Cross-Layer Design for Medium Access Control in CDMA Ad Hoc Networks

Amit Butala
Qualcomm Inc., San Diego, CA 92121, USA
Email: abutala@qualcomm.com

Lang Tong
School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA
Email: ltong@ee.cornell.edu

Received 7 August 2003; Revised 3 May 2004

A medium access control (MAC) protocol for spread-spectrum ad hoc networks with dynamic channel allocation (DCA) is presented. DCA can support large systems with a smaller number of channels by dynamically assigning channels only when a node has a packet to transmit. The protocol extends cross layer, with the scheduling at the MAC, and assignment of channels at the physical layer by means of a query. It is shown that DCA is collision free under ideal conditions. By assigning channels dynamically, DCA offers improved throughput normalized by available bandwidth. Analytical results are presented for the performance of the query detection and the throughput for a fully connected network.

Keywords and phrases: MAC, dynamic channel allocation, spread spectrum, query, hypothesis detection.

1. INTRODUCTION

There are several challenges in the design of medium access control (MAC) protocol for code division multiple access (CDMA) ad hoc networks. While it is possible to apply single-channel MAC protocols such as MACAW [1], DBTMA [2], and FAMA [3] to a multichannel CDMA network by treating channels independently, such approaches do not exploit the rich diversity of CDMA, nor do they offer an efficient utilization of available spectrum. Specifically, the classical problem of hidden/exposed nodes manifests itself differently in the presence of multiple access channels; multiple data channels and a control channel can coexist using different spreading codes. If the spreading codes have good cross-correlation properties, contention on one channel does not cause interference on the other channels. The selection of a channel, from a set of channels, to transmit upon, however, is an issue that has not been well addressed in literature.

Spread-spectrum protocols were introduced by Sousa and Silvester [4]. Based on the preassignment of codes, these protocols are identified as common transmitter (CT), common receiver (CR), and transmitter-receiver (TR). The CT protocol is the better suited protocol for ad hoc networks since it is less complex and requires a smaller set of spreading codes. In the CT protocol, a node may begin a transmission on the transmitters’s assigned code at any time. As there is no feedback on the status of the node, transmissions may be scheduled to nodes unable to receive. Moreover, an a priori assignment of transmit codes is assumed for all nodes in the network. This requires that the number of spreading codes be equal to the number of nodes in the system and necessitates the use of larger than necessary spreading sequences.

MACA-CT [5] improves on the CT scheme of code allocation by the use of a control sequence over the common channel. Medium access is time slotted. A node sends a request-to-send (RTS) at the beginning of a time slot and is scheduled to transmit data only if the intended receiver acknowledges the request with a corresponding clear-to-send (CTS). This prevents transmissions to busy nodes. Here too, an apriori assignment of transmit codes is assumed for all the nodes in the network.

In CHMA [6], on the other hand, all the nodes follow a common channel-hopping sequence with each hop duration equal to the amount of time needed for nodes to receive the control packet, either an RTS or a CTS, from a neighbor. The RTS-CTS is followed by data transmission on the same channel while all other nodes hop to another channel. CHMA performs better than the other protocols mentioned earlier under ideal circumstances, but a few factors need to be considered. The hopping channel length has to be at least as long as the length of the packet, which can be a significant penalty as the length of the data packets increases within
relatively small neighborhoods. Longer data packets may increase the network throughput but require a larger spreading gain to generate the larger number of spreading codes in the channel-hopping sequence. The problem of bandwidth utilization remains overlooked.

A common drawback in each of the above protocols is the need for large spreading gains, which imposes a severe penalty on the bandwidth utilization.

1.1. Contributions

Detection of signals is an integral part of MAC. All control signalling-based schemes require the detection of RTS and CTS. Protocols such as DBTMA must detect the presence of busy tones. In the presence of multipath fading, such detections cannot be assumed perfect; missed detections and false alarms may have an adverse effect on the protocol performance. Unfortunately, the problem of optimal detection for maximizing MAC throughput has not been considered.

In [7], we proposed a new MAC protocol to tackle the issue of efficient spectral utilization. Referred to as the dynamic channel allocation (DCA), this protocol requires only a fixed number of codes irrespective of the size of the network. Codes are dynamically assigned using a receiver-based request detector.

In this paper, an optimal design of the request detector is presented. Assuming a Rayleigh-fading model, a Neyman-Pearson detector is used with the detection threshold optimized for throughput. In order to perform such an optimization, a Markov chain analysis is used to obtain the relation between the detector level and normalized throughput.

Such a cross-layer design enables us to eliminate the dependence of the spreading gain on the number of nodes in the network and assign channels dynamically.

1.2. Structure

The structure of the paper is as follows. In Section 2, we discuss the model assumed for ad hoc networks. Section 3 elaborates the design of the new protocol and the receiver for DCA using a binary hypothesis model for channel occupancy and a busy tone backoff strategy. In Section 4, we build a Markovian representation of a fully connected ad hoc network. Analytical bounds on the throughput of the network are computed and compared with our implementation of the protocol. The results of comparisons between existing multichannel protocols and DCA are presented in Section 5. Relevant conclusions and foresights into the modeling of ad hoc networks are summarized in Section 6.

2. NETWORK MODEL

Consider a hypothetical multihop network as shown in Figure 1. We use the protocol model definition for the neighborhood of a node. Thus, each node within a fixed radius \( R \) of the transmitter is assumed to be contained in its neighborhood and can listen to the transmitter. The relationship is dual; a node is not affected by any transmission that originates outside its neighborhood. It is assumed that all the nodes transmit with a fixed transmit power.

The network consists of \( N \) nodes spatially distributed. Not all nodes are able to communicate with each other. The coverage areas for the nodes are represented by the circles centered at the respective nodes. Clearly, transmissions from A to B have to resolve potential contention with nodes C and E.

