Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2007, Article ID 12597, 14 pages
doi:10.1155/2007/12597

Research Article

Towards Scalable MAC Design for High-Speed Wireless LANs

Yuan Yuan,[1] William A. Arbaugh,[1] and Songwu Lu[2]

[1] *Department of Computer Science, University of Maryland, College Park, MD 20742, USA*
[2] *Computer Science Department, University of California, Los Angeles, CA 90095, USA*

Received 29 July 2006; Revised 30 November 2006; Accepted 26 April 2007

Recommended by Huaiyu Dai

The growing popularity of wireless LANs has spurred rapid evolution in physical-layer technologies and wide deployment in diverse environments. The ability of protocols in wireless data networks to cater to a large number of users, equipped with high-speed wireless devices, becomes ever critical. In this paper, we propose a token-coordinated random access MAC (TMAC) framework that scales to various population sizes and a wide range of high physical-layer rates. TMAC takes a two-tier design approach, employing *centralized, coarse-grained channel regulation,* and *distributed, fine-grained random access.* The higher tier organizes stations into multiple token groups and permits only the stations in one group to contend for the channel at a time. This token mechanism effectively controls the maximum intensity of channel contention and gracefully scales to diverse population sizes. At the lower tier, we propose an adaptive channel sharing model working with the distributed random access, which largely reduces protocol overhead and exploits rate diversity among stations. Results from analysis and extensive simulations demonstrate that TMAC achieves a scalable network throughput as user size increases from 15 to over 300. At the same time, TMAC improves the overall throughput of wireless LANs by approximately 100% at link capacity of 216 Mb/s, as compared with the widely adopted DCF scheme.

1. INTRODUCTION

Scalability has been a key design requirement for both the wired Internet and wireless networks. In the context of medium access control (MAC) protocol, a desirable wireless MAC solution should scale to both different physical-layer rates (from 1 second to 100 seconds of Mbps) and various user populations (from 1 second to 100 seconds of active users), in order to keep pace with technology advances at the physical layer and meet the deployment requirements in practice. In recent years, researchers have proposed numerous wireless MAC solutions (to be discussed in Section 7). However, the issue of designing a scalable framework for wireless MAC has not been adequately addressed. In this paper, we present our Token-coordinated random access MAC (TMAC) scheme, a scalable MAC framework for wireless LANs.

TMAC is motivated by two technology and deployment trends. First, the next-generation wireless data networks (e.g., IEEE 802.11n [1]) promise to deliver much higher data rates in the order of 100 seconds of Mbps [2], through advanced antennas, enhanced modulation, and transmission techniques. This requires MAC-layer solutions to develop in pace with high-capacity physical layers. However, the widely adopted IEEE 802.11 MAC [3], using distributed coordination function (DCF), does not scale to the increasing physical-layer rates. According to our analysis and simulations, (Table 4 lists the MAC and physical-layer parameters used in all analysis and simulation. The parameters are chosen according to the specification of 802.11a standard [4] and the leading proposal of 802.11n [2].) DCF MAC delivers as low as 30 Mb/s throughput at the MAC layer with the bit-rate of 216 Mbps, utilizing merely 14% of channel capacity. Second, high-speed wireless networks are being deployed in much more diversified environments, which typically include conference, enterprise, hospital, and campus settings. In some of these scenarios, each access point (AP) has to support a much larger user population and be able to accommodate considerable variations in the number of active stations. The wireless protocols should not constraint the number of potential users handled by a single AP. However, the performance of current MAC proposals [3, 5–8] does not scale as user population expands. Specifically, at user population of 300, the DCF MAC not only results in 57% degradation in aggregate throughput but also leads to starvation for most stations, as shown in our simulations. In summary, it is essential to design a wireless MAC scheme that effectively tackles the scalability issues in the following three aspects:

(i) *user population*, that generally leads to excessive collisions and prolonged backoffs,
(ii) *physical-layer capacity*, that requires the MAC-layer throughput scales up in proportion to the increases in physical-layer rate,
(iii) *protocol overhead*, that results in high signaling overhead due to various interframe spacings, acknowledgements (ACK), and optional RTS/CTS messages.

TMAC tackles these three scalability issues and provides an efficient hierarchical channel access framework by combining the best features of both reservation-based [9, 10] and contention-based [3, 11] MAC paradigms. At the higher tier, TMAC regulates channel access via a central token coordinator, residing at the AP, by organizing contending stations into multiple token groups. Each token group accommodates a small number of stations (say, less than 25). At any given time, TMAC grants only one group the right to contend for channel access, thus controlling the maximum intensity of contention while offering scalable network throughput. At the lower tier, TMAC incorporates an adaptive channel sharing model, which grants a station a temporal share depending on its current channel quality. Within the granted channel share, MAC-layer batch transmissions or physical-layer concatenation [8] can be incorporated to reduce the signaling overhead. Effectively, TMAC enables adaptive channel sharing, as opposed to the fixed static sharing notion in terms of either equal throughput [3] or identical temporal share [5], to achieve better capacity scalability and protocol overhead scalability.

The extensive analysis and simulation study have confirmed the effectiveness of the TMAC design. We analytically show the scalable performance of TMAC and the gain of the adaptive channel sharing model over the existing schemes [3, 5]. Simulation results demonstrate that TMAC achieves a scalable network throughput and high efficiency of channel utilization, under different population sizes and diverse transmission rates. Specifically, as the active user population grows from 15 to over 300, TMAC experiences less than 6% throughput degradation, while the network throughput in DCF decreases approximately by 50%. Furthermore, the effective TMAC throughput reaches more than 100 Mb/s at link capacity of 216 Mb/s, whereas the optimal throughput is below 30 Mb/s in DCF and about 54 Mb/s using the opportunistic auto rate (OAR),[1] a well-known scheme for enhancing DCF.

The rest of the paper is organized as follows. The next section identifies underlying scalability issues and limitations of the legacy MAC solutions. Section 3 presents the TMAC design. In Section 4, we analytically study the scalability of TMAC, which is further evaluated through extensive simulations in Section 5. We discuss design alternatives in Section 6. Section 7 outlines the related work. We conclude the paper in Section 8.

2. CHALLENGES IN SCALABLE WIRELESS MAC DESIGN

In this section, we identify three major scalability issues in wireless MAC and analyze limitations of current MAC solutions [2, 4]. We focus on high-capacity, packet-switched wireless LANs, operating at the infrastructure mode. Within a wireless cell, all packet transmissions between stations pass through the central AP. The wireless channel is shared among uplink (from a station to the AP) and downlink (from the AP to a station), and used for transmitting both data and control messages. APs connected to the wired may have connection directly to the wired Internet (e.g., in WLANs). Different APs may use the same frequency channel due to insufficient number of channels or dense deployment, and so forth.

2.1. Scalability issues

We consider the scalability issues in wireless MAC protocols along the following three dimensions.

Capacity scalability

Advances in physical-layer technologies have greatly improved the link capacity in wireless LANs. The initial 1 ∼ 11 Mbps data rates specified in 802.11b standard [3] have been elevated to 54 Mb/s in 802.11a/g [4], and to 100 seconds of Mb/s in 802.11n [1]. Therefore, MAC-layer throughput must scale up accordingly. Furthermore, MAC designs need to exploit the multirate capability offered by the physical layer for leveraging channel dynamics and multiuser diversity.

User population scalability

Another important consideration is to scale to the number of contending stations. The user population may range from a few in an office, to tens or hundreds in a classroom or a conference room, and thousands in public places like Disney Theme Parks [12]. As the number of active users grows, MAC designs should control contentions and collisions over the shared wireless channel and deliver stable performance.

Protocol overhead scalability

The third aspect in scalable wireless MAC design is to minimize the protocol overhead as the population size and the physical-layer capacity increase. Specifically, the fraction of channel time consumed by signaling messages per packet, due to backoff, interframe spacings, and handshakes, must remain relatively small.

2.2. Limitations of current MAC solutions

In general, both CSMA/CA [3] and polling-based MAC solutions have scalability limitations in these three aspects.

---

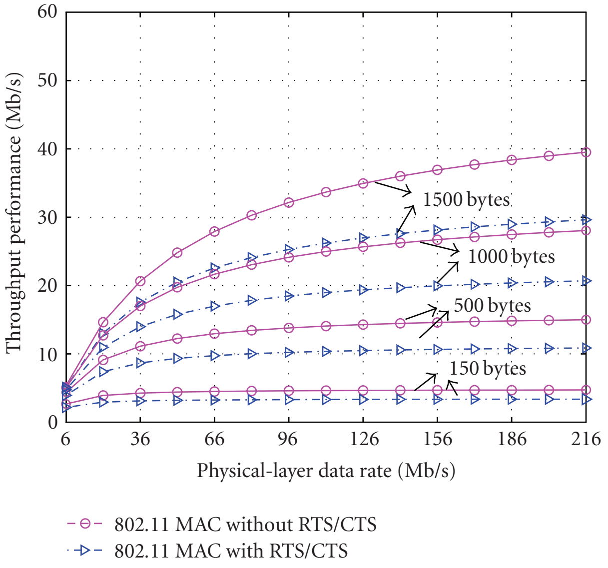
[1] OAR proposed to conduct multiple back-to-back transmissions upon winning the channel access for achieving temporal fair share among contending nodes.

