

# Techniques for Performance Improvement of IMS Voice Calls in UMTS Networks

Eduardo R. Vale, João Célio B. Brandão, Marco Grivet

**Abstract**—The introduction of IMS (IP Multimedia Subsystem) in UMTS (Universal Mobile Telecommunication Services) architecture has made possible the implementation of end to end VoIP calls, thus integrating mobile and fixed networks in the IP technology. However we know that, without the use of some techniques, VoIP transmission can demand more resources, in terms of bandwidth and power. This is not a serious problem for fixed networks, but it is of special concern in the case of mobile systems, mainly because of the natural limitations of the air interface and the power restrictions of the user cellular equipment. In this paper we address the performance of voice calls in terms of transmitted power efficiency. We show that making convenient choices of standard procedures and using specific optimization techniques it is possible to turn IMS VoIP competitive with the traditional voice telephony service performed by the CS (Circuit Switching) counterpart of the UMTS.

**Index Terms**—UMTS, IMS, Circuit Switching, voice calls.

## I. INTRODUCTION

IN recent years, internet and cellular applications have emerged as the major driving forces behind new developments in the area of telecommunications networks. Furthermore, there is a great interest in using IP technology to provide multimedia services in general.

Cellular systems are evolving from a purely telephony system to a multimedia service based on the IP technology which has been incorporated in 3G and particularly in the UMTS specifications. In fact, the UMTS system is in continuing evolution. There have been several releases, each of them providing a new version of specifications for vendors and operators to be used in the system implementation. Presently, the reference architecture specifies the main network components of UMTS including interfaces between access network, core network and also signaling network.

Two domains are delimited in the UMTS architecture: a Circuit Switched (CS) domain and a Packet Switched (PS) domain. CS domain includes the network components and protocols which are used in the Circuit Switched mode of network operation and are intended to support basically voice services. The PS domain, including the IMS (IP Multimedia Subsystem), is based on SIP [1] for signaling and is intended to support multimedia traffic in general. An administration interested in deploying an UMTS system can choose to install only CS or PS facilities or both of them simultaneously.

Eduardo R.Vale is with the Universidade Federal Fluminense, Rio de Janeiro, Brazil. E-mail:eduardo.vale@globo.com

João Célio B. Brandão, Marco Grivet are with the Center for Studies in Telecommunication of the Pontifical Catholic University of Rio de Janeiro, Brazil. E-Mail:jcelio@cetuc.puc-rio.br, mgrivet@cetuc.puc-rio.br

In both modes of transmission, the lower layers are responsible for a set of procedures as padding, coding, interleaving and rate matching. The 3GPP defines a set of Radio Access Bearers (RABs), each of them with the specification of one or more frames to carry the blocks of information coming from the upper protocol levels. In the case of the traditional voice service provided by the CS mode, the blocks of information coming from the AMR (Adaptive Multi Rate) codec fit perfectly in the frames defined by the proper RAB. On the other hand, for the transmission of VoIP in the IMS mode, there is a multitude of information packet sizes and it is not practical to define a RAB with the same number of different frame sizes. In this case padding is used to match the block sizes to the available frame sizes. This operation is performed at the RLC (Radio Link Control) level. At physical level, the Rate Matching units intend to adjust the coded information blocks of the different transport channels to the size of the physical channel. Additionally, this procedure may be used to obtain different QoS requirements for the channels, assuming that the error performance is improved with bit repetition and worsened with bit puncturing.

It is important to observe that, despite the increasing use of new forms of communication, voice continues to be the responsible for the major part of revenues, in cellular networks. In this paper we address the performance of UMTS voice call transmission in terms of transmitted power. It will be shown that, with the adoption of some procedures, it is possible to make VoIP IMS calls competitive with CS calls with respect to power efficiency. These procedures include mainly the optimization of Padding and Rate Matching operations, as well as the utilization of the following complementary techniques: RLC operating in the Transparent Mode (i.e. with null header)[2]; no transmission of UDP Checksum bits; use of Turbo Code also for the signaling channel; use of CRC with 8 bits for voice packets; PDCP (Packet Data Convergence Protocol) with additional padding function. The above techniques will be detailed throughout the paper. To the best of the author's knowledge, the approach here developed has not yet been addressed in the technical literature.

The paper is organized as follows. Section 2 describes the UMTS protocol architecture. Section 3 presents the fundamentals of Rate Matching optimization for CS and IMS modes and the utilization of complementary techniques for improving performance of voice calls. Section 4 describes Padding optimization for IMS and presents a numerical example of application for the techniques here introduced. Finally, Section 5 presents the conclusions.

II. UMTS PROTOCOL ARCHITECTURE

Figures 1 and 2 show the protocol architecture [3], [4], [5], for CS and PS modes, respectively. There we can see the three lower layers and the upper layers of the protocol stack. The Physical Layer (PHY) has the normal attributes common to data networks, including the transport of channel measurements to the RRC (Radio Resource Control). Layer 2 is composed of MAC (Medium Access Control) and RLC. In these figures we can see the inferior part of Layer 3, with RRC and BMC (Broadcast/Multicast Control). Also, in the case of PS mode, we have the PDCP, responsible for the compression/decompression of packet's headers [6]. RRC unit executes the control of the air interface and BMC is responsible for the broadcast/multicast of messages to the UE's (User Equipment). The connection between UTRAN (UMTS Terrestrial Access Network) and CN (Core Network) is performed by the Iu interface, with the signaling (Iu-CS and Iu-PS) and user parts (Iu-UP) for traffic. In the case of PS mode, both Iu-PS and Iu-UP have the same path (UTRAN-SGSN - Serving Gateway Support Node).

