FAIR DELAY OPTIMIZATION-BASED RESOURCE ALLOCATION ALGORITHM FOR VIDEO TRAFFIC OVER WIRELESS MULTIMEDIA SYSTEM

[1]M. R. M. Rizk, Senior Member, IEEE, [2]M. I. Dessouky, [3]Sami A. El-Dolil, and [4]Mohammed Abd-Elnaby

[1]Department of Electrical Engineering, Faculty of Engineering, Alexandria University, Alexandria, Egypt, mrmrizk@ieee.org.
[2,3,4]Department of Electrical Communication, Faculty of Electronic Engineering, Menoufia University, Menouf, B.O. BOX 32952 Egypt, dr_moawad@yahoo.com, msel_dolil@yahoo.com, mohetim2002@yahoo.com .

ABSTRACT

The major issue related to the realization of wireless multimedia system is the design of suitable medium access control (MAC) protocol. The design challenge is to maximize the utilization of the limited wireless resources while guaranteeing the various quality of service (QoS) requirements for all traffic classes especially for the stringent real-time constraint of real time variable bit rate (rt-VBR) video service. In this paper a novel resource allocation algorithm for video traffic is proposed. The proposed allocation algorithm aims to provide fair delay for video packets by minimizing the delay difference among transmitted video packets. At the same time it adaptively controls the allocated resources (bandwidth) for video traffic around the corresponding average bit rate, and has the ability of controlling the QoS offered for video traffic in terms of packet loss probability and average delay. A minimized control overhead of only two bits is needed to increase the utilization efficiency. Simulation results show that the proposed algorithm achieves very high utilization and provides nearly fair delay among video packets. Its efficiency is also investigated under traffic integration condition with voice and data traffic. Under traffic integration condition the data traffic increases the utilization while a good QoS is achieved for all traffic classes.

Keywords: MAC protocols, Resource Allocation, QoS , Performance Evaluation.

1 INTRODUCTION

In wireless multimedia system, the medium access control (MAC) protocol should play a central role in providing a statistically multiplexing capability over the wireless access interface and maximizing the utilization of limited wireless resources. Its design objective is to maximize the multiplexing gain over the wireless interface while guaranteeing the various QoS requirements for all multimedia traffic classes, especially for the stringent real-time constraint of real time variable bit rate (rt-VBR) video service [1].

The dynamic slot assignment become essential for achieving the statistically multiplexing gain among the various types of service classes with different traffic characteristics and has been recognized as a useful choice for the variable bit rate (VBR) traffic class, since it simply serves to coordinate the varying traffic demands among the independent and spatially distributed wireless terminals [2]-[4].

Existing MAC protocols are typically involved with an excessive overhead to transmit the buffer state information, e.g., instantaneous queue length, residual lifetime, or packet arrival rate as the dynamic parameter in a timely accurate manner. Two different types of access schemes are used to transfer the dynamic parameter information from the individual distributed wireless terminals to the access point: contention-based and contention less (contention-free) schemes.

In the contention-based scheme [5]-[8] reducing the collision probability and the size of control packet is necessary to reduce wastage of bandwidth in case of collision. On the other hand, the contention-free scheme[9]-[12] is based on either a piggyback mechanism, which piggybacks the dynamic parameters on the uplink data burst, or polling mechanism, which dynamically assigns an uplink dedicated slot for transferring the dynamic parameters. The piggybacking overhead should be minimized and the polling period should be efficiently adjusted to reduce the wastage of bandwidth.

In this paper a novel resource allocation algorithm for video traffic is proposed. The proposed allocation algorithm aims to provide fair delay for video packets by minimizing the delay difference among transmitted video packets. At the same time the proposed algorithm adaptively control the allocated resources (bandwidth) for video traffic around the corresponding average bit rate and has the ability of controlling the quality of service (QoS) offered for video traffic in terms of packet loss probability and average delay. A minimized control overhead of only two bits is needed to increase the utilization efficiency. Its efficiency is also investigated under traffic integration condition with voice and data traffic.

The paper is organized as follows. Section2 presents the system description and traffic models. In section 3, a detailed description of the proposed resource allocation algorithms for various traffic types are presented. Simulation results and discussion of protocol performance are given in Section 4. Finally, section 5 gives the conclusion.

2 SYSTEM DESCRIPTION AND SIMULATION ENVIRONMENT

In this section, we explain the simulation environment in detail. First, we provide a description of the air interface characteristics, and then the traffic models. A MATLAB simulator has been written to evaluate the performance of the proposed algorithms. The simulator exactly implements the allocation algorithms described in section 3.

2.1 Air Interface Frame Structure

A micro cellular environment with negligible propagation delay (in micro second) is investigated. The proposed protocol uses frequency division duplex (FDD) even if our proposed protocol could be also adopted for the time division duplex (TDD) scheme. The air interface frame has a fixed size (2 msec), in order to simplify the data scheduling. It is divided into equal size slots, each slot carries ATM packet of 53 byte (5 bytes header and 48 byte payload) resulting from ATM adaptation layer (AAL) type 5. The uplink channel bit rate is 4.9 Mbps with 24 slots per frame.

The uplink frame is divided into control and data transmission periods.The control period is further divided into contention and polling periods. Control slots assigned in the control period are further subdivided into four control minislots, some of them used as contention minislots for voice request access and the others used as polling minislots for data traffic. The contention minislots are used by voice users to send their reservation requests in contention mode at the beginning of talk spurt (when the connection becomes active). While, the poling minislot are allocated for data users to send their buffer length status to the base station.

2.2 Traffic Models

2.2.1 Video Model

The real video sources such as MPEG video encoders generate traffic flow, which discretely varies with time among a fixed number of possible rates. The discrete rate video source can be modeled by a superposition of 15 ON/OFF sources. Each ON/OFF source alternates between the ON and OFF states. Durations of ON and OFF states are exponentially distributed with means 100msec and 200 msec respectively [8], [13]. So, the resulting average bit rate is 250 Kbps. The maximum transfer delay (MTD) of video packet is set to be 50 msec.

2.2.2 Voice Model

The voice source generates a signal that follows a pattern of talkspurts separated by silent gaps [8]. Therefore, an ON/OFF model can describe a voice source: the source alternates between the ON state where the source generates packets at rate 16kbps, and the OFF state where no packets are generated. Durations of talkspurts and silent gaps are modeled as exponential distributions with mean values of 1 and 1.35 sec, respectively. If a voice packet is not sent within its maximum transfer delay (MTD=24 msec), it should be dropped.

2.2.3 Data Model

The data source generates messages of certain length arriving at the mobile station buffer at a certain time. The length of the message is exponentially distributed with mean 2 kbit, and the interarrival time between messages is exponentially distributed with mean 100 msec. The maximum transfer delay (MTD) of data packet is set to be 60 sec [8].

