Flow Level Performance Comparison of Packet Scheduling Schemes for UMTS EUL

D.C. Dimitrova[1⋆], J.L. van den Berg[1,2], G. Heijenk[1], and R. Litjens[2]

[1] University of Twente, Enschede, The Netherlands
[2] TNO ICT, Delft, The Netherlands

Abstract. The Enhanced Uplink (EUL) is expected to provide higher capacity, increased data rates and smaller latency on the communication link from users towards the network. A key mechanism in the EUL traffic handling is the packet scheduler. In this paper we present a performance comparison of three distinct EUL scheduling schemes (one-by-one, partial parallel and full parallel) taking into account both the packet level characteristics and the flow level dynamics due to the random user behavior. For that purpose, we develop a hybrid analytical/simulation approach allowing for fast evaluation of performance measures such as mean file transfer time and fairness expressing how the performance depends on the user's location.

1 Introduction

With the specification of the Enhanced Uplink (EUL) in 3GPP Release 6 of the UMTS standard [2] a next step in the evolution of WCDMA-based cellular networks is made. As the uplink counterpart of the HSDPA (High Speed Downlink Packet Access) technology standardised in 3GPP Release 5 [1] and currently being introduced by many mobile operators, EUL is primarily designed for better support of elastic data applications.

The enhanced uplink introduces a new transport channel called EDCH, see e.g. [7]. Channel access is coordinated by the base stations via packet scheduling based on time frames of fixed length (2 or 10 ms, termed TTI: Transmission Time Interval). Fast rate adaptation with an enhanced dynamic range and efficient time multiplexing through appropriate scheduling schemes enable higher data transfer rates than usually provided on DCHs in 'plain' UMTS. Other key benefits offered by the EUL technology are an enhanced cell capacity and a reduced latency. In contrast to HSDPA for the downlink, due to limited transmit powers of the user terminals, a single uplink user cannot always use the total available channel resource on its own when it is scheduled (which would optimize throughput, cf. [12]) depending on its distance to the base station. Hence, it makes sense to consider scheduling schemes with simultaneous transmissions on the uplink, see e.g. [7].

⋆ Corresponding author, d.c.dimitrova@ewi.utwente.nl, P.O. Box 217, 7500 AE Enschede, The Netherlands.

In the present research we compare the performance of different EUL scheduling schemes, as we are particularly interested in the influence of flow level dynamics due to flow (file) transfer completions and initiations by the users at random time instants, which leads to time varying number of ongoing flow transfers. We aim at quantifying performance measures such as file transfer times and fairness, expressing how the performance depends on the user's location in the cell.

Most EUL performance studies in literature are based on dynamic system simulations, see e.g. [13], [5], [10]. The underlying simulation models incorporate many details of the channel operations and traffic behaviour, but running the simulations tends to require a lot of time. Analytical modelling may overcome this problem by abstracting from system details yet allowing the same qualitative insights into the system performance. Most analytical studies focus on the performance of schedulers without taking into account the impact of the flow level dynamics, see e.g. [9]. Analytical studies on EUL performance capturing both the packet and flow level dynamics of the system are rare. Interesting references here are [4] and [11]. In particular, in [11] flow level performance metrics are analysed for two (rate-fair) scheduling disciplines assuming that the transmit powers of all mobiles are sufficient to reach the maximum bit rate.

In the present paper we extend the model in [11] to the practical situation that the transmit power of the users is a limiting factor and scheduling schemes are not rate-fair per se but access-fair which may, implicitly, favour users close to the base station over users at the cell edge. We consider a single cell scenario with two types of users: EUL (EDCH) users generating elastic traffic flows and DCH users generating traffic flows (e.g. speech calls) that require a constant bit rate. EUL traffic is of best effort type and adapts to the available resources left over by the interfering DCH traffic.

Our modelling and analysis approach is based on time scale decomposition and consists basically of three steps. The first two steps take the details of the scheduler's behaviour into account in a given state of the system, i.e. the number of EDCH and DCH users and their distance to the base station. In particular, in the first step the data rate at which a scheduled EDCH user can transmit is determined. The second step determines the user's average throughput by accounting for the frequency at which the user is scheduled for transmitting data. In the third step these throughputs and the rates at which new DCH and EDCH users become active are used to create a continuous-time Markov chain describing the system behaviour at flow level. From the steady-state distribution of the Markov chain the performance measures, such as mean file transfer time of a user, can be calculated.

Due to the complexity of the resulting Markov model (transition rates are dependent on the full state) an analytical solution is not feasible; only for some special cases explicit expressions can be obtained for the steady-state distribution. When such closed-form expressions are not available, standard techniques for deriving the steady-state distribution can be used, e.g. numerical solution of the balance equations or simulation of the Markov chain. As the jumps in the Markov chain only apply to the initiation or completion of flow transfers

(note that the packet level details are captured in the transition rates which are calculated analytically), simulation of the Markov chain is a very attractive option and does not suffer from the long running times of the detailed system simulations used in many other studies.

