Performance Evaluation of the PRMA Protocol for Voice and Data Transmissions in Low Earth Orbit Mobile Communication Systems

E. Del Re, R. Fantacci, G. Giambene, T. Pecorella

Dipartimento di Ingegneria Elettronica, Università degli Studi di Firenze
Via S. Marta, 3 - 50139 Firenze, ITALY
e-mail: delre@lenst.die.unifi.it

Abstract: The Universal Mobile Telecommunications System (UMTS) will provide multimedia services to users in whatever environment via uniform service access procedures. Efficient Medium Access Control (MAC) protocols are needed to guarantee high capacity. This paper deals with Packet Reservation Multiple Access (PRMA), an efficient MAC protocol that has been widely studied in terrestrial microcellular systems. We investigate the suitability of PRMA for supporting voice and data transmissions in a UMTS based on Low Earth Orbit - Mobile Satellite Systems (LEO-MSSs). It is shown here that, through a suitable choice of the system parameter values, the PRMA protocol may achieve a good quality of service also in LEO-MSSs.

Introduction:

The standardization of the Universal Mobile Telecommunications System (UMTS) will be completed by the European Telecommunications Standards Institute (ETSI) within 1998. UMTS should provide a wide range of services to mobile and stationary users. At present, some of these services are provided by various existing systems (e.g., cordless, cellular, satellite, etc.). However, UMTS aims to integrate in a unique system several networks and many services [1], [2]. ETSI describes UMTS as providing “personalized globally accessible high quality mobile communication services” meeting such objectives multimedia capability, uniform access through different interoperating networks, low cost and high frequency spectrum efficiency [3].

Present mobile communication systems provide telephony and a small set of data services. For instance, the Global System for Mobile communications (GSM), which has now a widespread success all over the world, allows data transmissions at low bit-rates (typically, 9.6 kbit/s) and the short message service (maximum 160 bytes). Of course, these capabilities are insufficient for multimedia applications envisaged by UMTS, where a bit-rate up to 2 Mbit/s is required.

Moreover, also presently implemented Mobile Satellite Systems (MSSs) allow low bit-rate data applications (e.g., a data bit-rate of 4.2 kbit/s is allowed by the IRIIDIUM system [4], whereas a data bit-rate up to 9.6 kbit/s is provided by the GLOBALSTAR system [5]). Only the TELEDESIC system [6], that is expected to start its operations by 2002, will provide data applications from 16 kbit/s to 2 Mbit/s (where the higher bit-rates will practically require fixed terminals) and will support the Asynchronous Transfer Mode (ATM).
As a matter of fact, UMTS needs to respond to the demand of consumers who want to combine mobility with multimedia applications (e.g., videoconference, remote learning, database access, multimedia mail, interactive transmissions, Internet access, telephony, etc.), each of them with specific Quality of Service (QoS) requirements. Within UMTS, the satellite component and the terrestrial one will provide the users with the same services and, possibly, with the same QoS of the terrestrial fixed network that will be ATM-based B-ISDN. ATM considers five traffic classes [7]:

- **Constant Bit-Rate (CBR) class**: delay sensitive, fixed bit-rate (e.g., voice, video),
- **Variable Bit-Rate, Real Time (VBR-RT) class**: delay sensitive, variable bit-rate (e.g., voice with silence detection, variable rate compressed video);
- **Variable Bit-Rate, Non-Real Time (VBR-NRT) class**: delay insensitive, variable bit-rate (e.g., data);
- **Available Bit-Rate (ABR) class**: delay insensitive, variable bit-rate (e.g., data);
- **Unspecified Bit-Rate (UBR) class**: no specific QoS is guaranteed for this traffic (e.g., Internet transactions).

It is well known that, MSSs will offer a stimulus to the take up of mobile communications by allowing the delivery of advanced communications services also in parts of the world where terrestrial networks are underdeveloped. Future broadband-MSSs within UMTS will employ packet switching techniques compatible with ATM used on terrestrial fixed networks (B-ISDN) and will support mainly VBR-RT, VBR-NRT and ABR traffics [7]. It will be important to use harmonized access procedures that take into account the different QoS of each traffic type.

It is anticipated that the satellite segment of future UMTS will be (partly or totally) based on non-geostationary satellites; in particular, Low Earth Orbit - Mobile Satellite Systems (LEO-MSSs) [8] represent an attracting solution, since they are characterized by low propagation delays and low propagation attenuations that allow the use of low-power hand-held terminals.

Efficient Medium Access Control (MAC) protocols have to be used, since the satellite channel is both power and bandwidth limited. This paper proposes Packet Reservation Multiple Access (PRMA) as an efficient MAC protocol to be used in LEO-MSSs.

The PRMA scheme is an improvement of Time Division Multiple Access (TDMA) that combines TDMA with Slotted-ALOHA [9]. A TDMA channel consists of a slot in subsequent frames. Each time slot can carry a packet belonging to a User Terminals (UTs). With PRMA, the assignment of time slots to UTs is not fixed, but it is dynamically handled on the basis of the presently active UTs. Each packet has a header which contains synchronization data, control data and the identification of both the source UT and the destination one.

