Bandwidth Allocation for Integrated MPEG-4 Video, WAP and MP3 Traffic over Next Generation Wireless Cellular Networks

P. Koutsakis
Dept. of Electrical and Computer Engineering
McMaster University, Canada
email: polk@ece.mcmaster.ca

Abstract
In this paper we explore, via an extensive simulation study, the performance of a Medium Access Control (MAC) protocol when integrating video, WAP and MP3 data packet traffic over a wireless picocellular system of very high capacity, with errors. Mobile terminals are considered to be high performance devices with extended storage capabilities which can act like cache memories streaming multimedia material. Our scheme achieves high aggregate channel throughput in all cases of traffic load, despite the very bursty nature of the examined traffic and the introduction of errors in the system.

Keywords- Medium Access Control, Multimedia Traffic, Channel Errors, QoS, Wireless Cellular Networks.

1. Introduction

In this work, we design and evaluate a multiple access scheme which efficiently integrates video and bursty data (WAP, MP3) traffic in high capacity picocellular environments. We focus on the uplink (mobiles to base station) channel, where a MAC scheme is required in order to resolve the source terminals contention for channel access.

The main concern in wireless communications is how to extend the broadband frontier to the end user. Therefore, the system capacity needs to grow as much as the demands of the various services provided in a wireless network grow. The use of an efficient MAC protocol to exploit the variations in access and service required by the disparate user sources helps to increase system capacity. Within the picocell, spatially dispersed mobile terminals (MTs) share a radio channel that connects them to a fixed Base Station (BS). Fixed-length packets arriving at the mobiles are buffered at the terminals until they are transmitted on the uplink to the BS. The BS allocates channel resources, delivers feedback information and transfers the packets to the wired networks and the Internet [1].

MTs are considered in our work to be high performance devices with extended storage capabilities which can act like cache memories, streaming multimedia material to other MTs, in order to reduce client start-up latencies and improve network performance. Even the whole picocell can be considered as a cache farm from the rest of the network. Therefore, we envision high quality stored video streaming on the uplink channel (MTs to the BS) in order for these streams to be delivered for playback to their destinations. The delivery of high quality stored video material is a service with very strict Quality-of-Service (QoS) requirements, hence the design of a MAC scheme which integrates this class of service with others, often requiring contradictory QoS guarantees, is necessary.

2. Traffic Models

A. MPEG-4 streams

MPEG-4 is an ISO standard that leverages existing digital video content by supporting the MPEG-1 and MPEG-2 coding standards, and enables richer development of new digital video applications and services. An MPEG-4 scene is composed by a number of audiovisual objects, which can be either static or time-varying in nature [5]. In addition MPEG-4 has a so-called Fine Granularity Scalability mode that allows transmission of the same video content at different bit rates from one coded version of the movie [6].

In our study, we use the trace statistics of actual MPEG-4 streams from [7, 8]. The video streams that we used were carefully selected in order to cover a broad range of movie characteristics. Table 1 shows the movies we used together with their mean and peak bit rates. A new video frame is generated every 40 msecs. Since the video streams are pulled out from MTs that act like caches and not dedicated cache engines with optimization or smoothing functionalities, no packet shaping mechanisms were used. Video content is delivered to the network in the initial form that was encoded. In our study we have made the assumption that all video packets of a

0-7695-2858-9/07 $20.00

video frame (VF) must be delivered before the next VF arrives. The allowed video packet dropping probability is set to 0.0001 from [9].

B. WAP Traffic Model

We adopt the WAP traffic model from [10], which corresponds to the WAP release 1.2.1. A WAP session consists of several requests for a deck performed by the user. The number of decks is modeled by a geometric distribution with mean equal to 20 decks and the packet size by a log2-normal distribution. To cover the influence of different applications, four different types of user profiles are introduced: email, news, m-commerce and common (referring to mixed traffic traced from a WAP server in real operation). The parameters used in the models presented in [10] were derived by measurements performed at a WAP gateway in test operation. These parameters are presented in Table 2. At the beginning of our simulation we let each WAP user choose randomly, with equal probability, one of the four user profiles presented above. We set an upper bound of 1 minute in the total transmission delay of a WAP session; given the high mean number of decks, the significant standard deviation and the presence of very bursty video traffic in the system, this is a quite strict upper bound.