3 RESOURCE ALLOCATION ALGORITHMS

Efficient resource allocation algorithm should be able to satisfy the required QoS and at the same time achieves high utilization of the wireless channel. The resource allocation algorithms for uplink transmission are only considered since the downlink transmission can be easily scheduled TDM mode. As different wireless multimedia services share the same resources, effective traffic integration is needed to maximize the utilization efficiency of the shared resources. So, the sequence of allocation process starting with voice traffic (highest priority) then video traffic and finally data traffic (lowest priority).

3.1 Resource Allocation Algorithms for Video Traffic

The base station broadcast upper

delay_threshold and lower delay_threshold to all video users every frame. Two bits piggybacking overhead are used to indicate the delay time of the transmitted packet relative to upper delay_threshold and lower delay_threshold.

The piggybacking overhead indicates three packet delay status:

1. Higher state when the packet delay is higher than the upper delay_threshold
2. Lower state when the packet delay is lower than the lower delay_threshold.
3. In_between state when the packet delay is lower than the upper delay_threshold and higher than the lower delay_threshold.

The allocation algorithm aiming to maximizing the number of in_between delay state users by converting the higher and lower delay state into in_between delay state users to achieve fair delay based allocation. So the allocation process performed on a frame-by-frame basis and proceeds as follows:

Step1: Increasing the number of allocated slots by one for higher delay state users.
Step2: Keeping the number of allocated slots unchanged for in_between delay state users.
Step3: Decreasing the number of allocated slots by one for lower delay state users.

The number of allocating slots for each user in the above three steps mainly depends on the availability of free slots in the current frame. So the lower delay state users put in the last step since they are less affected by the reduction of the available free slots than the other users.

When a user has no allocated slots inside the frame he goes to a wait state until he has allocated slots. If he stays in the wait state three frames the base station allocates one slot at least for him in the current frame to send one packet and updates its delay state information.

The users in step 3 are sorted according to their waiting states and the users of higher waiting time are served first for more efficient fair allocation.

The base station adjusts the upper and lower delay_threshold according to the following equations:

$$Up\_D\_Th = Lo\_D\_Th + Th\_Diff \qquad \text{(Second)} \quad (1)$$

$$Lo\_D\_Th = (R_{slot} * EXCESS * T_F)/(N_v * R_v) \qquad \text{(Second)} \quad (2)$$

$$EXCESS = EXCESS + \text{Number of allocated slots} - \eta MEAN \qquad \text{(slots)} \quad (3)$$

$$MEAN = (N_v * R_v)/R_{Slot} \qquad \text{(slots)} \quad (4)$$

Where:
Up_D_Th, Lo_D_Th: the upper and lower delay_threshold respectively.
Th_Diff: Threshold_difference.
N_v: the number of video user.
R_v: the average bit rate per video user.
R_{Slot}: the bit rate per slot.
MEAN: the number of slots equivalent to the total mean bit rate of all video users.
EXCESS: the excess number of slots allocated for video traffic than ηMEAN.

From the above equations we find that the resource allocation algorithm is aiming to adjust the average number of allocating slots for video traffic around ηMEAN by controlling the value of upper and lower delay_threshold.

So, when the EXCESS slots increase the base station increases controlling the value of upper and lower delay_threshold according to the above equation. This leads to reducing the number of higher delay state users and increase the number of lower delay state users, hence the number of allocated slots are reduced.

The value of threshold_difference is responsible of controlling the average difference of packet delay between any transmitted packets. Hence it should be very small as possible so we select it to be 4 msec (2 frames).

The lower delay_threshold has a minimum value of zero while the upper delay_threshold has a maximum value equals to the maximum transfer delay (MTD) of video packet (50 msec).

3.2 Resource Allocation Algorithms for Voice Traffic

At the beginning of talkspurt the voice user sends a reservation request through the contention minislot in contention mode by randomly selecting one of the available contention minislots with equal probability. Voice users that experience collisions will repeat the contention process in subsequent frames until they are successful.

When the base station successfully receives the request, it periodically allocates slots to the connection up to the end of talk spurt. At the end of talkspurt the user sets a one-bit flag in the last transmitted packet to indicate that the connection is no longer active.

3.3 Resource Allocation Algorithms for Data Traffic

Polling control minislots are used by data users to send their buffer length to the base station. The number of polling minislots is selected such that the polling period should be lower than or equal to the interarrival time between data messages (100 msec) to enable the base station to efficiently monitor the buffer length status of each data user.

At the base station, a token pool is introduced for each data user. Tokens are generated at a fixed rate that is equal to the mean packet generation rate of the data source and removed for every slot allocated to the user.

The number of slots allocated for a user is the minimum of the buffer length of that user and the number of token in its token pool, and the user with higher number of tokens in its token pool sends their packets earlier within the frame.

4. PERFORMANCE EVALUATION AND SIMULATION RESULTS

4.1 Performance Evaluation with Video Traffic

This set of simulations has been run with video users only to evaluate the performance of the novel video resource allocation algorithm that we have presented in this paper. First simulation results taken for a limited values of η (η= 1, 1.01, 1.02, 1.03, 1.04) and a wide rang of the number of users and all results are presented as a function of the number of video users.

Fig. 1 and Fig. 2 show that a very small increase in the value of η produces a significant reduction in the packet loss probability and average packet delay. Since as the value of η slightly increases the average values of delay_thresholds decreases as shown in Fig. 3 which means that the allocated slots for video traffic slightly increase. The allocation algorithm efficiently utilize the allocated bandwidth and produce a significant reduction in the packet loss probability and the average packet delay which follow the delay_thresholds values as long as the available bandwidth is sufficiently enough. As the number of video users increase the effect of η become more significant since the multiplexing gain increase and the total rate variation of video traffic reduces. So a higher value of η is needed with low number of users than with high number of users to achieve the same QoS.

Two factors could reduce delay_thresholds values: the first is increasing the value of η as we explain above, and the second when the allocated slots reduced due to reduction of the available free bandwidth. In this later case the number of higher delay state users significantly increases and the allocation algorithm tries to convert the most of them into in_between delay state as possible which leads to decrease the average packet delay but a greater number of packets will be lost due to reduction of the available free bandwidth and this situation could be observed at 18 video. Also, it should be considered that with very small values of delay_thresholds the allocation algorithm could not distinguish the more urgent delay users.

Fig. 4 presents the actual throughput, which is defined as the ratio of the average number of slots used for successful data packet transmission to the total number of slots per frame and it is approximately the same for all values of η.

The second simulation results are taken to specify the effect of η over a larger range (1: 1.2). We measure the performance at 4 users and 16 users as an example of low and high number of users and all results are presented as a function of η.

Figures 5, 6, and 7 Show that as the value of η increases the average values of delay_thresholds decrease, hence the allocated slots for video traffic gradually increase which in turn reduces the average packet delay and consequently the packet loss probability decreases. But this process continues until the delay_thresholds reaches their minimum values and as long as there is enough bandwidth to satisfy the allocation needs. Also, it is obvious that high number of video users highly effected by η than low number of video users due to increasing the multiplexing gain as the number of video users increase.

