Performance Evaluation and Optimization of a CDMA/PRMA Medium Access Protocol for WWW Users in Mobile Networks

Andri Kofmehl, Daniel Grob, Abbas Ibrahim and Samir Tohmé

***Abstract*— Future mobile communication networks are scheduled to integrate LEO (Low Earth Orbit) satellites as access points. These networks are conceived to provide various services, multiplexed on the same transmission medium. To meet the requirements for an efficient and flexible mobile access, a CDMA/PRMA protocol (Code Division Multiple Access / Packet Reservation Multiple Access) has been proposed. It is based on a so-called channel access function, which governs the access permission probability depending on the momentary user traffic load. In this paper, the proposed MAC (Medium Access Control) protocol is analyzed for WWW data transmission with a given maximum packet loss rate. To maximize the performance, an iterative simulation method was developed to optimize the channel access function. The results show very high efficiency for substantial traffic load variations, provided that the access function is adapted to the total number of users. The capacity is limited only by the maximum access delay accepted by the end-users.**

***Index Terms*— CDMA/PRMA, MAC protocols, Mobile Access, LEO Satellite Networks**

I. INTRODUCTION

WITH the deployment of sophisticated portable devices, the demand for wireless multimedia services has been rapidly increasing. The third generation mobile networks (3G) are designed to provide various services, such as voice, data and video transmission as well as enhanced messaging. In order to comply with the demand for high spectrum efficiency and multiple service support, a joint CDMA/PRMA (Code Division Multiple Access / Packet Reservation Multiple Access) structure in combination with an adaptive permission probability protocol has been proposed [1]. Its implementation is being discussed with regard to the satellite component of the UMTS network (S-UMTS) [2], [3], [4], [5].

Medium access control (MAC) is handled by a *channel access function*. This function determines the access permission probability depending on the number of active users. The channel access function has to restrict the number of simultaneously transmitting users in order to minimize packet loss due to *multiple access interference (MAI)*. To achieve the anticipated *quality of service (QoS)* in terms of packet loss, it is crucial to control the MAI, thereby setting constraints for system capacity.

Andri Kofmehl and Daniel Grob are with the Department of Information Technology and Electrical Engineering at the ETH Zürich.
Abbas Ibrahim and Samir Tohmé are with the Ecole Nationale Supérieure des Télécommunications, Paris.
Email: akofmehl@ee.ethz.ch, dgrob@ee.ethz.ch, abbas.ibrahim@enst.fr, samir.tohme@enst.fr.

The type of service analyzed corresponds to a source of WWW data traffic as specified by the 3GPP (3rd Generation Partnership Project) [6]. In order to maximize the capacity of a real system for a given QoS, the channel access function has to be chosen such that it provides efficient channel access for different traffic loads. By employing an iterative method based on successive simulations, an optimized channel access function can be obtained. The results of the optimization process are then subject to performance evaluations in terms of protocol efficiency and access delay.

The paper starts with a brief description of CDMA/PRMA and its characteristics followed by a section on data transmission, where first the source model is described and data traffic as well as system capacity are analyzed. Thereafter, a method for the optimization of the channel access function is presented and its performance is evaluated. The paper concludes with a brief discussion of the results.

II. CDMA/PRMA

Like for most wireless technologies, a major requirement for third generation mobile satellite networks is a highly efficient exploitation of the available frequency spectrum. A technique that meets this requirement is CDMA [7], which also offers a number of supplementary advantages [1].

For the transmission of different services on a shared medium, *packet switching* is very favorable. Packet Reservation Multiple Access (PRMA) means that terminals reserve a channel for the duration of the actual data transmission only. The idea of PRMA is inspired by the ATM technology and by the GSM/GPRS protocols.

