Effect of Payload Length Variation and Retransmissions on Multimedia in WLANs

S. Choudhury†, I. Sheriff‡, J. D. Gibson† and E. Belding-Royer‡
†Department of Electrical and Computer Engineering
‡Department of Computer Science
University of Califonia, Santa Barbara
Email: {sayantan,gibson}@ece.ucsb.edu,{isheriff,ebelding}@cs.ucsb.edu

***Abstract*— Multimedia transmission over wireless local area networks (WLANs) is a challenging task due to the varying nature of the wireless channel as well as the inherent difference between multimedia traffic and data traffic. In the MAC layer, a single bit error in the packet can lead to the entire packet being discarded. This results in a higher packet error rate for higher payload sizes. Retransmission due to packet errors causes the contention window to double, and this leads to a decrease in throughput if the wireless channel does not improve for the retransmitted packets. Hence, throughput is a function of packet payload length as well as the maximum number of allowable retransmissions. In this paper, we investigate the effect of payload length adaptation and retransmissions on the throughput and capacity of multimedia users. Numerical results and simulations reveal that careful payload adaptation significantly improves the throughput performance at low signal to noise ratios (SNRs). It is also observed that excessive retransmissions can reduce the effective throughput, thereby reducing the capacity of multimedia users in the presence of data users. Since multimedia traffic is more latency constrained and less error constrained, this suggests that by carefully selecting the payload length and maximum number of allowable retransmissions based on the channel conditions, a higher number of multimedia users can be supported.**

I. Introduction

Wireless local area networks (WLANs) are rapidly becoming part of the network infrastructure. These networks were primarily designed for data applications; however, there is currently great interest in conversational voice and video communications (two-way) over WLANs. Since present wireless networks were designed for data traffic, achieving efficient multimedia communications over wireless LANs is not straightforward. It has been observed that despite the low bandwith profile of voice traffic, coexistence of voice with data traffic leads to dropping calls and poor quality even with quality of service (QoS) features enabled [1].

In this paper, we investigate the effect of payload length variation for multimedia traffic on the effective single-user throughput under different channel conditions. We observe that dynamic adaptation of payload length with channel conditions can significantly improve throughput performance. An early investigation of the effect of payload size on throughput was conducted in [2]. Rate adaptation using a theoretical framework to evaluate the throughput has been investigated in [3]. We extend the above results by presenting a simple theoretical scheme that takes into account the tight coupling between payload length and data rate as well as the various protocol overheads. The theoretical scheme is then used to maximize the single-user throughput based on the channel conditions. Our theoretical formulation allows payload length to be varied continuously over a wide range and we provide a mathematical framework to dynamically adapt the payload length to maximize the throughput for additive white Gaussian noise (AWGN) channels. Our results indicate that both the payload length and the data rate are dependent on the channel under consideration. In particular, for a specified signal to noise ratio (SNR), we observe there is an optimal data rate and payload length that maximizes throughput. We validate our theoretical results by performing simulations using the Qualnet simulator [4].

In addition to payload length, we observe that the IEEE 802.11 overhead and retransmission scheme can have a significant effect on the number of multimedia users supported. Voice and video communications are latency constrained while data traffic cannot tolerate bit errors. Thus, the retransmission schemes aimed at minimizing packet losses can reduce the overall throughput. We study the impact of retransmissions under different scenarios. It is observed that in a lossy environment where packet losses are caused by bit errors in the packet, retransmissions result in a decrease in throughput if the wireless channel does not improve for the retransmitted packet. However, retransmissions result in a decrease in packet loss rate. Thus, a retransmission scheme can be devised which, instead of minimizing packet losses, can allow a certain fraction of packets to be lost (based on the application requirement) thereby leading to a higher throughput. We also observe that in a heavily loaded network consisting of multimedia and data users, controlling the maximum number of allowable retransmissions per traffic class results in the support of more multimedia users at the expense of data users. We investigate the impact of payload length and retransmissions because these parameters do not require a change in the MAC protocol and hence can be used with existing WLAN products. Moreover, our scheme can be also be used in addition to the QoS enhancements applicable for IEEE 802.11e.

