Scheduling for H.264 Video Traffic over GEO Satellite Networks

Polychronis Koutsakis[1], Michael Giannotta[1], Andrei Ianos[1] and Aggelos Lazaris[2]
[1]Dept. of Electrical and Computer Engineering, McMaster University, Canada
[2]Dept. of Electronic and Computer Engineering, Technical University of Crete, Greece
Tel: +1-905-525-9140 ext. 26542, +30-28210-37235
polk@ece.mcmaster.ca, giannomb@mcmaster.ca, ianosab@mcmaster.ca, alazaris@telecom.tuc.gr

ABSTRACT
The provision of acceptable Quality-of-Service (QoS) for integrated multimedia traffic over a geosynchronous earth orbit (GEO) satellite network demands the existence of a well-designed Medium Access Control (MAC) protocol. This paper proposes a new dynamic satellite bandwidth allocation technique (Predictive Resource Reservation Access, PRRA) which is based on accurate H.264 video traffic prediction. Our scheme is shown to provide very good throughput and delay results, when compared with other efficient schemes from the literature and with an "ideal" satellite bandwidth allocation scheme.

Categories and Subject Descriptors
C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design – *Wireless communication*.

General Terms
Algorithms, Performance

Keywords
Scheduling, H.264 video, MAC Protocol, GEO satellites, traffic modeling.

1. INTRODUCTION
In recent years, broadband satellite networks have attracted significant attention as a part of the global communications infrastructure. Satellite networks have specific characteristics such as including global coverage, providing access for remote users to terrestrial wide area networks from any location within the satellite coverage area, and broadcast/multicasting, which make them an attractive solution in order to complement terrestrial networks in providing worldwide access to the present and future generation multimedia services.

However, at the same time other characteristics of satellite networks, such as long propagation delays (270 ms round-trip delay for a GEO satellite network, i.e., at least 540 ms from the time the signaling information is sent from the station to the satellite until it reaches the receiver terminal), the limited bandwidth to be shared among many users, and limitation in power (which implies stringent use of buffer memory, transponder capacity and processing power) create the need for a well-designed Medium Access Control (MAC) protocol, especially in today's networks which are expected to handle bursty multimedia traffic. Among the above mentioned problems, the propagation delay is the most significant one, as it makes bursty users' (such as video users') current traffic profile rather useless for bandwidth allocation, since the profile will probably have changed significantly by the time the bandwidth allocation is made by the Network Control Center (NCC), which we consider to be integrated in the satellite on-board device.

Medium Access Control protocols enable communicating stations at diverse locations to regulate the movement of their packets and manage network bandwidth in order to utilize the network resources as efficiently as possible. Video traffic is expected to be a substantial portion of the traffic carried by emerging networks, hence presenting a challenge for satellite system designers. For Variable Bit Rate (VBR) coded video, statistical models are needed to design networks which are able to guarantee the strict Quality of Service (QoS) requirements of video traffic. Video packet delay requirements are strict, because delays are annoying to a viewer; whenever the delay experienced by a video packet exceeds the corresponding maximum delay, the packet is dropped, and the video packet dropping requirements are equally strict. Therefore, a good statistical model can be very useful in evaluating network performance under various videoconferencing loads.

H.264 is the latest international video coding standard. In this work we build a new model which accurately captures the behavior of multiplexed H.264 video movies. The model is used for predicting the behavior of video traffic in our new MAC protocol proposal for a Digital Video Broadcasting Return Channel Satellite (DVB-RCS) system. Our proposed scheme (Predictive Resource Reservation Access, PRRA) is shown, via simulations, to outperform two other efficient MAC schemes and to achieve video packet delays and video packet dropping probability only slightly worse than an ideal scenario in which the actual video user requirements for each video frame would be a priori known to the NCC.

2. H.264 VIDEO TRAFFIC MODELING

H.264 is the latest international video coding standard. It was jointly developed by the Video Coding Experts Group (VCEG) of the ITU-T and the Moving Picture Experts Group (MPEG) of ISO/IEC. It uses state-of-the-art coding tools and provides enhanced coding efficiency for a wide range of applications, including video telephony, video conferencing, TV, storage (DVD and/or hard disk based, especially high-definition DVD), streaming video, digital video authoring, digital cinema, and many others [1].

Similarly to our recent work on modeling H.263 videoconference traffic [4], our present work focuses on the accurate fitting of the marginal (stationary) distribution of video frame sizes of single video traces. More specifically, our work follows the steps of the work presented in [2], where Heyman et al. analyzed three videoconference sequences coded with a modified version of the H.261 video coding standard and two other coding schemes, similar to the H.261. The authors in [2] found that the marginal distributions for all the sequences could be described by a gamma (or equivalently negative binomial) distribution.