2.2.1. CSMA/CA-based MAC

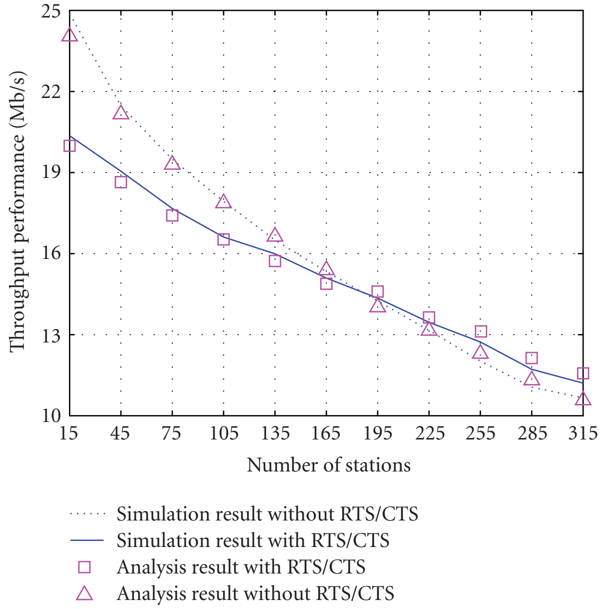
Our analysis and simulations show that DCF MAC, based on CSMA/CA mechanism, does not scale to high physical-layer capacity or various user populations. We plot the theoretical throughput attained by DCF MAC with different packet sizes in Figure 1(a).[2] Note that DCF MAC delivers at most 40 Mb/s throughput without RTS/CTS at 216 Mb/s, which further degrades to 30 Mb/s when the RTS/CTS option is on. Such unscalable performance is due to two factors. First, as the link capacity increases, the signaling overhead ratio grows disproportionately since the time of transmitting data packets reduces considerably. Second, the current MAC adopts a static channel sharing model that only considers transmission demands of stations. The channel is monopolized by low-rate stations. Hence the network throughout is largely reduced. Figure 1(b) shows results from both analysis[3] and simulation experiments conducted in *ns*-2. The users transmit UDP payloads at 54 Mb/s. The network throughput obtained with DCF reduces by approximately 50% as the user population reaches 300. The significant throughput degradation is mainly caused by dramatically intensified collisions and increasingly enlarged contention window (CW).

2.2.2. Polling-based MAC

Polling-based MAC schemes [3, 7, 14] generally do not possess capacity and protocol overhead scalability due to the excessive polling overhead. To illustrate the percentage of overhead, we analyze the polling mode (PCF) in 802.11b. In PCF, AP sends the polling packet to initiate the data transmission from wireless stations. A station can only transmit after receiving the polling packet. Idle stations respond to the polling message with NULL frame, which is a data frame without any payload. Table 1 lists the protocol overhead as the fraction of idle stations increases.[4] The overhead ratio reaches 52.1% even when all stations are active at the physical-layer rate of 54 Mb/s, and continue to grow considerably as more idle stations present. Furthermore, as the link capacity increases to 216 Mb/s, over 80% of channel time is spent on signaling messages.



(a) Throughput at different physical-layer data rates



(b) Network throughout at various user populations

FIGURE 1: Legacy MAC throughput at different user populations and physical-layer data rates.

3. TMAC DESIGN

In this section, we present the two-tier design of TMAC framework, which incorporates centralized, coarse-grained regulation at the higher tier and distributed, fine-grained channel access at the lower tier. Token-coordinated channel regulation provides coarse-grained coordination for bounding the number of contending stations at any time. It effectively controls the contention intensity and scales to various population sizes. Adaptive distributed channel access at the lower tier exploits the wide range of high data rates via the adaptive service model. It opportunistically favors stations under better channel conditions, while ensuring each station an adjustable fraction of the channel time based upon the perceived channel quality. These two components work together to address the scalability issues.

---

[2] Table 4 lists the values of DIFS, SIFS, ACK, MAC header, physical-layer preamble and header according to the specifications in [2, 4].

[3] We employ analytical model proposed in [13] to compute throughput, which matches the simulation results.

[4] The details of the analysis are listed in the technique report [15]. We computed the results using the parameter listed in Table 4.

Table 1: Polling overhead versus percentage of idle stations.

	0	15%	30%	45%	60%
54 Mb/s	52.1%	55.2%	59.1%	64%	70.3%
216 Mb/s	81.6%	83.2%	85.5%	87.3%	90.4%

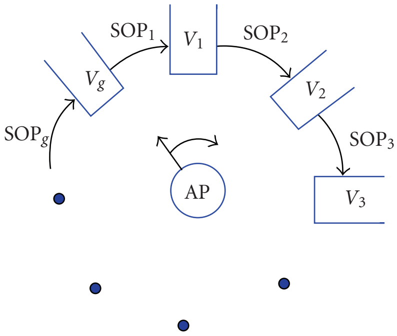


Figure 2: Token distribution model in TMAC.

3.1. Token-coordinated channel regulation

TMAC employs a simple token mechanism in regulating channel access at the coarse-time scale (e.g., in the order of 30 ∼ 100 milliseconds). The goal is to significantly reduce the intensity of channel contention incurred by a large population of active stations. The base design of the token mechanism is motivated by the observation that polling-based MAC works more efficiently under heavy network load [7, 16], while random contention algorithms better serve bursty data traffic under low load conditions [13, 17]. The higher-tier design, therefore, applies a polling model to multiplex traffic loads of stations within the token group.

Figure 2 schematically illustrates the token mechanism in TMAC. An AP maintains a set of associated stations, $S = \{s_1, s_2, \dots, s_n\}$, and organizes them into g number of disjoint token groups, denoted as $V_1, V_2, \dots, V_g$. Apparently, $\bigcup_{i=1}^{g} V_i = S$, and $V_j \cap V_j = \varnothing$ $(1 \le i, j \le g$ and $i \neq j)$. Each token group, assigned a unique Token Group ID (TGID), accommodates a small number of stations, N_{V_i}, and $N_{V_i} \le \overline{N}_V$, where $\overline{N}_V$ is a predefined upperbound. The AP regularly distributes a token to an eligible group, within which the stations contend for channel access via the enhanced random channel procedure in the lower tier. The period during which a given token group V_k obtains service is called *token service period*, denoted by TSP_k, and the transition period between two consecutive token groups is the *switch-over period*. The token service time for a token group V_k is derived using: $\mathrm{TSP}_k = (N_{V_k}/\overline{N}_V)\overline{\mathrm{TSP}}$, $(1 \le k \le g)$, where $\overline{\mathrm{TSP}}$ represents the maximum token service time. Upon the timeouts of TSP_k, the AP grants channel access to the next token group V_{k+1}.

To switch between token groups, the higher-tier design constructs a token distribution packet (TDP), and broadcasts it to all stations. The format of TDP, shown in Figure 3, is compliant with the management frame defined in 802.11b.

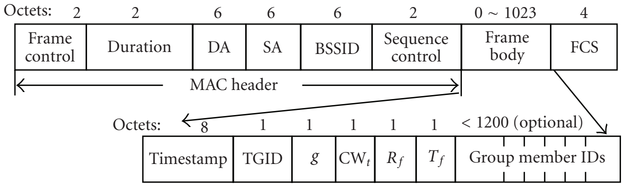


Figure 3: Frame format of token distribution packet.

In each TDP, a timestamp is incorporated for time synchronization, g denotes the total number of token groups, and the token is allocated to the token group specified by the TGID field. Within the token group, contending stations use CW_t in random backoff. The R_f and T_f fields provide two design parameters employed by the lower tier. The optional field of group member IDs is used to perform membership management of token groups, which can be MAC addresses, or dynamic addresses [18] in order to reduce the addressing overhead. The length of TDP ranges from 40 to 60 bytes ($\overline{N}_V = 20$, each ID uses 1 byte), taking less than 100 microseconds at 6 Mb/s rate. To reduce the token loss, TDP is typically transmitted at the lowest rate.

We need to address three concrete issues to make the above token operations work in practice, including membership management of token groups, policy of scheduling the access group, and handling transient conditions (e.g., when TDP is lost).

3.1.1. Membership management of token groups

When a station joins the network, TMAC assigns it to an eligible group, then piggybacks TGID of the token group in the association response packet [3], along with a local ID [18] generated for the station. The station records the TGID and the local ID received from the AP. Once a station sends a deassociation message, the AP simply deletes the station from its token group. The groups are reorganized if necessary. For performing membership management, the AP generates a TDP carrying the optional field that lists IDs of current members in the token group. Upon receiving the TDP with the ID field, each station with a matched TGID purges its local TGID. The station, whose ID appears in the ID field, extracts the TGID value from the TDP and updates its local TGID.