PRMA was initially proposed for terrestrial microcellular systems characterized by very low Round Trip Delays (RTD) ¹ with respect to the slot duration [10], [11], [12], [13]. PRMA has shown very interesting features in supporting both voice and data traffics [14], [15]. Moreover, several variants of the PRMA protocol have been proposed [16], [17], [18], [19].

The efficiency of the PRMA approach in managing voice sources relies on the use of voice activity detection: during a call, a voice source transmits a packet in a slot per frame only when it is actually active (i.e., during a talkspurt). When there is a silent phase, this slot can be destined to another active source (i.e., both a voice and a data source). In a multimedia scenario, PRMA can exploit the bit-rate variability of both voice UTs and data UTs to obtain a statistical multiplexing of several sources on the same carrier. High multiplexing gain values can be obtained owing to the burst behavior of voice and data sources.

---

¹ As a matter of fact, RTD is the time for the propagation of signals from a UT on the earth to the satellite and back to the UT. Therefore, RTD is the time a UT must wait before knowing the outcome of its transmission attempt.
In addition to this, PRMA is well suited for supporting ATM [20]. As a matter of fact, according to the ATM virtual circuit concept, there is no more a one-to-one association between time slots within a frame and transmissions as in TDMA. Packets are identified by the Virtual Channel Identifier (VCI) contained in the header. The VCI field indicates the destination UTs. Different UTs can recognize their packets broadcast on the downlink by checking the VCI field.

The access to a slot by PRMA is of the Slotted-ALOHA type: as soon as a talkspurt or a data message starts, the related UT attempts to transmit its packet in the first available slot on which it obtains the permission to transmit (the access to available slots is managed according to a permission probability which is used to randomize the attempts of different UTs needing to acquire a reservation). If no collision occurs, the UT attains a reservation for this slot in subsequent frames. Slots are marked available or busy by a feedback channel, broadcast by the controller (in this paper, the satellite). The feedback channel also inform the UTs about the results of their transmission attempts.

The main obstacle for applying the PRMA protocol to MSSs is the high RTD value. Some initial studies have highlighted that the RTD value in geostationary satellite systems (i.e., minimum 250 ms) prevents from any application of PRMA [21]. Conversely, this work shows the suitability of the PRMA protocol for supplying both voice and data traffics in LEO-MSSs, where the RTD values range from 10 to 30 ms. As regards voice traffic we consider a VBR-RT traffic, while for data traffic we considered two models with different characteristics. In particular, the first model assumes a Poisson traffic suitable for bulk data transfers; the second is a Modulated Markov Process (MMP) that has been used as a rough model for Internet-like traffic.

In the first case we consider a mobile phone with short data messages (GSM-like), while the second case we have assumed a palmtop terminal with Internet-like services.

We will prove that the PRMA protocol maintains in LEO-MSSs all the advantages of the classical terrestrial PRMA scheme. This is an interesting aspect in the light of UMTS where terrestrial cellular systems and MSSs will be part of a unique system with unified/compatible MAC procedures.

Two solutions have been considered for the cellular coverage of LEO-MSSs [22]: satellite-fixed cells (e.g., adopted by the IRIDIUM system [4]), where cells are fixed with respect to the satellites and fast move as regards the earth; earth-fixed cells (e.g., used by the TELEDESIC system [6]), where cells are fixed on the earth and satellite antennas are steered to point to the same area on the earth as long as possible.

This paper assumes an earth-fixed cell system. Therefore, aspects related to UT mobility are not considered in this work: we neglect UT cell changes during both the call lifetime and messages transmission time, owing to the large cell sizes obtained by satellite antennas on the earth. However, even if a satellite-fixed cell system would be assumed, the mobility management should not pose significant problems with PRMA. Let us refer to a voice UT with a call in progress: as soon as the UT goes into an adjacent cell a handoff procedure is started; this situation can be considered as a UT starting a talkspurt in the destination cell. Hence, no forced call termination is experienced by the UT, but, eventually, only a temporary speech quality degradation due to packet dropping, if the delay to obtain a reservation in the destination cell exceeds the maximum tolerable value.

This work is organized as follows: the following Section 2 describes the PRMA protocol to support voice and data applications in LEO-MSSs; Section 3 presents a study for the selection of optimum system parameter values. Finally, a performance comparison between the terrestrial PRMA protocol and its version for LEO-MSSs is carried out in Section 4.
The PRMA protocol in LEO-MSSs: voice and data applications:

We consider a given cell of the MSS, where all the UTs share a given carrier to transmit their information packets to the related satellite. Downlink channel traffic (from the satellite to the UTs on the earth) and uplink channel traffic are simultaneously transmitted by Frequency Division Duplex (FDD). The downlink transmission is a time division multiplex broadcast. The satellite with on board processing schedules the downlink traffic in order to avoid any contention. Therefore, we focus our attention on the PRMA protocol used on uplink of the LEO-MSS.