This work was supported by the California Micro Program, Applied Signal Technology, Dolby Labs, Inc. and Qualcomm, Inc., by NSF Grant Nos. CCF-0429884 and CNS-0435527, and by the UC Discovery Grant Program and Nokia, Inc..

The paper is outlined as follows. In the next section, we provide a brief description of the simulation setup. In Section III, we provide a theoretical model to evaluate single-user throughput in 802.11a networks. Using the theoretical model, we are able to plot effective throughput as a function of payload length for different data rates under different channel conditions. We validate our theoretical scheme with network simulation results. Section IV investigates the effect of retransmissions on throughput. It is shown that in a lightly loaded network where packet losses are caused primarily by bit errors and where the wireless channel does not change significantly, retransmissions reduce the overall packet loss rate but also reduce the effective throughput. However, in certain scenarios with good channel conditions and also in the presence of collisions, retransmissions improve the packet loss performance without causing a significant change in throughput. To evaluate the impact of payload size and the maximum number of allowable retransmissions on capacity, we perform a variety of simulations that vary these values. Conclusions and future work are presented in Section V.

II. Evaluation Environment Overview

Our experiments were conducted using the Qualnet simulator and leveraged IEEE 802.11a at the MAC layer. Two network configurations were used for the simulations - a single-user network with one user connected to a single access point, and a multi-user network with many users connected to a single access point. The first configuration represents a network with no packet collisions; packet loss is entirely due to the channel loss. The second setup represents a network with packet loss due to both collisions and channel loss. Simulations were done in a homogenous network with only multimedia users as well as a network that consisted of both multimedia and data users. The data user application consisted of FTP traffic while the multimedia traffic was 64 Kbps constant bit rate UDP traffic.

III. Effect of Payload Length on Throughput in IEEE 802.11a Wireless Networks

In this section we present a mathematical framework for single user throughput optimization by varying the payload length. We provide an integrated framework for link adaptation by considering various physical and MAC layer adaptation parameters. In particular, we show an adaptive frame length transmission algorithm that can be used to select an optimal payload size and data rate to maximize user throughput. We also show the dependency of throughput on payload length by peforming Qualnet simulations. We observe that the simulations closely match the theoretical results.

We define *throughput* as the number of payload bits per second received correctly. For simplicity of analysis, we consider the effects of payload variation in additive white Gaussian noise (AWGN) channels and assume that the acknowledgements from the receiver are error-free. We also assume a linear mapping between received signal strength (RSS) and SNR as in [5]. Thus based on the RSS of the ACK frames from the access points, the mobile station can estimate the SNR at the receiver. Throughput corresponding to PHY mode m in IEEE 802.11a is given by:

$$T(m) = \frac{L}{L + C_m} * R_m * P_s^m(\gamma_s, L) \tag{1}$$

where,
L: payload length in bits,
C_m: header and DCF overhead corresponding to rate 'm' in bits,
R_m: data rate corresponding to PHY mode m,
$P_s^m()$: packet success rate (PSR) defined as the probability of receiving a packet correctly corresponding to PHY mode m
γ_s: SNR per symbol.

C_m encapsulates both header overhead and CSMA/CA interframe spacing. The time delay is converted to bytes for the purpose of optimization by the following expression:

$$C_m = R_m * T_{ho} \tag{2}$$

where R_m is the transmission rate corresponding to PHY mode 'm' and T_{ho} is the total protocol overhead. T_{ho} can be evalauted as in [3].