The rest of the paper is organized as follows. Section 2 introduces the three different scheduling schemes we will analyse in this paper. In Section 3 we describe the network scenario considered in this paper and state the modeling assumptions. Subsequently, in Section 4 the anaytical performance evaluation approach is described in general terms, while the details of the analysis for each of the three scheduling schemes are given in Section 5. Section 6 presents and discusses numerical results illustrating their performance. Finally, in Section 7, conclusions and our plans for future work are given.

2 Scheduling Schemes for Enhanced Uplink

In this paper we focus on a class of scheduling schemes for the enhanced uplink, where the users get fair channel access independent of the actual channel conditions (channel 'oblivious' scheduling). Three different schedulers are investigated, termed one-by-one (OBO), partial parallel (PP) and full parallel (FP), which will be described in more detail below. The strategies mainly differ in the time scale on which the fair access is effectuated, in particular whether this is done within each TTI separately, or over a so-called scheduling cycle of multiple TTIs. Note, that fair channel access does not necessarily imply that each scheduling scheme yields equal bit rates to the different active EDCH users. The experienced bit rates depend on the received powers which can be different due to different distances to the serving base station.

A common notion in the three schemes is the available channel resource, termed total received power budget (B) at the base station. Expressed in linear units, B is the product of the noise rise target at the base station and the thermal noise. Part of the total budget B cannot be used by the EDCH users in a cell because of interference generated by other sources, e.g. thermal noise, intra-cell interference generated by DCH users and inter-cell interference generated by (E)DCH users in other cells. The budget left over for the intra-cell EDCH users (which varies over time) is termed the EDCH budget denoted by B'. We will now describe the scheduling schemes considered in this paper in more detail.

One-by-one (OBO) Scheduler

In this scheme, during a TTI, a single EDCH user is allowed to transmit and may use the entire available EDCH budget. The different EDCH users are selected for transmission in subsequent TTIs in a round robin fashion [11]. Figure 1(a) illustrates this scheme. Although it is generally beneficial to schedule only a single transmission during a TTI, due to high achievable instantaneous rate [12], there is also a downside, as a single EDCH user may not be able to fully utilize the available EDCH budget because of its power limitations. What part of the available resources is unused depends on the user's channel conditions. A

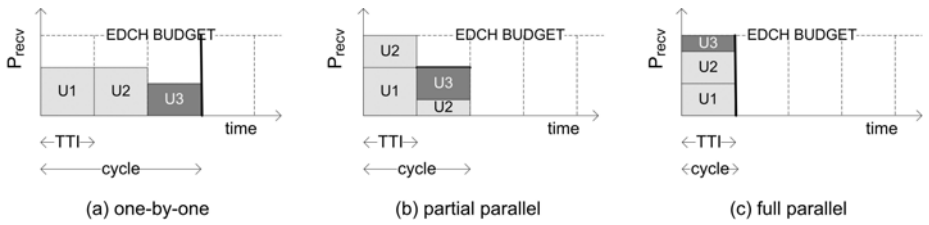


Fig. 1. Illustration of the packet handling of the considered EUL scheduling schemes.

user with good channel conditions, typically located close to the base station, is able to generate a relatively high received power level and hence leave less resources unutilised.

Partial Parallel (PP) Scheduler

The PP strategy attempts to optimize the strategy underlying the OBO scheme by selecting additional EDCH users for simultaneous transmission when the available EDCH budget cannot be fully utilised by a single transmission, see Figure 1(b). In fact, for a given TTI, EDCH users are added for simultaneous transmission as long as the sum of their maximum received powers does not exceed the available EDCH budget; the remainder of the budget is filled up by an additional EDCH user whose transmission is split over two consecutive TTIs[3]. Overall, the user selection in consecutive TTIs is done in a round robin fashion yielding fair channel access for the users.

Full Parallel (FP) Scheduler

The last scheduling scheme considered in this paper is the FP scheme (see e.g. [11]) which, like the PP scheme, also aims at full utilization of the channel resource. In this scheme all EDCH users are given simultaneous channel access in a given TTI, see Figure 1(c). If the total amount of resources requested by the EDCH users (when transmitting at their maximum power) is larger than the available EDCH budget, then the transmit powers are decreased proportionally.

