A Performance Analysis of the Basic Access IEEE 802.11 Wireless LAN MAC Protocol (CSMA/CA)

Shreyas Sadalgi
Department of Computer Science
Rutgers University, NJ 08854
sadalgi@cs.rutgers.edu

8 May, 2000

Abstract

The desire to provide universal connectivity for mobile computers and communication devices is fueling a growing interest in wireless packet networks. The 802.11 study group was formed under the IEEE to recommend an international standard for Wireless Local Area Networks (WLANs). A key part of standard are the Medium Access Control (MAC) protocols needed to support asynchronous and time bounded delivery of data frames. This Simulator aims at a performance evaluation of the asynchronous data transfer protocols that are a part of the proposed IEEE 802.11 standard. The Basic Access Method in the 802.11 MAC protocol is the Distributed Coordination Function (DCF), which is best described as the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). A detailed study is made of the Basic Access Method, taking into account the decentralized nature of communication between stations, with the possible presence of "hidden" stations. System throughput is computed over varying load conditions. The impact of “reachability” of nodes on system throughput is studied by making measurements over different “hidden” node probabilities. The results reveal that indeed, the problem of hidden nodes affects the system throughput, and so this method of access should be used only when nodes have a large degree of “reachability”, thus eliminating any hidden node conditions.

1 Introduction

In recent years there has been an increasing trend towards personal computers and workstations becoming "portable" and "mobile". This ever increasing group of mobile users have been demanding access to network services similar to their "tethered" counterparts. To meet these and other future communication needs it is expected that tomorrow's communication networks would employ wireless media in the local area and utilize high capacity wired media in the metropolitan and wide area environment. Wireless systems and networks will provide communication capability not only between mobile terminals but also permit these mobile devices to have access to "wired" networks.

In order to achieve the goal of offering broadband communication services and providing universal connectivity to mobile users it is important that (i) a suitable standard for Wireless Local Area Networks (WLANs) be designed and (ii) an approach to interconnect these WLANs to the existing wired Local Area Networks (LANs) and broadband networks be developed. A key design requirement for WLANs is that mobile hosts be able to communicate with other mobile and "wired" hosts (on other IEEE 802 LANs and/or networks) in a transparent manner, i.e., (i) WLAN should appear to the Logic Link Control (LLC) layer and those above as just another 802.x LAN 1 (for example, Ethernet and Token ring) and (ii) the response times should not be so large that the productivity of end-users is compro-
mised. To be able to achieve the above objectives it is imperative that mobility be handled at or below theMedium Access Control (MAC) layer (note that in wireless networks an "address" does not correspond to a fixed physical location as in wired networks). Furthermore, it is important that the performance available to mobile users be comparable to that available to wired hosts.

To satisfy the above mentioned needs of wireless data networking, study group 802.11 was formed under IEEE project 802 to recommend an international standard for WLANs. The scope of the 802.11 study group is to develop Medium Access Control (MAC) layer and Physical (PHY) layer standards for wireless connectivity of fixed, portable, and mobile stations within a local area. Specifically, the 802.11 standard will describe:

- Functions and services required by an 802.11 compliant device to operate within a wireless network as well as aspects of station mobility within these networks;
- MAC procedures to support asynchronous and time bounded delivery of data frames;
- Services required to provide security and privacy to 802.11 compliant devices.

The physical media used in WLANs have several key differences when compared to wired media. The physical media used in 802.11 WLANs is shared and has limited point-to-point connection range. Further, the media is significantly less reliable, i.e., error rates are higher, (when compared to wired media) (i) since it is unprotected from "outside signals" and (ii) due to the presence of non-stationary multipath fading. These two key properties of wireless media leads to WLAN networks having characteristics that are significantly different from traditional wired LANs. The design of WLANs is further complicated by the presence of hidden terminals and the possibility of capture.