We have assumed hard-decision Viterbi decoding at the receiver. For a packet L bits long, the probability of packet error can be bound by [6]:

$$P_e^m(L, \gamma_s) \leq 1 - (1 - P_u^m(\gamma_s))^L \tag{3}$$

where $P_u^m(\gamma)$ is the union bound of the first-event error probability corresponding to PHY mode m [7] and is given by:

$$P_u^m(\gamma_s) = \sum_{d=d_{free}}^{\infty} a_d \cdot P_d(\gamma_s) \tag{4}$$

Here, d_{free} is the free distance of the convolutional code selected in PHY mode m, and a_d is the total number of error events of weight d that can be obtained from [7]. For hard-decision decoding with probability of bit error ρ, $P_d(\gamma_s)$ is given by

$$P_d(\gamma_s) = \begin{cases} \sum_{k=(d+1)/2}^{d} \binom{d}{k} \cdot \rho^k \cdot (1-\rho)^{d-k} & \text{if d is odd} \\ \frac{1}{2} \cdot \binom{d}{d/2} \cdot \rho^{d/2} \cdot (1-\rho)^{d/2} & \\ + \sum_{k=d/2+1}^{d} \binom{d}{k} \cdot \rho^k \cdot (1-\rho)^{d-k} & \text{if d is even} \end{cases} \tag{5}$$

The bit error probability ρ computation for AWGN for the various PHY modes in IEEE 802.11a is described in [3].

The packet success rate (PSR) is given as:

$$P_s^m(\gamma_s, L) = 1 - P_e^m(L, \gamma_s) \tag{6}$$

where L is the packet length including the various overheads in bits.

In order to find the optimal payload length L^*, we assume the payload length L varies continuously. Differentiating Eq. (1) with respect to L and setting it to zero with packet success rate given by Eq. (6), the optimal payload length at a given SNR is:

$$L^* = -\frac{C_m}{2} + \frac{1}{2}\sqrt{C_m^2 - \frac{4 * C_m}{\log(1 - P_u^m(\gamma_s))}} \qquad (7)$$

Some key assumptions in our model are single transmission and no losses due to collisions, i.e. packet losses are caused only by bit errors in AWGN and transmissions of acknowledgments are error free.

The variation of throughput with payload lengths at different SNRs for 6 Mbps and 9 Mbps over an AWGN channel using the theoretical analysis as well as Qualnet simulations is shown in Figs. 1 and 2, respectively. We varied the payload lengths from 1-2000 bytes and added 40 bytes of RTP/UDP/IP header used for multimedia traffic. The plots show that there is an optimal payload length and data rate corresponding to the received SNR for throughput maximization. This is quite intuitive since at a given SNR, increasing the data rate, i.e. the bits per constellation or payload length, initially causes an increase in throughput. However, beyond an optimum throughput value (packet length), a higher packet error rate (Eq. (3)) results in a decrease in throughput. It should be noted that at SNRs in the range of 1.5-3.5 dB, only PHY mode 1, corresponding to 6 Mbps, is used. The higher data rates are not evident in the plot since the packet error rates at these higher transmission rates cause the effective throughput to be very close to zero. The optimal payload length using the theoretical model obtained at an SNR of 2 dB is approximately 300 bytes, corresponding to an effective throughput of 2.5 Mbps. Using a payload length different from the optimal value has a significant impact on throughput. For instance, the effective throughput achieved at an SNR of 2 dB in AWGN using a 20 byte payload is 0.6 Mbps, and with a 2000 byte payload, it is around 0.4 Mbps. Using Qualnet simulations at 1.5 dB, we observe that the optimal payload length is around 400 bytes corresponding to a throughput of 2.25 Mbps. It is interesting to note that the actual throughput obtained is significantly lower than the supported data rate of 6 Mbps.

Qualnet simulations confirm that the effective throughput obtained using the theoretical model is very close to that obtained by simulations. As can be seen from Figs. 1 and 2, there is a 0.5 dB difference in the channel SNR between the theoretical results and those obtained by simulation. For instance, Fig. 1 shows that the throughput vs. payload variation obtained at a 2 dB SNR using the theoretical model is similar to that obtained from simulation at an SNR of 1.5 dB. This difference can be attributed to the upper bound in Eq. (3).

