The Impact of RTS Threshold on IEEE 802.11 MAC Protocol

Fun Ye[1], Shiann-Tsong Sheu[1], Tobias Chen[1]and Jenhui Chen[2]

[1]*Department of Electrical Engineering*
Tamkang University
Tamsui,Taiwan 251, R.O.C.
E-mail: stsheu@ee.tku.edu.tw
[2]*Department of Computer Science and Information Engineering*
Tamkang University
Tamsui, Taiwan 251, R.O.C.

Abstract

Wireless technologies and applications received great attention in recent years. The medium access control (MAC) protocol is the main element that determines the efficiency in sharing the limited communication bandwidth of the wireless channel in wireless local area networks (WLANs). The request-to-send/clear-to-send (RTS/CTS) mechanism is an optional handshaking procedure used by the IEEE 802.11 wireless network to reduce the possibility of collision. The RTS_Threshold (RT) value which determines when the RTS/CTS handshaking mechanism should be used is an important parameter to investigate; since different RT values will produce different performance characteristics in data transmission. This paper presents an evaluation of the influence of the RT parameter on the IEEE 802.11 wireless network, and gives a guideline to dynamically adjust the RT value. Simulation results of this paper show that, in order to achieve the best performance, the RT should be dynamically adjusted according to the environment. However, we suggest to have the RTS/CTS always activated (RT = 0), saving complex work designing and implementing a dynamic RT mechanism, to obtain a near best performance.

Key Words: Ad Hoc, IEEE 802.11, MAC, RTS/CTS, Wireless

1. Introduction

Wireless communication is a rapidly emerging technology providing users with network connectivity without being restricted by a wired network. An *ad hoc* wireless local area network (WLAN) is a collection of mobile hosts, which forms a temporary network without the aid of any pre-established infrastructure or centralized administration. As a result, wireless applications are becoming more and more popular where wiring for conventional networking is difficult or not economic. The IEEE 802.11 Working Group provides detailed medium access control (MAC) and physical (PHY) layer specifications for WLANs [1]. Some characteristics of the IEEE 802.11 are discussed in [2,3], and a detailed analysis can be found in [4]. Since any transmission in a WLAN relies on a common and open radio medium, the MAC protocol in WLAN would be more important than in conventional wired networks.

The IEEE 802.11 WLAN standard includes a basic distributed coordination function (DCF) and an optional point coordination function (PCF) which is a centralized MAC protocol that supports collision free and time bounded services. The DCF uses carrier sense multiple access with collision avoidance (CSMA/CA) mechanism as the basic channel access protocol to transmit asynchronous data in the contention period. We will focus on DCF in this article.

The DCF employs two handshaking

techniques for packet transmission. The default scheme is a two-way handshaking technique called basic access mechanism. The other is the optional RTS/CTS four-way handshaking mechanism used to combat the effects of collisions. The RTS/CTS mechanism reserves the channel for transmissions involving larger data packets, with the desired effect that less bandwidth would be wasted when collision occurs. On the other hand, when an extremely short packet is of interest, we might not benefit but even consume extra bandwidth from the RTS/CTS mechanism. Therefore, the RTS_Threshold (RT) is a manageable parameter used to determine when an RTS/CTS handshake should precede a data packet.

The fraction of channel bandwidth used by successfully transmitted data packets excluding the MAC header gives a good indication of the overhead required by the MAC protocol to perform its coordination task among stations. This fraction is known as the utilization of the channel, and the maximum value it can attain is known as the throughput of the MAC protocol [5]. Collisions and packet retransmissions consume extra bandwidth that lower the throughput, therefore, the IEEE 802.11 could operate very far from the theoretical throughput [6].

This paper is outlined as follows. In section 2 we briefly review the DCF of the IEEE 802.11 MAC protocol including the RTS/CTS mechanism. In section 3 we describe the simulation environment which is followed by the discussion of simulation results in section 4. Finally, some conclusions are given in section 5.

