNR) of received GRTS packets to represent the estimated channel condition. Thus, the relative gain is the relative difference between instantaneous SNR and its average. In our current design, the

Gain subfield supports 2^{12} levels of relative gain. The utility of *Gain* will become clear after we discuss MAD scheduling algorithms in section 6.

To compute the relative gain at the receiver, average SNR values need to be estimated for every sender-receiver pair. For any receiver, the average SNR value regarding sender i is estimated via the following equation

$$\overline{SNR}_i = (1 - \alpha) \times \overline{SNR}_i + \alpha \times SNR_i$$

whenever a new SNR value is measured from the most recent GRTS packet sent from node i . α is a coefficient between 0 and 1, and we have fixed the value of α as 0.2 in the scope of this paper.

It is important to note that the task of monitoring channel conditions is assigned to the receiver. This choice was made because the receiver can make better estimations: First, the suggested rate is based on the reception of the most recent GRTS packets from sender; Second, the average SNR value can be updated as long as a GRTS is overheard by the receiver, even if that receiver is not in the polling list.

Given the above enhancements, the channel probing in IEEE 802.11 DCF works as follows. A sender transmits a GRTS packet including the RA list of receivers. Each receiver will transmit a CTS packet back to the sender at the assigned time slots after receiving GRTS. The time-slot duration equals the time to transmit the CTS control frame in base rate plus the Short Interframe Space (SIFS). In particular, the receiver waits for SIFS and then transmit its CTS packet in its designated time slot. Channel probing will be completed at the end of the time slot for the last receiver in GRTS's RA list. The sender will then pick a receiver based on its scheduling decision and move on to data transmission.

A virtual carrier-sensing mechanism referred to as the network allocation vector (NAV) is used by 802.11 MAC. The NAV maintains a prediction of future traffic on the medium based on duration information that is announced in RTS/CTS frames prior to the actual exchange of data [11]. However, with data rate negotiation between sender and receiver in MAD, it is impossible for sender to predict the exact remaining duration of channel occupation when it transmits (G)RTS, since it does not know the subsequent data rate yet. To resolve this issue, the sender would use existing packet output queue information to make a conservative estimation of channel occupation time and set that in (G)RTS's *Duration* field. This approach serves as a tentative channel reservation. Other nodes have to set their NAVs in accordance with the tentative reservation once they have overheard the (G)RTS. To facilitate later correction to the tentative reservation, the NAV of each node comprises tuples of form $(src, duration)$. The *src* corresponds to the MAC address of the node that initiates the data exchange sequence², and *duration* is updated from the *Duration* field of the most recently overheard packets containing same *src*. The current channel reservation time is the maximum *duration* of all tuples. To maintain consistency of channel reservation by both sender and receiver, *Duration* of CTS will be set according to sender's (G)RTS. After (G)RTS/CTS exchange, the sender can revise its reservation in the data packet. At the receiver side, ACK will also carry revised channel reservation. If multiple receivers are queried in channel probing,

²The address of *src* may be specified at different fields in different control packets. For example, it's in TA of RTS and RA of CTS.

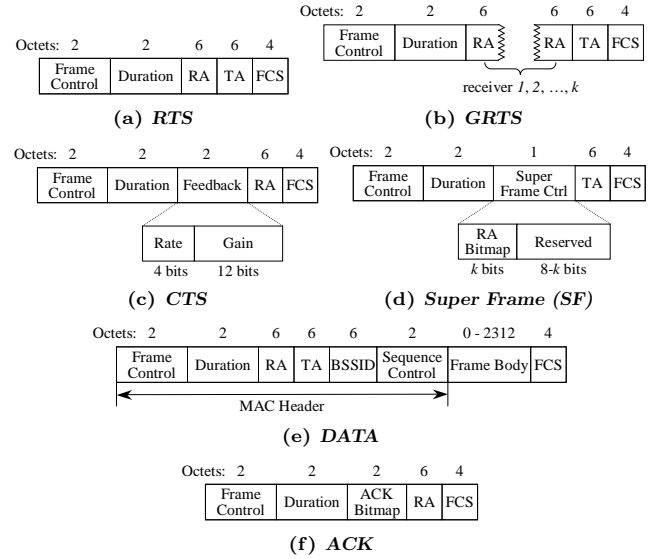


Figure 2: MAC frame formats used by MAD

any receiver not selected in data transmission will not be able to change its CTS reservation. This operation has a subtle impact on spatial reuse in wireless LANs, which we will study in the context of other ad hoc wireless networks in the future.

Note that there is an open question regarding the optimal number of receivers which a sender should query in order to achieve peak performance gains. We will discuss this issue in section 7. In next section, we will present realization of data transmission in MAD.

5. DATA TRANSMISSION

5.1 Overview

One of the design requirements for MAD is that it has to maximize the effective link throughput for wireless networks. A simple solution to meet this requirement is that during the data transmission phase of MAD, more data packets should be transmitted particularly when the current channel condition supports higher data rate. As a result, the expected overhead per packet can be reduced. Opportunistic Auto Rate (OAR) protocol proposed by Sadeghi et al. [23] is one of the recent efforts to better exploit durations of high quality channel conditions by such means. The key mechanism of the OAR protocol is to opportunistically send multiple back-to-back data packets to the same receiver whenever the channel quality is good. As channel coherence time typically exceeds the duration of multiple packet transmission in wireless LAN environment, OAR achieves significant throughput gains.

As a rate adaptation scheme, OAR can be directly employed by MAD during data transmission phase. Figure 3 gives an illustration to how OAR works within the MAD framework. After sensing the channel to be idle for DCF Inter-Frame Space (DIFS), the sender initiates channel probing by sending a GRTS. In this particular example, the sender decides to transmit packets to receiver 2 after receiving CTS feedbacks from all receivers because a higher data rate can be supported. The sender utilizes the IEEE

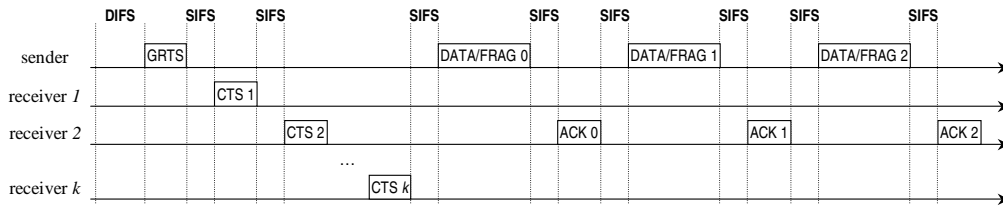


Figure 3: Illustration of Medium Access Diversity using OAR.

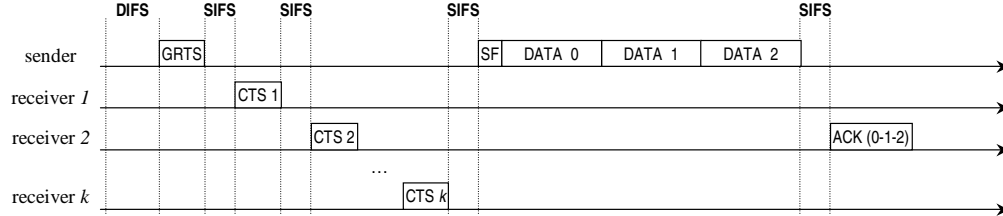


Figure 4: Illustration of Medium Access Diversity using PAC.

802.11 fragmentation support by treating multiple back-to-back packets as fragments of a “super-packet.” After receiving the ACK from receiver 2 for the first fragment, the sender will transmit the next fragment. This DATA and ACK packet exchange is repeated until a packet is lost or the sender uses up its time share. Note that in OAR, the maximum number of packets transmitted back-to-back cannot exceed the ratio between current feasible data rate and the base rate in order to maintain the temporal fairness property under single rate IEEE 802.11 [23].

However, OAR may not fully optimize the channel utilization even when the channel condition at a receiver supports the highest data rate. This sub-optimality is due to the fact that the MAC and physical layer overheads become progressively more dominant as the highest data rate increases in wireless LANs. We use IEEE 802.11a as an example to illustrate this point. In 802.11 wireless LANs, control packets like RTS, CTS and ACK must be transmitted at one of the base rates (typically much lower than the highest data rate) supported by all the communication devices in the network for three reasons: First, the standard has to support inter-operability among different wireless LAN card manufacturers, where many higher data rates are not mandatory [12]; Second, the communication range at a higher data rate is much smaller than that at the base data rate³, and thus control packets must be transmitted at a lower data rate to alleviate the hidden terminal problem [25] in wireless networks; Last, SIFS is a fixed duration (subject to how fast a transceiver can switch between receiving and transmitting modes) regardless of data rate. Consider that control packets are transmitted at 6Mbps. To transmit a single 1KB packet, the control overhead is 16.5% at 6Mbps and increases to 64.0% at 54Mbps for legacy 802.11. OAR reduces the overhead to 49.37% per packet in later case. However, there is still room for improvement, which motivates us to propose the concatenation of multiple packets in physical transmission.