A. Transmission Channels

The physical layer of the proposed system is a TD-CDMA/PRMA combination (TD for Time Division). The time axis is divided into slots of duration T_s. N_s slots are grouped into frames of duration T_f. A channel occupies the same slot in each frame alternating with the other channels on the other slots. Moreover, there are several CDMA channels in every time slot, i.e. the signal of a packet is modulated and spread to the whole bandwidth of the transmission by a certain code. These encoded signals are superposed and transmitted on the same slot. The receiver reconstructs the signals by demodulation with the corresponding code. The number of CDMA channels is limited by the desired Quality of Service (QoS) because superposed simultaneously transmitted signals are perceived as broad band noise by a single user. This so-called multiple access interference (MAI) corrupts the received signal. Therefore, the quality of the transmission decreases and the corruption probability P_{loss} of a transmitted packet increases with the number of users on the same time slot.

B. Channel Access

As traffic increases, access to the channel must be restricted in order to avoid excessive packet loss due to MAI. When reserving a channel for packet transmission, terminals must perform a Bernoulli experiment [1] with the current *access permission probability* P_a as input parameter. They are allowed to occupy a channel if the outcome of the Bernoulli experiment is positive. P_a is determined depending on the number of users n_s that are already transmitting in the corresponding slot. $P_a(n_s)$ is called *channel access function* (for an example see Figure 1). It is provided by the network connection management.

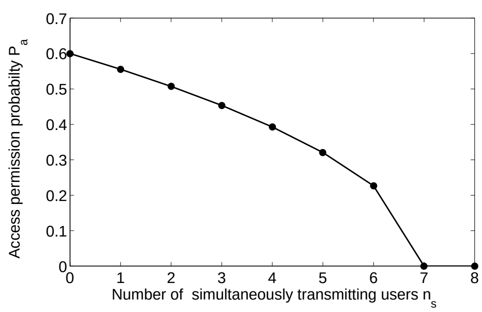

Fig. 1. Example of a channel access function.

C. Packet Loss

As mentioned previously, MAI increases the noise level and thus influences the transmission quality by corrupting a certain number of bits. Assuming that the MAI is Gaussian, the probability P_e that a bit is corrupted is, as shown in [1],

$$P_e \approx Q(\overline{SNR}), \tag{1}$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{u^2}{2}} du \ . \tag{2}$$

$\overline{SNR}$ is the average signal-to-noise ratio, obtained from

$$\overline{SNR} = \sqrt{\frac{3P_0 \cdot sf}{\underbrace{(n_s - 1)P_0}_{Intracell} + \underbrace{\sum_{k=1}^{n_s} \sum_{i=1}^{R} P_{(k,i)0}}_{Intercell}}}, \tag{3}$$

where P_0 is the power of the received signal, n_s the number of active senders, R the *frequency reuse factor* and sf the *spreading factor*.

[1]The outcome of the Bernoulli experiment is positive if the input parameter is less than a uniformly distributed random value $r \ \epsilon \ [0, 1]$.

If intercell interference is neglected, a straightforward approximation is obtained

$$\overline{SNR} = \sqrt{\frac{3 \cdot sf}{n_s - 1}} \ . \tag{4}$$

If $Q_e = 1 - P_e$ denotes the success probability of a single bit, the success probability of an entire packet is

$$Q_E = \sum_{i=0}^{t} \binom{L}{i} (1 - Q_e)^i (Q_e)^{L-i} \ , \tag{5}$$

where L is the packet length and t the number of errors that are corrected by the implemented code. As an example, we take an isolated cell with the values of Table I. For this case, a graph is depicted in Figure 2 with Q_E as a function of the number of simultaneously transmitting users.

TABLE I
PARAMETERS OF THE TRANSMISSION SYSTEM

Packet length L [bit]	160
Frame duration T_f [ms]	20
Slot duration T_s [ms]	2.5
Number of slots per frame N_s	8
Spreading factor sf	7
Data rate per channel d [bit/s]	8000
Number of correctable errors t	12

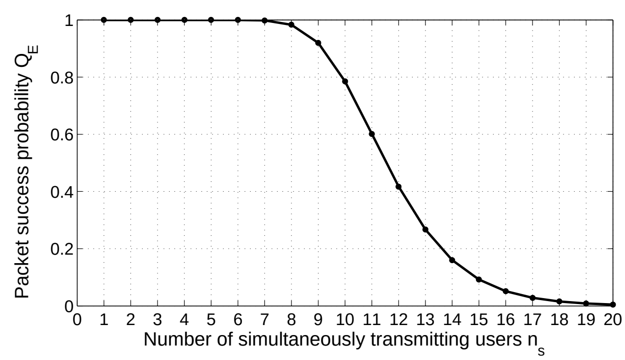

Fig. 2. Packet success probability as a function of the number of users in the slot.