The specific management functions are described in the pseudocode listed in Algorithm 1. Note that we evenly split a randomly chosen token group if all the groups contain $\overline{N}_V$ stations, and merge two token groups if necessary. In this way, we keep the size of token group above $\overline{N}_V/4$ to maximize the benefits from traffic load multiplexing. Other optimizations can be further incorporated into the management functions. At present, we keep the current algorithm for simplicity.

```
Function 1: On station s joining the network
if $g == 0$ then
    create the token group $V_1$ with $\mathrm{TGID}_1$
    $V_1 = s$, set the update bit of $V_1$
else
    search for $V_i$, s.t., $N_{V_i} < \overline{N}_v$,
    if $V_i$ exists then
        $V_i = V_i \cup s$, set the update bit of $V_i$
    else
        randomly select a token group $V_i$
        Split $V_i$ evenly into two token groups, $V_i$, $V_{g+1}$
        $V_i = V_i \cup s$
        set the update bit of $V_i$ and $V_{g+1}$, $g = g + 1$
    end if
end if
Function 2: On station s, $s \in V_i$, leaving the network
$V_i = V_i - s$
if $N_{V_i} == 0$ then
    delete $V_i$, reclaim $\mathrm{TGID}_i$, $g = g - 1$
end if
if $N_{V_i} < \overline{N}_v/4$ then
    search for $V_j$, s.t., $N_{V_j} < \overline{N}_v/2$,
    if $V_j$ exists then
        $V_j = V_j \cup V_i$
        delete $V_i$, reclaim $\mathrm{TGID}_i$
        set the update bit of $V_j$, $g = g - 1$
    end if
end if
```

Algorithm 1: Group membership management functions.

3.1.2. Scheduling token groups

Scheduling token groups deal with the issues of setting the duration of $\overline{\mathrm{TSP}}$ and the sequence of the token distribution.

The $\overline{\mathrm{TSP}}$ is chosen to strike a balance between the system throughput and the delay. In principle, the size of the $\overline{\mathrm{TSP}}$ should allow for every station in a token group to transmit once for a period of its temporal share T_i. T_i is defined in the lower-tier design and typically in the order of several milliseconds. The network throughput performance improves when T_i increases [19]. However, increasing T_i enlarges the token circulation period, $g * \overline{\mathrm{TSP}}$, thus affecting the delay performance. Consequently, $\overline{\mathrm{TSP}}$ is a tunable parameter in practice, depending on the actual requirements of throughput/delay. The simulation results of Section 6 provide more insights of selecting a proper $\overline{\mathrm{TSP}}$.

To determine the scheduling sequence of token groups, TMAC uses a simple round-robin scheduler to cyclicly distribute the token among groups. It treats all the token groups with identical priority.

3.1.3. Handling transient conditions

Transient conditions include the variation in the number of active stations, loss of token messages, and stations with abnormal behaviors.

The number of active stations at an AP may fluctuate significantly due to bursty traffic load, roaming, and power-saving schemes [16, 20]. TMAC exploits a token-based scheme to limit the intensity of spatial contention and collisions. However, potential channel wastage may be incurred due to underutilization of the allocated TSP when the number of active stations sharply changes. TMAC takes a simple approach to adjust the TSP boundary. The AP announces the new TGID for the next group after deferring for a time period TIFS $= (\mathrm{DIFS} + m * \overline{\mathrm{CW}}_t * \sigma)$, where $\overline{\mathrm{CW}}_t$ is the largest CW in the current token group, m is the maximum backoff stage, and σ is the minislot time unit (i.e., 9 microseconds in 802.11a). The lower-tier operation in TMAC ensures that TIFS is the maximum possible backoff time. In addition, if a station stays in the idle status longer than the defined idle threshold, the AP assumes that it enters the power-saving mode, records it in the idle station list, and performs the corresponding management function for a leaving station. When new traffic arrives, the idle station executes the routine defined in the second transient condition to acquire a valid TGID, and then returns to the network.

Under the second transient condition, a station may lose its transmission opportunity in a recent token service period or fail to update its membership due to TDP loss. In this scenario, there are two cases. First, if the lost TDP message informs group splitting, the station belonging to the newly generated group, continues to join TSP matches its original TGID. The AP, upon detecting this behavior, unicasts the station with the valid TGID to notify its new membership. Second, if the lost TDP message announces group merging, the merged stations may not be able to contend for the channel without the recently assigned TGID. To retrieve the valid TGID, each merged station sends out reassociation/reauthentication messages after timeouts of $g * \overline{\mathrm{TSP}}$.

We next consider the station with abnormal behaviors, that is, the station transmits during the TSP that it does not belong to. Upon detecting the abnormal activities, the AP first reassigns it to a token group if the station is in the idle station list. Next, a valid TGID is sent to the station to compensate the potentially missed TDP. If the station continues the behavior, the AP can exclude the station by transmitting it a deassociation message.

3.2. Adaptive distributed channel access

The lower-tier design addresses the issues of capacity scalability and protocol overhead scalability in high-speed wireless LANs with an adaptive service model (ASM). The proposed ASM largely reduces channel access overhead and offers differentiated services that can be *adaptively* tuned to leverage high rates of stations. The following three subsections describe the contention mechanism, the adaptive channel sharing model, and the implementation of the model.

3.2.1. Channel contention mechanism

Channel contention among stations within an eligible token group follows the carrier sensing and random backoff routines defined in DCF [3, 21] mechanism. Specifically, a station with pending packets defers for a DIFS interval upon sensing an idle channel. A random backoff value is then chosen from $(0, \overline{\mathrm{CW}}_t)$. Once the associated backoff timer expires, RTS/CTS handshake takes place, followed by DATA transmissions for a time duration specified by ASM. Each station is allowed to transmit *once* within a given token service period to ensure the validity of ASM among stations across token groups. Furthermore, assuming most of stations within the group are active, AP can estimate the optimal value of $\overline{\mathrm{CW}}_t$ based on the size of the token group, which will be carried in the CW_t field of TDP messages. $\overline{\mathrm{CW}}_t$ is derived based on the results of [13]:

$$\overline{\mathrm{CW}}_t = \frac{2}{\zeta\left(1 + p\Sigma_{i=0}^{m-1}(2p)^i\right)}, \tag{1}$$

where $p = 1-(1-\zeta)^{n-1}$ and the optimal transmission probability ζ can be explicitly computed using $\zeta = 1/(\overline{N}_V \cdot \sqrt{T_c^*/2})$, and $T_c^* = (\mathrm{RTS}+\mathrm{DIFS}+\delta)/\sigma$. m denotes the maximum backoff stage, which has marginal effect on system throughput with RTS/CTS turned on [13], and m is set to 2 in TMAC.

3.2.2. *Adaptive service model*

The adaptive sharing model adopted by TMAC extracts the multiuser diversity by granting the users under good channel condition proportionally longer transmission durations. In contrast, the state-of-the-art wireless MACs do not adjust the time share to the perceived channel quality, granting stations with either identical throughput share [3] or equal temporal share [5, 14, 22], under idealized conditions. Consequently, the overall network throughput is significantly reduced since these MAC schemes ignore the channel conditions when specifying the channel sharing model. ASM works as follows. The truncated function (2) is exploited to define the service time T_{ASM} for station i, which transmits at the rate of r_i upon winning the channel contention:

$$T_{\mathrm{ASM}}(r_i) = \begin{cases} \left(\dfrac{r_i}{R_f}\right) T_f & r_i \geq R_f, \\ T_f & r_i < R_f. \end{cases} \tag{2}$$

The model differentiates these two classes of stations, high-rate and low-rate stations, by defining the reference parameters, namely, the reference transmission rate R_f and the reference time duration T_f. Stations with transmission rates higher than or equal to R_f are categorized as high-rate stations, thus granted *proportional temporal share* in that the access time is roughly proportional to the current data rate. For low-rate stations, each of them is provided *equal temporal share* in terms of identical channel access time T_f. Thus, ASM awards high-rate stations with a proportional longer time share and provides low-rate stations equal channel shares. In addition, the current DCF and OAR MAC become the specific instantiations of ASM by tuning the reference parameters.

3.2.3. *Implementation via adaptive batch transmission and block ACK*

To realize ASM, AP regularly advertises the two reference parameters R_f and T_f within a TDP. Upon receiving TDP, stations in the matched token group extract the R_f and T_f parameters, and contend for the channel access. Once a station succeeds in contention, adaptive batch transmission allows for the station to transmit multiple concatenated packets for a period equal to the time share computed by ASM. The adaptive batch transmission can be implemented at either the MAC layer as proposed in OAR [5] or the physical layer as in MAD [8]. To further reduce protocol overhead at the MAC layer, we exploit the block ACK technique to acknowledge A_f number of back-to-back transmitted packets in a single Block-ACK message, instead of per-packet ACK in the 802.11 MAC. The reference parameter A_f is negotiated between two communicating stations within the received-based rate adaptation mechanism [23] by utilizing RTS/CTS handshake.

4. PERFORMANCE ANALYSIS

In this section, we analyze the scalable performance obtained by TMAC in high-speed wireless LANs, under various user populations. We first characterize the overall network throughput performance in TMAC, then analytically compare the gain achieved by ASM with existing schemes. Also, we provide analysis on the three key aspects of scalability in TMAC.