The main drawback of increasing η is shown in Fig. 8 since as η increases the wasted bandwidth (slots) increases. The wasted slots are the unused reserved slots, as η increases the allocated slots for video traffic increases and some of them become higher than the user needs. The wasted bandwidth reaches its maximum values when the delay_thresholds reaches their minimum values since at this time a maximum bandwidth is allocated. The wasted bandwidth at high number of video users is greater than at low number of video users since the allocated bandwidth is greater at high number of video users and this is a good property since at low number of video users a better QoS requires higher value of η than at high number of video users which means that at given QoS the wasted bandwidth is nearly the same at high and low number of users.

Fig. 9 shows that the throughput is independent on η and remains constant as the value of η increases. Fig. 10 Shows that the average of packet delay difference lower than 2 msec for all values of η and it slightly decreases as the value of η increase and since the delay_thresholds reaches their minimum values earlier at 16 user also the average of packet delay difference also reaches its minimum value earlier at 16 video users while stile decreasing at 4 video users.

To specify the fairing of the allocation algorithm Fig. 11 presents the probability that the packet delay difference lower than or equal to 4, 6, and 8 msec. It shows that the probabilities are higher than 0.9 and increases as η increases. Also, the probabilities at 16 users are higher than at 4 users and they become closer as the packet delay difference increases (e.g. at 8 msec). This means that the packet delay difference of greater than 90 % of the transmitted packets is lower than or equal to

Threshold_difference (4 msec) while the rest approximately has a packet delay difference lower than or equal to 8 msec, which indicates that the proposed allocation algorithm achieves near fair packet delay for video traffic.

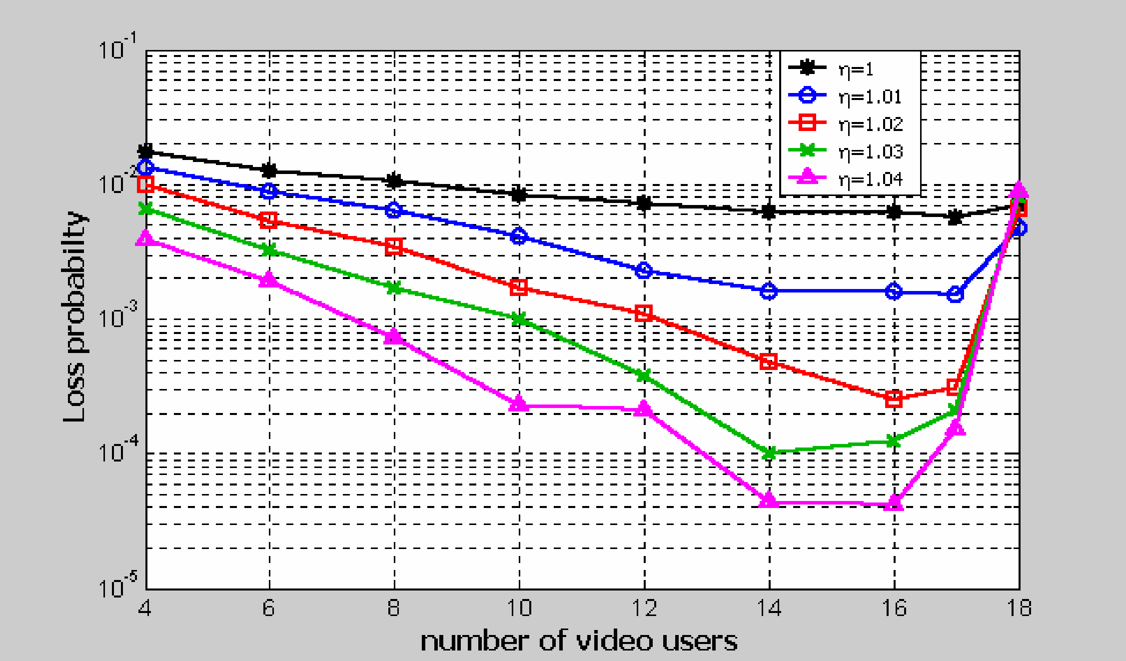

Figure 1: Packet loss probability as a function of the number of video users.

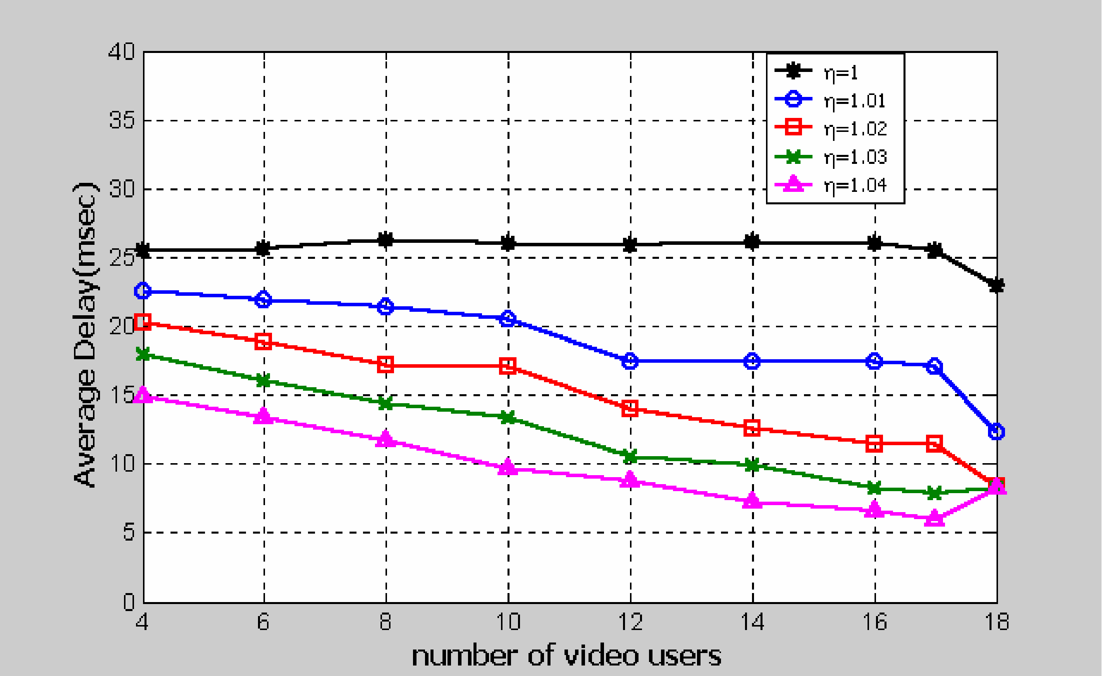

Figure 2: Average packet delay as a function of the number of video users.