We assume \( M + 1 \) distinct spreading codes available for transmission where \( M \) may be less than \( N \). The codes are designed with good correlation properties [8] such that transmissions using one code do not destroy reception on any of the other codes. As mentioned before, each code identifies a unique channel.

One of the channels is reserved for transmission of control sequences while the other \( M \) channels can support the data packets. Each node makes a choice of transmitting to a node in its neighborhood on any one of the \( M \) data channels. Issues related to routing are not considered. It is assumed that either the nodes know the routing tables a priori or the range of communication involves only neighboring nodes.

Nodes are half duplex and can tune to only one channel at any given time. In addition, nodes also have a frequency generator/receiver that may be used to transmit/receive a monotone on a preset frequency. This is used to specify a busy signal during packet reception.

Transmission time is slotted and the transmissions are packet synchronized. The data is broken up into minipackets that are transmitted in succession, with each minipacket requiring one time slot. The RTS and the CTS packets are assumed to be less than one half minipacket in length such that an RTS-CTS packet exchange between any two nodes in the network may be completed in a single minislot.

2.1. Normalized throughput

Since the number of channels in the system that satisfy the constraints on multiaccess interference is proportional to the spreading gain, the absolute performance cannot be inferred simply by observing the raw network throughput. The network throughput is expected to increase with an increase in spreading gain, and hence we introduce the concept of normalized throughput for comparison of different protocols.

The network throughput \( \Gamma \) is defined as the average number of packets successfully received in one time slot over the network when being in steady state. The spreading gain \( G \) is the ratio of the chip rate to the symbol rate of a spread-spectrum signal. Then the normalized throughput \( \eta \) can be defined as the ratio of the network throughput to the
spreading gain:

\[ \eta = \frac{\Gamma}{G}. \]  

Multiaccess interference can be largely eliminated if the codes are orthogonal to each other. In such codes, such as Walsh codes, the spreading gain is equal to the total number of channels available to the system. The normalized throughput would then be the ratio of the network throughput to the total number of channels. This metric is used in all subsequent discussions to compare protocol efficiencies.

3. DYNAMIC CHANNEL ALLOCATION

Fixed channel allocation schemes discussed so far increase the number of channels required in accordance with either the size of the network or the length of the data packet. A demand-driven dynamic allocation of channels is proposed as one solution for overcoming this constraint. DCA relies on the assignment of one of the available data channels to the nodes that get scheduled to transmit. Thus, the two basic requirements for packet exchange are scheduling of packets and allocation of channel.

Scheduling. For a successful transmission, there should be only one transmitter attempting to transmit to a node, and any such transmission must be destined to an idle node. This is effected by the transmission of the RTS-CTS on the control channel. Since the channel is a collision channel and multiple transmissions on the same channel result in packet collision, the RTS-CTS ensures proper scheduling of the transmissions.

Allocation. Given that two terminals are scheduled, there must be a channel available for transmission that does not interfere with any ongoing transmission.

This is effected by a new procedure called querying of channels.

3.1. Querying of channels

The RTS-CTS control packet exchange establishes the scheduling of packets over a particular channel, but it does not ascertain the availability of the channel. A channel is said to be available only if no node in the neighborhood of the intended receiver is transmitting on that channel, and no other node in the neighborhood of the intended transmitter is receiving on that channel. These are, respectively, the conventional exposed terminal and hidden terminal problems that need to be addressed in ad hoc networks. Thus, in our figure for a typical ad hoc network (Figure 1), node A may transmit to a node B on a specific channel L only if node C (the hidden node) is not transmitting on channel L and node E (the exposed node) is not receiving on channel L.

Overcoming the exposed terminal problem necessitates a response from other nodes in the neighborhood (i.e., E) if it is receiving data on the same channel. The hidden terminal problem necessitates a response from B to contention due to transmission from any node in the neighborhood of B from which A might be hidden (i.e., C).

The solution is the transmission of a query by the intended transmitter, A. The query is a known data packet and thus is a deterministic interference that may be estimated. Once a data transmission is scheduled using the RTS-CTS exchange, the transmitter sends out the query on the selected channel.

In response to contention, if any, caused by the query, the receiver transmits a busy tone. The busy tone is a sinusoid sent on an out-of-band frequency and intimates the transmitter that the channel is in use. A query is successful only if no busy tone is heard by the transmitter. This represents the case that no exposed terminal is receiving, and no hidden terminal is transmitting, on the selected channel. A node may transmit only if its query is successful.

With the introduction of the query, in each time slot, all the nodes may be classified into the following four states.

1. Idle (or backlogged) state: nodes that are not engaged in packet reception or transmission.
2. Query state: nodes that get scheduled and are transmitting the query in the current time slot.
3. Data state: nodes involved in transmission or reception of data packets. Only nodes in the data state successfully transmit data over the network.
4. Locked state: an extra state that tracks nodes involved in data packet collisions. This occurs due to a misdetection of the query and will be discussed in more detail in the next section.

3.2. The protocol

The DCA protocol is defined below and has been illustrated in Figures 2, 3, and 4.

1. Any idle node (e.g., A) that has a packet to transmit to any of its immediate neighbors (e.g., B) attempts to establish a communication by broadcasting an RTS on the common channel at the beginning of the minislot (Figures 3 and 4).
2. The RTS contains the following information: the destination node (B) identifier, the transmitting node (A) identifier, and the selected channel (Q) on which the data will follow. The channel Q is randomly chosen from the set of available channels.
3. If the destination node B receives the RTS, it responds immediately in the same time slot, with a CTS on the common channel (Figures 3 and 4). B transitions from the idle state to the query state for the next time slot and tunes its receiver to the selected channel Q.
4. If A does not receive a CTS in the same time slot, it times out and reverts back to the idle state. A retransmission is attempted according to the backoff strategy. If A does receive the CTS, it moves from the idle state to the query state. This completes the scheduling.
5. In the next minislot, A transmits a query on the selected channel. The query is successful if no busy tone is generated (Figure 3).
Figure 2: query for different network states: (a) success, (b) failure, and (c) failure.