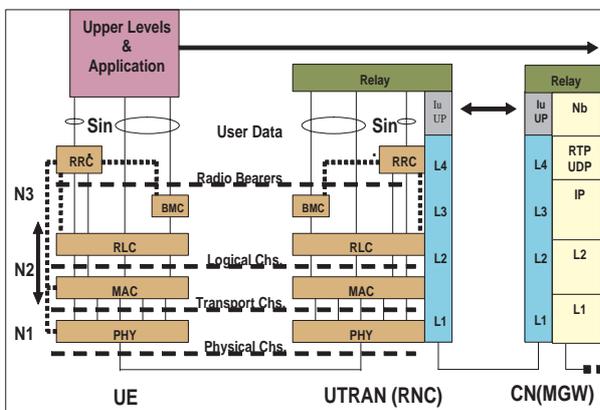


Fig. 1. Protocol Architecture (CS mode)

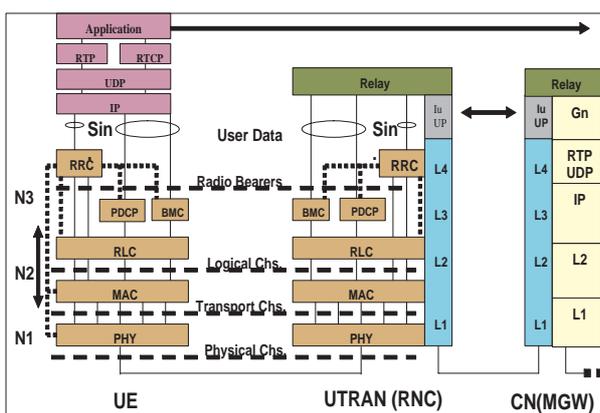


Fig. 2 Protocol Architecture (PS mode)

CRC attachment, the blocks of data are either concatenated or segmented, to fit the allowable limits specified for the block sizes. Next, the channel encoding is performed, with four different types of coding: 1/2 convolutional coding (CC2), 1/3 convolutional coding (CC3), 1/3 turbo coding (TC) or no coding. The RF equalization and RF segmentation have the purpose to adapt the data to the radio frames, when more than one radio frame is transmitted. The first interleaving reduces the probability of an occurrence of error bursts and the rate matching unit has basically two purposes: to fit the data blocks in the specified physical channel and to adapt the transport physical channel and to adapt the transport physical channels for the different levels of QoS required. Finally, the different channels are serial multiplexed, as show in Figure 4, resulting in the formation of the Common Composite Transport Channel (CCTrCH).

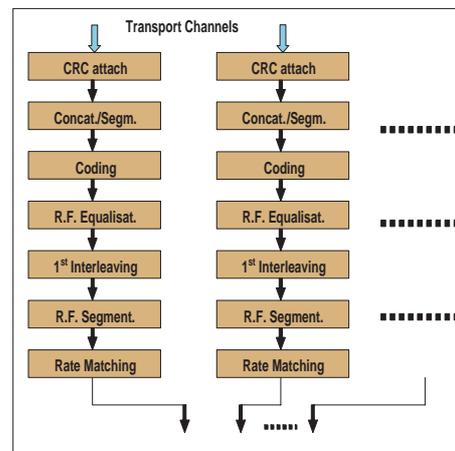


Fig. 3. Channel Structure

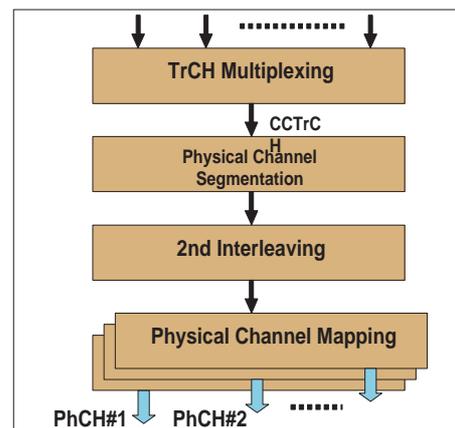


Fig. 4. Multiplexing Structure

Considering now the Physical Layer, Figure 3 shows the process of creation and multiplexing of transport channels in the case of uplink transmission [7]. For each channel, after the

The described architecture provides the fundamentals for comparing the performance of voice calls in CS and IMS modes. This is done in the next two sections.

### III. PERFORMANCE EVALUATION OF VOICE TRANSMISSION WITH RATE MATCHING OPTIMIZATION

In this section the main characteristics of voice transmission in CS and PS modes are presented and particular scenarios are chosen to numerically evaluate the transmission performance in terms of power efficiency. The evaluation considers novel procedures for improving the performance, including Rate Matching optimization and others.

#### A. CS Mode

The main characteristics of CS mode transmission is now presented according to the current specification of the UMTS and assuming that the AMR codec is operating at 12.2 kbit/s. In this paper it is considered the UMTS FDD (Frequency Division Duplex) mode of operation. Four channels are defined in [8]:

Channel A: carries the most sensitive information bits (81 bits) coded by 1/3 Convolutional Code (CC3) and requires  $BER = 5 \cdot 10^{-4}$ .

Channel B: carries the medium sensitive information bits (103 bits) coded by 1/3 Convolutional Code (CC3) and requires  $BER = 10^{-3}$ .

Channel C: carries the least sensitive information bits (60 bits) coded by 1/2 Convolutional Code (CC2) and requires  $BER = 5 \cdot 10^{-3}$ .

Channel D: carries signaling data (148 bits) coded by 1/3 Convolutional Code (CC3) and requires  $BER = 10^{-4}$ .

In Figure 5 we see the formation of the radio frame (10ms) with the data of the four transport channels. For each channel, after the CRC attachment, the blocks of data are either concatenated or segmented, to fit the allowable limits specified for the block sizes. Next, the channel encoding is performed, with four different types of coding: 1/2 convolutional coding (CC2), 1/3 convolutional coding (CC3), 1/3 turbo coding (TC) or no coding. The RF equalization and RF segmentation have the purpose to adapt the data to the radio frames, when more than one radio frame is transmitted. The first interleaving reduces the probability of a occurrence of error bursts and the rate matching unit has basically two purposes: to fit the data blocks in the specified physical channel and to adapt the transport channels for the different levels of QoS required. The tail bits are used for the initialization of the respective coders. It is important to note that the final number of bits depends of the Rate Matching unit.

