MAC FOR WATM AIR INTERFACE: IMPACT OF ERROR CONTROL SCHEMES ON PROTOCOL DESIGN

Flaminio Borgonovo, Antonio Capone, Luigi Fratta, Mario Marchese

Politecnico di Milano, Dipartimento di Elettronica e Informazione, Italy
e-mail: {borgonov, capone, fratta, mmarches}@elet.polimi.it

Abstract – The extension of the ATM (Asyncronous Transfer Mode) protocol stack to the wireless segment requires the introduction of a MAC (Medium Access Control) layer able to provide reliable connection services to the network layer. Two basic functions of the MAC layer are resource allocation and error control. In this paper we analyze these issues and consider a MAC scheme for TDMA systems based on ARQ (Automatique Repeat reQuest) and a multi-level priority scheduling. The impact of radio channel characteristics and QoS requirements on the protocol design has been studied and a comparison of the ARQ-based approach and the classical approach based on FEC (Forward Error Correction) and interleaving has been performed.

I. INTRODUCTION

ATM technique was firstly designed for wired networks. Recently, the growing interest in telecommunication services for mobile users has pushed the research community to consider the use of the ATM technique into wireless access networks [1]. The basic system architecture is composed by some base stations (BS) which provide the access to a set of mobile terminals (MT) by means of a shared radio channel. Each BS is connected through a wired link to an ATM switch. The extension of the conventional ATM protocols stack into the wireless medium requires to take into account some new issues [2]: the unreliability of the radio medium, the user terminals multiple access to the common radio resources, and the user terminals mobility. The first two issues rely on the design of a MAC layer below the ATM layer [3] and are the objects of the study presented in this paper.

In order to guarantee a Quality of Service (QoS) to end-to-end ATM connections, the MAC layer must be able to exploit radio resources and provide reliable links to the ATM layer. The services offered by the MAC layer can differ according to a set of QoS parameters and are used by the ATM layer to meet the traffic profile and QoS requirements of the active connections. At least two parameters must be considered in multimedia environments: the maximum Cell Loss Rate (CLR) and Cell Transfer Delay (CTD). The elements the MAC layer must include in order to offer services with guaranteed CLR and CTD are a multiple access scheme and an error control mechanism. The first enables to subdivide the available bandwidth among active connections according to their traffic profile, while the second provides services with maximum CLR lower than those provided by the raw channel.

Different multiple access schemes have been considered for WATM [2][3]. The most suitable schemes seems to be those based on time division multiple access (TDMA) with centralized dynamic slot assignment (DSA), as confirmed by the work carried out in ETSI BRAN (European Telecommunications Standards Institute, Broadband Radio Access network) [6] and ATM Forum [7]. The common accepted framework considers a TDMA structure based on a time-frame composed by a Header (H), a Downlink Period (DP) and an Uplink Period (UP), slotted and of variable lengths [9]. The H is used by the BS to allocate uplink slots to connections, according to a scheduling algorithm. Each uplink and downlink slot can carry a MAC PDU, which contains the information of an ATM cell.

The error control scheme must be able to cope with the characteristics of the radio channel in order to meet the constraints on the CLR. The radio channel usually presents a quality, measured in terms of Bit Error Rate (BER), poor and time varying. In fact, the limited bandwidth is used by different base stations (BS) for the connection to mobile terminals (MT) and the interference noise is therefore high. Moreover, because of multi-path fading and shadowing, the received signal is characterized by random attenuation factors. The time correlation of these factors mainly depends on terminal speed and makes the channel characterized by blocks of errors. In fact, during time intervals in which the received signal falls at a level comparable with that of the noise, the BER is high and many errors occur. The cell loss rate (CLR) experienced over the air interface can be very high if no error control technique is adopted. In order to provide the CLR required by the ATM connections, the use of FEC and/or ARQ is mandatory. Most common adopted solutions are hybrid schemes: a minimum acceptable CLR is guaranteed by a proper channel coding (FEC) implemented at the physical layer and ARQ techniques are adopted only for non real-time services due the delay introduced by packet retransmissions. This approach has prevented to perform a joint optimization of the two error control techniques since they are usually ana-

