On the Impact of IEEE 802.11 MAC on Traffic Characteristics

Omesh Tickoo, *Student Member, IEEE,* and Biplab Sikdar, *Member, IEEE*

***Abstract*—IEEE 802.11 medium access control (MAC) is gaining widespread popularity as a layer-2 protocol for wireless local-area networks. While efforts have been made recently to evaluate the performance of various protocols in wireless networks and to evaluate the capacity of wireless networks, very little is understood or known about the traffic characteristics of wireless networks. In this paper, we address this issue and first develop an analytic model to characterize the interarrival time distribution of traffic in wireless networks with fixed base stations or ad hoc networks using the 802.11 MAC. Our analytic model and supporting simulation results show that the 802.11 MAC can induce pacing in the traffic and the resulting interarrival times are best characterized by a multimodal distribution. This is a sharp departure from behavior in wired networks and can significantly alter the second order characteristics of the traffic, which forms the second part of our study. Through simulations, we show that while the traffic patterns at the individual sources are more consistent with long-range dependence and self-similarity, in contrast to wired networks, the aggregate traffic is not self-similar. The aggregate traffic is better classified as a multifractal process and we conjecture that the various peaks of the multimodal interarrival time distribution have a direct contribution to the differing scaling exponents at various timescales.**

***Index Terms*—IEEE 802.11, medium access control (MAC), traffic modeling.**

I. Introduction

In the recent past, the use of mobile computing has increased tremendously and there has been a lot of work on improving the performance of various wireless protocols (both ad hoc and cellular with stationary access points). Evolving from networks which provide rudimentary services, the thrust of current development efforts is toward providing a richer set of quality sensitive applications. The performance of various mechanisms and policies which have been proposed to achieve these goals depends, to various extents, on the network's traffic characteristics. Thus, accurate models for the traffic and an understanding of the impact of various factors on the traffic characteristics are necessary for improving the capability of wireless networks in general and developing efficient schedulers, admission policies, etc., in particular. In this paper, we develop a model for traffic in IEEE 802.11 medium access control (MAC) based wireless networks. We then characterize the impact of transport and application layer protocols on the traffic characteristics. The insights from these results can lead to designs for improved network performance, as well as provide the modeling tools for more accurate performance evaluation.

While various protocols exist for the MAC layer, one of the protocols gaining popularity in wireless local area networks is the IEEE 802.11 MAC [13]. In this paper, we focus on the implications of this protocol on the traffic characteristics in a wireless network. We first develop analytic models for characterizing the interarrival time distribution in wireless networks. Our models are applicable to both ad hoc, as well as infrastructure networks (networks with fixed base stations). Our results show that in contrast to wired networks, the interarrival time distribution in wireless networks shows evidence of pacing induced by the MAC protocol and the interarrival times have a multimodal distribution. Our models are verified using extensive simulations.

While the interarrival time distribution gives important insights into the traffic characteristics, the scaling, burstiness, and long-range dependence associated self-similar and multifractal nature of traffic can lead to a number of undesirable effects like high buffer overflow rates, large delays, and persistent periods of congestion [5], [6], [20]. In this paper, we also investigate the impact of the 802.11 MAC on the second order scaling of the traffic. Our simulation results show that while individual TCP sources in wireless networks show evidence of self-similarity, the aggregate traffic is no longer self-similar and shows multifractal properties. We believe that the multimodal interarrival time distribution has a significant contribution to this phenomena. Also, the self-similarity at the individual nodes suggests that transport protocols dominate the user traffic patterns while at the base stations or intermediate nodes, the MAC protocols dominate. Additionally, we investigate how upper layer factors like the choice of the transport protocol and the application layer characteristics interact with the MAC protocol to induce power laws like behavior in the interarrivals and identify the critical time scales over which these layers dominate in their influence of the traffic characteristics.

The results of this paper provide a new insight into the behavior of wireless traffic and an analytic model for the impact of the 802.11 MAC on the interarrival times. These results can be used to develop more accurate queuing models and provide guidelines for developing schedulers for delay sensitive traffic, admission control policies, etc. Additionally, the paper also investigates the fractal nature of wireless traffic which gives important insights on the behavior of traffic across all time scales, important for accurate prediction of network performance as shown in [5]. The critical timescales over which various protocols influence the traffic are also identified in this paper. These timescales give important guidelines for buffer dimensioning

Manuscript received April 10, 2002; revised October 7, 2002. This work was supported in part by the Defense Advanced Research Projects Agency (DARPA) under Contract F30602-00-2-0537.

The authors are with the Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: tickoo@poisson.ecse.rpi.edu; bsikdar@poisson.ecse.rpi.edu).

Digital Object Identifier 10.1109/JSAC.2002.807346

and system design as specified by the network requirements [11], [5]. The complex interactions between the various layers of the protocol stack and their overall impact on system characteristics are also highlighted by these results.

The rest of the paper is organized as follows. Section II presents the recent work in the related fields, in Section III, we give a brief introduction of concepts used in this paper. An analytic model of the 802.11 MAC behavior for the transmission control protocol (TCP) traffic is presented in Section IV. We extend our analytical results to ad hoc systems in Section IV-B. We present our simulation results in Section V. Section VI summarizes the paper's contributions.

II. Related Work

Since the standardization of IEEE 802.11 MAC [13] many research efforts have concentrated around modeling the performance and capacity of 802.11 MAC protocol [3], [26] in particular and CSMA protocols in general [24]. Work has also been conducted on improving the 802.11 MAC as reported in [2], [29], and the references therein.

There has been a lot of work going on in the field of network and TCP traffic self-similarity [8], [9], [18], [20], [27], [28], [30]. A trace based study of traffic self-similarity in wireless networks is presented in [16]. Causes behind traffic self-similarity are investigated in [4], [9], [19], [22], [27], [28], [30], and in particular for TCP traffic, the authors of [27] suggest that the self-similarity of TCP traffic is due to the chaotic nature of the TCP traffic. The self-similarity of traffic leads to long range dependence that in turn makes it difficult to accurately predict the buffer requirements for such flows and, hence, leads to higher losses or gross resource underutilization [18], [20].

III. Introduction to Concepts

A. 802.11

The IEEE 802.11 standard [13] covers both the physical (PHY) and the MAC layer of wireless networks. The standard specifies that a network can be configured in two different ways: infrastructure and ad hoc. In infrastructure mode, the network is "structured," the hosts are fixed, and communicate through physical links. In an ad hoc network, the computers are brought together to form a network dynamically. There is no definite structure and any two computers can communicate directly as long as they are within "hearing" distance from each other.

The PHY layer can use any of the following modes: 1) direct sequence spread-spectrum (operation in the 2.4 to 2.4835 GHz frequency band); 2) frequency-hopping spread-spectrum; and 3) infrared pulse modulation (operation in the 300 to 428 GHz frequency band). The channel data rate can either be 1 Mb/s or 2 Mb/s. For more information on the 802.11 PHY please refer [13]. We now present a brief description of 802.11 MAC in the following subsection.

1) IEEE 802.11 MAC: The IEEE 802.11 MAC layer is responsible for a structured channel access scheme and is implemented using a distributed coordination function (DCF) based on the carrier sense medium access with collision avoidance (CSMA/CA) protocol. An alternative to the DCF is also provided in the form of a point coordination function (PCF) which is similar to a polling system for determining the user with the right to transmit. We now describe the relevant details of the DCF access method and refer the reader to [13] for details on the IEEE 802.11 standard.

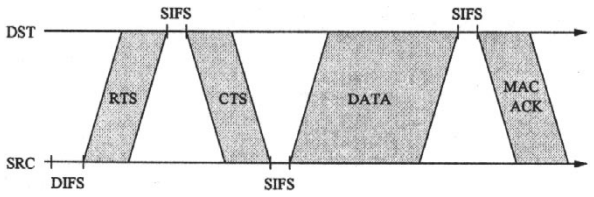


Fig. 1. Basic operation of the CSMA/CA protocol.

The CSMA/CA based MAC protocol of IEEE 802.11 is designed to reduce the collisions due to multiple sources transmitting simultaneously. In a network with a CSMA/CA MAC protocol, each node with a packet to transmit first senses the channel to ascertain whether it is in use. If the channel is sensed to be idle for an interval greater than the distributed interframe space (DIFS), the node proceeds with its transmission. If the channel is sensed as busy, the node defers transmission till the end of the ongoing transmission. The node then initializes its *backoff timer* with a randomly selected *backoff interval* and decrements this timer every time it senses the channel to be idle. The timer has the granularity of a *backoff slot* and is stopped in case the channel becomes busy and the decrementing is restarted when the channel becomes idle for a DIFS again. The node is allowed to transmit when the backoff timer reaches zero. Since the backoff interval is chosen randomly, the probability that two or more stations will choose the same back off value is very low. The details of the exact implementation of the backoff mechanism are described in Section IV. Along with the collision avoidance, 802.11 uses a positive acknowledgment (ACK) scheme. All the packets received by a node implementing 802.11 MAC must be acknowledged by the receiving MAC. After receiving a packet the receiver waits for a brief period, called the short interframe space (SIFS), before it transmits the ACK.

There is another particular feature of wireless local-area networks (LANs), known as the "hidden node" problem, that 802.11 MAC specification addresses. Two stations that are not within hearing distance of each other can lead to collisions at a third node which receives the transmission from both sources. To take care of this problem, 802.11 MAC uses a reservation based scheme. A station with a packet to transmit sends an ready to send (RTS) packet to the receiver and the receiver responds with a clear to send (CTS) packet if it is willing to accept the packet and is currently not busy. This RTS/CTS exchange, which also contains timing information about the length of the ensuing transaction, is detected by all the nodes within hearing distance of either the sender or receiver or both and they defer their transmissions till the current transmission is complete.

The basic operation of the CSMA/CA based MAC protocol of IEEE 802.11 is shown in Fig. 1 and shows the exchange of various packets involved in each successful transmission and the spacing between these packets.

B. Network Simulator ns-2

ns-2 is a discrete event simulator developed at University of California at Berkeley under the VINT project [7]. The wireless model is a very recent addition to the *ns-2* package. Apart from introducing the 802.11 channel model, the *ns-2* wireless capabilities include a variety of routing protocols, applications and physical layer models.

C. Destination Sequenced Distance Vector (DSDV) Protocol

DSDV protocol [21] is a modification of distance vector algorithms like the Bellman Ford algorithm. It is a hop-by-hop, non on-demand protocol. The modifications in the DSDV aim to introduce loop freedom that is present in the classical Distance Vector protocols primarily due to the propagation of "stale" routing information. We use DSDV as the routing protocol for all the simulations reported in this paper. However, we have verified our results with another routing protocol, the dynamic source routing or DSR [17], and the results we obtained match the ones reported in this paper for DSDV.

D. Discrete Wavelet Transform and Multiscaling

The effectiveness and accuracy of fractal models to capture the complex behavior of networks over a range of timescales has been shown under a number of scenarios in wireline networks [8], [18], [20], [27], [28], [30]. This paper addresses similar issues for wireless traffic and in this section we provide a brief introduction to the related concepts. We present the discrete wavelet transform based method for evaluating the scaling properties [8] and differentiate between the various types of scaling behavior studied in literature [5].