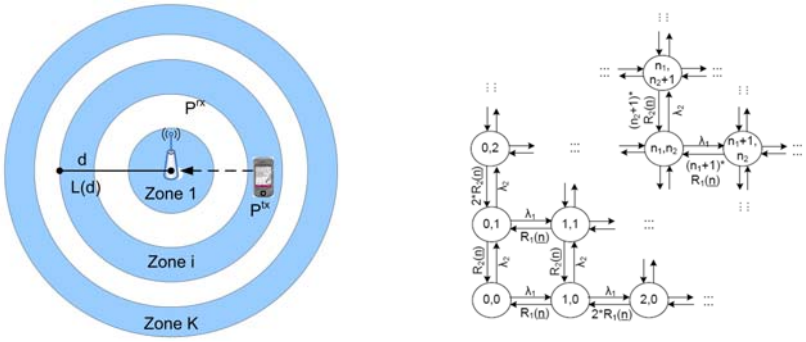
In a preliminary qualitative performance comparison of the three scheduling schemes we expect the PP scheduler to perform best. Which of the other two schedulers is best primarily depends on the available budget: if the budget is relatively low, the OBO scheduler is expected to outperform the FP scheduler since it experiences no interference from other EUL users; if the budget is high, the FP scheduler is likely to better utilize the available budget. In terms of operational complexity and computation the three schedulers differ - OBO being the least complex and PP the most. However, compared to other EUL functionality, e.g. power control, the level of complexity is relatively low.

[3] Obviously, other possible strategies exist to deal with filling up the last part of the available budget, but the differences between various strategies appear to be very small.

3 Modelling Assumptions

In this section we describe the modelling assumptions underlying the presented analysis. At the *system* level, we consider the uplink of a single cell with an omnidirectional base station, serving both DCH and EDCH calls. As illustrated in Figure 2(a), the considered cell is split in K concentric zones, where zone i is characterized by a distance d_i to the base station and a corresponding path loss denoted $L(d_i)$, $i = 1, \cdots, K$. Derived from an operator-specified noise rise target, the total received power budget at the base station is denoted B. This budget is partially consumed by the constant thermal noise level N, while the remainder is consumed by a varying amount of intra-cell interference originating from either DCH or EDCH calls. The DCH budget is the maximum part of the total budget that may be used by DCH calls. At any time, the EDCH calls may fully use that part of the budget that is not claimed by the thermal noise or on-going DCH calls: this is referred to as the EDCH budget, which is denoted $B'(n_D)$, where n_D denotes the number of existing DCH calls.

A number of additional assumptions are made at the *user* level. Calls are generated according to spatially uniform Poisson arrival processes with rates λ (EDCH calls) and λ_D (DCH calls). For the performance of EDCH calls it matters in which zone they appear. As a direct consequence of the uniformity assumption, the probability q_i that a generated EDCH call appears in zone i, is calculated as the ratio of the area of zone i and the total cell area, so that the EDCH call arrival rate in zone i is equal to λq_i, $i = 1, \cdots, K$. EDCH calls are characterised by a file that needs to be uploaded, whose size is exponentially distributed with mean F (in kbits). All calls have the same maximum transmit power $P^{tx}_{\max}$ but different maximum received power at the base station $P^{rx}_{i,\max}$ due to the zone-dependent path loss. As no user mobility is considered, users keep their positions in the cell during the file transmission. The bit rate at which an EDCH call is served depends on the experienced signal-to-interference ratio C/I. Given a prefixed E_b/N_0 (energy-per-bit to interference-plus-noise-density ratio) requirement, the attainable bit rate is equal to $r = r_{chip}\,(C/I)\,/\,(E_b/N_0)$, where $r_{chip} = 3840$ kchips/s denotes the system chip rate. The signal level C is determined by the call's transmit power and the zone-dependent path loss. The interference level I comprises several distinct components: *(i)* the thermal noise level N; *(ii)* the self-interference modelled by parameter ω, which is due to the effects of multipath fading; *(iii)* the interference $I_{EDCH}(\underline{n})$ originating from EDCH calls; and *(iv)* the interference $I_{DCH}(n_D)$ originating from DCH calls. DCH calls model e.g. speech telephony or video streaming calls and are characterised by a constant bit rate and hence a prefixed consumption P_D of the base station's received power budget, regardless of the specific location of the user. The applied value of P_D is based on a worst-case assumption that the noise rise target is fully utilised, which is indeed the objective of the EUL scheduler and hence a rather harmless assumption. Using P_D the above-mentioned DCH budget is readily translated to a maximum m on the number of admissible DCH calls, where m is increasing in the DCH budget and decreasing in the bit rate (which determines P_D). Also considering the single cell focus of our study, note



(a) (b)

Fig. 2. Modelling approach. Figure (a) shows the split up of the cell into K concentric zones; figure (b) shows the state transition diagram of the continuous-time Markov chain describing the system at flow level for the case $K = 2$ (DCH dimension is excluded for clarity of presentation).

that it suffices to keep track of the *aggregate* number of DCH calls in the cell. The DCH call duration is exponentially distributed with mean τ (in seconds). At a given time, the system state $\underline{n} \equiv (n_1, n_2, \cdots, n_K, n_D)$ is described by the number of EDCH calls n_i in zone i, $i = 1, \cdots, K$, and the total number of DCH calls n_D.