As seen from Figure 2, the packet success probability decreases rapidly, when n_s exceeds a certain threshold. Therefore, it is necessary to strictly limit the number of users per slot in order to protect the transmission and ensure the required QoS.

III. DATA TRANSMISSION

The type of data traffic analyzed in this paper corresponds to traffic generated by a user surfing the WWW. This can be considered as a client-server process consisting of requests and responses. Whenever the user clicks on a link, a small amount of data, representing a request for the transmission of a web page is transmitted. Thereupon, the requested data is provided via the satellite link. As the requests' contribution to overall traffic is negligible, only the response transmission will be considered. In order to assure a satisfactory QoS without packet retransmission, the error probability P_{loss} is assumed to be limited to 0.1 %. However, the user is supposed to tolerate a certain delay in the arrival of the requested page.

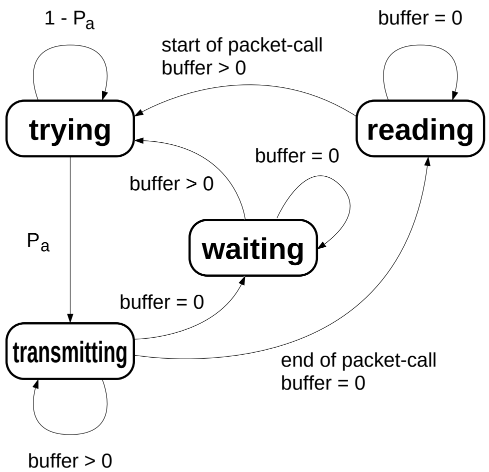

Fig. 3. State diagram of a data user.

A. Source Description

The web source is described by a state diagram as depicted in Figure 3. When a new page is retrieved, its content is transmitted during a packet-call. A packet-call consists of several data bursts, which are each split into small packets to be transmitted by the TD-CDMA/PRMA system. Before their transmission, the packets of arriving bursts are written to a buffer memory. When the first burst of a packet-call arrives in the buffer, the transmission unit switches to the *trying* state where it contends for a channel. A successful reservation allows to transmit data as long as the buffer is not empty. Whenever the buffer is empty again, the reservation ends. If the end of the packet-call is not reached yet, the system then switches to the *waiting* state. As soon as new data is written to the buffer, it returns to the *trying* state. Upon completion of the transmission of the web page, the user spends a certain time reading its content, represented by the *reading* state. In Figure 4, an example for this process is given as a function of time.

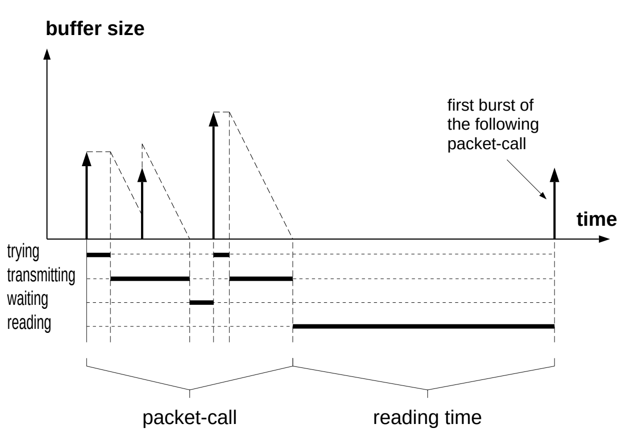

Fig. 4. Buffer size during a cycle.

According to the model proposed by [6], the burst size of a packet-call s_b follows a Pareto distribution with cut-off. The probability density function of the normal Pareto distribution with parameters α and k is given by

$$f_{\alpha,k}(s_b) = \begin{cases} \frac{\alpha k^\alpha}{s_b^{\alpha+1}} & , \quad s_b \geq k \\ 0 & , \quad s_b < k \end{cases} . \tag{6}$$

Table II contains the values of the parameters that apply for the burst size.