³For example, the D-Link DWL-A520 5GHz (802.11a) adaptor has a range of 300 feet at 6Mbps and 40 feet at 54Mbps (www.dlink.com).

5.2 Packet Concatenation (PAC)

The key mechanism of PAC is to transmit a sequence of physical data frames to the same receiver back-to-back after MAD has chosen that receiver for data transmission. Figure 4 demonstrates how PAC operates under the MAD framework. Once a receiver is selected after channel probing, PAC will transmit a *Super Frame* (SF) control frame (see Figure 2(d)), immediately followed by a chain of data packet frames. After the reception of the string of frames, the receiver will wait for SIFS duration and reply with an ACK to selectively acknowledge the received data packets. In contrast to OAR (see Figure 3), PAC can eliminate many ACKs and SIFS, which potentially will improve channel utilization for MAD. Note that the maximum number of packets in a single transmission cannot exceed the ratio between the current feasible data rate and the base rate, so *PAC has the same temporal fairness characteristics as OAR (and single rate 802.11 MAC)*. Similar to OAR, PAC also takes advantage of the fact that channel coherence time typically exceeds multiple packet transmission times. Details on coherence time estimation for modern digital communications can be found in [22]. Next, we describe how PAC is incorporated with the IEEE 802.11 MAC and PHY layers.

5.2.1 Incorporation with IEEE 802.11 MAC

As mentioned previously, a SF control frame is introduced in the MAC to precede the concatenated data packet frames, declaring the beginning of super-packet. The format of SF is shown in Figure 2(d). The *Frame Control* field in SF is the same as those in other packet frames [11]. The *Duration* field of SF specifies the estimated duration before the current data transmission phase will end, including the time for the transmission of the super-packet plus a SIFS and ACK duration (see Figure 4). It can be used by other nodes to update their NAVs for virtual carrier sensing. The *Super Frame Control* field contains a k -bit RA bitmap where each bit corresponds to the receivers polled by the previous GRTS packet. A nonzero bit notifies the corresponding receiver that it is an intended recipient of the subsequent data packets after SF. The same node is of course expected to reply with an ACK after reception of the data packets.

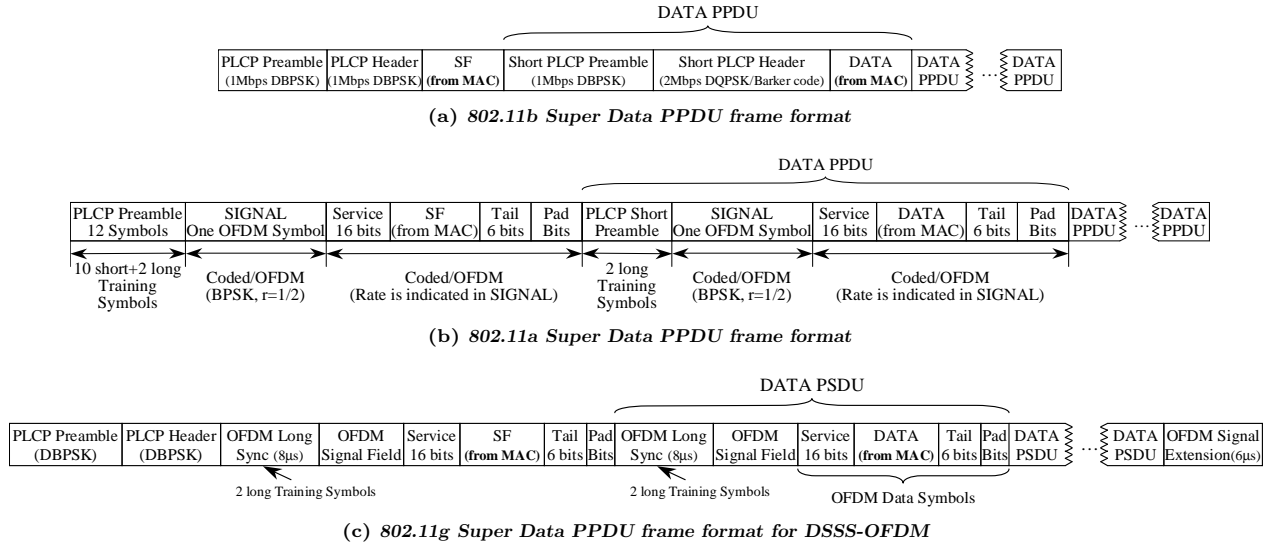


Figure 5: Super Data PPDU frame formats for PAC

The ACK control frame needs to be modified to provide feedback for the exact packets received in the super-packet. As shown in Figure 2(f), The *ACK Bitmap* field (2 octets) is inserted into the original ACK control frame. Each bit of the *ACK Bitmap* field corresponds to the ordered data packet received in the previous super-packet.

We also make a minor modification to the *DATA* packet frame (see Figure 2(e)). The *Duration* field of a *DATA* packet frame is originally required for NAV updates. As the same information has been provided by the SF preceding the data packets of a super-packet. This field is reused as a counter for the super-packet sender to identify the order of a data packet in the super-packet. The modifications to *DATA* and *ACK* allow the receiver to selectively acknowledge correctly received data packets.

5.2.2 Incorporation with IEEE 802.11 PHY

Some modification to the IEEE 802.11 PHY layer is also necessary to support the notion of the super-packet in PAC. The 802.11 based wireless networks currently have specifications for PHY at 2.4 GHz (802.11b) or 5 GHz band (802.11a), and a higher data rate extension at 2.4 GHz band (802.11g). Here we briefly describe how *Super Data* PLCP protocol data unit (PPDU) is designed for each of these PHY specifications. Due to space constraints, we only show the *Super Data* PPDU frame formats. The original formats can be found in [11] and its amendments [12].

IEEE 802.11b PHY has long and short PLCP PPDU formats for 802.11b PHY. In the long format, the *Sync* field contains 128 bits of scrambled 1's. Both the PLCP preamble and PLCP header are transmitted at 1Mbps with DBPSK modulation. In the short format, *Sync* contains 56 bits of scrambled 0's. The preamble is transmitted at 1Mbps while the header is transmitted at 2Mbps with Barker code and DQPSK modulation. We introduce the *Super Data* PPDU format (See Figure 5(a)) which begins with a long PPDU including SF (see Figure 2(d)) as its PLCP service data unit (PSDU). It is then followed by multiple short PPDUs each containing a data packet in the PSDU from the MAC layer. The SF is a special control packet transmitted at the base

rate independent of the rate used for subsequent data packet transmissions.

Similarly, *Super Data* PPDU frame formats are also introduced to the IEEE 802.11a/g PHYs, which use Orthogonal Frequency Division Multiplexing (OFDM) [8]. The major difference between the *Super Data* PPDU of 802.11g (see Figure 5(c)) and that of 802.11a (see Figure 5(b)) is that an 802.11b PLCP preamble and header are added at the beginning of *Super Data* PPDU in 802.11g for operational compatibility with 802.11b.

Note that although the *Super Data* PPDU looks like a very long frame, it is in fact a concatenation of multiple self-contained normal-sized physical frames. Channel estimation and fine frequency acquisition can still be performed at the beginning of each smaller frame in the *Super Data* PPDU during reception. Thus, when the duration of the super-packet is within the limits of channel coherence time, the reception should be as good as that when packets are transmitted individually. Please note that the MAC layer often sets the maximum packet size to limit the transmission duration such that the channel estimation by the PHY layer of the receiver remains accurate until the end of packet reception. The PAC concatenation does not contradict this MAC guideline because channel retraining for each concatenated packet is facilitated in the *Super Data* PPDU design.

5.2.3 Other PAC Issues

There are also some minor issues regarding maintenance of packet retransmission and contention window. In 802.11 MAC, a sender maintains a separate retry counter for each receiver, and the head-of-line packet for a receiver will be dropped if its retry counter reaches the limit. In PAC, we extend it by letting the sender maintain a retry counter for each retransmitted packet. It can be seen that the number of retry counters for each receiver will not exceed the maximum number of data packets in a single transmission. Another issue is that in legacy 802.11, the contention window is enlarged whenever a packet is retransmitted. To preserve such behavior in PAC, we use a simple heuristic in which the contention window is enlarged whenever more than half

of the packets in a super-packet need retransmission and is reset if more than half of the packets are acknowledged. The reason is that if more than half of the packets are acknowledged, the loss is most likely not due to contention. The reverse case implies occurrence of either collision or strong interference.

6. MAD SCHEDULING

The objective of the MAD scheduler is to improve the utilization of the channel while maintaining temporal fairness among multiple back-logged flows. In this section, we first discuss a theoretical scheduling policy in Rayleigh fading channels, then present the scheduling algorithms approximating the ideal policy behavior in MAD.