2. IEEE 802.11 Distributed Coordination Function

This section briefly summarizes the DCF as standardized by the IEEE 802.11 Working Group. For a more complete and detailed presentation, please refer to the IEEE 802.11 standard [1]. A station with a new packet for transmission needs to monitor the channel activity first. If the channel is idle for a period of time equal to the distributed inter-frame space (DIFS), the station starts to transmit instantly. Otherwise, if the channel is sensed busy, the station persists to monitor the channel until it is measured idle for a DIFS. At this point, the station generates a random backoff interval before transmitting in order to minimize the probability of multiple stations simultaneously starting transmission. Furthermore, to avoid channel capture a station must wait a random backoff time between two consecutive packet transmissions, even if the medium is sensed idle for a DIFS time period^2 after the previous transmission.

For efficiency, DCF employs a discrete-time backoff scale scheme. The time right after an idle DIFS is slotted, and a station is only allowed to transmit at the beginning of each time slot.

Since the CSMA/CA can not rely on the stations to detect a collision by listening to their own transmission, as it is done in IEEE 802.3 wired networks, an ACK is transmitted by the destination to signal the source of the successful packet reception. An ACK is transmitted after a short inter-frame space (SIFS) at the end of the received packet.

The two-way handshaking technique for packet transmission described above is called basic access mechanism. DCF also defines an optional four-way handshaking technique for packet transmission. This mechanism, also known as RTS/CTS, is shown in Figure 1. A station that has a packet queued for transmission follows the backoff rules explained above, but instead of transmitting the packet it preliminarily transmits a special short frame called request to send (RTS). When the destination detects an RTS frame it responds after a SIFS time period with a clear to send (CTS) frame. The source is only allowed to transmit the data packet if the CTS frame is correctly received within a duration called CTS_Timeout.

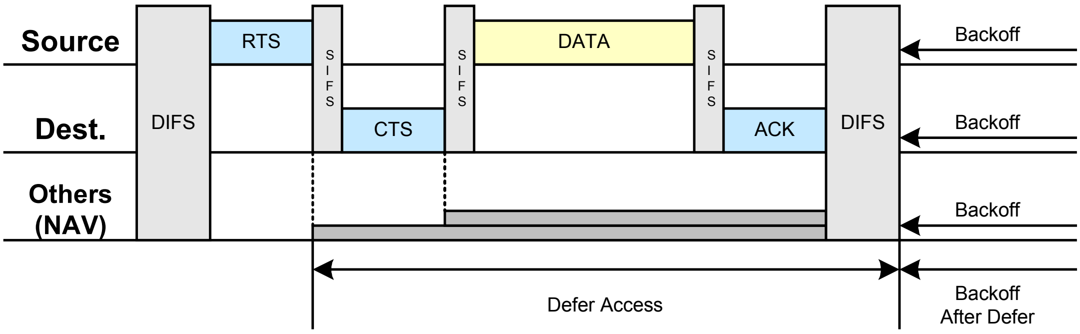

Figure 1. An illustration of RTS/CTS and backoff mechanism of DCF

The RT is a switch for the RTS/CTS mechanism; four way handshaking is used when the packet for transfer is larger than the RT value, otherwise two way handshaking is used. Packets that can be sent with a collision probability less than or equal to the collision probability of a RTS packet should be sent directly (without RTS/CTS handshaking). Because the RTS/CTS mechanism consumes extra bandwidth which has a negative effect on the performance of the network, packets with collision probability slightly greater than the probability of a RTS packet should also be sent directly. The RT should be balanced between higher collision penalty and extra bandwidth consumption. The RTS/CTS mechanism is very effective in terms of system performance, especially when large packets are considered, as it reduces the length of the frames involved in the contention process. In fact, assuming perfect channel sensing by every station, collision may only occur when two or more stations start transmission within the same time slot. If both sources employ the RTS/CTS mechanism, collisions would only occur while transmitting the RTS frames and would promptly be detected by the source lacking the CTS responses.

Now, we will investigate how the RT affects network performance in the IEEE 802.11 WLAN.