TABLE II
PARAMETERS OF THE PARETO DISTRIBUTION

α	1.1
k	81.5
cut-off [bytes]	1500

The number of bursts per packet-call n_b is distributed geometrically, whereas the time intervals between the bursts t_b and the reading time t_{read} are distributed exponentially. Their mean values are given by Table III. For practical reasons, they have been adjusted for the simulations presented below. The traffic generated by a single user according to the adjusted parameters is comparable to traffic resulting from the original parameters.

TABLE III
PARAMETERS OF A DATA USER

	original	adjusted
$\overline{t}_{read}$ [s]	412	40
$\overline{n}_b$	25	2.5
$\overline{t}_b$ [s]	0.5	0.5

B. Traffic per User

The traffic generated by a data source can be defined as the fraction of time a channel is occupied [2]. Given the model described above, the Erlang traffic of a single data source during one cycle (consisting of a packet-call and a reading period) can be calculated as

$$a = \frac{\frac{1}{d}\sum_{i=1}^{n_b} s_{b,i}}{t_{read} + \sum_{i=1}^{n_b-1} t_{b,i} + s_{b,n_b} \cdot \frac{1}{d}} , \tag{7}$$

where d represents the data rate of a single channel. In this formula, the time required to reserve a transmission channel is neglected and the transmission of a burst is assumed completed before the arrival of the next burst. The mean of a amounts to 0.0183 Erlang per user according to the (adjusted) source parameters.

In case of congestion (Figure 5), the time required for the reservation of a channel increases and can thus no longer be neglected. If the congestion intensifies, it is assumed that the buffer is never empty during a packet-call. In this case, the traffic is approximately

[2]The traffic is expressed in Erlang, with 1 Erlang corresponding to a continuous occupation of single channel.

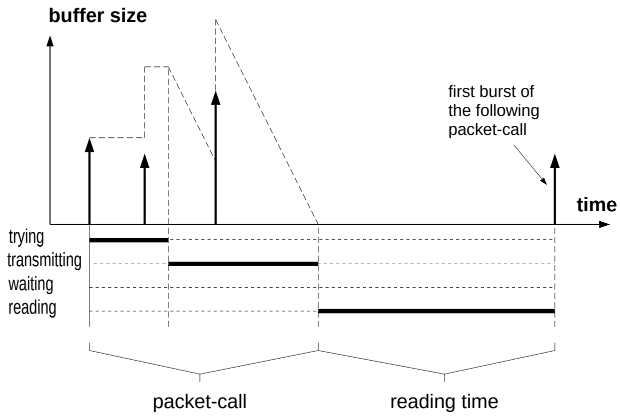

Fig. 5. Buffer size under congestion.

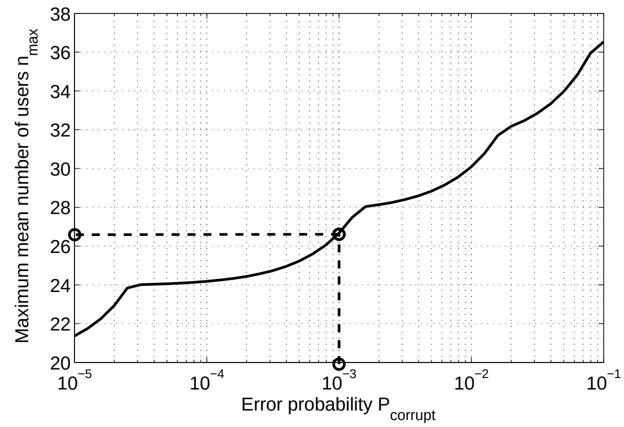

Fig. 6. Maximum number of users as a function of the error probability.

$$a = \frac{\frac{1}{d}\sum_{i=1}^{n_r} s_{r,i}}{t_{read} + t_{res} + \frac{1}{d}\sum_{i=1}^{n_r} s_{r,i}} , \quad (8)$$

with t_{res} being the time needed to reserve a channel. An increase of t_{res} causes a decrease of the traffic a. Thus, the number of users sharing a link with a certain capacity is not strictly limited due to the fact that the traffic per user will adapt to the degree of congestion encountered.