Consider an arrival process $A(0,\, t)$ which counts the cumulative traffic arrival in the time interval $(0,\, t)$ and its associated increment process $X_\Delta(i)$ defines as

$$X_\Delta(i) = A(0,\, i\Delta) - A(0,\, (i-1)\Delta). \tag{1}$$

The basic hypothesis associated with scaling of process is that the moments of the increment process behave as

$$\sum_i X_\Delta(i)^q \sim C(q)\Delta^{-\tau(q)} \qquad \text{as } \Delta \to 0. \tag{2}$$

While in theory, this behavior should hold over all time scales for the scaling hypothesis to be satisfied, in practice, the hypothesis can be said to be reasonable if the behavior is satisfied over a range of timescales. The function $\tau(q)$ is called the *structure* function and for mono-fractal or self-similar processes, $\tau(q)$ is linear in q. For multifractal processes, $\tau(q)$ is nonlinear in q and for both the processes, in general, is decreasing in q. The discrete wavelet transform represents a one dimensional signal $X(t)$ in terms of shifted and dilated versions of a bandpass wavelet function $\psi(t)$ and shifted versions of a low-pass scaling function $\phi(t)$. For the choices of $\psi(t)$ and $\phi(t)$ which allow us to form an orthonormal basis, the signal can be represented as

$$X(t) = \sum_k \langle X(t)\phi_{0,\,k}(t)\rangle \phi_{0,\,k}(t) + \sum_{j=0}^{\infty}\sum_k \langle X(t)\psi_{j,\,k}(t)\rangle \psi_{j,\,k}(t) \tag{3}$$

where $\phi_{j,\,k}(t) = 2^{-j/2}\phi(2^{-j}t - k)$ and $\psi_{j,\,k}(t) = 2^{-j/2}\psi(2^{-j}t - k)$ are the shifted and dilated version of the scaling and wavelet function, respectively. The quantity $d_{j,\,k} = \langle X\psi_{j,\,k}\rangle$ is referred to as the *wavelet coefficient* at scale j and time $2^j k$. Also, $|d_{j,\,k}|^2$ measures the energy in the signal at time $2^j k$ and about the frequency $2^{-j}\lambda_0$, where the reference frequency λ_0 depends on the wavelet ψ. If a process has scaling in some second-order statistic, then it will very often have scaling for all moments. The *partition function* is defined as

$$S_q(j) = \frac{1}{n_j}\sum_k |d_{j,\,k}|^q \sim C(q) j^{\alpha(q)} \tag{4}$$

and specifically the nature of the function $\alpha(q)$ can then be used to determine the nature of the scaling phenomena. For self-similar processes $\alpha(q)$ is linear and given by $\alpha(q) = Hq + q/2$. On the other hand, if $\alpha(q)$ nonlinear, we say that the process shows multiscaling. It is common in multifractal theory to define the exponents slightly differently: $\zeta(q) = \alpha(q) - q/2$ which results in $\zeta(q) = Hq$ for self-similar processes. An estimate of the self-similarity can also be obtained through a linear regression of $\log_2(S_q(j))$ on $\log_2(2^j) = j$. In this paper, we plot $y_j \equiv \log_2(S_q(j)) - g_j$ against the octave j to determine the presence of self-similarity in a log–log plot where g_j is defined as $g_j = \psi(n_j/2)/\ln 2 - \log_2(n_j/2)$.

IV. Modeling Interarrival Times

The aim of this paper is two-fold. First, we investigate the traffic characteristics of wireless networks while specifically focusing on the impact of the MAC protocol. Second, we develop a model for the traffic characteristics in these scenarios. In this section, we develop the analytic model to characterize the interarrival times of packets arriving at the base station or an intermediate point ("router") in a wireless network. We use the term router in this paper to denote a node which is capable of forwarding packets to other destinations. These nodes themselves may be sources or destinations of other traffic.

This paper focuses on modeling the interarrival times at an intermediate node in an ad hoc network or the access node in a fixed wireless network. The generalized topology assumed for the fixed wireless network is shown in Fig. 2(a) where a set of N hosts use the base station B to communicate. For the ad hoc network case, we consider a topology shown in Fig. 2(b) with the N sources using the two nodes marked B to forward their packets to their respective destinations. Note that this is a generalized scenario and the N destinations could be arbitrarily distributed, and may use more than two intermediate nodes to forward their packets to their destinations. The transmission rate of the wireless links is denoted by C b/s. Also, we are not concerned with issues of mobility here since we are interested in the traffic dynamics at a node which is being used by a set of other nodes. Our interest is in characterizing these interarrival times given a set of nodes are routing their packets through B. As long as this set of nodes stays within the transmission range of node B the medium access and control pattern is not affected. Of course if any node moves out of the range of B then the packet interarrival times change, but this is not the focus of this paper. However, even with mobility, the qualitative nature of the traffic characteristics, specially at lower time scales, does not change since the

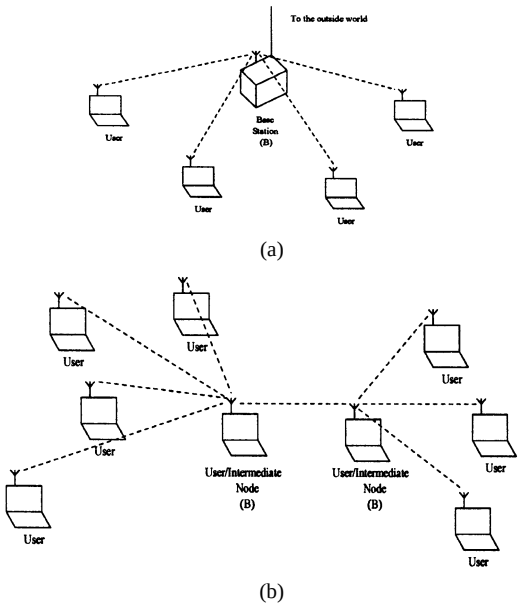


Fig. 2. The network models assumed in this paper. (a) Shows a BSS with a fixed base station. (b) Considers an ad hoc network.

channel access characteristics at the intermediate node essentially stays the same. This is verified using simulation results in Section V-B2.

The set of nodes N can send packets of different lengths and in the following derivation we restrict the possible packet lengths to two values: L_1 and L_2 bits to correspond to data and acknowledgment (ACK) packets, respectively. However, this model can be easily extended to account for arbitrary packet size distributions and transmission rates as shown in Section IV-C. Consider a network in operation for T_{sim} s and during this period, let the total number of data and ACK packets being forwarded to node B be given by N_1 and N_2, respectively. We consider that all nodes have packets to send at all points in time and the throughput is limited by other network related factors.

Before we deal with the basic service sets (BSS) of Fig. 2(a) and (b), we first analyze the backoff mechanism associated with the exponential backoff mechanism of the 802.11 MAC protocol's CA mechanism. The full details of the backoff scheme are given in [13]. In Fig. 3, we show the details of this backoff mechanisms. With multiple nodes contending for the channel, once the channel is sensed idle for a DIFS, each node with a packet to transmit decrements its backoff timer. The node whose timer expires first begins transmission and the remaining nodes stop their timer and defer their transmission. Once the current node finishes transmission, the process repeats again and the remaining nodes start decrementing their timer from where they left off. We now proceed with our analytic model which characterizes this behavior and use it to model the interarrival process at the base station or intermediate nodes.

Once a node goes into collision avoidance or the exponential backoff phase, we denote the number of slots that it waits beyond a DIFS period before initiating transmission by BC. This backoff counter is calculated from

$$\mathrm{BC} = \mathrm{int}(\mathrm{rnd}(\cdot) \cdot \mathrm{CW}(k)) \tag{5}$$

where the function rnd$(\cdot)$ returns a pseudorandom number uniformly distributed in $[0, 1]$ and $CW(k)$ represents the contention window after k unsuccessful transmission attempts. Note that in case the int operation is done using a ceil$(\cdot)$ function, the effective range for BC becomes $1 \leq \mathrm{BC} \leq \mathrm{CW}(k)$ since the probability of rnd$(\cdot) = 0$ is 0 assuming a continuous distribution. For the rest of this paper, we assume that a ceil$(\cdot)$ function is used to do the int$(\cdot)$ operation. The binary exponential backoff mechanism is manifested in the equation which relates $\mathrm{CW}(k)$ to the number of unsuccessful transmission attempts k already performed for a given packet. The first attempt at transmitting a given packet is performed assuming a CW value equal to the minimum possible value of $\mathrm{CW}_{\min}$ [13]. For each unsuccessful attempt, the value of CW is doubled until it reaches the upper limit of $\mathrm{CW}_{\max}$ specified by the protocol. Then, at the end of k unsuccessful attempts, $\geq 1, \mathrm{CW}(k)$ is given by

$$\mathrm{CW}(k) = \min(\mathrm{CW}_{\max}, 2^{k-1}\mathrm{CW}_{\min}). \tag{6}$$

Also, let the probability that a transmission attempt is unsuccessful, i.e., the probability of a collision be denoted by p. Then, the probability that $\mathrm{CW} = W$ is given by

$$\Pr\{\mathrm{CW} = W\} = \begin{cases} p^{k-1}(1-p), & \text{for } W = 2^{k-1}\mathrm{CW}_{\min} \\ p^n, & \text{for } W = \mathrm{CW}_{\max} \end{cases} \tag{7}$$

where $n = \log_2(\mathrm{CW}_{\max}/\mathrm{CW}_{\min})$. The probability that backoff counter $\mathrm{BC} = i$, $1 \leq i \leq \mathrm{CW}_{\max}$, is then given by

$$\Pr\{\mathrm{BC} = i\} = \begin{cases} \left[\sum_{k=0}^{n-1} \dfrac{p^k(1-p)}{2^k\mathrm{CW}_{\min}} + \dfrac{p^n}{\mathrm{CW}_{\max}}\right] \\ \qquad 1 \leq i \leq \mathrm{CW}_{\min} \\ \left[\sum_{k=j}^{n-1} \dfrac{p^k(1-p)}{2^k\mathrm{CW}_{\min}} + \dfrac{p^n}{\mathrm{CW}_{\max}}\right] \\ \qquad 2^{j-1}\mathrm{CW}_{\min} + 1 \leq i \leq 2^j\mathrm{CW}_{\min} \\ \dfrac{p^n}{\mathrm{CW}_{\max}} \\ \qquad 2^{n-1}\mathrm{CW}_{\min} + 1 \leq i \leq \mathrm{CW}_{\max}. \end{cases} \tag{8}$$

The expected value of the backoff counter is then given by

$$\begin{aligned} E[\mathrm{BC}] &= \sum_{j=0}^{n-1} p^j(1-p)\,\frac{2^j\mathrm{CW}_{\min}}{2} + \frac{p^n}{\mathrm{CW}_{\max}} \\ &= \frac{1-p-p(2p)^n}{1-2p}\,\frac{\mathrm{CW}_{\min}}{2}. \end{aligned} \tag{9}$$

At the end of each successful transmission, each node with a packet to transmit senses the channel and once it is sensed idle for a duration corresponding to a DIFS, the nodes start decrementing their backoff counters. The node(s) which decrements its backoff counter to 0 first sends out an RTS packet to reserve the channel. To find the interarrival times at the base station or any intermediate node, we first need to characterize the distribution of the number of slots it takes till the first node initiates transmission once the channel is idle for a DIFS.