3. Simulation Environment

The simulator is custom made and the simulation model follows the IEEE Standard 802.11b-1999 using direct sequence spread spectrum (DSSS) at the physical layer with the long PLCP PPDU (PLCP refers to physical layer convergence protocol and PPDU refers to PLCP protocol data unit) format and DCF at the MAC layer. Most of the parameters were taken from the standard and are listed in Table . Poisson distribution was used to determine the number of MAC service data unit (MSDU) arrivals and the lengths of the MSDUs were decided by the exponential distribution function. Several assumptions were made to reduce the complexity of the simulation model: 1) all stations support the 2 Mb/s data rate, 2) all data and control frames were sent at 2 Mb/s, 3) the propagation delay was neglected, 4) the channel was error-free, 5) there was no interference from nearby basic service sets (BSS). In the simulation all nodes had direct radio contact which means that each source to destination had only 1 hop distance.

4. Simulation Results

The load of the network is determined by three factors: the number of contending nodes (denoted as N) in the BSS, the packet arrival rate per slot time per node (denoted as l), and the mean data length (MDL) of the packets. The network load is equal to (N×l×MDL)/(aSlotTime×Data rate). In order to investigate the influence of the RT value on network performance, intensive simulations were performed by considering the three factors of load. The performance of the network in this article is measured by goodput, which excludes all control and management overhead of the MAC and physical layers, only the data in the frame body of a successfully transmitted MAC frame is considered and accumulated. In other words, the goodput is the ratio of the pure data service rate and the network data rate. The term throughput refers to goodput throughout this article. The simulations plots show the throughput as a function of RT while the three factors vary. Given the three factors of network load, the optimal RT value is defined as the RT value that makes the WLAN reach the maximum throughput. The simulations are split into four parts:

1. Influence of various mean data lengths: Find the optimal RT value for different MDLs while keeping the packet arrival rate and the number of contending nodes fixed. (in Figure 2-4)
2. Influence of various packet arrival rates: Find the optimal RT for different packet arrival rates while keeping the mean packet length and the number of contending nodes fixed. (in Figure 5-6)
3. Influence of various contending nodes: Find the optimal RT for different numbers of contending nodes while keeping the packet arrival rate and the mean packet length fixed. (in Figure 7-8)
4. Trend of throughput: Show the trend of throughput for different RT values while keeping the three load factors the same. (in Figure 9-10)

In Figure 2, the packet arrival rate is considerably low (l = 0.0001) and the throughput curves with mean data length (MDL) shorter than 2k bytes are completely flat. This is because the load of the network is far too low, therefore, giving different RT values does not effect the performance of the network; all packet arrivals can be serviced successfully. For the throughput line with MDL = 2k bytes the load reaches 1 (= (25 ×2k ×8 ×0.0001) / (20 ×2)), now contention is higher and collisions happen more frequently. We can clearly see the throughput falling as the RT value increases. Actually this curve has a peak and it is located at RT = 135 octets; this phenomenon can be seen more clearly on curves with smaller MDLs in figures that have heigher l. Because the lines other then the line with MDL = 2k bytes are completely flat, we infer that the optimal RT equals 135 octets when l = 0.0001 packets/slot/node and N = 25.

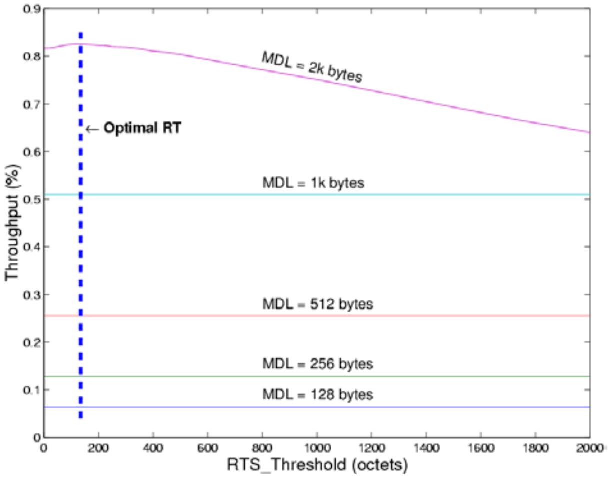

Figure 2. Throughput vs. RT for various MDLs as N = 25 and l = 0.0001 packets/slot/node