0-7803-5565-2/99/$10.00 © 1999 IEEE

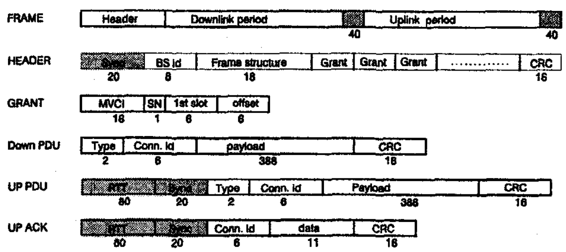

Figure 1: Frame architecture.

lyzed in different phases of the system design (Physical and MAC layer).

Old discussions [8] and recent results [10][11] suggest that ARQ techniques can be an appropriate solution also for delay constrained applications and they are even performing better than FEC ones. In our scheme we propose an approach based on ARQ also for real-time services. The proposed ARQ scheme, completely implemented in the BS, schedules transmissions and retransmissions for both uplink and downlink connections. Therefore, the acknowledgment and retransmission procedures have impact on the design of the scheduling algorithm and frame structure.

In this paper we evaluate the performance of the proposed scheme with the aim of showing how some parameters, such as the channel characteristics and the QoS constraints, influence the behavior of the error control mechanism. Moreover, we present some preliminary results on the comparison of a completely different approach based on classical FEC and bit interleaving. The channel model assumed is the Gilbert-Elliott one [5], as commonly considered for the mobile radio channel [4].

The paper is organized as follows. In section II we describe the structure of the considered MAC protocol. In section III the maximum number of connection (system capacity) which can be served with a fixed CLR versus the frame length is evaluated considering constant bit rate (CBR) connections and different value of the cell rate, the channel parameters and the maximum CTD. In section IV the MAC architecture is modified in order to be used with the classical FEC plus interleaving approach and an overall protocol efficiency is defined. Then the ARQ-based and FEC-based approaches are compared varying the system parameters. Finally, some concluding remarks and a look at further studies are given in section V.

II. MAC ARCHITECTURE

Frame structure

The complete frame structure, which specifies the parameter values adopted in the simulations, is shown in Figure 1. The frame length is fixed in order to allow an easier synchronization procedure at the physical layer. The DP contains the PDUs transmitted by the BS to the MTs in a unique radio burst. The UP is divided in a data period, an ack period and a signaling period. The MTs transmit their scheduled PDUs to the BS during the data period, and the acknowledgements of the PDUs received during the DP in the ack period. The signaling period is used to transmit MAC signaling information as for example that needed by the scheduler to know the status of traffic sources or MT transmitting queues. Reliable and fast signaling procedures are fundamental with traffic profiles characterized by a variable rate, but their definition is out of the scope of this paper. As shown in Figure 1 a guard time (RTT) is adopted to avoid the overlap of uplink transmissions and a syncronization field is used at the beginning of each radio burst. Between, the DP and the UP, and the UP and the H, there is a Turn Around Time (TAT) which allows the radio elements to switch from/to the receiving/transmitting mode. The lengths of all the internal periods of a frame are variable and are defined in the H. The H also contains the uplink grants used by the BS to solicit the PDUs transmission from active MTs. Though not necessary, even downlink grants are included in the H to allow MT to save energy during the DP.

The protocol is connection oriented. Each active connection is identified at MAC layer by a MVCI (Mobile Virtual Circuit Identifier) assigned during setup phase. Multiple connections can be setup by a MT. The MVCI is included in the grants which assign a set of slots to the connection, and allow the MT to transmit or receive a set of PDUs in the subsequent frame. The 1st slot and offset fields identify the position of the first assigned slot and the number of slots, while the SN (Sequence Number) field is used from the ARQ mechanism. To allow the receiver to detect corrupted data units, a parity field (CRC) is included in the H, PDU and ACK.

ARQ scheme

We have defined a MAC layer Real-Time ARQ technique to control errors on both real-time and non real-time applications. With such a scheme a cell loss occurs only when a MAC PDU can not be correctly delivered within its life-time, a parameter related to the maximum CTD. The performance of this technique depends on the number of retransmissions which can be performed during the life-time. For this reason the acknowledgments for downlink traffic are immediately transmitted by the terminals at the end of each frame, so that the scheduling optimizes the slot assignment in the next frame according to an up-to-date information and taking into account the remaining life-time for each packet.

