PREPRINT VERSION

PERFORMANCE EVALUATION OF UMTS PACKET SERVICES

Flaminio Borgonovo
Antonio Capone
Matteo Cesana
Luigi Fratta

Politecnico di Milano,
Dipartimento di Elettronica e Informazione, Italy

The provision of efficient high speed packet data services is probably the most important challenge for third generation systems (3G).
The efficiency of such services mainly depends on the radio interface.
The UMTS W-CDMA radio interface is characterized by great flexibility and a variety of different physical and logical channel types: for example, on the downlink, the Dedicated Channels (DCH) offer circuit switching, and the Downlink Shared Channel (DSCH) and Forward Access Channel (FACH) use packet switching, the former with closed loop power control. Furthermore, the choice of suitable parameters, such as spreading factors, code rates and ARQ schemes make possible several user rate and protection options. In this paper, we evaluate the performance of UMTS packet data services over dedicated channels and shared channels through detailed simulations. In particular, we study the effect of several parameters (spreading factor, code rate, channel set-up delay, etc.) on the system capacity by means of the delay-throughput curves and show that the setting of parameters may be critical with respect to system performance and stability. Finally, in order to improve the quality of the offered packet service, we propose a simple but effective flow control mechanism which is able to limit interference on the shared resource.

1. INTRODUCTION

The recent success of mobile cellular systems has completely changed the way in which users access telecommunication services. However, even if cellular systems are now widely adopted for telephone services, the number of users who use them to access mobile data services as well remains very low. This is due to the technological limitations of bandwidth and service flexibility of the second generation systems. As an example, the GSM (Global System for Mobile communications) only provides low rate circuit switched data services unsuitable even for web or e-mail applications due to high latency times and high costs.

To fill the gap between user needs and offered services, second generation systems are being enhanced to include packet data services, such as GPRS (General Packet Radio Service), which allow a more flexible use of radio resources and higher peak rates [1]. The commercial introduction of these new services (usually referred to as 2.5 generation services) is coming while the standardization of third generation systems is being completed. This is a further challenge to third generation systems since they must prove their ability to provide access for a wide range of multimedia applications and services in an efficient and cost effective way.

Within the third generation systems (3G), 3GPP, the Third Generation Partnership Project [5], has standardized the Universal Mobile Telecommunications System (UMTS) [2, 3, 4], which ITU (International Telecommunication Union) considers among the standards for the IMT-2000 (International Mobile Telephone standard 2000) family [6, 7].

One of the main advantages of UMTS with respect to the second generation is the capability of providing radio access to multimedia services. The traffic being transferred within 3G mobile networks can be composed of different information flows with various constraints on the required QoS (bit rate, delays, etc...). In order to do the job, the radio interface of UMTS is characterized by great flexibility and a variety of different physical and logical channel types [8]. For instance, several user rates and channel code protections can be selected by choosing suitable parameters, such as spreading factors, FEC (Forward Error Correction) rates, and ARQ (Automatic Repeat reQuest) schemes. This approach is quite different from the one adopted by second generation systems, where a small set of services can be implemented by vendors and provided by operators.

If from one side, this added flexibility is an advantage of UMTS, from the other, it makes the real service implementation more complex. In fact, the complexity of detailed services definition and system parameters optimization has been moved out of specifications and left to UMTS vendors and operators.

Among the new services offered by UMTS, the packet data service is one of the most critical and challenging mainly because of the characteristics of the code division access scheme adopted at the radio interface [9]. Up to now, only a few studies that thoroughly investigate the effects of the different parameters settings on the performance of UMTS data services have been published [10, 11]. In this scenario, the optimization of radio interface parameters is of utmost importance, both in circuit switched service and in packet switched ones [12].

In this paper, we discuss the topic of packet service and present a detailed simulation study of the UMTS downlink segment. After a brief overview of UMTS radio interface, aimed to provide the basics to readers unfamiliar with the system (Section 2), we present, in Section 3, the problems related to the implementation of packet data services. In Section 4, we describe the system model adopted for simulations and introduce the flow control mechanism to prevent the system from working in an instable region. In Section 5, we discuss the performance results obtained for packet service both on shared and dedicated channels. We present our conclusion in Section 6.

2. UTRA BASICS

In January 1998, ETSI, the European Telecommunications Standard Institution, selected two access schemes for the UMTS radio interface, the W-CDMA (Wideband CDMA) scheme and the TD-CDMA (Time Division CDMA), to be used respectively in the paired part of the spectrum assigned to UMTS, 60 MHz from 1920 to 1980 MHz (uplink) and 60 MHz from 2110 to 2170 MHz (downlink), and in the unpaired part, 35 MHz from 1900 to 1920 MHz and from 2010 to 2025 MHz, respectively.

The W-CDMA scheme adopts a QPSK modulation and a chip rate of 3,840 Mchip/s for the in-phase and quadrature channels. It presents a carrier separation of 5 MHz, so that up to 12 carriers can be defined in the available bandwidth. For the downlink direction, a QPSK modulation is adopted after spreading, while for the uplink direction, the in-phase and quadrature channels are used to transmit two BPSK flows with different spreading codes [13].

The spreading process is based on two codes, namely the spreading code and the scrambling code. The spreading code increases the flow bit rate to the chip rate of the air interface according to the Spreading Factor (SF). Different values of SF ranging from 4 to 512 are available and are obtained using a tree of orthogonal Hadamard-Walsh codes. The tree has the characteristic that two codes, even with different SF, are orthogonal if they are located in a different branch of the tree. Multiple trees can be generated using a scrambling code which shuffles the order of chips. The channels transmitted by the same station (base or mobile) can use codes in the same tree so that the mutual interference is greatly reduced, while channels transmitted by different stations should use different scrambling codes so that the mutual interference is almost independent of the delay offset at the receiver due to different propagation paths.

Physical channels are defined by the associated spreading and scrambling codes. The bit flow is divided into time-slot (666 μs). During a time-slot, both physical control bits and data bits can be transmitted, and, while the number of chips is fixed, the total number of information bits depends on the SF. The minimum transmission unit offered by the physical layer to the upper layers is the Transmission Time Interval (TTI), also called frame in the following, and is composed of multiples of 15 slots.

The transport services provided by the physical layer to the upper layers are based on transport channels which are mapped into the physical channels [14]. Transport channels are divided into dedicated and common channels. Dedicated Channels (DCH) are used to transmit user and control information in the uplink and downlink direction and are devoted to the connection between a single mobile station and the UTRA Network (UTRAN). A DCH is mapped into two physical channels, namely the DPDCH (Dedicated Physical Data Channel) and the DPCCH (Dedicated Physical Control Channel). The DPDCH carries user data while the DPCCH carries physical signaling used to control the channel. In particular, in the DPCCH, the pilot symbols are transmitted for channel response and interference estimation, the TFCI (Transport Format Combination Indicator) symbols describe the format adopted for the channel (mainly SF and error protection code), the TPC (Transmit Power Control) symbols are used to transmit the commands for the closed loop power control algorithm, and the FBI (FeedBack Information) symbols are used to perform a closed-loop signal-quality control. In the downlink direction, the DPDCH and the DPCCH are time multiplexed into each slot, while, in the uplink direction, they are transmitted into the in-phase and quadrature channels.

Common transport channels are used to transmit both control and user information. Among the channels mainly used for control are the BCH (Broadcast Channel), used to broadcast system information, the PCH (Paging Channel), used to transmit downlink control information into a location area, and the SCH (Synchronization Channel) used for mobile synchronization control. The RACH (Random Access Channel) and the FACH (Forward Access Channel) can be used in each cell to transmit control information or packets in uplink and downlink direction, respectively. Finally, the CPCH (Common Packet Channel) in the uplink and the DSCH (Downlink Shared Channel) in the downlink are used for packet transfer only.

The physical layer directly controls the transmission power of physical channels [15]. The power control exerted on the DCH is based on a closed-loop signaling, outlined in the following. The TPC symbols in each slot carry a command for increasing or decreasing the transmission power in each direction. The power step is fixed and is basically equal to 1 dB. On the receiving side, if the estimated SIR (Signal-to-Interference Ratio) after despreading is lower than a SIR target value, an increase command is sent in the subsequent slot. A decrease command is sent otherwise. The SIR target value is controlled by an outer control loop which is based on the quality of the decoded bit flow. Other channels adopt different power control mechanisms. As an example, the DSCH is not directly power controlled, but its transmission power is computed on the basis of the power of the DCH associated to the user actually transmitting on the DSCH. This implies that each mobile station involved in the DSCH packet transmission has an active DCH.

The information received by higher layers can be protected by the physical layer using FEC (Forward Error Correction) codes [16]. The basic coding schemes use convolutional codes with rate 1/3 or 1/2, or a turbo code with rate 1/3. Different rates can be obtained using the rate matching process, which can increase the code rate by means of puncturing.