In Figure 3 the packet arrival rate is slightly higher than in Figure 2. Both the lines with MDL = 2k bytes and 1k bytes are curves which peak at RT = 135 octets while the lines with MDL = 128 bytes and 256 bytes remain flat. As for the line with MDL = 512 bytes, it is straight before degrading at the point where RT = 850 octets. This line is at a critical state where a RT > 850 octets overloads the network because of high collision penalty. If a collision happens at this point, bandwidth is wasted waiting for a corrupted packet to finish transmission and at the same time the buffer of other nodes start to pile up which may increase the contention at the next contention window. Excluding the straight lines in Figure 3, lower RT values have better performance than higher ones.The peaks of the lines with MDL = 2k bytes and 1k bytes are approximately at RT = 135 octets. According to Figure 3 we infer that the optimal RT = 135 octets when l = 0.0002 packets/slot/node and N = 25. The last figure of this series Figure 4 shows the throughput as a function of RT when the network is fully saturated for all five MDLs. The curves still have a nearly common peak approximately at RT = 135 octets.

Each of the statements above reveal that for a fixed number of contending nodes and packet arrival rate there exists a unique optimal RT. Combining the statements we can deduce that for different packet arrival rates the optimal RT is approximately the same, thus it should be appropriate to say that given a fixed number of contending nodes there exists a unique optimal RT value.

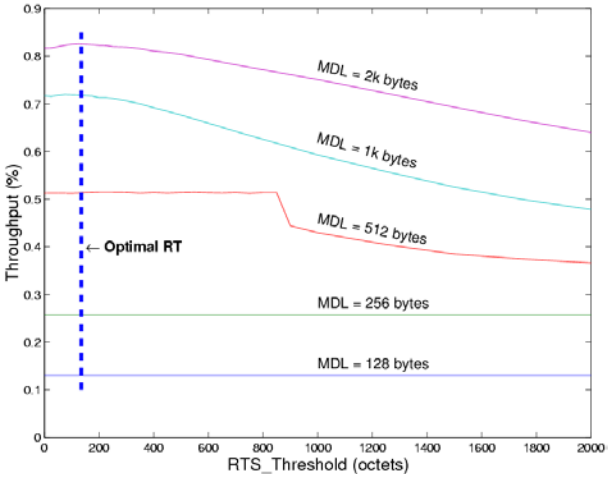

Figure 3. Throughput vs. RT for various MDLs as N = 25 and l = 0.0002 packets/slot/node

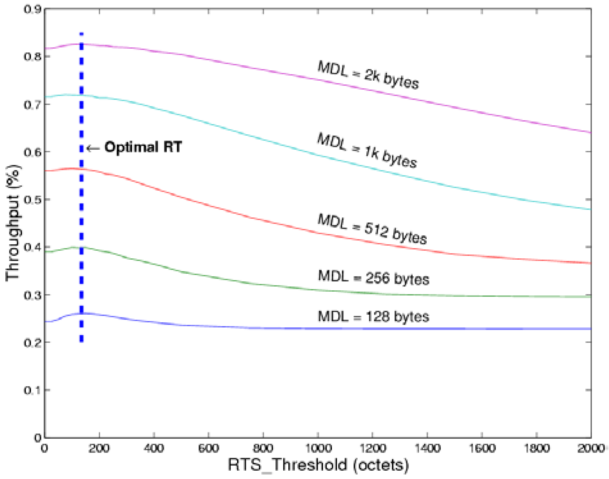

Figure 4. Throughput vs. RT for various MDLs as N = 25 and l = 0.001 packets/slot/node