Also for the uplink transmissions the ARQ procedure is performed by the BS which in the frame header indicates to the terminals which PDUs must be (re)transmitted. A single ACK per connection is sent at each frame even if more than one PDU has been transmitted. The positive ACK is sent only if all packets are correctly received, otherwise they are all retransmitted. For this we can define our technique a Multiple-Stop-and-Wait. The efficiency of this approach is in principle lower than that of a selective-repeat. However, due to the strong correlation of channel errors (bursts) and the heavy overhead required by the selective-repeat we have observed a negligible efficiency loss.

Scheduling mechanism

Slot allocation for both uplink and downlink transmissions is performed by the BS scheduling algorithm. This task is more complex than in ATM switches because the BS must consider also retransmissions and must have knowledge of the exact status of the queues in the terminals. For the latter we assume the existence of a proper signaling protocol that informs the BS about the connections queues status.

In our algorithm the retransmissions are taken into account by sorting transmission requests (for uplink and downlink packets) according to two priority fields. The first field is the number of transmissions: packets with less retransmission are scheduled first. The second field is the remaining life-time: packet with smaller residual life-time are scheduled first. If more than a request for the same connection is present in the scheduling queue, the sequence order is forced by the algorithm.

The rationale of the proposed priority-based scheme is this. The first field prevents the head-of-the-line effect which dramatically reduces system performance when a connection experiences very poor channel quality. Moreover, since the retransmissions do not steal resources to normal trasmissions, the scheduler behaviour can be easily predicted by the call admission controller. The second field allows to assign timely resources to connections with stringent delay requirements.

III. SIMULATION RESULTS

The system model used for simulations is based on the TDMA structure shown in Figure 1 and the procedures described in the previous section. We have considered a raw channel of 40 Mb/s, a RTT of 2 μs and a TAT of 1 μs. We model the physical channel by means of a two state (ON/OFF) Markov chain (Gilbert-Elliott model). When the state is ON the channel is error free, while when the state is OFF

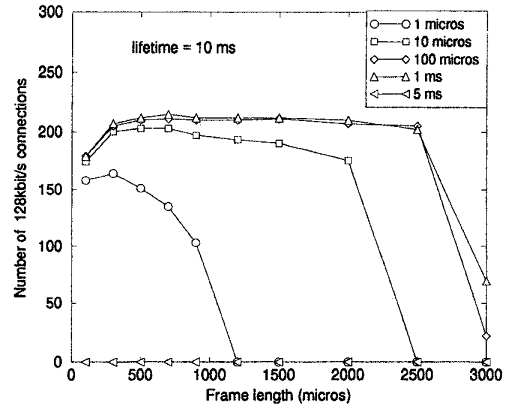

Figure 2: Maximum number of connections versus the frame length for different values of T_{off} and a lifetime equal to 10 ms.

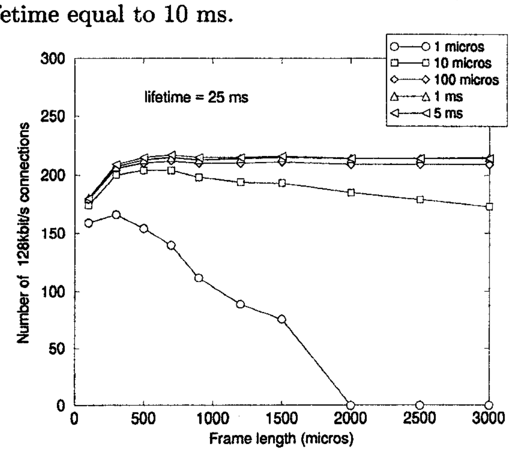

Figure 3: Maximum number of connections versus the frame length for different values of T_{off} and a lifetime equal to 25 ms.

the bit error probability is equal to 0.5. We assign the characteristics of the chain by means of the stationary probability, Pb=0.01 of being in the OFF state, and the average time spent in that state, T_{off}. These parameters are related by the equation:

$$P_b = \frac{T_{off}}{T_{off} + T_{on}} \tag{1}$$

where Ton is the average time spent in the ON state.