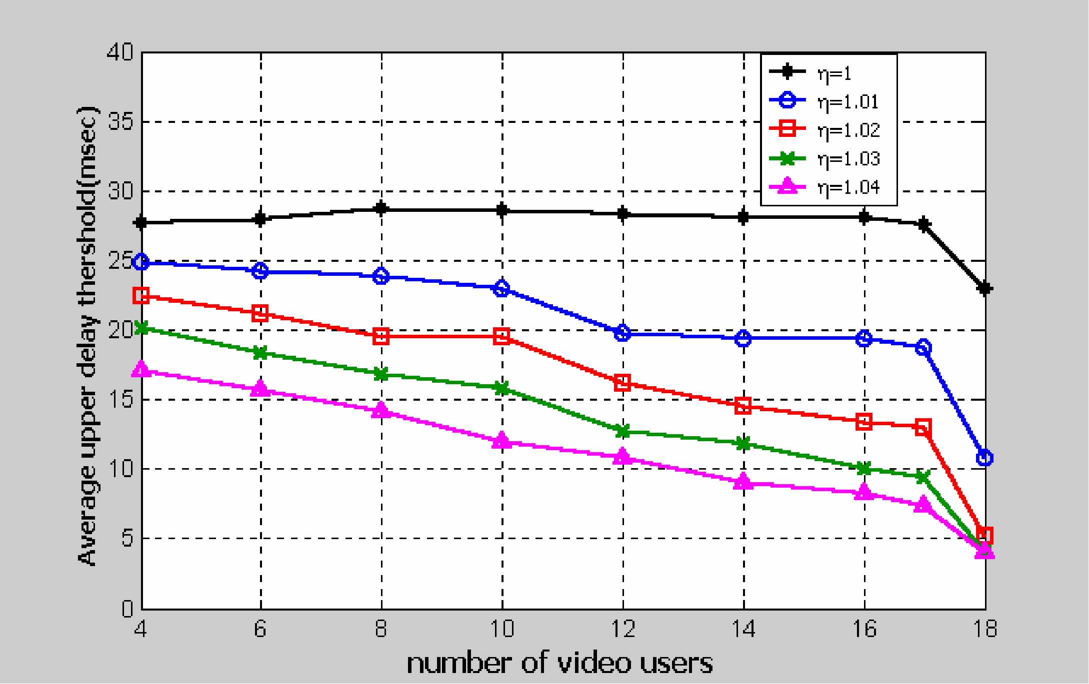

Figure 3: Average of upper delay_threshold as a function of the number of video users.

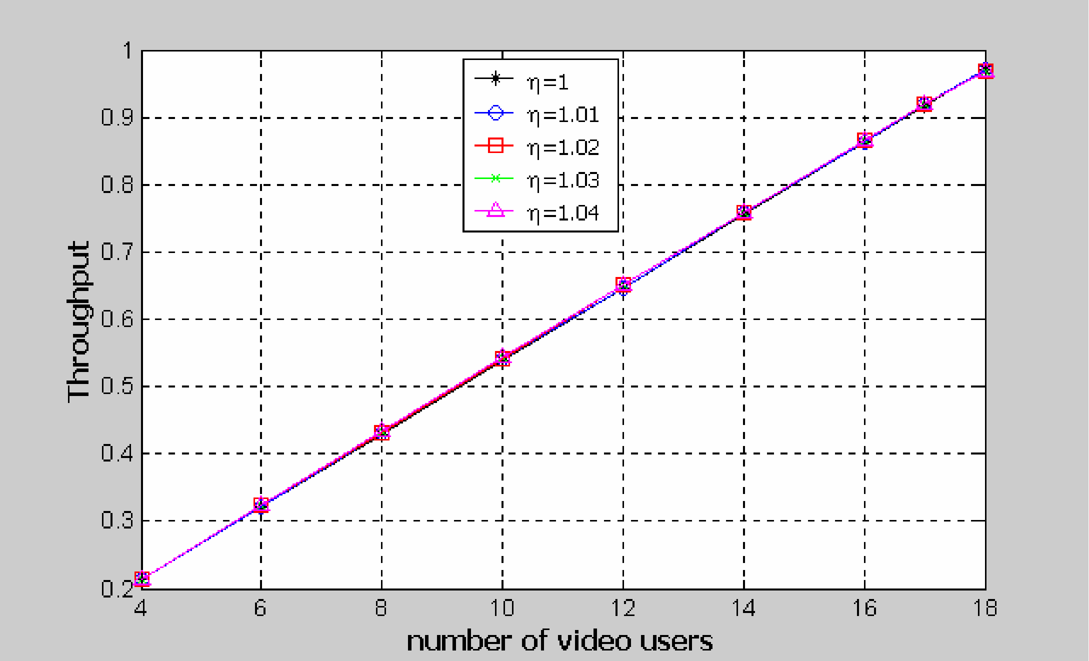

Figure 4: Throughput as a function of the number of video users.