Once the backoff counter value is selected after each unsuccessful transmission attempt, this value is decremented every

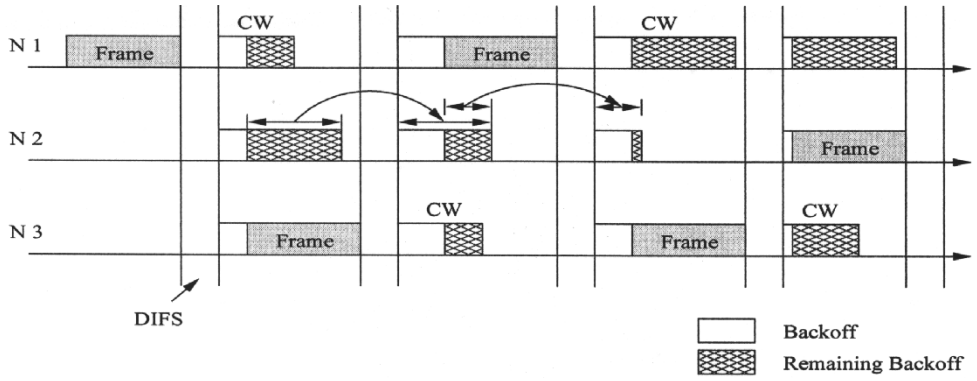


Fig. 3. The backoff mechanism of 802.11 MAC (based on [29, Fig. 3]). The frame transmission time includes the RTS/CTS exchange and the MAC layer ACK. CW: Contention window.

time the node senses an idle DIFS in the channel. At the end of the first idle a DIFS after the unsuccessful attempt, the node has to wait till the slots corresponding to the entire BC value expire. However, if some other node transmits before the BC value of the tagged node is decremented to 0, the counter is stopped. At the end of the next idle DIFS interval, the tagged node only has to wait till the *residual* value of BC is decremented to zero. Note that the generation of values for the BC at every unsuccessful transmission constitute a discrete renewal process. Using the results for the residual time distribution for discrete time processes [23], the residual time distribution for the backoff counter is given by

$$\Pr\{R = k\} = \frac{\Pr\{\mathrm{BC} > k\}}{E[\mathrm{BC}]} \tag{10}$$

where R represents the residual time and $1 \leq k \leq \mathrm{CW}_{\max}$. Note that while this is the residual time distribution only after the first idle DIFS, we use the above expression to approximate the residual time of the backoff counter at the end of the subsequent idle DIFSs.

With N nodes contending for the channel and decrementing their backoff counters simultaneously, the number of backoff slots from the end of the first idle DIFS period till the next packet transmission is the minimum of the residual times of the contending nodes. Note that since we assume that all nodes have a packet to transmit at all times, and at the end of a successful transmission the successful node delays his next transmission by a random number of slots, it is easy to see that all transmissions are preceded by a backoff. To find the distribution of the number of slots till the next transmission is initiated, we need to find the distribution of the minimum of the residual times of the competing nodes. This distribution is given by

$$\begin{aligned}\Pr\{M \leq k\} &= 1 - \Pr\{R_1 > k, R_2 > k, \ldots, R_N > k\} \\ &= 1 - \prod_{i=1}^{N} (1 - \Pr\{R \leq k\})\end{aligned} \tag{11}$$

where R_i, $1 \leq i \leq N$ represents the residual time at node i and in the final expression, we assume that all the R_is are independent and identically distributed. With these basic results in place, we now proceed to model the interarrival times at the access point or intermediate nodes. We add here that from the results of [26], the probability of collision p can be obtained by solving the following equation:

$$p\,\frac{1 - p - p(2p)^n}{1 - 2p} = \frac{2}{\mathrm{CW}_{\min}}\left(1 + \frac{2}{3}N\right)\frac{N-1}{N}. \tag{12}$$

A. Wireless Networks With Fixed Base Stations

We now consider the network depicted in Fig. 2(a) which corresponds to a wireless LAN. In this scenario, a number of wireless stations, which may be either fixed or mobile, use the central base station (B) to communicate with the outside world. We assume that the connection to the outside world from the base station is a wired link. Note that the results in the next section can be applied to cases where the outgoing link from the base station is through a wireless connection.

The packets sent or received by the nodes in the network are assumed to be either data or ACK packets corresponding to lengths of L_1 and L_2 bits, respectively, and the number of such packets arriving at B are N_1 and N_2, respectively. To derive the distribution of the interarrival times at B, we first start with the simple case where all nodes send *only* data packets and receive *only* ACK packets. In other words, we assume unidirectional data transfer. The duration for which the channel is busy when a data packet is successfully transmitted can be calculated as

$$T_1 = \mathrm{DIFS} + 3\mathrm{SIFS} + \tau_{\mathrm{RTS}} + \tau_{\mathrm{CTS}} + \tau_{L_1} + \tau_{\mathrm{ack}} \tag{13}$$

where τ_{RTS}, τ_{CTS}, τ_{L_1} and τ_{ack} correspond to the transmission times of RTS, CTS, data packet of length L_1 bits, and the MAC layer acknowledgment packet. The terms on the right-hand side (RHS) of (13) comes from the times associated with the sequence of packet transmissions associated with each successful transaction and is also shown in Fig. 1. Note that τ_{L_1} is given by $\tau_{L_1} = L_1/C$. Similarly, the time to successfully transmit an ACK packet is given by

$$T_2 = \mathrm{DIFS} + 3\mathrm{SIFS} + \tau_{\mathrm{RTS}} + \tau_{\mathrm{CTS}} + \tau_{L_2} + \tau_{\mathrm{ack}} \tag{14}$$

with $\tau_{L_2} = L_2/C$. Additionally, in the case of a collision, i.e., when two or more of contending nodes waiting to transmit decrement their backoff counter to zero simultaneously, the

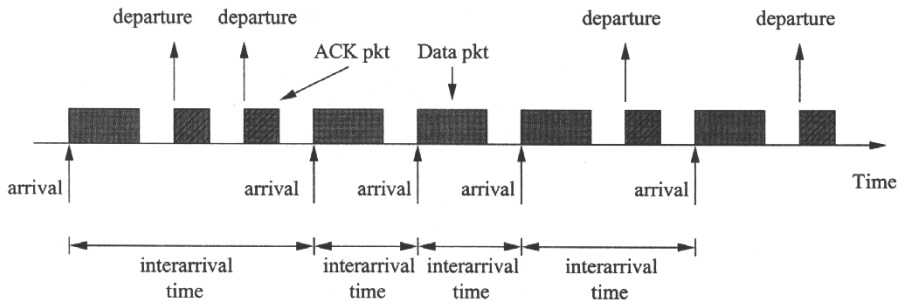


Fig. 4. Packet arrivals and departures at node B with uni-directional data flow.

channel stays busy for a duration of T_C before contending nodes start decrementing their backoff counters again. This duration corresponds to a DIFS period and $\tau_{\rm RTS}$ when the RTS messages from the colliding nodes overlap. Thus

$$T_C = \text{DIFS} + \tau_{\rm RTS}. \tag{15}$$

In the current scenario, the packets arriving at node B from the wireless medium correspond to the data packets. And all the packet transmissions out of this node into the wireless channel are of length L_2 corresponding to ACK packets. These packet arrivals and transmissions form a sequence of events as shown in Fig. 4. As depicted in the figure, the interarrival times at B correspond to the difference of the arrival instants of two successive data packets. This duration between the arrival of successive data packets depends on the number of ACK packets which arrive between them. Also, with each data or ACK packet, in addition to the times T_1 and T_2 which are associated with their transmissions, we have the times associated with the backoff counter slots which contribute to the interarrival times. Now, since we have N_1 and N_2 packets of lengths L_1 and L_2, respectively, the probabilities p_1 and p_2 that any arbitrary event at node B corresponds to a data or ACK packet, respectively, are given by

$$p_1 = \frac{N_1}{N_1 + N_2} \qquad p_2 = \frac{N_2}{N_1 + N_2}. \tag{16}$$

To evaluate the interarrival times, we need to evaluate two components: 1) the number of ACK packet transmissions between two successive data packets and 2) the random number of backoff slots associated with each data or ACK packet transmission. The presence of each ACK packet adds a time of T_2 s to the interarrival time while the first of the two data packets under consideration adds T_1 s. The expected value of the contribution due to the backoff slots associated with each ACK packet is given by $E[M]$ where the mass function for the number of backoff slots M is obtained using (11). Then, the probability that we wait for T's between two successive arrivals at node B is given by

$$\Pr\{T = t\} = p^i(1-p)^{j+1}p_1p_2^j \Pr\{M = k\} \tag{17}$$

where

$$t = T_1 + jT_2 + iT_C + (i+j)E[M] + k\Delta \tag{18}$$

and Δ represents the width of a backoff slot in seconds and $i \geq 0$, $j \geq 0$, and $1 \geq k \geq \text{CW}_{\max}$. The term p^i in the expression accounts for the increase in the interarrival times by $i(T_C + E[M])$ in the case of i collisions. The term $p_1p_2^j$ denotes the probability of j ACK packets being forwarded by B before receiving a data packet. The final term $\Pr\{M = k\}$ denotes the probability of having k backoff slots before the data packet transmission begins once the sender senses that the channel is IDLE for DIFS seconds. Note that the same value of t can be obtained through a number of possible choices of i, j and k, and the probability corresponding to each of these cases should be summed up obtain the probability of an interval having a length of t. We also note that the above expression has an approximation in the fact that we approximate the random contribution due to the backoff slots associated with the ACK packets by the expected duration of these backoff slots.

The generalization of the above model to consider the case when the user nodes are allowed to transmit, as well as receive data can be done through the derivations in the next section which considers the intermediate nodes in an ad hoc network. In this cases, since the users send both data and ACK packets instead of only data packets, node B has to forward packets of both lengths into the wireless medium. This reduces the access point B to a special case of the intermediate nodes of Fig. 2(b) as far as modeling the interarrival times is considered. We elaborate more on this in the following section.

B. Ad Hoc Networks

The general structure for the ad hoc network considered in this paper is shown in Fig. 2(b). The intermediate nodes (B) forward traffic from other nodes in addition to sending or receiving data themselves. Also, we do not put any limitations on the users in terms of the types (data or ACK) of packets they transmit and assume bidirectional traffic at the user nodes. Thus, the intermediate nodes send, receive and forward both data and ACK packets. The access point B of the previous section can, thus, be considered a special case of these intermediate nodes if we allow bidirectional traffic in the BSS.