On top of the physical layer in the user plane, the link layer is split into the MAC (Medium Access Control), the RLC (Radio Link Control) and PDCP (Packet Data Convergence Protocol) [17]. The MAC layer [18] provides logical channels to the RLC and maps logical channels into transport channels. On common channels, the MAC provides addressing of user equipment and scheduling of PDUs (Packet Data Units). The RLC layer [19] can offer an Acknowledged or Unacknowledged data transfer mode. With the Acknowledged mode, it adds control information on each transmitted PDU and an error check on each received PDU. If the unacknowledged mode is selected, erred PDU are simply discarded, while, with the acknowledged mode, an ARQ (Automatic Repeat Request) mechanism is adopted. Finally, the PDCP maps each network layer instance into one RLC entity and performs higher layers header compression, if required (for example TCP/IP header compression).

3. DOWNLINK PACKET DATA SERVICES

As mentioned in the previous section, the provision of transport services at the radio interface of UMTS systems is a complex task since many possible configurations are available and a large number of parameters are involved.

In UMTS W-CDMA radio interface, the packet data service can adopt dedicated or shared channels for the downlink direction [8]. Dedicated Channels (DCH) are assigned to single users through set-up and tear down procedures and are power controlled according to a closed loop mechanism that adjusts transmission power to keep the SIR at a target value. The common channels time multiplex packets of different users on the same physical channels. Two different common channels are available for packet transmissions: DSCH (Downlink Shared Channel) and FACH (Forward Access Channel). To access the DSCH, users must have an associated active DCH on the downlink whose power control mechanism is also used to control the power of the shared channel itself. The FACH is shared by many users to transmit short bursts of data, but, unlike DSCH, no closed-loop power control is exerted, and no associated DCH is needed.

A further packet service named High Speed Downlink Packet Access (HSDPA) is currently being standardized by the 3GPP for release 5/6 of the UMTS. The new service will include features such as: adaptive modulation and coding schemes, hybrid ARQ, fast cell selection algorithms, and MIMO solutions, which will enhance the spectral efficiency of the downlink packet access with respect to the older release of the standard.

Since for each one of the above channels and services different combinations of spreading factor and code rate can provide the bandwidth and the protection required for different services and environments, the problem to select the best choice arises.

Well known results for real-time circuit traffic show that CDMA with closed-loop power control is very effective in spectrum exploitation [21]. Using powerful FEC codes, which have proven more effective than spreading codes, can further enhance efficiency [22]. Closed-loop power control is able to track interference variations, thereby stabilizing the BER (bit error rate) and optimizing CDMA performance [23].

With packet data traffic, the system may behave very differently, and the effect of error control schemes and power control mechanisms is not easily predictable. Due to bursts of intermittent traffic, interference levels are highly variable. When the power control fails due to interference variations or power limitations, the SIR usually decreases to unacceptable values and introduces bursts of errors into the bit stream. Even powerful low rate FEC codes can hardly cope with long error sequences. The ARQ procedure, which is usually more effective with burst errors than FEC, can retransmit the lost packets [24, 25]. However, retransmissions increase channel traffic and interference for users with unfavorable propagation conditions, therefore leading to further potential loss. This introduces a positive feedback which can make the system behavior unstable, depending on traffic and propagation conditions.

The aim of our investigation, presented in the remaining part of this paper, is to provide a quantitative evaluation of the performance of the different alternatives outlined thus far and to support the UMTS Radio Resource Management (RRM).

4. SIMULATOR DESCRIPTION

The software we have developed in C++ emulates a real UMTS system with 49 hexagonal cells organized in a wrap–around domain to avoid border effects in the interference calculation. Each cell has an omni-directional antenna with unit gain located at the center.

4.1. Propagation model

According to the guidelines of ETSI [26], the received power P_r is given by

$$P_r = P_t 10^{\frac{\epsilon}{10}} L, \tag{1}$$

where P_t is the transmitted power, and L is the path loss, $10^{\frac{\epsilon}{10}}$ accounts for the loss due to slow shadowing, with ϵ a normal variety with zero mean and σ^2 variance.

In the following, we refer to a macro-cellular environment for which the cell radius is 300 m and the path loss L is expressed as

$$10 \log L = -(128.1 + 37.6 \log r)(dB),$$

where r (in kilometers) represents the distance between the mobile and the base. We assume no fading as well as a shadowing standard deviation equal to 5 dB.

When a new user is generated, its position is chosen at random over the torus surface, and the path losses for the radio links toward all BSs are determined. The user is assigned to the BS with the minimum attenuation (Best Site). We are not considering user mobility at this stage.

Each cell is assigned a single tree of orthogonal variable spreading factors (OVSF) so that channels in the same cell are always orthogonal. The loss of orthogonality of the received signals due to the multipath effect is accounted for in the receiver model, as specified in section 4.4.

4.2. Traffic model

We have adopted the model for Web traffic proposed by ETSI as a common framework for evaluation studies of UMTS [26]. In such a model, users become active according to a Poisson Point Process of intensity λ, as described in [10]. Each user, upon activation, generates a flow of packets whose length is negative exponentially distributed with a mean 3840 bit. The packet flow is composed of a number of packets, geometrically distributed with mean $N_p = 25$, and packets arrive according to a Poisson Point Process whose intensity is chosen to match a given source speed.

A user leaves the system as soon as the last packet has been transmitted.

4.3. Transmission model

The packets destined to each user are delivered to the RLC layer at the UTRAN (UMTS Terrestrial Radio Access Network) side, where they are processed before being queued for transmission. In details, the RLC layer offers an Acknowledged Transfer service subdividing the upper layer packets in transmission blocks, RLC Service Data Units (RLC–SDU), and adding a 16 bit header that contains the information useful for ARQ purposes [19].

The RLC blocks are delivered to the MAC (Medium Access Control) layer, where they are buffered without any header addition (Transparent MAC Mode), as far as transmission over DCHs is concerned. In DSCHs transmission, the MAC layer adds some overhead information (Non Transparent MAC Mode), basically containing users' identifiers [18]. At each frame, the MAC layer sends a number of blocks up a number of blocks up to the physical layer to fill the space available in the frame. Before transmission starts, the physical layer adds redundancy bits due to the Cyclic Redundancy Check procedure and to the adopted coding scheme [16]. The number of MAC–PDU contained in each physical frame depends on the speed of the physical channels (the spreading factor) and on the code rate of the FEC code adopted.

UMTS supports several coding schemes. In some cases, the rate matching procedure can modify the size of the transmission block at the physical layer in order to fill exactly the TTI with an integer number of blocks. reducing the code redundancy bits with puncturing techniques or adding further bits.

Our simulator adds the parity bits required by Convolutional Codes standardized by the 3GPP, with 256 states, Constraint Length $K = 9$. Optimal puncturing, whose Bit Error Probability (BER), obtained through link level simulations [27], is shown in Fig. 1. In particular, we have considered code rates, spreading factors and block sizes such that the bits introduced by rate matching, which add overhead without increasing error protection, are very few.

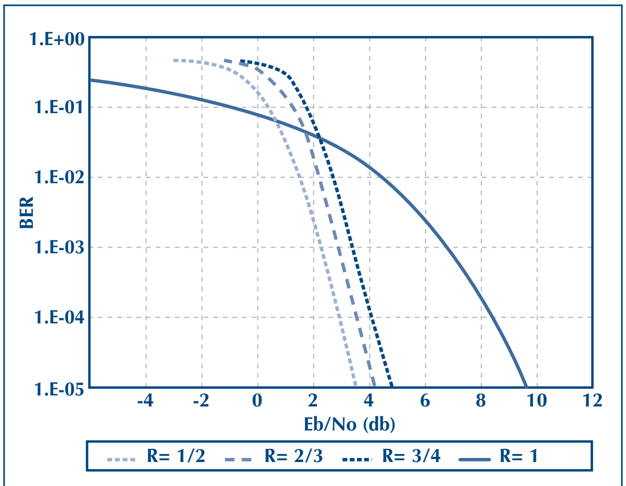

FIGURE 1: BIT ERROR RATE OF THE CONVOLUTIONAL CODES ADOPTED IN UMTS AS FUNCTION OF THE BIT NORMALIZED ENERGY.

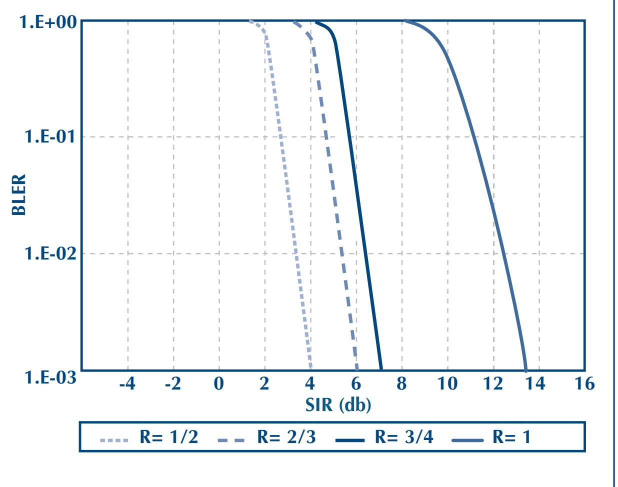

FIGURE 2: BLOCK ERROR RATE OF THE CONVOLUTIONAL CODES ADOPTED IN UMTS AS FUNCTION OF THE MEASURED SIR.