A pair of stations in the same WLAN is referred to as being hidden from each other if the transmissions from one station cannot be heard by the other. Given the presence of hidden stations in the WLAN, carrier sensing is not reliable since stations can sense the state (busy or idle) of the wireless channel differently (even when propagation delays are negligible). It should also be mentioned that carrier sensing in wireless environments requires significantly longer time (tens of microseconds) when compared to wired networks. Capture refers to the ability of a receiver to successfully receive a transmission from a given station when multiple stations are transmitting simultaneously. For instance, if the received power from stations transmitting simultaneously is sufficiently different, the transmission that has the highest power at the receiver could be received successfully. The ability to capture depends on the modulation/demodulation used and the receiver characteristics.

1.1 Architecture of an 802.11 System

An 802.11 network, in general, consists of Basic Service Sets (BSS) that are interconnected with a Distribution System (DS); see Figure 1.

![Figure 1: Components of an 802.11 WLAN System](image-url)
(CF) --- the logical function that determines when a station transmits and receives via the wireless medium. Stations in a BSS gain access to the DS and to stations in "remote" BSSs through an Access Point (AP). An AP is an entity that implements both the 802.11 and the DS MAC protocols and can therefore communicate with stations in the BSS to which it belongs and to other APs (that are connected to the DS). Before a station can access the wireless medium it has to be associated with an AP. A station can be associated with only one AP at any given time. The DS supports mobility by providing services necessary to handle the address 1 to destination mapping and the integration of BSS's in a manner that is transparent to stations, i.e., hosts (either mobile or wired) do not need to know the physical location of other hosts for communication. A network of interconnected BSSs, as in Figure 1, in which mobiles can roam without loss in connectivity is frequently referred to as an Extended Service Set (ESS).

1.2 Physical Layer

Before describing the MAC layer protocols in detail it is relevant to expand upon the underlying physical layer. As noted earlier, the physical layers being considered for IEEE 802.11 WLANs include direct sequence and frequency hopping spread spectrum as well as diffused infrared. When using DSSS a single spreading sequence is used in any given BSS; thus all stations use the same spreading sequence. The frequency bands used when employing DSSS in neighboring BSSs (which are potentially overlapping) are chosen to be different in order to minimize the interference between BSSs. However, in a FHSS system different hopping sequences are employed in neighboring BSS's while all stations in a BSS use identical ones.

1.3 The MAC Layer Protocol

The 802.11 MAC layer protocol provides asynchronous, time-bounded, and contention free access control on a variety of physical layers. These functions are provided independent of the characteristics of the underlying physical layers and/or data rates.

The basic access method in the 802.11 MAC protocol is the Distributed Coordination Function (DCF) which is best described as the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). In addition to the DCF the 802.11 also incorporates an alternative access method known as the Point Coordination Function (PCF) --- an access method that is similar to "polling" and uses a point coordinator (usually the AP) to determine which station has the right to transmit. Further, an optional Distributed Time Bounded Service (DTBS) may be provided by the DCF. DTBS is a "best effort" service that provides bounded delay and bounded delay variance. The DCF is now described in detail.

1.4 Distributed Coordination Function – Basic Access Method

When using the DCF, a station, before initiating a transmission, senses the channel to determine if another station is transmitting. The station proceeds with its transmission if the medium is determined to be idle for an interval that exceeds the Distributed Inter-Frame Space (DIFS). In case the medium is busy, the transmission is deferred until the end of the ongoing transmission. A random interval, henceforth referred to as the backoff interval, is then selected and is used to initialize the backoff timer. The backoff timer is decremented only when the medium is idle; it is frozen when the medium is busy. After a busy period the decrementing of the backoff timer resumes only after the medium has been free longer than DIFS. A station initiates a transmission when the backoff timer reaches zero. To reduce the probability of collisions, after each unsuccessful transmission attempt the expected value of
the random backoff interval is increased exponentially until a predetermined maximum is reached.