C. MP3 traffic

The authors in [2] present the packet size distribution for MP3 traffic downloaded from "Napster". The distribution is, to a large extent, a discrete one, where packets of sizes equal to 40 bytes, 576 bytes and 1500 bytes constitute the largest amount of the overall packet sizes. More specifically,

- 34.98% of the total number of packets have a length of 40 bytes.
- 45.54% of the total number of packets have a length of 576 bytes.
- 4.18% of the total number of packets have a length of 1500 bytes.
- the remainder 15.3% of the total number of packets have lengths uniformly distributed between 40 bytes and 1500 bytes.

In our study, we consider all MP3 packets as messages comprising of a number of ATM-length packets equal to the MP3 packet length, in bytes, divided by 53 (ATM-packet length). Message interarrival times follow an exponential distribution (from [3], where the same MP3 traffic model is used). The mean file size is 4.37 MB [4], and MP3 file downloads vary in size according to a geometric distribution. The mean for the geometric distribution is equal to

$$\mu=(1-p)/p, \qquad (1)$$

where $0<p<1$. The variance for the geometric distribution is given by

$$\sigma^2 = (1-p)/p^2, \qquad (2)$$

and by using (1) we find that the standard deviation $\sigma = 4.83$ MB.

Today's Internet companies offering MP3 downloads to their subscribers aim at an average download time per MP3 file of 3-4 minutes, while cutting-edge wireless technology companies in Japan offer services of 6-8 minute downloads of 3-minute songs. In our study, we set the user QoS requirement for average time of an MP3 file download equal to a moderate upper bound of 10 minutes. If this limit is exceeded, we consider that the system has failed in offering the desired QoS to the user.

3. Frame structure, Base Station Scheduling and Actions of the Mobile Terminals

In our scheme, the uplink channel time is divided into time frames of equal length. Each frame consists of two types of intervals. These are the request intervals and the information intervals. Each of these intervals is divided into a number of time slots.

Within an information interval, each slot accommodates exactly one, fixed length, packet that contains video or data (WAP, MP3) information and a header.

Each request slot is subdivided into two mini-slots and each mini-slot accommodates exactly one, fixed length, request packet. By using more than one minislot per request slot, a more efficient usage of the available request bandwidth is possible, as explained in [11]. In that work, we studied the integration of videoconference streams (not actual movies as in this work, as MTs were considered simple devices which could not act as cache memories) with voice and email data traffic (instead of the much burstier MP3 and WAP traffic, as in this study) over a lower capacity wireless channel without errors (a burst-error model is adopted in this work) and with a different frame structure and BS scheduling policy.

The base station broadcasts a short binary feedback packet at the end of each mini-slot indicating only the presence or absence of a collision within the mini-slot (collision (C) versus non-collision (NC)). Upon successfully transmitting a request packet the terminal waits until the end of the corresponding request interval to learn of its reservation slot (or slots). If unsuccessful within the request intervals of the current frame, the terminal attempts again in the request intervals of the next frame. A terminal with a reservation transmits freely within its reserved slot.

In contrast to [11], where no contention was assumed among video terminals (they were considered to "live" permanently in the system), in this work *we propose the idea that video and data (WAP, MP3) terminals can share the request slots* (WAP terminals enter contention after the end of the video contention period, followed by MP3 terminals; the reason for the higher priority of WAP terminals is that users are expected to be more tolerant to MP3 download delays than to WAP access delays). The reason we require video terminals to contend in order to reserve information slots is that in this work we are dealing with high quality stored video content which, in contrast to videoconference streams, generates highly bursty information. Video requirements change frequently over time and it is impossible to ensure that their demands will reach the BS successfully if we use techniques similar to the ones described in [11]. The two-cell stack reservation random access algorithm [12] is used to resolve the collisions, among video request packets. The two-cell stack random access algorithm [12] is used to resolve collisions among WAP or MP3 packets.