A proper selection of data rate along with the payload size is crucial for maximizing spectral efficiency. Constant selection of the lowest data rate, say 6 Mbps using BPSK with rate 1/2 convolutional coding, is too conservative an approach, wasting system resources, while selecting a much higher data rate can cause a severe degradation in overall throughput. Another interesting observation is that the region of optimal payload length is very narrow at low SNRs and using different payload lengths leads to a significant lowering of effective throughput. We observe from Figs. 1 and 2 that there are relatively sharp peaks in throughput for lower SNRs while a more gradual transition is observed for higher SNRs. For the 2 dB case in Fig. 1, there is a small range of payload lengths around 300 bytes that can be selected for optimal performance, whereas for the 5.5 dB case, there is a much broader range of payload lengths around 650 bytes. This suggests that the payload length adaptation is more crucial at lower SNRs.

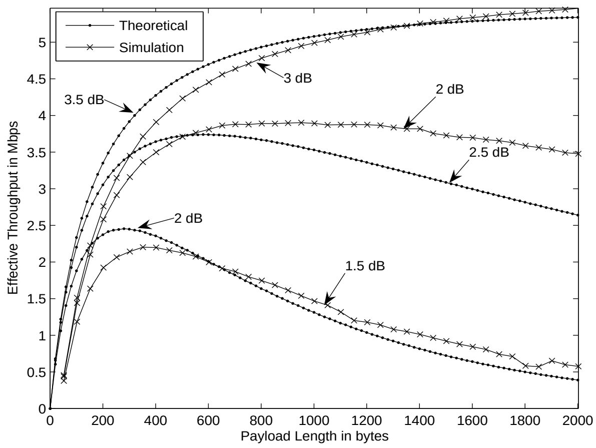


Fig. 1. Throughput versus payload length at 6 Mbps in AWGN channel.

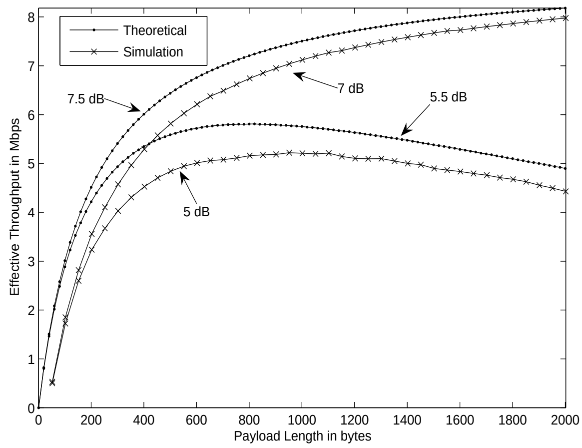


Fig. 2. Throughput versus payload length at 9 Mbps in AWGN channel.

The variable payload scheme discussed can be used to maximize the number of voice and video users supported at different channel conditions. The lower payload lengths of 80 and 160 bytes correspond to 10 ms and 20 ms G.711 frames, respectively, while the higher payload lengths of 200 to 2000 bytes correspond to low bit rate video applications. The number of bits consumed by each frame of a video sequence depends heavily on the video context. The first frame of each scene (Intra-coded, or I frame) usually needs many more bits than the following frames (predicted, or P frames). The ratio varies from about 1.5 for high motion videos to as high as 100 for low motion videos. The typical size of P-frames varies from 100 to 2000 bytes for low bit rate videos. In the next section, we vary the payload length for 64 Kbps multimedia users and study the impact on the capacity of multimedia users with a varying number of retransmissions.

IV. Effect of Retransmissions on Throughput in IEEE 802.11a Wireless Networks

In this section, we extend the analysis of the previous section to incorporate the effect of the IEEE 802.11 retransmission limit on throughput and capacity. We observe that the use of retransmissions in a lossy environment reduces packet loss but can lead to a decrease in the effective throughput. Since multimedia traffic can tolerate a certain amount of packet loss (typically 5-10%), the selection of an appropriate retransmission limit can significantly increase the effective throughput. We investigate the use of selecting different retransmission limits for multimedia and data users in a mixed traffic scenario and observe its effect on the capacity of the network.