The Rate Matching unit run a standard procedure intended to match the number of data bits to the available number of bits in the physical channel and with this goal, data bits may be duplicated or punctured. Furthermore, this procedure may be used to attain possibly different QoS requirements for the channels, assuming that the error performance is improved with bit repetition and worsened with bit puncturing. The selection of the bits to be repeated or punctured is a task of superior protocols.

There are several papers in the literature [9], [10], [11] considering the rate matching operation, but none of them gives a sound treatment to the performance subject. The performance improvement allowed by bit repetition is attained

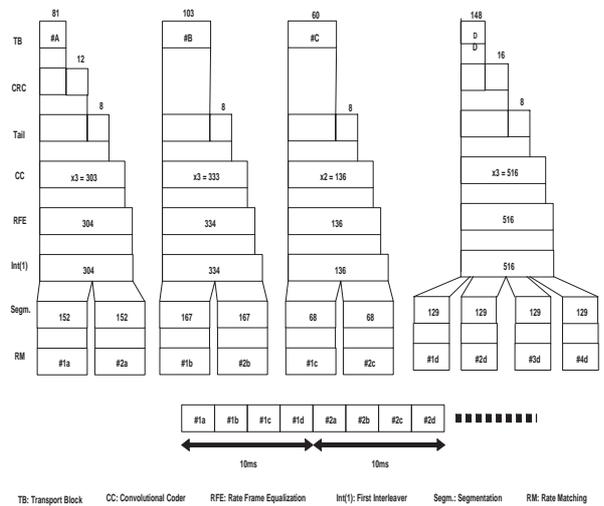


Fig. 5. Transmission of Voice Signal (CS mode)

once the receiver is capable of using the extra bit energy in the detection process. This may be easily calculated for uncoded transmission in AWGN channels but, for a transmission using convolutional or turbo codes, one has to rely on simulation to generate performance curves.

In [12] we present a methodology for optimally defining the duplicated or punctured bits such that the required transmitted power is minimized while the QoS requirements are satisfied. The procedure presented consists of obtaining a set of BER versus  $E_b/N_0$  curves by simulating data transmission in an AWGN channel with the standard convolutional and turbo coders and considering different figures for bit repetition and puncturing. These curves were adjusted by polynomial functions and an optimization problem was formulated and analytically solved. The main steps of the development are summarized in Appendix I.

As an application example, suppose the case illustrated in this section. For transmitting voice at 12.2 kbit/s with the AMR codec, associated with signaling, it is interesting to work with a spreading factor (SF) of 64, which results in an uplink channel rate of 60 kbit/s. With the four channels previously defined we have (prior to the rate matching units) a total of 516 bits ( $152+167+68+129$ ) to be allocated in a radio frame of 10ms, as shown in Figure 5. Considering that the capacity of a 10ms radio frame is 600 bits per slot, we have 84 bits ( $600-516$ ) to distribute among to the four channels, using bit repetition. By implementing the optimization procedure proposed in [12] relative to the distribution of the extra bits in the channels, we obtain the results presented in Table I, which also summarizes the above assumptions. After optimized distribution of the extra bits, 14 bits were allocated to channel A, 3 to channel B and 67 to channel D, which is the most stringent channel. No extra bit is given to channel C.

If no repetition bits were given to any of these transport channels, the minimum  $E_b/N_0$  values that meet the desired BER are respectively, 2.22 dB, 1.98 dB, 1.74 dB and 2.76 dB, being the last the most severe. With the performance

TABLE I  
TRANSPORT CHANNELS DATA (CS MODE)

Channel	Coder	BER	Bits	Rep.Bits
A	CC3	$5.10^{-4}$	152	+14
B	CC3	$1.10^{-3}$	167	+3
C	CC2	$5.10^{-3}$	68	0
D	CC3	$1.10^{-4}$	129	+67
Total			516	+84

improvement achieved with the extra bits, the required BER can be satisfied in all the channels with an  $E_b/N_0$  of 1.88 dB.

### B. IMS Mode

As stated before, in order to extend VoIP utilization on fixed networks, the 3GPP has been working in methods to introduce this technology in cellular networks, with transmission of voice packets all the way from the cellular equipment to another systems and vice-versa. When designing the architecture and related procedures for voice transmission in IMS mode, several alternatives are available in 3GPP. Furthermore, different types of packets may appear in a VoIP session. Before addressing the performance evaluation, we first summarize the different procedures and packet types found in 3GPP recommendations.

#### *Type of Header Compression*

Among several mechanisms of header compression, 3GPP has chosen ROHC (Robust Header Compression) as the more suitable for UMTS applications. IETF defined three methods of operation for ROHC [13]: The Unidirectional Mode, only applicable if there is no return channel available, the Optimistic Mode and the Reliable Mode. The Optimistic Mode transmits information in a return channel with limited resources while the Reliable Mode considers the existence of a return channel capable to transmit more information than the previous mode. The Reliable Mode is more robust than the others and has been chosen for use in UMTS systems.

#### *Support for the RTP payload*

The IETF specifies two modes for the transport of voice RTP payloads: the Octet-Aligned Mode and the Bandwidth Efficient Mode [14]. The Octet-Aligned Mode is more robust, because it admits interleaving, robust sorting and CRC operations at voice frame level. These advantages are obtained at the expense of the inclusion of more bits (used mainly for padding), therefore increasing the power consumption for frame transmission. 3GPP has chosen the Bandwidth Efficient Mode for this kind of operation.

#### *Operation of RLC*

The RLC has three modes of operation [2]: the Transparent Mode, where there is no changes in the information blocks received; the Acknowledged Mode that uses a return channel for improving the transmission and is not suitable for VoIP applications; the Unacknowledged Mode, recommended by 3GPP for VoIP applications. This last mode appends to the information block some bytes indicating the SN (Sequence

Number) and the LI (Length Information) of the information packet.

#### *UDP Checksum*

The utilization of IPv6 is recommended by 3GPP. In the case of IPv4, the transmission of the UDP Checksum is considered optional by the IETF, but it is compulsory for IPv6.

#### *Coding of the Signaling Channel*

The 3GPP recommends the utilization of CC3 for the signaling channel.