6.1 Basic Scheduling Model

The global fairness among different senders is not guaranteed by the IEEE 802.11 CSMA/CA protocol [26, 17]. We do not intend to change or improve fairness of this contention based channel access mechanism, but rather focus on the throughput improvement and fairness issues related to the scheduling phase of MAD within a sender. To simplify the analysis, let us consider a network of a single sender and N receivers. The set of instantaneous channel conditions (observed in terms of perceived signal reception power) from one sender to N receivers $\{\varphi_i(t)\}$ at any time t is always available and φ_i is wide-sense stationary and ergodic. The highest instantaneous data transmission rate $r_i(t)$ for receiver i at time t can be expressed by Shannon's law [24] as

$$r_i(t) = W \log_2 \left(1 + \frac{\varphi_i(t)}{\Gamma_i} \right) \quad (1)$$

where W is the bandwidth of the communication channel; Γ_i is the expected power of the additive white Gaussian noise (AWGN).

We propose a maximum relative gain scheduling policy which chooses a node with maximum gain relative to others during each data transmission phase of MAD. The *relative gain* function for receiver i at time t can be defined as

$$G_i(t) = c_i \left[\frac{\varphi_i(t) - \bar{\varphi}_i}{\Gamma_i} \right] \quad (2)$$

where c_i is a positive (nonzero) relative scaling factor for receiver i and $\bar{\varphi}_i$ is the mean of φ_i . Note that $\{\bar{\varphi}_i\}$ is not homogeneous among different receivers. Assuming flows to all receivers are always backlogged, the policy is expressed as the following:

$$i(t) = \arg \max \{G_i(t)\} \quad (3)$$

where $i(t)$ indicates that node i is chosen for data transmission at time t . If a tie occurs among multiple receivers in (3), it will be broken with equal probability. We assume rate selection is perfect so that all packets are received correctly. We further assume the data transmission duration is always fixed regardless of the data rate $r_i(t)$, since the super-packet size is proportional to the data rate. We find that maximum relative gain scheduling has the following nice properties when $c_i = r_i / \bar{\varphi}_i$ ⁴. This c_i value will be assumed when we refer to G_i in (2) in other sections.

⁴In this case relative gain function has range $(-1, +\infty)$

PROPOSITION 1. *With independent Rayleigh fading channels, asymptotic (long term) temporal fairness is guaranteed by maximum relative gain scheduling if $c_i = \frac{r_i}{\bar{\varphi}_i}$.*

PROOF. Given the above scheduling algorithm, the temporal share that receiver i will obtain in the long term can be expressed by

$$\phi_i = Pr(G_i > \max\{G_{j \neq i}\}) \quad (4)$$

According to Rayleigh fading channel, the c.d.f. for relative gain of receiver i is

$$\begin{aligned} P_i(\xi) &= Pr(G_i < \xi) = Pr(\varphi_i < \bar{\varphi}_i[1 + \xi]) \\ &= 1 - e^{-[1 + \xi]} \end{aligned} \quad (5)$$

the corresponding p.d.f. $p_i(\xi)$ is

$$p_i(\xi) = \begin{cases} e^{-[1 + \xi]} & (-1 < \xi) \\ 0 & (\xi \leq -1) \end{cases} \quad (6)$$

Thus, substituting into the rhs of (4) we get,

$$\begin{aligned} \phi_i &= \int_{-1}^{\infty} \prod_{j \neq i} [P_j(\xi)] p_i(\xi) d\xi \\ &= \int_{-1}^{\infty} \{1 - e^{-[1 + \xi]}\}^{N-1} e^{-[1 + \xi]} d\xi = \frac{1}{N} \end{aligned} \quad (7)$$

Hence, we prove that when $c_i = \frac{r_i}{\bar{\varphi}_i}$ maximum relative gain scheduling has asymptotic temporal fairness. \square

PROPOSITION 2. *With independent Rayleigh fading channels, the difference between asymptotic throughput produced by optimal temporal fair scheduling and that by maximum relative gain scheduling ($c_i = \frac{r_i}{\bar{\varphi}_i}$) is upper bounded by $W \log_2^{[1/\delta]}$, and $\delta = \min \left\{ \frac{1 + \bar{\varphi}_i / \Gamma_i}{1 + \bar{\varphi}_j / \Gamma_j}, \frac{1 + \Gamma_i / \bar{\varphi}_i}{1 + \Gamma_j / \bar{\varphi}_j} \mid \forall i, j \right\}$.*

PROOF. Suppose there is an optimal scheduling algorithm that generates highest overall throughput to all N receivers while maintaining asymptotic temporal fairness. For any receiver i let us define

$$\bar{r}_i = W \log_2^{[1 + \frac{\bar{\varphi}_i}{\Gamma_i}]} \quad (8)$$

and

$$\Delta_i(t) = r_i(t) - \bar{r}_i \quad (9)$$

The overall throughput for both the optimal and the maximum relative gain scheduling algorithms can be represented as

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T r_i(t) &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \bar{r}_i(t) \\ &+ \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \Delta_i(t) \end{aligned} \quad (10)$$

Since the first term in the rhs of (10) must be equal for both algorithms due to temporal fairness guarantees, the difference in overall throughput generated by the two algorithms can be decided by the second term.

Suppose the transmission duration is fixed regardless of the rate, both algorithms become time division schemes. Consider a situation where the maximum relative gain scheduling transmits to receiver a while the optimal algorithm

transmits to receiver b at time slot t . We know $G_a(t) \geq G_b(t) > -1$. Furthermore,

$$\begin{aligned} \Delta_a - \Delta_b &= W \log_2 \frac{\Gamma_a + \varphi_a}{\Gamma_a + \varphi_a} - W \log_2 \frac{\Gamma_b + \varphi_b}{\Gamma_b + \varphi_b} \\ &= W \log_2 \frac{\left[1 + (1 + G_a) \frac{\varphi_a}{\Gamma_a}\right] \left[1 + \frac{\varphi_b}{\Gamma_b}\right]}{\left[1 + (1 + G_b) \frac{\varphi_b}{\Gamma_b}\right] \left[1 + \frac{\varphi_a}{\Gamma_a}\right]} \\ &\geq W \log_2^{\delta(a,b)} \end{aligned} \quad (11)$$

where

$$\delta(a, b) = \min \left\{ \frac{1 + \varphi_b / \Gamma_b}{1 + \varphi_a / \Gamma_a}, \frac{1 + \Gamma_b / \varphi_b}{1 + \Gamma_a / \varphi_a} \right\}$$

If $\delta = \min \{ \delta(i, j) \mid \forall i, j \}$, $\Delta_a - \Delta_b \geq W \log_2^\delta$ holds for any time slot t where a stands for the receiver chosen by maximum relative gain scheduling in that slot. Hence, the overall throughput difference between optimal scheduling algorithm and maximum relative gain scheduling is upper bounded by $W \log_2^{[1/\delta]}$. \square

COROLLARY 1. *With independent Rayleigh fading channels, if $\forall i, j : \frac{\varphi_i}{\Gamma_i} = \frac{\varphi_j}{\Gamma_j}$, maximum relative gain scheduling achieves asymptotically highest overall throughput while maintaining temporal fairness.*

PROOF. The result is intuitive from proposition 2, as $\delta = 1$ which further implies $W \log_2^{[1/\delta]} = 0$. \square

6.2 Approximation Algorithms

In the previous subsection we have analyzed that ideally using maximum relative gain as the policy of scheduling decision is reasonably good when all channels are wide-sense stationary and ergodic with Rayleigh (fast) fading and time-invariant nonhomogeneous slow fading. However, in practice many assumptions may not hold. The channel condition is not only subjected to fast fading (such as Rayleigh fading), but also time-varying slow fading (path loss). The mean signal reception power to a particular receiver is no longer stationary but rather time-location dependent. Furthermore, in the presentation of channel probing design for MAD (see section 4) one can see that the overhead of inquiring instantaneous channel conditions is directly proportional to the number of receivers queried. Not all receivers may be queried for their instantaneous relative gains. Hence, the basic scheduling model described above cannot be directly implemented in MAD. Next, we propose two approximations to the above ideal scheduling algorithm. The two algorithms differ from existing scheduling algorithms exploiting multiuser diversity in that the scheduling decisions must be made based on partial feedbacks from receivers.