In the considered *traffic handling* scheme the fixed rate DCH calls are treated with priority within the DCH budget, while at any time the EDCH calls are allowed to utilise the remaining part of the uplink budget, including any part of the DCH budget that is not used by DCH calls. Consequently, the dynamics of the DCH calls can be described by an $M/M/m/m$ queueing model (Erlang loss model), which is independent of the EDCH dynamics. For this Erlang loss model explicit expressions are known that relate the traffic load and the channel capacity to the induced blocking probability.

4 Generic Analysis

We now move on to present how the EDCH performance is analysed by applying our proposed three-step approach. The approach here is generic in the sense that it covers all proposed schedulers. In the next section, it will be 'filled in' with the specifics of the different scheduling schemes in order to complete the analysis.

4.1 Instantaneous Rate

In the first step of the analysis we determine the so-called instantaneous bit rate $r_i(\underline{n})$, i.e. the transmission rate a call in zone i can achieve *when* it is scheduled for transmission. The instantaneous rate is defined within the boundaries of a TTI and depends on the zone i where the user is located and the interference

experienced from all other calls that are scheduled simultaneously, which in turn depends on the current state $\underline{n}$ and the scheduling scheme. We define $r_i(\underline{n})$ as a generalization of [6] eq. 8.4

$$r_i(\underline{n}) = \frac{r_{chip}}{E_b/N_0} \cdot \frac{C}{I} = \frac{r_{chip}}{E_b/N_0} \cdot \frac{P_i^{rx}}{I_{EDCH}(\underline{n}) - \omega P_i^{rx} + I_{DCH}(n_D) + N}. \quad (1)$$

Since $I_{EDCH}(\underline{n})$ is defined to include the reference call's own signal, a fraction ω of the own signal must be subtracted from $I_{EDCH}(\underline{n})$ to model the effects of self-interference properly. Since P_i^{rx}, $I_{EDCH}(\underline{n})$ and $I_{DCH}(n_D)$ depend on the state $\underline{n}$ and/or the scheduling scheme, so does the instantaneous rate $r_i(\underline{n})$.

4.2 State-dependent Throughput

Knowledge of $r_i(\underline{n})$ is not sufficient to determine a call's throughput in system state $\underline{n}$ since a call often has to wait several TTIs between actual data transmissions (scheduling cycle; see also Figure 1). The result is a decreased effective transmission rate, which we term state-dependent throughput $R_i(\underline{n})$. More precisely, $R_i(\underline{n})$ is the average transmission rate an active call achieves during one scheduling cycle, given that the system is and remains in state $\underline{n}$. Denoting with $c(\underline{n})$ the cycle length, which depends on the number of ongoing EDCH calls and the applied scheduling scheme, we have

$$R_i(\underline{n}) = \frac{r_i(\underline{n})}{c(\underline{n})}. \quad (2)$$

4.3 Markov Chain Modelling

Now that the packet level analysis is completed we can introduce flow level dynamics. This is done in the third step of the analysis with the creation of a continuous-time Markov chain model describing the dynamics of EDCH and DCH call initiations and completions in the cell. The states in the Markov model are given by $\underline{n} = (n_1, n_2, \cdots, n_K, n_D)$, i.e. the distribution of the EDCH calls over the different zones in the cell and the total number of on-going DCH calls. Hence the Markov model itself has $K+1$ dimensions, with K dimensions covering the EDCH calls in the different zones and an additional dimension to cover the DCH calls in the cell. Each of the K 'EDCH dimensions' is unlimited in the number of admissible calls, while the 'DCH dimension' is limited to m simultaneous calls. The transition rates of the Markov model are as follows, cf. Figure 2(b):

$\underline{n} \to (n_1, \cdots, n_i+1, \cdots, n_K, n_D)$	at rate	λ_i	(EDCH call arrival)
$\underline{n} \to (n_1, \cdots, n_{i+1}, \cdots, n_K, n_D+1)$	at rate	λ_D	(DCH call arrival)
$\underline{n} \to (n_1, \cdots, n_i-1, \cdots, n_K, n_D)$	at rate	$\frac{n_i}{F} R_i(\underline{n})$	(EDCH call completion)
$\underline{n} \to (n_1, \cdots, n_i+1, \cdots, n_K, n_D-1)$	at rate	$\frac{n_D}{\tau}$	(DCH call completion)

From the steady-state distribution of the Markov model we can easily derive the desired performance measures, viz. the mean file transfer times for the different zones and the fairness index.

5 Scheduler-specific Analysis

The generic three-step approach described in Section 4 can be used to analyse various schedulers. Applied to the three scheduling schemes of our interest the approach results in three different Markov models that can be classified as complex processor sharing type of queueing models. The differences are a consequence of the different expressions of $r_i(\underline{n})$ and $c(\underline{n})$. Due to the specifics of the schedulers, the Markov chains generated here are too complex for the steady-state distribution to be obtained analytically. Therefore we have chosen to simulate the Markov model in order to find the steady-state distributions.