Let the instantaneous packet error rate (PER) for a payload size of L bytes at a given SNR γ_s and PHY mode m be given by $P_e^m(L,\gamma_s)$ and N_{max} be the maximum number of allowable retransmissions. In order to meet the packet loss requirement, the overall packet loss after N_{max} retransmissions should be less than the packet loss threshold P_{MAX}:

$$P_e^m(L,\gamma_s)^{N_{max}+1} \leq P_{MAX} \tag{8}$$

The effective throughput after N_{max} retransmissions (under the assumption of an identical wireless channel for retransmitted packets as original transmissions) can be evaluated as:

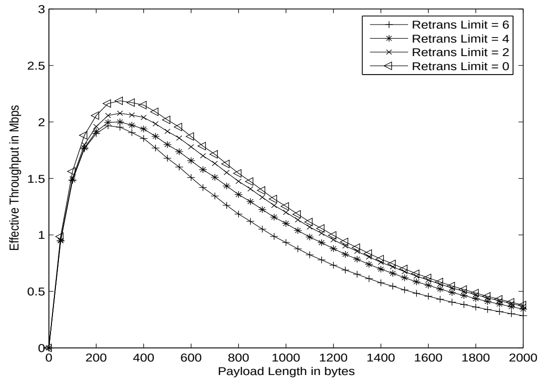
$$T(m) = \frac{L * (1 - P_e^m(L,\gamma_s)^{N_{max}+1})}{E[Time]} \tag{9}$$

$E[Time]$ is the expected packet transmission time for $N_{max}+1$ transmissions and can be approximated as:

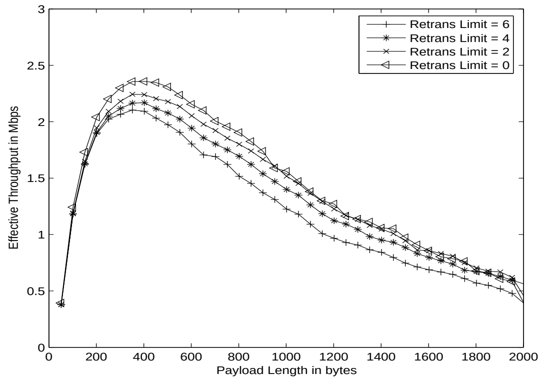
$$\begin{aligned} E[Time] \quad = \quad & \sum_{i=1}^{N_{max}+1} P_e^m(L,\gamma_s)^{(i-1)} * (1 - P_e^m(L,\gamma_s)) * T_i \\ & + (P_e^m(L,\gamma_s))^{N_{max}+1} T_{fail} \end{aligned} \tag{10}$$

where T_i is the time taken to successfully transmit a packet after i transmissions and T_{fail} is the time spent in N_{max} unsuccessful retransmissions.

IEEE 802.11 utilizes retransmissions to reduce unicast packet loss in the wireless medium. The retransmission limit thus affects the packet loss in the medium. However, every transmission attempt that results in a loss causes an exponential increase in the contention window size. We show that in a network with little packet loss due to collisions (few users connected to an access point), decreasing the retransmission limit can provide significant throughput gains at the expense of packet loss in the medium. Multimedia applications, such as voice and video streaming, can typically tolerate 5-10% packet loss while still maintaining acceptable quality. Thus an appropriate choice of retransmission limit can provide throughput gains. However, as the number of users increases, decreasing the retransmission limit provides little or no gain in throughput and can lead to intolerable packet loss in the network. We analyze the impact of setting different retransmission limits at multimedia and data users and show that the multimedia user capacity depends on the retransmission limit at a given data rate. Significant gains in the multimedia user capacity can be obtained with an appropriate choice of retransmission limit.