We consider voice UTs and data UTs: they are part of two sub-systems with different characteristics for the packet generation process and the required QoS. In the light of the future realization of multi-application terminals, we could consider that a voice UT and a data UT are physically integrated in the same multimedia UT. However, for the sake of generality, we will distinguish between data UTs and voice UTs, taking into account that some data UTs may be implemented together with voice UTs.

When a voice or a data UT needs to transmit its packets, it enters the contending state (CON) [10], [11] to obtain the reservation of a slot per frame: as soon as there is an available slot and the UT has the permission, it transmits the first packet. Different permission probabilities have been considered for different traffics in order to take into account their specific QoS requirements. We denote by $p_v$ the permission probability for voice UTs and by $p_d$ the permission probability for data UTs.

When a contending UT is successful, that is the header of the packet sent on a given slot has been successfully decoded at the satellite, the slot is assigned to the UT in each subsequent frame. The reservation mechanism is made possible by a feedback channel broadcast by the satellite (with on board processing capabilities) which informs all the UTs within a cell about the state of each slot of the PRMA carrier (i.e., idle/reserved slot and, for the reserved slots, the UTs which are currently using them).

If two or more UTs have attempted to send their packets in the same slot: there is a collision. If we neglect the capture effect [9], the satellite can not recognize any UT, so it leaves the slot unreserved. Therefore, the involved UTs re-schedule randomly their transmission attempts on free slots, according to the contention procedure.

Note that we consider that a multimedia UT (i.e., a voice UT integrated with a data UT) needs two distinct reservations when it has contemporaneously voice and data packets to be transmitted.

In a LEO-MSS, the RTD depends on the satellite constellation altitude, the minimum elevation angle envisaged for UT and the relative position between UT and satellite. Moreover, the RTD usually varies during the service of a talkspurp/message due to the satellite constellation dynamics. Hence, we have considered that RTD is always equal to its maximum value for a given constellation, $RTD_{max}$ (conservative assumption), which corresponds to the satellite seen by the UT according to the minimum elevation angle.

Let $T_f$ denote the frame duration and $T_c$ denote the slot duration. We have selected $T_f = RTD_{max} + \varepsilon T^2$. This means that a UT is informed about the outcome of a transmission attempt at the beginning of the homologous slot in the subsequent frame. We have assumed that a UT can perform only one access attempt per frame. If the attempt has been successful the UT can send the next packet on this slot. Otherwise, if the attempt has been unsuccessful, the UT goes back in the CON state and the reservation procedure restarts.

We describe below the characteristics of both voice UTs and data UTs and their related sub-systems.

---

$^2$ Parameter $\varepsilon$ denotes the time needed to transmit only the header of a packet. The satellite recognizes the UT requesting a reservation on the basis of the header of the packet received; as soon as the header is received, the satellite broadcasts a suitable acknowledgement on the feedback channel. Therefore, the UT requesting a reservation will know the outcome of its attempt on a given slot before the homologous slot in the subsequent frame, if $T_f = RTD_{max} + \varepsilon$. For practical considerations, $\varepsilon$ can be neglected with respect to $RTD_{max}$ and we may consider $T_f = RTD_{max}$. In general, we could consider $T_f = n RTD_{max} + \varepsilon$, with $n \geq 1$, but this study is beyond the scope of this work.
**Voice sub-system**

The frame structure is designed so that a speech source exactly generates one packet per frame. A time slot carries the bits produced by a voice source in $T_f$ plus a header of $H_v$ bits. As discussed above, $T_f$ is considered (approximately) equal to $RTD_{max}$; moreover, the source bit-rate, $R_s$, depends on the voice codec and the channel bit-rate, $R_c$, is another important system parameter which depends on the transmission technique and the available bandwidth. According to this, we have that the number of slots per frame, $N_s$, and the slot duration, $T_s$, are obtained as follows:

$$N_s = \left\lfloor \frac{R_s T_f}{R_s T_f + H_v} \right\rfloor \text{ slots per frame}$$

$$T_s = \frac{T_f}{N_s} \text{ [ms]}$$

(1)

where $\left\lfloor x \right\rfloor$ denotes the biggest number less than or equal to $x$.

We have assumed a slow speech activity detector [23] which reveals only principal gaps within the conversation. A fast speech activity detector [23] can not attain a good performance in an MSS, since it entails a greater contention rate, which has to be avoided in the presence of high RTD values (there is the risk of CON state congestion). Owing to the slow speech activity detector, a conversation is divided in talking phases and silent phases; the times spent in these states have been modeled by exponential distributions with expected values 1 s and 1.35 s, respectively for the talking phase and the silent one.