5.1 One-by-one (OBO) Scheduler

In case of OBO scheduling only one EDCH call may be scheduled per TTI. Hence the scheduled user may in principle utilise the entire EDCH budget but may very well be limited by its own maximum received power level: $P_i^{rx} = \min\left\{P_{i,\max}^{rx}, B'(n_D)\right\}$. Having only a single active user per TTI yields a cycle length of $n \equiv n_1 + n_2 + \cdots + n_K$ (for state $\underline{\mathrm{n}}$), hence, $c(\underline{n}) = n$. Under OBO scheduling the sources of interference are thermal noise, interference from DCH calls and self-interference. $I_{EDCH}(\underline{n})$ in this case consists only of the own signal power, which allows expression (1) to be rewritten and the resulting state-dependent throughput $R_i(\underline{n})$ is given by:

$$R_i(\underline{n}) = \frac{r_{chip}}{E_b/N_0} \cdot \frac{P_i^{rx}}{(1-\omega)P_i^{rx} + I_{DCH}(n_D) + N} \cdot \frac{1}{n}. \tag{3}$$

Expression (3) shows that $R_i(\underline{n})$ depends on the current state $\underline{n}$ only via $c(\underline{n})$ and n_D. Considering the third step of the analysis, in the special case where $\lambda_D = 0$, the Markov model for the OBO scheduler is effectively a multi-class $M/M/1$ processor sharing model, which is well examined and an explicit expression for the steady-state distribution is available, see e.g. [3].

5.2 Partial Parallel (PP) Scheduler

As the PP scheduler allows parallel transmissions of multiple EDCH calls in a single TTI, $I_{EDCH}(\underline{n})$ comprises the interference contributions from all scheduled EDCH calls in a TTI. Consequently, the instantaneous rate depends on $\underline{n}$, see (1). The opportunity for simultaneous transmissions also results in a shorter cycle than under OBO scheduling which can be expressed as the ratio of the aggregate resource requested by all present EDCH calls and the available EDCH budget $B'(n_D)$. The cycle length is then given by

$$c(\underline{n}) = \max\left\{1, \frac{\sum_{i=1}^{K} n_i P_{i,\max}^{rx}}{B'(n_D)}\right\} \tag{4}$$

and hence the state-dependent throughput $R_i(\underline{n})$ is equal to

$$R_i(\underline{n}) = \frac{r_{chip}}{E_b/N_0} \cdot \frac{P^{rx}_{i,\max}}{B'(n_D) - \omega P^{rx}_{i,\max} + I_{DCH}(n_D) + N} \cdot \frac{B'(n_D)}{\sum_{i=1}^{K} n_i P^{rx}_{i,\max}}, \quad (5)$$

if $\sum_{i=1}^{K} n_i P^{rx}_{i,\max} \geq B'(n_D)$. In the alternative case that the sum of the power levels of the active users is lower than the budget, i.e. if $\sum_{i=1}^{K} n_i P^{rx}_{i,\max} < B'(n_D)$, each active user sends in each TTI and the cycle length is $c(\underline{n}) = 1$. That simplifies the state-dependent throughput expression to

$$R_i(\underline{n}) = \frac{r_{chip}}{E_b/N_0} \cdot \frac{P^{rx}_{i,\max}}{I_{EDCH}(\underline{n}) - \omega P^{rx}_{i,\max} + I_{DCH}(n_D) + N}, \quad (6)$$

where $I_{EDCH}(\underline{n}) = \sum_{i=1}^{K} n_i P^{rx}_{i,\max}$.

5.3 Full Parallel (FP) Scheduler

Under FP scheduling, an active user transmits in each TTI and therefore the cycle length $c(\underline{n})$ is equal to 1 for all states $\underline{n}$. Hence the state-dependent throughput is equal to the instantaneous rate (calculated for the appropriate received power level), i.e. $R_i(\underline{n}) = r_i(\underline{n})$. In the expression for $r_i(\underline{n})$, $I_{EDCH}(\underline{n})$ comprises contributions from all EDCH calls. We distinguish between two cases. In the first case the number of EDCH calls is such that the sum of their (received) maximum powers is lower than the EDCH budget $B'(n_D)$. In that case the EDCH calls use their maximum transmit power and the state-dependent throughput is the same as under PP scheduling, see (6). The second case is when the summed maximum received power from all users is higher than the EDCH budget. Since all users are assigned to transmit in parallel, the transmit power levels have to be decreased such that the summed received powers fit in the EDCH budget. The resulting received power levels P^{rx}_i are derived from the maximum received power via a proportional decrease:

$$P^{rx}_i = \frac{P^{rx}_{i,\max}}{\sum_{i=1}^{K} n_i P^{rx}_{i,\max}} \cdot B'(n_D), \quad (7)$$

so that $I_{EDCH}(\underline{n}) = \sum_{i=1}^{K} n_i P^{rx}_i = B'(n_D)$. Using P^{rx}_i we can rewrite expression (1) for this second case of FP scheduling as

$$R_i(\underline{n}) = r_i(\underline{n}) = \frac{r_{chip}}{E_b/N_0} \cdot \frac{P^{rx}_i}{B'(n_D) - \omega P^{rx}_i + I_{DCH}(n_D) + N}. \quad (8)$$