In Figure 5, the lines with l <= 0.0003 packets/slot/node could not have many collisions due to their low netwok loads. The RTS/CTS mechanism does not have much effect on these curves, thus we can see that these lines are completely flat. The network is saturated when l > = 0.0005 packets/slot/node, thus all curves overlap. A l greater than the l that saturates the network will only make more packets queue in buffer, therefore, the contention of the network is not effected and the curves remain exactly the same. The load is at a critical state with l = 0.0004 packets/slot/node as we can see the throughput curve does not have a peak for maximum throughput. The rise of threshold as RT shifts from 0 to 135 octets is because less extra bandwidth is consumed by RTS/CTS mechanism as the RT value gets larger. As RT > 135 octets and RT < 600 octets the bandwidth used is just a little lower than the maximum capacity of the network thus forming a straight line. After RT > 600 octets, throughput falls as a result of directly sending over large packets. Having a few straight lines, a few overlapping curves, and a highland shaped curve we can easily determine an optimal RTS value by taking the common maximum value RT = 135 octets. In Figure 6 where we used a larger MDL the same phenomenon can be seen. Accordingly, while the number of contending nodes and MDL are fixed there exists an unique optimal RTS value. This statement agrees with the statement in the previous paragraph.

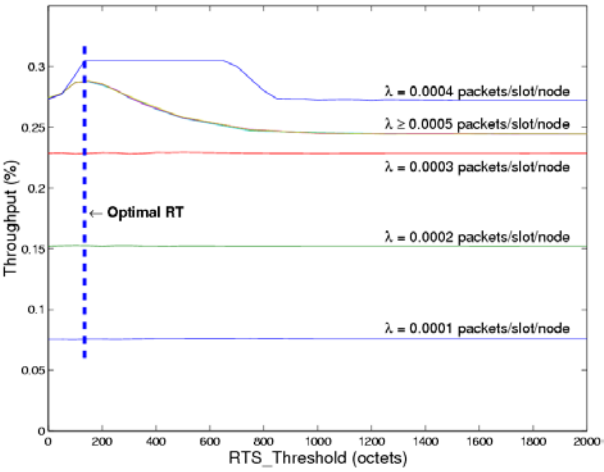

Figure 5. Throughput vs. RT for various ls as N = 25 and MDL = 150 bytes

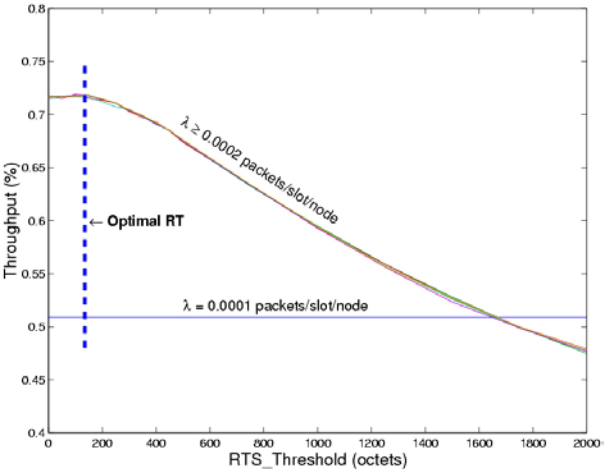

Figure 6. Throughput vs. RT for various ls as N = 25 and MDL = 1k bytes

The curves of Figure 7 and Figure 8 are more interesting. Here we investigate the influence of the number of contending nodes on the optimal RT value while the packet arrival rate and the MDL stay the same. Looking at Figure 7, the line with N = 5 is flat because of low load, and the line with N = 10 grows up and then becomes straight indicating critical load. With the critical load as N = 10 the throughput is lower with a small RT value; the RTS/CTS mechanism is activated more frequently, and the extra bandwidth required by the RTS/CTS mechanism saturates the network thus lowering the performance. For N > =15 we see a curve with a unique peak, and as the number of nodes gets larger it is easy to see that the value corresponding to the peak of the curve gets smaller. The same phenomenon is also observed, only with a larger curve and a more drastic fall when the RT increases, in Figure 8 where we used a larger value for the MDL. Interpreting the implications of these two plots we infer: for a fixed packet arrival rate and MDL, the number of contending nodes is inverse proportional to the optimal RT value. This is also true for N = 5 and N = 10 in Figure 7 since on a straight line any RT corresponds to the same throughput. Here N = 10 is also considered as a flat line because the bent front of the line will always be smaller than the optimal RT value.