A voice UT in the CON state discards the first packet from the buffer if the time to obtain a reservation exceeds a maximum value, $D_{max}$. Hence, as in [14], [23] we have assumed $D_{max}$ equal to 32 ms. When a packet is discarded, the voice UT tries to obtain a reservation with the next packet. The quality of the voice transmission with PRMA is measured by the probability $P_{drop}$ that a packet is dropped from the buffer of a UT, because it has waited for a time longer than $D_{max}$. Obviously, the greater the number of UTs with a call in progress that share a PRMA carrier, $N_v$, the greater the average number of UTs in the CON state, the greater $P_{drop}$. With contemporary speech codecs, it is required $P_{drop} \leq 1\%$, in order to cause a minimal degradation in the perceivable speech quality [14]. Therefore, let $N_{max}$ denote the maximum number of voice UTs with a call in progress that may share a PRMA carrier with $P_{drop} \leq 1\%$; $N_{max}$ depends on the parameters of both the voice sub-system and the data one.

The throughput of the voice sub-system, $\eta_v$, (i.e., the average number of packets carried out per slot) is given by [23]:

$$\eta_v = \frac{N_v \psi (1 - P_{drop})}{N_s} \left\lfloor \frac{\text{pkts}}{\text{slot}} \right\rfloor$$

(2)

where $\psi$ is the voice activity factor obtained as $1/(1+1.35) = 0.425$.

**Data sub-system**

Each type of data traffic is characterized by specific characteristics and QoS requirements. The advantage of using LEO satellites instead of geostationary ones is that they are more close to the earth, so allowing a lower end-to-end delay. A typical value for the maximum cell transfer end-to-end delay is 400 ms for voice applications (VBR-RT) in terrestrial networks [25]. However, in this work we have assumed an ABR-like data traffic. This type of traffic is suitable for modeling e-mail messages and data file transfer. It could also include the signaling traffic associated with the
management of voice communications. Note that the PRMA protocol for the data sub-system is equivalent to a Reservation-ALOHA [11].

We consider \( N_d \) data UTs which produce messages according to independent Poisson processes with mean rates \( \lambda \). Every message has a random length and it is segmented in packets each of them has a header of \( H_p \) bits. We denote by \( L_o \) the average length in bits of a message (only information part); the corresponding average number of packets \( L_s \) can be obtained as described below, by assuming the same header size for voice and data packets:

\[
L_s = \frac{L_o}{R_s T_f} \left\lfloor \frac{\text{pkts}}{\text{msg}} \right\rfloor \tag{3}
\]

We have considered that the number of packets of a message has a geometric distribution with expected value \( L_s \). The average input data traffic on a PRMA carrier due to all data UTs can be expressed as follows:

\[
r_d = \lambda T_s L_s N_d \left\lfloor \frac{\text{pkts}}{\text{slot}} \right\rfloor \tag{4}
\]

This traffic type, usually assumed as a rough model for data traffic, does not effectually describe internet-like traffic, where the communication is composed by a high variable burst of messages, each message being of a random length.

A good approximation of this traffic type can be obtained with a Modulated Markov Process (MMP) generator, which has two message sub-generators with rates \( \lambda_A \) and \( \lambda_B \); each sub-generator produces messages of a variable number of packets, geometrically distributed with means \( L_A \) and \( L_B \).

The two sub-generator are alternatively activated according to a two-state Markov chain with average permanence times \( T_A \) and \( T_B \).³

By properly choosing the right values for the 6 parameters of this traffic model, the generator can reproduce the effects of a highly bursty communication traffic.

We have assumed low values for both the mean traffic rate and the average message length. This assumption is based on two considerations:

- the satellite segment within UMTS will offer only reduced bandwidth services;
- the data traffic of an Internet connection on the uplink channel is usually very low (i.e., the Internet traffic is asymmetric).

We have assumed that when a data UT needs to transmit a message it must obtain the reservation of a slot in subsequent frames for exclusive transmissions as for the voice packet transmissions by active voice UTs. The reservation is maintained by a data UT until it transmits all messages in the buffer according to an exhaustive discipline [26].

The parameter which measures the performance of the data sub-system is the average message transmission time, \( T_{msg} \), that is the mean time from the message arrival to the buffer of a UT to the instant when all packets of the message have been transmitted.

³ Several models for Internet traffic has been proposed [27, 28]. Our model is a simplified one, but it takes in account the bursty characteristics of this type of traffic.
Under the assumption of a stable data sub-system, the average number of packets in the buffers of each UT is finite. Hence, the data throughput, $\eta_d$, is equal to the input data traffic, $r_{d^i}$, expressed in packets/slot. The total throughput considering both voice and data traffic is $\eta = \eta_v + \eta_d$.

**Simulation results**

A simulation tool has been developed which reproduces the behavior of $N_v$ voice UTs and $N_d$ data UTs which share the access to a PRMA carrier with a round trip delay $RTD_{\text{max}}$. As in [23], we assume the system parameter values given in Table 1.

**Table 1:** System parameters values used in this work.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_c$</td>
<td>channel bit rate</td>
<td>720 kbit/s</td>
</tr>
<tr>
<td>$R_s$</td>
<td>speech source bit rate</td>
<td>32 kbit/s</td>
</tr>
<tr>
<td>$H_v = H_d$</td>
<td>header size of a packet</td>
<td>64 bit</td>
</tr>
</tbody>
</table>

For instance, according to (1) and on the basis of the values shown in Table 1, we obtain $N_v = 20$ slots/frame and $T_f = 0.8$ ms for $T_f = 16$ ms. Moreover, each packet has 512 information bits and 64 header bits (this packet size could be compatible with an extended ATM cell used for mobile systems [20]).