C. System Capacity

Given a desired mean error probability P^*_{loss}, the system capacity for an ideal MAC can be determined. The error probability will be lowest when the occupation of the slots is uniform, i.e. the number of users per slot takes two neighboring values n_1 and n_2 [3]. If the total number of users is a multiple of the number of slots N_s, n_1 equals n_2. Starting from the packet success probability Q_E, the total error probability P_{loss} is calculated as

$$P_{loss} = \frac{N_1 n_1 (1 - Q_E(n_1)) + N_2 n_2 (1 - Q_E(n_2))}{N_1 n_1 + N_2 n_2}, \quad (9)$$

where N_1 and N_2 denote the number of slots containing n_1 and n_2 users respectively. Taken into account that $N_1 + N_2 = N_s$, a system of two equations is obtained, which allows to determine N_1 and N_2 [4].

The specified satellite communication link is assumed to be shared in time division mode between two different services, one of which is web data transfer at 8 kb/s as described above. Accordingly, 4 out of 8 slots are available for the web users, whereas the remaining slots are at the disposal of other services such as voice transmission [8].

Figure 6 depicts the maximum number of users n_{max} as a function of the error probability P_{loss}, as it can be computed using the considerations above with $n_{max} = N_1 n_1 + N_2 n_2$. If an error probability of 0.1% is designated, an ideal protocol would support an average number of 26.7 users transmitting simultaneously. Thus, the part of the link available for web users supports a traffic of 26.7 Erlang. If the traffic per user is assumed to be 0.0183 Erlang, as calculated in the preceding section, the system is expected to manage approximately 1400 users.

[3]If the total number of users corresponds to n, it is $n_1 = \lceil \frac{n}{N_s} \rceil$ and $n_2 = \lfloor \frac{n}{N_s} \rfloor$, N_s being the total number of slots.

[4]N_1 and N_2 are average values and therefore not necessarily integers.

IV. Channel Access Function

A. Characteristics

The channel access function is used to control the medium access of users that want to transmit data. To assure efficient system operation, it must comply with the following major requirements:

- Channels must be assigned as soon as they are needed, which keeps the channel access time and the transmission delay on a low level. Thus, the values of the channel access function are chosen as high as possible for user loads n_s that correspond to low error rates $P_E = 1 - Q_E$.
- As the probability of packet loss for all users transmitting simultaneously increases drastically when a certain number of users per slot is exceeded, it is necessary to protect the transmission from excessive access and therefore data loss. As a consequence, the values of the channel access function have to be chosen such that excessive access becomes very unlikely.
- Since the total number of users is not as strictly limited, as for example for voice transmission [8], the access function will have to deal with varying degrees of congestion. This imposes the need for system adaptibility to provide efficient access control for changing load conditions.

B. Optimization of the Channel Access Function

In this section, we present a method that permits to determine an optimized channel access function for a given total number of users and a designated packet loss probability of $P^*_{loss} = 0.1\%$.

If medium access is controlled by a channel access function, there is a chance for excessive channel reservation. Whenever the sum of the users trying to reserve a slot n_t and the users already transmitting in the respective slot n_s exceeds the maximum number of users per slot $n_{s,max}$ [5] , there is a certain probability P_{ER} of excessive reservation. By choosing appropriate values for the channel access function, P_{ER} can be adjusted to ensure that the packet loss probability P_{loss} meets the desired level P^*_{loss}.

If the distribution of n_t as a function of n_s, which will be denoted $P_t(n_t, n_s)$, is known, the probability of excessive reservation P_{ER} for a given channel access function $P_a(n_s)$ can be calculated as

$$P_{ER}(n_s) = \sum_{n_t}^{\infty} P_t(n_t, n_s) \sum_{i}^{n_t} C^i_{n_t} P_a(n_s)^i (1 - P_a(n_s))^{n_t - i} \tag{10}$$

with $n_t = n_{s,max} - n_s + 1$ and $i = n_{s,max} - n_s + 1$ as the first indices of the corresponding sums and $C^i_{n_t} = \binom{n_t}{i}$.