6.2.1 k -Set Round Robin

A simple way to mimic the behavior of maximum relative gain scheduling is to leverage the extra information of maximum relative gain among the k receivers (that have been queried in the channel probing phase of MAD) on top of basic round robin scheduling. The algorithm shown in Figure 6 works as the following. At the sender, the backlogged receivers are arranged in a round robin queue. In each iteration of the loop, the first k receivers will be queried in the channel probing phase to get both the feasible data rates and the relative gains with current channel conditions. The receiver with the maximum relative gain among the k queried

Definition:

- Q_1 : Servicing queue of backlogged receivers
- Q_2 : Serviced queue
- R : Set of receivers for channel probing

Routine:

```
while( $Q_1 \neq \emptyset$ )
{
    // dequeue at most  $k$  receivers
     $R = Q_1.dequeue(k)$ ;
    // channel probe dequeued receivers
     $R.probe()$ ;
    // service receiver with maximum relative gain
    among the dequeued ones
     $R_i = \arg \max \{G_i \mid R_i \in R\}$ ;
     $R_i.service()$ ;
    // push non-serviced back to head of  $Q_1$ 
     $Q_1.push(R - R_i)$ ;
    // append serviced receiver to tail of  $Q_2$ 
     $Q_2.append(R_i)$ ;
}
swap( $Q_1, Q_2$ );
```

Figure 6: k -Set Round Robin Scheduling

ones will be dequeued and put in a waiting queue for data transmission. The process repeats until all receivers are dequeued in current round. When the current queue is empty, it will be swapped with the waiting queue. Note that as more and more receivers are dequeued, it is possible that there will be less than k receivers to query in the channel probing phase of MAD. The reason not to put serviced receivers immediately to the back of the current queue is to enforce fairness, so that each receiver will be served exactly once in each round.

6.2.2 Revenue Based Scheduling

Prior research efforts on revenue (credit) based wireless fair scheduling [3, 20] have looked into exploiting time varying channel conditions in wireless fading environment. They implicitly assume that no cost is associated with channel probing in time division multiple access (TDMA) or CDMA cellular systems. However, we must reemphasize that multiuser diversity gain in the IEEE 802.11 wireless LAN is subject to the constraints of the feedback mechanism in MAD. First, due to CSMA/CA behavior of the network nodes, probing channel conditions immediately before data transmission is important as the obtained information may be stale after the transmitter resumes from random backoff. Second, increasing cost of channel probing and limited coherence time prevent the sender from querying all receivers (see section 7). Given these constraints, we propose a variation of the revenue based temporal fair scheduling for MAD. We will compare performance of this algorithm with that of k -set round robin algorithm in section 8.

The difference between traditional revenue based scheduling algorithms and the one proposed here is that in MAD, the receiver selection decision has to be made in two places: selecting the candidates for channel probing and selecting the final candidate for data transmission after channel probing. Figure 7 shows pseudo code for our revenue based scheduling algorithm, in which every receiver should receive a fair share of time in the long term. Initially every receiver

<u>Definition:</u>	
K_i^t	revenue of receiver i after t^{th} transmission
$R(G_i^t)$	reward function
U_i^t	time consumed by t^{th} transmission
Q_i^t	queue for receiver i after t^{th} transmission
C^t	candidates for channel probing
c^t	candidate for t^{th} transmission
<u>Receiver selection rule:</u>	
channel probing candidates:	
$C^{t+1} = \{i \text{high_to_low_rank}(K_i^t) \leq k\};$	
data transmission candidate:	
$c^{t+1} = \arg \max\{K_i^t + R(G_i^{t+1}) i \in C^{t+1}\};$	
<u>Revenue update rule:</u>	
for($i = 1; i \leq N; i++$)	
{	
if($Q_i^{t+1} \neq \emptyset$)	
if($i \neq c^{t+1}$)	
$K_i^{t+1} = K_i^t + \max(U_{c^{t+1}}^{t+1} - K_{c^{t+1}}^t, 0);$	
else	
$K_i^{t+1} = \max(K_i^t - U_i^{t+1}, 0);$	
else	
$K_i^{t+1} = 0;$	
}	

Figure 7: Revenue Based Scheduling

has revenue of zero. The transmitter will pick k nodes with highest revenue (ties broken randomly) for channel probing. After the channel probing, each probed receiver may receive an instant revenue reward defined by

$$R(G_i) = \max[\beta(1 + G_i), 0] \quad (12)$$

where β is a multiplicative constant. The ultimate receiver chosen for data transmission will be the receiver with highest sum of revenue and its instant reward. After the data transmission finished, the time taken will be deducted from the revenue of the serviced receiver. If the balance is negative, it will be reset to zero, while the difference will be credited to all other backlogged receivers. Note that the instant reward can only be used for final receiver selection but is not accounted for in the revenue updates to avoid long term unfairness. In depth discussions on why statistical fairness can be guaranteed by revenue based scheduling approach can be found in [3].

Note that there may be conditions where other metrics may achieve better performance. For instance, considering the number of packets in the queue destined for a receiver may improve the scheduling decision. Alternative scheduling metrics in various network conditions are to be evaluated in our future work.

7. PERFORMANCE ANALYSIS

In this section, we analyze the impact of MAD on the expected network throughput in a wireless LAN. We begin with derivation of expected transmission rates for a single transmitter and discuss the optimal number of receivers (k) to query during the channel probing of MAD to maximize the goodput of the channel. The single transmitter results will then be used to deduce the overall network capacity. We only use IEEE 802.11a for numerical results, but the analysis

is also applicable if MAD uses other existing 802.11 physical specifications. We assume free space pathloss and Rayleigh fading propagation models and AWGN in all derivations. Unless otherwise mentioned, the derivations are based on MAD using PAC data transmission scheme.

7.1 Optimal Number of Receivers to Query (k)

As mentioned at the end of section 4, one open question is what is the optimal number of receivers a sender should query in MAD. In this subsection, we attempt to analytically answer this in the context of using PAC in MAD. We consider a generic network model where a sender ν always has backlogged packets of constant length l to numerous potential receivers. ν is located at center of disk area \mathcal{D} of radius X while each receiver is randomly distributed in \mathcal{D} . X must not be larger than radio communication range at base rate in order to maintain connectivity. All channels between the sender and receivers have i.i.d. fast Rayleigh fading. For simplicity of analysis, we assume all packets are received correctly except for collisions. In this model, ν always randomly selects k receivers $\{\omega_i\}$ to probe channel conditions and transmits data to a receiver with the highest relative gain (see section 6).

7.1.1 Derivation of p.d.f. of MAD Transmission Rate

Suppose the physical radio supports M data rates denoted as r_1, r_2, \dots, r_M ($M = 8$ for 802.11a). A receiver can only suggest a data rate r_m if the estimated SNR is above a threshold η_m ($\eta_1 < \eta_2 < \dots < \eta_M$). For simplicity of formulation, we add rate $r_{M+1} > r_M$ and corresponding threshold $\eta_{M+1} = +\infty$ so that the probability for receiver to suggest rate r_{M+1} is 0. Thus, the probability for sender ν to transmit data using rate r_m , denoted as $P(r_m)$ can be expressed as follows.

$$P(r_m) = P(\eta_{r_m} \leq \eta^* < \eta_{r_{m+1}}) \quad (13)$$

where η^* denotes the SNR value of the receiver with the highest relative gain among those probed by the sender ν . If we denote q_i as a function of relative gain G_i (see (2)) for receiver ω_i : $q_i = 1 + G_i$, η^* is given as

$$\eta^* = P_t PL^* q^* / \Gamma \quad (14)$$

in which P_t is the transmission power; PL^* is the pathloss value at the receiver (denoted as ω^*) with highest relative gain; $q^* = q_{\omega^*}$; Γ is the energy level of the noise. Since relative gain is not a function of pathloss and vice versa, PL^* and q^* are statistically independent. To obtain the probability distribution of η^* , we first derive p.d.f. of PL^* and c.d.f. q^* .

Let x represent the distance between ν and ω^* . Using free space pathloss model [22], PL^* is given as:

$$PL^* = \frac{G_t G_r \lambda^2}{(4\pi)^2 L} \frac{1}{x^2} = Z \frac{1}{x^2}, \text{ where } Z = \frac{G_t G_r \lambda^2}{(4\pi)^2 L} \quad (15)$$

G_t and G_r represent the sender and receiver's antenna gains; λ is wavelength; L is a system loss factor not related to propagation. Since every receiver is randomly distributed in \mathcal{D} , x has the following c.d.f.

$$P(x < X) = (X/d)^2 \quad (16)$$

Hence, the c.d.f. and p.d.f of PL^* are

$$P(PL^* < Y) = P(x > \sqrt{Z/Y}) = 1 - Z/(Yd^2)$$

$$p(PL^*) = Z/(d \cdot PL^*)^2 \quad (17)$$

From (5) we can derive the c.d.f. for q^* as

$$P(q^* < Q) = (1 - e^{-Q})^k \quad (18)$$

Now we can derive the c.d.f of η^* (see (14)) via joint probability distribution of PL^* and q^* , given in (17) and (18):

$$\begin{aligned} P(\eta^* < \eta) &= P\left(\frac{P_t PL^* q^*}{\Gamma} < \eta\right) = P\left(PL^* q^* < \frac{\eta \Gamma}{P_t}\right) \\ &= \int_{Z/d^2}^{\infty} p(PL^*) \cdot P\left(q^* < \frac{\eta \Gamma}{P_t PL^*}\right) d(PL^*) \\ &= \int_{Z/d^2}^{\infty} \frac{Z \left(1 - e^{-\frac{\eta \Gamma}{P_t PL^*}}\right)^k}{(d \cdot PL^*)^2} d(PL^*) \\ &= 1 + \sum_{i=1}^k \frac{(-1)^i Z P_t}{i \eta \Gamma d^2} \binom{k}{i} \left(1 - e^{-\frac{i \eta \Gamma d^2}{Z P_t}}\right) \end{aligned} \quad (19)$$