Figure 3: Successful querying: case (a).

Figure 4: Failed querying: cases (b) and (c).

Table 1: DCA algorithm.

<table>
<thead>
<tr>
<th>Data state: transmitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>T2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>T3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Idle nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>T0</td>
</tr>
<tr>
<td>T1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>T2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>T3</td>
</tr>
</tbody>
</table>

3.3. Detection of the query

In the presence of noise and multiaccess interference, the detection of the query is not perfect and is contingent on the operating characteristics of the receivers. At every receiver, the interference due to the query may be missed or a false alarm may be raised in response to a query that does not interfere. This results in a probabilistic model for the acceptance of the query.
Cross-Layer Design for MAC in CDMA Ad Hoc Networks

Missed detection. In the case of a missed detection of the query, there will now be two (or more) nodes transmitting on the same channel within the vicinity of the receiver. This results in a packet collision at the receiver and it is unable to detect either packet. The receiver and the corresponding transmitters are assumed to be in the locked state. The throughput of node pairs in the locked state is zero.

Transition of node pairs out of the locked state would depend on the coding scheme used and the higher-layer scheduling. Without imposing any additional constraints, we assume that the pair remains in the locked state till the end of the current data-packet transmission, after which they reset again.

False alarm. The false alarm induces less damage, since it merely results in the node (i.e., A) aborting the transmission and reverting back to the idle state. A retransmission is attempted in accordance with the protocol. This too would lead to a decrease in the throughput of the network.

The two parameters are related, thus the optimization of the throughput requires an analysis of the receiver operating characteristics (ROC). Two types of nodes need to process a query:

1. data state: nodes currently in reception;
2. query state: nodes attempting to tune to the transmitter.

3.3.1 Detection of the query during the data state

We make the following assumptions:

(1) Each minipacket has a fixed packet size of \( K \) bits.
(2) A header of pilot training bits \( (κ ≪ K) \) is embedded in each data packet to aid channel estimation and timing synchronization.
(3) The total number of data channels is \( M \).
(4) The channel undergoes slow Rayleigh fading. The amplitude of the fade \( (A) \) can be assumed complex, circularly Gaussian, and constant over one time slot: \( A \sim \mathcal{CN}(0, \phi^2) \).

Then, for any particular receiver in the data state, the received signal can be written as [9]

\[
y(t) = \sum_{m=1}^{M} \sum_{k=0}^{K-1} A_mb_m[k]s_m(t - kT - \tau_m) + \eta(t),
\]

where \( A_m \) denotes the signal power on the \( m \)th channel, \( b_m[k] \) denotes the \( k \)th bit on the \( m \)th channel, \( s_m(\cdot) \) is the signature waveform of the \( m \)th channel, \( \tau_m \) is the timing offset of the \( m \)th channel, and \( \eta(t) \) is the additive white Gaussian noise (AWGN) at time \( t \).

Let the receiver be tuned into some channel \( L \). Transmissions on all other channels is a secondary interference and under the interference model assumed is treated as AWGN. Transmissions on the same channel, however, cannot be ignored.

By the definition of the protocol, no other data transmission can be on the same channel as long as the receiver is in the data state. Thus, the primary interference, if it exists, is due to the transmission of a query. Let the query be transmitted on channel \( Q \) which may or may not be the same as \( L \). We use \( \delta_{QL} \) to denote the interference of the query at the receiver. Thus,

\[
\delta_{QL} = \begin{cases} 1 & \text{if query is present in the current slot} \\ 0 & \text{and transmitted on channel } L, \\ \text{otherwise.} \end{cases}
\]

Then, the received signal at the output of a matched filter that is synchronized to channel \( L \) can be represented as

\[
y[k] = A_Lb_L[k] + A_Q[b_Q[k]\rho_{QL}(\tau) + b_Q[k+1]\rho_{QL}(\tau)]\delta_{QL} + n[k] \quad \forall k = 1, \ldots, K,
\]

where \( \rho_{QL} \) and \( \rho_{QL} \) are the cross-correlation between the channels \( L \) and \( Q \) on the interval over which the bits \( b_Q[k] \) and \( b_Q[k+1] \), respectively, overlap bit \( b_L(k) \), and \( n[k] \) is the filtered output of the secondary interference and the noise in the \( k \)th bit.

Detection of the query is a binary hypothesis testing problem, hence for simplicity of the receiver, we set all the bits in the query to 1, that is, \( b_Q[k] = 1 \). Also, assuming good correlation properties on the channels, the output signal at any receiver in the data state is

\[
y[k] = A_Lb_L[k] + A_Qn[k].
\]

Thus, two hypotheses can be formulated as below.

(i) The null hypothesis \( (H_0) \): the query is not on the same channel \( (\delta_{QL} = 0) \):

\[
H_0 : y[k] = A_Lb_L[k] + n[k] \quad \forall k = 1, \ldots, K.
\]

(ii) The alternative hypothesis \( (H_1) \): the query is sent on the same channel as the data packet \( (\delta_{QL} = 1) \):

\[
H_1 : y[k] = A_Lb_L[k] + A_Qn[k] \quad \forall k = 1, \ldots, K.
\]