When a video terminal wants to decrease its bit rate (because of a decrease in the number of bits needed for the transmission of the next video frame), it releases all the slots that were previously allocated to it and are no longer necessary. The BS consequently allocates the newly released slots to any other requesting terminals, after the end of the request interval of the next frame.

In our scheme, when the bit rate of a video terminal increases then the terminal must enter the contention process in order to acquire the additional information slots it needs. In such a case, the terminal releases all of its currently reserved information slots before entering the contention period. When it passes the contention period successfully, then the required information slots are reserved and the video terminal can freely transmit within them. The rationale behind this policy is fairness. We prefer to let each video terminal compete with other video terminals at an equal basis (since video has absolute priority), in order to "spread" the video packet losses uniformly.

The BS allocates channel resources at the end of the corresponding request interval. Video terminals have highest priority in acquiring the slots they demand, due to their strict Quality-of-Service requirements. If a full allocation is possible, the BS then proceeds to allocate any still available information slots to the requesting data (WAP, MP3) MTs. In the case that a new VF arrives and the number of data slot reservations is such that the video terminal can not be fully serviced, *the BS in this study preempts WAP and MP3 reservations* in favor of video terminals waiting for transmission (MP3 terminals, which have the lowest priority, are preempted first). If a full allocation to the video terminal is still not possible, the BS in our scheme grants to the video terminals as many of the slots they requested as possible (i.e., the BS makes a partial allocation). The BS allocates the earliest available information slots to the video terminals, which, if needed, keep these slots in the following channel frames, until the next video frame (VF) arrives.

There are two reasons for proposing the partial allocation policy, described above, for video users. The first is that *in our system there are ten available request slots (twenty minislots) in each channel frame*, which practically means that almost all of the possible collisions among the (few) request packets of video terminals which request an increase in their bit rate can be resolved within the contention period of one channel frame. The second reason is that video terminals in this study are transmitting high quality bursty video. The size of a video frame is usually of the order of a few kilobytes, which means that after segmentation it corresponds to a few hundred channel cells (of ATM size). Especially in MPEG-4 coding, the correlation factor between two adjacent video frames, is extremely high. Therefore, even in the case that a few tens of cells do not reach their destination, the decoder can successfully reconstruct the video frame. Additionally, MPEG-4 incorporates very capable error correction techniques and can therefore suffer relatively higher packet loss rates compared to other media standards without serious degradation of the playback quality. Consequently, we try to conform to the QoS requirement on packet dropping while spreading the packet losses in time. This spreading of losses can lead to an imperceptible blurring during video playback, which is expected to be acceptable when watching a movie, instead of concentrating the lost cells in time, and consequently receiving video "hiccups" and severe losses in the decoder.

4. Error Model

We adopt in our study a robust error model for wireless channels, presented in [13]. This model, with the use of the short and long error bursts makes accurate predictions on the long term correlation of wireless channel errors. The error model consists of a three-state discrete-time Markov chain, where one state is the "good state" (error-free) and the other two states are the "bad states", the *long bad* and the *short bad* state respectively (the Markov chain is shown in Figure 1). A transmission is successful only if the channel is in the "good state" (G); otherwise it fails. The difference between the long bad (LB) and short bad (SB) states is the time correlation of errors: LB corresponds to long bursts of errors, SB to short ones.

The parameters of the error model are presented in Table 3. The average number of error bursts, in slots, experienced when the states LB and SB are entered, are respectively given by: $B_{LB}=1/p_{bg\_L}$ and $B_{SB}=1/p_{bg\_S}$, where p_{bg_S} is the transition probability from state SB to G, and p_{bg_L} is the transition probability from state LB to G. Similarly, the average number of consecutive error-free slots is given by $B_G=1/p_{gb}$, where p_{gb} is the probability to leave state G. The parameter k is the probability that the Markov chain moves to state LB given that it leaves state G; k also represents the probability that an error burst is long (i.e., the fraction of long bursts over the total number of error bursts).