Using (13) and (19) we can derive $P(r_m)$, the probability that ν will transmit at rate r_m , as follows:

$$\begin{aligned} P(r_m) &= P(\eta^* < \eta_{r_{m+1}}) - P(\eta^* < \eta_{r_m}) \\ &= \sum_{i=1}^k \frac{(-1)^i Z P_t}{i \eta_{r_{m+1}} \Gamma d^2} \binom{k}{i} \left(1 - e^{-\frac{i \eta_{r_{m+1}} \Gamma d^2}{Z P_t}}\right) \\ &\quad - \sum_{i=1}^k \frac{(-1)^i Z P_t}{i \eta_{r_m} \Gamma d^2} \binom{k}{i} \left(1 - e^{-\frac{i \eta_{r_m} \Gamma d^2}{Z P_t}}\right) \end{aligned} \quad (20)$$

7.1.2 Derivation of Expected Goodput per Data Transmission

Let us define transmission dialogue as a sequence of packet exchanges between ν and $\{\omega_i\}$ from (G)RTS till ACK for each data transmission. The goodput per data transmission is defined as the ratio of payload size over duration of a transmission dialogue. With the results of (20), we now derive the expected goodput of ν per transmission dialogue for MAD.

First, the goodput in a transmission dialogue where ν transmits at rate r_m , denoted as S_{r_m} can be expressed as follows:

$$S_{r_m} = \frac{H_{r_m}}{T_c^{grts/cts} + T_{sp}(r_m) + T_{SIFS} + 2\tau + T_{ACK}} \quad (21)$$

where H_{r_m} is the payload size for ν to transmit at rate r_m ; $T_c^{grts/cts}$ is the time for GRTS/CTS exchange; $T_{sp}(r_m)$ is the time to transmit a PAC Super Data frame at r_m ; T_{ACK} is the time for ACK transmission. They are given as follows:

$$\begin{aligned} H_{r_m} &= l(r_m/r_1) \\ T_c^{grts/cts} &= T_{GRTS} + \tau + T_{SIFS} + k(T_{CTS} + \tau + T_{SIFS}) \\ T_{GRTS} &= T_{RTS} + (k-1)l_{ADDR}/r_1 \\ T_{sp}(r_m) &= \begin{cases} T_{PH} + T_{MH} + \frac{l}{r_1} & (m=1) \\ T_{sp}(r_1) + T_{SF} + T_{SPH} \frac{r_m}{r_1} & (1 < m \leq M) \end{cases} \end{aligned}$$

where τ is propagation delay. l_{ADDR} is the length of MAC address in bits. All T s in the rhs of the above equations

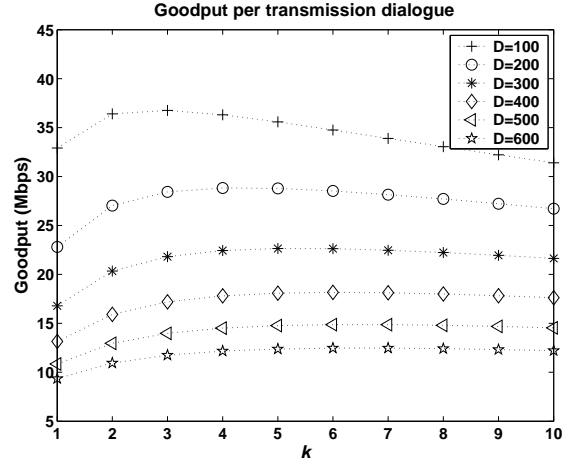


Figure 8: Expected Goodput per Data Transmission with Difference Values of k

refer to the time to transmit corresponding physical or MAC header, or MAC control frame (transmitted at base rate r_1), denoted by their subscripts. PH , SPH , MH and SF refers to 802.11 physical header, short physical header, MAC header and Super Frame (see Figure 2(d)) respectively.

With (20) and (21), the expected goodput per data transmission by ν , denoted as S can be computed by

$$S = \sum_{m=1}^M P(r_m) S_{r_m} \quad (22)$$

7.1.3 Optimal Value of k

As we mentioned, this paper will only use IEEE 802.11a PHY as an example to analyze and evaluate the performance of MAD, though MAD is applicable to other 802.11 physical layer specifications. Based on the above derivation results in this subsection, we can plot the expected goodput per data transmission vs. the value of k , if given the specifications of radio devices including receiving sensitivity, transmission power, the modulation and coding schemes for each transmission rate. Using the 802.11a standard [12] and an actual 802.11a compliant network card product specification found in [29], we set up the parameters as follows. l is set as 1K bytes. Since the transmission range of a node is calculated as approximately 600m with free space pathloss alone, d is varied from 100m up to 600m in the plot to represent difference cases in wireless LAN environment. We must find the optimal value of k that produces close-to-highest goodput in all cases. The plotted graph is shown in Figure 8, where we make several observations. First, the optimal value of k increases with d . Second, the second moment of S decreases with increasing k and S improves very little if any, when k is larger than 3. Third, S decreases in all cases when k is greater than 6, this shows that the overhead of channel probing overshadows the multiuser diversity gains. Finally, the figure also demonstrates that when receivers are all close to the sender ($d = 100$), the probability of the sender using peak data rate for data transmission is high. As a result, multiuser diversity gain is overshadowed by the overhead of MAD even with smaller value of k (3 in this case).

In practice, the overhead by PAC may be even higher as some of the packets in a super-packet may not be re-

ceived successfully. Furthermore, the coherence time will pose another constraint on the value of k . If k is large, by the time the last receiver reports its channel condition, the channel conditions from earlier reports may have been obsolete. Considering these factors, the optimal value of k in MAD is clearly 3 when PAC is used. Though not shown here, we conducted similar analysis for MAD using OAR in data transmission and found that 3 is also reasonable for k in that context.

7.2 Network Throughput

Now we analyze the overall network throughput in an 802.11a wireless LAN that employs MAD. We assume all senders in the wireless LAN operate independently with same probability distribution of data transmission rates derived in subsection 7.1 and have backlogged queues of packets of equal length. Each sender operates the same way as described in subsection 7.1. As GRTS serves the same purpose as RTS for channel contention, the network undergoes the same state transitions as in single rate 802.11 DCF. Thus, results from [6] can be used to derive the expected network throughput C :

$$C = \frac{P_{tr}P_sE[H]}{(1 - P_{tr})\sigma + P_{tr}P_sT_s + P_{tr}(1 - P_s)T_c}$$

$$P_{tr} = 1 - (1 - \zeta)^n$$

$$P_s = \frac{n\zeta(1 - \zeta)^{n-1}}{P_{tr}} = \frac{n\zeta(1 - \zeta)^{n-1}}{1 - (1 - \zeta)^n}$$

$$\zeta = \frac{2(1 - 2p)}{(1 - 2p)(CW_{min} + 1) + p \cdot CW_{min}(1 - (2p)^m)}$$

$$p = 1 - (1 - \zeta)^{n-1}$$

$$m = \log(CW_{max}/CW_{min})$$

n is the total number of senders in the network. σ is a slot time. ζ can be solved using the last three equations above; CW_{min} , CW_{max} is the initial and maximum value of contention window; $E[H]$ is the expected payload size; T_s is the average time when the channel is busy due to a successful transmission dialogue; T_c is the duration in which the channel is sensed busy by a non-transmitting node due to the collision of a GRTS packet.

When MAD employs PAC for data transmission, the values of $E[H]$, T_c , and T_s are calculated as follows.

$$E[H] = \sum_{m=1}^M P(r_m) l \frac{r_m}{r_1}$$

$$T_c = T_{GRTS} + T_{DIFS} + \tau$$

$$T_s = T_c + kT_{CTS} + T_{SD} + T_{ACK} + (k + 2)(T_{SIFS} + \tau)$$

$$T_{SD} = T_{MH} + \frac{l}{r_1} + T_{PH} + \sum_{m=2}^M P(r_m) \left(T_{SF} + \frac{r_m}{r_1} T_{SPH} \right)$$

$P(r_m)$ is the probability that a node transmit at rate r_m after the channel probing as given in (20).

If OAR is used instead of PAC by MAD, $E[H]$ and T_c remain the same but T_s is larger. Due to space limitation, the derivation of T_s for MAD using OAR will not be shown here.