We assume that in the presence of a slow block fading channel, a node in the data state has already estimated the channel fade on the data link \((A_L)\). The fading on the query \((A_Q)\) is also a constant but cannot be assumed to be known by the receiver. Then, for the duration of the pilot training symbols in each packet, we can define a metric \( \tilde{y} \) as given below:

\[
\tilde{y} = \sum_{k=0}^{K} (y[k] - A_Lb_L[k]).
\]
This simplifies our hypotheses (6) and (7) as given:

\[ H_0 : \tilde{y} = \sum_{k=0}^{\infty} n[k] \Rightarrow \tilde{y} \sim N(0, \kappa \sigma^2), \]

\[ H_1 : \tilde{y} = \sum_{k=0}^{\infty} (A_Q + n[k]) \Rightarrow \tilde{y} \sim N(0, \kappa^2 \phi^2 + \kappa \sigma^2). \]  

This is a standard energy detector problem. For an \( \alpha \)-level receiver (i.e., probability of false alarm \( P_a = \alpha \)), hypothesis \( H_1 \) is selected by the Neyman-Pearson detector if

\[ |\tilde{y}|^2 > \kappa^2 \left[ Q^{-1} \left( \frac{P_a}{2} \right) \right]^2. \]  

If the signal-to-noise ratio (SNR) is known, the power of the detector is given by

\[ P_D = 2Q \left[ \frac{Q^{-1}(P_a/2)}{\sqrt{\kappa \text{SNR} + 1}} \right]. \]  

### 3.3.2. Detection of the query during the query state

For a receiver in the query state, the queried channel is rejected if any of the neighboring nodes transmit on the same channel synchronized with it. Analogous to (5), the model of the received signal at the receiver in the query state is

\[ y[k] = A_Q + A_L b_L[k] \delta_{Q,L} + n[k] \quad \forall k = 1, \ldots, K. \]

Over the interval of the query, the channel \( A_Q \) is a constant and known at the receiver. Thus, the signal error over the interval of the pilot training bits is

\[ \tilde{y} = y - A_L b_L. \]

The binary hypothesis thus simplifies as

\[ H_0 : \tilde{y} = n \sim N(0, \sigma^2 I), \]

\[ H_1 : \tilde{y} = A_L b_L + n \sim N(0, \phi^2 b_L b_L^T + \sigma^2 I), \]

which again is differentiable only in one of the singular-value components. This yields exactly the same detector from the previous part.

### 3.4. Selection of the threshold for the detector

For the Neyman-Pearson detector, the threshold of the detector affects the probability of missed detection. We assume that all the nodes use the same detector operating at a threshold of detection. Thus, the probability of false alarm and missed detection are constant over the network and known a priori.

The throughput is a function of both the parameters, hence the optimal value of the threshold is the one that maximizes this throughput over the ROC of the detector.
of the query, we model the performance based on the ROC from state $l_{mn}$ by Cross-Layer Design for MAC in CDMA Ad Hoc Networks.

Since the system is affected by the detection probability of the query, we model the performance based on the ROC of the query detector, namely, the probability of false alarm ($P_a$) and the probability of missed detection ($P_b$).

### 4.1. The state transition probabilities

The Markov chain is completely described by the number of nodes in each state. Given that the total number of nodes in the network is $N$, we describe each state by the three identifiers described above, namely $l, m, n$. Consider the transition from state $l_{mn}$ to state $xyz$:

$$P_{l_{mn},xyz} = P(x, y, z|l, m, n) = P(x|l, m, n)P(y, z|x|l, m, n)$$

where step three follows from the knowledge that the number of nodes in the query state is determined only by the state of the network in the preceding slot, or more precisely, only on the number of idle nodes in the previous slot.

#### 4.1.1. Computation of $P(x|l, m, n)$

A node pair reaches the query state if the RTS/CTS communication is successful. In a fully connected network, since a maximum of one RTS can be successful in a time slot, the CTS can be granted to be always successful. An RTS is assumed successful if it is transmitted to an idle node. Let this event be denoted by $l$. Under this assumption, we compute $P(x|l, m, n)$, which represents the probability of an RTS/CTS exchange succeeding in the current time slot. A successful RTS/CTS exchange implies a query is attempted in the next time slot, that is, $x = 1$:

$$P(x = 1|l, m, n) = P(\text{one RTS is transmitted } \cap l) = B(N', p, 1)N' - 1 \frac{N - 1}{N - 1}$$

where $B(n, p, k) = \binom{n}{k} p^k (1 - p)^{n-k}$ is the binomial distribution of selection of $k$ from a set of $n$ when each individual probability of selection is $p$.

The probability of no success, that is, $x = 0$, is

$$P(x = 0|l, m, n) = 1 - P(x = 1|l, m, n).$$

#### 4.1.2. Computation of $P(y, z|l, m, n)$

Each time slot can be classified on the basis of the occurrence of four events.

1. Query ($Q$): this corresponds to the event that a node transmits a query packet over one of the data channels.
2. Interference ($I$): this corresponds to the event that the query transmitted is on the same channel as that of one of the data transmissions.
3. Missed detection ($M$): this is the event that the interference of the query on the data transmission is missed by the receiving node.
4. False alarm ($F$): this is the event that any one of the receivers in the data state raises the busy tone even though the query is not transmitted on the channel it is tuned to.

We compute the transitions conditioned on the presence of the query.

If $l = 0$, no query was sent in the previous time slot, and no new node pair starts transmitting. Assume $i$ node pairs in the data state and $j$ node pairs in the locked state complete the transmission, and thus revert back to idle.

$$P(y, z|l = 0, m, n) = \sum_{i=0}^{m} \binom{m}{i} P(\text{i data pairs become idle}) \sum_{j=0}^{n} P(\text{j locked pairs become idle}),$$

$$P(y, z|l = 0, m, n, i, j) = \sum_{i=0}^{m} \sum_{j=0}^{n} B(m, 1 - q, i)B(n, 1 - q, j) \delta(y - (m - i)) \delta(z - (n - j))$$

where $B(n, p, k)$ is the binomial distribution and $\delta(0) = 1$, and $\delta(x) = 0$ for all $x \neq 0$ represent the acceptable state transitions.