In our work, we have studied three different long sequences of H.264 VBR encoded videos, from the publicly available High Definition Video Trace Library and the Main Profile Video Trace Library of [3]. We have investigated the possibility of modeling the traces with a number of well-known distributions and our results have shown that the best fit among these distributions for modeling a single movie is achieved for all traces examined with the use of the Weibull distribution.

The three traces are, respectively:

1. A demo from the Sony Digital HD Video Camera.

2. A documentary ("KAET's From Mars to China").

3. A talk-show program ("KAET's Horizon").

In the work presented in this paper, we will show that, based on the good fit of the Weibull distribution for modeling a single movie, the behavior of *both single and multiplexed H.264 video traces* from VBR coders can be accurately captured and used for efficient proactive resource management in satellite systems *(the case of modeling multiplexed video streams is especially significant for our study, since numerous sources are multiplexed in the uplink channel)*. To do this, we will first briefly explain our results on modeling a single movie.

For each one of the three videos under study we have used the I-P-B frames quantization version of 28-28-30. New video frames arrive every 33.3 msecs. The length of the videos varies from 10 to 30 minutes and the data for each trace consists of a sequence of the number of cells per video frame. Table 1 presents the trace statistics for each trace. The Probability Density Function (PDF) of a Weibull distribution with parameters (α, β) is:

$f(x)= [\alpha\beta^{-\alpha}x^{(\alpha-1)}\, e^{-(x/\beta)\alpha}]/ [\beta^{-\alpha}\,\Gamma(\alpha)]$, for all x>0, and zero otherwise.

The mean and variance are given by the equations:
$\text{Mean}=(\beta/\alpha)\,\Gamma(1/\alpha)$
$\text{Variance}=(\beta^2/\alpha)\{2\Gamma(2/\alpha)-(1/\alpha)[\Gamma(1/\alpha)]^2]\}$.
These formulas are used in order to compute the (α, β) parameters of the Weibull distribution which models each trace, based on the mean and the variance of the actual traces.
Autoregressive models have been used in the past to model the output bit rate of VBR encoders, e.g. [5, 6]. A Discrete Autoregressive model of order p, denoted as DAR(p) [7, 8], generates a stationary sequence of discrete random variables with an arbitrary probability distribution and with an autocorrelation structure similar to that of an Autoregressive model.

Table 1. Trace Statistics.

Movie	Mean Bit Rate (Mbps)	Peak Bit Rate (Mbps)	Standard Deviation (Mbps)
Sony Demo	0.422	6.67	6.65
Horizon	1.53	24	22.85
From Mars to China	4.85	78.45	64.36

DAR(1) is a special case of a DAR(p) process and it is defined as follows: let $\{V_n\}$ and $\{Y_n\}$ be two sequences of independent random variables. The random variable V_n can take two values, 0 and 1, with probabilities 1-ρ and ρ, respectively. The random variable Y_n has a discrete state space S and $P\{Y_n = i\} = \pi(i)$. The sequence of random variables $\{X_n\}$ which is formed according to the linear model:
$X_n = V_n X_{n-1} + (1- V_n)\, Y_n$
is a DAR(1) process.
A DAR(1) process is a Markov chain with discrete state space S and a transition matrix:
$\mathbf{P} = \rho\mathbf{I} + (1-\rho)\,\mathbf{Q}$
where ρ is the autocorrelation coefficient, **I** is the identity matrix and **Q** is a matrix with $Q_{ij} = \pi(j)$ for i, j ϵ S.
Autocorrelations are usually plotted for a range W of lags. The autocorrelation can be calculated by the formula:
$\rho(W)= E[(X_i - \mu)(X_{i+w} - \mu)]/\sigma^2$
where μ is the mean and σ^2 the variance of the frame size for a specific video trace.
As in [2], where a DAR(1) model with negative binomial distribution was used to model the number of cells per frame of VBR teleconferencing video, we want to build a model based only on parameters which are either known at call set-up time or can be measured without introducing much complexity in the network. DAR(1) provides an easy and practical method to compute the transition matrix and gives us a model based only on four physically meaningful parameters, i.e., the mean, peak, variance and the lag-1 autocorrelation coefficient ρ of the offered traffic. The DAR(1) model can be used with any marginal distribution [7].
More specifically, in our model the rows of the **Q** matrix consist of the Weibull probabilities ($f_0, f_1, \ldots f_k, F_K$), where $F_K = \Sigma_{k>K} f_k$, and K is the peak rate. Each k, for $k<K$, corresponds to possible source rates less than the peak rate of K.
We proceeded with testing our model statistically in order to study whether it produces a good fit for the trace superposition. For this reason we have used Q-Q plots. The Q-Q plot is a powerful goodness-of-fit test [2,9], which graphically compares two data sets in order to determine whether the data sets come from populations with a common distribution (if they do, the points of the plot should fall approximately along a 45-degree reference line). More specifically, a Q-Q plot is a plot of the quantiles of the data versus the quantiles of the fitted distribution (a z-quantile of X is any value x such that $Pr(X \leq x) = z$).