The loss of the header causes the complete loss of PDUs on both channels. A PDU must be correctly received within its lifetime L, otherwise it is dropped. The maximum CLR, equal to 10^{-3}, is the QoS requirement that must be satisfied.

We considered traffic sources characterized by a constant rate. For each connection the scheduling mechanism introduces requests in the virtual queue at

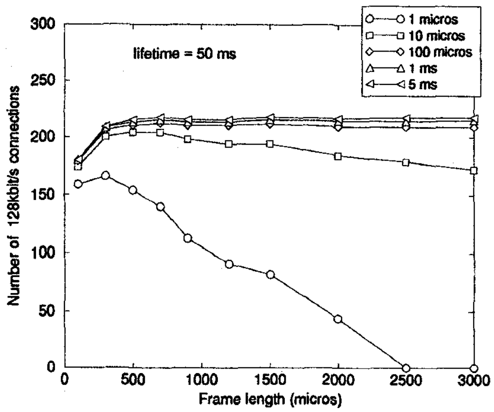

Figure 4: Maximum number of connections versus the frame length for different values of T_{off} and a lifetime equal to 50 ms.

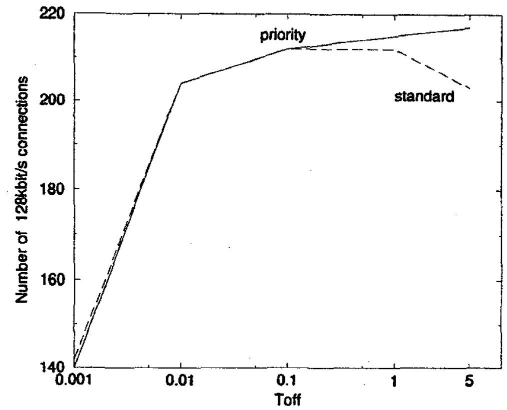

Figure 5: Maximum number of connections versus T_{off} for the priority-based and the *standard* schedulers, lifetime equal to 25 ms.

regular intervals according to the source rate. The results presented have been obtained with connection of 128 kbit/s. Similar results, obtained for other rates, are not shown for the sake of semplicity. Figures 2 shows the maximum number of connections that can be served with the required maximum CLR versus the frame length for different values of T_{off} and a lifetime of 10 ms. The curves with T_{off} in the range 1 μs - 1ms show a maximum. This behaviour is due to two opposite effects: the frame overhead and the ARQ efficiency. The number of connections is lower than the maximum with short frames because the overhead is heavier, and with long frames because few retransmissions can be performed to recover channel errors before lifetime expires. The highest curves have been obtained with a T_{off} equal to 100 μs and 1 ms. Channels with a faster dynamics (T_{off} lower that a timeslot) present a higher probability that a PDU is received corrupted and, therefore, the number of connections is lower. On the contrary, with a T_{off} equal to 5 ms the number of connection which can be served is almost zero. This is due to the impossibility of ARQ to recover a sufficiently large number of errors since the probability that the channel remains in the OFF state for a time longer than the lifetime is too high. Figures 3 and 4 show the results obtained with a lifetime equal to 25 and 50 ms respectively. We observe that the system capacity increases with the lifetime. In fact, for a given frame length the number of possible retransmissions increases as the lifetime increases. Even with long frames the number of connections remains high. The flatness of the curves evidentiates a range of frame length values in which the capacity is near the maximum. Considering all the obtained results we suggest to adopt a frame lengh of 0.5 - 1 ms.

To evaluate the efficiency of the proposed priority-based scheduling mechanims we considered also a standard scheduler which performs the queue sorting only according to the remaining lifetime. For the priority-based and the standard schedulers, Figure 5 shows the maximum number of connections versus T_{off} with a frame length equal to 0.7 ms and a lifetime equal to 25 ms. With small values of T_{off} the two schemes offer almost the same capacity, while with large values of T_{off} the priority based algorithm show a better performance. In fact, when T_{off} is high and an error occurs, the probability that the subsequent frame be in the same OFF period is high. The standard scheme performs an immediate retransmission which results often useless, while the priority-based scheme first transmits the other PDUs and then performs the retransmission with a higher success probability.