If $l = 1$, a query was sent in the previous slot. There are four outcomes for the query: success, interference detection, false alarm, or missed detection. The probability of successfully establishing a data channel depends on the number of available channels. Let $\psi_D$ node pairs be added to the data state and $\psi_L$ node pairs end up in the locked state in the given time slot:

$$P(y, z|l = 1, m, n) = \sum_{i=0}^{m} \sum_{j=0}^{n} P(\text{i data pairs become idle}) \sum_{j=0}^{n} P(\text{j locked pairs become idle}),$$

$$P(y, z|l = 1, m, n, i, j) = \sum_{i=0}^{m} \sum_{j=0}^{n} B(m, 1 - q, i)B(n, 1 - q, j) \times \sum_{j=0}^{n} P(\psi_D, \psi_L) \delta(y - (m - i + \psi_D)) \delta(z - (n - j + \psi_L)).$$
In the case of a missed detection, there are two nodes transmitting on the same channel within the vicinity of the receiver. This results in a packet collision at the receiver and it is unable to detect either packet. The receiver and the corresponding transmitted are assumed to be in the locked state for the duration of the transmission.

Since our model of the state space does not carry the information about the channel that gets assigned to the transmitter, this case needs to be tackled independently of the knowledge of the number of node pairs involved in the packet collision.

4.1.3. The upper bound

Since the nodes are half duplex, there can be no feedback from the receiver to the transmitter. An upper bound can be constructed under the assumption that a "Genie" informs the transmitter involved in a missed detection, in which case they immediately stop transmitting. In other words, a missed detection causes nodes to move to the idle state instead of the locked state. Hence, $z = n = 0$ always.

Only one pair of nodes can be in the query state in any time slot. On the basis of the occurrence of the above four events, it is clear that $\mathcal{M} \subset \mathcal{I} \subset \mathcal{Q}$. Also, since the network is fully connected, false alarms that are on another channel would also cause the query to fail. Thus, $\mathcal{F}$ and $\mathcal{Q}$ can be considered as independently occurring events (Figure 6).

(i) A query corresponds to the combination of events

$$\{Q \cap F c \cap I c\} \implies (\psi_D = 1, \psi_L = 0). \quad (20)$$

(ii) False alarm and interference detection both result in the generation of a busy tone and the query fails. This is the combination of events

$$\{(Q \cap F c \cap I \cap M c) \cup (Q \cap F)\} \implies (\psi_D = 0, \psi_L = 0). \quad (21)$$

(iii) Missed detection is the event set

$$\{Q \cap F c \cap I \cap M\} \implies (\psi_D = -1, \psi_L = 0). \quad (22)$$

Conditioned on the arrival of the query, the probabilities for false alarm, missed detection, interference are

$$P_F = P\left\{\text{false alarm at least one data receiver}\right\} = 1 - (1 - P_a)^{m-i-1},$$

$$P_M = P\left\{\text{missed detection at contended data receiver}\right\} = P_pP_b,$$

$$P_I = P\{\text{nonidle channel selected}\} = \frac{m-i}{M}. \quad (23)$$

Thus, we have, from (19),

$$P(y, z|l = 1, m, n)$$

$$= \sum_{i=0}^{m} B(m, 1-q,i)$$

$$\times[\delta(y - (m-i+1))P(\psi_D = 1, \psi_L = 0)$$

$$+ \delta(y - (m-i))P(\psi_D = 0, \psi_L = 0)$$

$$+ \delta(y - (m-i-1))P(\psi_D = -1, \psi_L = 0)]$$

$$= \sum_{i=0}^{m} B(m, 1-q,i)$$

$$\times[\delta(y - (m-i+1))(1 - P_F)(1 - P_I)$$

$$+ \delta(y - (m-i))\{(1 - P_F) P_I (1 - P_M) + P_F\}$$

$$+ \delta(y - (m-i-1))(1 - P_F) P_I P_M]. \quad (24)$$

4.1.4. The lower bound

In the absence of feedback from the “Genie,” when a collision occurs, the transmitter does not stop transmitting. The node-pair transitions to the locked state are unavailable until the transmitter has completed its transmission.

In addition, since the state space does not carry the information of which channel the locked transmitter is transmitting upon, we assume that every locked node pair occupies a different channel. Clearly, this is a very conservative estimate and provides us with a lower bound for the system in the presence of missed detection.

Again there are four outcomes for the query: success, interference detection, false alarm, and missed detection. Missed detection causes transition of node pairs from the data state to the locked state. For receivers in the locked state, since there are multiple simultaneous transmissions on the same channel, the interference is nondeterministic. The hypothesis detector fails to identify the query sent on the channel. Hence, receivers in the locked state do not raise a busy flag, irrespective of the contention.
Thus, depending on whether the missed detection was with a node pair in the data state or already in the locked state, \((\psi_D, \psi_L) = (-1, 2)\) or \((0, 1)\). Using \(I_T\) to indicate that the interference was with a channel assigned to a data node pair and \(I_L\) to indicate interference with a locked node pair, we have the following (Figure 7):

(i) A query success corresponds to the events
\[
\{Q \cap F^c \cap I_T \cap I_L^c\} \implies (\psi_D = 1, \psi_L = 0).
\]  

(ii) False alarm and interference detection both result in the generation of a busy tone and the query fails. This is the combination of events
\[
\{(Q \cap F \cap I_T \cap M') \cup (Q \cap F)\} \implies (\psi_D = 0, \psi_L = 0).
\]  

(iii) A missed detection involving a channel assigned to a node pair in the transmit state is the event
\[
\{Q \cap F^c \cap I_T \cap M\} \implies (\psi_D = -1, \psi_L = 2).
\]  

(iv) A missed detection when the channel chosen is in the locked state is the event
\[
\{Q \cap F^c \cap I_T^c\} \implies (\psi_D = 0, \psi_L = 1).
\]