6 Numerical Results

In this section we present and discuss numerical results on the performance of the three scheduling schemes under various cell load conditions. The comparison of the three scheduling schemes is based on performance measures such as mean file transfer time for the EDCH users and the fairness index.

6.1 Parameter Settings

In the numerical experiments we assume a system chip rate r_{chip} of 3840 kchips/s, a thermal noise level N of -105.66 dBm and a noise rise target η at the base station of 6 dB. From these parameters the total received power budget B can be calculated: $B = \eta \cdot N$. A self-interference of 10% of the own signal is considered, i.e. $\omega = 0.9$. The assumed path loss model is given by $L(d) = 123.2 + 35.2 \log_{10}(d)$ (in dB).

The considered cell is split in $K = 10$ zones[4]. Given an E_b/N_0 target of 1.94 dB for EUL transmissions, a maximum transmission power of $P^{tx}_{\max} = 0.125$ Watt and a worst case interference level (where the received power budget B is fully used), we applied straightforward link budget calculations to determine the zone radii corresponding to a set of ten bit rates between 256 kbit/s (zone 10) and 4096 kbit/s (zone 1). EDCH calls consist of file transfers of mean size $F = 1000$ kbit; the aggregate rate at which new file transfers are initiated is $\lambda = 0.4$ unless stated otherwise.

DCH users are assumed to generate voice calls with requested bit rate of 12.2 kbit/s, an activity factor of 50% and an E_b/N_0 of 5.0 dB. The mean duration τ of the voice calls is 120 seconds. Given a noise rise target of 6 dB, this translates to a P_D of $0.0729 \cdot 10^{-14}$ Watt. A DCH budget of 70% of the totally available channel resource B was used as a default value, implying a maximum of 77 simultaneous speech calls. Given a target blocking probability of 1%, this translates to a supported speech traffic load of about 62 Erlang, and hence $\lambda_D \approx 0.52$ calls/s. The DCH call arrival rate λ_D is always chosen such that the DCH call blocking probability equals 1%. The use of other than default values will be explicitly indicated where applicable.

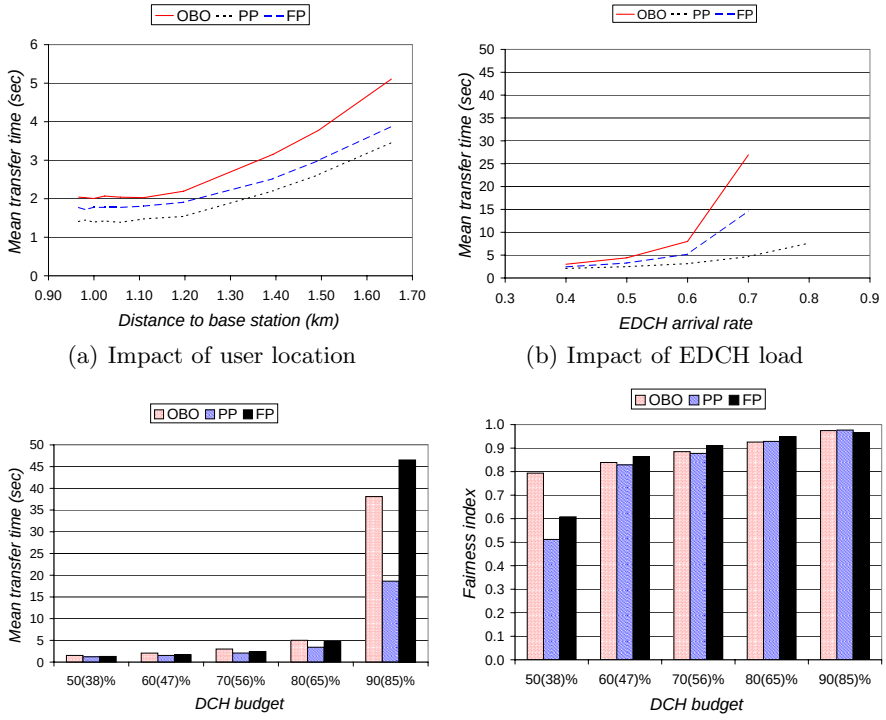
We created a generic simulator in MatLab for deriving the steady-state distribution of the multi-dimensional Markov chain describing the system behaviour at flow level (Step 3 of the analysis approach). This requires relatively short running times. For example, our simulations with confidence intervals of about 1% took typically 2.5 minutes.