In Figure 1, we have plotted the 0.01-, 0.02-, 0.03-,… quantiles of the actual Horizon trace versus the respective quantiles of a superposition of 10 pseudo-traces generated by the DAR(1) model. As shown in the Figure, the points of the Q-Q plot fall almost completely along the 45-degree reference line, for about 90% of the quantiles, and are very close to reference line for the rest of the quantiles as well. The very good fit shows that the superposition of the actual traces can be modeled very well by a respective superposition of data produced by the DAR(1) model.
The same conclusions were deducted by our results for the superposition of various number of sources transmitting the other two traces; these results show that our model captures with great accuracy the behavior of H.264 video traffic.
In the following Sections we will show that our accurate H.264 video modeling can be used very efficiently for proactive resource management in satellite systems.

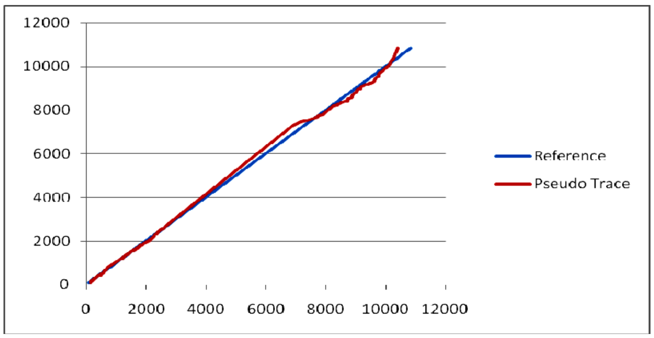

Figure 1. Q-Q plot of DAR(1) model versus the actual Horizon trace for 10 superposed sources.

3. PREDICTIVE RESOURCE RESERVATION ACCESS (PRRA)

3.1 The Proposed Scheduling Approach

The Digital Video Broadcasting Return Channel Satellite (DVB-RCS) standard [14] develops a communication system for the return channel (uplink channel), i.e., the link from the user terminal to the network gateway. Due to the expected services features and large delay-bandwidth product of satellite networks, DVB-RCS represents a proper test bed for the proposed resource allocation schemes. The most relevant elements of the DVB-RCS Network are: a) RCSTs (Return Channel Satellite Terminals), i.e., a generic access terminal, b) the NCC (Networks Control Center), a device in charge of managing the access and bandwidth allocation for RCSTs, c) Gateways and Feeders, which are the elements that receive and transmit information outside the network [10].

As in [10], our proposed satellite medium access scheme is based on a Multi-Frequency Time Division Multiple Access (MF-TDMA) approach, according to which a carrier is divided in timeslots (grouped in frames and superframes). MF-TDMA schemes are capable of providing efficient and flexible bandwidth utilization [10, 15]. The system parameters are taken from [10] and are presented in Table 2. We use a packet size of 48 information bytes (ATM cell size) throughout this work, but our mechanism can be used equally well with packets of other sizes, as the nature of our modeling results would not be altered at all.

The NCC allocates to each active RCST a set of timeslots, each characterized by a frequency, bandwidth, start time and duration time. The DVB-RCS standard provides five allocation request types [14, 16] (i.e., these types of requests do not concern user access to the channel, but only ways of providing slots to users) which can be joined in order to satisfy the QoS requirements:

Continuous Rate Assignment (CRA) is a fixed capacity negotiated between the RCST and the NCC. It is maintained across frames until a new negotiation.

Rate Based Dynamic Capacity (RBDC) is a capacity allocated to the RCST based on its rate request (bytes/frame), and is subject to a maximum rate limit negotiated between the RCST and the NCC. The last request from a RCST overwrites all previous RBDC requests from the same RCST.

Volume Based Dynamic Capacity (VBDC) is an assignment strategy in which a terminal signals its request in terms of total number of slots required to empty its queue. The request remains effective as long as not all of the requested time slots have been granted. This strategy is especially suited for bursty traffic [15].

Absolute Volume Based Dynamic Capacity (AVBDC) request is a request similar to VBDC, with the difference that the last request from a RCST overwrites all previous AVBDC requests from the same RCST.