Using the above, (19) simplifies as follows:
\[
P(y, z | l = 1, m, n) = \sum_{i=0}^{m} \sum_{j=0}^{n} B(m, 1 - q, i)B(n, 1 - q, j) \times \left[ \delta(y - (m - i + 1))\delta(z - (n - j)) \times (1 - P_F)(1 - P_{T_{y,1}}) \right.
\]
\[
+ \delta(y - (m - i))\delta(z - (n - j)) \times \left\{ (1 - P_F)P_{T_{y,1}}(1 - P_M) + P_F \right\}
\]
\[
+ \delta(y - (m - i - 1))\delta(z - (n - j + 2)) \times (1 - P_F)P_{T_{y,1}}P_M
\]
\[
+ \delta(y - (m - i))\delta(z - (n - j + 1)) \times (1 - P_F)P_{T_{y,1}} \bigg].
\]  

Using the above, (25) simplifies as follows:
\[
\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} B(m, 1 - q, i)B(n, 1 - q, j) \times \left[ \delta(y - (m - i + 1))\delta(z - (n - j)) \times (1 - P_F)(1 - P_{T_{y,1}}) \right.
\]
\[
+ \delta(y - (m - i))\delta(z - (n - j)) \times \left\{ (1 - P_F)P_{T_{y,1}}(1 - P_M) + P_F \right\}
\]
\[
+ \delta(y - (m - i - 1))\delta(z - (n - j + 2)) \times (1 - P_F)P_{T_{y,1}}P_M
\]
\[
+ \delta(y - (m - i))\delta(z - (n - j + 1)) \times (1 - P_F)P_{T_{y,1}} \bigg]
\]  

5. RESULTS AND SIMULATIONS

The maximum throughput of DCA is prone to the operating characteristics of the detector for the query. Peak throughput depends on both the probability of false alarm as well as missed detection. Each receiver may pick up its operating point based on its individual requirements. For simplicity, however, we assume that all receivers operate at the same point on the ROC. The throughput then relates to the ROC as shown earlier in Figure 5. Once the system SNR is computed at the receiver, the threshold of the detector is set at the point on the ROC curve that maximizes the throughput.

A comparison of the three schemes discussed earlier; MACA-CT, CHMA, and DCA, is made in Figure 8 for a fully connected 20-node network carrying data packets geometrically distributed in length and with a mean length of 10 minipackets. The number of data channels depends on the protocol. For DCA, we randomly choose 5 data channels. MACA-CT has 20 channels, determined by the size of the network. For CHMA, this number would have to be greater than the largest data packet in the network, which is infinity. We compare against the normalized throughput of modified CT, which illustrates the best case performance of CHMA for a channel-hopping sequence that is twice as long as the length of the average data-packet length (see Appendices B and C).

The normalized throughput of the 3 protocols are plotted below. The query detector for DCA is assumed to be operating at 2 dB SNR.

Figure 9 shows the maximum normalized throughput of DCA, at various SNR levels, compared with that of MACA-CT and CHMA. Significant performance gains are observed for the parameters indicated. Clearly, the scheme performs uniformly better for any probability of transmission and channel interference over the given bandwidth expansion available.

5.1. Parameter selection

The efficiency of the protocol depends upon the length of the data packets and the number of data channels. Increasing the number of channels increases the success rate of the query and thus the overall throughput per slot. However, this would also require an increase in the spreading gain, thus wiping out the advantages of DCA. Increasing the length of the data packet should increase the protocol efficiency by reducing the fraction of the number of control packets per packet of data. At the same time, larger data packets are more prone to collisions which would result in the data channel being locked for longer intervals. Clearly, there are tradeoffs involved in selection of both parameters.
5.1.1. Performance as a function of the number of data channels

From Figure 10, the normalized throughput appears to be almost monotonically decreasing beyond the addition of the first few channels. The best case performance is for systems with 2 to 4 data channels. The results are not totally surprising since one might expect the control channel to be the bottleneck as more channels are made available for data. Increasing the available number of channels does not yield to a proportionate increase in data traffic. Interestingly, performance
of DCA is superior till the number of channels equals 8. Fixed channel allocation schemes would yield better throughput than DCA if more channels might be made available.

5.1.2. Performance as a function of the length of the data packet

As seen from Figure 11, the throughput increases with an increase in the length of the data packet and then drops off. Again, this is not unexpected since longer data streams are more likely to be involved in missed detections of the query and result in locked states.

This seems to suggest that the average data packet should be kept approximately at 15 slots. Networks with differing traffic requirements might be able to achieve better performance by assigning some channels for longer data packets and maintaining a nonuniform probability for selection of channels. This would entitle successful transmission of longer data streams without increasing the latency on the shorter transmissions.

Thus, the gains by DCA are more significant for networks with short data packets and fewer channels.

5.2. Transmission delay

The system delay depends, along with other factors, on the performance of the query detection and retransmission. We can however estimate the minimum delay in packet reception by assuming perfect detection of the query. The retransmission policy is defined with a buffer of one packet at each node. The packet arrivals are Bernoulli with a probability \( p \) for every idle node.

Similar to the argument given in [6], we use Little’s theorem to calculate the average delay. The average delay \( D \) is the time taken for a new arriving packet to be transmitted and successfully received by the intended receiver. For a stabilized system, the arrival rate is equal to the throughput of the system \( (\Gamma) \). The total number of nodes \( (B) \) in the system are the nodes that are either receiving, transmitting, or having a packet to transmit:

\[
B = \sum_{l,m,n} [(N - 2m - 2l)p + m + l]S_{l,m,n}. \tag{31}
\]

Thus the average delay per minipacket is \( \bar{D} = B/\Gamma \). Since the average packet length is \( \hat{L} = 1/(1 - q) \), the average system delay is

\[
D = \bar{D}\hat{L} = \frac{\bar{D}}{(1 - q)}. \tag{32}
\]

For light loads \( (p < 0.1) \), the protocols appear to have bounded delays. Delay for DCA is increased due to the additional overhead required for the resolution of the query. The best-case performance of DCA would in fact be the curve for MACA-CT and would occur in the event that every query was successful.