We now consider an arbitrary intermediate node in the ad hoc network which we represent by B. Let the number of end users and other intermediate nodes forwarding packets through node B be denoted by N. In contrast to the model in the previous section with uni-directional traffic, allowing for the case when node B can forward data packets or receive ACK packets from the wireless channel introduces complexities in the traffic patterns. A possible sequence of packet arrivals and departures from node B is shown in Fig. 5. In this case, instead of only successive data packet arrivals representing an interarrival period, interarrivals can correspond to successive arrivals of the following possible order of

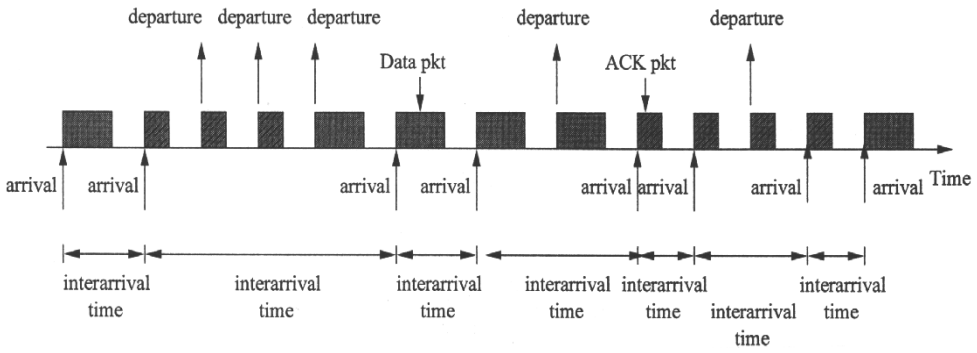


Fig. 5. Packet arrivals and departures at node B with bi-directional data flow.

packets: 1) data packet and data packet; 2) data packet and ACK packet; 3) ACK packet and data packet; and 4) ACK packet and ACK packet. Note that between the successive arrivals of either data or ACK packets and combinations thereof, we can have an arbitrary number of data or ACK packets being forwarded by the node which will contribute to the interarrival times.

Let the number of data and ACK packets forwarded successfully by the users and other intermediate nodes to node N be denoted by N_1 and N_2, respectively. Note that if there are no buffer overflows at the intermediate node, all of these packets are forwarded by node B. We denote the probability that at any point of time, node B is receiving data or ACK packets by p_1 and p_2, respectively, and the probability that it is forwarding data or ACK packets by p_3 and p_4, respectively. These probabilities are, thus, given by

$$p_1 = p_3 = \frac{N_1}{2(N_1+N_2)} \qquad p_2 = p_4 = \frac{N_2}{2(N_1+N_2)}. \quad (19)$$

To find the probability mass function for the interarrival times, we first need to find the probabilities corresponding to the various ways in which two successive arrivals can be spaced by various sequences of forwarded packets between them. Given that we have an arrival at node B, the probability that it is a data or ACK packet, p_1' and p_2', is given by

$$p_1' = \frac{p_1}{p_1+p_2} \qquad p_2' = \frac{p_2}{p_1+p_2}. \quad (20)$$

Now, if the first of the two arrivals constituting an interarrival is a data packet, it adds T_1 s to the interarrival time while an ACK packet adds T_2 s. Each data or ACK packet forwarded by node B between the two arrivals constituting the interarrival adds another $T_1+E[M]$ or $T_2+E[M]$ s, respectively. Then, the probability that we wait for T's between two successive arrivals at node B is given by

$$\Pr\{T=t\} = \begin{cases} \left[(1-p)^{i+j+1}\dfrac{p_1}{p_1+p_2}\,p^l p_3^i p_4^j \cdot (p_1+p_2)\Pr\{M=k\}\right] & \text{first arrival: data packet} \\ \left[(1-p)^{i+j+1}\dfrac{p_2}{p_1+p_2}\,p^l p_3^i p_4^j \cdot (p_1+p_2)\Pr\{M=k\}\right] & \text{first arrival: ACK packet} \end{cases} \quad (21)$$

where

$$t = \begin{cases} [(i+1)T_1 + jT_2 + lT_C & \text{first arrival:} \\ \quad +(i+j+l)E[M] + k\Delta] & \text{data packet} \\ [(j+1)T_2 + iT_1 + lT_C & \text{first arrival:} \\ \quad +(i+j+l)E[M] + k\Delta] & \text{ACK packet} \end{cases} \quad (22)$$

where $i \geq 0$, $j \geq 0$, $l \geq 0$ and $1 \geq k \geq \mathrm{CW}_{\max}$. The term p^l in (21) represents the cases where l data or ACK packet transmissions are attempted unsuccessfully before a second packet is correctly received at node B. These add $l(T_C + E[M])$ s to the interarrival time. The terms $p_1/(p_1+p_2)$ and $p_2/(p_1+p_2)$ represent the probability that the first of the two arrival constituting the interarrival is a data or ACK packet, respectively. This adds T_1 or T_2 s to the interarrival time. The term $p_3^i p_4^j$ corresponds to cases where i data and j ACK packets are forwarded by node N between the two arrival contributing $iT_1+jT_2+(i+j)E[M]$ s. The term (p_1+p_2) represents the probability that the i data and j ACK packet forwarded by the intermediate node is followed by a packet reception. Finally, the term $\Pr\{M=k\}$ denotes the probability of having k backoff slots before the data or ACK packet transmission corresponding to the second arrival begins once the sender senses that the channel is idle for DIFS seconds. Again, the same value of t can be obtained through a number of possible choices of i, j, k, l, and m and the probability corresponding to each of these cases should be summed up to obtain the probability of an interval having a length of t.

The model can also be extended to find the interarrival times considering the arrivals from only the data or ACK packets in the stream. For the case when we consider *only* the data packets arrivals to constitute the arrival process that we are interested in, between two such successive arrivals, we can have data or ACK packets being forwarded or ACK packets being received at node B. Each of these events, since they are not considered "arrivals," contributes to the interarrival times. The probability mass function for the interarrival time can be expressed as

$$\Pr\{T=t\} = p^i(1-p)^{j+l+m+1} p_1 p_2^j p_3^l p_4^m \Pr\{M=k\} \quad (23)$$

where

$$t = (l+1)T_1 + (j+m)T_2 + iT_C + (i+j+l+m)E[M] + k\Delta \quad (24)$$

where $i \geq 0$, $j \geq 0$, $1 \geq k \geq \mathrm{CW}_{\max}$, $l \geq 0$, and $m \geq 0$. Comparing this expression for $\Pr\{T=t\}$ with the one in (17), we note that there is an additional term $p_2^j p_3^l$ (the terms p_4^m in (23) and p_2^j in (17) represent the same thing). This reflects the

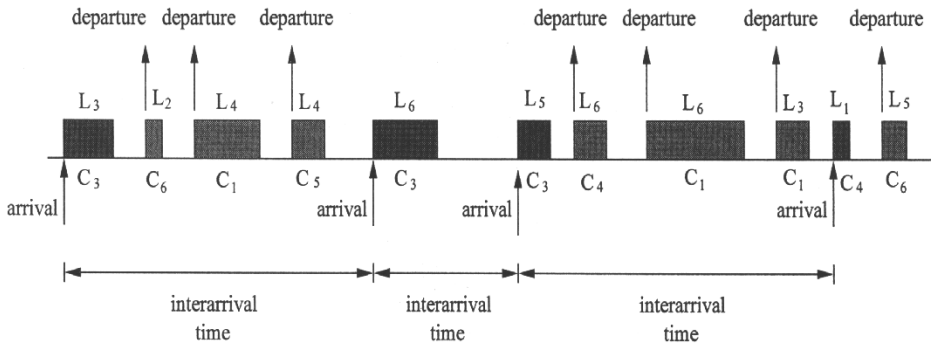


Fig. 6. Packet arrivals and departures at node B with multirate data flow.

fact that now node B in addition to receiving ACKs which do not count as arrivals in the current definition, also forward data packets which increases the interarrival times. The increase in the interarrival time corresponds to $lT_1+mT_2+(l+m)\Delta$. Also, a simple modification to (23) allows us to model the interarrival times when we consider *only* the ACK arrivals at B to constitute our arrival process. The interarrival time mass function is given by

$$\Pr\{T=t\}=p^j(1-p)^{i+l+m+1}p_1^i p_2 p_3^l p_4^m \Pr\{M=k\} \quad (25)$$

where

$$t=(i+l)T_1+(m+1)T_2+jT_C+(i+j+l+m)E[M]+k\Delta \quad (26)$$

where only the term $p^{j-1}(1-p)$ changes to reflect the fact that arrivals now correspond to ACK packets and $j \geq 0$, $i \geq 0$, $1 \geq k \geq \mathrm{CW}_{\max}$, $l \geq 0$, and $m \geq 0$.

C. Multirate Operation in 802.11

Wireless systems compliant with IEEE 802.11 MAC specifications may use different modulations schemes depending on existing channel characteristics and, thus, achieve different data rates. While the original specification allowed for data rates of 1 and 2 Mb/s, IEEE 802.11b [14] specifications allows for additional rates of 5.5 and 11 Mb/s while IEEE 802.11a specifications target a range of data rates from 6 to 54 Mb/s [15]. Under such circumstances, the operation of the network becomes more complex and traffic interarrival times can be expected to show more variability. We now present an extension of the model developed in the previous sections to account for multirate operation.

Consider the case when the MAC is capable of supporting r different transmission rates $C_1, C_2, \ldots, C_r$. We also generalize the allowable packet lengths from just two in the previous sections to s possible values $L_1, L_2, \ldots, L_s$. Let the number of packets of size L_i, $1 \leq i \leq s$, transmitted at rate C_j, $1 \leq j \leq r$, forwarded successfully by the users to an arbitrary intermediate node B be denoted by $N^r_{i,j}$ with $\sum_{i=1}^{s}\sum_{j=1}^{r} N^r_{i,j}=N$. Now node B forward each of these packets to their next hop though it is not necessary that the forwarding transmission rate is the same as the receiving rate. The probability that at any point of time, node B is receiving a packet of length L_i, $1 \leq i \leq s$, at rate C_j, $1 \leq j \leq r$, is given by

$$p^r_{i,j}=\frac{N^r_{i,j}}{2\sum_{i=1}^{s}\sum_{j=1}^{r} N^r_{i,j}}=\frac{N^r_{i,j}}{2N}. \quad (27)$$

Similarly, we define $N^s_{i,j}$ and $p^s_{i,j}$ as the number of packets of size L_i transmitted at rate C_j at node B and the probability of a transmission at rate C_j of a packet of length L_i, respectively. The transmission time of a packet of size L_i, $1 \leq i \leq s$, transmitted at rate C_j, $1 \leq j \leq r$, is given by $\tau_{L_i,C_j}=L_i/C_j$. Then, following the arguments of Section IV-A the time to successfully transmit a packet of size L_i and rate C_j is given by

$$T_{i,j}=\mathrm{DIFS}+3\mathrm{SIFS}+\tau_{\mathrm{RTS}}+\tau_{\mathrm{CTS}}+\tau_{L_i,C_j}+\tau_{\mathrm{ack},C_j}. \quad (28)$$

where τ_{ack,C_j} is the time required to transmit the MAC layer ACK packet at the rate C_j.