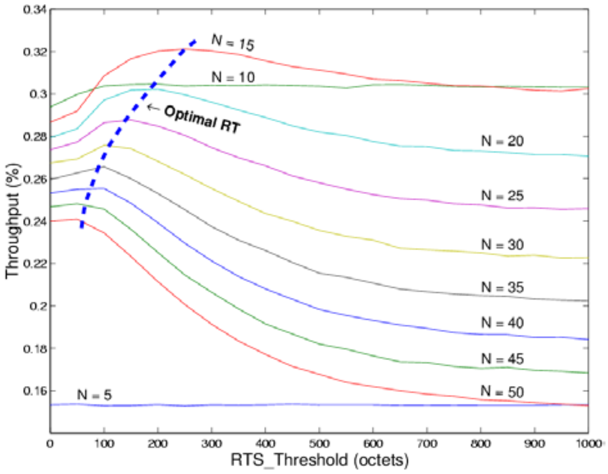

Figure 7. Throughput vs. RT for various numbers of nodes as l = 0.001 packets/slot/node and MDL = 150 bytes

In the last group of plots we investigate the performance of different RT values as a function of load. Figure 9 shows the great performance gap between RT << 00 and RT = 00 especially when the load is high, therefore, it is clear that using the

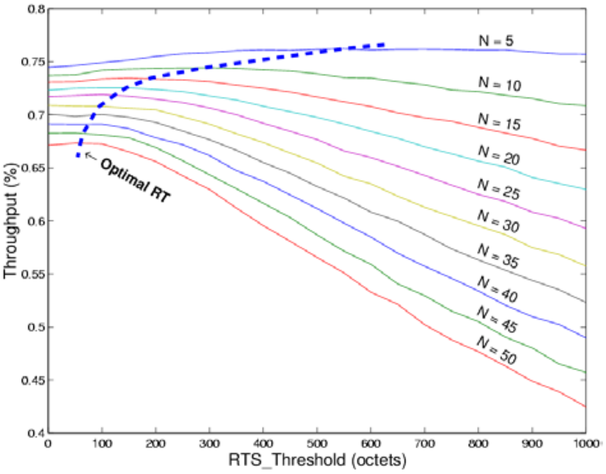

Figure 8. Throughput vs. RT for various numbers of nodes as l = 0.001 packets/slot/node and MDL = 1k bytes

RTS/CTS mechanism benefits network performance. The performance of RT < 500 octets is better than RT > 500 octets, that implies a small RT is better than a large one. Considering the RT = 0 ~ 175 octets, we actually do not see much difference in performance between them, hence it is appropriate to use RT = 0 with any network configuration. Figure 10 which has a lower packet arrival rate shows linear growth in throughput for all RT curves before the curves with larger RT branch off one after another. Although the plot look a little different, the curves with a small RT have a better performance under all circumstances thus revealing the same fact as Figure 9.

The results of the simulations determine that the number of contending stations are the main factor that influence the optimal RT while the packet arrival rate and the length of the packets have only minimal effect which could be ignored. The RTS/CTS mechanism is superior to the basic access method in most cases [1]. It is also clear that a small RT value has better performance than a large one, and a certain small value for RT (RT != 0) has the optimal performance, but among all the small RTs (including RT = 0) only minimal difference is observed. Therefore, we suggest to simply use RT = 0 for any network configuration.