The distribution $P_t(n_t, n_s)$ that is obtained from a simulation can be used to compute P_{ER} as a function of P_a for the values of $n_s \leq n_{s,max}$ as depicted in Figure 7, plot A. For a chosen value of P_{ER} (e.g. 10^{-3}), the intersections of the corresponding horizontal with the array of curves yield a channel access function that would produce the chosen probability of excessive reservation. The obtained channel access function as shown in plot B of Figure 7 ($P_a(n_s)$ will be set to zero for $n_s \geq n_{s,max}$).

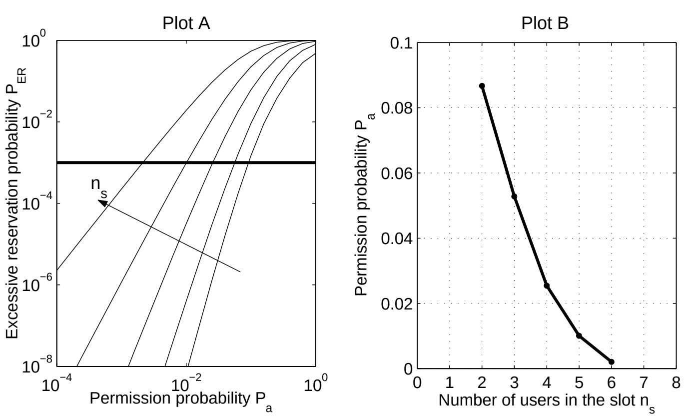

Fig. 7. Computation of the channel access function.

However, the method as it has been presented so far, is insufficient since it is based on uncertain data: The distribution $P_t(n_t, n_s)$ relies on the channel access function $P_a(n_s)$, whereas, the newly deduced channel access function would result in a different distribution. Therefore, the distribution $P_t(n_t, n_s)$ used for optimization will only be reproduced by the new channel access function if the latter equals the channel access function that was employed to obtain it. This is generally not the case. To overcome this problem, an iterative method is deployed. This means that several consecutive simulations are performed, while the value of P_{ER} for the computation of each improved channel access function is adapted such that the packet loss probability P_{loss} approaches P^*_{loss}. The iteration process ends when the channel access function remains the same, while the packet loss probability corresponds to P^*_{loss}.

[5]The maximum number of users per slot $n_{s,max}$ is determined as the smallest value of n_s for which $1 - Q_E(n_s) \geq P^*_{loss}$, Q_E being the packet success probability. For the parameters chosen in this paper, $n_{s,max}$ equals 7.

V. Simulation Results

In order to evaluate the performance of the protocol for a set of critical numbers of users, three scenarios representing different degrees of congestion have been selected for closer analysis:

Scenario I	1000 users
Scenario II	1250 users
Scenario III	1500 users

A. Optimized Channel Access Functions

Employing the iterative method described in the preceding section, three individual optimized channel access functions are obtained. They are depicted in Figure V-A. Obviously, the number of users has a very strong impact on the values of the channel access function, which is a consequence of the accumulation of contending users in case of congestion (Figure 9, middle row). Empirical results show that for a number of users close to or above the congestion threshold, the channel access function corresponds approximately to a parabolic function.

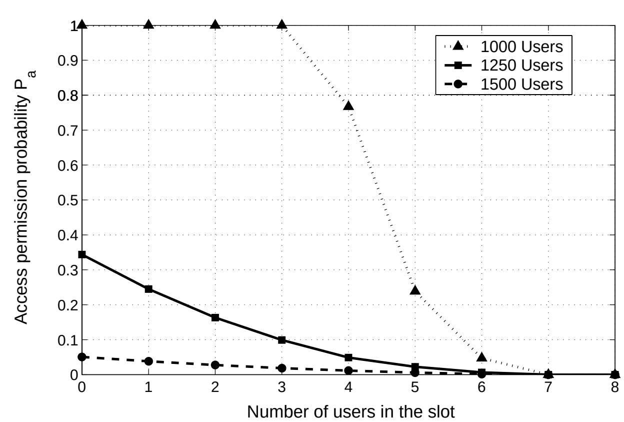

Fig. 8. Access function for the three scenarios.