We now consider an arbitrary intermediate node B in the ad hoc network. The interleaving of packets of different lengths transmitted at different rates results in much more complex interarrival patterns at node B and an example of a possible sequence of packet arrivals and departures is shown in Fig. 6. Note that now between any two successive arrivals, we can have an arbitrary number of packets of all possible combinations of packet lengths and transmission rates. Extending the derivation of Section IV-B, the probability that we wait for T's between two successive arrivals at node B is given by

$$\Pr\{T=t\}=\frac{p^l(1-p)^{m+1}p^r_{i,j}}{\sum_{i=1}^{s}\sum_{j=1}^{r} p^r_{i,j}} \cdot\left[\prod_{i=1}^{s}\prod_{j=1}^{r}(p^s_{i,j})^{m_{i,j}}\right]\Pr\{M=k\}$$

where

$$t=T_{i,j}+\sum_{i=1}^{s}\sum_{j=1}^{r} m_{i,j}T_{i,j}+lT_C+(m+l)E[M]+k\Delta \quad (30)$$

and $m_{i,j}$ represents the number of transmissions of packets of length L_i at rate C_j at node B between any two arrivals and $m=\sum_{i=1}^{s}\sum_{j=1}^{r} m_{i,j}$. The explanations for the various terms in the equations above is analogous to those in Section IV-B. Note that the above equations are for the case when the first received packet corresponding to the interarrival have length L_i

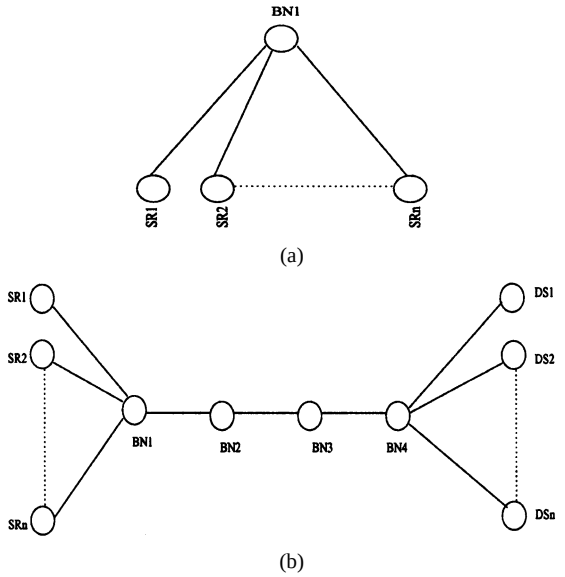


Fig. 7. Simulation topologies. (a) BSS with a fixed base station. (b) Ad hoc network.

and was transmitted at rate C_j. Thus, (29) needs to be summed for all possible combinations of L_i and C_j and also all possible values of $m_{i,j}$ which result in an interarrival time of t. The above analysis can also be extended to find the interarrival times considering the arrivals from only a particular packet size.

V. Simulation Results

To verify and evaluate the model presented in Section IV, we conducted extensive simulations using the *ns-2* simulator and in this section, we present these results. In addition to the validation results, we also investigate the scaling behavior of wireless traffic in various scenarios.

A. Simulation Model

To verify our analytical model and also to study higher order characteristics of traffic under 802.11 MAC protocol, we used the topologies shown in Fig. 7. The distances between the nodes were chosen such that all the source nodes are at a "hearing" distance from each other, as well as the node BN1. Thus, BN1 can be considered as the base station B of Section IV-A and the intermediate node B in Section IV-B. The results presented in this section correspond to the traces and statistics collected at node BN1. In Fig. 7(b), none of the source nodes are in the range of BN2, BN3 and BN4, or any of the destination nodes DS1, DS2, etc. Similarly, the destination nodes are within the hearing distance of each other and only BN4. The data packets are, thus, forwarded to BN1 and subsequently to BN2, BN3, and BN4 before finally reaching its destination node. The sources use TCP as the transport protocol and the data packet size was 1000 B while the ACK packets were 40 B long. Each simulation was run for 3600 s of simulated time giving us hour long traces. The interface queues used a Droptail policy and the interface queue length was set at 50 packets. All the simulation results reported use DSDV as the routing protocol. We have also verified our results using the DSR routing protocol. All sources and receivers have an OmniAntenna of height 1.5 m and transmitter and receiver gains of one. All other parameter settings for these simulations are given in Table I.

TABLE I
Simulation Settings

Physical Layer		802.11 MAC	
Propagation	2 ray gnd	RTS size	44 bytes
Channel	Wireless	CTS size	38 bytes
Rx Threshold	3.652e-10	DIFS	50 μsec
Bandwidth	2 Mbps	SIFS	10 μsec
Frequency	914 MHz	Slot size	20 μsec
Loss Factor	1.0		

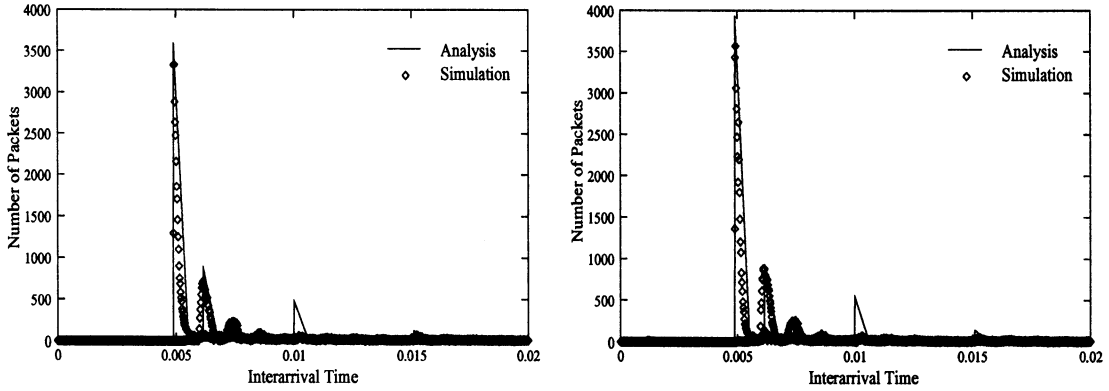
In our simulation models, the nodes have no mobility. This is primarily because our interest in this paper is to characterize and understand how 802.11 MAC on its own can influence the traffic interarrivals and second order scaling. Also, as described in Section IV, as long as a set of nodes are within hearing range of each other, their mobility pattern is not an issue in the channel access and usage dynamics. In the presence of mobility, the set of nodes forwarding their packets through BN1 will keep changing. We validate the generality of the conclusions of our paper for scenarios with mobility using simulations in Section V-B2. This implies that our analysis can also be used to approximate the traffic characteristics in a wireless network with mobile hosts if we interpret N of Section IV as the "average" number of nodes forwarding packets through the base station or intermediate node.

We consider the following scenarios for the purpose of our analysis: **A)** Simulations for ad hoc networks (modeled in Section IV-B) and **B)** Simulations for wireless networks with fixed base stations (modeled in Section IV-A). For both cases, we consider the following scenarios: 1) 40 sources; 2) 20 sources; and 3) 10 sources. We now present the results for the 40 and 40 sources cases in Section V-B. The results for the 20 nodes case are similar and thus omitted.

B. Results

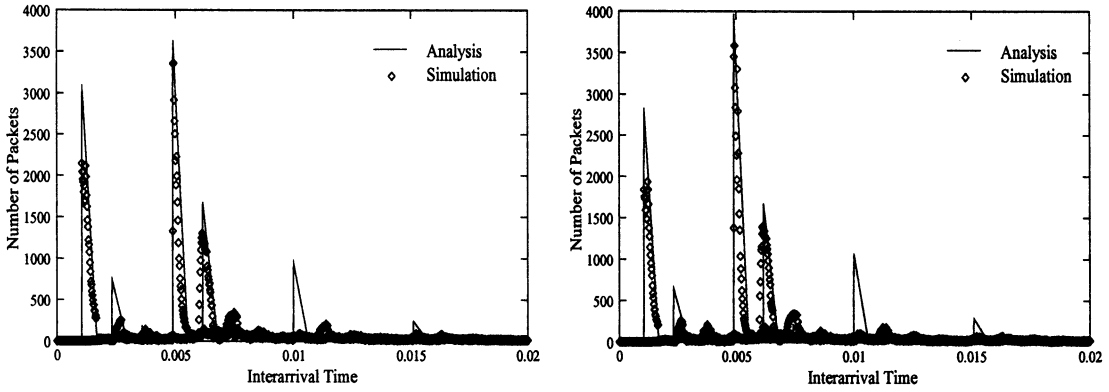
In our study of the traffic characteristics under a 802.11 MAC, we first investigate the distribution of the interarrival times at the base stations and intermediate nodes. In this section, we also present the simulation results to validate our analytic model for the interarrival time distributions. Our second traffic characteristic of interest is the second order scaling or self-similarity of the traffic at the base stations or the intermediate nodes. In this context, we use the wavelet based analysis of [1] to characterize the second order scaling since in addition to estimating the degree of long range dependence, wavelet plots are well suited for characterizing nonlinearities in the scaling exponents. We also used the absolute value and R/S statistic methods to estimate the degree of similarity [25] though we only report the results from the wavelet based method due to space constraints.

1) Interarrivals Times: Fig. 8 shows the interarrival time distribution at node BN1 when the interarrival process corresponds to just the data packet arrivals. these results correspond to the topology shown in Fig. 7(b). The figure also shows the analytic results as given by (23) and we note the close match with the simulation results. The first peak corresponds to two TCP packets coming next to each other while the next two peaks signify cases where we have one or two ACK packets between



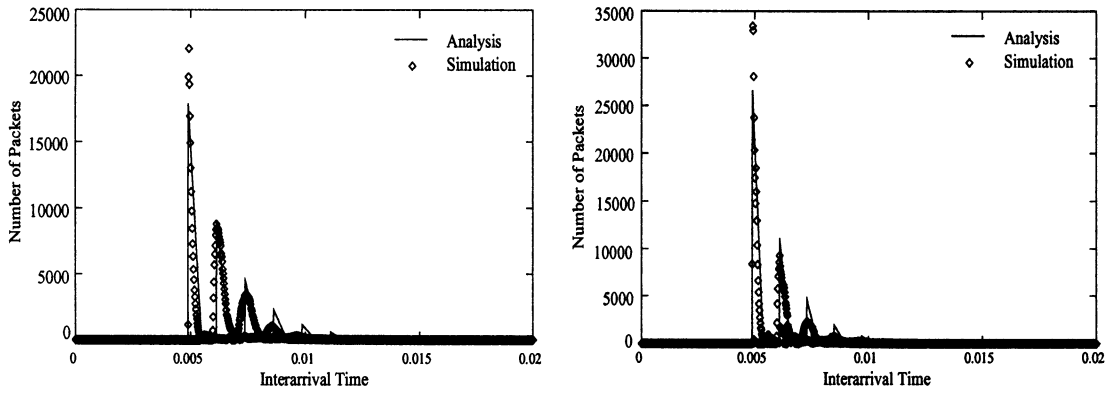
(a) (b)

Fig. 8. Interarrival times for only TCP packets for the ad hoc network. (a) 10 nodes. (b) 40 nodes.



(a) (b)

Fig. 9. Interarrival times for both TCP and ACK packets for the ad hoc network. (a) 10 nodes. (b) 40 nodes.