6.2 Discussion of Numerical Results

Performance Impact of the User Location

Figure 3(a) shows, for each of the three schedulers, the mean file transfer time as a function of the user's distance from the base station. The system and traffic parameters are set according to their default values indicated above. As expected, the partial parallel (PP) scheduling scheme outperforms the two other schemes, cf. the discussion at the end of Section 2; in the current situation the full parallel (FP) scheme performs second best and the one-by-one (OBO) scheme shows the worst performance. For all three schedulers, when moving away from

[4] Extensive numerical experiments showed that this granularity is sufficient for our purposes. Finer granularities, e.g. $K = 20$ or $K = 40$, did not provide essentially different results.



(a) Impact of user location

(b) Impact of EDCH load

(c) Impact of DCH budget (load)

(d) Fairness for different DCH budgets

Fig. 3. Performance comparison of the three scheduling schemes

the base station, the mean flow transfer times remain more or less constant until a distance of about 1.1 km. Apparently, in these central zones the combination of the available budget and the maximum attainable received power allows the same (maximal) data rates. At larger distances the effect of the increasing path loss, consequently lower attainable received powers and hence lower bit rates becomes clearly visible through rapidly increasing flow transfer times.

Note, that even for users close to the base station, which are able to fill up the whole EDCH budget on their own, the PP scheme performs considerably better than the OBO (and FP) scheme. This is due to the fact that the particular disadvantage of OBO for users at the cell edge (who cannot fill the budget on their own, and, hence, waste resources compared to PP) also has a disadvantageous effect on the throughputs obtained by the users at the centre of the cell (as the channel access is fairly shared among all users in the cell).

Performance Impact of the Effective EDCH Load

In Figures 3(b)-3(c) the effective EDCH load is varied in two distinct ways. In Figure 3(b), the aggregate EDCH flow arrival rate λ is varied directly. In Figure 3(c), the available capacity for EDCH transfers is varied (which effectively

corresponds with an inverse variation of the EDCH load) by varying the DCH budget and load. As we will see, the key difference between these distinct approaches in varying the effective EDCH load is visible in the relative performance of the OBO and FP schedulers.

In Figure 3(b) the mean flow transfer time (appropriately averaged over all zones in the cell) is given as a function of λ. The other parameters are the same as in Figure 3(a). Observe that the performance difference between the schedulers shown by Figure 3(a) becomes even more pronounced when λ increases. In particular, the mean flow transfer time for the OBO (and also FP) scheme increases very rapidly when λ becomes larger than, say, 0.6 flow initiations/sec, while the system becomes saturated for λ's between 0.7 and 0.8. The growth of the mean flow transfer time under PP scheduling remains moderate. Apparently, the relatively inefficient traffic handling in the OBO and FP schemes leads to a considerable reduction of the cell capacity compared to the PP scheme.

As argued before, and illustrated by Figure 3(a) and Figure 3(b), the PP scheduling scheme performs always better than (or at least as good as) the OBO and FP schemes. It is however interesting to consider how the performance gain of PP over the OBO and FP schedulers depends on the available EDCH budget. In particular, it is expected that when the available EDCH budget is small enough to be filled up by a single user, OBO is more efficient than FP, since it yields lower intra-cell interference, higher signal-to-interference ratios and hence higher bit rates. In order to investigate this we have evaluated the scenario of Figure 3(a) under various DCH budgets/loads affecting the (remaining) EDCH budget available for the EDCH users. More specifically, we have varied the DCH budget between 50% and 90% of the total budget B and in each case determined the DCH arrival rate λ_D such that the DCH traffic experiences a blocking probability of 1%. Besides the available DCH budget, the horizontal axis indicates (between brackets) the (average of the) actually used budget by the DCH calls.

Figure 3(c) shows, for each of the schedulers, the resulting mean EDCH flow transfer time versus the DCH budget/load. Obviously, for all three schedulers, the mean EDCH flow transfer times increase when the DCH load increases, since the resources remaining for EDCH transfer decreases. For small values of the DCH load, the performance of the three scheduling schemes is quite similar, in particular for PP and FP. This is due to the fact that when the number of simultaneously ongoing flow transfers is small (typically when the overall system load is small) PP and FP (and to a lesser extent also OBO) handle them effectively in the same way. For higher DCH budgets/loads the performance gain of PP (compared to FP and OBO) increases, while for the highest considered DCH budget/load, the OBO scheme indeed performs better than FP.