The plots of the third row of Figure 9 show the distributions of users occupying a slot. They illustrate that the occurence of excessive reservation (bars at $n_s \geq 8$) remains low as the user load increases.

B. Protocol Performance

1) Efficiency: The efficiency E of an access protocol can be defined as the ratio of the actual amount of data transferred to the maximum amount of data the system could transfer in a given time interval. This corresponds to the transmitted traffic A_t divided by the system capacity A_{max}. In order to evaluate the performance of the protocol, the value of E is

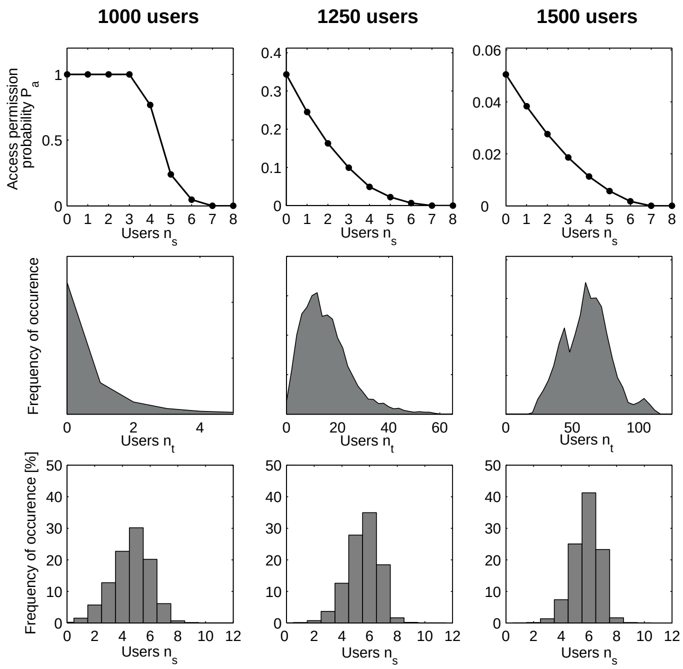

Fig. 9. Simulation results.

compared to E_i, which represents the ideal efficiency, defined as the proportion of the offered traffic A_o to the capacity of the system A_{max}. We have

$$E_i = \frac{A_o}{A_{max}} \quad \text{and} \quad E = \frac{A_t}{A_{max}} \, . \qquad (11)$$

As illustrated by Table IV and Figure 10, the protocol is quasi ideal for a number of users clearly below the congestion threshold, whereas efficiency remains very high as the number of users increases.

TABLE IV

PROTOCOL EFFICIENCY

Number of users	1000	1250	1500
System capacity A_{max} [Erlang]	26.7		
Offered traffic per user a [Erlang]	$0,0183$		
Offered traffic A_o [Erlang]	18.3	22.9	27.4
Ideal efficiency E_i [%]	69	86	100
Simulations			
Transmitted traffic A_t [Erlang]	18.3	22.3	23.3
Real efficiency E [%]	69	83	87

2) Access Delay: For this evaluation, the access delay is defined as the mean time a packet spends in the buffer before it is transmitted. For a given data packet, the access delay depends on the time spent queueing behind packets that have been written to the buffer earlier and possibly on the time needed to obtain a transmission channel. If the traffic load is low, the reservation component is negligible, whereas in case of congestion, it increases considerably. This effect appears in Table V and is illustrated in Figure 11. It should be noted that the access delay limits the system capacity in terms of the

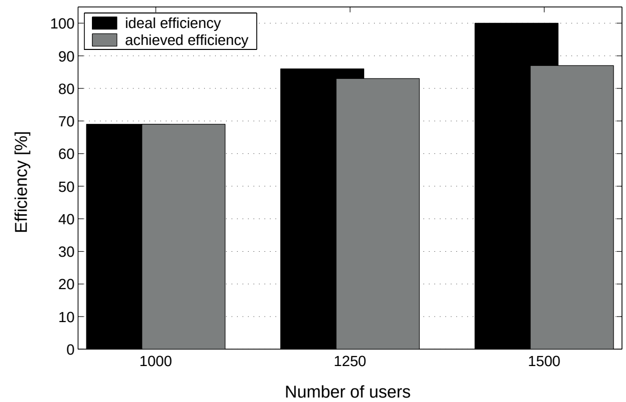

Fig. 10. Efficiency of the protocol for the three cenarios.

acceptable number of users. This is due to end-users tolerating only a certain latency when surfing the web, which makes the access delay an important criterion for the quality of service provided.