It may also be noted that the delay increases exponentially and is much steeper. Thus, proper selection of probabilities for transmission is very critical. Packet delay at the point of maximum throughput, denoted in Figure 12 by “*”,
5.3. Multihop networks

All the above analysis is for a fully connected single-hop scenario. Modeling of a multihop network is difficult. However, a few reference cases were simulated to postulate the applicability of DCA to multihop networks and to exhibit its performance gain over existing protocols.

Figure 13a shows a fully connected network in which all the traffic is directed to the base station. Figure 13b is a multihop network of 16 nodes with each node having 4 neighbors [10]. The lines between the nodes show the connectivity between the nodes. The throughputs using DCA and MAC-CT are shown in Figure 14.

It is interesting to note the structural dependence on the requirement of the number of spreading codes for the other protocols. In case (b), MACA-CT can be designed using a minimal of 11 data channels by taking advantage of spatial separation. For either situation, CHMA would still require as many channels as the maximal data-packet length. Both problems can be avoided by a dynamic allocation of channels.

The parameters used in the simulations are identical to those used previously. We consider 5 data channels with one common control channel. Mean data-packet length is 10 slots with a geometric distribution. Nodes have a single packet buffer. The network throughput is recorded with a constant probability of packet arrival ($p$).

As can be seen for case (b), since the contention neighborhood is much smaller, the throughput of DCA is significantly greater than that for a fully connected network of the same size. Also the gains of DCA over MACA-CT are clearly visible.

Thus, in the context specified, DCA is superior to the other protocols and offers significant advantage. The penalty is the increased complexity of the receiver and the need for proper parameter selection. These could either be set a priori or kept variant, depending on the network load.

6. CONCLUSIONS

Medium access control is a critical issue in ad hoc networks. One of the biggest stumbling blocks that remains is the proper scheduling and reception of data packets in the absence of a central controller. Contention of data packets occurs at the receivers, and hence proper scheduling of data packets requires the propagation of the contention information from the receivers to the transmitters. This is particularly interesting for multichannel ad hoc networks since the contention information can also be used in channel allocation.

In multichannel ad hoc networks, the channel assignment has conventionally been regarded as a separate issue and isolated from the MAC. The spreading gain and consequent loss in the data rate are mostly overlooked.

Our objective here has been to propose a MAC protocol for multichannel ad hoc networks based on the feedback of channel contention at the receiver. A channel is selected for transmission only if it does not cause any contention at any of the receivers in the neighborhood. The protocol is proposed in Section 3.

The salient features of the protocol include the fact that channel allocation is included as a part of the MAC and the introduction of a feedback mechanism to propagate channel contention. This not only results in a tighter reuse of channels over a multihop network but also makes the spreading gain independent of the size of the neighborhood.
We propose a novel method for the dynamic allocation of channels to nodes by means of querying the channel. Querying is a binary hypothesis detection and it is shown that the detection of the query can be modeled in terms of a Neyman-Pearson detector. The success of the hypothesis is quantized in terms of two quantities based on the signal-to-noise ratio at the receiver, the probability of false alarm, and missed detection of the query.

The throughput of the protocol is analysed for a fully-connected network in Section 4. Our analysis and simulations reveal that the network throughput is a convex function of the spreading codes, data-packet length, and the probability of transmission. The operating threshold of the query detection also has significant impact on the network throughput. Proper selection of network parameters is crucial in order to maximize the throughput.

Performance of the system for different parameters is analysed in Section 5. It is seen that for low noise conditions, DCA is superior to other protocols. DCA also manages to reduce the dependence of the protocol on the network topology thus being more versatile.

Before we conclude, it is perhaps important to note that most of the losses in DCA are the result of improper querying. We believe that the efficiency of the protocol can be further improved with the use of a "smart," nonrandom channel selection policy as well as by optimizing the data state detectors and the query state detector to independent thresholds.

APPENDICES

A. MODIFICATIONS TO MACA-CT

The improvement in throughput in CHMA is due to the relocation of the CTS from the common channel to the transmitter’s assigned channel. Unfortunately, due to the relation between the maximum data-packet length and the hopping sequence length, it is not easy to calculate the normalized throughput of CHMA. The same technique can however be implemented without channel hopping. This may be considered as an extension of MACA-CT. We call this protocol modified CT (to acknowledge it as an extension of the CT protocols) and is introduced primarily to obtain an estimate on the maximum normalized throughput achievable by CHMA.

B. MODIFIED CT: THE PROTOCOL

Consider a time-slotted system with $N$ nodes. Each node has a preassigned channel on which it transmits all the data packets. Thus, there are $N$ fixed data channels. In addition, there is a common control channel. Any node that has a packet to transmit sends an RTS on the control channel. The RTS specifies the transmitter, the receiver, and the transmitter’s assigned channel. This part of the protocol is exactly identical to MACA-CT.

Since all the idle nodes are tuned to the common channel, if the RTS is received successfully by the intended receiver, it sends a CTS to the source node over the transmitter’s assigned channel. This is the basic difference between MACA-CT and modified CT (Figure 15). At that time, the two given nodes will proceed to exchange data over the transmitter’s assigned channel. When the transmission of the data is completed, the sender and the receiver reset and tune back to the common channel.

If either multiple RTSs are sent or the destination does not receive the RTS, no CTS is sent, and consequently the source node reverts back to idle. In the absence of detection errors, the CTS always succeeds. Since the channel chosen for transmission of the data packet depends upon the transmitter and is not dependent on the slot number, the normalized throughput can be calculated for this case and is simply the network throughput divided by the total number of channels ($N + 1$).