Based on the above derivations, the duration of the contention period is invariant of k , and the expected network throughput reaches the maximum when the goodput per transmission dialogue is maximized, assuming $T_{GRTS} \approx T_{RTS}$.

Table 1: Comparison of MAD, OAR and Base-rate 802.11 DCF (with 802.11a PHY)

	$C(Mbps)$	$T_s(\mu s)$	$E[H](bits)$
MAD _{PAC}	21.41	1844.33	40414
MAD _{OAR}	17.40	2284.25	40414
OAR	14.35	1968.97	28792
802.11-DCF	4.78	1635.33	8192

This implies the network throughput improvement by MAD is near its maximum when $k = 3$. Next, we compare the network throughput obtained by MAD, OAR and original 802.11 DCF MAC, based on above derivation and results in [23], [6]. The parameters used to generate the numerical results are as follows: n equals 6; maximum distance between receiver and sender is 300m; other parameters are the same as those used in subsection 7.1 (k is 3). The numerical results are shown in Table 1. In this specific case, the expected payload size ($E[H]$) in MAD is improved by 40% over OAR (i.e. multiuser diversity gain). Within MAD, PAC uses 20% less time (T_s) to transmit same payload than OAR. Thus, MAD employing PAC as the data transmission scheme achieves the best network throughput, a 49% increase over original OAR. Note that a simulation experiment approximating this specific case has been conducted to compare with the numerical result of network throughput. That simulation result will be given later in the next section.

8. PERFORMANCE EVALUATION

We conducted an extensive suite of simulation experiments to evaluate the performance of MAD. The main performance metrics considered in our studies include network throughput and temporal fairness. We compare existing rate adaptation MAC protocols proposed for wireless LAN operated in 802.11 DCF mode, such as OAR and ARF, with four variants of MAD. Each variant of MAD employs a specific combination of data transmission (PAC or OAR), and scheduling (k -set round robin or revenue based scheduling) schemes. In all four variants k is fixed at 3, with which the overall network throughput approximates the maximum based on our analysis results in the previous section. β in MAD_{OAR-Rev} or MAD_{PAC-Rev} is set as 5000us.

The simulation experiments are conducted using QualNet (a commercial network simulator derived from GloMoSim [1]), which provides a set of detailed physical layer models [28]. Based on the IEEE 802.11a specifications, a node can choose one of the eight data rates, 6, 9, 12, 18, 24, 36, 48, and 54Mbps. The transmission power and radio sensitivities for different data rates are configured according to the manufacturer's specification in [29]. We use Free space and Rayleigh fading as the propagation model in all the experiments conducted. The propagation and PHY layer models in our simulation account for the effects of interference as well as stale channel estimation due to the time lag between channel probing and packet reception.

Five sets of simulation experiments have been carried out, in each of which one of the following five parameters is varied: density of each traffic flow, number of traffic flows per transmitter, distance between a transmitter and receiver, duration of channel coherence time, and network topology. To isolate effects of these parameters on network perfor-

D :	distance between a transmitter and receiver
D_{max} :	maximal value of D
N_f :	number of traffic flows per transmitter
N_{tr} :	number of transmitters in the network
$MAD_{OAR-k-Set}$:	MAD (OAR and k -Set round robin scheduling)
$MAD_{PAC-k-Set}$:	MAD (PAC and k -Set round robin scheduling)
$MAD_{OAR-Rev}$:	MAD (OAR and Revenue based scheduling)
$MAD_{PAC-Rev}$:	MAD (PAC and Revenue based scheduling)

Figure 9: Notations

mance, a star topology is used except in the last set of experiments. This star topology mimics a wireless LAN comprising an access point surrounded by a set of mobile stations of equal distance to the access point. The mobile stations do not transmit data packets themselves but passively listen to the access point and receive data packets from the access point if any. The access point has a number of CBR traffic flows. The destinations of any two traffic flows are different. Unless otherwise stated, the packet length is fixed at 1K bytes, and traffic flows are always backlogged. In the last set of experiments, a star topology with mobile stations located at randomly chosen distance to the access point, and a random topology are used to study the impacts of network topology on the performance of MAD. Figure 9 lists the notations introduced for brevity.

8.1 Impact of Traffic Density

Figure 10 depicts the network throughput with ARF, OAR, PAC and four variants of MAD under various traffic load. In this experiment, N_f is set as 9 and the traffic rate of each CBR session is varied from 25 to 275 pkts/s, corresponding to an aggregate traffic load of 1.8Mbps up to 20.3Mbps. D is set as 300m. When the traffic load is low and the channel is under-utilized, the network throughput is identical among different MAC mechanisms, with packet delivery ratio equal to 100%. With ARF or OAR, the network saturates at a traffic load of 2.7Mbps or 7Mbps. As the volume of traffic increases, the throughput gains of MAD rise and $MAD_{PAC-Rev}$ performs the best among the four variants of MAD. With $MAD_{PAC-Rev}$, the network saturates at 14.4Mbps, which is twice as high as that with OAR.

8.2 Impact of D

Figure 11 compares the network throughput for ARF, OAR, PAC, and four MAD variants when D is varied from 50 to 500m. N_f is set to 3 or 15. It is noted that the network throughput drops as D increases. This result is intuitive since the probability of a transmitter using high data rates declines.

PAC vs. OAR

When used alone, PAC improves the network throughput significantly compared to OAR when D is small. This result again demonstrates that MAC and PHY overhead dominates transmission time and is aggravated with high data rates. Such overhead is reduced significantly by PAC. It is worth noting that OAR slightly outperforms PAC when D is relatively large. While larger D (lower SNR) reduces the performance benefits of both PAC and OAR as fewer packets can be transmitted consecutively, the performance difference between them also decreases because PAC can remove less Interframe Space (IFS) and ACK time when fewer packets can be concatenated. Moreover, a single transmission in PAC may include multiple erroneous packets while OAR al-

ways stops upon first packet error. Thus, PAC has higher retransmission overhead than OAR when channel condition is not favorable.

k -set Round Robin vs. Revenue Based Scheduling

It is observed in Figure 11 that revenue based scheduling generally outperforms k -set round robin. k -set round robin enforces fairness among individual flows at fine time granularity with the flows transmitting in strict order. With revenue based scheduling, a flow corresponding to a receiver with good channel condition is rewarded with instant credit and thus allowed to occupy a channel more often in the short term than other flows. Hence a transmitter can opportunistically exploit favorable channel conditions pertaining to individual flows at the price of the short term fairness among flows and achieve better network throughput.

MAD

Among all the mechanisms investigated, $MAD_{PAC-Rev}$ performs the best. The benefits of revenue based scheduling are best exploited when PAC is used in MAD since PAC can enhance the goodput per transmission dialogue the most among all the data transmission schemes in presence of good channel condition. Therefore, despite slightly poorer performance of PAC against OAR with larger D , the network throughput is always maximized with $MAD_{PAC-Rev}$, as shown in Figure 11. Compared to ARF, $MAD_{PAC-Rev}$ improves the network throughput by 136% to 452%. Compared to OAR, the improvement is 30% to 120%. If we compare four MAD variants with OAR and PAC, we can conclude that the throughput improvement of $MAD_{PAC-Rev}$ is largely attributed to reduced MAC and PHY layer overhead with small D and to improved data transmission rates with large D .

8.3 Impact of Number of Flows

Figure 12 plots the network throughput obtained with different MAC mechanisms when N_f is varied from 3 to 18. D is set as 300m. These results again demonstrate that $MAD_{PAC-Rev}$ performs best among all the mechanisms investigated, achieving 70% to 100% throughput improvement over OAR. Since a transmitter can always query three receivers during channel probing, $MAD_{PAC-Rev}$ can significantly increase the probability of using high data rates with even a small number of flows. As the channel conditions among all receivers are statically independent, the average transmission rate remains relatively constant with increased N_f . However, in $MAD_{OAR-k-Set}$ or $MAD_{PAC-k-Set}$, the probability of the transmitter not querying three receivers drops when N_f increases. Therefore, the average data transmission rates used by the transmitter improves with larger N_f , resulting in higher network throughput.