TABLE V

ACCESS DELAY INDUCED BY THE PROTOCOL

Number of users	1000	1250	1500
Access delay [s]	0.46	0.82	2.11
reservation [s]	0.01	0.30	1.47
transmission [s]	0.45	0.52	0.64
reservation component [%]	2	37	70

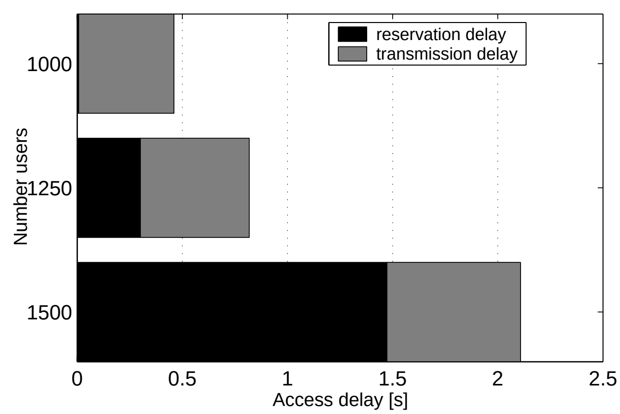

Fig. 11. Access delay of the three scenarios.

VI. CONCLUSIONS

In this paper, we show that an iterative simulation method may be employed to optimize the performance of the channel access function, which is the core element of the proposed MAC for a joint CDMA/PRMA protocol for LEO satellite networks. A strong dependency of the channel access function on the total number of users has been stated. This demands a channel access function which is adapted to the traffic load. The results of the simulations performed after the optimization process demonstrate that the protocol guarantees very high efficiency, regardless of the total number of users. System overload merely causes an increase in access delay, but does not considerably affect data throughput. Consequently, the limiting factor in terms of capacity is the maximum delay accepted by end-users when retrieving a web page. A so-called connection admission control (CAC) function will have to be implemented to ensure the appropriate QoS.

The results make the investigated MAC protocol a promising candidate for future wireless communication networks providing services.

REFERENCES

[1] A. E. Brand and A. H. Aghvami, "Performance of a joint CDMA/PRMA protocol for mixed voice/data transmission for third generation mobile communications," *IEEE J. Select. Areas Commun.*, 1996.

[2] A. Ibrahim and S. Tohmé, "CDMA/PRMA analytical model for voice users in Satellite-UMTS systems," *ICC New York*, 2002.

[3] ——, "A modified CDMA/PRMA medium access control protocol for integrated services in LEO satellite systems," *Mobicom2000 Boston*, 2000.

[4] E. J. Bostic and I. Mertzanis, "Network and MAC issues for mobile satellite multimedia networks," *COST-252/253/255 Joint Workshop, Toulouse, France*, 1999.

[5] G. G. Benelli, R. Fantacci and C. Ortolani, "Performance analysis of a PRMA protocol suitable for voice and data transmissions in low earth orbit mobile satellite systems," *IEEE Transactions on wireless communications*, 2002.

[6] ETSI, "Universal mobile telecommunication system (UMTS), selection procedure for the choice of radio transmission technologies of the UMTS," ETSI, Tech. Rep. TR 101 112 v3.2.0, 1998.

[7] R. Prasad and T. Ojanper, "An overview of CDMA evolution toward Wideband CDMA," *IEEE Communications Surveys*, 1998.

[8] A. Kofmehl and D. Grob, "Analyse et optimisation d'un protocole CDMA/PRMA pour des services voix/data dans une constellation de satellites LEO," March 2002.