Free Capacity Assignment is not a true request, allocating the otherwise unused capacity. It is automatic and does not involve signaling from the RCST to the NCC.

Table 2. System Parameters

Frame Duration	26.5 ms
Carriers	4
Slots/frame/carrier	128
Bytes/slot	53
System global rate	8 Mbps

The modeling approach proposed in [10] for self-similar data traffic (such as World Wide Web traffic) lets an RBDC request correspond to the data traffic prediction, plus a corrective factor ζN. The first difference between the work presented in [10] and PRRA is that, additionally to the MF-TDMA frame structure, we adopt the idea that after all requests have been satisfied the bandwidth left is distributed freely following a certain algorithm (this approach is named in the literature as a Combined Free and Demand Assignment Scheme, CFDAMA scheme); the algorithm implemented in our scheme is a simple round-robin assignment algorithm to all RCSTs which are currently active. This idea of a hybrid protocol (according to the satellite MAC protocols' classification in [13]) was first proposed in [17] and is especially useful in allocating slots to video users, as the difficulty in providing them with adequate bandwidth due to their frequent changes in bandwidth needs could be somewhat alleviated by their acquiring the unused channel bandwidth freely in a round-robin manner. We use the Combined Free and Demand Assignment Scheme with Piggybacking (CFDAMA-PB) version of the protocol, which was shown in [18] to be the most efficient way of making reservations. According to the PB strategy, user stations send their capacity requests embedded in the header of their packets. The free capacity distribution performed by the protocol brings the end-to-end delay performance at low loads close to that obtained with random access protocols, while the demand-based bandwidth allocation at the beginning of each frame guarantees the protocol's stability, robustness and efficient utilization of transmission bandwidth at high loads. Free capacity distribution in an MF-TDMA frame structure was also used in [15], which is one of the three schemes with which we will conceptually compare our work in Section 3.2.

The second and most important difference of PRRA with [10] is related to the fact that an approach similar to the one used in [10] for data traffic (letting the RBDC request correspond to the data traffic prediction) is not enough for real time Variable Bit Rate (rt-VBR) video traffic, which has strict QoS requirements in terms of average video packet dropping (*set to a maximum of 0.1% in our work) and average video packet delay (set to a maximum of 0.6 seconds in our work, which is especially strict considering that, for each possible failure of our prediction due to underassignment, the respective packets which would have to wait for a new assignment will have a minimum round-trip video packet delay of 0.54 seconds*). For this reason, we propose the following different approach in our MAC scheme.

As explained in Section 2, the Weibull fit provides a good (but not perfect) fit for a single H.264 video trace. Also, a logical assumption for next generation networks is that video users will be allowed to adopt one of a set of specific "modes" (each corresponding to a set of traffic parameters), for call admission control and traffic policing purposes. Therefore, in PRRA we consider that a video user can adopt one of the "modes" presented in Table 1 (of course, a larger pool of modes would have to be used in a real satellite system scenario). Based on the good model for single video traces and the highly accurate model of multiplexed traffic, *we propose that the "burden" of traffic prediction for the RCSTs should fall on the NCC instead of the RCSTs*. More specifically, the NCC should run a real-time simulation, both for single and for multiplexed video sources. Hence, based on the "mode" declared by the RCSTs at call establishment, the NCC does not need to wait for a request from the RCSTs every channel frame (which would arrive with a delay of more than 10 channel frames, due to the propagation delay); instead, it can start allocating resources to the video terminals, by simulating the single source models with the sources' mean rate as a simulation start point, and by computing the free slots in each channel frame (using the DAR(1) models for multiplexed video traffic, and subtracting the estimated used slots from the total number of slots in the system) in order to allocate the estimated number of free slots in a round-robin manner to all active RCSTs. With this slot allocation scheme, the RCST will not need to send frequent requests to the NCC but it will only need to send a "corrective" AVBDC request every superframe (defined in our work as equal to 11 channel frames, to account for the propagation delay). The reason for sending this request will be for the RCST to help the NCC correct any mistakes (due to either slots overassignment or underassignment) of the models produced at the NCC via online simulation. After receiving the AVBDC request, the NCC will resume its simulation with the current RCST state (in terms of bandwidth requirements) as a start point. As it will also be shown from our results, this approach helps waive the three main disadvantages of hybrid protocols:

a. No large processing times are needed (our approach minimizes the need for requests from RCSTs and signaling among the video RCSTs and the NCC),

b. There is no difficulty in specifying the ideal amount of fixed bandwidth to offer to terminals (video users' needs are accurately predicted), and

c. choosing the correct weights according to the stations' priority in the round-robin allocation (again, the round-robin allocation of free slots is done based on the estimation of current user needs in terms of bandwidth).