(a) Throughput versus payload length at an SNR of 2 dB with 6 Mbps data rate using the theoretical analysis.



(b) Throughput versus payload length at an SNR of 1.5 dB with 6 Mbps data rate using Qualnet simulations.

Fig. 3. Effect of retransmissions on throughput using theoretical analysis and Qualnet simulations

We first analyze the impact of the IEEE 802.11 retransmission limit on the throughput obtained in a single-user network. Since the network has a single user, there are no collisions and the packet loss is entirely due to channel conditions. Fig. 3 shows the effective throughput vs payload length for different retransmission limits using the theoretical analysis and simulations. Fig. 4 shows the loss variation using different retransmission limits for a data rate of 6 Mbps at an SNR of 1.5 dB obtained by simulations. As seen in Fig. 3(b), there is a gain of up to 20% in throughput with a lower retransmission limit in a lossy channel (1.5 dB). This is because each loss results in an exponential backoff until the packet is either successfully delivered or the retransmission limit is reached.

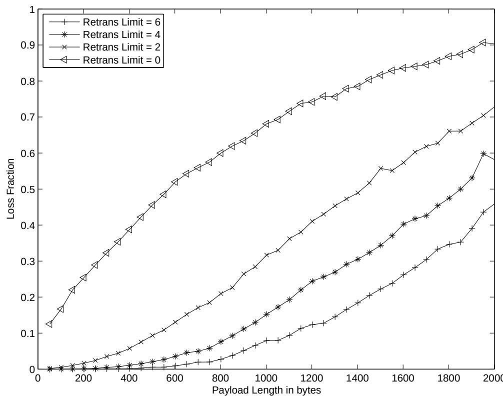


Fig. 4. Loss versus payload length at an SNR of 1.5 dB with 6 Mbps data rate.

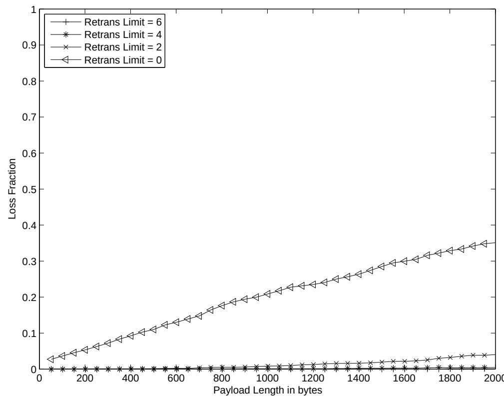


Fig. 6. Loss versus payload length at an SNR of 8 dB with 18 Mbps data rate.

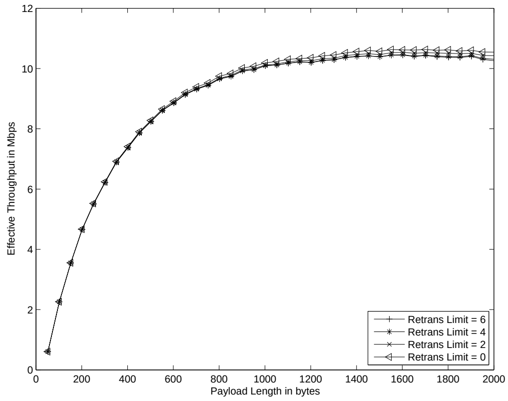


Fig. 5. Throughput versus payload length at an SNR of 8 dB with 18 Mbps data rate.

With a low retransmission limit, there are fewer backoffs and hence less idle time in the medium. This results in more packets transmitted per unit time. However, there is a trade-off between the throughput and loss. The gain in throughput is at the expense of packet loss in the network. Fig. 4 shows the loss at different retransmission limits. With no retransmissions, the loss rate is high and can be intolerable to many applications. However, voice and video streams are likely to be able to tolerate certain degree of loss (up to 5 to 10%). Fig. 4 also shows that for a given payload length used by the application, an appropriate retransmission limit can be chosen to suit the application requirements and obtain the highest throughput in the network. The network can thus obtain higher throughput with the use of a low retransmission limit that meets the packet loss constraint of the application. The payload length also determines the gain in throughput obtained. At very small payload lengths (under 100 bytes), there is little gain in throughput obtained with a low retransmission limit.