The only difference between [13] and the error model used in this work is that we do not consider the "good" channel state to be perfectly error-less, or the "bad" channel state to correspond to 100% transmission errors. The error probability in the good state is taken to be equal to $0.2*10^{-4}$, and the error probability in the bad state (SB or LB) to be equal to 0.8.

We have chosen in our study the value of the probability P_{bad}, i.e., the steady-state probability that the channel is in bad state, to be equal to $8*10^{-5}$; this value has been chosen in order to test an "almost worst" case scenario for our system, as the video packet dropping probability is set to 10^{-4} and by choosing a value of bad state probability larger than the upper bound on video packet dropping, the strict QoS requirement of video users would certainly be violated. The value for p_{gb} has been taken from [13], while the bad state probabilities P_{SB} and P_{LB} are considered equal (i.e., $4*10^{-5}$ each); p_{bg_L} is derived from the steady-state behavior of the Markov chain, and depends on the parameter k, which is taken equal to 0.05 [13].

5. System Parameters

The channel rate is 20 Mbps (from [14]), and the frame duration equal to 12 ms (from [9, 11]). The 12 ms of frame duration accommodate 566 slots (556 information slots plus 10 request slots).

6. Simulation Results and Discussion

Via an extensive simulation study we studied system behavior and performance under all possible movie loads from 1 to 6 movies. For each load, all different combinations were examined (63 in total), and each simulation point is the result of an average of 10 independent runs, each simulating one hour of network operation (Monte Carlo method).

Table 4 presents the results (maximum arrival rate of WAP sessions/second for which the system satisfies the QoS of all types of users) for each movie separately, in the case of 40 MP3 uploads (corresponding to an average load of 388 Kbps). The video clips trace turned out to be the most demanding stream in contrast to the lecture trace, which was the less demanding. The lowest and highest throughputs were achieved, respectively, by the system for these two movies.

Table 5 shows the average results of the simulation runs under all possible movie loads. As the number of video terminals increases, the aggregate bit rate gets smoother, since the superposition of the movies is known from the literature to always be less bursty than a single movie, and the average bandwidth of the superposition is close to the sum of the mean bit rates of the movies. This is proven once more in our study, as the case of the system accommodating only one movie is shown to provide the smallest throughput than any other case of movie load.

7. Conclusions

A Medium Access Control scheme is proposed in this work for the integration of MPEG-4 video, WAP and MP3 traffic from high performance devices, over high capacity wireless networks. The results of our study show that the proposed mechanism achieves very high aggregate channel throughput (more than 80% in almost all the studied cases), while preserving the QoS requirements of all traffic types.