4.1. *Network throughput*

To derive the network throughput in TMAC, let us consider a generic network model where all n stations are randomly located in a service area Ω centered around AP, and stations in the token groups always have backlogged queues of packets at length L. Without loss of generality, we assume each token group accommodates N_V number of active stations, and there are total g groups. We ignore the token distribution overhead, which is negligible compared to the TSP duration. Thus, the expected throughput S_{TMAC} can be derived based on the results from [13, 24],

$$\begin{aligned} S_{\mathrm{TMAC}} &= \frac{P_{tr}P_sE[P]}{(1-P_{tr})\sigma + P_{tr}P_sT_s + P_{tr}(1-P_s)T_c}, \\ P_{tr} &= 1-(1-\zeta)^{N_V}, \\ P_s &= \frac{N_V\zeta(1-\zeta)^{N_V-1}}{1-(1-\zeta)^{N_V}}. \end{aligned} \tag{3}$$

$E[P]$ is the expected payload size; T_c is the average time the channel is sensed busy by stations due to collisions; T_s denotes the duration of busy channel in successful transmissions. σ is the slot time and ζ represents the transmission probability at each station in the steady status. The value of ζ can be approximated by $2/(\mathrm{CW}+1)$ [24], where CW is the contention window chosen by the AP. Suppose that the physical layer offers M options of the data rates as $r_1, r_2, \ldots, r_M$, and $P(r_i)$ is the probability that a node transmits at rate r_i. When TMAC adopts the adaptive batch transmission at the

Table 2: Comparison of TMAC, DCF, and OAR.

	Analysis			Simulation	
	S (Mb/s)	T_s (μs)	$E[P]$ (bits)	S (Mb/s)	S_f (Mb/s)
DCF MAC	18.41	404.90	8192	18.79	20.24
OAR MAC	31.50	781.24	20760	32.11	26.52
$\text{TMAC}_{R_f=108}$	38.46	2119.42	83039	38.92	39.31
$\text{TMAC}_{R_f=54}$	41.64	1763.27	75093	42.13	42.59
$\text{TMAC}_{R_f=24}$	46.31	1341.61	62587	46.85	47.37

MAC layer, the values of $E[P]$, T_c, and T_s are expressed as follows:

$$\begin{gathered} E[P]=\sum_{i=1}^{M}P(r_m)\cdot L\cdot\frac{T_{\text{ASM}}(r_i)}{T^{\text{EX}}(r_i)},\\ T_c=T_{\text{DIFS}}+T_{\text{RTS}}+\delta,\qquad(4)\\ T_s=T_c+T_{\text{CTS}}+\sum_{i=1}^{M}P(r_i)T_{\text{ASM}}(r_i)+T_{\text{SIFS}}+2\delta. \end{gathered}$$

$T^{\text{EX}}(r_i)$ is the time duration of the data packet exchange at rate r_i, specified by $T^{\text{EX}}(r_i)=T_{\text{PH}}+T_{\text{MH}}+L/r_i+2\cdot T_{\text{SIFS}}+T_{\text{ACK}}$, with T_{PH}, T_{MH} being the overhead of physical-layer header and MAC-layer header, respectively. δ is the propagation delay.

Next, based on the above derivations and results in [5, 13], we compare the network throughput obtained with TMAC, DCF, OAR. The parameters used to generate the numerical results are chosen as follows: n is 15; g is 1, and L is 1 K; T_f is set to 2 milliseconds; the series of possible rates are 24, 36, 54, 108, and 216 in Mb/s, among which a station uses each rate with equal probability; other parameters are listed in Table 4. The results from numerical analysis and simulation experiments are shown in Table 2 as the R_f parameter in ASM of TMAC varies. Note that TMAC, with R_f set to 108 Mb/s, improves the transmission efficiency, measured with $S_f = E[P]/T_s$, by 22% over OAR. On further reducing R_f, the high-rate stations are granted with the proportional higher temporal share. Therefore, TMAC with $R_f = 24$ Mb/s achieves 48% improvement in network throughput over OAR, and 84% over DCF. Such throughput improvements demonstrate the effectiveness of ASM by leveraging high data rates perceived by multiple stations.

4.2. *Adaptive channel sharing*

Here, we analyze the expected throughput of ASM, exploited in the lower tier of TMAC, as compared with those of the equal temporal share model proposed in OAR [5] and of the equal throughput model adopted in DCF [3].

Let ϕ_i^{ASM}, ϕ_i^{OAR} be the fractions of time that station i transmits at rate r_i in a time duration T using the scheme of ASM and OAR, respectively, where $0 \leq \phi_i \leq 1$. During the interval T, n denotes the number of stations in the equal temporal sharing policy, and n' is the number of stations transmitting within the adaptive service model, clearly $n' \geq n$. Then, we have the following equality:

$$\sum_{i=1}^{n}\phi_i^{\text{OAR}}=\sum_{i=1}^{n'}\phi_i^{\text{ASM}}=1. \qquad (5)$$

Therefore, the expected throughput achieved in ASM is given by $S_{\text{ASM}}=\sum_{i=1}^{n'}r_i\phi_i^{\text{ASM}}$. We obtain the following result, using the above notations.

Proposition 1. *S_{ASM}, S_{OAR}, and S_{DCF} are the total expected throughput attained by* ASM, OAR, *and* DCF, *respectively. One has*

$$S_{\text{ASM}} \geq S_{\text{OAR}} \geq S_{\text{OAR}}. \qquad (6)$$

Proof. From the concept of equal temporal share, we have $\phi_i^{\text{OAR}} = \phi_j^{\text{OAR}}$, $(1 \leq i, j \leq n)$. The expected throughput in equal temporal share is derived as

$$S_{\text{OAR}}=\sum_{i=1}^{n}r_i\phi_i^{\text{OAR}}=\frac{1}{n}*\sum_{i=1}^{n}r_i. \qquad (7)$$

Thus, by relations (5) and Chebyshev's sum inequality, we can have the following result:

$$S_{\text{OAR}}\leq\frac{1}{n}\sum_{i=1}^{n}\phi_i^{\text{ASM}}\sum_{i=1}^{n}r_i\leq\sum_{i=1}^{n}\phi_i^{\text{ASM}}r_i\leq S_{\text{ASM}}. \qquad (8)$$

Similarly, we can show that $S_{\text{DCF}} \leq S_{\text{OAR}}$. □

4.3. *Performance scalability*

We analytically study the scalability properties achieved by TMAC, while we show that the legacy solutions do not possess such appealing features.

4.3.1. *Scaling to user population*

It is easy to show that TMAC scales to the user populations. From the throughput characterization of (3), we observe that the throughput of TMAC is only dependent on the token group size N_V, instead of the total number of users n. Therefore, the network throughput in TMAC scales with respect to the total number of stations n.

To demonstrate the scalability constraints of the legacy MAC, we examine the DCF with RTS/CTS handshakes. Note that DCF can be viewed as a special case of TMAC, in which all n stations stay in the same group, thus $N_V = n$. We measure two variables of ζ and T_w. ζ is the transmission probability of a station at a randomly chosen time slot and can be approximated by $2/(CW + 1)$. T_W denotes the time wasted on the channel due to collisions per successful packet transmission, and can be computed by

$$T_W=\left(T_{\text{DIFS}}+T_{\text{RTS}}+\delta\right)\left(\frac{1-(1-\zeta)^n}{n\zeta(1-\zeta)^{n-1}}-1\right), \qquad (9)$$

where δ denotes the propagation delay.

Table 3: Analysis results for ζ and T_W in DCF.

n	15	45	105	150	210	300
ζ	0.0316	0.0177	0.0110	0.0090	0.0075	0.0063
T_W (μs)	21.80	43.24	72.78	92.75	119.61	163.34

Table 4: PHY/MAC parameters used in the simulations.

SIFS	16 μs	DIFS	34 μs
Slot time	9 μs	PIFS	25 μs
ACK size	14 bytes	MAC header	34 bytes
Peak datarate (11a)	54 Mb/s	Basic datarate (11a)	6 Mb/s
Peak datarate (11n)	216 Mb/s	Basic datarate (11n)	24 Mb/s
PLCP preamble	16 μs	PLCP header length	24 bytes

As the number of stations increases, the values of ζ and T_W in the DCF are listed in Table 3 and the network throughput is shown in Figure 1(b). Although ζ decreases as the user size expands because of the enlarged CW in exponential backoff, the channel time wasted in collisions, measured by T_W, increases almost linearly with n. The considerable wastage of channel time on collisions leads to approximately 50% network throughput degradation as the user size reaches 300, as shown by simulations.