(a) (b)

Fig. 10. Interarrival times at the base station of the BSS. (a) 10 nodes. (b) 40 nodes.

the TCP packet, respectively. Fig. 9 shows the interarrival times when we consider both TCP and ACK packets to constitute the arrival process. The analytic model, which again shows a close match to the simulation results is expressed in (21). Note that in this case the first peaks appear much earlier and correspond to back to back ACK packets. The smaller peaks correspond to the increase in interarrival times due to the ACK and TCP packets being forwarded by the node.

In Fig. 10, we show the interarrival time distribution at the base station BS1 of the topology shown in Fig. 7(a). The analytic model corresponds to (17) and the arrival process at BN1 corresponds to the TCP packets being received. We note that there are more peaks in this case with the first peak corresponding to two TCP packets being received back to back. The subsequent four discernible peaks correspond to 1, 2, 3, or 4 ACK packets arriving between two data packets.

The most significant observation from these figures is the pacing or clustering of the interarrival times around various modes in the distribution. The times T_1 and T_2 associated with the transmission times and the intermingling of the reception and forwarding events at the base station or intermediate node leads to the formation of these peaks.

2) Results With Mobile Nodes and Fading Channels: Though our model does not account for mobility,

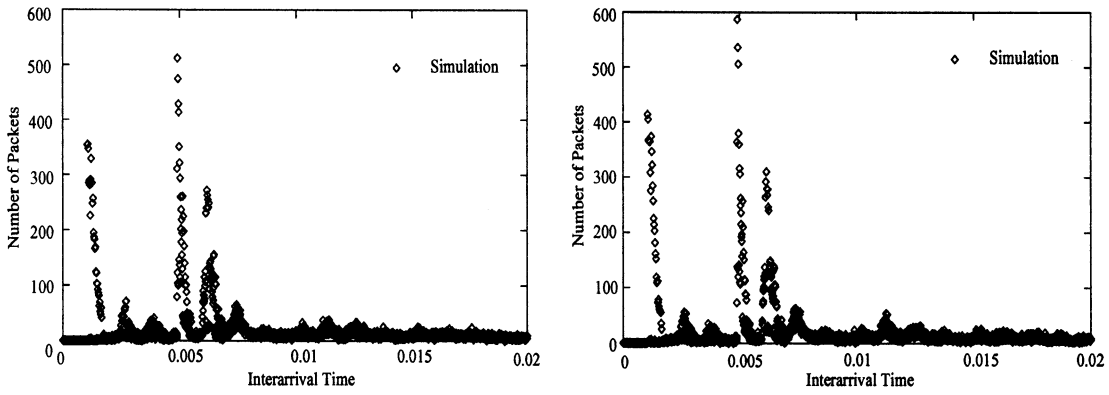


Fig. 11. Interarrival times with mobility. (a) 10 sources. (b) 40 sources.

its general conclusions on the impact of the 802.11 MAC on the traffic characteristics, for example pacing and multiple modes, are still valid in scenarios with mobile nodes. With mobility, nodes move in and out of range of any given intermediate node, and the set of nodes using this intermediate node to forward their packets may change over time. However, as long as the number of nodes in this sets does not change very rapidly, the traffic characteristics remain more or less the same since the medium access and control pattern is not significantly affected by these variations.

In Fig. 11, we plot the simulation results for the interarrival times at an intermediate node when the nodes are mobile. The simulation was conducted assuming an area of 500 $\times$ 1500 m. The random way-point model was used to generate the mobility patterns and all nodes had a speed of 5 m/s. We note that for both 10 and 40 node case, the interarrival time distributions has characteristics similar to those in Fig. 9 and shows the pacing introduced by the MAC. Thus, even with mobility, the qualitative nature of the traffic characteristics, specially at lower time scales, does not change since the channel access characteristics at the intermediate node essentially stays the same. We do note that with mobility, various other factors like the mobility model and routing protocol can affect the interarrival time characteristics. However, characterizing these effects is beyond the scope of this paper.

Wireless networks frequently operate under harsh and unpredictable channel conditions due to a variety of factors like multipath and shadow fading, Doppler spread and time dispersion, or delay spread. These factors are all primarily caused by the variability introduced by the mobility of the user or objects in its vicinity. The most common and significant of these, multipath fading, occurs when the transmitted signal is reflected by objects in the environment between the base station and a user. The distinct paths followed reflected signals results in random phase offsets at the sender which causes random fading of the signal when the reflections destructively (and constructively) superimpose on one another. The degree of cancellation, or fading, depends on the delay spread of the reflected signals, as embodied by their relative phases, and their relative power.

The statistics describing the fading signal amplitude are usually characterized as either Rayleigh or Ricean. Rayleigh fading occurs when there is no line-of-sight (LOS) component present in the received signal while in its presence, the fading follows a Ricean distribution. The probability density function for the Ricean envelop is given by

$$p(r)=\frac{r}{\sigma^2}\,e^{-((r^2+A^2)/2\sigma^2)}I_0\left(\frac{Ar}{\sigma^2}\right), \quad \text{for } A\geq 0,\ r\geq 0. \tag{31}$$

The parameter A denotes the peak amplitude of the dominant signal and $I_0(\cdot)$ is the modified Bessel function of the first kind and zero order. The Ricean distribution is also described in terms of a parameter K with $K = A^2/(2\sigma^2)$ where σ^2 is the variance of the amplitude of the dominant signal. Also, $A \rightarrow 0$, $K \rightarrow -\infty$, the Ricean distribution degenerates to a Rayleigh distribution.

In Fig. 12, we plot the interarrivals time distribution at the base station BS1 of the topology shown in Fig. 7(a) for both Ricean and Rayleigh fading models. In both cases, the interarrival characteristics are similar with that observed in Fig. 10. However, in the presence of fading, increased frame errors leads to considerably less packets being received successfully thereby increasing the interarrival times. Also, with the Ricean model we have line of sight and, thus, better signal reception. Thus, the magnitudes of the peaks in this case is considerably higher than that for the case with Rayleigh fading.

3) Traffic Characteristics With Point Coordination Function (PCF): While DCF is the basic medium access in 802.11 and must be implemented by all nodes, the specifications also allow for a polling based access mechanism, the PCF, primarily intended for real-time traffic. PCF provides a contention free transfer mechanisms which coexists with the DCF at the MAC layer. The access point performs the functions of the point coordinator by gaining control of the medium at periodic intervals (the contention free repetition interval or superframe) and starting the contention free period (CFP). During the CFP, nodes which are on the polling list of the access point are polled for a *single* pending frame transfer according to a polling list. On receiving the poll, stations with data to send transmit their frames. Once all nodes in the polling list finish transmitting, the time remaining in the superframe is spent in the DCF mode. The superframe size, as well as the number of nodes being polled have an important impact on the traffic characteristics and in this section we investigate these in more detail.

With periodic polling of nodes, we expect the traffic to become more regular in the presence of PCF based operation.

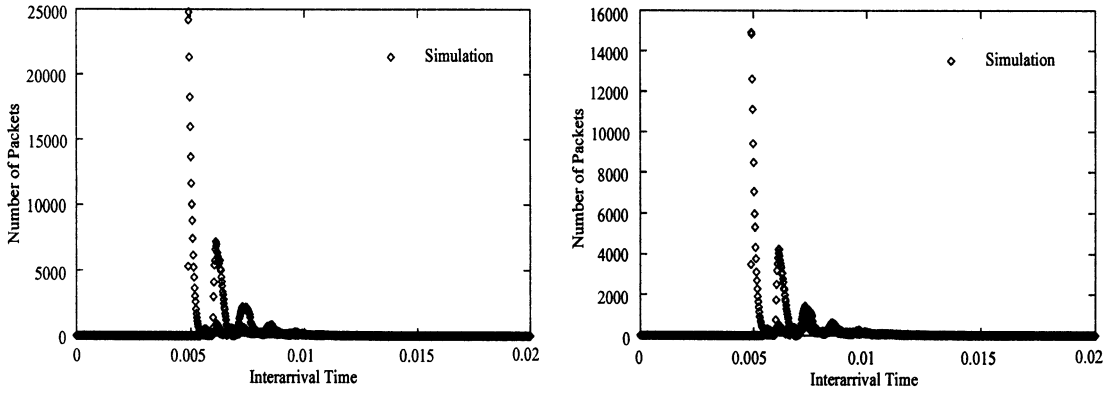


Fig. 12. Interarrival times with fading. (a) Rayleigh, $K = -\infty$. (b) $K = 6$.

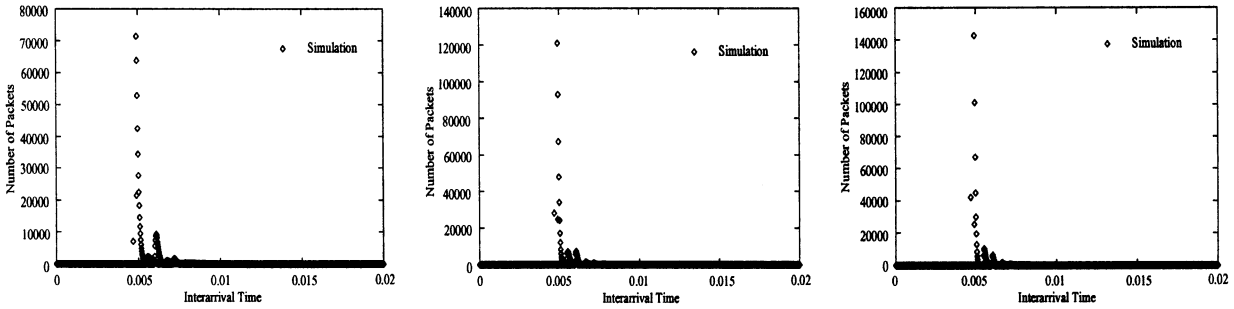


Fig. 13. Interarrival times with PCF operation with ten nodes. (Superframe size = 500 ms, 1000-B packets.) (a) Two polled nodes. (b) Five polled nodes. (c) Seven polled nodes.

Also, as the number of nodes in the polling list increases, with fixed length packets, traffic becomes more and more periodic. In Fig. 13, we show the simulation results for the interarrival time distribution at the base station BS1 of the topology shown in Fig. 7(a). The figure shows the results for 10 nodes and the results for the other topologies are similar. The polled nodes generate constant bit rate (CBR) UDP traffic while the nonpolled nodes generate TCP traffic. The choice of CBR UDP traffic for the polled nodes was made keeping in mind the fact that the polled mode is primarily for real-time traffic. We note that even with two nodes polled, though the traffic characteristics have some similarity to those observed in Fig. 10, now we have more traffic which arrives periodically as seen by the relative sizes of the peaks. As the number of polled nodes increases, we see that the relative magnitude of the first peak increases, showing that the traffic becomes more periodic. With a larger number of polled nodes, the fraction of time spent in the CFP increases. This results in a larger fraction of incoming packets arriving at periodic intervals.