For these reasons, our scheme will be shown to provide highly satisfactory results in terms of video users' QoS requirements satisfaction.

3.2 Conceptual Comparison with Relevant Work

The subject of allocating bandwidth to video users in a satellite system has not been widely studied in the relevant literature. Most of the research on satellite MAC protocols has focused on handling data traffic and on integrating voice and data traffic over satellite uplink channels of limited capacity, and, to the best of our knowledge, no previous work has studied the problem of transmitting H.264 video traffic over GEO satellite links. However, there are a few schemes in the literature which have addressed the issue of providing acceptable QoS to video users. We will compare PRRA conceptually with three such schemes, the ones proposed in [11, 12, 15].

In [11], the authors propose a scheme which uses n levels of allocation to MPEG-2 video users. Each of the levels represents a threshold for the user's throughput; if the data currently transmitted by the user is below a threshold, the user gives up a portion of its allocated bandwidth in favor of other users which are currently in need of more bandwidth resources; if the user's throughput exceeds a threshold, the user regains the bandwidth which was initially allocated to it. The disadvantages of this scheme concern: a) the signaling delay for giving up and regaining bandwidth, b) the large number of levels needed for ideal control of slot allocation and deallocation (the authors produce results with various numbers of levels, reaching up to 100 levels), which add a very significant computational load to the system c) the authors use for their results only one, synthetically generated MPEG-2 compressed movie (whereas we use actual video traces in our work); the implementation of their mechanism for a large number of movies (i.e., for users transmitting various movies at various qualities) would need a definition of a different large number of levels for each one of the movies, as the good choice of thresholds is of vital importance to the proposed mechanism, hence yielding this approach impractical.

In [12], the authors proposed a satellite MAC protocol for transmitting MPEG-1 video over satellite channels. Their work is based on the assumption that the video stream at the Group of Pictures (GOP) level could be thought of as a sequence of scenes, and a scene change occurs when a significant change occurs in the GOP size; therefore, a new bandwidth allocation should take place only at scene changes, hence decreasing the frequency of signaling between user terminals and the NCC. For this reason, the authors propose the extraction of statistics of the slow time process and the fast time process of MPEG-1 movies in order to derive a metric for allocating bandwidth to video users. However, the packet flow metric used by the authors depends on certain parameters (e.g., the time window in which the metric is measured), the ideal value of which will actually depend on the burstiness of the movie, therefore these parameters need to be different for different movies, as pointed out in [12]. This complexity limits the practicality of the authors' scheme, at least in its ideal form (it could be implemented with a compromise on a fixed time window length for all movies transmitted and at all qualities for each movie, but this would lead to inferior system performance). Finally, the authors in [12] conclude that if video packet loss as high as 3% is tolerable, their scheme achieves considerably lower delays than other known satellite MAC protocols. However, the maximum video packet dropping of 3% is quite high; our scheme will be shown to provide very good throughput results for a significantly stricter maximum allowed video packet dropping of 0.1%.

In [15], the authors propose a satellite MAC protocol based on a MF-TDMA access scheme (i.e., they use a similar approach to our work). After receiving the terminals' requests, the scheduler determines the bandwidth which should be allocated to each terminal. The protocol, however, proposes the use of a CRA-type assignment for rt-VBR traffic such as the videoconferencing traffic used in our work, with the difference that the assignment is fixed for the duration of the connection (no new negotiation is needed), and equal to the rt-VBR user's peak transmission rate; the reason for this choice in [15] is that rt-VBR traffic has strict delay and packet dropping constraints, and no accurate traffic prediction mechanism is provided in [15], hence leading to the "defensive" choice of peak cell rate assignment, which leads to significant bandwidth waste, as the assignment is most of the time larger than the rt-VBR user's actual needs (the other problem with this type of assignment is that if the assignment was made for less than the peak transmission rate, it would at times lead to severe packet dropping; the use, in [15], of a free slot assignment policy, could help to partially alleviate this problem, but the authors chose the "safer" solution of peak transmission rate assignment). The scheme in [15] achieves good performance results, but takes into consideration a case where all rt-VBR users steadily offer 25% of their peak rate, and therefore Available Bit Rate (ABR) traffic can take advantage of the rest of the bandwidth assigned to a terminal, by statistical multiplexing done within terminals. If, however, the terminal does not need the remaining bandwidth for other types of traffic than the rt-VBR one, this bandwidth is lost; this choice needs to inferior performance, especially since all three traces used in our study have a peak-to-mean ration larger than 5.4, which means that the constant allocation of the peak rate to all video sources leads to significant loss of valuable bandwidth resources. This will also be shown through the simulation comparison of PRRA and [15] in Section 4. Finally, it should be noted that video traffic in [15] is generated with the use of a Markov Modulated Poisson Process (MMPP) model, whereas we use actual video traces in our work.