#### *CRC Length*

The voice transmission in the CS mode uses CRC (12 bits) only for Class A bits, while Classes B and C bits use no CRC at all. In turn, IMS applications usually employ CRC with 16 bits.

#### *Packet Types*

In IMS mode, it is necessary to transmit packets with varying sizes and other different characteristics. 3GPP is currently developing an adequate support for the transmission of the following different categories of packets:

- RTP voice packets with no header compression (99 bytes): these occur mostly at the beginning of the call (the so called IR packets, in the ROHC nomenclature);
- RTP voice packets with intermediary degrees of header compression: occurring whenever it is necessary to adjust the mechanism of header compression;
- RTP voice packets with a high degree of header compression (40 bytes): they are called R-0 packets in the ROHC nomenclature and constitute the major part of the voice packet transmission;
- RTCP packets with no header compression (140 bytes): they occur not very frequently, during the realization of the call. 3GPP is evaluating the possibility of not using these packets at the air interface;
- SIP packets, possibly with header compression (until 7500 bytes): they occur mostly at the beginning and the end of the call;
- Voice call signaling packets: for signaling purpose. They carry the same information of the channel D specified before;

At this point it is interesting to note that one of the main functions of the RLC is to allocate the information packets in the available RLC frames, with sizes previously defined. If the size of the packet is less than the nearest RLC frame size, it will be necessary to use padding. The RLC frames sizes are defined by the applicable RAB (Radio Access Bearer), selected from a set of available RABs already standardized by 3GPP. Here there is a problem, because it is not practical to define RLC frames in the same number of the multitude of information packets sizes. Therefore, it is important to define a RAB with an adequate number of frames with size values chosen to reduce the waste produced by the padding operation and 3GPP is now working on in this subject. As it will be seen in section 4, we have developed an algorithm to design the RAB frame which can minimize the padding waste.

1) *Simple Model Evaluation:* As we have done before for transmission in the CS mode, we present now an example of performance evaluation of the power efficiency in IMS mode. For the sake of simplicity, let us consider the VoIP transmission during those time intervals when we only have voice packets with a high degree of header compression, simultaneously with voice call signaling packets.

Figure 6 describes this situation. We have channel A carrying voice multiplexed with channel B (like channel D of the CS mode example). The compressed voice packet will have a total length of 40 bytes (320 bits), corresponding to:

- RTP payload in the Bandwidth Efficient mode [14]: 32 bytes;
- Compressed header of the R-0 packet [13]: 1 byte;
- Return Information (ACK) appended in the voice packet [13]: 2 bytes;
- Checksum of the UDP header: 2 bytes;
- RLC header: SN: 1 byte; LI: 2 bytes;

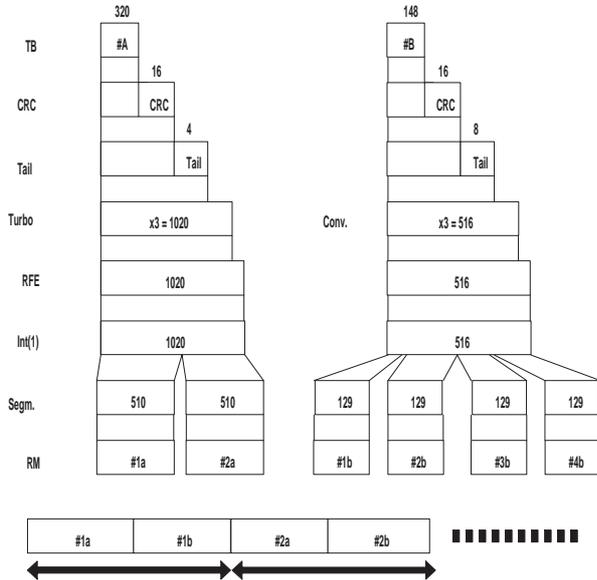


Fig. 6. Transmission of Voice Signal (IMS mode)

It was considered ROHC operating in the Reliable Mode [13].

As we can see in Figure 6, before the rate matching unit we have 639 bits (510+129) to be accommodated in the physical channel. There are two options. The first is to use a spread factor of 32, corresponding to the rate of 120 kbit/s. Then, to fit 639 bits in a radio frame of 10ms (1200 bits) we have 561 bits (1200-639) to use as repetition bits, distributed over the two channels. The second one is to use a spread factor of 64 (as in the CS case), corresponding to the rate of 60 kbit/s. Now we have to use puncturing, for fitting 639 bits in radio frames of 600 bits.

In order to maintain the same rate correspondence with the CS case, the second option has been selected for the comparison, and the corresponding specification is showed in Table II. Note that the BER value for channel A is equal to

the one specified for the most stringent traffic channel in CS mode and for channel B (signaling channel) is the same value specified for channel D (signaling channel) in CS mode. The optimization algorithm, described in [12], in this case allocates 45 punctured bits to channel A and 6 repeated bits to channel B.

TABLE II  
TRANSPORT CHANNELS DATA (PS MODE)

Channel	Coder	BER	Bits	Rep. Bits
A	TC3	$5.10^{-4}$	510	-45
B	CC3	$1.10^{-4}$	129	+6
Total			639	-39

Without optimization, channels A and B require  $E_b/N_0$  values equal to 1.74 dB and 2.76 dB, respectively. After the optimization procedure, the  $E_b/N_0$  value required is 2.53 dB, for both channels. Although there is not much room for an optimization here, the less stringent requirement of channel A allows the puncture of 45 bits, but only six bits are used to improve the performance of channel B, the worst channel, resulting an energy saving of 0.23 dB (2.76 - 2.53). Comparing these results with the previous one for the CS case, we see that the voice transmission in IMS mode requires an increase of 0.65 dB (16,1%) in transmitted power relative to the CS mode case.

2) *Performance with alternative procedures:* Now we evaluate the performance of the simple model presented before, considering the use of some alternative procedures that can increase the power efficiency of IMS transmission:

- RLC operating in the Transparent Mode (i.e. with null header);
- UDP Checksum calculated in the destination (no more sent aggregated to the voice packet);
- Use of Turbo Code also for the signaling channel;
- Use of CRC with 8 bits for voice packets, protecting only the associated headers and calculated by the PDCP;

Figure 7 shows the physical layer components, for this situation. These options are not specifically recommended by 3GPP and are discussed below.