4.3.2. Scaling of protocol overhead and physical-layer capacity

Within a token group, we examine the protocol overhead at the lower tier as compared to DCF. At a given data rate r, the protocol overhead T_o denotes the time duration of executing the protocol procedures in successfully transmitting a $E[P]$-bytes packet, which is given by

$$\begin{aligned} T_o^{\text{DCF}} &= T_o^p + T_{\text{idle}} + T_{\text{col}}, \\ T_o^{\text{ASM}} &= \frac{T_o^{\text{DCF}}}{B_f} + T_o^{\text{EX}}. \end{aligned} \tag{10}$$

T_{idle} and T_{col} represent the amount of idle time and the time wasted on collisions for each successful packet transmission, respectively. T_o^p specifies in DCF the protocol overhead spent on every packet, which is equal to $(T_{\text{RTS}} + T_{\text{CTS}} + T_{\text{DIFS}} + 3T_{\text{SIFS}} + T_{\text{ACK}} + T_{\text{PH}} + T_{\text{MH}})$. T_o^{EX} denotes the per-packet overhead of the adaptive batch transmission in ASM, which is calculated by $(2T_{\text{SIFS}} + T_{\text{ACK}} + T_{\text{PH}} + T_{\text{MH}})$. B_f is the number of packets transmitted in T_{ASM} interval and $B_f = T_{\text{ASM}}/T^{\text{EX}}$. From (10), we note that the protocol overhead in ASM is reduced by the factor of B_f as compared with DCF, and B_f is a monotonically increasing function of data rate r. Therefore, TMAC effectively controls its protocol overhead and scales to the channel capacity increase, while DCF suffers from fixed per-packet overhead, throttling the scalability of its network throughput. Moreover, T_o^{EX} is the fixed overhead in TMAC, incurred by physical-layer preambles, interframe spacings, and protocol headers. It is the major constraint to further improve the throughput in the MAC layer.

4.3.3. Scaling to physical-layer capacity

To demonstrate the scalability achieved by TMAC with respect to the channel capacity R, we rewrite the network throughput as the function of R, and obtain

$$\begin{aligned} S_{\text{DCF}} &= \frac{L}{R \cdot T_o^{\text{DCF}} + L} \cdot R, \\ S_{\text{TMAC}} &= \frac{L}{\left(T_o^{\text{DCF}}/T_{\text{ASM}} + 1\right)\left(R \cdot T_o^{\text{EX}} + L\right)} \cdot R. \end{aligned} \tag{11}$$

Note that T_{ASM} is typically chosen in the order of several milliseconds, thus having $T_{\text{ASM}} \gg T_o^{\text{DCF}}$. Now, the limiting factor of network throughput is $L/(R \cdot T_o^{\text{DCF}})$ in DCF, and $L/(R \cdot T_o^{\text{EX}})$ in ASM. Since $T_o^{\text{EX}} \ll T_o^{\text{DCF}}$ and T_o^{EX} is in the order of hundreds of microseconds (e.g., $T_o^{\text{EX}} = 136$ microseconds in 802.11a/n), ASM achieves much better scalability as R increases, while the throughput obtained in DCF is restrained by the increasingly enlarged overhead ratio. In addition, the study shows transmitting packets at larger size L can greatly improve network throughput. Therefore, the technique of packet aggregation at the MAC layer and payload concatenation at the physical layer is promising in next-generation high-speed wireless LANs.

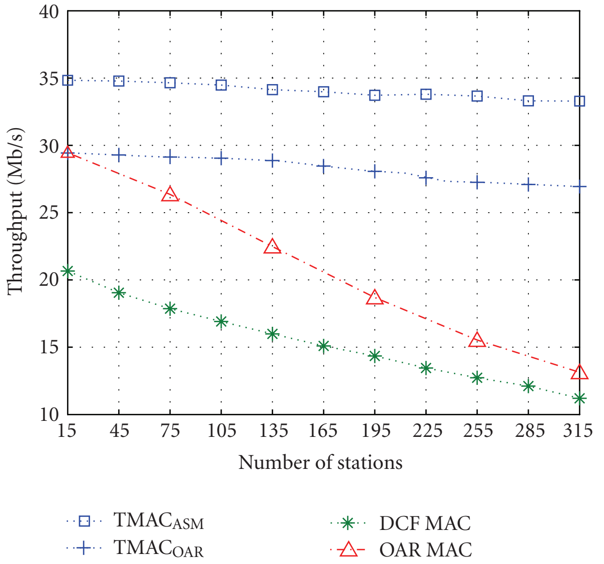
5. SIMULATION

We conduct extensive simulation experiments to evaluate scalability performance, channel efficiency, and sharing features achieved by TMAC in wireless LANs. Five environment parameters are varied in the simulations to study TMAC's performance, including user population, physical-layer rate, traffic type, channel fading model, and fluctuations in the number of action stations. Two design parameters, T_f and A_f, are investigated to quantify their effects (R_f has been examined in the previous section). We also plot the performance of the legacy MACs, 802.11 DCF and OAR, in demonstrating their scaling constraints. We use TMAC_{DCF} and TMAC_{OAR} to denote TMAC employing DCF or OAR in the lower tier, which are both specific cases of TMAC.

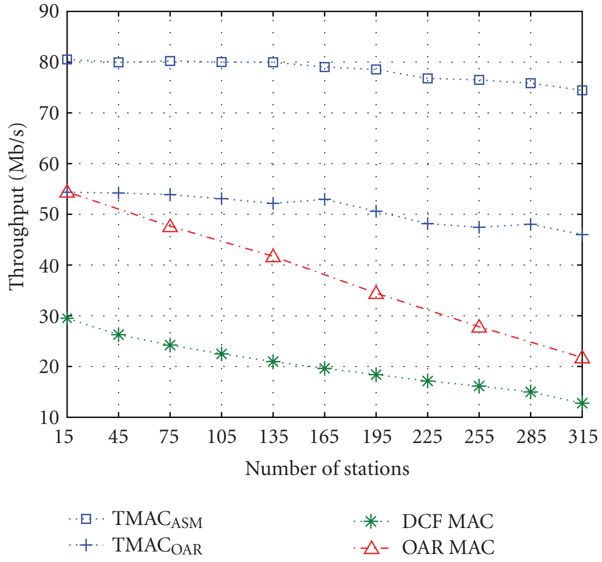
The simulation experiments are conducted in *ns*-2 with the extensions of Ricean channel fading model [25] and the receive-based rate adaptation mechanism [23]. Table 4 lists the parameters used in the simulations based on IEEE 802.11b/a [3, 4] and the leading proposal for 802.11n [2]. The transmission power and radio sensitivities of various data rates are configured according to the manufacturer specifications [26] and 802.11n proposal [2]. The following parameters are used, unless explicitly specified. Each token group has 15 stations. T_f allows 2 milliseconds batch transmissions at MAC layer. Each block ACK is sent for every two packets (i.e., $A_f = 2$). Any packet loss triggers retransmission of two packets. Token is announced approximately every 35 milliseconds to regulate channel access. Each station generates constant-bit-rate traffic, with the packet size set to 1 Kb.

5.1. Scaling to user population

We first examine the scalability of TMAC in aspects of network throughput and average delay as population size varies.



(a) Network throughput at 54 Mb/s link capacity



(b) Network throughput at 216 Mb/s link capacity

FIGURE 4: Network throughput versus the number of stations.

5.1.1. *Network throughput*

Figure 4 shows that both TMAC$_{\mathrm{ASM}}$ and TMAC$_{\mathrm{OAR}}$ achieve scalable throughput, experiencing less than 6% throughput degradation, as the population size varies from 15 to 315. In contrast, the network throughput obtained with DCF and OAR does not scale: the throughput of DCF decreases by 45.9% and 56.7% at the rates of 54 Mb/s and 216 Mb/s, respectively, and the throughput in OAR degrades 52.3% and 60%, in the same cases. The scalable performance achieved in TMAC demonstrates the effectiveness of the token mechanism in controlling the contention intensity as user population expands. Moveover, TMAC$_{\mathrm{ASM}}$ consistently outperforms TMAC$_{\mathrm{OAR}}$ by 21% at 54 Mb/s data rate, and 42.8% at 216 Mb/s data rate, which reveals the advantage of ASM in supporting high-speed physical layer.

TABLE 5: Average delay (s) at 216 Mb/s.

Num.	15	45	75	135	165	225	285
DCF MAC	0.165	0.570	0.927	1.961	3.435	4.539	5.710
TMAC$_{\mathrm{DCF}}$	0.163	0.822	1.039	1.654	2.400	2.590	2.870
TMAC$_{\mathrm{ASM}}$	0.053	0.169	0.359	0.620	0.760	0.829	1.037

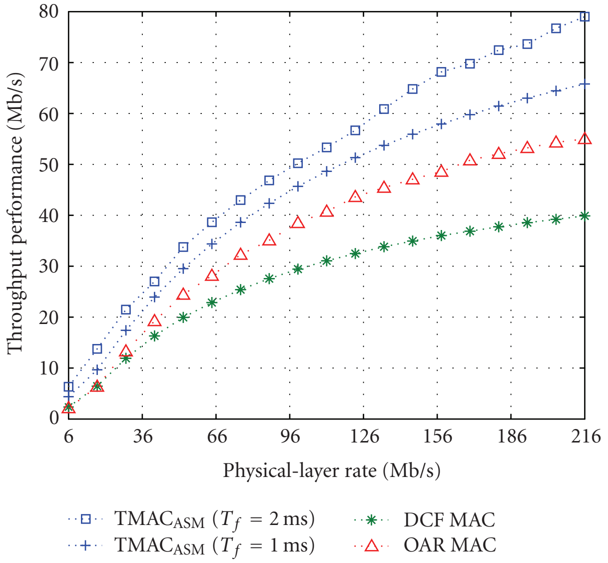


FIGURE 5: Network throughput versus physical-layer data rates.