4.4. Receiver model

At the receiving side and at each transmission block, the carrier to interference ratio is evaluated for each transmission as

$$\frac{C}{I} = \frac{P_r}{\alpha I_{intra} + I_{inter} + P_N}, \qquad (2)$$

where P_N is the thermal noise assumed equal to –99 dBm, I_{inter} is the sum of the signal powers received from the other cells, Iintra is the sum of the signal powers received from other users in the same cell, and α is the loss-of-orthogonality factor that, according to [28], is assumed equal to 0.4. All the received powers are obtained by eq. (1).

From the *C/I* evaluated as in (2), the normalized energy per information bit is obtained as

$$\frac{E_b}{N_0} = \frac{1}{2R} \times \frac{SF}{I} \times \frac{C}{1} \qquad (3)$$

where *R* is the coding rate. For each value of the ratio E_b/N_0, the curves in Fig. 1 give the corresponding BER values which allow us to obtain the BLock Error Rate (BLER) as

$$\mathrm{BLER} = 1 - (1 - \mathrm{BER})^l, \qquad (4)$$

l being the transmission block length.

By testing the value of a normalized random variable against the BLER, we are able to decide the correctness of the transmission. Fig. 2 shows the BLER of blocks of 750 bits versus the SIR after despreading, which is defined as SIR=SF x C/I.

Our simulator assumes an ideal ARQ procedure, i.e., the transmitted block is kept in the transmitting queue in case of error and is canceled otherwise. After 10 failed transmissions, the block is dropped, and the user is declared in outage.

4.5. Power control model

The closed loop power control mechanism adopted for DCHs uses two loops. The inner loop controls the transmitted power to maintain the SIR at the target value, whereas the outer loop controls the SIR target to provide a target BLER. This last control mechanism has been envisaged to provide different qualities to different services. In our simulation, where we investigate a service at a time, the corresponding BLER requirement can be assumed constant; therefore, we have implemented the inner loop only, treating the BLER and consequently the target SIR.

Though the UMTS specifications state that power-update requests of ±1 dB are transmitted every time slot (0.666 ms), we have assumed transmit power updates every frame (10 ms) to simplify and speed up the simulator. To meet the dynamic of the real mechanism, the power updates are in the range of ±15 dB. This simplification allows all blocks in a frame, which belong to the same user, to be transmitted at the same level. However, no impact is expected on the convergence of the mechanism though some slight differences in the number of retransmitted blocks are possible. Therefore, in our model, DCH power updates, limited within the ±15 dB range, are requested at each frame, depending on the difference between the SIR target and the SIR evaluated on the last frame.

Physical power constraints are also added, as specified in [28]. No channel may exceed a transmitted power of 30 dBm, whereas the overall power transmitted by a base station is limited to 43 dBm. The constraint on the channel power is first enforced by setting the transmission power of channels exceeding the limit to 30 dBM and then checking the base station power constraint. If it is not satisfied, the transmission power of all channels is proportionally reduced to obtain a global transmitted power equal to 43 dBm.

4.6. The Flow Control Mechanism

For several values of system parameters, we have observed throughput instability both for DSCH and DCH service. Instability occurs when the mean interference level increases and the power control drives the transmitted power to saturation. In such a condition, the power control cannot prevent transmission errors since it is forced to request more power than what is available. As a consequence, the curve of throughput versus *G*, the average fraction of frames used for transmissions and retransmissions shows a convex behavior and decreases after reaching a maximum in G < 1.

In unstable systems, for any value of the offered traffic, the traffic G increases in time until it reaches $G = 1$, yielding a throughput lower than the maximum possible.

Since the main cause of instability is the high interference and since we are considering packet date service, we can adopt an additional flow control (FC) mechanism in order to adjust the load on the active channels dynamically, therefore generating the average interference, according to traffic and propagation conditions. Such an FC scheme must be able to limit the transmission rate only when necessary, reducing the intensity and frequency of interference peaks.

We have proposed and implemented a dynamic FC mechanism which is based on a feedback provided by mobile terminals in our simulator, both for the DSCH service and the DCH service. The basic idea is to reduce the transmission rate on the channel when one or more consecutive transmissions fail. However, since a transmission can also fail due to interference variations and not due to a real traffic congestion (we could say interference congestion), the mechanism is triggered only when a transmission is performed at the maximum power. More precisely, a flow-control counter (N_i) is started and increased at each transmission failure if the maximum transmission level is used. The counter is set to 0 whenever the transmission is successful. Any erred packet is retransmitted after a number of frames K randomly chosen between 1 and the value of the counter N_i. In such a way, we control the traffic on the channels (G) and consequently limit the mean interference level. To notify the transmission failure/success to the transmitter, one of the FBI (Feedback Information Bits) bits defined in the uplink transport block format [14] is used. Fig. 3 and 4 report the flow chart of the FC mechanism.

5. SIMULATION RESULTS

Due to the complexity of the overall system and the interaction among parameters and performance variables, a simple and straightforward analysis is impossible. Thus, we have been forced to split the study into several sub-problems.

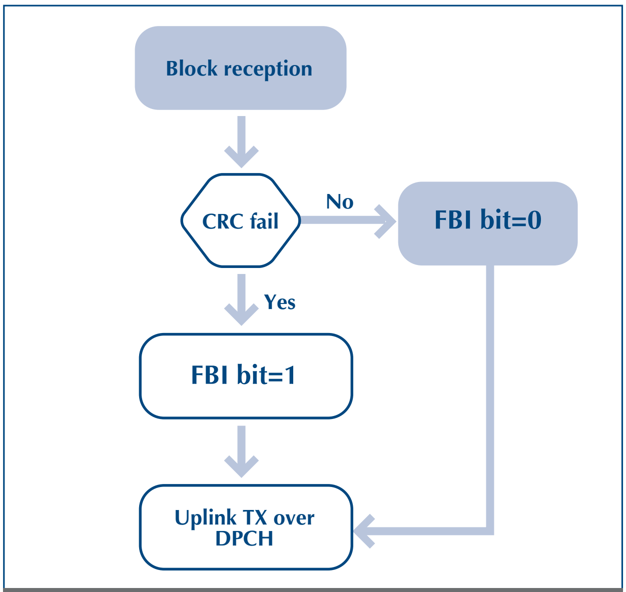

FIGURE 3: FLOW CONTROL SCHEME MOBILE EQUIPMENT SIDE.

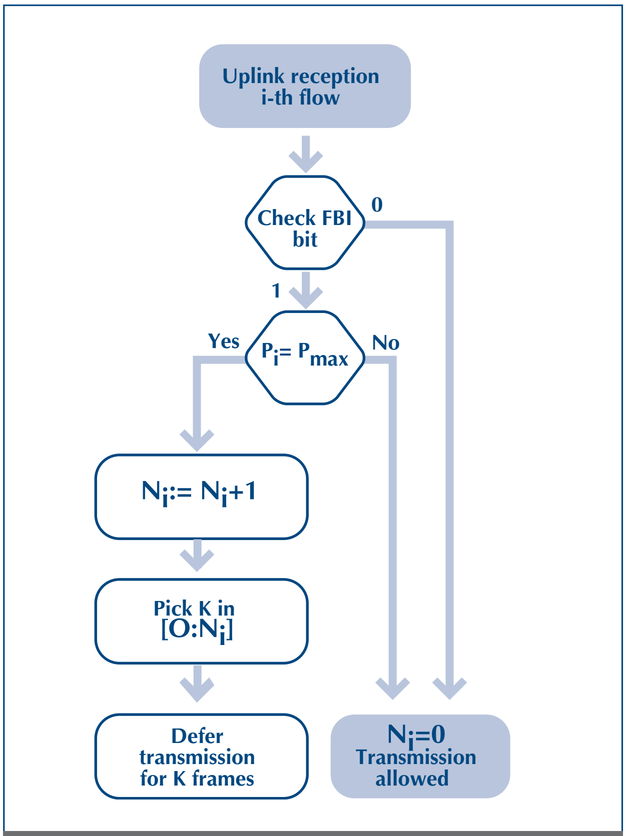

FIGURE 4: FLOW CONTROL SCHEME MOBILE BASE STATION SIDE.

Section 5.1 reports the simulation analysis of the shared service performances. In particular, in section 5.1.1, we discuss the impact of different FEC codes on the performance of a single PDSCH per cell, while, in section 5.1.2, we search for the maximum throughput achievable when the use of multiple PDSCH is allowed. In Section 5.2, the performances of a DCH service are evaluated as a function of the physical channel speed (Section 5.2.1) and of the set-up delays (Section 5.2.2).