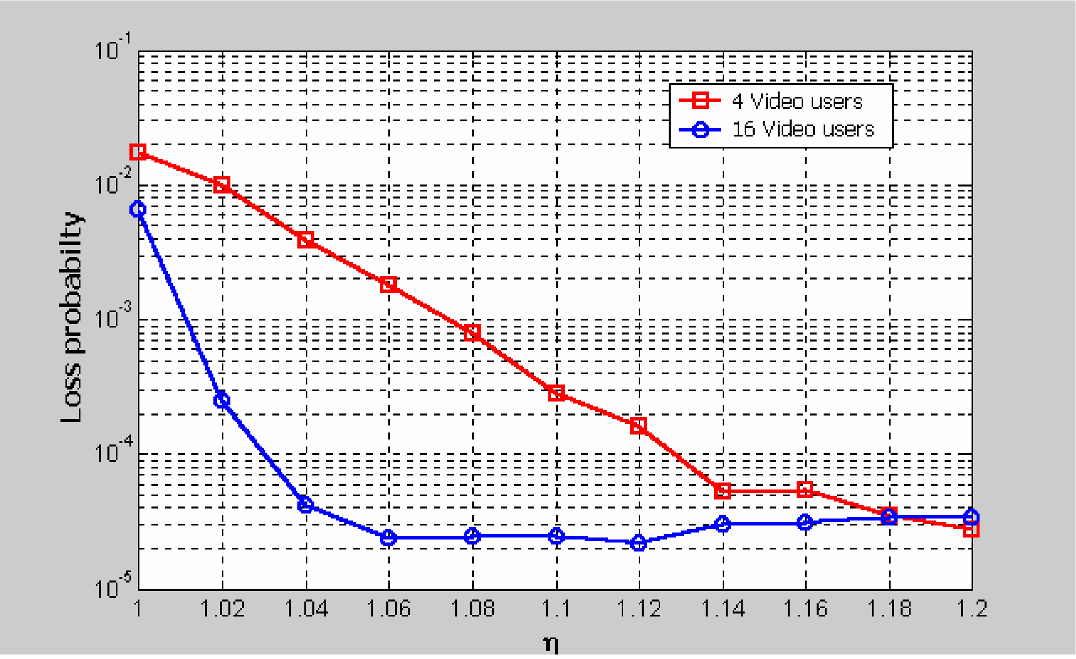

Figure 5: Packet loss probability as a function of η

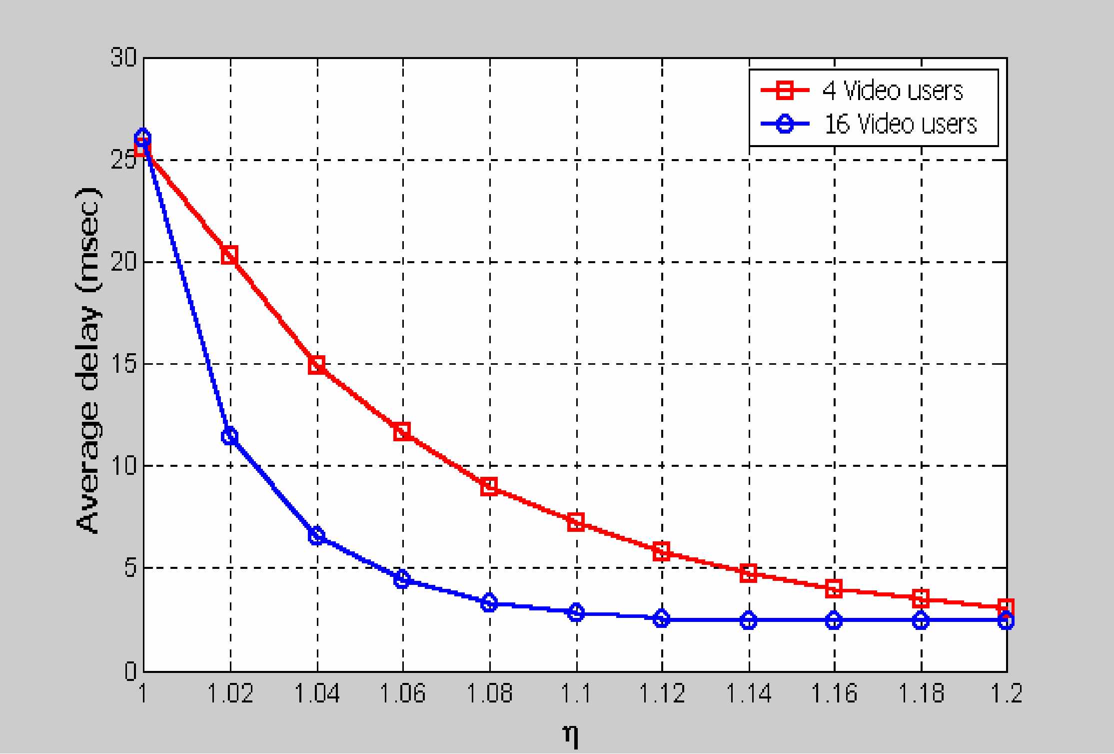

Figure 6: Average packet delay as a function of η

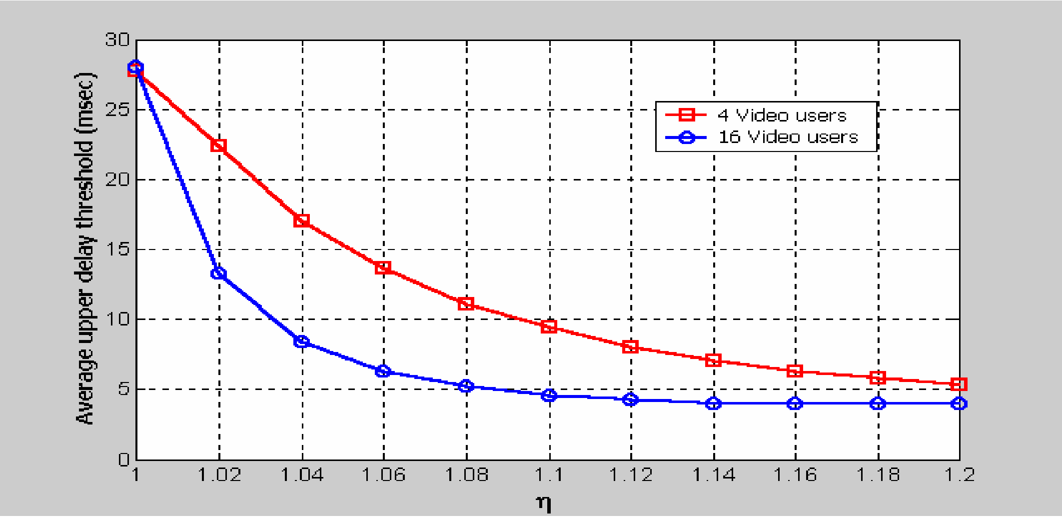

Figure 7: Average of upper delay_threshold as a function of η

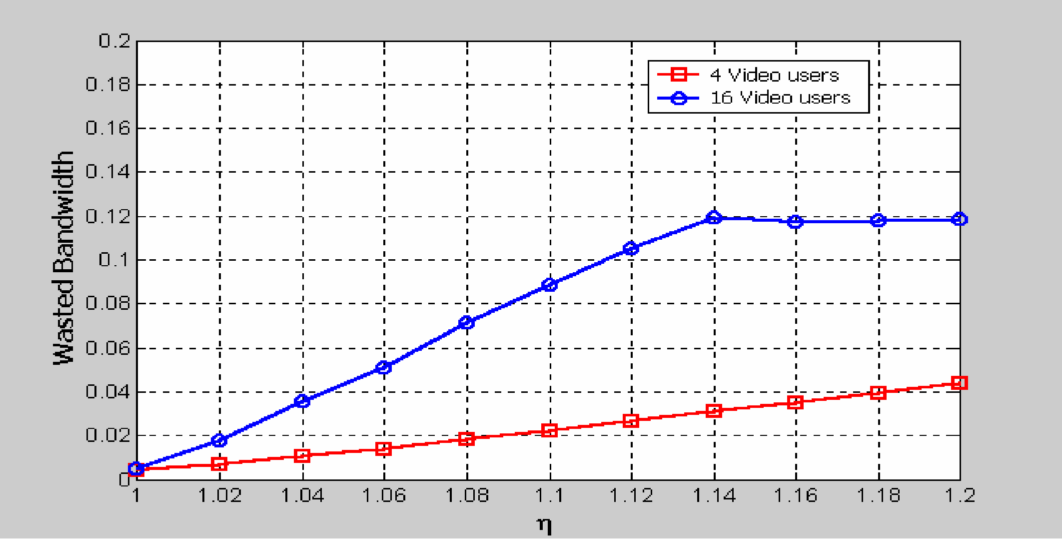

Figure 8: The wasted bandwidth as a function of η