Fig. 1 shows simulation results concerning the parameters $P_{\text{drop}}$ and $T_{\text{msg}}$ as a function of the input data traffic for: $T_f = RTD_{\text{max}} = 16$ ms, $N_v = N_d = 15$ stations/carryer, $p_v = p_d = 0.5$ (i.e., the same permission probabilities for voice and data) and $L_a = 25$ kbit/message (according to (3), $L_a = 48.8$ packets/message). It is evident from Fig. 1 that the maximum input data traffic, $r_{d^i}$, that can be supported by fulfilling $P_{\text{drop}} \leq 1\%$ is 0.47 packets/slot; correspondingly, the voice sub-system throughput is 0.31 packets/slot and $T_{\text{msg}}$ is about 2700 slots (i.e., 2.16 s which correspond to $T_{\text{serv}} = 0.06$ s and $T_{\text{delay}} = 2.1$ s; note that $RTD_{\text{max}} = 0.016$ s). In order to improve this performance, a suitable selection of system parameters values (i.e., $p_v$, $p_d$ and $T_f$) is required, as highlighted in the next Section.

**Selection of system parameters**

The criterion adopted to select suitable value for both $p_v$, $p_d$ and $T_f$ is that of minimizing $P_{\text{drop}}$, without significantly increasing $T_{\text{msg}}$. Fig. 2 shows $P_{\text{drop}}$ as a function of both $p_v$ and $p_d$. The numerical values have been obtained by computer simulation by assuming: $T_f = RTD_{\text{max}} = 16$ ms, $N_v = N_d = 15$ stations/carryer, $r_{d^i} = 0.4$ packets/slot and $L_a = 25$ kbit/message. In Fig. 2 we note that:

- $P_{\text{drop}}$ mainly depends on $p_v$; e.g., if $p_v$ is too low, $P_{\text{drop}}$ increases owing to the growth of the access delay to obtain a reservation. Likewise, if $p_v$ is close to 1, $P_{\text{drop}}$ increases, because collisions are more frequent among the attempts of UTs which need to transmit their talkspurts. In both these extreme cases, the CON state is congested.

- $P_{\text{drop}}$ increases with $p_d$; i.e., if $p_d$ increases more data UTs attempt their transmissions on available slots, so hindering the reservation of voice UTs. However, it must be avoided an excessive reduction of $p_d$ in order to prevent high values of $T_{\text{msg}}$. 
From Fig. 2 it appears that $p_v = 0.6$ is the best choice. Moreover, we have verified that $T_{msg}$ is practically independent of $p_v$ and $p_d$, provided that $p_d \geq 0.2$; whereas, if $p_d < 0.2$, $T_{msg}$ suddenly increases when $p_d$ decreases. Therefore, $p_d = 0.2$ is a good choice that allows low values for both $T_{msg}$ and $P_{drop}$.

We have obtained the same optimized values for both $p_v$ and $p_d$ also for different values of $T_f = RTD_{max}, r_d$ and $L_b$.

Let us focus on the selection of the optimum value of $T_f$. The frame duration is a very important parameter which determines the number of slots per frame, the average number of packets per talkspurt and the average number of packets per message. Moreover, according to our assumption, $T_f$ corresponds to the $RTD_{max}$ of the satellite system.

By assuming fixed values for $R_c, R_s, H_c = H_d, L_b$, if $T_f$ increases, the average number of packets per talkspurt decreases as well as the average number of packets per message; whereas, according to (1), $N_s$ increases up to the maximum value $\left\lfloor \frac{R_c}{R_s} \right\rfloor$, that is equal to 22 slots/frame on the basis of the values shown in Table 1.

Parameter $T_{msg}$ has a slight dependence on $T_f = RTD_{max}$ in the LEO range: a high $T_f$ value allows a slight increase of $N_s$ that, in turn, increases the transmission attempt rate. However, a variation of $T_f$ has a greater impact on $P_{drop}$:

- An excessive value of $T_f$ causes a high $P_{drop}$ value due to the maximum acceptable delay $D_{max}$: a reduced number of attempts is available for the contending voice UTs in order to obtain a reservation; many packets experience an access delay greater than $D_{max}$ without obtaining a reservation.

- With a low value of $T_f$, there are few slots per frame which may be contented among the active UTs. This leads to an increase of the congestion for the CON state and, hence, to high values for $P_{drop}$.