*RLC operating in the Transparent Mode*

In this case, the SN and LI fields are not transmitted. Furthermore, considering the characteristics of RLC when operating in Transparent Mode [2], it can no longer execute the padding operation. The SN information is already contained in packets when ROHC compression is used.

The LI information is important for padding operation, but the proposal is to use the PDCP, having some functions extended for this purpose. In this sense, at the reception of a VoIP packet from the Application Level, the PDCP operation could be described by the following steps:

- Read the value of RTP Payload Number of this first packet;
- If this value indicates Codec AMR, proceed reading the FT (Frame Type) field of the voice packet (the PDCP

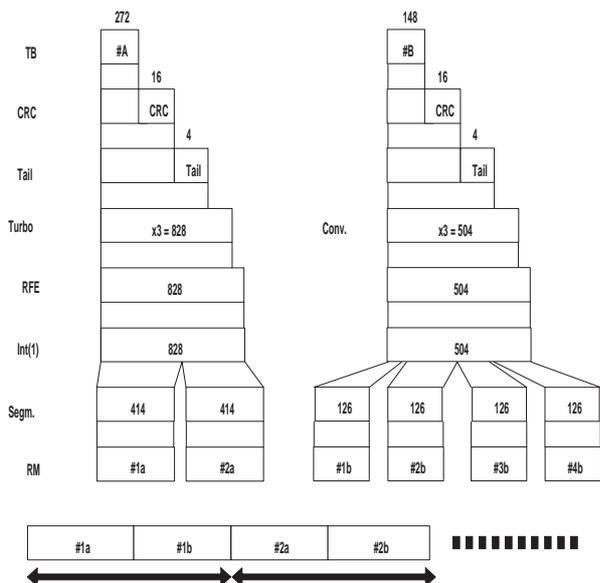


Fig. 7. Transmission of Voice Signal with Optimization (IMS)

must have a table containing, for each type of codec AMR, the sizes of the different compressed packets as well as the frame sizes used by the RLC);

- After consulting the table, calculate the CRC and execute the padding operation;
- Send the packet to the RLC; The same procedure will be applied, in the case of packet reception from the air interface.

*UDP Checksum Removal*

It has been shown [16], [17] that the checksum in the IPv6 header for UDP transmission may be eliminated with no impact in the error performance at the UMTS air interface. Furthermore, it may be noted that the main function of the checksum, in the case of IPv6, is to protect the IP addresses and port numbers. The calculation of the UDP checksum is a very simple procedure, with limited results and only necessary when other error control methods do not exist in the lower levels. This is not the case of UMTS transport system where a CRC is appended at the physical level. In fact, it has been demonstrated in [12] through computer simulation that in a channel with a typical BER the checksum practically does not detect additional errored frames, after the packets have passed the physical layer CRC. Besides, note that the integrity of the whole IP header is strongly protected since in the initial phase of the call, the first packets (that contains the IP addresses and port numbers) are transmitted several times in the air interface (normally seven times, for a RTT of 140ms).

*Use of Turbo Code also for the signaling channel*

The use of the turbo code for IMS mode comes from the fact that this type of coder is more effective than CC3 in terms of the power required for a given BER and employes less tail bits (four bits) in comparison with CC3 (8 bits);

*Use of CRC with 8 bits for voice packets, protecting only the associated headers*

Reference [16] shows that, for the voice packet sizes of interest, it is possible to use CRC-8 instead of CRC-16 without impacting the detection of errored frames. Based on that conclusion and considering also that it is better to send the voice packets with errored payloads, instead of discarding them [17], we suggest to use the CRC-8 for protecting only the following associated packet headers: R-0 header (containing the SN field), CMR (Codec Mode Request) and TOC (Table of Contents). Therefore, only the packets containing errors in those headers would be discarded. Figure 8 shows the format of the R-0 packet, in this condition.

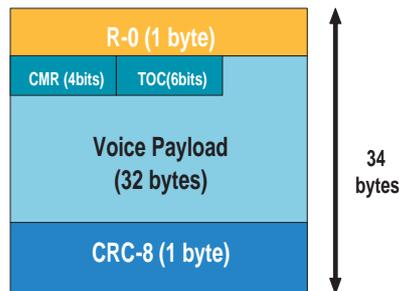


Fig. 8. IMS VoIP Packet Format

With the consideration of the alternative procedures previously described and using the same calculation procedure of the previous section, Table III is obtained. Without rate matching optimization, channels A and B require  $E_b/N_0$  values equal to 1.94dB and 2.87dB, respectively. After the optimization procedure, the  $E_b/N_0$  value required is 1.89dB, for both channels. It is easy to verify the very close results for the required value of  $E_b/N_0$  in CS (1.88dB) and IMS (1.89dB) modes. Note that in both cases the bandwidth used is the same.

TABLE III  
TRANSPORT CHANNELS DATA WITH OPTIMIZATION (PS MODE)

Channel	Coder	BER	Bits	Rep.Bits
A	TC3	$5 \cdot 10^{-4}$	414	+4
B	CC3	$1 \cdot 10^{-4}$	126	+56
Total			540	+60

IV. VOICE TRANSMISSION IN IMS MODE WITH RATE MATCHING AND PADDING OPTIMIZATION

Now we present a numerical performance evaluation of a typical IMS voice call considering the occurrence of all types of packets involved in the call. Here, besides the rate matching optimization, we also utilize a technique to minimize the bit waste in padding operation.

The characteristics of the various packet types involved in a typical IMS voice call may be found in [13] and are listed below (with the respective header's length in brackets):

- IR (61 bytes): used for context initialization;

- IR-DYN (13 bytes): used for updating the dynamic part of the context;
- R-0 (1 byte): transport the six least significant digits of the SN;
- R-0-CRC (2 bytes): sent  $m$  times, since the last ACK received; the operator is free to chose the  $m$  value;
- R-2 (3 bytes): used to change the value of the RTP marker (M) bit;
- R-2-Ext3 (5 bytes): used for signaling of more bits for the Sequence Number (SN) parameter;
- R-0-ACK (3 bytes): used to acknowledge a burst of R-0-CRC received;

All the above packets may contain voice information data but only IR-DYN, R-0, R-0-CRC, R-2-Ext3 and R-0-ACK packets may carry Silence Descriptor (SID) information.