We have obtained all the presented results by running steady-state simulations 600 seconds long [29]. We use the first 100 seconds as warm–up time, that is to say, we neglect the statistical results collected during this period. We divide the remaining 500 seconds into 5 simulation runs. During each run, we collect the results and use them to calculate one sample of each statistical quantity used for evaluation before checking the output results according to the t-student statistical test. For all the measures reported in the following (throughput, BLER, etc.), the confidence interval is under the 5%, given a confidence level of 95%.

5.1. DSCH Performances

5.1.1. Effect of Codes

Fig. 5 shows the average delay suffered by a packet versus that suffered by the throughput when one PDSCH is adopted with SF=4 and for protecting codes with R= 1, 3/4, 2/3, 1/2. In all cases, we have adopted the FC mechanism described above. We define the throughput per communication as the ratio between the data bits correctly received (L1/L2 - overhead free) and the simulated time. We obtain the best performance with light codes (R= 3/4 and R=2/3). Heavier codes (R=1/2) achieve a poorer performance since the added redundancy bits reduce the throughput without yielding any benefit from the excess protection. The low throughput measured in the case of no error correction (R= 1) is due to the high interference and shows that the protection of the spreading process with SF= 4 is not sufficient to fight interference.

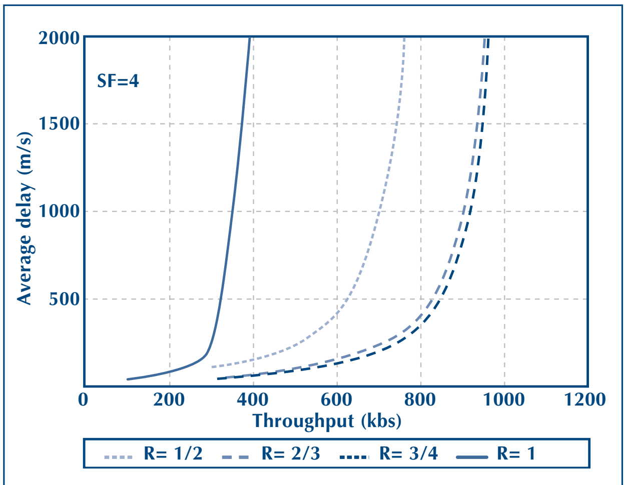

Figure 5: Average delay versus throughput for SF=4 and different code rates with the basic traffic model.

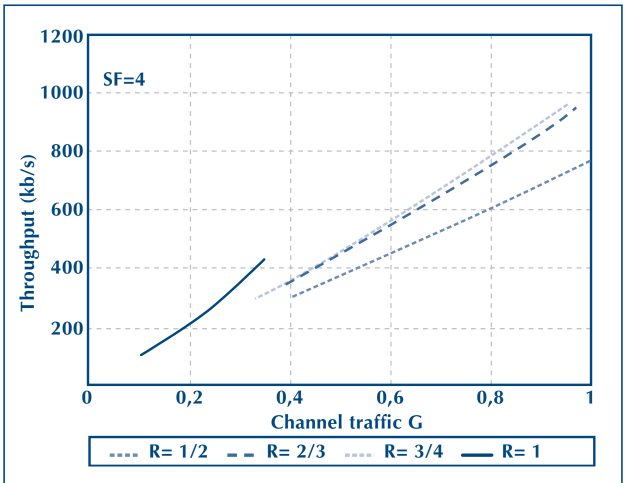

Figure 6: Throughput versus the channel traffic G for the cases reported in Figure 4.

For a deeper understanding of the curves shown, we need to consider additional performance figures figures, such as those shown in Fig. 6, 8, 7. These figures show, respectively, the throughput, the BLER, and the average fraction of transmission at the maximum power (saturation fraction) versus the channel traffic *G*.

In the *R* = 1 case, we have adopted a SIR target of 13 dB, since Fig. 2 shows that this SIR value provides a BLER low

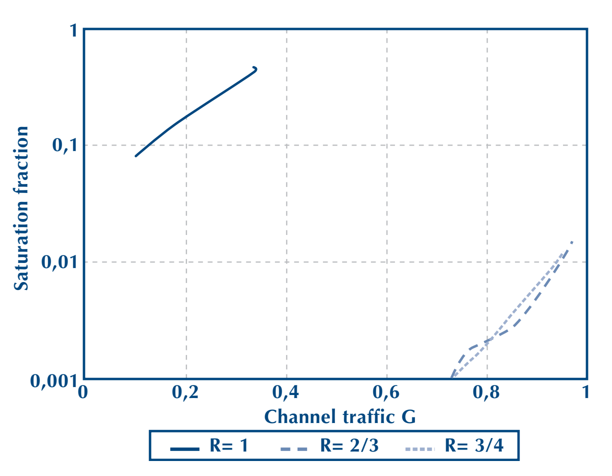

Figure 7: Throughput (kb/s) Fraction of user in saturation versus the channel traffic G for the cases reported in Figure 4.

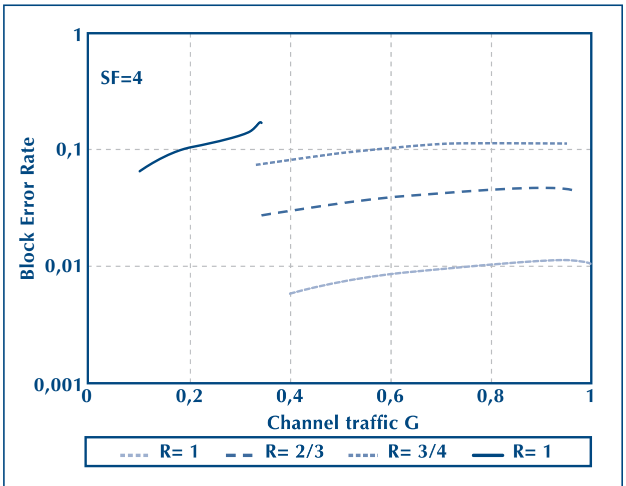

Figure 8: Block Error Rate versus the channel traffic G for the cases reported in Figure 4.

enough to be suitable for ARQ operation. Unfortunately, due to the power limits, the power control is not able to reach this value at high loads, and the BLER results are indeed much higher.

Fig. 7 explains such an effect. As channel transmissions *G* increase from small values, the power control saturates many sources, causing a sharp increase in the BLER (Fig.8). At this point, the FC mechanism intervenes and drastically limits the maximum *G* as shown in Fig. 6.

Since the high SIR target is responsible for the bad performance at *G* = 1, to increase throughput, we must adopt correcting codes that require a lower SIR. Referring again to Fig. 2 we see that, with an R = 3/4 code, a 9 dB SIR is enough to guarantee a very low BLER. Even in this case, the performance is too poor in *G* = 1, and the FC mechanism intervenes limiting *G* to 0.955. Similar behavior shows the code *R* = 2/3, though the *G* limit is further increased to 0.97. The delay curves for the two latter cases overlap almost perfectly, showing the ability of the FC to keep the system very close to capacity.

Although in the latter cases, the reduced SIR target reduces the saturation fraction and the BLER so that *G* can be increased almost to one, the obtained BLER is still much higher than what Fig. 2 predicts. The reason is that the power control mechanism is not able to keep the SIR constant when traffic and interference are bursty. In our simulations, we have measured SIR standard deviation values in the range 3.7 - 4.3 dB (note that we evaluate all interference and SIR statistics throughout the paper considering logarithmic values). As a consequence of this behavior, the errors introduced by the radio channel are not independent and occur in bursts when the SIR is too low.

With more powerful codes, (R= 1/2), *G* = 1 can be reached, and the saturation fraction and the BLER is further reduced, thanks to the further reduction in the SIR target. However, the benefits of the reduced number of retransmissions do not compensate the loss of throughput due to the increased code redundancy. Furthermore, the maximum observed throughput is remarkably less than that obtained with R= 2/3. We have observed a similar effect by increasing the protection with spreading factors higher than SF= 4; the corresponding results are not shown for the sake of brevity.

In Figures 9 and 10, we show the effect of the SIR target on the maximum throughput and the BLER, respectively. If we choose a SIR target that is too small, too many errors occur since the SIR fluctuations around the target value quite often drive the

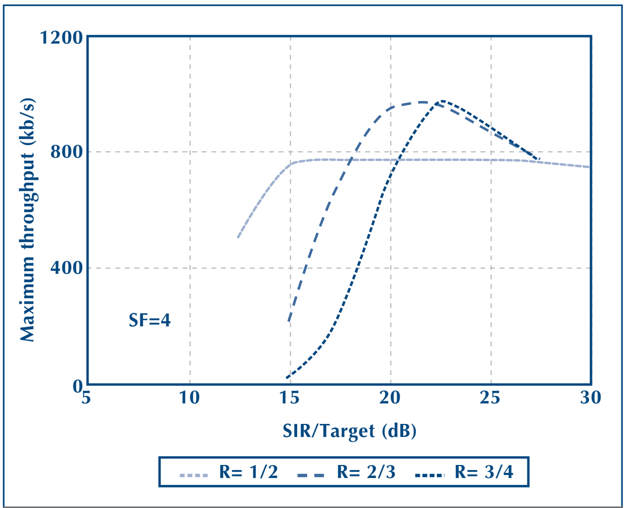

FIGURE 9: MAXIMUM THROUGHPUT AS FUNCTION OF THE SIR TARGET FOR THE CASES R = 3=4; 2=3; 1=2.