The above experiments were carried out with a fixed transmission rate. When the network is using rate adaptation, a less conservative retransmission scheme that allows a certain amount of packet loss can benefit the network. An application-aware rate adaptation scheme that operates at higher rates even in presence of packet loss, as allowed by the application, can provide even more significant gains in throughput.

Figs. 5 and 6 represent the throughput and loss variation of a single-user network at an SNR of 8 dB with 18 Mbps. In this scenario, there is little difference in the throughput obtained with different retransmission limits. Thus, the retransmission limit plays little role in a network with low loss. However, it has to be noted that the experiments were carried out in a collision free environment. The same experiment was repeated in a network with 50 users. It was observed that beyond the capacity limit, the maximum achievable throughput with different retransmission limits is approximately the same. However, the loss in the network with no retransmissions is as high as 10-15%, while the loss in the network with a retransmission limit of three or greater is 0%. This highlights the importance of a higher retransmission limit in a network operating under load with many users.

We next analyze the impact of different retransmission limits for data users and multimedia users. We consider a mixed user network, with data users running FTP and multimedia users running a constant bit rate stream of 64 Kbps. Data transfer applications such as FTP cannot tolerate any loss, and thus any loss at the MAC layer results in end-to-end retransmissions at the upper layer (TCP). Since end-to-end retransmissions are expensive, we use the highest retransmission limit at the MAC layer for data users. Multimedia applications such as voice and video streaming, however, can tolerate some loss in the network. The tolerable loss is specific to the application and depends on the codec used. Assuming the tolerable limit to be 10%, we study the impact of varying the retransmission limit at the multimedia users on the capacity of the network.

Figs. 7 and 8 show the capacity results for multimedia users with one and five data users, respectively. The number of multimedia users supportable with an average loss of less than 10% represents the capacity of the network. As seen in Fig. 7, in a network with a single user running an FTP application, the capacity variation with different retransmission limits at multimedia users is minimal. However, with five data users in the network, as seen in Fig. 8, there is a significant gain in

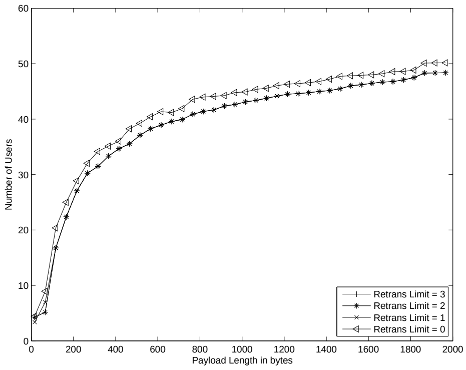


Fig. 7. Capacity gain with 1 data user in the network (SNR is 8 dB and data rate is 6 Mbps).

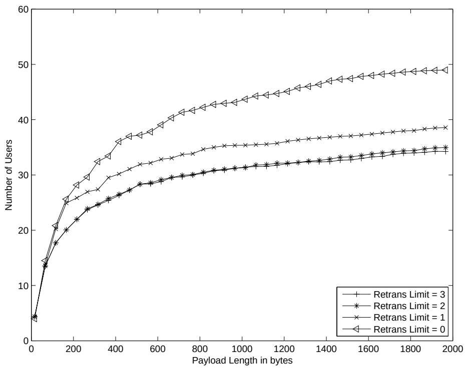


Fig. 8. Capacity gain with 5 data users in the network (SNR is 8 dB and data rate is 6 Mbps).

capacity with the use of a lower retransmission limit at the multimedia users. With a payload size of 1000 bytes, there is a 50% gain in capacity if no retransmissions are allowed compared to setting a higher retransmission limit of 4 or higher. With a lower retransmission limit at the multimedia users, there are fewer exponential backoffs, and thus there is prioritized channel access for the multimedia users over the data users. This is at the cost of packet loss, which is assumed to be tolerable up to 10% for these users. As mentioned earlier, these experiments were carried out for fixed data rate, and the choice of retransmission limit depends on the data rate chosen. A rate selection scheme that adjusts the retransmission limit appropriately for different users can provide significant capacity gains for the multimedia users.