Immediate positive acknowledgments are employed to determine the successful reception of each data frame (note that explicit acknowledgments are required since a transmitter cannot determine if the data frame was successfully received by listening to its own transmission as in wired LANs). This is accomplished by the receiver initiating the transmission of an acknowledgment frame after a time interval *Short InterFrame Space* (SIFS), that is less than DIFS, immediately following the reception of the data frame. Note that the acknowledgment is transmitted without the receiver sensing the state of the channel. In case an acknowledgment is not received the data frame is presumed lost and a retransmission is scheduled (by the transmitter). This access method, henceforth referred to as Basic Access, is summarized in Figure 2.

![Figure 2: Basic Channel Access Method](image)

### 1.5 Backoff Mechanism

A station desiring to initiate transfer of data and management frames uses both the physical and the virtual carrier sense function to sense the state of the medium. If the medium is busy, the station defers until after a DIFS is detected and then generates a random backoff period for an additional deferral time before transmitting. This minimizes collisions during contention between multiple stations. If any carrier activity is detected during the backoff interval, then the backoff procedure is suspended, that is, the backoff timer will not be decremented. The medium should be sensed idle for a duration of a DIFS period before the backoff is allowed to resume. Transmission commences whenever the backoff timer reaches zero.

The backoff interval is chosen as:

\[
\text{int}(\text{CW}[\text{random()}]\times\text{Slot time})
\]

where:

- \(\text{CW} = \) An integer between \(\text{CW}_{\text{min}}\) and \(\text{CW}_{\text{max}}\) representing the Contention Window;
- \(\text{random()}\) = Pseudo-random number between 0 and 1
- \(\text{Slot time} = \) Clear channel assessment time + Rx Tx turnaround time + Air propagation time

The contention window takes an initial value of \(\text{CW}_{\text{min}}\) for each frame queued for transmission. If a frame collides, the transmitter times out and re-transmits the frame. The time out interval is the nominal time taken by the frame to reach the receiver plus the time taken by the corresponding response frame to reach the transmitter. The contention window takes the next value in the series at every retry to send a particular frame until it reaches \(\text{CW}_{\text{max}}\). The contention window will remain at \(\text{CW}_{\text{max}}\) for the remaining retries. The set of \(\text{CW}\) values are 7 (\(\text{CW}_{\text{min}}\)), 15, 31, 63, 127, 255 (\(\text{CW}_{\text{max}}\)).

The remaining report is organized as follows: In Section 2, the implementation details of the simulator are described; Section 3 describes the simulation models and performance results; finally Section 4 makes concluding remarks.
2 Simulation Methodology

Various new connector-derived objects were designed and used in the implementation of the simulator. This section describes each of them and how their implementation correctly models the actual protocol.

2.1 Physical Channel

The Channel object implements the physical medium for the wireless protocol. It stores parameters specific to the channel as private variables: bandwidth reflects the transmission capacity of the channel; reachableDistance reflects the “reachability” factor – which specifies the radius within which the node can successfully communicate to its peers. The Boolean variable busy tells other components whether the channel is in use or not. In addition, a list of nodes that are waiting for access to the channel to transmit data is maintained in waitList.

To gain access to the channel, the MAC object at every node must generate a ChannelWait event. On receipt of this event, the channelWait method adds this request to its waitList, before which it checks for the “hidden” node condition. This condition checks if any other node wants to transmit to the same destination node as that of the new transmit request. The condition is satisfied only if the initiation time of such a packet transmission falls inside the “vulnerability interval” of the new transmit event. When a hidden node problem is detected, a collision is ascertained, and the packets are dropped into the StatisticCollector of the channel – which collects statistics about hidden node collisions.

