Integrated Voice, Video, and Bursty Data Wireless Access in TDMA-Based Mobile Networks

P. Koutsakis and M. Paterakis

Dept. of Electronics and Computer Engineering
Telecommunications and Information & Computer Networks Laboratories
Technical University of Crete, Chania, Greece
tel.: +30 821 37343, e-mail:polk@telecom.tuc.gr, pateraki@telecom.tuc.gr

ABSTRACT

A new medium access control (MAC) protocol for mobile wireless communications is presented and investigated. We explore, via an extensive simulation study, the performance of the protocol when integrating voice, video and bursty data packet traffic over a wireless picocellular system of high capacity.√

I. SYSTEM MODEL

In this work, we design and evaluate a multiple access scheme which efficiently integrates voice (Constant Bit Rate, CBR On/Off Traffic), video (Variable Bit Rate, VBR) and bursty data 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.

A. Channel Frame Structure

The uplink channel time is divided into time frames of equal length. The frame duration is selected such that a voice terminal in talkspurt generates exactly one packet per frame. Each frame consists of two *types* of slots. These are the *request* slots and the *information* slots.
Each information slot accommodates exactly one, fixed length, packet that contains voice, video or data information and a header. As for the request slots, *we introduce the idea* [6] *that the request slots can be shared by voice and data terminals (first by voice terminals and, after the end of voice contention, by data terminals)*, in order to optimize the request bandwidth. The request slots are subdivided into mini-slots and each mini-slot accommodates exactly one, fixed length, request packet. Since we assume that all of the voice source state transitions occur at the frame boundaries, we place all request intervals at the beginning of the frame, in order to minimize the voice packet access delay.
Voice and data terminals do not exhaust their attempts for a reservation within the request intervals. *Any other free, at the time, information slot of the frame can be temporarily used as an extra request slot (ER slot)* for voice or data terminals [1] (with priority given to voice terminals).
The concept of reserving a minimum bandwidth for terminals to make reservations helps to keep the voice access delay within relatively low limits and gives clearly better performance than the PRMA [2] and quite a few PRMA-like algorithms, such as DPRMA [4].
No request slots are used for the video terminals, because of two reasons, which will be be analyzed in paragraph I.D.

B. Voice and Video Traffic Models

Our primary voice traffic model assumptions are the following:

1) Voice terminals are equipped with a voice activity detector [2]. Voice sources follow an alternating pattern of talkspurts and silence periods (on and off).
2) All of the voice source transitions (e.g., talk to silence) occur at the frame boundaries. This assumption is reasonably accurate, taking into consideration that the duration of a frame is equal to 12 ms here, while the average duration of the talkspurt and silence periods exceeds 1 sec.
3) The voice delay limit is equal to 40 ms.
4) The channel is error-free and without capture.
5) Reserved slots are deallocated immediately.

We adopt the same video traffic model with the one in [4]. This model is based upon work done by Heyman, et al [3]. In this study of actual videoconferencing traffic, video frames (VFs) were found to be generated periodically and to contain a varying number of cells in each frame. The distribution of the number of cells per VF was found to be described by a gamma (or equivalently negative binomial) distribution.
The statistics for video conferencing traffic in [3], were the result of coding a video sequence with a modified version of the H.261 standard and are presented in Table 1. New VFs are assumed to arrive every 40 msecs.

C. Data traffic models

Data traffic has low priority compared to voice traffic, and data messages are generated by a large unknown number of data terminals (theoretically infinite). Data messages vary in length according to a geometric

√ *This research has been funded by the Greek General Secretariat of Research and Technology and the European Social Fund, through PENED '99 grant no 99ED 192.*

distribution with mean B packets per message, and the steady state data packet arrival rate is equal to λB packets/frame, where λ is the data message arrival rate. (B is equal to 8 in our study).