We now look at the impact of different retransmission limits on the throughput of data users. Fig. 9 shows the impact of varying the retransmission limit for the multimedia users on the throughput of data users for a payload length of 1200 bytes and a data rate of 6 Mbps. The number of multimedia users supported at different retransmission limits is as follows: 31 multimedia users with a retransmission limit of 3, 32 multimedia users with a retransmission limit of 2, 35 multimedia users with a retransmission limit of 1, and 43 multimedia users with a retransmission limit of 0. As observed in Fig. 9, the capacity gain of multimedia users with lower

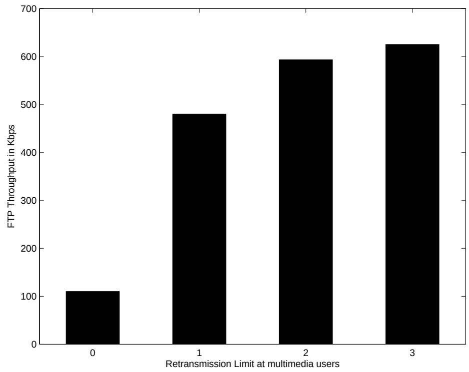


Fig. 9. Effect of retransmission limit on the throughput of data users.

retransmission limits is at the expense of the throughput of the data users.

It is thus clear that setting an appropriate retransmission limit results in the support of more multimedia users in the network. This comes at the expense of the throughput of data users and depends on the allowable packet loss by the multimedia stream. As discussed earlier, the retransmission limit selection depends on the data rate and the current channel loss in the network. The selection of this limit can be integrated with the data rate selection to suit the application requirements in terms of latency and loss.

V. Conclusions

In this paper, we explore schemes to improve the effective throughput of multimedia users over a wireless link. Careful adaptation of PHY and MAC layer parameters, namely, data rate, payload length and maximum number of allowable retransmissions has a significant impact on the effective throughput. Varying the data rate and payload length based on the channel conditions can significantly improve the effective throughput, thereby leading to a greater number of supported multimedia users. We also observe that by controlling the maximum number of retransmissions, a greater number of multimedia users can be supported in the presence of data users.

References

[1] D. Newman, "Voice over Wireless LAN." Available on the Internet: http://www.networkworld.com/reviews/2005/011005rev.html, January 2005.
[2] P. Lettieri and M. B. Srivastava, "Adaptive Frame Length Control for Improving Wireless Link Throughput, Range and Energy Efficiency," in *Proceedings of IEEE INFOCOM'98*, pp. 564–571, March 1998.
[3] D. Qiao, S. Choi, and K. G. Shin, "Goodput Analysis and Link Adaptation for IEEE 802.11a Wireless LANs," *IEEE Transactions on Mobile Computing (TMC)*, vol. 1, pp. 278–292, Oct-Dec 2002.
[4] Qualnet, http://www.scalablenetworks.com, 2005.
[5] J. P. Pavon and S. Choi, "Link Adaptation Strategy for IEEE 802.11 WLAN via Recieved Signal Strength Measurement," in *Proceedings of IEEE ICC '03*, pp. 1108–1113, May 2003.
[6] M. B. Pursley and D. J. Taipale, "Error Probabilities for Spread-Spectrum Packet Radio with Convolutional Codes and Viterbi Decoding," in *IEEE Transactions on Communications*, vol. COM-35, pp. 1–12, Jan. 1987.
[7] C. Lee and L. H. C. Lee, *Convolutional Coding: Fundamentals and Applications*. Artech House Publishers, 1997.