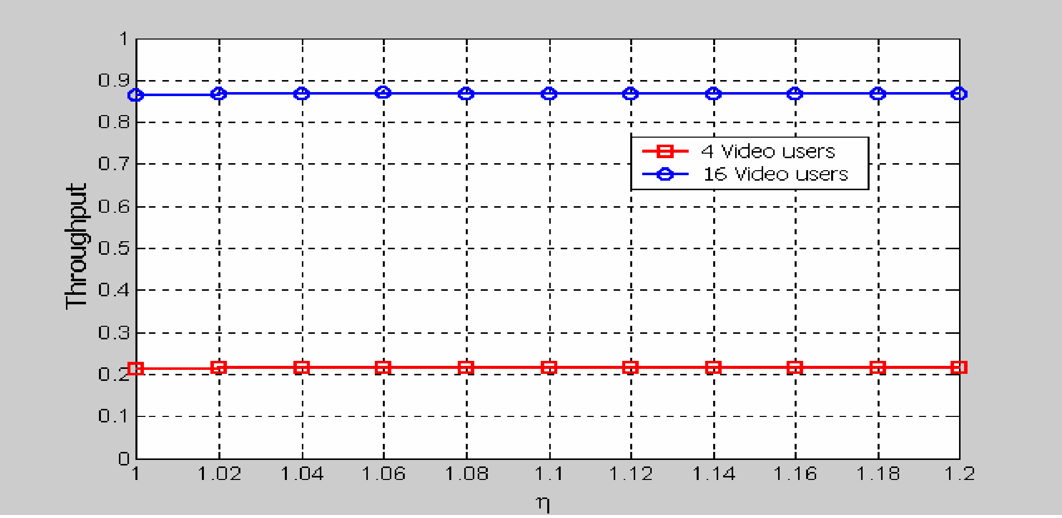

Figure 9: Throughput as a function of η

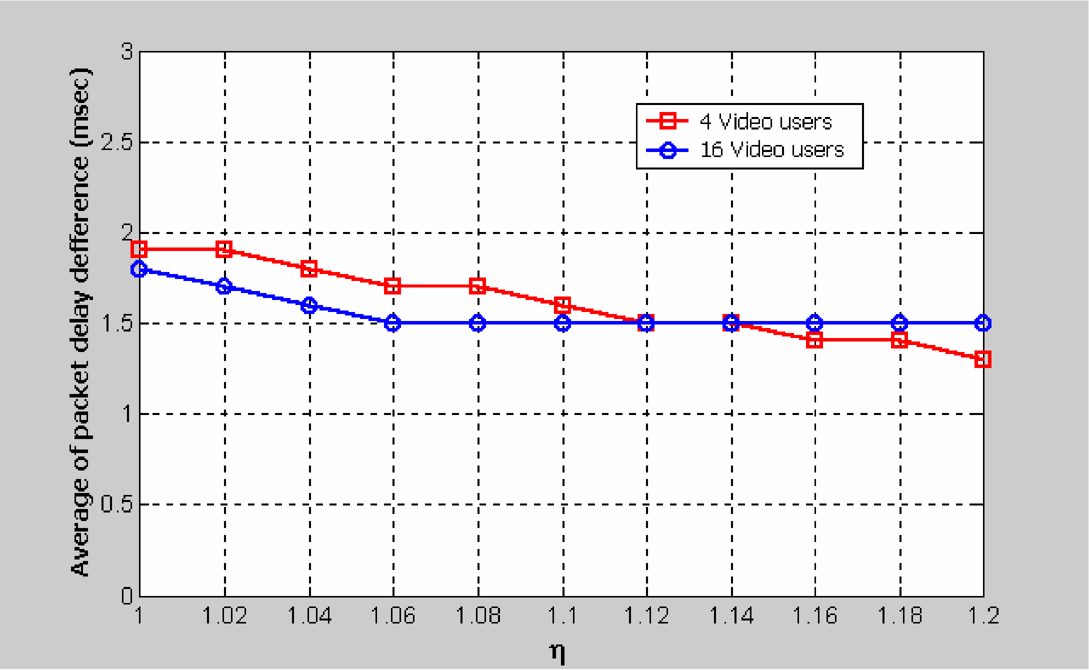

Figure 10: Average of packet delay difference as a function of η.

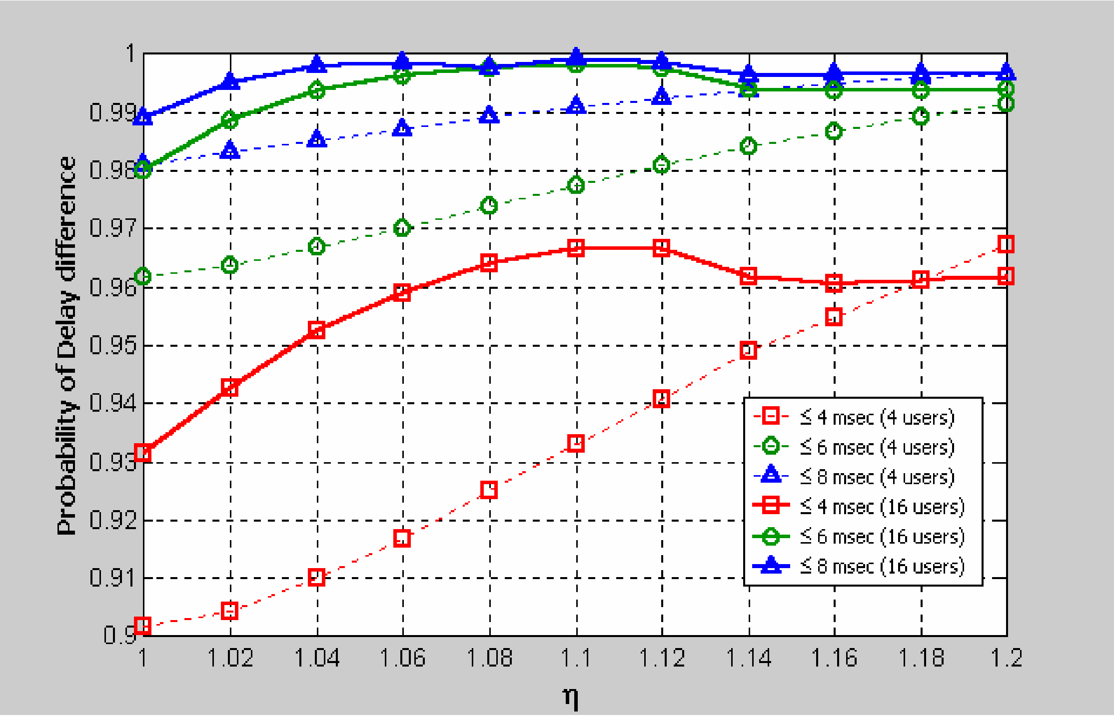

Figure 11: Probability of packet delay difference lower than or equal to 4, 6, and 8 msec as a function of η.