The optimum value of $T_f$ is that which allows the maximum capacity of voice UTs for a PRMA carrier (i.e., the maximum value of $N_{(max)}$). In this investigation we assume the optimum permission probabilities values obtained at the previous step. Fig. 3 presents the behavior of $N_{(max)}$ as a function of $T_f$ by assuming: $p_v = 0.6, p_d = 0.2, N_d = 15$ data stations/carrier, $r_d = 0.4$ packets/slot and $L_b = 25$ kbit/message. Hence, $N_{(max)}$ has a maximum value equal to 18 voice stations/carrier for $T_f$ ranging from 16 to 19 ms; we have selected $T_f = 16$ ms in order to be consistent with [23], where terrestrial microcellular systems are considered. This is an interesting result on the light of the expected integration between terrestrial and satellite systems: the optimization of $T_f$ is quite insensitive to the RTD value, on the condition that RTD be sufficiently low as in LEO-MSSs. Finally, from Fig. 3 we note that Medium Earth Orbit (MEO) satellites, which are at higher altitudes than LEO ones and entail RTD values greater than 70 ms, can not guarantee an acceptable PRMA performance.

Let us focus on the multiplexing gain achievable for the voice sub-system by PRMA. $\mu$: the multiplexing gain is the ratio between the maximum number of voice UTs supported per PRMA carrier, $N_{(max)}$, and the equivalent number of TDMA channels devoted to voice UTs (an ideal TDMA without packet overhead is considered). The number of TDMA slots is given by $\left\lfloor \frac{R_c}{R_s} \right\rfloor$, but only a part of these slots will be used by voice UTs in the equivalent ideal TDMA for the voice sub-system: $\lambda L_b N_d \text{bit/s}$ of the total channel capacity $R_c$ are (on average) destined to data transmissions. Therefore, the number of equivalent TDMA channels is $\left\lfloor \left( R_c - \lambda L_b N_d \right) / R_s \right\rfloor$. In conclusion, the multiplexing gain $\mu$ is obtained as follows:
Parameter $\mu_v$ expresses the capability of the MAC protocol to exploit efficiently the talkspurt and silent phases of voice sources. Only when $\mu_v > 1$, PRMA is advantageous as regards TDMA in managing voice UTs. Fig. 4 shows the behavior of $\mu_v$ as a function of $r_d$ for $L_b = 25$ kbit/message, $N_d = 20$ data stations/carrier and the selected optimum parameters (i.e., $p_c = 0.6$, $p_d = 0.2$, $T_f = 16$ ms). We can note that $\mu_v$ decreases as $r_d$ increases. This behavior can be justified as follows: $\mu_v$ depends on the number of slots devoted per frame to support voice transmissions; when $r_d$ increases we have a reduced number of slots for voice transmissions and, therefore, a low $\mu_v$ value. For $r_d \leq 0.6$ data packets/slot, we have that $\mu_v \geq 1$. In these conditions, PRMA allows a more efficient management of voice sources than TDMA. Moreover, when $r_d = 0$ (i.e., no data traffic) we obtain $\mu_v = 1.6$ voice conversations/voice channel, that is the same $\mu_v$ value as in terrestrial microcellular systems with only voice sources and optimized parameter values [23]. This result highlights that the high RTD value in LEO-MSSs as regards that experienced in microcellular systems does not change the multiplexing capabilities of the PRMA protocol.

Since a given input data traffic $r_d$ can be obtained either with short messages and a high arrival rate or with long messages and a low arrival rate, we study in Fig. 5 the impact of different messages length $L_b$ for $r_d = 0.4$ data packets/slot, $N_c = N_d = 15$ stations/carrier and optimized system parameter values (i.e., $p_c = 0.6$, $p_d = 0.2$, $T_f = 16$ ms). In the graphs of Fig. 5 we note that $P_{\text{drop}}$ has a flat behavior for $L_b \in [5, 40]$ kbit/message; moreover, $P_{\text{drop}}$ shows two opposite trends for extreme values of $L_b$:

- **When $L_b$ is low, $P_{\text{drop}}$ increases, because we have a high arrival rate of messages.** Since messages are short, it is likely that the service of a message ends before a new message needs to be sent by the same data UT (probably, at each message arrival a new reservation must be obtained according to the contention procedure). Hence, the contention rate is high and causes a worse performance for the delay sensitive voice sub-system.

- **If $L_b$ is high, $P_{\text{drop}}$ increases, because the transmission of the message of a UT requires a long time; during this interval, new messages may arrive at the buffer of the UT and they are managed according to the exhaustive policy: the slot reservation at the end of the present message is not released, because other messages wait to be transmitted in the buffer of the UT.** Therefore, the number of slots available for UTs with new talkspurts is practically reduced and $P_{\text{drop}}$ increases.

Through other simulation runs not shown in this work, we have verified that the impact of these two trends depend on the value of the input data traffic $r_d$:

- **If $r_d$ is low, the first trend prevails: $P_{\text{drop}}$ decreases if $L_b$ increases.** When $L_b$ is high $P_{\text{drop}}$ decreases, because the message arrival rate is not sufficiently high to have the arrival of messages during the transmission time of a given message.

- **Whereas, if $r_d$ is very high, the second trend prevails: $P_{\text{drop}}$ increases with $L_b$.** Even if $L_b$ is low, $P_{\text{drop}}$ increases with $L_b$, because when a data UT completes the transmission of a message, it is quite likely that another message is ready to be transmitted in its buffer.