C. ANALYSIS OF THROUGHPUT FOR MODIFIED CT

The modified CT protocol is analysed for a single-hop fully connected network under the same assumptions made in CHMA.

For any time slot, the network can be described by

1. $k$: the number of nodes transmitting an RTS in the current minislot;
2. $l$: the number of nodes that sent an RTS in the previous time slot but failed the contention;
3. $m$: the number of node pairs communicating on the transmitter’s assigned channel. As seen from Figure 15, the packet will be either CTS or data.

Given a network with $N$ nodes, any combination of these parameters ($k, l, m$) completely describes the current state of the network. Also, let $(w, x, y)$ represent identical parameters for the previous time slot.

We assume that the length of the data packet has a geometric distribution with a probability $q$ of the data transmission continuing to the next time slot. Thus, the length ($D$) of the packet is $P(D = d) = q^{d-1}(1 - q)$. Then, the state in the next time slot $(w, x, y)$ would depend only on the current state $(k, l, m)$ and the states form a Markov chain.
Let $\mathcal{T}$ represent the event that the transition from $(k, l, m)$ to $(w, x, y)$ occurs, $\mathcal{I}$ the event that exactly one RTS is sent (i.e., $k = 1$) and it is sent to an idle node, and $\mathcal{B}$ the event that exactly one RTS is sent (i.e., $k = 1$) but it is sent to a busy node. The transition probabilities for the state in the Markov chain can be computed as

$$P_{klm, wxy} = P(w, x, y | k, l, m) = \sum_{i=0}^{m} P(i \text{ data pairs}) \times \left[ P(\mathcal{T} \cup \mathcal{I}) + P(\mathcal{T} \cup \mathcal{B}) + P(\mathcal{T} \cup (k \neq 1)) \right] = \sum_{i=0}^{m} B(m, 1 - q, i) \times \left[ \delta(m' - 1) \delta(x) \delta(k - 1) \times B(N' - 1, p, w) \left( \frac{N''}{N - 1} \right) + \delta(m') \delta(x - 1) \times (k - 1)B(N', p, w) \left( \frac{N - N'' - 1}{N - 1} \right) + \delta(m') \delta(k - x) (1 - \delta(k - 1)) B(N', p, w) \right],$$

(C.1)

where $B(m, 1 - q, i)$ is the binomial distribution and represents the probability that $i$ out of the $m$ data streams terminate, $N' = N - 2(m - i) - k$ is the number of nodes that are not transmitting or receiving at the end of the slot, $N'' = N - 2m - l - k$ is the number of idle nodes for the duration of the slot, and $m' = y - (m - i)$ is the number of new node pairs that start transmitting.

The chain is aperiodic and irreducible, thus a steady-state distribution $P_{klm}$ exists. Since the CTS is also transmitted on the data channel, it needs to be subtracted from our computation of the average number of packets carried per slot. The network throughput is given by

$$\Gamma = \sum_{k, l, m} m S_{klm} - \sum_{k=1, l, m} S_{klm} \left( \sum_{w, x, y} P_{klm, wxy} \right),$$

(C.2)

where the first term on the right-hand side is the average number of packets carried over the data channels, and the second term represents the average number of RTS successful in one time slot. Since for every successful RTS, the CTS is always successful, the difference denotes the raw data throughput. Also important to note is that the slot length for modified CT is one half that of MACA-CT.

Numerical values for the throughput of modified CT are compared against that of MACA-CT and CHMA for fully connected networks of different sizes and with a mean data-packet length that is 20 times the length of the RTS (Table 2). It is seen that the network throughput of CHMA and modified CT is the same. This substantiates our claim that the normalized throughput of modified CT represents a limit on the performance achievable by CHMA.

### Table 2: Throughput of MACA-CT, CHMA, and modified CT for networks of different sizes.

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACA-CT</td>
<td>1.7669</td>
<td>2.1521</td>
<td>2.4131</td>
<td>2.5981</td>
</tr>
<tr>
<td>CHMA</td>
<td>2.4148</td>
<td>3.2190</td>
<td>3.7832</td>
<td>4.3363</td>
</tr>
<tr>
<td>Modified CT</td>
<td>2.4148</td>
<td>3.2190</td>
<td>3.7832</td>
<td>4.3363</td>
</tr>
</tbody>
</table>

### ACKNOWLEDGMENTS

The authors would like to acknowledge the suggestions of Gokhan Mergen, Atul Maharsi, and Mamata Desai for their constructive feedback in the development of this paper. This work was supported in part by the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564 and by the Army Research Office under Grant ARO-DAAB19-00-1-0507.

### REFERENCES


Amit Butala received the B.S. degree from the Indian Institute of Technology, Mumbai, India, in 1999, and the M.S. degree in electrical engineering in 2001 from Cornell University, Ithaca, New York. His areas of interest include ad hoc communications, coding theory, and spread-spectrum communications. He is currently with Qualcomm Inc., Campbell, California.

Lang Tong is a Professor in the School of Electrical and Computer Engineering, Cornell University, Ithaca, New York. He received the B.E. degree from Tsinghua University, Beijing, China, in 1985, and the M.S. and Ph.D. degrees in electrical engineering in 1987 and 1990, respectively, from the University of Notre Dame, Notre Dame, Indiana. He was a Postdoctoral Research Affiliate at the Information Systems Laboratory, Stanford University, in 1991. He was also the 2001 Cor Wit Visiting Professor at the Delft University of Technology. Dr. Tong received the Young Investigator Award from the Office of Naval Research in 1996, and the Outstanding Young Author Award from the IEEE Circuits and Systems Society. His areas of interest include statistical signal processing, wireless communications, communication networks and sensor networks, and information theory.