For comparison purposes, the user and channel parameters used for simulations with video traffic have been set similar to those used in [8]. The slot allocation algorithm for video traffic in [8] requires a higher piggybacking overhead of five bits to predict the packet generation rate of the video source and when a packet arrive from the video source, a rate control algorithm is required to determine if the packet conform with the video traffic parameter. If it conforms it will be tagged as a guaranteed packet and will receive priority service which guarantees a high QoS using time-to-expiry algorithm for slot allocation. Otherwise, the packet is tagged as a best effort packet and will receive a low QoS.

Fig. 12 to Fig. 14 present the simulation results for the loss probability, average packet delay, and throughput obtained for video user with average rate of 250 Kbps and at η=1.04. Fig. 12 shows that our protocol achieves a great reduction in loss probability than that in [8] which means a significant better QoS (with only 2% reduction in allocation efficiency). Also, if we take the target loss probability of 0.004 our protocol can support 26 video users with packet delay lower than 7 msec while [8] can support 24 users. Although the protocol in [8] provides slightly lower packet delay than our protocol as shown fig. 13; the achieved packet delay by our protocol is still considerably low. Fig. 14 shows that, both protocols achieve the same throughput but when the system become overloaded our protocol achieve slightly higher throughput.

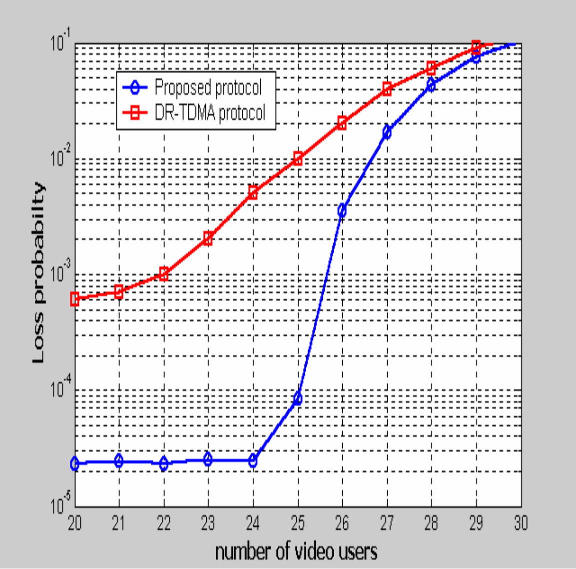

Figure 12: Packet loss probability as a function of the number of video users

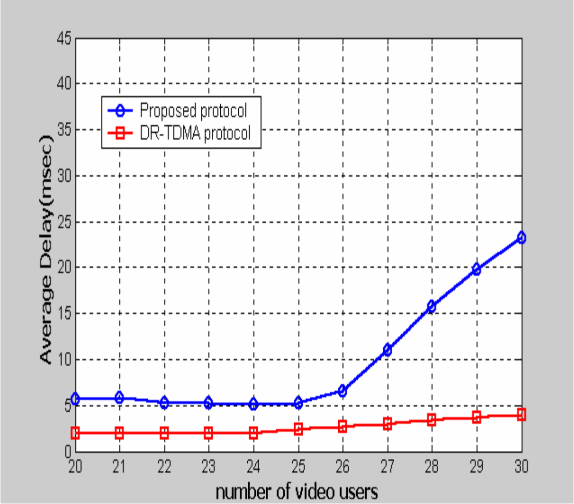

Figure 13: Average packet delay as a function of the number of video users.

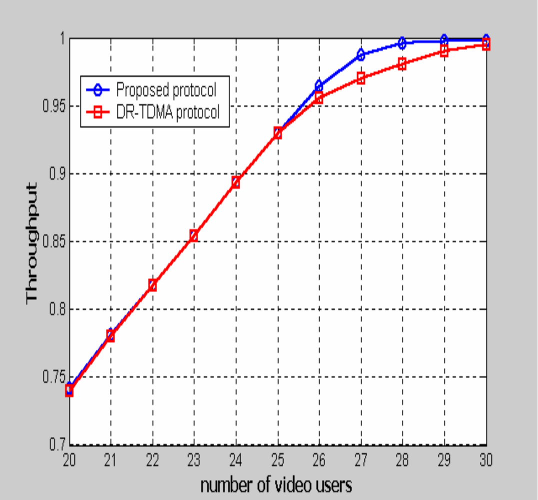

Figure 14: Throughput as a function of the number of video users.

4.2 Performance Evaluation with Integrated Traffic

This set of simulations has been run with 45 voice users and 12 video users while the data users are added gradually. All results are taken at η=1.02 and presented as a function of the number of data users.

Fig. 15 and Fig. 16 show that adding the data user to the system has no effect on the voice and video traffic in term of packet loss probability and average delay. But, the average packet delay of data traffic increases as the number of data users increase as shown in fig. 17 due to the reduction of the available resources (bandwidth) for each data user. Fig. 18 shows the increase in the achieved throughput by adding data traffic.

The actual throughput defined as the ratio of the average number of slots used for successful data packet transmission to the total number of slots per frame. While the total throughput defined as the ratio of the average number of slots used for control and data transmission to the total number of slots per frame. Only one control slot (4 minislots) is use for control purpose which means 4.1 % if the total available bandwidth.

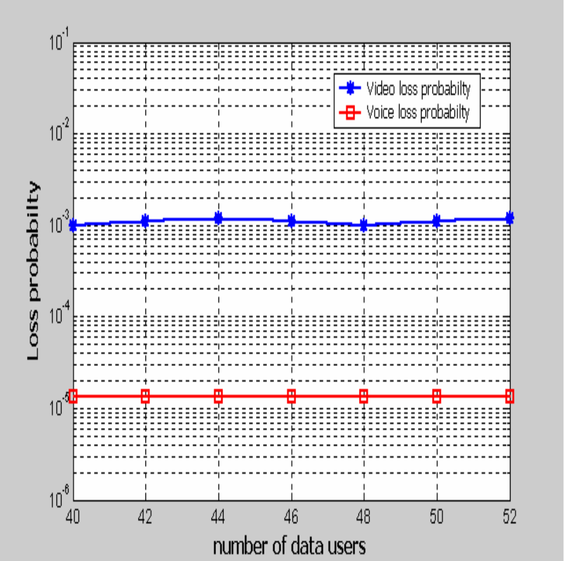

Figure 15: Packet loss probability of voice and video traffic as a function of the number of data users at (12 video user, and 45 voice user).