For the performance evaluation, an interval of 10 minutes of conversation has been considered, and the following observations are made to determine the average number of each type of packet in the conversation:

- assuming equally mean time values of 1 second for voice / silence and the transmission of one SID packet at each eight intervals of 20ms during one silence period implies the transmission of an average number of 15000 voice packets and 1875 SID packets;
- considering that, in the beginning of the transmission, an IR packet is transmitted at each 20ms interval during an RTT of 140ms between UE and RNC, then 7 IR packets will be generated;
- considering that the number of R-2-Ext3 and IR-DYN packets are 1% of the total number of transmitted packets, there will be 150 R-2-Ext3-Voice / IR-DYN-Voice and 19 R-2-Ext3-SID / IR-DYN-SID packets;
- the other quantities may be obtained by considering the total number of voice packets involved (15000), the number of all SID type packets (1875) and the parameter  $m=3$  which sets the number of times the R-0-CRC packet is sent, since the last ACK is received; in this case, we have  $(26 - m) = 61$  R-0 packets (plus R-0-ACK) and 3 R-0-CRC packets sent in an alternate fashion; this leads to the following equations:

$$\begin{aligned}
 & - R-0-ACK-Voz = R-0-CRC-Voz / 3 \\
 & - R-0-ACK-SID = R-0-CRC-SID / 3 \\
 & - (R-0-Voz + R-0-ACK-Voz) / (R-0-CRC-Voz) = 61/3 \\
 & - (R-0-SID + R-0-ACK-SID) / (R-0-CRC-SID) = 61/3
 \end{aligned}$$

Once the statistics of packet occurrence has been determined, we could produce a table describing the characteristics of each type of packet after the successive operations of the transport channel procedures, including rate matching and also a padding optimization. As we have mentioned before, it is not practical to define frames with all these different sizes in order to accommodate the packets. Then we have to use padding. Since, as stated before, the RLC must operate in the Transparent Mode, the padding function has to be developed by the PDCP. For the purpose of minimizing the waste, the authors have proposed in [15] an algorithm to allocate the packets into frames in the best way. This algorithm may be used to increase the performance not only of voice calls but

of general multimedia calls in IMS systems. Its fundamentals are presented in Appendix II.

The performance results are presented in Table IV. The total transmitted energy during the call are showed in the last column considering the final number of bits of each packet in the call and the respective  $E_b/N_0$  after the whole processing, including the number of bits to be repeated.

Note that:

- the algorithm for padding optimization [15] yields the values for the column RLC Frame Size;
- Weighted Padding Ratio is the product of the columns Ratio of Appearance and Padding Ratio. The summation of the values of this column (Totals) gives the mean value of padding during the connection. Here it is possible to verify the performance of the algorithm developed (only 2.27% of waste);
- Size Before Rate Matching represents the frame size at the physical level just before the Rate Matching unit. These values can be obtained by the following expression:  $\text{Size Before Rate Matching} = [\text{RLC Frame Size} + 4(\text{Tail})] \times 3/2$
- the column SF & Frame Size gives the values of Spread Factor and the resulting length of the physical frame;
- the column Repeat. of Bits shows the number of bits to be repeated, for each packet type;

In order to make a comparison, Table V gives the results for a 10 minutes CS connection performed under the same conditions.

Comparing the total values of the last columns of Tables IV and V, we see that with the described procedures, the total energy in an IMS VoIP call will exceed the CS call by only 1.5%. It is important to observe that in the previous example we are considering the transmission, in time, of only one information channel for the IMS case and three information channels (carrying the bits of classes A, B and C) for the CS case. Therefore, the padding optimization procedure is applied to both cases but the rate matching optimization process is used only in the CS case.

## V. CONCLUSIONS

Undoubtedly the IMS brings several benefits for the UMTS, especially in the multimedia transmission arena. In parallel we are witnessing the growing of VoIP technology utilization in fixed networks. Then it is natural the ever growing trend to use VoIP also in the cellular networks. However, the air interface imposes severe limitations for multimedia traffic, especially VoIP, if compared with the traditional circuit switching technologies.

This paper showed that, with the adoption of two algorithms developed by the authors for optimization of the rate matching component and to discover the best frames sizes (with consequent minimization of waste in padding) and the introduction of some minor changes in the UMTS recommendations it is possible to turn VoIP performance comparable with the CS counterpart for voice calls.

Because we are interested in make comparisons between voice transmission in the two modes (CS and IMS), all the

TABLE IV  
STATISTICS FOR A VOIP SESSION (12.2KBPS-AMR)

Pckt Type	Pckt Size	Appearance	Ratio Appea.	RLC Frame	Padd. Ratio	Weight. Padding	Size pre RM	Size & Pr. Size	Repeat. of bits	$E_b/N_0$ (dB)	Number TX bits	$E_b/N_0$ TX bits
IR(voice)	752	7	0.00%	752	0.00%	0.00%	1134	32/1200	+66	1.10	8400	10821
IR-DYN(voice)	368	150	0.89%	752	51.06%	0.45%	1134	32/1200	+66	1.10	180000	231885
R-0(voice)	272	13212	78.29%	272	0.00%	0.00%	414	64/600	+186	0.87	7927200	9685450
R-0-CRC(voice)	280	661	3.92%	304	7.89%	0.31%	462	64/600	+138	1.15	396600	516836
R-2(voice)	288	600	3.56%	304	5.26%	0.19%	462	64/600	+138	1.15	360000	469140
R-2-Ext3(voice)	304	150	0.89%	304	0.00%	0.00%	462	64/600	+138	1.15	90000	117285
R-0-ACK(voice)	288	220	1.30%	304	5.26%	0.07%	462	64/600	+138	1.15	132000	172018
IR-DYN(SID)	168	19	0.11%	272	38.24%	0.04%	414	64/600	+186	0.87	11400	13929
R-0(SID)	72	1722	10.20%	80	10.00%	1.02%	126	256/150	+24	2.35	258300	443736
R-0-CRC(SID)	80	86	0.51%	80	0.00%	0.00%	126	256/150	+24	2.35	12900	22161
R-2-Ext3(SID)	104	19	0.11%	272	61.76%	0.07%	414	64/600	+186	0.87	11400	13929
TOTAL						2.27%					9405600	11718449