4. SIMULATION RESULTS AND DISCUSSION

We use computer simulations executed on Pentium-IV workstations to study the performance of PRRA. Each simulation point is the result of an average of 10 independent runs (Monte-Carlo simulation).

Since the system global rate is 8 Mbps, it is clear by the trace parameters presented in Table 1 that the traffic generated by the High Definition trace of the documentary "From Mars to China" cannot be accommodated by the network. For this reason, we only use in our simulation study the other two traces from Table 1 (the fact that the peak bit rate for the Horizon trace is 24 Mbps is not forbidding its use, given that this peak is very seldom reached and the trace's mean bit rate is much smaller than the global rate). It should be noted that, given the very good modeling results from Section 2, PRRA would work equally well with three or more modes in a network with higher capacity.

At the start of our simulation study, we let each video user choose one of the two traffic parameter sets ("modes") which are presented in Table 1, with equal probability.

Four MAC schemes will be compared in this Section.

The first is our proposed scheme, PRRA.

The second is the scheme proposed in [15] (the MMPP traffic generation is substituted here by the real video traces), allocating to each video terminal its peak rate (declared at call establishment).

The third is the Scheduled-Retransmission Multiaccess (SRMA) Protocol [19], which improved on the well-known Announced Retransmission Random Access (ARRA) [20] satellite MAC protocol. More specifically, in ARRA the frame is divided into a number of K slots and each slot is divided into a data slot and K minislots; minislots have a much smaller length than the data slot. When a terminal attempts to transmit in a data slot, it also sends a reservation request by choosing one of the minislots of that slot at random; if the data transmission is successful, the reservation that has been made through the use of the minislot will be cancelled; otherwise, the reservation made is valid and the terminal will transmit again its failed data packet in the data slot corresponding to the used minislot. Additionally, each frame begins with an additional group of K minislots which serve as the common minislot pool. SRMA achieved improved throughput performance results in comparison to ARRA by not using the common minislot pool (and thus saving this bandwidth for transmissions) and, instead, assigning a status vector to each packet to eliminate the reservation collisions from different slots in the frame. Two versions of SRMA were designed and evaluated in [19], one with a fixed frame duration (SRMA-FF) and one with a dynamic frame duration (SRMA-DF), containing a fixed-length Aloha frame and a variable-length reserved frame with length equal to the number of successfully reserved packets in the corresponding Aloha frame. Since SRMA-DF was shown to provide higher throughput and smaller packet delay results than SRMA-FF, we choose SRMA-DF for comparison with our scheme.

The fourth is an "ideal" scheme, as we want to compare PRRA with a similar protocol in which the NCC would "magically" know, without any information exchange (therefore, no contention is necessary among video RCSTs), exactly what the video RCSTs' bandwidth demands for the next video frame would be.

Figure 2 presents our simulation results for the average video packet dropping metric versus the system utilization. Utilization indicates the traffic load normalized to the uplink capacity, e.g., a traffic load equal to 40% represents 40% of the 8 Mbps uplink capacity, i.e., 3.2 Mbps system throughput. As it is shown in the Figure, the difference in video packet dropping between our scheme and the "ideal" case is so small that it can be considered almost negligible for all normalized video traffic loads. Our scheme can handle up to 63% system load while at the same time satisfying the strict QoS requirement of maximum video packet dropping equal to 0.1%; the respective maximum system load which the "ideal" scheme can handle is 72%, while SRMA-DF achieves only a 29% maximum throughput and [15] a 36% maximum throughput for the same QoS requirement.

The reason that neither PRRA nor the "ideal" scheme can achieve a higher throughput is the high burstiness of video traffic; in certain channel frames, video bursts from more than one RCST happen to take place simultaneously in the uplink channel.

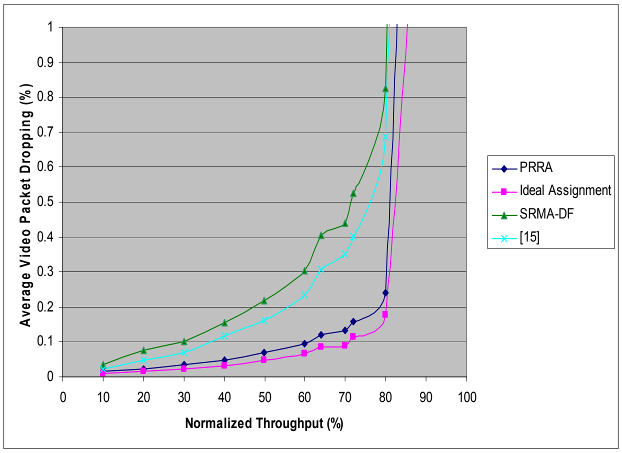

Figure 2. Average video packet dropping vs. System Utilization.