We consider two different cases for the data message arrivals:

1) The data message arrivals are Poisson distributed, and
2) The interarrival times of the data message arrival process are assumed independent, and identically distributed according to a Pareto distribution with shape parameter α and location parameter k [5]:

The location parameter k corresponds to the minimum interarrival time between packets. It can be easily shown that if α<=2, the distribution has infinite variance, while if α<=1 the mean is infinite as well. Therefore, the Pareto distribution is heavy tailed with infinite mean and variance. This mathematical property accounts for the strong burstiness of the data message arrival process in this case.

Finally, an upper limit on the mean data message delay, equal to 200 ms, is assumed.

D. Actions of Voice and Video Terminals, and Base Station Scheduling

Voice terminals with packets, and no reservation, contend for channel resources using a random access protocol to transmit their request packets only during the request intervals. 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.

The same reservation procedure is followed by the data terminals, right after the end of the voice contention period, which is indicated by two consecutive NC feedback packets. Since data traffic is the most tolerant of delays, data terminals have the lowest priority in acquiring the information slots they demand. Each data user is allowed to reserve just one slot per frame. This “restriction” is explained later in this section.

Video terminals do not have any request slots dedicated to them. This happens for two reasons:

1) Video sources “live” permanently in the system, they do not follow an ON-OFF state model like voice sources.
2) Video traffic follows a multi-state Markov model, in which however state transitions do not occur very often.

Thus, there is no need for granting request bandwidth to the video terminals, as it would be wasted in most cases. *Video terminals convey their requirements to the base station by transmitting them within the header of the first packet of their current video stream.*

Upon successful receipt of a voice or data request packet, the BS provides an acknowledgment and queues the request. The BS allocates channel resources at the end of the corresponding request interval. Video terminals have absolute priority in acquiring the slots they demand. If a full allocation is possible, the BS then proceeds to allocate any still available information slots to the requesting voice terminals. Otherwise, if a full allocation is not possible, the BS grants to the video users as many of the slots they requested as possible (i.e., the BS makes a partial allocation). The BS keeps a record of any partial allocations so that the remaining requests can be accommodated whenever the necessary channel resources become available. In either allocation type case, 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.

Voice terminals which have successfully transmitted their request packets do not acquire all the available (after the servicing of video terminals) information slots in the frame. If this happened, voice terminals would keep their dedicated slots for the whole duration of their talkspurt (mean of 1 sec here), and thus video terminals, would not find enough slots to transmit in, and the particularly strict video QoS requirements would be violated. Consequently, *the BS allocates a slot to each requesting voice terminal with a probability* ***p*****. The requests of voice terminals which ″fail″ to acquire a slot, based on the above, remain queued. Within each priority class, the queuing discipline is assumed to be First Come First Served (FCFS).

The above comments on why we need to use **p***, stand for the explanation of our “defensive” slot allocation policy to data users, as well.

Finally, in order to preserve the video QoS, *we enforce a scheduling policy for the video terminals* which prevents unnecessary dropping of video packets in channel frames within which the arrival of a new VF of a video user takes place (the details of this reshuffling policy can be found in [6].

E. Transmission Protocol

In our study, we adopt the *two-cell stack* reservation random access algorithm for use by contending voice terminals, due to its operational simplicity, stability and relatively high throughput when compared to the PRMA (Aloha-based) ([1]) and PRMA-like algorithms, such as [1,4].

The actions of the data terminals, in order to acquire access to the channel resources, are slightly different from those of voice terminals.

II. SYSTEM PARAMETERS

The simulations were conducted with the parameters contained in Table I. Each simulation point is the result of an average of 10 independent runs, each simulating 305,000 frames (the first 5,000 of which are used as warmup period).

The number of shared voice and data request slots depends on the number of video sources admitted into the system[1]*, and it varies accordingly between 1 and 5 slots (see Table II, where the values of the probability* ***p**** *are presented as well).*

Even for the case where 5 request slots are needed, this corresponds to a 1.95% request bandwidth only. We should also note that:

1) In our design, we chose the number of minislots per request interval (4), to allow for guard time and synchronization overheads, for the transmission of a generic request packet and for the propagation delay within the picocell.
2) Since the maximum transmission delay for video packets is set to 40 msecs, all video packets of a VF must be delivered before the next VF arrives.