TABLE V  
STATISTICS FOR A CS SESSION (12.2KBPS-AMR)

Packet Type	Transport Channel	RLC Frame Size	Size Before Rate Matching	Appearance	SF & Frame Size	Repeat. of Bits	$E_b/N_0$ (dB)	Number of TX bits	$E_b/N_0$ x TX bits
Voice	A	81	152	15000	64/600	+95	0.961	9000000	112290037
Voice	B	103	167	15000	64/600	+89	0.961	9000000	112290037
Voice	C	60	68	15000	64/600	+29	0.961	9000000	112290037
SID	A	39	89	1875	256/150	+61	0.539	281250	318414
TOTAL								9281250	11547451

simulations were developed assuming an AWGN channel. In fact,

All the simulations were developed assuming an AWGN channel. In fact, another channel characterized by a distribution of uncorrelated errored bits would serve either, because we are interested only in make comparisons between voice transmission in the two modes (CS and IMS). Consequently, the  $E_b/N_0$  values here indicated are better than those found in real mobile systems, normally subjected to interference and Rayleigh fading.

#### APPENDIX I

##### THE RATE MATCHING OPTIMIZATION

Consider the situation where we have  $M$  transport channels and we wish to transmit  $N_i$  bits per radio frame in the  $i$ -th channel. The sum  $N_T = \sum N_i$  is smaller than  $N$ , the total number of bits that are capable to be transmitted. Hence  $\Delta N = N - N_T$  bits can be selected for repetition, thus reducing the power requirements without sacrificing QoS. We define as  $\Delta N_i$  the number of repeated bits in the  $i$ -th channel. Each channel is also characterized by its bit error probability  $P_i$  that is a function of  $N_i$ ,  $\Delta N_i$  and  $(E_b/N_0)_i$ . QoS demands  $P_i \leq \alpha_i$ , a prefixed value. Since  $N_i$  is known,  $P_i$  depends only on  $x_i = \Delta N_i$  and  $y_i = (E_b/N_0)_i$ . We assume that  $P_i(x_i, y_i)$  is a monotonic decreasing function on both arguments and the following equations should be obeyed:

- $x_i \geq 0$  and  $y_i \geq 0$  for all values of  $i$
- $x_1 + x_2 + \dots + x_M = \Delta N$
- $P_i(x_i, y_i) \leq \alpha_i$  for all values of  $i$

Among all possible ways of defining a quality criterion, one that seems more suitable for the problem in hand, is the minimization of the maximum  $(E_b/N_0)$  value that meets the requirements. This is equivalent to the maximization of the artificial variable  $w$  subjected to  $w \geq y_i$  for all  $i$ . In the optimization theory parlance, we are faced with a  $(2.M + 1)$

dimensional mixed-integer optimization problem with linear objective function but with linear and nonlinear constraints. If we momentarily drop the integrality condition imposed on  $x$  variables and if we employ the well known Karush-Kuhn-Tucker theorem, it is possible to prove that the optimization problem can be solved by an alternative algorithm that does not require conventional optimization tools at all. This algorithm involves only a bisection scheme very simple to implement as far as computer programs are concerned. For more details, please see [12].

We are well aware that in mixed-integer programming, the optimal solution rarely is the truncated or rounded version of the optimal real solution. Nevertheless, due to the smoothness of the function  $P_i(x_i, y_i)$ , all simulations revealed that the error due to truncation was never greater than 2%. Although we cannot strictly speak of optimality, we are fairly confident the results produced this way will not deviate more than few per cent of the optimal solution. It is interesting to observe that a similar algorithm was developed for the case of puncturing (instead of bit repetition), applicable when  $\Delta N$  is negative (i.e.  $N < N_T$ ).

#### APPENDIX II

##### ALGORITHM FOR PADDING OPTIMIZATION

Consider the process of transmitting packets with sizes in bytes belonging to the set  $\Omega = \{x_1, x_2, \dots, x_N\}$  where for the sake of simplicity, we assume that i)  $x_1 > 0$  ii)  $x_{k+1} > x_k$  for all admissible  $k$ . We here adopt a probabilistic approach where we denote by  $\alpha_k$  the probability that a given packet has size  $x_k$ . We wish to determine the sizes (also in bytes) of  $M$  ( $M < N$ ) transmission frames that can accommodate the above packets. We represent these frame sizes by a vector  $\underline{X} = \{X_1, X_2, \dots, X_M\}$  where we also assume that i)  $X_1 > 0$  ii)  $X_{m+1} > X_m$  for all admissible  $m$ . Since  $M$  is smaller than  $N$ , irrespective of the choice made, a certain "waste" is

produced. Frames will carry unused bits that probably will be unnecessarily protected in other stages of transmission. In order to minimize this quantity, packet of size  $x_k$  must be accommodated in frame  $X_m$  if and only if  $X_{m-1} < x_k \leq X_m$ . To avoid mathematical inconveniences, we define  $X_0 = 0$ . Here, this minimal waste is given by  $(X_m - x_k)$  and consequently the total average waste will be given by a sum of terms like  $(X_m - x_k) \cdot \alpha_k$ . Formally speaking, let  $\Omega_m(\underline{X})$  be the subset of all elements of  $\Omega$  that belongs to the range  $(X_{m-1}, X_m]$ . Then the total average waste  $D$  depends on  $\underline{X}$  and it can be expressed as:

$$D(\underline{X}) = \sum_{m=1}^M \sum_{x_k \in \Omega_m(\underline{x})} [X_m - x_k] \cdot \alpha_k \quad (1)$$