Although our traffic modeling scheme makes good predictions of such bursts, the total amount of requested bandwidth in certain channel frames may surpass the system's available capacity; this will lead to inevitable video packet dropping, as some of the packets may not be sent within the (33.3/26.5=1.26) channel frames which pass before the arrival of the next video frame (when a new video frame arrives, all packets of the previous video frame which have not yet been sent are discarded). The bursty nature of video traffic is also responsible for the much lower throughput results achieved by the other two schemes. The most significant reasons for which PRRA excels in comparison to SRMA-DF, which is also a hybrid protocol, are explained in Section 3.1. The reasons for which PRRA excels in comparison to [15] were explained in Sections 3.1 and 3.2. Additionally, it should be pointed out that in comparison to [15], SRMA-DF has the advantage that requests from video RCSTs are made based on the terminals' specific needs in terms of bandwidth at that moment (i.e., they do not receive a constant peak rate allocation, which causes significant waste of bandwidth); on the other hand, in comparison to [15] SRMA-DF has the disadvantages that

a. no bandwidth is allocated to RCSTs until their requests are correctly received (i.e., without collision) by the scheduler,

b. it is not optimized for accommodating video traffic, therefore requests are constantly generated by RCSTs for every new video frame (every 1.51 channel frames), hence increasing the probability of collisions.

For these reasons, [15] achieves higher maximum throughput, and as shown in Figure 3, lower mean video packet delay than SRMA-DF.

Figure 3 presents our simulation results for the average video packet delay versus the system utilization. The results show again that our scheme's efficiency is very close to the one achieved by the "ideal" scheme and much better than those of SRMA-DF and [15]. However, it should be pointed out that:

a. as the system load increases, the "ideal assignment" scheme achieves a lower delay of about 0.2-0.3 seconds in comparison to our scheme, due to the lack of contention (and therefore, lack of collisions) in the "ideal" scenario,

b. by studying Figures 1 and 2, it is clear that for a 64% system load only the video packet dropping metric surpasses its maximum set value in our scheme (the average video packet delay remains lower than 0.6 seconds for the specific load); similarly, for a 72% system load only the video packet dropping metric surpasses its bound in the "ideal" scheme (the average video packet delay remains again lower than the upper bound of 0.6 seconds for the specific load). Therefore, since the strictest metric is shown to be that of the maximum mean video packet dropping, our future work will focus on the subject of introducing an efficient Forward Error Control (FEC) mechanism, which will allow for larger maximum video packet dropping without adding a large overhead to the information transmitted in the uplink.

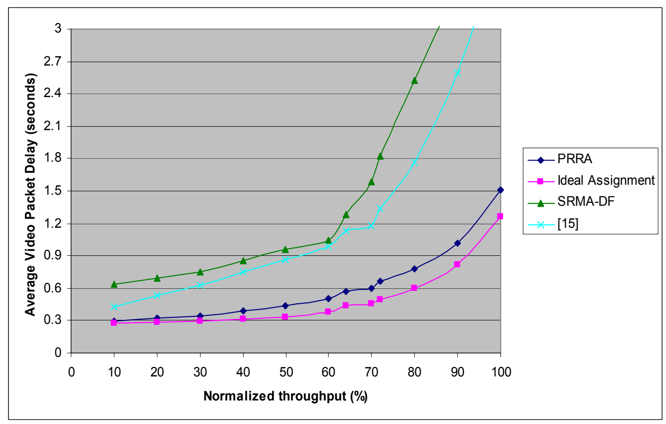

Figure 3. Average video packet delay vs. System Utilization.

Finally, it should also be noted that as shown in Figure 3, given the very strict set QoS requirement of 0.6 seconds maximum mean video packet delay, SRMA-DF would only be able to achieve 7% throughput (instead of the 29% which is achieved if only the mean video packet dropping bound is taken into consideration) and [15] would be able to achieve 28% throughput (instead of the 36% which is achieved if only the mean video packet dropping bound is taken into consideration). This means that for both these protocols the strictest metric is the maximum allowed video packet delay.