- **For intermediate values of $r_d$ the behavior is like that shown in Fig. 5.**
To test the influence of a internet-like traffic type on the system performance, we chose a slightly different configuration. Figures 7-9 refer to a 20 voice terminal situation and a variable number of data terminals each offering a MMP traffic, as outlined before; generator parameters are in Table 2.

Table 2: MMP generator parameters.

<table>
<thead>
<tr>
<th></th>
<th>Permanence time in the A state</th>
<th>6000 slots</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_A$</td>
<td>message interarrival time in the A state</td>
<td>1200 slots</td>
</tr>
<tr>
<td>$I_A$</td>
<td>mean packet per message in the A state</td>
<td>50 pkt</td>
</tr>
<tr>
<td>$T_B$</td>
<td>Permanence time in the B state</td>
<td>15000 slots</td>
</tr>
<tr>
<td>$I_B$</td>
<td>message interarrival time in the B state</td>
<td>173.68 slots</td>
</tr>
<tr>
<td>$L_B$</td>
<td>mean packet per message in the B state</td>
<td>4 pkt</td>
</tr>
</tbody>
</table>

This assumption (a variable number of terminals each offering a fixed traffic type) is derived from the peculiarity of the internet traffic type.

We must stress that, at the present, it is impossible to know the exact amount of traffic given by a single data terminal accessing internet over the UMTS satellite layer, since we don’t know either which type of services will be used, or if these services will be similar to those used at present in Internet. Remember the change of the Internet traffic shape since the introduction of the world wide web service; this service, which now is responsible of a lot of Internet traffic, shifted the Internet traffic type from mainly a bulk one constituted by large messages to an interactive one constituted by a large number of small messages grouped in large busts.

From figure 7 however, we can highlight that the effect on voice system performance of this traffic type is not different from a Poisson traffic type; the only real difference is due to the combined effects of a short and a long message length, which, as outlined before, differently affects the performance of the voice sub-system.

This combined effects may cause a performance degradation worse than the one estimated by using a Poisson-like distribution with similar parameters derived by the observation of mean message length and data traffic rate.

In Fig. 8 we can see the obvious dependance of the $T_{pk}$ (the packet transmission time) from the channel occupation, but we can observe the delay change is small due to the high number of packets that makes up a message. This gives raise to an high message transmission time, which affects the $T_{pk}$.

The most interesting observation comes from Fig. 9, which shows an histogram of the packet transmission time. We must stress that the waiting time distribution shows that a considerable number of packets suffers from high transmission time and so the variance of the message delay (here not shown) is very high.

This problem must be carefully considered especially in building TCP/IP implementations for the terminals and the base stations, since the high transmission delay can affect the performances of TCP retransmission scheme.

Before concluding this Section, we introduce some qualitative considerations about system stability. The voice sub-system has no problem of stability, but it may experience a low throughput when the CON state (which is shared with the data-subsystem) is congested: voice packets queuing in the buffers of UTs is limited by the dropping mechanism; when the access delay exceeds $D_{max}$, voice packets are dropped. In a congested situation $P_{drop}$ is very high. The limit situation before congestion
is when the throughput of the voice sub-system has a maximum. Whereas, the data sub-system may have problems of unstability: if the input data traffic becomes too high, the throughput of the data sub-system may become insufficient. We have unstability when the mean number of data packets in a buffer of a data UT tends to infinity (i.e., it is unbounded as the simulation time increases).

The CON state is common among voice UTs and data UTs, since both types of UTs share the access to the same PRMA carrier and with the same procedures. Therefore, it is expected that the congestion of the voice sub-system occurs before the unstability of the data sub-system. Parameters primarily affecting system congestion are: $p_v$, $p_d$, $r_d$, $N_v$, and $N_d$. Of course, values of $p_v$, $p_d$ and $r_d$ close to 1 causes congestion. An in-depth study on system stability/congestion is beyond the scope of this paper which aims to demonstrate the suitability of the PRMA protocol for supporting voice and data applications in LEO-MSSs. However, we have checked by simulation that all the results shown in this work are related to an uncongested system behavior.

The performance of the PRMA protocol in terrestrial microcellular networks and LEO-MSSs

Let us consider a terrestrial microcellular system, where RTD can be neglected with respect to the slot duration (e.g., RTD = 1 μs for a cell with 1 km radius): a UT attempting a transmission on a free slot immediately knows the outcome of its attempt from the base station of its cell. Therefore, if the attempt has been unsuccessful the UT can reschedule the transmission even from the next slot (provided that this slot is available and the UT attains the permission to transmit). Whereas, in a satellite system the UT will know the outcome of its attempt only after an RTD time. Therefore, RTD limits the number of attempts that a voice UT can carry out in the CON state before dropping a packet.