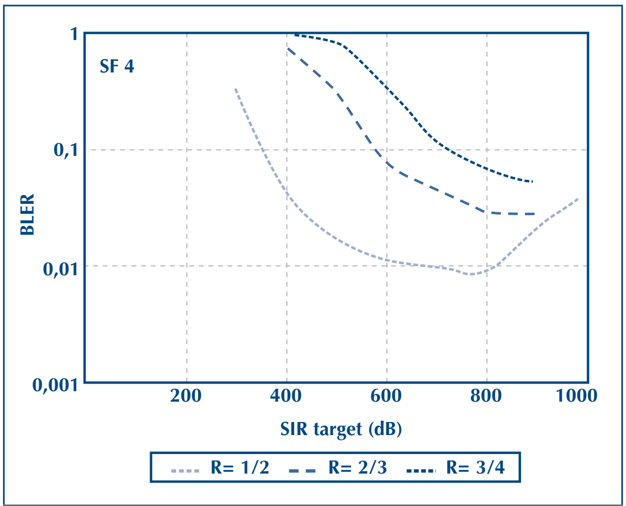

FIGURE 10: BLOCK ERROR RATE AS FUNCTION OF THE SIR TARGET FOR THE CASES R = 3=4; 2=3; 1=2.

system in a condition where the code protection is useless. On the other hand, with a high SIR target, the power requirement increases. Too many transmissions tend to be driven to saturation with a consequent increase of the BLER. For each code, an optimum value of SIR target exists. In our simulation, we have assumed these optimal values, namely 7 dB for codes with R= 3/4 and R= 2/3 and 6 dB for the code with R= 1/2.

5.1.2. MULTIPLE PHYSICAL CHANNELS

While in the previous sections we have considered a DSCH mapped into a single physical channel (PDSCH), we now investigate the performance of the system when using several concurrent PDSCHs in each cell. First we consider the case BLER in which we obtain the same overall channel rate by doubling both the number of channels and the SF. If the two channels with SF= 8 were orthogonal at the receiver, the system would achieve the same capacity as the single channel with SF= 4 and should show an increased average delay because of the lower speed of the channels.

Since in our model we adopt an orthogonality loss, $\alpha = 0.4$, as suggested in [26], one would expect a capacity decrease, since each new channel adds some intra-cell interference.

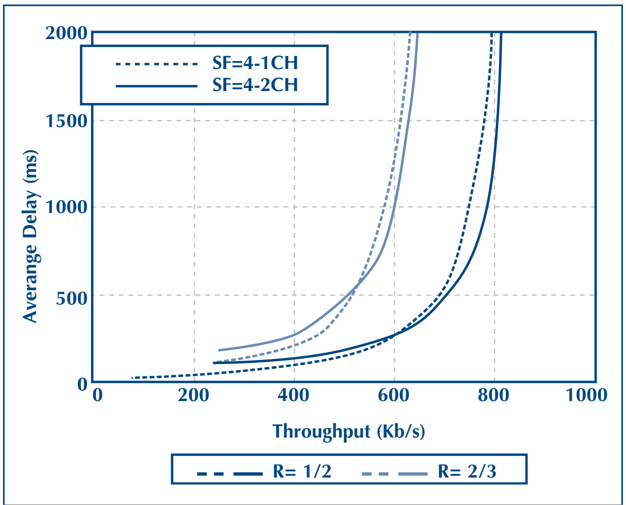

FIGURE 11: COMPARISON OF THE AVERAGE DELAYS ACHIEVED WITH TWO CHANNELS AND DOUBLED SPREADING FACTOR FOR DIFFERENT CODE RATES.

On the contrary, we have measured that two channels with SF= 8 always provide a higher throughput than the single channel with SF= 4, as shown in Fig. 11 for R= 2/3, 1/2 (the case R= 3/4 overlaps with R= 2/3). This unexpected behavior is due to two main reasons. First, the use of slower channels achieves higher frame filling degree (0.98 instead of 0.95 for R = 1/2). Second, the power control works better as long as the traffic burstiness is reduced, having doubled the packet transmission time. This is

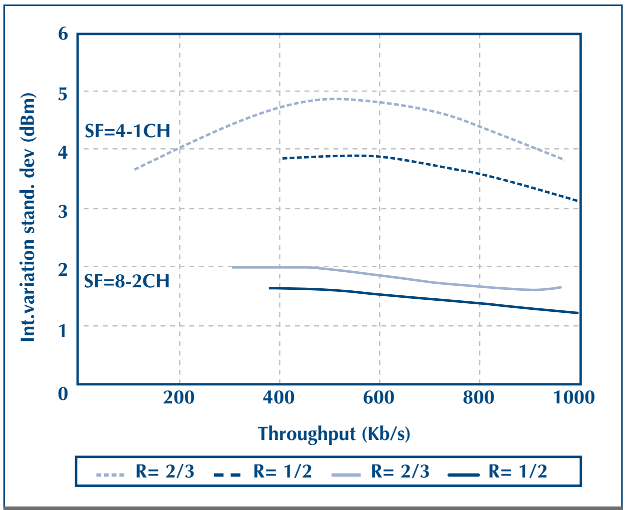

Figure 12: Standard deviation of the interference changes in two consecutive frames for the cases reported in Figure 11.

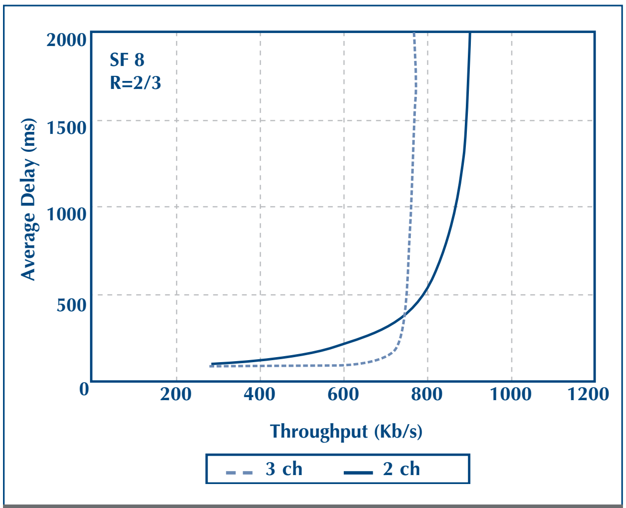

Figure 13: Average delay versus throughput when a different number of physical channels are used in parallel for SF= 8 R = 2=3.

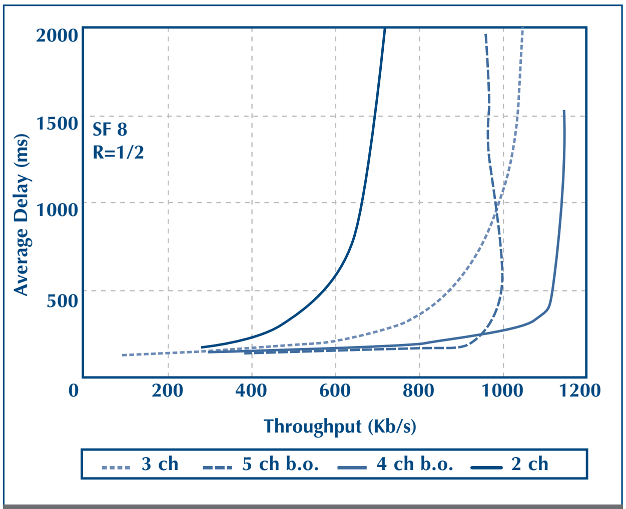

Figure 14: Average delay versus throughput when a different number of physical channels are used in parallel for SF= 8 R = 1=2.

proved by the reduced standard deviation of interference variation that occurs in adjacent frames measured in the case of 2 PDSCHs (Fig. 12). We have observed that a further splitting of the channel speed no longer improves the maximum throughput since the de-orthogonalization effect prevails.

We now investigate the maximum throughput as function of the number of physical channels.

For the code R = 2/3, no further improvement is achieved by increasing to 3 PDSCHs, as shown in Fig. 13. In fact, in this case, the increase in the interference prevails. The FC mechanism limits the channel traffic G, but, when it intervenes, it sharply increases the variance of the interference, which causes a reduction in the maximum achievable throughput.
On the contrary, with R = 1/2, the performance is significantly improved (Fig. 14) by using more PDSCHs. The case of 4 channels provides the maximum throughput (1240 kb/s) among those examined, with a SIR target equal to 4 dB, which, in this case, has shown to be optimal.