Once a ChannelFree event is generated, the channelFree method searches the waitlist of nodes, to give access to the one which has the lowest value of the backoff timer. While doing so, a check is also performed for collisions which might occur as a result of similar values of backoff timers. If such a collision is detected, a new value is set for the Backoff timer of these nodes – by incrementing it exponentially. Once the node is selected, to whom access to the channel will be granted, all other backoff timers are decremented by the appropriate values. The access is indicated by generating a PacketServe event at the appropriate MAC object of the source node.

On receipt of a PacketArrival event (when a node transmits after having access to the channel), the PacketTransmit event is in turn triggered. This activates the send method, which sends out a copy of the packet to each node that is “reachable” from the source node.

2.2 MAC Layer – CSMA/CA protocol

The MacCSMA object implements the MAC layer protocol. It stores a list called queue to keep track of those packets for which no ACK has been received as yet.

On a PacketArrival event, which is generated to denote arrival of a packet from the Link Layer, the checkAndWait method is activated. This method senses the channel; if found free (which is true only in case of the first access to the Channel in the simulation), then it sets the channel to busy & transmits the packet onto the Channel; else if busy, then it sets a random value for the backoff timer of the node (whose value is stored in the packet for convenience) and generates a ChannelWait event at the Channel for this node.

On a PacketServe event, which is generated when access to the Channel is granted, the send method adds this packet to the queue of unACKed packets, and sends out the packet for transmission onto the Channel. In addition, a TimeOut event is also scheduled at the MacCSMA object at time = (PackSize+ACKsize)/ChannelBandwidth
On a PacketReceive event, the receive method checks the packet if it is destined for this node. If it is destined for this node, then (i) if it is a DATA packet, it schedules a PacketGenAck event at this MAC object, (ii) if it is an ACK packet, it records the ACK by deleting its entry from the unACKed queue and it sets the ChannelFree event at the Channel. If the packet is not destined for this node, then it is dropped into the sink that collects junk packets.

On a PacketGenAck event, the genAck method generates an ACK packet and schedules its arrival at the Channel. Also, a PacketArrival event is scheduled immediately at the Link Layer to send up the received packet to the node.

On receipt of a TimeOut event, the retransmit method checks if it really is a timeout by checking for an appropriate entry in the unACKed list – if an entry is found then it is a timeout, and so the Channel is set free and again a transmit process is carried out for the packet.; else nothing is done.

2.3 Node Object

Each Node is an independent entity that stores its own position and has its set of the network layer components. It has a MacCSMA object through which it is connected to the Channel. Figure 3 shows the schematic of the node and clearly outlines the various components of a node.

The node generates packets through its CBRGenerator source, on receipt of a PacketGenerate event. It receives a packet from the Channel on a PacketArrival event, by triggering a PacketReceive event at the MacCSMA object.

2.4 StatsCollector

This is the object that the sinks set as target to collect statistics about the Simulation. It is an object similar to the PacketSink object.

2.5 Modifications to Packet Object

The control field specifies whether this is a DATA or ACK frame. SendTime and recvTime are used to record the latency of the transmission. The location of the node is denoted by x and y – respectively denoting

Figure 3: Schematic of a node and overall setup of Simulator components.
the co-ordinate values. \textit{InitTime} is used to store the time at which the transmission was initialized – to detect hidden node problems. \textit{MacID} is used to keep a reference to the MAC object when a Channel wants to grant access to the node. \textit{Seq} stores the sequence number of packets transmitted to a particular destination. The backoff timer value is stored in \textit{backOff} – and the Contention Window index is stored in \textit{CW}, which gives the exponential increment in case of a transmission failure.

3 Simulation Models and Results

Performance evaluation of 802.11 MAC protocol poses several new challenges that require innovative solutions. The presence of hidden stations, use of carrier sensing in a wireless environment, and decentralized nature of communication, i.e., without the intervention of the AP, are some of the key factors whose impact on system performance needs to be carefully evaluated. In WLAN systems the performance observed by all stations is not identical; it is expected to be a function of not only their location but the location of other stations as well. Therefore, the impact of the spatial characteristics, for example, station location and traffic patterns, on system performance has to be determined. In this study, the performance of 802.11 has been evaluated only within a single BSS. This means that no AP is needed for communication – the network is peer-to-peer.