IV. EFFICIENCY COMPARISON

To traslate the obtained capacity results in a figure of merit of the adopted scheme we defined the overall protocol efficiency as the ratio between the number of connections that can be served and the number of connections that could be served with an error-free channel and no overhead for the error control scheme. This last number can be easily evaluated considering a frame structure without acknowledments and error detection codes. The efficiency of the considered approach versus T_{off} and the lifetime is shown in Figure 6 for a frame length equal to 0.7 ms.

Further simulations have been run to compare the efficiency of the proposed approach and that of the classical approach based on FEC and bit interleaving. Traditionally adopted correcting codes are optimized for memory-less channels and, therefore, bit interleaving is used to spread bursty errors over the coded words. The simplest interleaving scheme is obtained using an array with N raws and N columns. The coded words, divided into N blocks, are stored in the array by raws and transmitted by columns.

The higher is the value of N, named interleaving depth, the larger is the errors spreading. However, the interleaving process introduces a delay linearly increasing with N and its maximum value is limited by the lifetime.

To evaluate the performance of the approach purely based on FEC and interleaving we modified the frame structure. In the PDUs the CRC is replaced by the FEC redundancy, and the ack period in the UP is no more present. For the simulations we assumed that the content of each PDU (k= 396 bits), as transferred by the MAC layer to the physical layer, is increased by a parity check of n-k bits to form a code block of length n. For each connection, N consecutive blocks undergo a diagonal interleaving process to provide N channel blocks that are transmitted in the assigned time slots in the frame. The codes used belong to the BCH family. Because of interleaving, the maximum value of N must be less than L/T, where T is the PDU inter-generation time and L the lifetime. The efficiency of this FEC plus interleaving scheme is shown in Figure 7. These results have been obtained assuming an error-free frame header. When errors in the header are also considered we measured a really poor performance. This proves how critical is the problem of the header protection in the classical approach: interleaving can hardly be used and the overhead required by a correcting code should be much heavy. However, even if the efficiency values of Figure 7 have been obtained with error-free headers, they are lower than those of Figure 6 for most of the T_{off} values. Only with a T_{off} equal to one bit period, i.e. with independent errors, the FEC approach outperforms the ARQ approach.

This result is not surprising. The interleaving tries to transform the memory channel into a memoryless channel with error rate ε equals to $P_b/2$. The capacity of such a channel is given by the well known shannon limit:

$$C = 1 + \varepsilon ln\varepsilon + (1+\varepsilon)ln(1+\varepsilon) \qquad (2)$$

However, for a given maximum CLR, practial coding schemes achieve a throughput far from the limit since the interleaving must have a finite depth due to the delay constraints and the code has limited complexity. On the other hand, the capacity of the considered Gilbert channel is higher than C [5], and it is equal to $(1-P_b)$ if the channel status is known at the receiver. The redundancy required to provide this information reduces the capacity to $(1-P_b)\eta$, with $\eta < 1$. The throughput of a packet system with a pure ARQ can be $(1-P_b)\eta\gamma$, where γ is an additional reducing factor due to the ON/OFF periods not equal to an exact multiple of a slot. When T_{off} is not too small but enough smaller than the lifetime, $\eta\gamma$ is near 1 and the throughput higher than that achievable with FEC and interleaving. When T_{off} is comparable to a timeslot, $\eta\gamma$ is much low while the efficiency of FEC schemes is high since errors are almost independent. Finally, when T_{off} is high, even longer than lifetime, no scheme can recover errors due to delay constrains.