5.2. DCH Performances

As outlined in section 2, UMTS allows us to set transport channel parameters in a flexible way in order to optimize system performance according to traffic characteristics and interference conditions. The dedicated channels are the same as those adopted for circuit switched services such as voice, but, with packet service, their behavior changes significantly due to high variable interference.

As in the shared service 5.1.1, the optimal value of the SIR target for the closed loop power control procedure comes from a trade-off choice. Too small a SIR target is chosen, too many errors occur since the SIR fluctuations around the target value often drive the system into a condition where the code protection is useless. On the other hand, with a high SIR target, the power Average delay (msec) requirement increases and too many transmissions tend to be driven into saturation. We have obtained all the results that we are presenting with optimized SIR target values with respect to the maximal throughput and BLER.

5.2.1. Effect of Spreading Factor

In this subsection, we study the performance of the DCH service when varying the spreading factor of the physical channel to be used. To characterize the performance, we consider again the curves of the average packet-delivery delay versus the achieved throughput.

Fig. 15 shows the average delay vs throughput curves when the service uses 11, 12 and 13 SF= 32 channels per cell and the information is protected with a R= 1/2 convolutional code.

No FC mechanism is adopted. This configuration provides a neat bit rate entering the physical layer of 105.9 kb/s. The best performance is achieved with 12 channels per cell, with a maximal throughput of 1050 kb/s.

We observe that, if the contemporary use of 13 dedicated channels per cell is allowed, the system is driven to instability, and the average delay vs. throughput curve bends backwards.

Naturally, when we increase the number of dedicated channels in each cell, we increase also the interference. The power control procedure tries to counteract such an interference increase by requesting base stations to transmit more power. However, if a new equilibrium point (a power level assignment which gives all SIR values equal to the target one) does not exist, the closed loop control mechanism increases all transmission powers until some of them reach the upper limit seen in section 4.5. In these conditions, most of the connections cannot reach the SIR target value, and, as a result, the fraction of erred transmission increases, and the throughput decreases. Moreover, due to the ARQ mechanism which tries to recover

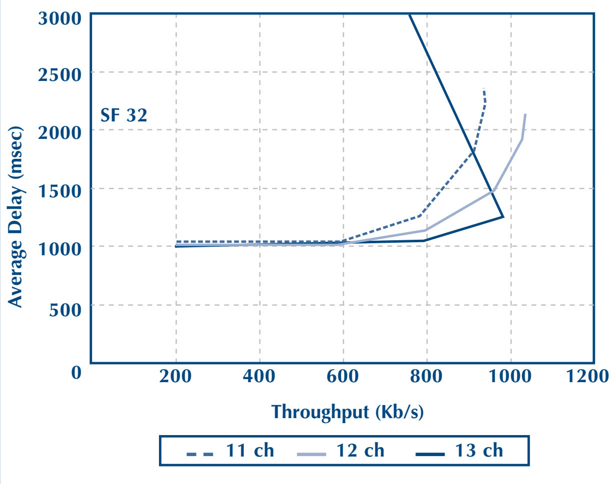

FIGURE 15: AVERAGE DELAY VS. THROUGHPUT FOR SF = 32 AND DIFFERENT NUMBER OF DCHS.

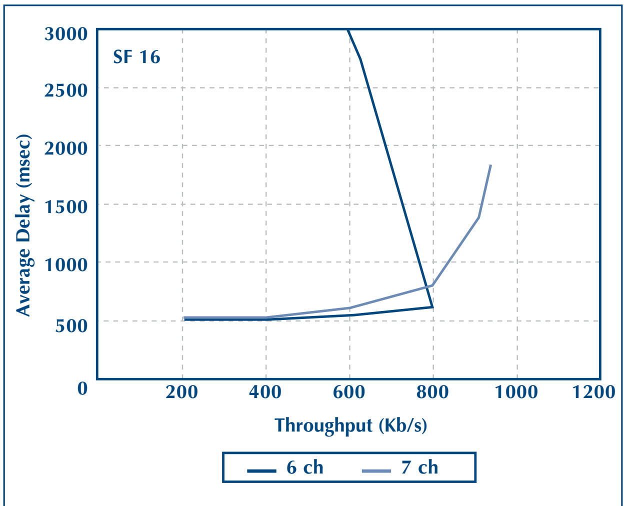

FIGURE 16: AVERAGE DELAY VS. THROUGHPUT FOR SF = 16 AND DIFFERENT NUMBER OF DCHS.

lost packets, the traffic on the channel becomes higher, and the interference further increases.

This behavior shows that with packet data service, we also need to limit the number of resources available for the service, since there exists a limit on the interference which is tolerable by the system. If not, the power control procedure is not effective in facing interference at high loads, and the system efficiency may

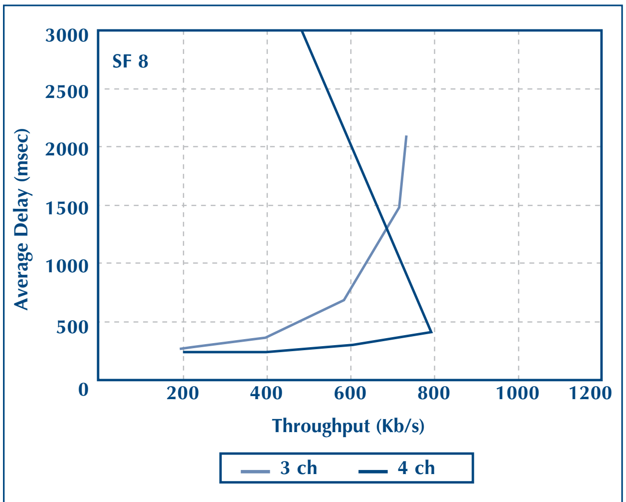

FIGURE 17: AVERAGE DELAY VS. THROUGHPUT FOR SF = 8 AND DIFFERENT NUMBER OF DCHS.

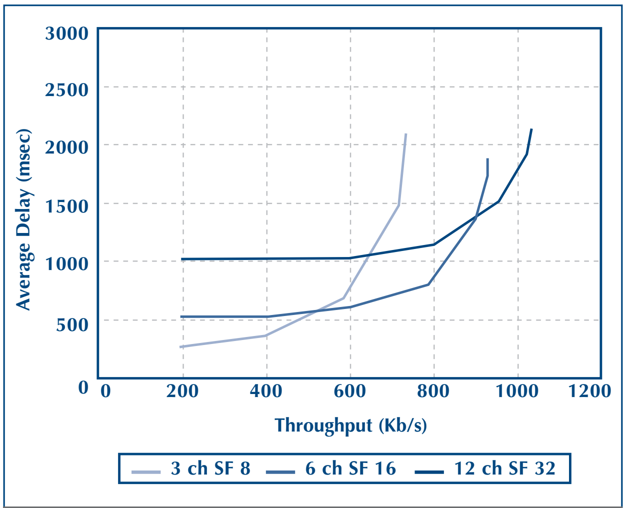

FIGURE 18: AVERAGE DELAY VS. THROUGHPUT. COMPARISON OF THE BEST CURVES FOR THE SF = 8; 16; 32 CONFIGURATIONS.

become very poor. Since we are dealing with dedicated channels which are assigned to single users through a set-up procedure, we can say that such a limitation is a kind of call admission control even if, from the service point of view, there is no call (the setup and tear-down should be transparent to packet service users).

The limitation of the number of contemporary active dedicated channels per cell is a simple control that can be adopted, even if more flexible mechanisms (out of the scope of this paper) could be designed [30].

We have observed a similar behavior using different spreading factor services. Fig.16 and 17 shows the average delay versus the throughput when using respectively 6, 7 SF= 16 channels and 3, 4 SF= 8 channels per cell. Also, in these two cases, there is a maximum level of interference affordable by the system, and a maximum number of channels to be fixed to prevent instability, which is 3 in the SF= 8 configuration and 6 in the SF= 16 one.

Fig. 18 continues with the results of figures 15, 16 and 17, showing the best delay-throughput curves obtained for the three configuration of the dedicated service (SF= 8, 16, 32). The configuration with SF= 32 channels is shown to be optimal from the maximal throughput point of view, but it presents high packet delivery delays, even for low throughput values. On the other hand, the configuration with SF= 8 achieves lower throughput values but grants limited delay for lower load values. This result confirms that, with low rate channels, the power control mechanism can better track SIR target due to the reduced interference variance, even if the average level increases due to the added intra-cell interference.

To show the effect of the proposed FC mechanism, we have reported in Fig. 19 the Block Error Rate (BLER) and the Block Error Rate when the transmission is at maximum power (BLER max) versus the offered load in the two cases with and without the FC mechanism. The physical layer configuration used is the one with 7 SF= 16 channels. Without FC, the BLER is very high and almost equal to the BLER max (about 0.7 at high loads). This shows that most of the errors occur when the power control cannot reach an equilibrium point, and bringing the power to its maximum value. In contrast, in the case with the FC mechanism, the BLER is lower (0.1 at high loads) and, furthermore, only a relatively small fraction of errors are due to power limitations. We have observed a similar behavior with the other physical layer configurations (SF= 8; 32) as well.