5.1.2. *Average delay*

Table 5 lists the average delay of three protocols, DCF, TMAC$_{\mathrm{DCF}}$, and TMAC$_{\mathrm{ASM}}$ in the simulation scenario identical to the one used in Figure 4(b). The table shows that the average delay in TMAC increases much slower than that in DCF, as the user population grows. In specific, the average delay in DCF increases from 0.165 second to 5.71 seconds as the number of stations increases form 15 to 285. TMAC$_{\mathrm{DCF}}$, adopting token mechanism in the higher tier, reduces the average by up to 39%, while TMAC$_{\mathrm{ASM}}$ achieves approximately 70% average delay reduction over various population sizes. The results demonstrate that the token mechanism can efficiently allocate channel share among a large number of stations, thus reducing the average delay. Moveover, ASM improves channel efficiency and further decreases the average delay.

5.2. *Scaling to different physical-layer rates*

Within the scenario of 15 contending stations, Figure 5 depicts the network throughput obtained by DCF, OAR, and TMAC with the different settings in the lower tier, as the physical-layer rate varies from 6 Mb/s to 216 Mb/s. Note that

TMAC$_{\text{ASM}}$, with T_f set to 1 millisecond and 2 milliseconds, achieves up to 20% and 42% throughput improvement over OAR, respectively. This reveals that TMAC effectively can control protocol overhead at MAC layer especially within the high-capacity physical layer. Our study further reveals that the overhead incurred by the physical-layer preamble and header is the limiting factor for further improving the throughput achieved by TMAC.

5.3. *Interacting with TCP*

In this experiment, we examine the throughput scalability and the fair sharing feature in TMAC when stations, exploiting the rate of 54 Mb/s, carry out a large file transfer using TCP Reno. The sharing feature is measured by Jain's fairness index [27], which is defined as $(\sum_{i=1}^{n} x_i)^2/(n\sum_{i=1}^{n} x_i^2)$. For station i using the rate of r_i,

$$x_i = S_i * \frac{T_f}{(r_i * T_{\text{ASM}}(r_i))}, \qquad (12)$$

where S_i is the throughput of station i. Figure 6 plots the network throughput and labels the fairness index obtained with DCF, OAR, and TMAC$_{\text{ASM}}$ in various user sizes. TMAC demonstrates scalable performance working with TCP. Note that both OAR and DCF experience less than 10% throughput degradation in this case. However, as indicated by the fairness index, both protocols lead to severe unfairness in channel sharing among FTP flows as user size grows. Such unfairness occurs because in DCF and OAR, more than 50% of FTP flows experience service starvation during the simulation run, and 10% flows contribute to more than 90% of the network throughput, as the number of users grows over 75. On the other hand, TMAC, employing the token mechanism, preserves the fair sharing feature while attaining scalable throughput performance at various user sizes.

5.4. *Ricean fading channel*

We now vary channel fading model and study its effects on TMAC with the physical layer specified by 802.11a. Ricean fading channel is adopted in the experiment with $K = 2$, where K is the ratio between the deterministic signal power and the variance of the multipath factor [25]. Stations are distributed uniformly over 400 m × 400 m territory (AP is in the center) and move at the speed of 2.5 m/s. The parameter R_f is set at rate of 18 Mb/s. Figure 7 shows the network throughput of different MAC schemes. These results again demonstrate the scalable throughput achieved by TMAC$_{\text{ASM}}$ and TMAC$_{\text{OAR}}$ as the number of users grows. TMAC$_{\text{ASM}}$ consistently outperforms TMAC$_{\text{OAR}}$ by 32% by offering adaptive service share to stations in dynamic channel conditions. In contrast, OAR and DCF experience 72.7% and 68% throughput reduction, respectively, as the user population increases from 15 to 255.

5.5. *Active station variation and token losses*

We examine the effect of variations in the number of active stations caused and of token losses. During the 100-second simulation, 50% stations periodically enter 10-second sleep mode after 10-second transmission. Receiving errors are manually introduced, which causes loss of the token message in nearly 20% of active stations. The average of network throughput in TMAC and DCF is plotted in Figure 8 and the error bar shows the maximum and the minimum throughput observed in 10-second interval. When the user size increases from 15 to 255, DCF suffers from throughput reduction up to approximately 55%. It also experiences large variation in the short-term network throughput, indicated by the error bar. In contrast, TMAC achieves stable performance and scalability in the network throughput, despite the

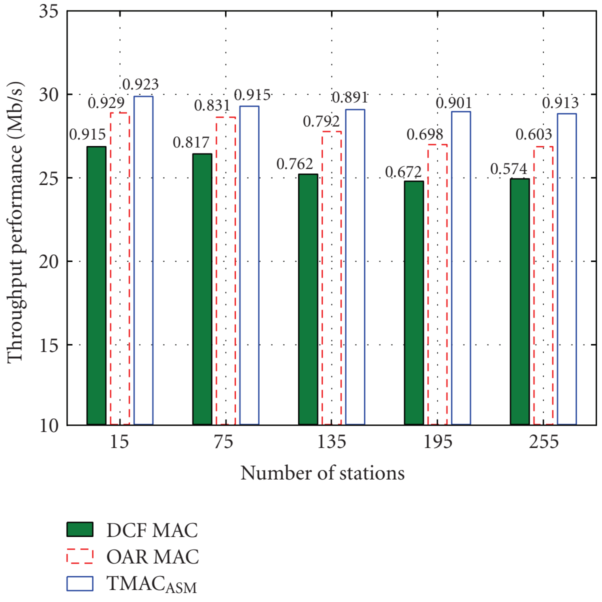


Figure 6: Network throughput in TCP experiments.

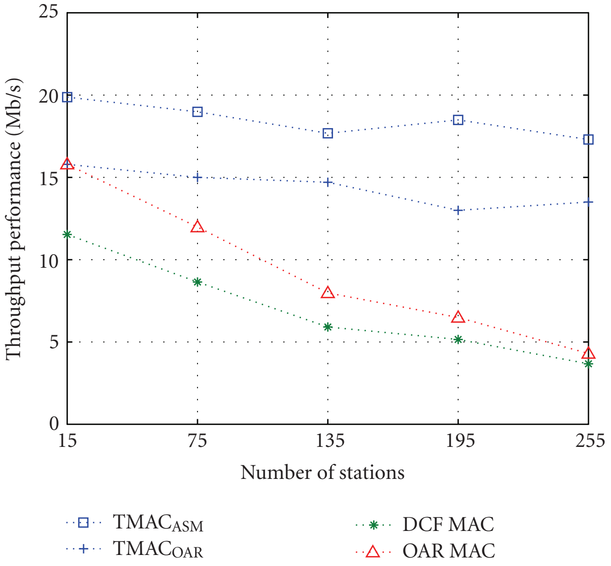


Figure 7: Network throughput in Ricean fading channel.

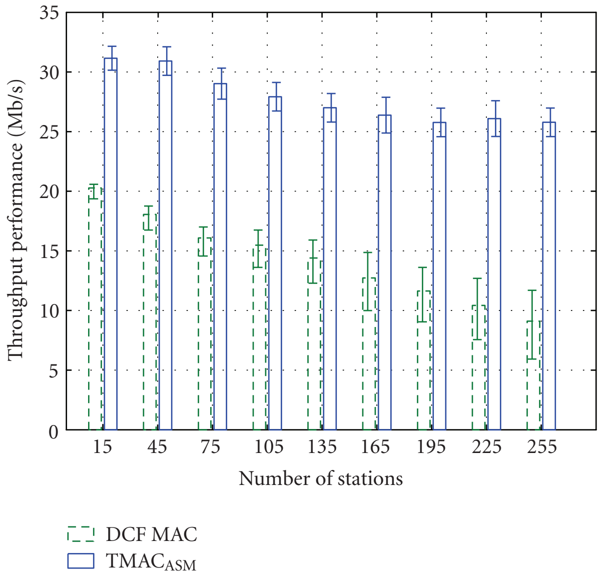


FIGURE 8: Network throughput versus the number of stations.

fact that the throughput degrade by up to 18% in the same case. Several factors that contribute to the throughput reduction in TMAC include the wastage of TSP, the overhead of membership management and the cost of token loss.

5.6. *Design parameters* A_f *and* T_f

We now evaluate the impacts of the design parameters T_f and A_f. We adopt scenarios similar to the case A, and fix the number of users as 50. The reference transmission duration T_f varies with A_f set to 1, where T_f of 0 millisecond grants one packet transmission as in the legacy MAC. Next, to quantify the effect of the block ACK size, we tune A_f from 1 to 6, with 3 milliseconds T_f.

Table 7 presents the network throughput obtained with TMAC as the design parameters of T_f and A_f vary. When T_f changes from 0 millisecond to 5 milliseconds, the aggregate throughput improves by 63.7% at 54 Mb/s data rate, and 127% with 216 Mb/s rate. Tuning the parameter A_f can further improve the throughput to more than 100 Mb/s. The improvements show that the overhead caused by per-packet contention and acknowledgement has been effectively reduced in TMAC.

5.7. *Exploiting rate diversity*

In the final set of experiments, we demonstrate that TMAC can adaptively leverage multirate capability at each station to further improve the aggregate throughput. We use the fairness index defined in Section 5.3. We consider the simulation setting of eight stations in one token group. Each station carries a UDP flow to fully saturate the network. There are four transmission rate options, 24 Mb/s, 54 Mb/s, 108 Mb/s, and 216 Mb/s. Each pair of stations randomly chooses one of the four rates. The results are obtained from averaging over 5 simulation runs.