III.VOICE-VIDEO-DATA INTEGRATION RESULTS AND DISCUSSION

The two parameters of the Pareto distribution are realated through the following equation:

$$(\alpha-1)/\alpha = \lambda^* k$$

which means that the ratio $(\alpha-1)/\alpha$ is equal to the ratio of the minimum data message interarrival time (k) to the mean data message interarrival time ($1/1/\lambda$).

We have chosen to investigate the cases when the above ratio is equal to 0.2, 0.3, and 0.4 respectively. Thus, our simulations have been conducted for three different values of the shape parameter α, 1.25, 1.429 and 1.667, respectively. Therefore, the data message arrival Pareto distributions considered have finite mean and infinite variance. The closer the value of α is to 1, the burstier the distribution is.

Tables III-VI present the results for the maximum voice capacity obtained for various values of the data message arrival rate λ, and for various numbers of video users (0,1,3, and 5, respectively), when all the QoS requirements for the three traffic types are satisfied.

We can make the following observations:

1) Regarding the Poisson distribution, the smoothest transitions for the maximum voice capacity, as λ increases, take place when no video users exist in the system. This is due to the fact the number of voice terminals has to be drastically decreased as the data message arrival rate increases, in order to cope with the burstiness of the video traffic and still be able to preserve the QoS requirements.
2) The results confirm that, for the smaller values of the Pareto shape parameter α, i.e. for the "burstier" Pareto distribution, the decrease in the maximum voice capacity is the greatest compared to the results for the Poisson distribution. Still, we observe from our results in Tables III,V, and VI that as λ increases, and especially for λ>2, the differences in the maximum voice capacity achieved for different α values are smoother than the ones for smaller λ values. This means that the burstiness given by the smaller α values are overcome by the fact that, for λ>2, the data load becomes "heavy" for the system to accommodate. The only exception seems, by the results presented in Table IV, to be the case where 1 video user is admitted in the system. In this case, we observe that, as λ increases, the differences in the maximum voice capacity achieved by the system for the different α values increase as well. This can be explained by the fact that 1 video source alone behaves in a quite more bursty way than 3 or 5 video sources which are simultaneously active do.
3) From Tables IV and VI, we observe, when 5 video users are admitted in the system, that the decrease in the maximum voice capacity, when using the Pareto data message interarrival distribution instead of the Poisson message arrival distribution, is very large compared to the decrease when 1 video user is admitted in the system. This is a general conclusion: as the video load increases, the maximum voice capacity achieved with the Pareto data traffic model suffers a greater % decrease, for all the λ values, in comparison to the maximum voice capacity achieved when using the Poisson data traffic model. The very small allowed video dropping probability is responsible for this result.
4) As shown in Tables III and V (this also stands for the other two cases), our scheme achieves quite high throughput for all the cases investigated in this paper, i.e. not only in the case of the Poisson, but also in the case of the very bursty Pareto data traffic model. The very good results achieved by our scheme are a consequence of the combination of three factors: a) our voice slots allocation policy, b) our video slots scheduling policy, and c) the use of the unused information slots as extra request slots. These three factors help us ″exploit″ all the available information slots within each frame in the best possible way.

REFERENCES

1. N. M. Mitrou, Th. D. Orinos and E. N. Protonotarios, ″ A Reservation Multiple Access Protocol for Microcellular Mobile-Communication Systems ″, *IEEE Transactions on Vehicular Technology*, Vol. 39, No. 4, Nov. 1990, pp. 340-351.
2. S. Nanda, D. J. Goodman and U. Timor, ″ Performance of PRMA: A packet voice protocol for cellular systems ″, *IEEE Transactions on Vehicular Technology*, Vol. 40, pp. 584-598, 1991.
3. D. Heyman, T. V. Lakshman, A. Tabatabai and H. Heeke, ″ Modeling teleconference traffic from VBR video coders″, *International Conference on Communications '94,* vol. 3, pp. 1744-1748, 1994.