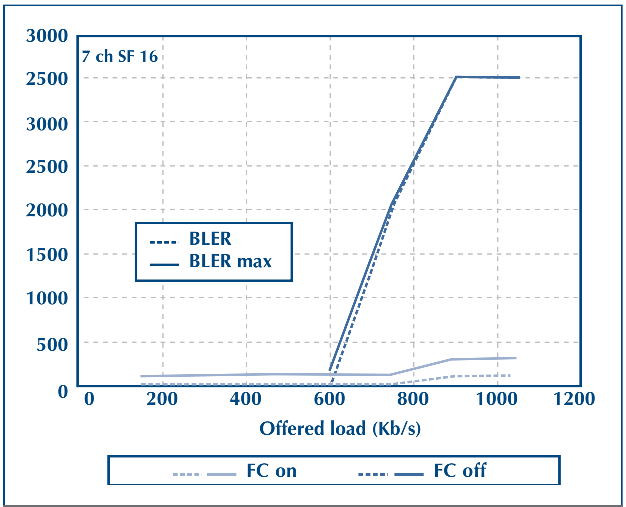

FIGURE 19: BLER VS. OFFERED LOAD WITH SF = 16. COMPARISON OF THE CASES WITH AND WITHOUT THE FC MECHANISM.

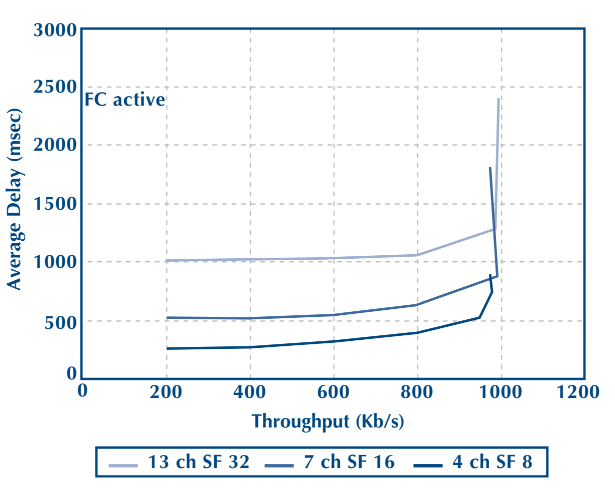

FIGURE 20: AVERAGE DELAY VS. THROUGHPUT WITH THE FC MECHANISM. COMPARISON OF THE BEST CURVES FOR THE SF = 8; 16; 32 CONFIGURATIONS.

The throughput vs average delay curves obtained with the FC mechanism have been summarized in Fig. 20, which reports for SF= 8, 16, 32 only the cases with the highest maximum throughput. In contrast, with the results obtained without FC (see Fig. 18), the maximum throughput achieved with the three SF values is almost the same. We observed that, at high loads, the FC mechanism forces the system with all the spreading factor configurations to work with the same average interference levels by modulating the traffic on the channels (G). This also allows configurations with high rate channels (small SF) to reach a high throughput while maintaining a low delay at high loads. Moreover, in all configurations, the system is able to handle one more channel than the corresponding cases without the FC mechanisms.

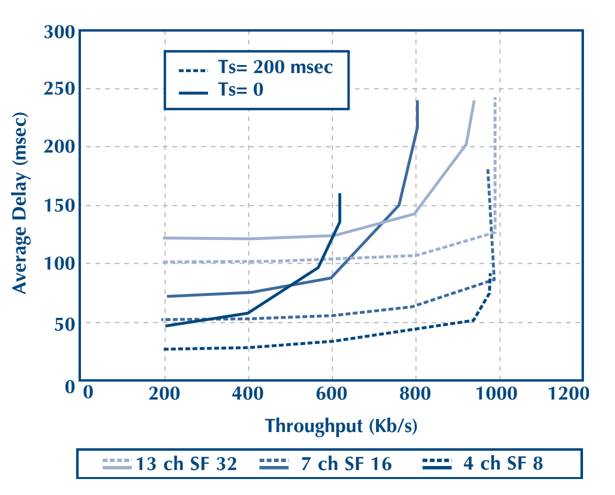

FIGURE 21: AVERAGE DELAY VS. THROUGHPUT WITH DIFFERENT SPREADING FACTORS AND DIFFERENT SET–UP DELAYS.

5.2.2. IMPACT OF THE SET–UP DELAY

The dedicated channel is assigned to the requesting user through a RRM set–up procedure [31]. Higher layer signaling needs a certain interval of time to be completed, so the starting point of the transmission session is delayed with respect to the request time.

In the previous sections, we have considered an ideal scenario where the set–up time is equal to zero, that is to say, a dedicated resource is immediately available to any user requesting for a data transfer. Here we analyze the impact of channel activation delay on the performances of the dedicated channel service, considering a constant set-up time.

Fig. 21 shows a comparison in terms of average delay vs. throughput between the cases of ideal set–up delay (equal to zero) and fixed set–up delay equal to 200 msec in the three

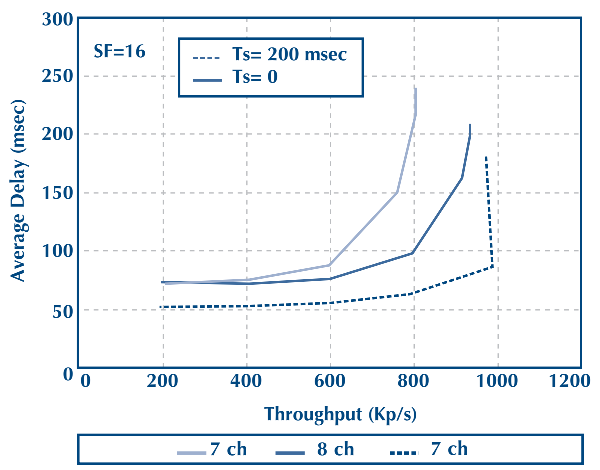

FIGURE 22: AVERAGE DELAY VS. THROUGHPUT WITH SF = 16, SET–UP DELAY EQUAL TO 200 MS AND DIFFERENT NUMBER OF CHANNELS.

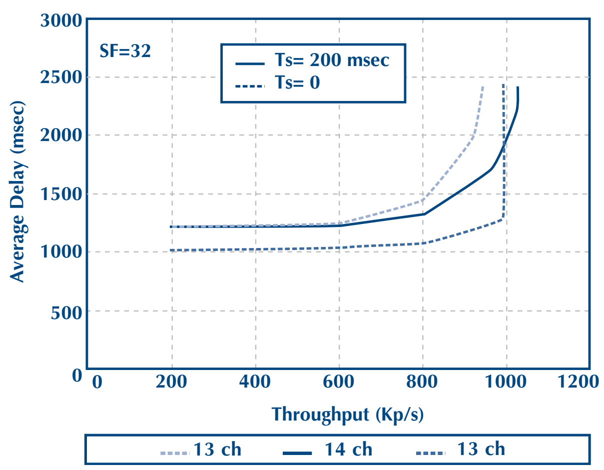

FIGURE 23: AVERAGE DELAY VS. THROUGHPUT WITH SF = 32, SET–UP DELAY EQUAL TO 200 MS AND DIFFERENT NUMBER OF CHANNELS.

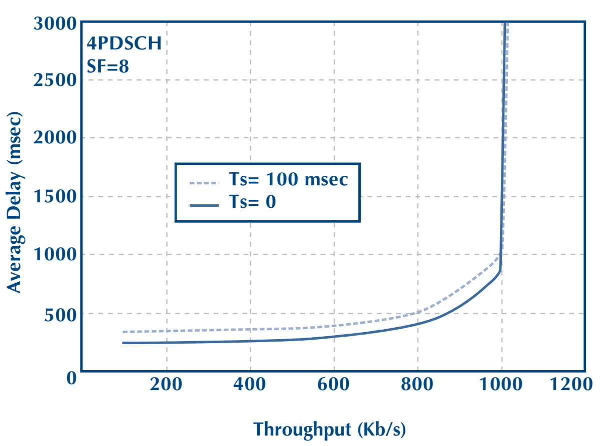

FIGURE 24: AVERAGE DELAY VS. THROUGHPUT WITH SF = 8, SET–UP DELAY EQUAL TO 200 MS AND DIFFERENT NUMBER OF CHANNELS.

physical configurations, using respectively 4, 7 and 13 dedicated channels per cell with SF= 8, 16, 32. As expected, the set–up delay affects the throughput performance since, during the time needed to set–up the physical channel, the radio resources are not fully utilized. The performance impairment is higher with high rate channels (*SF* = 8), while it is almost negligible with low rate channels (*SF* = 32). The efficiency loss mainly depends on the ratio between the set–up time and the call duration, that is the time needed to complete the data transmission. With high rate channels, the data transmission time is shorter than with low rate channels, and, therefore, the efficiency loss is higher.