3.1 Parameters and Assumptions

The various constants used for the simulation have been mentioned in Table 1. The different parameters that can be varied to study the performance of the MAC protocol are:

\begin{itemize}
  \item \textit{G} - Load on system
  \item \textit{P_h} - Hidden node probability – the probability that 2 nodes cannot “hear” each other. This parameter is set by the spatial characteristics of the nodes.(location).
  \item \textit{N} - Number of stations in the scenario
  \item \textit{Reachability} – the radius within which a node can transmit or “hear” about other transmissions.
\end{itemize}

In this simulation, for the purpose of simplification, the DIFS and SIFS intervals are neglected – so the channel access is based purely on the backoff timers – if the channel is sensed busy. This is so because, in the case of my simulation, where the progress of the clock is event-based, rather than time-period based, the need of a DIFS and SIFS wait is redundant.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth of Channel</td>
<td>1 Mbytes/sec</td>
</tr>
<tr>
<td>SLOT time</td>
<td>100 ms</td>
</tr>
<tr>
<td>Length of ACK packet</td>
<td>14 bytes</td>
</tr>
<tr>
<td>LinkLayer Delay</td>
<td>0.1ms</td>
</tr>
<tr>
<td>Frame/Packet Size</td>
<td>526 bytes</td>
</tr>
<tr>
<td>ContentionWindow Slot sizes</td>
<td>{7,15,31,63,127,255}</td>
</tr>
<tr>
<td>(ReachableDistance:P_h)</td>
<td>{(1.0:0.0),(0.7,0.01),(0.5,0.05),(0.3,0.1)}</td>
</tr>
</tbody>
</table>

Table 1: Various constants used in Simulation

The physical channel normally takes into account the power level of the transmitter and receiver to determine a proper packet reception or a garbled message. Here in this simulation model, I have simplified the issue by converting the above problem in terms of distances – a packet is sent out from the channel only to those sources which are “reachable” – based on the \textit{reachableDistance} parameter.

3.2 Performance Metric

\textit{S} - Throughput: is defined as the number of successful transmissions over the time elapsed between packet generation and termination of simulation. It is also defined as the number of successful transmissions between successful “renewal points”, divided by the the length of the time interval between the renewal points.
Renewal points in a WLAN occur when all stations simultaneously sense the channel to be idle for more than a DIFS. Unlike wired multiple access systems, the completion of a successful transmission or a collision is not a renewal point due to the presence of hidden stations. Obtaining statistics of the renewal interval in a WLAN environment in the presence of hidden stations is intractable; I therefore assume that the instance of the completion of a successful transmission and/or a collision is a renewal point.

3.3 Model 1

In the first simulation model, the effect of hidden node probability on the system throughput is studied. This means studying the effect of (i) Spatial characteristics, (ii) Reachability of nodes.

First, an estimate is made to reflect a particular scenario for the hidden node probability – $P_h$ is modeled according to 2 parameters:

(i) ratio of number of nodes not reachable from a source node to total number of nodes – this is set by varying the reachableDistance parameter.

(ii) Spatial characteristics are varied by changing the destination nodes of transmissions—rather than actually changing the positions of nodes (which is tedious).

So, to study the effect of throughput, $P_h$ is increased by:

(i) decreasing reachableDistance (see Table 2 for the used values)

(ii) Changing the destination for a particular node so that it acts as “hidden” node now, if it did not in an earlier scenario – eg. If node A and node C are not reachable, but they did not both earlier have a node B as a destination, which is within reach of both nodes separately. Now, setting B as a destination for both A and C creates a hidden node scenario.