8.4 Impact of Channel Coherence Time

In this experiment, we vary the channel coherence time to evaluate the impact of node mobility on network throughput in a wireless LAN. Figure 13 compares the network throughput obtained with ARF, OAR, and MAD. The channel coherence time is varied such that it corresponds to mobile speed of 2, 4, 6, 8, or 10m/s. Figure 13 shows that as channel coherence time decreases, the network throughput drops due to more packet retransmissions. Among all the mechanisms examined, $MAD_{PAC-Rev}$ performs the best and improves the network throughput by up to 25% with

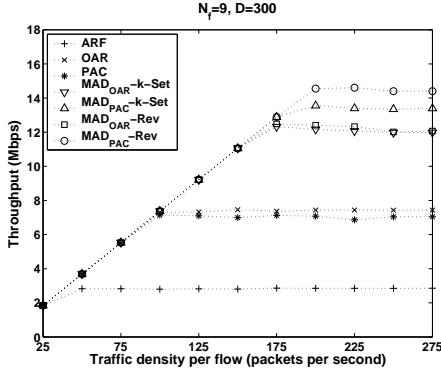


Figure 10: Network throughput vs. traffic density

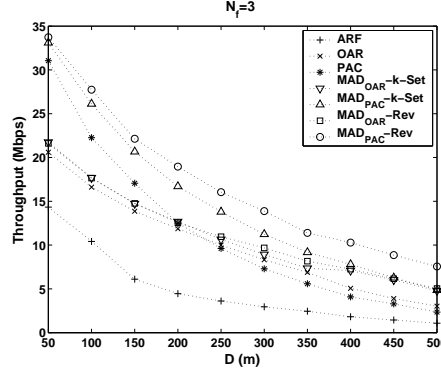


Figure 11: Network throughput vs. D

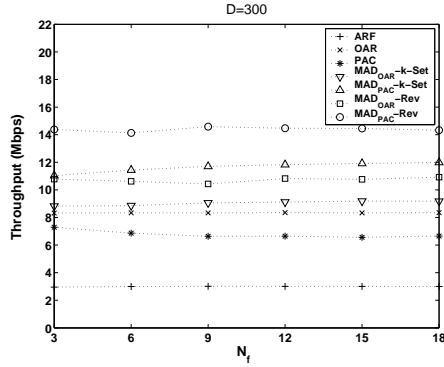
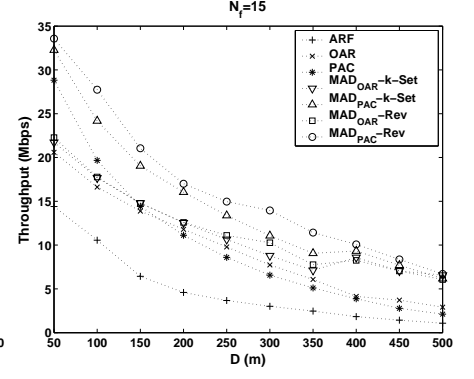


Figure 12: Network throughput vs. N_f .

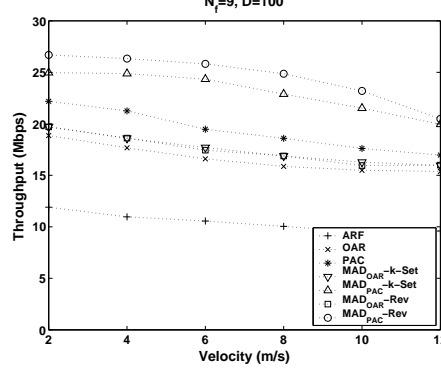
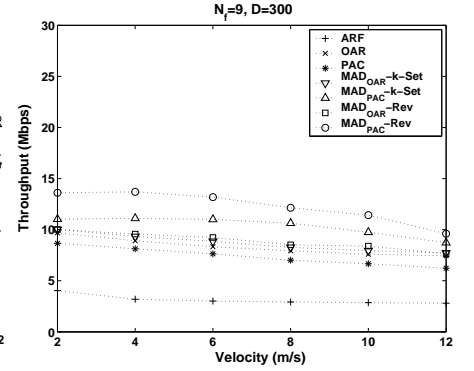


Figure 13: Network throughput vs. coherence time



respect to $MAD_{PAC-k-Set}$. The throughput improvement by $MAD_{PAC-Rev}$ is more evident when the channel coherence time is longer. Such improvement occurs because with $MAD_{PAC-Rev}$ can take advantage of a slow varying channel and select a flow with consistent good channel conditions for data transmission a few times in a short period of time.

8.5 Network Performance with Different Topologies

In this set of experiments, we evaluate the network throughput in two other topologies: a star topology with variable D and a random topology.

8.5.1 Star Topology with Variable D

In the previous experiments, we used a star topology with constant D . In this experiment, we let D be uniformly distributed among 0 and a constant denoted as D_{max} . D_{max} is varied from 100 to 600 meters. Figure 14(a) compares the network throughput of different MAC mechanisms. N_f is set as 9. Figure 14(a) shows that $MAD_{PAC-Rev}$ also works well when individual receivers of a transmitter observe uneven time-averaging channel conditions. Similarly as before, $MAD_{PAC-Rev}$ performs best and significantly improves the network throughput by up to 13Mbps compared to OAR. As in star topology with constant D , the relative ratio of throughput improvement by $MAD_{PAC-Rev}$ grows with D_{max} .

Figure 14(b) plots the temporal share of each traffic flow when D_{max} is equal to 300m. The temporal share for a

traffic flow is defined as the ratio of time used to transmit data packets belonging to this flow over the total time consumed for data transmission at the access point for the entire simulation run. We exclude contention time and MAC and PHY layer overhead when measuring temporal share of flows. Figure 14(b) shows that $MAD_{PAC-k-Set}$ achieves slightly better temporal fairness than $MAD_{PAC-Rev}$. With $MAD_{PAC-Rev}$, each flow has a temporal share ranging from 10% to 11.5% (11.1% for 9 flows with extreme temporal fairness), despite the fact that receivers for these flows are not evenly far apart from the transmitter and observe different time averaging channel conditions.

8.5.2 Random Topology

In this set of experiments, we set up a random topology which comprises 25 nodes uniformly distributed in a two dimensional terrain of size 200×200 . In this terrain, D_{max} is 283m. There are randomly chosen N_{tr} transmitters in this network, where each transmitter has 5 traffic flows, each to a randomly chosen receiver. Figure 15(a) shows the network throughput with different MAC schemes when N_{tr} is varied from 1 to 6. We note that early observations in star topologies are also applicable here. $MAD_{PAC-Rev}$ consistently outperforms all the others, resulting in significant throughput improvement. It increases the network throughput by up to 332% with respect to ARF and 71% with respect to OAR. The network throughput does not vary much with N_{tr} since contention time remains rather constant when N_{tr} is small. The network throughput of $MAD_{PAC-Rev}$ (shown in

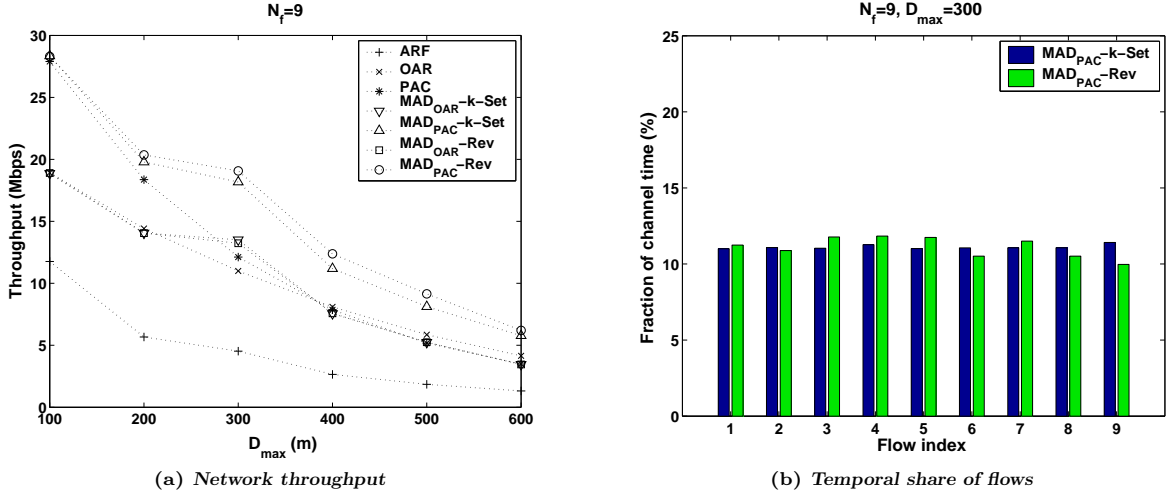


Figure 14: Star topology with variable D

Figure 15(a)) also matches well with our analysis in Table 1 (21.41Mbps). Note that the network throughput observed in simulation is slightly lower than predicted by the analysis as not all packets are received correctly in the simulation.

Figure 15(b) depicts the temporal share of flows at each transmitter in a random topology with three transmitters. The figure demonstrates that although $MAD_{PAC-Rev}$ sacrifices short term temporal fairness to exploit slow varying channel conditions, it can ensure temporal fairness in the long run. With $MAD_{PAC-Rev}$, the temporal share of flows range from 19% to 21% for the three transmitters as opposed to 20% with strict temporal fairness.