However, this result does not tell the complete story. As a matter of fact, the set–up delays reduce the fraction of time the channels are used for transmission and therefore also reduce the mean interference. With a reduced interference, the system can tolerate additional channels per cell with respect to the optimum configuration in the ideal case with no set–up delay. Figures 22 and 23 show that with a set–up delay equal to 200 msec, it is possible to increase the number of channels with *SF* = 16 and *SF* = 32. In these two scenarios, the maximum throughput is very close to that achieved with zero set–up delay and, in the case of *SF* = 32, it is even slightly higher since the flow control mechanism can better control the reduced interference on the channel. We have obtained similar results in the *SF* = 8 configuration.

The packet service based on dedicated channels needs some time to activate the physical resource to be used for data transmission. In contrast, with packet services over shared channels, the channels for data transmissions are time multiplexed and do not need to be set–up for each user. However, since, to access the Downlink Shared Channel, a DCH carrying signaling must be set–up, it is also worth considering the effect of set–up delays on the performance of this type of packet service. Fig. 24 shows the average delay vs. throughput curves obtained, considering a downlink shared channel mapped onto 4 $SF= 8$ physical channels (PDSCH) per cell and a R= 1=2 convolutional code. We observe that, in this case, the time needed to set up the dedicated channel of each user has no impact on transmission throughput, since the radio resources for data transmission are shared and during this time can be used by other users.

6. CONCLUSIONS

Due to the effort of the standardization bodies, the radio interface of the UMTS is characterized by great flexibility and a variety of different physical and logical channel types. If, from one side, this added flexibility is an advantage of UMTS, from the other it makes the task of real service implementation and Radio Resource Management (RRM) more complex and challenging.

In this paper, we have investigated the performance of the UMTS radio interface with packet service mainly evaluating the maximum throughput achievable by the DSCH and DCH services with different physical channel configurations.

In order to prevent system instability, we have proposed a flow control algorithm, which dynamically controls traffic on the channels and forces the system to work with a tolerable average interference level.

As far as concerned packet transfer over DSCH, the results show that, when the packet service can use only a single physical channel, the maximum throughput is attained with the smallest available spreading factor, $SF= 4$, and a light code, $R = 2/3$. The other cases characterized by a higher channel protection (lower R or higher SF) present a lower throughput since the loss due to the added overhead is not compensated by the reduced BLER.

If the use of multiple physical channels is allowed, the maximum throughput is attained by using up to four channels with $SF= 8$ and $R = 1/2$, despite the new intra-cell interference introduced. This is mainly due to the improved efficiency of the closed-loop power control that, taking advantage of the longer transmission time and of the reduced interference burstiness, better tracks the SIR target. This confirms the common belief that CDMA characteristics are better exploited by circuit services with constant interference and, therefore, by using many small channels rather than a single big one.

If we use dedicated channels for packet data transfer for all the three spreading factor configurations considered (SF=8,16,32), we have proved that there exists a limit on the mean interference which the system can tolerate. Beyond this limit, the closed loop power control can't provide the SIR target with many active connections, and the system is driven to instability. We have shown that, with the proposed flow control mechanism, the maximum throughput achievable by the DCH service is almost the same, regardless of the channel speed, even if, of course, the delay is higher with slower channels.

Eventually, we evaluated the performance degradation of the dedicated service due to the channel set–up delay and pointed out that the throughput loss is more remarkable with low rate channels. However, since the average interference decreases due to the reduced utilization factor of radio resources, the maximum number of channels can be increased, and the maximum throughput can be brought back to the value achieved with no set–up delay in most of the cases.

Future work will expand the study to include the High Speed Downlink Packet Access (HSDPA) and will consider different traffic models to validate the generality of the results shown here.

REFERENCES

[1] R. Kalden, I. Meirick, M. Meyer, "Wireless Internet access based on GPRS", IEEE Personal Communications, vol. 7, no. 2, April 2000, pp. 8-18.

[2] A. Samukic, "UMTS universal mobile telecommunications system: development of standards for the third generation", IEEE Transactions on Vehicular Technology, vol. 47, no. 4, Nov. 1998, pp.1099-1104.

[3] K.W. Richardson, "UMTS overview", Electronics & Communication Engineering Journal, vol. 12, no. 3, June 2000, pp. 93-100.

[4] M. Gallagher, W. Webb, "UMTS the next generation of mobile radio", IEE Review, vol. 45, no. 2, March 1999, pp. 59-63.

[5] UMTS Technical Specifications, public documents available at http://www.3gpp.org.

[6] T. Ojanpera, R. Prasad, "An overview of air interface multiple access for IMT-2000/UMTS", IEEE Communications Magazine, vol. 36, no. 9, Sept. 1998, pp. 82-95.

[7] D. O'Mahony, "UMTS: the fusion of fixed and mobile networking", IEEE Internet Computing, vol.2, no. 1, Jan.-Feb. 1998, pp. 49-56.

[8] H. Holma, A. Toskala, **WCDMA for UMTS**, Second Edition, John Wiley & Sons, West Sussex (UK), 2001.

[9] T. S. Rappaport, A. Annamalai, R.M. Buerher, W.H. Tranter, "Wireless Communications: past events and future perspective", IEEE Communications Magazine, vol. 40, no. 5, May. 2002, pp. 148-161.

[10] F. Borgonovo, A. Capone, M. Cesana, L, Fratta, "Packet service in UMTS: delay-throughput performance of the downlink shared channel (DSCH)", Journal of Computer Networks, vol. 38, no. 1, Jan. 2002, pp. 43-59.

[11] A. Capone, M. Cesana, "Performance of UMTS packet service over dedicated channels (DCH)", The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications 2002 (PIMRC 2002), Sep 15-18 2002, Volume: 1, Page(s): 349 -353.

[12] E. Berruto, M. Gudmundson, R. Menolascino, W. Mohr, M. Pizarroso, "Research activities on UMTS radio interface, network architectures, and planning", IEEE Communications Magazine vol. 36, no. 2, Feb. 1998, pp. 82-95.

[13] 3rd Generation Partnership Project, Spreading and modulation (FDD), 3G TS 25.213, December 2001.

[14] 3rd Generation Partnership Project, Physical channels and mapping of transport channels onto physical channels (FDD), 3G TS 25.211, December 2001.

[15] 3rd Generation Partnership Project, Physical layer - General description, 3G TS 25.201, December 2001.

[16] 3rd Generation Partnership Project, Multiplexing and channel coding (FDD), 3G TS 25.212, December 2001.

[17] 3rd Generation Partnership Project, Radio Interface Protocol Architecture, 3G TS 25.301, December 2001.

[18] 3rd Generation Partnership Project, MAC protocol specification, 3G TS 25.321, December 2001.

[19] 3rd Generation Partnership Project, RLC Protocol Specification, 3G TS 25.322, December 2001.

[20] 3GPP Technical Report 25.848, Physical layer aspects of UTRA High Speed Downlink Packet Access, version 4.0.0, March 2002.

[21] A.M. Viterbi, A.J. Viterbi, "Erlang capacity of a power controlled CDMA system", IEEE Journal on Selected Areas in Communications, vol. 11, no. 6, Aug. 1993, pp. 892-900.

[22] J. Y. N. Hui, "Throughput analysis for Code Division Multiple Access of the Spread Spectrum Channel", IEEE Journal on Selected Areas in Communications, vol. 2, no. 4, July 1984.

[23] A. Chockalingam, P. Dietrich, L.B. Milstein, R.R. Rao, "Performance of closed-loop power control in DS-CDMA cellular systems" IEEE Transactions on Vehicular Technology, vol. 47, no. 3, Aug. 1998, pp. 774-789.

[24] R.J. McEliece, W.E. Stark, "Channels with block interference", IEEE Trans. on Information Theory, Vol. 30, No. 1, January 1984.

[25] F. Borgonovo, A. Capone, L. Fratta, "Retransmissions Versus FEC Plus Interleaving for Real-Time Applications: A Comparison Between CDPA and MC-TDMA Cellular Systems", IEEE Journal on Selected Areas in Communications, vol. 17, no. 11, Nov. 1999.

[26] UMTS 30.03, Annex B: Test environments and deployment models, TR 101 1112 v.3.2.0, April 1998.

[27] A. Bellini, M. Ferrari, **Personal communication**, Politecnico di Milano, 2000.

[28] 3rd Generation Partnership Project, RF system scenarios, 3G TR 25.942, December 2001.

[29] K. Pawlikowski, H.J. Jeong, J. Lee, "On credibility of simulation studies of telecommunication networks", IEEE Communications Magazine, Vol. 40, no 1, Jan. 2002, pp. 132-139

[30] A. Capone, S. Redana, "Call Admission Control Schemes for UMTS", IEEE VTC 2001 Fall, Atlantic City, NJ, Oct. 2001.

[31] 3rd Generation Partnership Project, RRC Protocol Specification, 3G TS 25.331, December 2001.

■ Contact: st.journal@st.com ■