Table 6 enumerates the aggregate throughput and the fairness index for flows transmitting at the same rate, using the 802.11 MAC, and TMAC with different R_f settings. TMAC enables high-rate stations to increasingly exploit their good channel conditions by granting the high-rate nodes more time share than the low-rate stations. This is realized by reducing a single parameter R_f. TMAC acquires 65%, 87%, 111% and 133% overall throughput gains compared with the legacy MAC as adjusting R_f to 216 Mb/s, 108 Mb/s, 54 Mb/s, and 24 Mb/s, respectively.

Moreover, The fairness index for TMAC design is close to 1 in every case, which indicates the effectiveness of the adaptive sharing scheme. The fairness index of DCF MAC is 0.624 in temporal units. DCF results in such a severe bias because it neglects the heterogeneity in channel quality experienced by stations and offers them equal throughput share. In summary, by lowering the access priority of low-rate stations that are nevertheless not in good channel conditions, TMAC provides more transmission opportunities for high-rate stations perceiving good channels. This feature is important for high-speed wireless LANs and mesh networks to mitigate the severe aggregate throughput degradation incurred by low-rate stations. The lower channel sharing portion by a low-rate station also motivates it to move to a better spot or to improve its reception quality. In either case, the system throughput is improved.

6. DISCUSSIONS

In this section, we first discuss alternative designs to address the scaling issues in high-speed wireless LANs. We then present a few issues relevant to TMAC. We will discuss the prior work related to TMAC in detail in Section 7.

TMAC employs a centralized solution to improve user experiences and provide three scaling properties, namely user population scaling, physical-layer capacity scaling, and protocol overhead scaling. The design parameters used in TMAC can be customized for various scenarios, which is especially useful for wireless Internet service providers (ISP) to improve the service quality. One alternative scheme for supporting large user sizes is to use the distributed method to tune CW. Such a method enables each node to estimate the perceived contention level and thereafter choose the suitable CW (e.g., AOB [28], Idle Sense [24]). Specifically, the slot utilization, or the number of idle time slots, is measured and serves as the input to derive CW in DCF MAC.

The distributed scheme for adjusting CW will have difficulty in providing scaling performance especially in high-speed wireless LANs. First, the distributed scheme derives CW by modeling the DCF MAC. The result cannot be readily applied to high-speed wireless networks. The MAC design in high-speed wireless networks, such as IEEE 802.11n,

Table 6: Throughput (Mb/s) and fairness index.

MAC type	802.11MAC	TMAC ($R_f = 216$ Mb/s)	TMAC ($R_f = 108$ Mb/s)	TMAC ($R_f = 54$ Mb/s)	TMAC ($R_f = 24$ Mb/s)
24 Mb/s flows	6.649	4.251	1.922	1.198	0.910
54 Mb/s flows	6.544	8.572	11.282	5.004	4.695
108 Mb/s flows	6.655	12.660	15.489	20.933	10.649
216 M/bs flows	6.542	17.795	20.986	28.811	45.136
All flows	26.490	43.278	49.679	55.946	61.390
Fairness index	0.6246	0.9297	0.9341	0.9692	0.9372

Table 7: Network throughput (Mbits/s) versus T_f and A_f.

T_f	0 ms	1 ms	2 ms	3 ms	4 ms	5 ms
54 Mb/s	20.40	25.33	28.91	32.10	32.93	33.40
216 Mb/s	35.16	70.70	76.19	78.35	79.31	79.88
A_f	1	2	3	4	5	6
216 Mb/s	78.35	93.92	95.91	97.29	98.94	101.72

largely adopts the existing schemes proposed in IEEE 802.11e [14]. In 802.11e, several access categories are defined to offer differentiated services in supporting various applications. Each access category uses different settings of the deferring time period, CW, and the transmission duration. The new MAC protocol inevitably poses challenges to the distributed schemes based on modeling DCF, the simpler version of 802.11e. Second, the distributed scheme mainly considers tuning CW to match the contention intensity. The scaling issues of protocol overhead and physical-layer capacity are not explicitly addressed. Moreover, the distributed scheme requires each node to constantly measure contention condition for adjusting CW. The design incurs extra management complexity at APs due to lack of the control of user behaviors.

The problem with the distributed scheme of CW tuning may be solvable in high-speed wireless LANs, but it is clear that a straightforward approach using the centralized control prevents several difficulties. TMAC can support access categories by announcing the corresponding parameters, such as CW, in the token messages. More sophisticated schedulers (e.g., weighted round robin) can be adopted to arrange the token groups in order to meet the quality-of-service (QoS) requirements of various applications. The adaptive service model enables packet aggregation and differentiates the time share allocated to the high-rate and low-rate stations to leverage data-rate diversities. In addition, most computation complexity in TMAC occurs at APs, while client devices only require minor changes to handle received tokens. The design parameters offers wireless ISPs extra flexibility to control system performance and fairness model. The two-tier design adopted by TMAC extracts benefits of the random access and the polling mechanism, hence provides a highly adaptable solution for the next generation, high-speed wireless LANs.

We now discuss several issues relevant to the TMAC design.

(a) Backward compatibility

TMAC so far mainly focuses on operating in the infrastructure mode. Since the fine-grained channel access is still based on CSMA/CA, TMAC can coexist with stations using the current 802.11 MAC. AP still uses the token distribution and reference parameter set to coordinate channel access among stations supporting TMAC. However, the overall MAC performance will degrade as a larger number of regular stations contend for channel access.

(b) Handling misbehaving stations

Misbehaving stations expose them by acquiring more channel share than its fair share during its batch transmission or contending for channel access when it does not possess the current TGID. We can mitigate these misbehaving stations by monitoring and policing them via central AP. Specifically, the AP can keep track the channel time each station received, and calculates its fair share based on the collected information of the station transmission rate and other reference parameter settings. When the AP detects an overly aggressive station, say, access time beyond certain threshold, it temporarily revokes channel access right of the station. This can be realized via the reauthentication mechanism provided by the current 802.11 MAC management plane.

(c) Power saving

TMAC supports power saving and also works with the power saving mechanism (PSM) defined in 802.11. In TMAC, time is divided into token service periods, and every node in the network is synchronized by periodic token transmissions. So every node will wake up at beginning each token service period at about the same time to receive token messages. The node that does not belong to the current token group can save energy by going into doze mode. In doze mode, a node consumes much less energy compared to normal mode, but cannot send or receive packets. Within the token service period, PSM can be applied to allow a node to enter the doze mode only when there is no need for exchanging data in the prevailing token period.

7. RELATED WORK

A number of well-known contention-based channel access schemes have been proposed in literature, starting from the early ALOHA and slotted ALOHA protocols [6], to the more recent 802.11 DCF [3], MACA [21], MACAW [11]. These proposals, however, all face the fundamental problem that their throughput drops to almost zero as the channel load increases beyond certain critical point [29]. This issue leads to the first theoretical study of network performance as the user population varies [29]. The study further stimulates the recent work [16, 17, 20, 24, 28, 30] on dynamically tuning the backoff procedure to reduce excessive collisions within large user populations. However, backoff tuning generally requires detailed knowledge of the network topology and traffic demand, which are not readily available in practice. TMAC differs from the above work in that it addresses the scalability issues in a two-tier framework. The framework incorporates a higher-tier channel regulation on top of the contention-based access method to gracefully allocate channel resource within different user populations. In the meantime, TMAC offers capacity and protocol overhead scalability through an adaptive sharing model. Collectively, TMAC controls the maximum intensity of resource contention and delivers scalable throughput for various user sizes with minimal overhead.

A number of enhanced schemes for DCF have been proposed to improve its throughput fairness model [3] in wireless LANs. Equal temporal share model [5, 31] and throughput propositional share model [22] generally grant each node the same share in terms of channel time to improve the network throughput. In 802.11e and 802.11n, access categories are introduced to provide applications different priorities in using the wireless medium. In TMAC, the existing models can be applied to the lower-tier design directly. To offer the flexibility of switching the service model, we exploit the adaptive service model, which allows administrators to adjust the time share for each station based on both user demands and the perceived channel quality. To further reduce the protocol overhead, TMAC renovates the block ACK technique proposed in 802.11e [14] by removing the tedious setup and tear-down procedures, and introduces an adjustable parameter for controlling the block size. More importantly, TMAC is designed for a different goal—it is to tackle the three scalability issues in next-generation wireless data networks.

Reservation-based channel access methods typically exploit the polling model [7, 10] and dynamic TDMA schemes in granting channel access right to each station. IBM Token Ring [32] adopts the polling model in the context of wired network by allowing a token to circulate around the ring network. Its counterpart in wireless network includes PCF [3] and its variants [14]. The solutions of HiperLAN/2 [9, 33] are based on dynamic TDMA and transmit packets within the reserved time slots. All these proposals use reservation-based mechanisms in fine-time-interval channel access for each individual station. In contrast, the polling model applied in TMAC achieves coarse-grained resource allocation for a group of stations to multiplex bursty traffic loads for efficient channel usage.