5. Conclusions

In this paper, we have performed a series of detailed simulation to evaluate the effects of RT values on the IEEE 802.11 protocol. The results determine that the number of contending stations are the main factor that influence the optimal RT while the packet arrival rate and the length of the packet have only minimal effect which could be ignored. Compared to the basic access method, network performance is improved by the RTS/CTS mechanism. The best performance is obtained with RT set to a small value that is dynamically adjusted according to the number ofcontending stations. Although a certain small value for RT produces a better performance than always using the RTS/CTS mechanism (RT = 0), the improvement of performance is very trivial. Thus, instead of trying to figure out the number of contending nodes and having a dynamic RT, we suggest always using the RTS/CTS mechanism

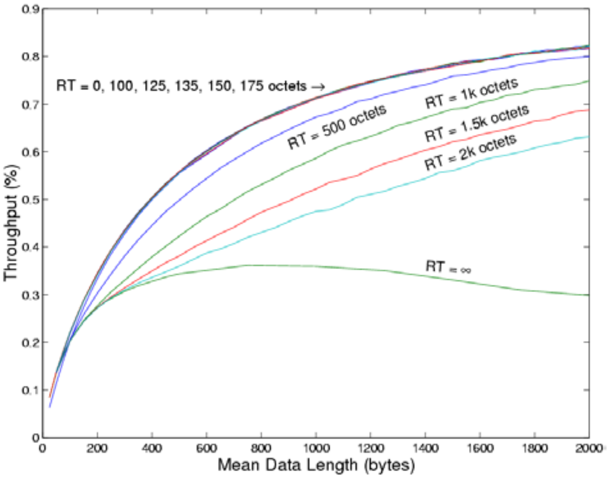

Figure 9. Throughput vs. DML for various RTs as N = 25, l = 0.001 packets/slot/node.

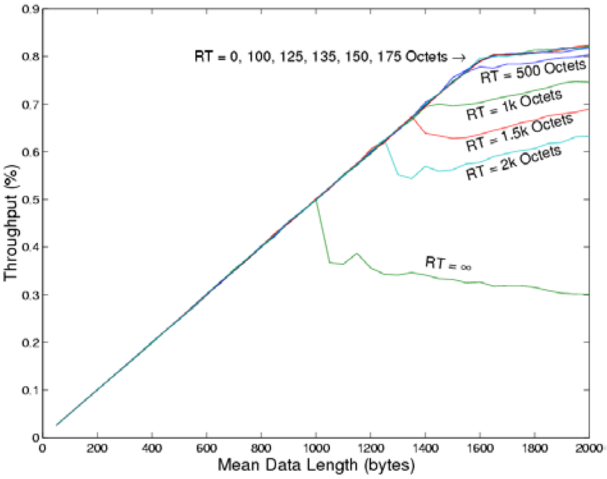

Figure 10. Throughput vs. DML for various RTs as N = 25, l = 0.0001 packets/slot/node.

References

[1] IEEE 802.11 Working Group, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY)

Specifications," ANSI/IEEE Std. 802.11 (1999).

[2] Crow, B. P., Widjaja, I., Kim, L. G. and Sakai, P. T., "IEEE 802.11 Wireless Local Area Networks," *IEEE Commun. Mag.*, Vol. 35 (1997).

[3] Tobagi, F. A. and Kleinrock, L., "Packet Switching in Radio Channels: Part II-The Hidden Terminal Problem in Carrier Sense Multiple-Access and the Busy-Tone Solution," *IEEE Trans. Commun.*, Vol. COM-23, pp. 1417-1433 (1975).

[4] Chhaya, H. S. and Gupta, S., "Performance Modeling of Asynchronous Data Transfer Methods in the IEEE 802.11 MAC Protocol," *ACM/Baltzer Wireless Networks*, Vol. 3, pp. 217-234 (1997).

[5] Kurose, J. F., Schwartz, M. and Yemini, Y. "Multiple-Access Protocols and Time-Constrained Communication," *ACM Computing Surveys*, Vol. 16, pp. 43-70 (1984).

[6] Cali, F., Conti, M. and Gregori, E., "Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit," *IEEE/ACM Trans. Networking*, Vol. 8, pp. 785-799 (2000).

[7] Bianchi, G., "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE J. Select. Areas Commun.*, Vol. 18 (2000).

Manuscript Received: Oct. 1, 2002
and Accepted: Dec. 30, 2002