4) Interarrival Times at Larger Time-Scales: In the results presented in the previous section, note that the X axis has a maximum value of 20 ms. Thus, these results take a fine grained look at the interarrival times. The use of RTS and CTS messages to reserve the channel usage reduces the time for which the channel remains unutilized and, thus, also affects the interarrival times. Also, the dynamics of the IEEE 802.11 MAC are such that its impact on the interarrival times are confined to the smaller time scales. However, the impact of upper layer protocols like TCP can also modulate the packet interarrival times even if the application always has data to send. In this section, we investigate the large time scale behavior of the traffic characteristics. The results in this section correspond to simulations of 20 sources. The trends in the results for the other topologies are similar and are omitted due to space constraints.

In Fig. 14(a), we plot the complementary cumulative distribution function (CCDF) of the packet interarrival times to time scales of tens (the graphs show the total number of arrivals with interarrival times greater than a certain value rather than its normalized value). We see that the CCFD plot obtained from the simulation has a linear slope in the log–log plot till time scales of few hundred milliseconds. However, looking at the characteristics beyond a few hundred milliseconds, it is apparent that the interarrival times are not truly heavy-tailed as the curve drops sharply. The figure also shows the CCDF of the interarrival times as predicted by the analytic model of Section IV and we note that the contribution of the MAC layer to the interarrival times die out after a few tens of milliseconds. Thus, it must be the upper layer protocols which contribute to the linear behavior of the CCDF at higher time scales and we now explore these in detail.

The "heavy-tailed" behavior over the time scales of 10s milliseconds to 100s of milliseconds can be attributed to the impacts of TCP, and specifically to the timeout and retransmission policies on the packet transmission times as has been noted for the wired networks [10], [12], [22]. To verify that this part of the tail is indeed due to TCP related causes, in Fig. 14(b) we plot the interarrival times for the same topology with the transport protocol changed to constant bit rate (CBR) UDP sources of rate 40 kb/s. We see that with CBR UDP sources, the tails of the CCDF plots drop sharply much earlier as there are no

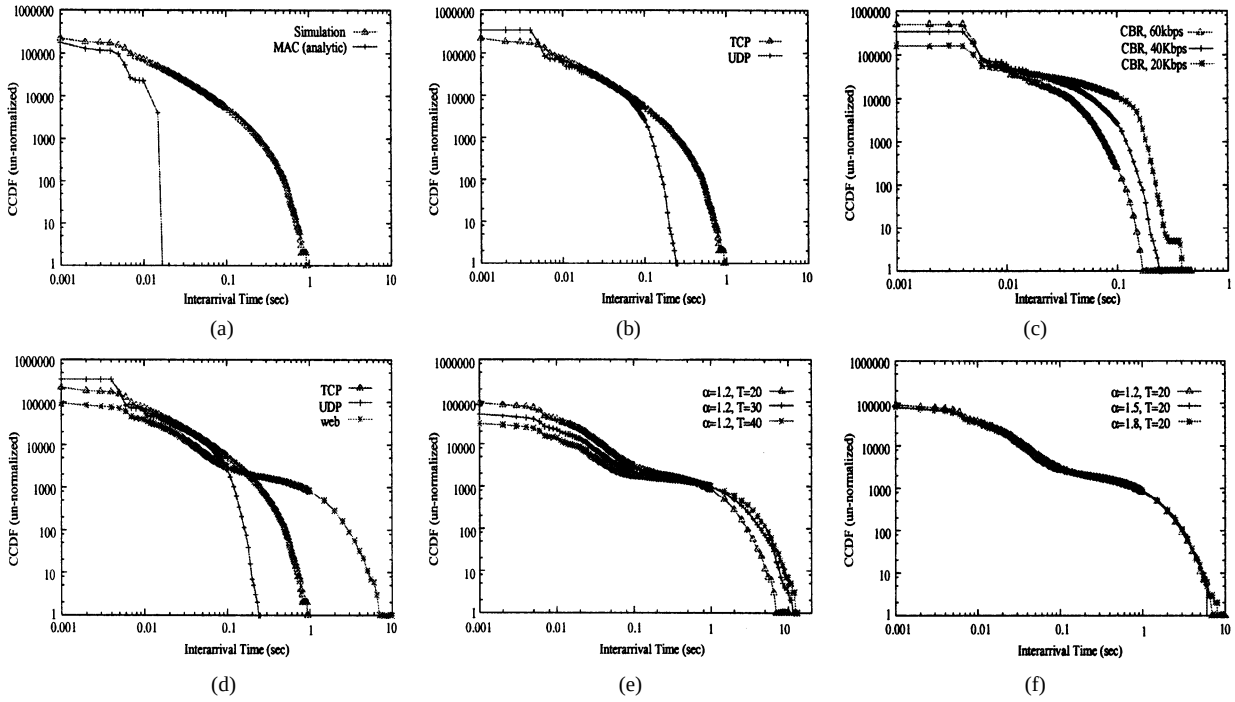


Fig. 14. Power laws in the interarrival times at the intermediate nodes (α = shape parameter, T = average think time). (a) MAC versus TCP. (b) TCP versus UDP (40 kb CBR). (c) UDP with different rates. (d) UDP versus TCP versus WEB. (e) WEB with different think times. (f) WEB with different file sizes.

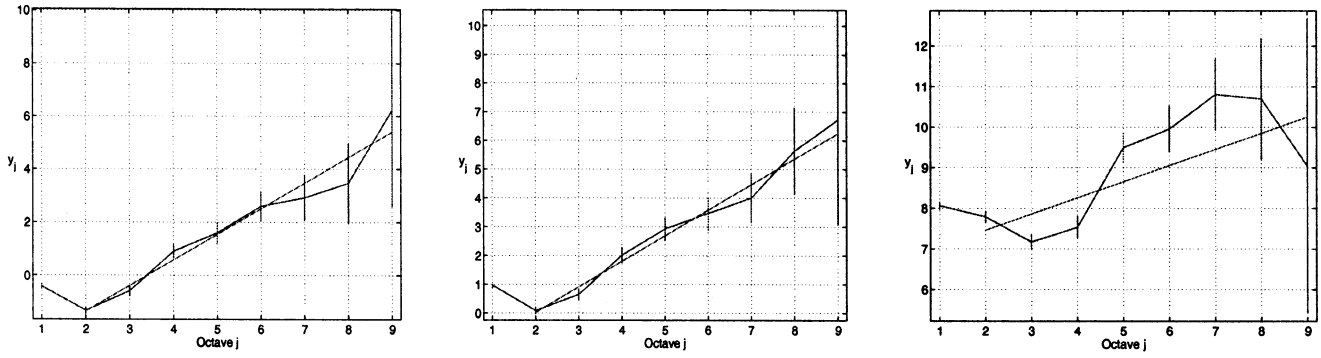
timeouts and exponential backoffs. Since the CBR source continuously produces packets, its rate introduces an upper bound on the time till which a source can remain idle with the interpacket time being inversely proportional to the rate. The effect of this on the interarrival times is visible in Fig. 14(c) where sources with rates 20, 40, and 60 kb are used. As expected, the scenarios with higher rates have quicker fall in the tail of the interarrival times.

With transport protocol related causes, the time scales over which we can observe linear regions in the log–log CCDF plots is limited to a few seconds. However, it is still possible to observe the linear behavior at higher time scales due to application and human effects. Heavy tailed file sizes and human think times, which are inherently larger time scale phenomena, have been pointed out as possible contributors to the self-similarity of network traffic [4]. In Fig. 14(d), we show the CCDF of the interarrival times when the nodes in the network transfer web traffic and compare it with the TCP and UDP cases. We see that with web traffic, the tails of the CCDF extend further, primarily due to the impact of interpage request times (i.e., the human think times). The web traffic was generated using the specifications defined in [9]. In these simulations, each source requests a number of pages from the server (with exponentially distributed interpage request times representing human think times). Each page contains a number of objects with interobject request times being exponentially distributed with mean 0.01 s. The size of each object is drawn from a heavy tailed (Pareto) distribution with an average size of 10 kb and various shape parameters.

The impact of different think times is shown in Fig. 14(e) where we plot results corresponding to average interpage times of 20, 30, and 40 s. As expected, with larger average interpage times, the tails move to the right. The impact of the degree of heavy tail in the file sizes is shown in Fig. 14(f) where we plot the curves for shape parameters of 1.2, 1.5, and 1.8 for the Pareto file size distribution. It is interesting to see that, contrary to expectations, the degree of heavy tail in the file size does not affect the interarrival times. This is because the file sizes only determine the durations for which TCP flows are active. While the flows are active, only the MAC and TCP related causes control the interarrival times. Also, we note three distinct regions in the CCDF plots in Fig. 14(e) and (f), each characterized by a different slope. The difference in the slope results from the fact that in each of these regions, either the MAC or the transport or the application level causes are the dominant contributors.

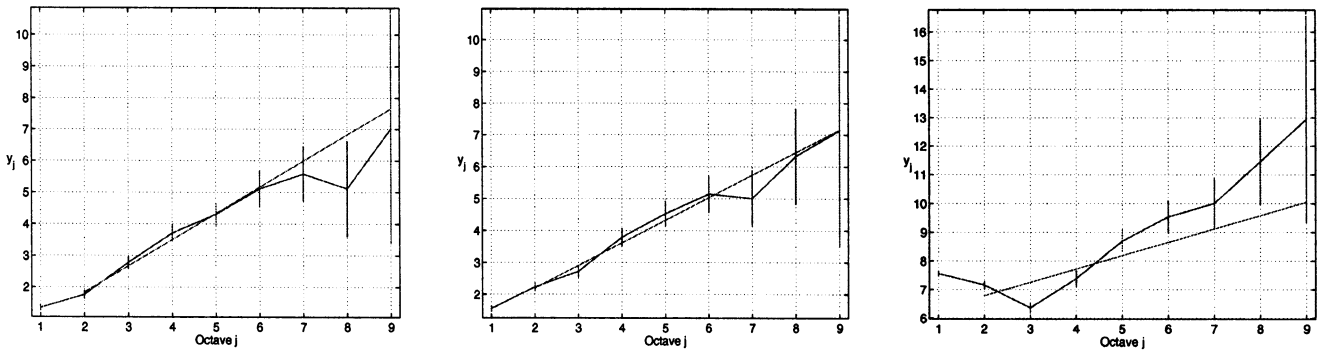
Thus, the packet interarrival times are governed by a host of factors and it is important to understand the exact impact of each on the characteristics and the time scales they operate on. Depending on the buffer size of the system under consideration, it is then possible to evaluate its critical time scale [11]. This time scale then determines which of the three layers plays the most crucial role in the system's performance.

5) Second Order Scaling: For the topology of Fig. 7(a) with 40 nodes, Fig. 15(a) and (b) shows the wavelet plots for two randomly selected nodes and plot 15(c) shows the wavelet plot for the intermediate node. These graphs plot y_j (as defined in Section III-D) as a function of the octave scale j and if the traffic is self-similar, it is manifested in these plots in the form of a straight line. We note that while the individual sources show linear scaling consistent with long range dependence and self-similarity, the bottleneck has a more complex behavior and cannot be considered self-similar. The nonlinear behavior of the energy in each time scale suggest a behavior more consistent with multifractal processes. Similar results are obtained for the 20-source simulation and the 10-source simulation (Fig. 16). The values of the Hurst parameter or the degree of self-similarity, as calculated by the wavelet method are tabulated in Table II. We also note that for most of the wavelet plots, the



(a) (b) (c)

Fig. 15. Second order scaling at various sources and the base station (40 nodes). (a) Source node 27. (b) Source node 33. (c) Base station.