which is equivalent to:

$$\bar{D}(\underline{X}) = \sum_{m=1}^M X_m \cdot [F(X_m) - F(X_{m-1})] \quad (2)$$

where  $F(W) = \sum_{n=1, x_n \leq W}^N \alpha_n$

Simple considerations applied to all components of  $\underline{X}$  allow us to state that the optimal solution  $\underline{X}^*$  has all its components belonging to the set  $\Omega$ . Furthermore, it is also easy to observe that  $X_M$  must be equal to  $x_N$ . Hence the real unknowns of this problem are  $X_1, X_2, \dots, X_{M-1}$  and they must be selected from the set  $\{x_1, x_2, \dots, x_{N-1}\}$ . Should brute force method be feasible, we are facing a search among all combinations of  $N - 1$  elements in groups of size  $M - 1$ . Even for moderate values of  $M$  and  $N$ , brute force procedure can be prohibitive. This problem can be circumvented if we notice that our objective function is "additive" and the Bellman's optimization principle can be applied. These observations do indicate that it is possible to define a graph whose shortest path between two specific nodes corresponds to the optimal solution of the problem here discussed. Hence the optimal solution for our problem can be obtained by finding the shortest path between node 0 at level 0 and node  $N - M$  at level  $M$ . This shortest path can be found by the classical algorithm of dynamic programming, also known in the realm of telecommunications, as Viterbi algorithm. For more details please see [15].

ACKNOWLEDGMENTS

Marco Grivet acknowledges fellowship support from CNPq.

REFERENCES

[1] IETF RFC 3261, "SIP: Session Initiation Protocol", June 2002.  
 [2] 3GPP Document, "Radio Link Control (RLC) Protocol Specification", TS 25.322, v6.0.0, December 2003.  
 [3] 3GPP Document, "Network Architecture", TS 23.002 v6.3.0, Dec.2003.  
 [4] 3GPP Document, "Radio Interface Protocol Architecture", TS 25.301 v6.0.0, Dec.2003.  
 [5] 3GPP Document, "UTRAN Iu User Plane Protocols", TS 25.415 v6.0.0, Dec.2004.  
 [6] 3GPP Document, "Packet Data Convergence Protocol (PDCP) Specification", TS 25.323 v6.0.0, Dec.2003.  
 [7] 3GPP Document, "Multiplexing and Channel Coding", TS 25.212 v6.2.0, June 2004.  
 [8] 3GPP Document, "Common Test Environments for User Equipment (UE): Conformance Testing", TS 34.108 v5.1.0, Jun. 2004.

[9] Baey S., Dumas M. and Dumas M.-C., "QoS Tuning and Resource Sharing for UMTS WCDMA Multiservice Mobile", IEEE Transactions on Mobile Computing, vol.1, No. 3, pp. 221-235, July-Sept. 2002.  
 [10] Sohn I. and Bang S.C., "Performance Studies of Rate Matching for WCDMA Mobile Receiver", IEEE VTC 2000, 52nd, vol.6, pp.2661-2665, 2000.  
 [11] Aftelak S. and Bhatoolaul, "Rate Matching Attribute Settings and Error Rate Performance Sensitivity for Selected UMTS FDD Services", IEEE Semiannual Vehicular Technology Conference Final Program, October 6-9, 2003.  
 [12] Vale E., Brandão J.C., Grivet M., "UMTS Rate Matching Optimization", International Association of Science and Technology for Development - IASTED, CSA 2005, Banff, Canada, July 2005.  
 [13] IETF RFC 3095, "Robust Header Compression (ROHC): Framework and four profiles: RTP, UDP, ESP, and uncompressed", July 2001.  
 [14] IETF RFC 3267, "Real-Time Transport Protocol (RTP) Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs", June 2002.  
 [15] Vale E., Brandão J.C., Grivet M., "Padding Optimization in UMTS", IEEE International Symposium on Personal, Indoor and Mobile Radio Communications - PIMRC 2006, Helsinki, Finland, September 2006.  
 [16] Vale E., Brandão J.C., Grivet M., "Reducing the Voice Packet Overhead In UMTS Interface", submitted to IEEE Communications Magazine.  
 [17] 3GPP Document, "UDP Checksum Handling for VoIP Packets", TDOC R2-060371.TSG-RAN2 Meeting 51, Denver, USA, Feb. 2006.  
 [18] Hammer F., Reichl P., Nordstrom T., Kubin G., "Corrupted Speech Data Considered Useful", Acta Acoustica, vol. 90, no.6, pp 1052-1060, Nov/Dec. 2004.



**Eduardo Rodrigues Vale** was born in Rio de Janeiro, Brazil, in 1949. He received the B.S. degree in 1972, the M.Sc. degree in 1975 and the Ph.D. degree in 2006, all of them from the Pontifical Catholic University of Rio de Janeiro (PUC-RIO), Brazil, in Telecommunications Engineering. Presently he is a Professor in the Telecommunications Engineering Department of Universidade Federal Fluminense (UFF). His current research interests include the area of mobile communications, particularly the Third Generation UMTS systems.



**João Célio Brandão** received an MSc. degree in Electrical Engineering from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Brazil, in 1973. Since then he has worked at the Center for Studies in Telecommunications of PUC-Rio (CETUC) where he is presently Professor and Vice-Director. At CETUC, he has been involved with teaching and research in the area of digital transmission and network performance with particular application to satellite and cellular systems.



**Marco Antonio Grivet Mattoso Maia** was born in Rio de Janeiro, Brazil, in 1951. He received the B.S degree in 1974 and the M.Sc. degree in 1977 from the Pontifical Catholic University of Rio de Janeiro (PUC-RIO), Brazil, both in telecommunications engineering. He received the Ph.D. degree in 1983 from the University of Kent, Canterbury, England. He was associate lecturer at the Military Institute of Engineering from 1983 to 1987 and presently is at the Center for Studies in Telecommunications of PUC-RIO. His current research interests include mobile communication, location-based systems, image recovery and visualization and digital image processing.

mobile communication, location-based systems, image recovery and visualization and digital image processing.