But the actual value of $P_h$ cannot be determined correctly from the above 2 parameters – because we cannot establish a proper analytical formula to get a value for $P_h$; so we calculate $P_h$ from the hiddenStats Statistic Collector. $P_h$ is then calculated after a simulation as:

\[
\frac{\text{number of packets collected at hiddenStats}}{\text{Total number of packets generated during the simulation}}
\]

The simulation was run in a scenario which consisted of a finite number of nodes – 25 nodes were chosen with locations shown in Table 2: the locations were picked to be uniformly distributed in a 2-dimensional BSS of unit radius.

The load on the system is reflected by $G$ – which is determined by the rate of the CBRGenerator of the nodes.

System throughput is normalized over the bandwidth of the channel.

The nodes were setup to transmit such that one node transmits at most to one other node during a simulation run.

The results are plotted in Figure 4. The results show that throughput decreases, with increase on the load of the system for a fixed $P_h$ value. It initially increases from a low value and then drops down. The nature of this initially increasing characteristic of the curve is because the packet generators do not generate packets as fast as the sources can consume them – (ie) once the channel is sensed idle and a time interval DIFS has elapsed, the time until a frame is generated at a node $i$ that is destined for node $j$ is large.
The throughput is also observed to decrease on increasing $P_h$ values. Further, the rate of decrease in system throughput for large load increases with increase in hidden node probabilities $P_h$. Therefore system stability decreases with increasing $P_h$.

3.4 Model 2

Another simulation model was studied where the effect of number of nodes on throughput was observed. This simulation consisted of changing the number of stations for each run, with varying loads for each case.

The nodes were setup to transmit such that one node transmits at most to one other node during a simulation run.

---

**Figure 4**: Basic Access Method : Effect of Hidden nodes

**Figure 5**: Basic Access Method : Effect of Number of nodes
Table 2: Mobile Node Coordinates for Simulation Model 1

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(.206, .142)</td>
<td>(.275, .408)</td>
<td>(.441, .444)</td>
<td>(.095, .314)</td>
<td>(.479, .175)</td>
</tr>
<tr>
<td>(.271, .133)</td>
<td>(.629, .696)</td>
<td>(.447, .246)</td>
<td>(.619, .177)</td>
<td>(.098, .320)</td>
</tr>
<tr>
<td>(.682, .766)</td>
<td>(.867, .057)</td>
<td>(.268, .671)</td>
<td>(.346, .012)</td>
<td>(.775, .523)</td>
</tr>
<tr>
<td>(.562, .037)</td>
<td>(.919, .730)</td>
<td>(.857, .202)</td>
<td>(.546, .811)</td>
<td>(.853, .462)</td>
</tr>
<tr>
<td>(.499, .340)</td>
<td>(.156, .323)</td>
<td>(.619, .539)</td>
<td>(.320, .539)</td>
<td>(.983, .270)</td>
</tr>
</tbody>
</table>

The results are shown in Figure 5. The results show that as expected, system throughput degrades on increasing number of nodes.

4 Conclusion

In this report, I analyzed the throughput of the asynchronous data transfer methods of the IEEE 802.11 protocol. In particular, the Basic Access Method (CSMA/CA) was studied for nodes communicating within a single BSS. The study was done to reflect the effect of hidden nodes on the system throughput; and why the need for Virtual Carrier Sense (RTS/CTS) mechanism arises.

The simulation of the CSMA/CA method was carried out successfully and its correctness was verified with the results obtained. The components developed to reflect the appropriate Carrier Sense and Collision Avoidance mechanisms are valid and correct, as can be seen in the explanation of their implementation.

The effect of the increasing hidden node probabilities is to degrade the system performance. So we see that, if the number of hidden nodes are small, Basic CSMA/CA method is robust enough for large loads, whereas if the number of hidden nodes are large in number, then there is a need for a better MAC protocol that resolves the collision scenarios; more specifically, the Virtual Carrier Sense mechanism that uses RTS/CTS method is suitable then.

5 References