(a) (b) (c)

Fig. 16. Second order scaling at various sources and the base station (10 nodes). (a) Source node 7. (b) Source node 9. (c) Base station.

curves do not increase linearly at higher or lower octaves which might indicate that the self-similarity is not expressed at all time scales (i.e., the process is not truely self-similar).

Note that the absence of self-similarity at the base stations or intermediate nodes is in sharp contrast to the results at routers or wired LANs [18], [30]. We conjecture that this is at least partially due to the pacing effects of the 802.11 MAC and the resulting multimodal interarrival time distribution. The presence of different modes in the traffic also contributes to the varying scaling exponents at different time scales giving a more multifractal feel to the traffic. The almost self-similar behavior of the individual TCP sources is consistent with similar observations in wired networks [27], [28]. This suggests that the TCP related dynamics dominate the source traffic characteristics while at the base station or intermediate nodes, 802.11 MAC has a dominant role to play.

VI. Conclusion

In this paper, we studied the implications of IEEE 802.11 MAC protocol on the traffic characteristics of a wireless network. Our study concentrated on two characteristics: the interarrival time distribution and the second order scaling behavior of the traffic. We also developed an analytic model to characterize the interarrival time distribution of packets arriving at the base station of a wireless BSS or the intermediate nodes acting as routers in ad hoc networks. Our model was validated using extensive *ns-2* simulations.

The results of this paper provide a new insight into the behavior of wireless traffic and can be used to develop more accurate queuing models, traffic schedulers for achieving quality of service, admission control policies, etc. The paper also presents a study of the fractal properties of wireless traffic, critical for understanding the behavior of traffic across all time scales and necessary for accurate prediction of network performance. We also identify the critical timescales over which protocols at different layers affect the traffic characteristics. These timescales determine the buffer requirements and the corresponding delays and loss rates introduced in the system. The paper also studies the complex interactions between the various layers of the protocol stack and their overall impact on the traffic characteristics.

TABLE II
H Values From Wavelet Method

40 Nodes		20 Nodes		10 Nodes	
Node	H	Node	H	Node	H
3	0.991	3	0.929	2	0.957
14	0.993	7	0.921	4	0.953
27	0.983	13	0.889	7	0.916
33	0.945	17	0.946	9	0.854

The main result of our analytic model and the simulation results in its support is to show the pacing of traffic induced by the 802.11 MAC protocol. The resulting interarrival times have a multimodal distribution with the position of the peaks determined by the transmission times associated with packets of different sizes. This has serious implications for designing interface buffers at wireless nodes and traditional traffic models can no longer be used to evaluate performance. Our model provides a basis for further performance studies and accurate analysis of 802.11 based networks. Also, the region of impact of the MAC is

till the time scales of a few tens of milliseconds while transport or application layer dominate the characteristics beyond that.

This paper also investigated the presence of heavy tails in the interarrival times and the contribution of the MAC, transport and application layer protocols to such behavior. We showed that while the small time scale (up to few 10s of ms) behavior of the interarrivals is governed by the MAC, moderate (10s to 100s of ms) and large (hundreds of millisecond and higher) time scale behavior is dominated by the transport and application layer protocols, respectively. An interesting observation here was that in scenarios with web traffic transfer, the degree of heavy tailedness of the distribution used to generate file sizes did not have any effect on the interarrival time distribution.

The second part of our study focuses on the second order scaling of TCP traffic in 802.11 networks. Our simulation results show that while the traffic patterns at the individual sources are more consistent with long-range dependence and self-similarity, in contrast to wired networks, the aggregate traffic is not self-similar. The aggregate traffic is better classified as a multifractal process and we conjecture that the various peaks of the multimodal interarrival time distribution have a direct contribution to the differing scaling exponents at various timescales.

REFERENCES

[1] P. Abry and D. Veitch, "Wavelet analysis of long range dependent traffic," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2–15, Jan. 1998.

[2] F. Cali, M. Conti, and E. Gregori, "IEEE 802.11 protocol: Design and performance evaluation of an adaptive backoff mechanism," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1774–1786, Sept. 2000.

[3] H. Chhaya and S. Gupta, "Performance modeling of asynchronous data transfer methods of IEEE 802.11 MAC protocols," *Wireless Networks*, vol. 3, pp. 217–234, 1997.

[4] M. Crovella and A. Bestavros, "Self-similarity in World Wide Web traffic: Evidence and possible causes," *IEEE/ACM Trans. Networking*, vol. 5, pp. 835–846, Dec. 1997.

[5] A. Erramilli, O. Narayan, A. Neidhardt, and I. Sainee, "Performance impacts of multiscaling in wide area TCP/IP traffic," in *Proc.IEEE INFOCOM*, Tel Aviv, Israel, Mar. 2000, pp. 352–359.

[6] A. Erramilli, O. Narayan, and W. Willinger, "Experimental queuing analysis with long-range dependent packet traffic," *IEEE/ACM Trans. Networking*, vol. 4, pp. 209–223, Apr. 1996.

[7] K. Fall and K. Varadhan, Eds., "ns notes and documentation," in *The VINT Project*: Univ. California, Berkely, LBL, USC/ISI, and Xerox PARC, Nov. 1997.

[8] A. Feldman, A. C. Gilbert, and W. Willinger, "Data Networks as cascades: Investigating the multifractal nature of Internet WAN traffic," *Comput. Commun. Rev.*, vol. 28, no. 4, pp. 42–58, 1998.

[9] A. Feldmann, A. C. Gilbert, P. Huang, and W. Willinger, "Dynamics of IP traffic: A study of the role of variability and the impact of control," in *Proc. ACM SIGCOMM*, Boston, MA, Aug. 1999, pp. 301–313.

[10] D. R. Figueiredo, B. Liu, V. Misra, and D. Towsley, "On the autocorrelation structure of TCP traffic," Univ. Massachusetts, Comput. Sci. Dept., Amherst, MA, Tech. Rep. TR 00-55, 2000.

[11] M. Grossglauser and J.-C. Bolot, "On the relevance of long-range dependence in network traffic," *IEEE/ACM Trans. Networking*, vol. 7, pp. 629–640, Oct. 1999.

[12] L. Guo, M. Crovella, and I. Matta, "How does TCP generate pseudo-self-similarity?," in *Proc. Int. Workshop on Modeling, Analysis and Simulation of Computer and Telecommunications Systems*, Cincinnati, OH, Aug. 2001, pp. 215–223.

[13] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std. 802.11, Jan. 1997.

[14] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher Speed Physical Layer (PHY) Extension in the 2.4GHz Band*, IEEE Std. 802.11b/D5.0, Apr. 1999.

[15] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher Speed Physical Layer (PHY) Extension in the 5GHz Band*, IEEE Std. 802.11a/D5.0, Apr. 1999.

[16] M. Jiang, M. Nikolic, S. Hardy, and L. Trajkovic, "Impact of self-similarity on wireless data network performance," in *Proc. IEEE ICC*, 2001, pp. 477–481.

[17] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*, T. Imielinsky and H. Korth, Eds. Norwell, MA: Kluwer, 1996, ch. 5, pp. 153–181.

[18] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic (Extended version)," *IEEE/ACM Trans. Networking*, vol. 2, pp. 1–15, Feb. 1994.

[19] K. Park, G. Kim, and M. Crovella, "On the relationship between file sizes, transport protocols and self-similar network traffic," in *Proc. Int. Conf. Network Protocols*, Columbus, OH, Oct. 1996, pp. 171–180.

[20] V. Paxon and S. Floyd, "Wide area traffic: The failure of Poisson modeling," *IEEE/ACM Trans. Networking*, vol. 3, pp. 226–244, June 1995.

[21] C. E. Perkins and P. Bhagwat, "Highly dynamic Destination Sequenced Distance-Vector routing (DSDV) for mobile computers," in *Proc. ACM SIGCOMM*, Aug. 1994, pp. 234–244.

[22] B. Sikdar and K. S. Vastola, "On the contribution of TCP to the self-similarity of network traffic," in *Evolutionary Trends of the Internet, Springer-Verlag Lecture Notes in Computer Science*, Sept. 2001, pp. 596–613.

[23] H. Takagi, *Analysis of Polling Systems*. Cambridge, MA: MIT Press, 1986.

[24] H. Takagi and L. Kleinrock, "Throughput analysis for persistent CSMA systems," *IEEE Trans. Commun.*, vol. COM-33, pp. 627–638, 1985.

[25] M. S. Taqqu and V. Teverovsky, "On estimating long-range dependence in finite and infinite variance series," in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, R. J. Adler, R. E. Feldman, and M. S. Taqqu, Eds. Boston: Birkhauser, 1998, pp. 177–217.

[26] Y. Tay and K. Chua, "A capacity analysis for the IEEE 802.11 MAC protocol," *Wireless Networks*, vol. 7, pp. 159–171, Mar. 2001.

[27] A. Veres and M. Boda, "The chaotic nature of TCP congestion control," in *Proc. IEEE INFOCOM*, Tel-Aviv, Israel, 2000, pp. 1715–1723.

[28] A. Veres, Z. Kenesi, S. Molnar, and G. Vattay, "On the propagation of long-range dependence in the Internet," in *Proc. ACM SIGCOMM*, Stockholm, Sweden, Sept. 2000, pp. 243–254.

[29] J. Weinmiller, H. Woesner, J.-P. Ebert, and A. Wolisz, "Modified backoff algorithms for DFWMAC's distributed coordination function," in *Proc. 2nd ITG Fachtagung Mobile Kommunikation*, Neu-Ulm, Germany, Sept. 1995, pp. 363–370.

[30] W. Willinger, M. S. Taqqu, R. Sherman, and D. V. Wilson, "Self-similarity through high-variability: Statistical analysis of Ethernet LAN traffic at the source level," *IEEE/ACM Trans. Networking*, vol. 5, pp. 71–86, 1997.

Omesh Tickoo (S'01) received the B.E. degree in electronics and communication from Karnataka Regional Engineering College, Surathkal, India, the M.S. degree in computer and systems engineering from Rensselaer Polytechnic Institute, Troy, NY, and is currently working toward the Ph.D. degree in the Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute.

His research interests include traffic modeling and multimedia streaming over wireless networks.

Biplab Sikdar (S'98–M'02) received the B.Tech. degree in electronics and communication engineering from North Eastern Hill University, Shillong, India, the M.Tech. degree in electrical engineering from Indian Institute of Technology, Kanpur, India, and the Ph.D. in electrical engineering from Rensselaer Polytechnic Institute, Troy, NY, in 1996, 1998, and 2001, respectively.

He is currently an Assistant Professor in the Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute. His research interests include traffic modeling, sensor networks, queuing theory, and switch performance evaluation.

Dr. Sikdar is a Member of Eta Kappa Nu and Tau Beta Pi.