This Section compares a terrestrial system and a satellite one in order to evaluate the impact of RTD on the performance of the PRMA protocol in terms of both $P_{drop}$ and $T_{msg}$. In particular, we have selected in both cases the parameter values shown in Table 1, $T_j = 10$ ms (then, $N_v = 18$ slots/frame), $N_v = 8$ voice stations/carryer, $N_d = 10$ data stations/carryer, $p_v = 0.35$, $p_d = 0.15$, $L_v = 10$ packets/message. Fig. 6 presents the behavior of $P_{drop}$ and $T_{msg}$ as a function of the input data traffic, $r_d$. Provided that $P_{drop} \leq 1\%$, the differences between the two situations are small: the maximum $r_d$ value is 0.3 packets/slot in the terrestrial case, and 0.25 packets/slot in the LEO satellite case. Parameter $T_{msg}$ is in both cases on the order of hundreds of slots with a difference of few tens of slots between the terrestrial case and the satellite one; this difference is due to the fact that in the satellite case a data UT stops contending after a transmission attempt to wait for the response from the satellite that will be received after an RTD time.

From the obtained results it is possible to state that the PRMA protocol manages voice and data traffics in LEO-MSSs with a QoS very close to that obtained in terrestrial microcellular systems.

Conclusions

LEO-MSSs will provide (partly or totally) the satellite component of future UMTS. This system will have a global coverage and will provide the users with multimedia services. The terrestrial component and the satellite one will use as far as possible the same protocols, and will be part of an integrated system.

This work is concerned with MAC protocols for uplink access to the satellite. An attracting solution which was initially proposed for terrestrial microcellular systems is the PRMA protocol, which exploits silent periods of a voice source to multiplex more conversations on the same channel. Moreover, PRMA may support voice and data traffics.
The results shown in this paper have proved that PRMA is well suited for application in LEO-MSSs, where it maintains all the interesting features described for terrestrial systems. The following bearer services have been envisaged: a VBR-RT voice and an ABR-like data transfer. The related performance parameters are: the packet dropping probability for voice transmissions and the average transmission time for data messages. A system parameter optimization has been carried out which has allowed the selection of optimum values for both permission probabilities and frame duration. Moreover, the impact of the input data traffic and the message length on both the packet dropping probability and the average message transmission time has been investigated.

We have shown that PRMA is advantageous with respect to TDMA in managing integrated voice and data traffics for a wide range of input data traffic. In LEO-MSSs, PRMA attains a performance close to that of terrestrial microcellular systems. Therefore, we may conclude that PRMA is a good candidate as a MAC protocol for future UMTS, since it is well suited not only for the terrestrial component, but also for the LEO satellite one.

REFERENCES


Figures

![Graph 1](image1.png)

**Fig. 1** Behavior for both $P_{\text{drop}}$ and $T_{\text{msg}}$ as a function of $r_d$ in a LEO-MSS with $T_f = RTD_{\text{max}} = 16$ ms, $N_v = N_d = 15$ stations/carryer, $p_v = p_d = 0.5$ and $L_b = 25$ kbit/message.

![Graph 2](image2.png)

**Fig. 2** $P_{\text{drop}}$ performance as a function of both $p_v$ and $p_d$ by assuming $T_f = RTD_{\text{max}} = 16$ ms, $N_v = N_d = 15$ stations/carryer, $r_d = 0.4$ packets/slot and $L_b = 25$ kbit/message.
Fig. 3 Behavior of $N_{\text{max}}$ as a function of $T_f$ with $p_t = 0.6$, $p_d = 0.2$, $N_d = 15$ data stations/carrier, $r_d = 0.4$ packets/slot and $L_b = 25$ kbit/message.

Fig. 4 Behavior of $\mu_v$ as a function of $r_d$ for $L_b = 25$ kbit/message, $N_d = 20$ data stations/carrier and the optimized system parameters (i.e., $p_v = 0.6$, $p_d = 0.2$, $T_f = 16$ ms).
Fig. 5 Impact of different messages length $L_b$ for $r_d = 0.4$ data packets/slot, $N_v = N_d = 15$ stations/carrier and the optimized system parameters (i.e., $p_v = 0.6$, $p_d = 0.2$, $T_f = 16$ ms).

Fig. 6 Behavior of $P_{drop}$ and $T_{msg}$ as a function of the input data traffic, $r_d$, in a terrestrial microcellular system (RTD = 0), dashed line, and in a LEO-MSS (RTD = 10 ms), continuous line, for $T_f = 10$ ms, $N_v = 8$ voice stations/carrier, $N_d = 10$ data stations/carrier, $p_v = 0.35$, $p_d = 0.15$, $L_s = 10$ packets/message.
Fig. 7. Behavior of $P_{\text{drop}}$ as a function of the input data traffic $r_d$ ($T_f = 16$ ms, $N_v = 20$ voice stations, 20 slots/frame)

Fig. 8. Behavior of $T_{\text{pkt}}$ (in slots) as a function of the input data traffic $r_d$ ($T_f = 16$ ms, $N_v = 20$ voice stations, 20 slots/frame)

Fig 9. Histogram of the distribution probability of $T_{\text{pkt}}$ (in slots), for $r_d = 0.287$; ($T_f = 16$ ms, $N_v = 20$ voice stations, 20 slots/frame)