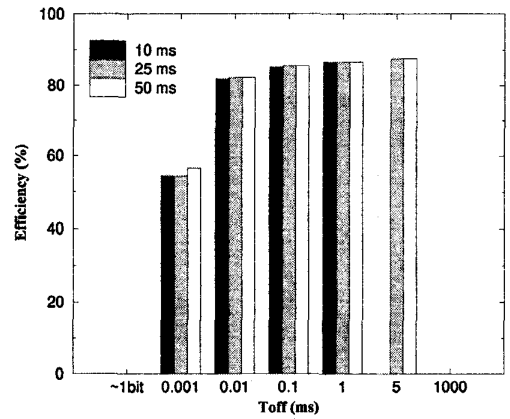

Figure 6: Efficiency of the ARQ approach versus T_{off} and versus the lifetime.

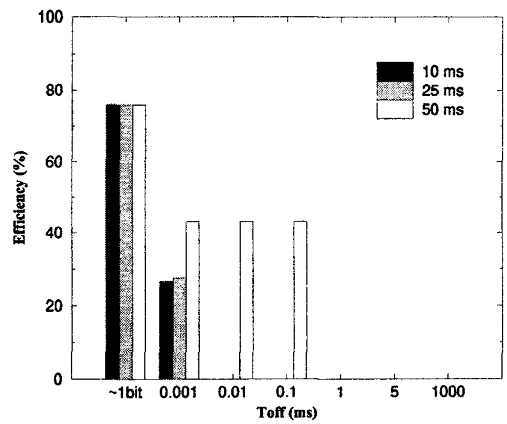

Figure 7: Efficiency of the FEC approach versus T_{off} and versus the lifetime.

V. CONCLUSIONS

In this paper we have considered a MAC protocol for Wireless ATM systems based on ARQ. The capacity of the proposed scheme has been measured and the design parameters have been optimized according to the channel and connections characteristics. In particular, the optimum frame length has been evaluated in varius scenarios adopting the Gilbert channel model. Moreover, preliminary results on the comparison of the overall protocol efficiency achievable with the ARQ approach and the classical FEC plus interlaving one have been presented. The results show that, within the range of values commonly adopted for mobile-radio channel, the number of connections served with the ARQ approach is higher than that obtained by FEC. FEC scheme shows a high efficiency only when errors are almost incorrelated. Further studies are in progress to consider more accurate channel models and to take into account the signaling traffic with variable rate connections.

ACKNOWLEGMENTS

The authors wish to thank graduate students A. Bacchi and E. Monticelli for their fundamental help in obtaining simulation results.

REFERENCES

(1) IEEE Communications Magazine, Special Issue "Introduction to Mobile and Wireless ATM, November 1997,

(2) O. Kubbar, H.T. Mouftah, "Multiple Access Control Protocols for Wireless ATM: Problems Definition and Design Objectives", IEEE Communications Magazine, vol. 35, no. 11, Nov. 1997., pp. 93-99.

(3) J. Sanchez, R. Martinez, M.W. Marcellin, "A Survey of MAC Protocols Proposed for Wireless ATM", IEEE Network, Nov-Dic 1997, pp. 52-62 .

(4) M. Zorzi, R.R. Rao, L.B. Milstein, "On the Accurancy of a First-order Markov Model for Data Block Transmission on Fading Channels", IEEE ICUPC '95, pp. 211-215, Nov. 1995.

(5) E.N. Gilbert, "Capacity of a Burst-Noise Channel", Bell System Tech. Journal, vol. 39, pp. 1253-1266, Sept. 1960.

(6) ETSI BRAN Home page, http://www.etsi.fr/bran/

(7) ATM Forum, WATM Home Page, http://www.atmforum.com/

(8) G.D. Forney, "Burst-Correcting Codes for the Classic Bursty Channel", IEEE Trans. on Communications, vol. 19, no. 5, Oct. 1972, pp. 772-781.

(9) N. Passas, S. Pascalis, D. Vali, L. Meracos, "Quality-of-Service-Oriented Medium Access Control for Wireless ATM Networks", IEEE Communications Magazine, vol. 35, no. 11, Nov. 1997., pp. 42-50.

(10) F. Borgonovo, A. Capone, L. Fratta, "ARQ vs FEC in mobile radio systems", Technical Report, Politecnico di Milano 1998.

(11) M. Zorzi, "Performance of FEC and ARQ Error Control in Bursty Channels under Delay Constraints", IEEE VTC 98, Ottawa (CA), May 1998.