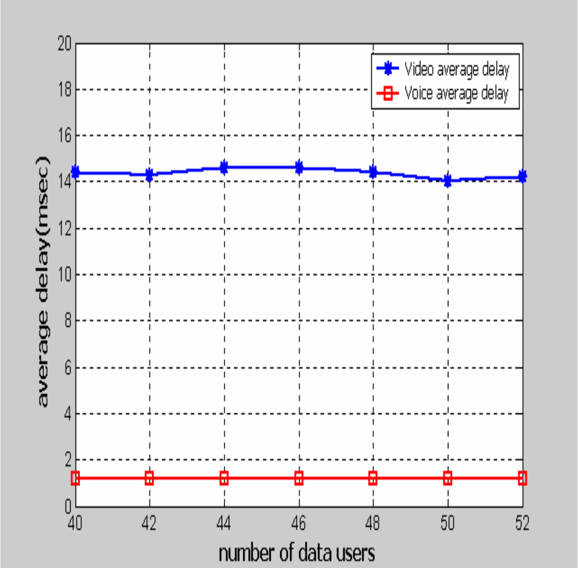

Figure 16: Average packet delay of voice and video traffic as a function of the number of data users at (12 video user, and 45 voice user).

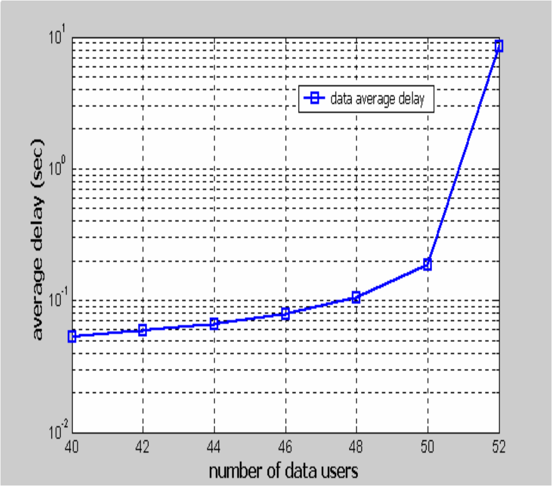

Figure 17: Average packet delay of data traffic as a function of the number of data users at (12 video user, and 45 voice user).

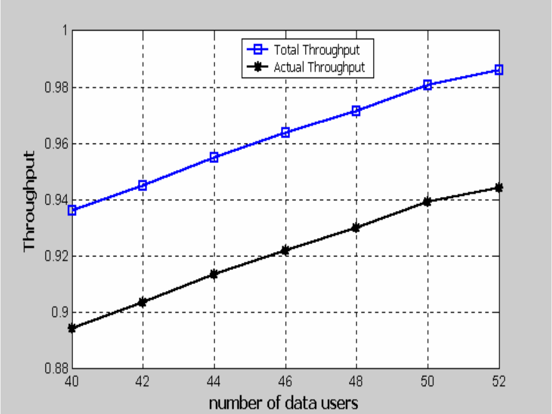

Figure 18: Total and actual throughput as a function of the number of data users at (12 video user, and 45 voice user).

5. CONCLUSION

In this paper, a novel fair delay optimization based resource allocation for video traffic over wireless multimedia system is presented. The proposed allocation algorithm is trying to provide fair delay for video packets by minimizing the delay difference among transmitted video packets and adaptively control the allocated resources (bandwidth) for video traffic around the corresponding average bit rate while respecting the QoS requirements of video traffic. A minimized control overhead of only two bits is needed which increases the utilization efficiency. Simulation results show that the proposed resource allocation algorithm achieves very high utilization and provides nearly fair delay among video packets. In addition, it has the ability of controlling the QoS offered for video traffic in terms of packet loss probability and average delay to provide a better QoS as required. Under traffic integration condition the data traffic increases the utilization to 98 % by using the remaining bandwidth after voice and video traffic while a good QoS is offered to voice and video traffic, and a delay of 200 msec offered to data traffic.

6. REFERENCES

[1] O. Kubbar and H. T. Mouftah, "Multiple Access Control Protocols for Wireless ATM: Problems Definition and Design Objectives, " *IEEE Communication Magazine*, pp. 93-99, November 1997.

[2] Y. Kwok and V. K. N. Lau, "A quantitative Comparison of Multiple Access Control Protocols for Wireless ATM," *IEEE Transaction on Vehicular Technology*, vol. 50, no. 3, pp. 796-815, May 2001.

[3] I. Akyildz, J. McNair, L. Carrasco and R. Puigjaner, "Medium Access Controls for Multimedia Traffic in Wireless Networks," *IEEE Network Magazine*, vol. 13, no. 4, pp. 39-47, July 1999.

[4] J. Wen, J. Lain, and Y. lai, "Performance Simulation of Wireless Multimedia Systems Using NC-PRMA/DA and SNC-PRMA/DA Protocols," *IEEE Transactions on Systems, MAN, AND Cybernetics Part A: Systems and Humans*, vol. 32, no. 6, pp. 780-787, November 2002.

[5] S. Elnoubi and A. M. Alsayh, "A Packet Reservation Multiple Access (PRMA)-Based Algorithm for Multimedia Wireless System " *IEEE Transaction on Vehicular Technology*, vol. 53, no. 1, pp. 215-222, January 2004.

[6] J. Kuri and M. Gagnaire, "ATM Traffic Management in an LMDS Wireless Access Network" *IEEE Communication Magazine*, pp. 128-133, September 2001.

[7] N. Passas, S. Paskalis, D. Vali and L. Merakos, "Quality-of-Service-Oriented Medium Access Control for Wireless ATM Networks," *IEEE Communication Magazine*, pp. 42-50, 1997.

[8] J. F. Frigon, H. C. B. Chan and V. C. M. leung, "Dynamic Reservation TDMA Protocol for Wireless ATM Networks," *IEEE J.Selected Areas in Communication*, vol.19, no.2, pp 370-383, Feb. 2001.

[9] C. G. Kang, C. W. Ahn, K. H. Jang, and W. S. Kang, "Contention-Free Distributed Dynamic Reservation MAC Protocol With Deterministic Scheduling (C-FD^3 R MAC) for Wireless ATM Networks, " *IEEE J. Selected Area Communications*, vol. 18, no. 9, pp. 1623-1635, September 2000.

[10] L. Musumeci, P. Giacomazzi, and L. Fratta, " Polling- and Contention -Based Schemes for TDMA-TDD Access to Wireless ATM Networks, " *IEEE J. Selected Area Communications*, vol. 18, no. 9, pp. 1597-1607, September 2000.

[11] C.S. Chang, K.C Chen, M.Y. You, and J.-F Chang, "Guaranteed Quality-of-Service Wireless Access to ATM Networks, " *IEEE J. Selected Area Communications*, vol. 15, no. 1, pp. 106-118, January 1997.

[12] L. Fratta, P. Giacomazzi, and L. Musumeci, "PRAS: A Mac Protocol for Wireless ATM Networks" *Globecom'99*, pp. 2743-2715, 1999.

[13] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson, and J. D. Robbins, "Performance Models of Statistical Multiplexing in Packet Video Communication" *IEEE Trans. Commun.*, vol. 36, pp. 834-844, July 1988.