[1] The channel bandwidth consumed by each video source is large, and thus, when we examine cases with a small number of video sources, the system can accommodate a significantly larger number of voice sources, therefore more voice request slots are needed.

4. D. A. Dyson and Z. J. Haas, , "A Dynamic Packet Reservation Multiple Access Scheme for Wireless ATM", *ACM/Baltzer MONET Journal, 1999.*

5. Z. Harpantidou and M. Paterakis, "Random Multiple Access of Broadcast Channels with Pareto Distributed Packet Interarrival Times", *IEEE Personal Comminications, Vol.5, No.2, pp.48-55, April 1998.*

6. P. Koutsakis and M. Paterakis, ″On Multiple Traffic Type Integration over Wireless TDMA Channels with Adjustable Request Bandwidth″, *International Journal of Wireless Information Networks,* Vol. 7, No.2, pp.55-68, 2000.

Design Parameters	
Channel Rate (Mbps)	9.045 [4]
Speech Codec Rate (Kbps)	32
Frame Duration (ms)	12
Slots per Frame	256
Minislots per request slot	4
Packet size (bytes)	53 (5 header)
Voice and video delay limit (ms)	40
Mean talkspurt (silence) duration (s)	1.0 (1.35)
Maximum voice (video) dropping probability	0.01 (0.0001)
Peak (average) video bit generation rate (Mbps)	2.112 (1.006)
Standard deviation of video bit rate (Mbps)	0.285

Table I. Experimental System Parameters

Number of Video users	**Number of request slots**	**Probability p***
6	1	0.0072
5	1	0.03
4	2	0.06
3	2	0.085
2	3	0.128
1	4	0.18
0	5	1

Table II. Adjustable request bandwidth depending on the number of video users.

λ (mes./frame)	α=1.25		α=1.429		A=1.667		Poisson	
	V. Cap.	%Throu.	V. Cap.	%Throu.	V. Cap.	%Throu.	V. Cap.	%Throu.
0.5	521	89.9	529	91.3	531	91.6	557	96.0
1	518	91.0	527	92.5	529	92.9	549	96.3
1.5	510	91.2	514	91.9	524	93.6	539	96.2
2	488	89.1	491	89.6	495	90.3	530	96.2
3	449	85.7	453	86.4	454	86.5	510	96.0
4	432	86.0	434	86.3	435	86.5	494	96.5

Table III. Maximum Voice Capacity and Throughput when 0 video users are admitted in the system.

λ (mes./frame)	α=1.25	α=1.429	α=1.667	Poisson
0.5	409	413	414	433
1	392	396	397	410
1.5	378	381	384	400
2	366	371	373	385
3	336	347	354	362
4	310	322	327	328

Table IV. Maximum Voice Capacity when 1 video user are admitted in the system.

λ (mes./frame)	α=1.25		α=1.429		α=1.667		Poisson	
	V. Cap.	%Throu.	V. Cap.	%Throu.	V. Cap.	%Throu.	V. Cap.	%Throu.
0.5	221	75.7	232	77.6	241	79.1	264	82.7
1	201	74.0	214	76.1	222	77.5	244	81.0
1.5	185	72.8	195	74.5	203	75.9	229	80.0
2	170	71.9	178	73.3	188	74.9	215	79.3
3	158	73.1	165	74.2	174	75.7	192	78.7
4	138	72.9	144	73.9	147	74.4	173	78.7

Table V. Maximum Voice Capacity and Throughput when 3 video users are admitted in the system.

λ (mes./frame)	α=1.25	α=1.429	α=1.667	Poisson
0.5	19	29	38	90
1	14	18	22	70
1.5	12	16	20	57
2	9	11	14	44
3	1	2	4	27
4	0	1	3	6

Table VI. Maximum Voice Capacity when 5 video users are admitted in the system.