9. CONCLUSION

The increasing number of users and new wave of applications have created strong demand for high bandwidth wireless LANs. In this paper, we proposed a medium access mechanism, MAD, to improve network throughput by exploiting multiuser diversity effects in wireless LANs, which can be implemented as an extension to 802.11a MAC/PHY with minor modifications. We have identified and addressed the following three main design issues of MAD: channel probing, data transmission for link goodput optimization, and receiver scheduling. We analyzed the tradeoff between throughput gain and the overhead of channel probing. We also improved link utilization, which is crucial to the success of MAD, by proposing a new rate adaptation scheme, PAC. PAC itself can achieve up to 50% link goodput gain over existing rate adaptation schemes in 802.11a wireless LANs, when channel conditions are favorable. Further, we derived the asymptotic performance bound of MAD by two scheduling algorithms that maintain temporal fair share among multiple data flows.

The experiment results show that by using OAR as the data transmission scheme, MAD achieves up to 70% network throughput improvement over original OAR, which proves that exploiting multiuser diversity in wireless LANs can enhance overall network performance substantially. Combining the superior benefits of PAC and multiuser diversity gains, MAD consistently improves the overall network throughput over OAR in the range of 30% ~ 120%. When compared with ARF, the first rate adaptation scheme used

in Lucent's WaveLAN II, MAD's improvement is from 120% up to 452%. The simulation results also show that temporal fairness among individual flows is well maintained in MAD.

Multiuser diversity is a common effect and can be leveraged in many mobile wireless network environments including wireless LANs and mobile ad hoc networks. While MAD needs to rely on the existence of backlogged traffic to multiple receivers to improve the network throughput, it has been observed that downlinks carry more traffic than uplinks in typical campus wireless LANs [18], making heavily loaded access point an ideal location to deploy MAD. In mobile ad hoc networks, we conjecture similar situations to be observed at intermediate routers, but the design and performance metrics of MAD may be affected by other factors, including the interaction of MAD scheduling and routing protocols or the tradeoff between end-to-end delay and throughput. Recent research studies including [13, 7] have proposed utilization of path diversity, which can be orthogonal to MAD. One important piece of future work is to identify and address a different set of issues and challenges to explore multiuser diversity in various mobile networking scenarios.

10. REFERENCES

- [1] L. Bajaj, M. Takai, R. Ahuja, and R. Bagrodia. Simulation of large-scale heterogeneous communication systems. In *MILCOM'99*, November 1999.
- [2] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana, and A. Viterbi. CDMA/HDR: A bandwidth-efficient high-speed wireless data service for nomadic users. *IEEE Communications Magazine*, 38(7), 2000.
- [3] B. Bensaou, D. H. K. Tsang, and K. T. Chan. Credit-based fair queueing (CBFQ): a simple service-scheduling algorithm for packet-switched networks. *IEEE/ACM Transactions on Networking*, 9(5), 2001.
- [4] F. Berggren and R. Jäntti. Multiuser scheduling over rayleigh fading channels. In *IEEE GLOBECOM'03*, San Francisco, CA, Dec 2003.
- [5] P. Bhagwat, P. Bhattacharya, A. Krishna, and S. K.

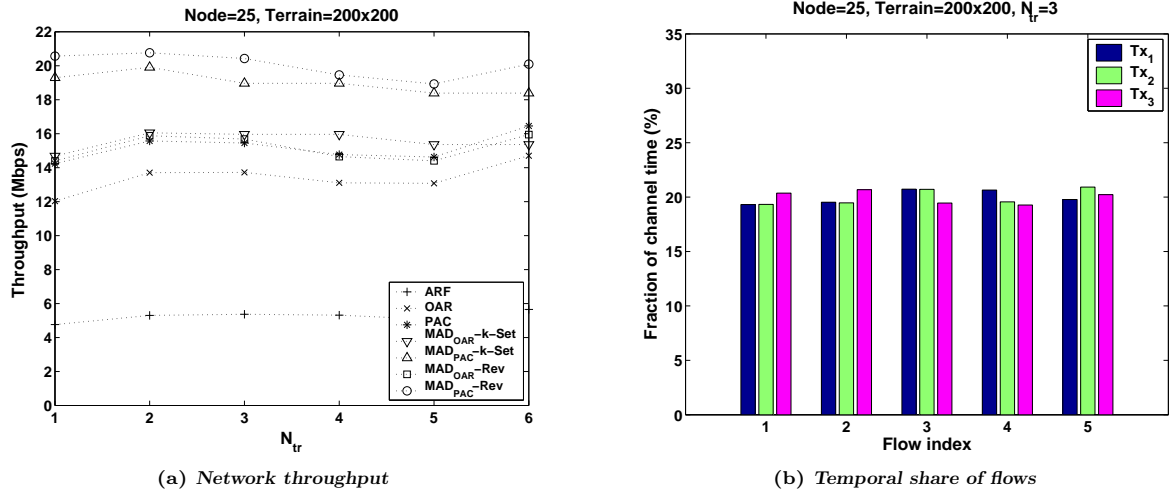


Figure 15: Random topology

- Tripathi. Enhancing throughput over wireless LANs using channel state dependent packet scheduling. In *IEEE INFOCOM'96*, San Francisco, CA, Mar 1996.
- [6] G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3), Mar 2000.
- [7] R. R. Choudhury and N. H. Vaidya. MAC-layer anycasting in ad hoc networks. *Computer Communication Review*, 34(1):75–80, 2004.
- [8] J. Heiskala and J. Terry. *OFDM Wireless LANs: A theoretical and practical guide*. SAMS, 2001.
- [9] G. Holland, N. Vaidya, and P. Bahl. A rate-adaptive MAC protocol for wireless networks. In *ACM MOBICOM'01*, Rome, Italy, July 2001.
- [10] H. Holma and A. Toskala. *WCDMA for UMTS, Radio access for third generation mobile communications*. Wiley, 2nd edition, 2002.
- [11] IEEE Computer Society LAN/MAN Standards Committee. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *ANSI/IEEE Std 802.11, 1999 Edition*.
- [12] IEEE Computer Society LAN/MAN Standards Committee. IEEE Std 802.11a-1999 (Supplement to IEEE Std 802.11-1999). 1999.
- [13] S. Jain and S. R. Das. Exploiting path diversity in the link layer in wireless ad hoc networks. Technical report, SUNY Stony Brook, WINGS Lab, July 2003.
- [14] N. Joshi, S. R. Kadaba, S. Patel, and G. S. Sundaram. Downlink scheduling in CDMA data networks. In *ACM MOBICOM'00*, Seattle, WA, Jun 2000.
- [15] A. Kamerman and L. Monteban. WaveLAN II: A high-performance wireless LAN for the unlicensed band. *Bell Labs Technical Journal*, Summer 1997.
- [16] R. Knopp and P. A. Humblet. Information capacity and power control in single-cell multiuser communications. In *IEEE ICC'95*, Seattle, WA, Jun 1995.
- [17] C. E. Koksal, H. Kassab, and H. Balakrishnan. An analysis of short-term fairness in wireless media access protocols. In *ACM SIGMETRICS'00*, Santa Clara, CA, Jun 2000.
- [18] D. Kotz and K. Essien. Analysis of a campus-wide wireless network. In *ACM MOBICOM'02*, Atlanta, GA, Sep 2002.
- [19] X. Liu, E. Chong, and N. Shroff. Opportunistic transmission scheduling with resource-sharing constraints in wireless networks. *IEEE Journal on Selected Areas in Communications*, 19(10), 2001.
- [20] Y. Liu, S. Gruhl, and E. Knightly. WCFQ: an opportunistic wireless scheduler with statistical fairness bounds. *IEEE Transactions on Wireless Communication*, 2(5), Sep 2003.
- [21] X. Qin and R. Berry. Exploiting multiuser diversity for medium access control in wireless networks. In *IEEE INFOCOM'03*, San Francisco, CA, Mar 2003.
- [22] T. S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall, 2nd edition, 2001.
- [23] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly. OAR: An opportunistic auto-rate media access protocol for ad hoc networks. In *ACM MOBICOM'02*, Atlanta, GA, Sep 2002.
- [24] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656, July and October 1948.
- [25] F. A. Tobagi and L. Kleinrock. Packet switching in radio channels: Part II - the hidden terminal problem in carrier sense multiple-access modes and the busy-tone solution. *IEEE Transactions on Communications*, COM-23(12), 1975.
- [26] N. H. Vaidya, P. Bahl, and S. Gupta. Distributed fair scheduling in a wireless LAN. In *ACM MOBICOM'00*, Boston, MA, Aug 2000.
- [27] P. Viswanath, D. N. C. Tse, and R. Laroia. Opportunistic beamforming using dumb antennas. *IEEE Transactions of Information Theory*, 48(6), 2002.
- [28] QualNet User Manual v3.7. Scalable Network Technologies, Culver City, CA. <http://www.qualnet.com>.
- [29] Aries2 Product Spec. http://www.senao.com.tw/english/product/product_wireless01_outdoor_1.asp?pgt1=Wireless&tp1id=02&tp2id=03&proid=000029#3.