5. CONCLUSIONS

In this paper, we have proposed the use of a model for VBR H.264 video traffic for predicting video user needs in a new dynamic bandwidth allocation scheme in a satellite network system based on the DVB-RCS standard. Our Discrete Autoregressive (DAR) model was shown to be highly accurate, hence providing significant advantages for the on-board satellite scheduler in terms of being able to achieve high bandwidth utilization and to satisfy the strict video users' QoS requirements. We also explain why our proposed scheduling scheme, PRRA, excels both conceptually and in simulation results in comparison with other schemes for satellite bandwidth allocation.

6. REFERENCES

[1] T. Wiegand, G. Sullivan, and A. Luthra, "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496 -10 AVC)," May 2003.

[2] D. P. Heyman, A. Tabatabai and T. V. Lakshman, "Statistical Analysis and Simulation Study of Video Teleconference Traffic in ATM Networks", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 2, No. 1, pp. 49-59, 1992.

[3] [Online] http://trace.eas.asu.edu/hd/index.html

[4] P. Koutsakis, "A New Model for Multiplexed VBR H.263 Videoconference Traffic", *in Proceedings of the IEEE GLOBECOM 2006,* San Francisco, USA.

[5] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson, and J. D. Robbins, "Performance Models of Statistical Multiplexing in Packet Video Communications", *IEEE Transactions on Communications*, Vol. 36, No.7, pp. 834-844, 1988.

[6] C. Shim, I. Ryoo, J. Lee and S. Lee, "Modeling and Call Admission Control Algorithm of Variable Bit Rate Video in ATM networks", *IEEE Journal on Selected Areas in Communications*, vol.12, No.2, pp. 332-344, 1994.

[7] A. Adas, "Traffic Models in Broadband Networks", *IEEE Communications Magazine,* Vol. 35, No.7, pp. 82-89, 1997.

[8] P. A. Jacobs and P. A. W. Lewis, "Time Series Generated by Mixtures", *Journal of Time Series Analysis*, Vol. 4, No. 1, pp. 19-36, 1983.

[9] A. M. Law and W. D. Kelton, "Simulation Modeling & Analysis", 2nd Ed., McGraw Hill Inc., 1991.

[10] F. Chiti, R. Fantacci and F. Marangoni, "Advanced Dynamic Resource Allocation Schemes for Satellite Systems", *in Proceedings of the IEEE International Conference on Communications (ICC) 2005,* Vol. 3, pp. 1469-1472*,* Seoul, Korea.

[11] N. Celandroni, E. Ferro and F. Potorti, "A Multi-Level Satellite Channel Allocation Algorithm for Real-Time VBR Data", *International Journal of Satellite Communications*, Vol. 20, No. 1, pp. 47-61, 2002.

[12] D. Connors, B. Ryu, G. J. Pottie and S. Dao, "A Medium Access Control Protocol for Real Time Video over High Latency Satellite Channels", *Mobile Networks and Applications,* Vol. 7, No. 1, pp. 9-20, 2002.

[13] H. Peyravi, "Medium Access Control Protocols Performance in Satellite Communications", *IEEE Communications Magazine,* Vol. 37, No. 3, pp. 62-71, 1999.

[14] Std., ETSI EN 301 790 V1.2.2 (2000-12).

[15] A. Iuoras, P. Takats, C. Black, R. DiGirolamo, E. A. Wibowo, J. Lambadaris and M. Devetsikiotis, "Quality of Service-Oriented Protocols for Resource Management in Packet-Switched Satellites", *International Journal of Satellite Communications,* Vol. 17, No. 2-3, pp. 129-141, 1999.

[16] Std., ETSI EN 101 790 V1.1.1 (2001-09).

[17] S. V. Krishnamurthy and T. Le-Ngoc, "Performance of CF-DAMA Protocol with Pre-Assigned Request Slots in Integrated Voice/Data Satellite Communications", *in Proceedings of the IEEE International Conference on Communications (ICC) 1995,* Vol. 3, pp. 1572-1576*,* Seattle, USA.

[18] T. Le-Ngoc and S. V. Krishnamurthy, "Performance of Combined Free/Demand Assignment Multiple Access Schemes in Satellite Communications", *International Journal of Satellite Communications,* Vol. 14, No.1, pp. 11-21, 1996.

[19] T-S. Yum and E. W. M. Wong, "The Scheduled-Retransmission (SRMA) Protocol for Packet Satellite Communications", *IEEE Transactions on Information Theory*, Vol. 35, No. 6, 1989, pp. 1319-1324.

[20] D. Raychaudhuri, "Announced Retransmission Random Access Protocols", *IEEE Transactions on Communications*, Vol. COM-33, No. 11, 1985, pp. 1183-1190.