Some recent work has addressed certain aspect of the scalable MAC design. The work by [34] recognized the impact of scalability in MAC protocol design, but did not provide concrete solutions. Commercial products [12, 35] have appeared in the market that claimed scalable throughput in the presence of about 30 users for their 802.11b APs. ADCA [19], our previous work, is proposed to reduce the protocol overhead as the physical-layer rate increases. The method of tuning CW based on idle slots [24] have been explored to manage channel resource and fairness for large user sizes. Multiple-channel [36] and cognitive radios [37] offer the promise of spectrum agility to increase the available resources by trading off the hardware complexity and cost. Inserting an overlay layer [38] or using multiple MAC layers [39, 40] has been exploited to increase network efficiency. However, an effective MAC framework that is able to tackle all three key scalability issues has not yet been adequately addressed.

8. CONCLUSION

Today wireless technologies are going through similar development and deployment cycles that wired Ethernet has been through in the past three decades—driving the speed to orders of magnitude higher, keeping low protocol overhead, and expanding deployment in more diversified environments. To cater to these trends, we propose a new scalable MAC solution within a novel two-tier framework, which employs coarse-time-scale regulation and fine-time-scale random access. The extensive analysis and simulations have confirmed scalability of TMAC. The higher-tier scheduler of TMAC that arbitrates token groups can be enhanced to provide sustained QoS for various delay- and loss-sensitive applications, which is our immediate future work.

REFERENCES

[1] IEEE 802.11n: Wireless LAN MAC and PHY Specifications: Enhancements for Higher Throughput, 2005.

[2] IEEE 802.11n: Sync Proposal Technical Specification, doc. IEEE 802.11-04/0889r6., May 2005.

[3] B. O'Hara and A. Petrick, *IEEE 802.11 Handbook: A Designer's Companion*, IEEE Press, Piscataway, NJ, USA, 1999.

[4] IEEE Std 802.11a-1999—part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY).

[5] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly, "Opportunistic media access for multirate ad hoc networks," in *Proceedings of the 8th Annual International Conference on Mobile Computing and Networking (MOBICOM '02)*, pp. 24–35, Atlanta, Ga, USA, September 2002.

[6] L. Kleinrock and F. A. Tobagi, "Packet switching in radio channels—part I: carrier sense multiple-access modes and their throughput-delay characteristics," *IEEE Transactions on Communications*, vol. 23, no. 12, pp. 1400–1416, 1975.

[7] F. A. Tobagi and L. Kleinrock, "Packet switching in radio channels—part III: polling and (dynamic) split-channel reservation multiple access," *IEEE Transactions on Communications*, vol. 24, no. 8, pp. 832–845, 1976.

[8] Z. Ji, Y. Yang, J. Zhou, M. Takai, and R. Bagrodia, "Exploiting medium access diversity in rate adaptive wireless LANs," in *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking (MOBICOM '04)*, pp. 345–359, Philadelphia, Pa, USA, September-October 2004.

[9] Hiperlan/2 EN 300 652 V1.2.1(1998-07), Function Specification, ETSI.

[10] H. Levy and M. Sidi, "Polling systems: applications, modeling, and optimization," *IEEE Transactions on Communications*, vol. 38, no. 10, pp. 1750–1760, 1990.

[11] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, "MACAW: a media access protocol for wireless LAN's," in *Proceedings of the Conference on Communications Architectures, Protocols and Applications (SIGCOMM '94)*, pp. 212–225, London, UK, August-September 1994.

[12] http://www.computerworld.com/mobiletopics/mobile/story/0,10801,65816,00.html.

[13] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, 2000.

[14] IEEE Std 802.11e/D8.0—part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY).

[15] W. Arbaugh and Y. Yuan, "Scalable and efficient MAC for next-generation wireless data networks," Tech. Rep., Computer Science Department, University of Maryland, College Park, Md, USA, 2005.

[16] Y. Kwon, Y. Fang, and H. Latchman, "A novel MAC protocol with fast collision resolution for wireless LANs," in *Proceedings of the 22nd Annual Joint Conference on the IEEE Computer and Communications Societies (INFOCOM '03)*, vol. 2, pp. 853–862, San Francisco, Calif, USA, March-April 2003.

[17] H. Kim and J. C. Hou, "Improving protocol capacity with model-based frame scheduling in IEEE 802.11-operated WLANs," in *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MOBICOM '03)*, pp. 190–204, San Diego, Calif, USA, September 2003.

[18] V. Bharghavan, "A dynamic addressing scheme for wireless media access," in *Proceedings of IEEE International Conference on Communications (ICC '95)*, vol. 2, pp. 756–760, Seattle, Wash, USA, June 1995.

[19] Y. Yuan, D. Gu, W. Arbaugh, and J. Zhang, "High-performance MAC for high-capacity wireless LANs," in *Proceedings of the 13th International Conference on Computer Communications and Networks (ICCCN '04)*, pp. 167–172, Chicago, Ill, USA, October 2004.

[20] F. Cali, M. Conti, and E. Gregori, "IEEE 802.11 protocol: design and performance evaluation of an adaptive backoff mechanism," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 9, pp. 1774–1786, 2000.

[21] P. Karn, "MACA: a new channel access method for packet radio," in *Proceedings of the ARRL/CRRL Amateur Radio 9th Computer Networking Conference*, pp. 134–140, Ontario, Canada, September 1990.

[22] D. Tse, "Multiuser diversity in wireless networks: smart scheduling, dumb antennas and epidemic communication," in *Proceedings of the IMA Wireless Networks Workshop*, August 2001.

[23] G. Holland, N. Vaidya, and P. Bahl, "A rate-adaptive MAC protocol for multi-hop wireless networks," in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking (MOBICOM '01)*, pp. 236–250, Rome, Italy, July 2001.

[24] M. Heusse, F. Rousseau, R. Guillier, and A. Duda, "Idle sense: an optimal access method for high throughput and fairness in rate diverse wireless LANs," in *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM '05)*, pp. 121–132, Philadelphia, Pa, USA, August 2005.

[25] T. S. Rapport, *Wireless Communications: Principles and Practice*, Prentice Hall, Upper Saddle River, NJ, USA, 2nd edition, 2005.

[26] Cisco Aironet Adapter, http://www.cisco.com/en/US/products/hw/wireless/ps4555/products_data_sheet09186a00801ebc29.html.

[27] D.-M. Chiu and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks," *Computer Networks and ISDN Systems*, vol. 17, no. 1, pp. 1–14, 1989.

[28] L. Bononi, M. Conti, and E. Gregori, "Runtime optimization of IEEE 802.11 wireless LANs performance," *IEEE Transactions on Parallel and Distributed Systems*, vol. 15, no. 1, pp. 66–80, 2004.

[29] F. A. Tobagi and L. Kleinrock, "Packet switching in radio channels—part IV: stability considarations and dynamic control in carrier sense multiple access," *IEEE Transactions on Communications*, vol. 25, no. 10, pp. 1103–1119, 1977.

[30] F. Cali, M. Conti, and E. Gregori, "Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit," *IEEE/ACM Transactions on Networking*, vol. 8, no. 6, pp. 785–799, 2000.

[31] G. Tan and J. Guttag, "Time-based fairness improves performance in multi-rate WLANs," in *Proceedings of the USENIX Annual Technical Conference*, pp. 269–282, Boston, Mass, USA, June-July 2004.

[32] IEEE 802.5: Defines the MAC layer for Token-Ring Networks.

[33] I. Cidon and M. Sidi, "Distributed assignment algorithms for multihop packet radio networks," *IEEE Transactions on Computers*, vol. 38, no. 10, pp. 1353–1361, 1989.

[34] R. Karrer, A. Sabharwal, and E. Knightly, "Enabling large-scale wireless broadband: the case for TAPs," in *Proceedings of the 2nd Workshop on Hot Topics in Networks (HotNets-II '03)*, Cambridge, Mass, USA, November 2004.

[35] Scalable Network Technologies, http://scalable-networks.com/.

[36] N. Vaidya and J. So, "A multi-channel MAC protocol for ad hoc wireless networks," Tech. Rep., Department of Electrical and Computer Engeneering, University of Illinois, Urbana-Champaign, Ill, USA, January 2003.

[37] C. Doerr, M. Neufeld, J. Fifield, T. Weingart, D. C. Sicker, and D. Grunwald, "MultiMAC—an adaptive MAC framework for dynamic radio networking," in *Proceedings of the 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN '05)*, pp. 548–555, Baltimore, Md, USA, November 2005.

[38] A. Rao and I. Stoica, "An overlay MAC layer for 802.11 networks," in *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services (MobiSys '05)*, pp. 135–148, Seattle, Wash, USA, June 2005.

[39] A. Farago, A. D. Myers, V. R. Syrotiuk, and G. V. Zaruba, "Meta-MAC protocols: automatic combination of MAC protocols to optimize performance for unknown conditions," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 9, pp. 1670–1681, 2000.

[40] B. A. Sharp, E. A. Grindrod, and D. A. Camm, "Hybrid TDMA/CSMA protocol for self managing packet radio networks," in *Proceedings of the 4th IEEE Annual International Conference on Universal Personal Communications (ICUPC '95)*, pp. 929–933, Tokyo, Japan, November 1995.