References

1. M. Conti, C. Demaria, and L. Donatiello, "Design and Performance Evaluation of a MAC Protocol for Wireless Local Area Networks", *Mobile Networks and Applications (MONET) Journal,* Vol. 2, pp. 69-87, 1997.
2. A. Klemm, C. Lindemann and M. Lohmann, "Traffic Modeling and Characterization for UMTS Networks", *in Proceedings of the IEEE Globecom 2001,* San Antonio, USA, pp. 1741-1746.
3. R. Garcia, V. Garcia, X. G. Paneda, D. Melendi and A. Neira, "Aggregated Traffic Generation in FTTX Networks", *in Proceedings of the International Conference on Applied Computing,* Lisbon, Portugal, 2004.
4. H.-C. Kim, D. Lee, J. Lee, J. J. Suh and K. Chon, "A Measurement Study of Storage Resource and Multimedia Contents on a High-Performance Research and Education Network", *in Proceedings of the 6th IEEE International Conference on High-Speed Networks and Multimedia Communications (HSNMC),* Estoril, Portugal, 2003, pp. 108-117.
5. A. Basso, S. Varakliotis, and R. Castagno, "Transport of MPEG-4 over IP/RTP", i*n Proceedings of the IEEE International Conference on Multimedia and Expo*, New York, U.S.A., 2000.
6. L. Chiariglione, "MPEG-4, why use it?", Telecom Italia Lab – Italy, [Online] http://leonardo.telecomitalialab.com/paper/mpeg-4/
7. [Online] http://peach.eas.asu.edu/index.html
8. F. H. P. Fitzek and M. Reisslein, "MPEG-4 and H.263 Video Traces for Network Performance Evaluation", *IEEE Network,* Vol. 15, No. 6, 2001, pp. 40-54.
9. D. A. Dyson and Z.J. Haas, "A Dynamic Packet Reservation Multiple Access Scheme for Wireless ATM", *Mobile Networks and Applications Journal*, Vol. 4, No.2, pp. 87-89, Jan. 1999.
10. P. Stuckmann and C. Hoymann, "Dimensioning GPRS Networks for Coexisting Applications based on WAP and Conventional Internet Protocols", *in Proceedings of the European Wireless Conference 2002,* Florence, Italy.
11. P. Koutsakis, S. Psychis and M. Paterakis, "Integrated Wireless Access for Videoconference from MPEG-4 and H.263 Video Coders with Voice, Email and Web Traffic", *IEEE Transactions on Vehicular Technology,* Vol. 54, No. 5, 2005, pp. 1863-1874.
12. M. Paterakis and P. Papantoni – Kazakos, "A Simple Window Random Access Algorithm With Advantageous Properties", *IEEE Transactions on Inform. Theory*, Vol. IT- 35, No. 5, September 1989, pp.1124-1130.
13. M. Bottigliengo, C. Casetti, C.-F. Chiasserini and M. Meo, "Short-term Fairness for TCP Flows in 802.11b WLANs", *in Proceedings of the IEEE Infocom 2004,* Hong Kong, China.
14. G. Colombo, L. Lenzini, E. Mingozzi, B. Cornaglia and R. Santaniello, "Extended Performance Evaluation of PRADOS: A Scheduling Algorithm for Traffic Integration in a Wireless ATM Network", *Wireless Networks*, Vol. 8, No. 2-3, 2002, pp. 265-274.

Movie Name	**Mean Bit rate (Kbps)**	**Peak Bit rate (Kbps)**
Lecture	210	1500
Jurassic Park	770	3300
Video Clips	1000	3700
Aladdin	440	3100
Formula 1	840	2900
Die Hard III	700	3400

Table 1. Trace Statistics.

Type of application	Size of "Content" (bytes)	
	Mean	Stand. Deviation
News	638.4	389.9
Email	582.9	260.0
M-commerce	641.0	342.9
Common	511.0	368.0

Table 2. Parameters of the WAP traffic model.

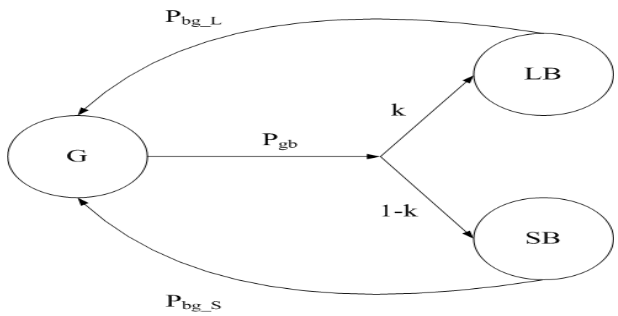

Figure 1. Channel Error Model.

P_{good}=0.99992
$B_G=1/p_{gb}$=65160 slots
$B_{SB}=1/p_{bg\_S}$=2.5 slots
$B_{LB}=1/p_{bg\_L}$

Table 3. Error Model Parameters.

Movie	λ (WAP sessions/second)	Channel Throughput (%)
Lecture	121.36	85.44
Jurassic Park	108.97	82.17
Video Clips	64.58	65.59
Aladdin	88.02	76.27
Formula 1	111.31	83.41
Die Hard III	106.24	80.55
Average	**100.08**	**78.91**

Table 4. Integration Results for each movie (40 MP3 terminals present in the system).

System Load	Average λ (WAP sessions/second)	Average Channel Throughput (%)
1 Movie	100.08	78.91
2 Movies	87.09	80.22
3 Movies	73.18	82.30
4 Movies	56.31	87.86
5 Movies	29.74	91.22
6 Movies	6.86	91.98

Table 5. Integration Results for the superposition of movies (40 MP3 terminals present in the system).