Fairness Issues

Finally, for the same scenarios as consider in Figure 3(c), we investigate in some more detail the fairness of the three schedulers with respect to their performance as observed by users at different locations in the cell, cf. Figure 3(a). The fairness can be defined in different ways. We have used the fair-

ness index applied by e.g. Jain [8], which, in the present context, is defined as $(\sum_{i=1}^{K} D_i)^2/[K \sum_{i=1}^{K} (D_i)^2]$, where D_i denotes the mean flow transfer time for users in zone i, $i = 1, ..., K$. The maximum value of the fairness index equals 1, which refers to a perfectly fair scenario in which the mean flow transfer times are the same for all zones. The smaller the fairness index the larger the (relative) differences among the mean flow transfer times in the different zones.

Figure 3(d) shows the fairness results for the three schedulers. The general impression is that the fairness performance is more or less the same for the three schemes, see also Figure 3. However, for the case that the DCH budget/load is small (i.e. the available EDCH budget is relatively large) the OBO scheme appears to be significantly more fair than the two other schemes, in particular when compared to the PP scheme. This can be explained as follows. Under PP scheduling, users near the base station (with a high maximum received power) are more likely to be served alone compared to users at the cell edge, which are mostly served in parallel with others (because they are unable to utilise the available EDCH budget on their own) and will consequently experience lower signal-to-interference ratios and hence lower bit rates. This will lead to relatively large differences between the throughputs observed by users close to the base station and users at the cell edge. Under OBO scheduling, all users are scheduled in a one by one fashion, including remote users, which therefore do not suffer from the additional intra-cell interference as would be imposed on them under FP scheduling, thus establishing a greater degree of fairness. Observe that the fairness of the schedulers improves when the DCH budget/load increases (i.e. available EDCH budget decreases). This is due to the fact that when the available EDCH budget becomes smaller, the maximum (received) power that can be achieved by the users (i.e. their mobile equipment) at different locations is less and less a limiting factor for remote users, and hence the disadvantageous effect that remote users suffer from added intra-cell interference vanishes.

7 Conclusions

We have presented a modelling and analysis approach for comparing the flow level performance of three scheduling schemes for UMTS EUL, viz. one-by-one (OBO), partial parallel (PP) and full parallel (FP) scheduling, in a single cell scenario taking into account the users' limited uplink transmission power and interference factors such as thermal noise and intra-cell interference from UMTS R'99 uplink users. This hybrid analytical/simulation approach allows for fast evaluation of mean flow transfer times of users at different cell locations.

The numerical results show that the PP scheduling scheme clearly outperforms the two other schemes. Besides delivering higher user throughputs (i.e. smaller flow transfer times), PP also yields a considerably higher system capacity by exploiting the available channel resources in a more efficient way. The FP scheme performs mostly better than OBO, but when the available channel resource for EUL users is low (e.g. when there are many UMTS R'99 users getting preference over the 'best effort' EUL users) then OBO yields lower flow transfer

times than FP. Additionally, we observed that the three schedulers do not differ that much in (un)fairness with respect to the performance experienced by users at different locations in the cell. Only in scenarios where the available EDCH budget is relatively high, the schedulers differ significantly in the established fairness. Although in such cases the PP scheduler appears to be unfairest, it is noted that still all users are best off when compared to OBO and FP.

Currently we are extending our flow level performance modelling and analysis approach to scenarios with multiple cells taking into account mutual interactions due to inter-cell interference. One step further towards a more exhaustive research is including user mobility in the analysis approach. In addition, besides the channel oblivious scheduling schemes studied in the present paper, we will also consider channel aware schedulers in order to investigate their potential for performance enhancement and impact on e.g. fairness.

References

1. 3GPP TS 25.308. High Speed Downlink Packet Access (HSDPA); Overall Description.
2. 3GPP TS 25.309. FDD Enhanced Uplink; Overall Description.
3. J.W. Cohen. The multiple phase service network with generalized processor sharing. In *Acta Informatica*, volume 12, pages 254–284, 1979.
4. G. Fodor and M. Telek. Performance analysis of the uplink of a CDMA cell supporting elastic services. Networking 2005, Waterloo, Canada, 2005.
5. K.W. Helmersson, E. Englund, M. Edvardsson, C. Edholm, S. Parkvall, M. Samuelsson, Y-P.E. Wang, and J-F. Cheng. System performance of WCDMA enhanced uplink. IEEE VTC '05 (Spring), Stockholm, Sweden, 2005.
6. H. Holma and A. Toskala. *WCDMA for UMTS*. John Wiley & Sons Ltd, 2001.
7. H. Holma and A. Toskala. *HSDPA/HSUPA for UMTS*. John Wiley & Sons Ltd, 2006.
8. R. Jain. *The Art of Computer Systems Performance Analysis: techniques for experimental design, measurement, simulation, and modeling*. John Wiley & Sons Ltd, 1991.
9. K. Kumaran and L. Qian. Uplink scheduling in CDMA packet-data systems. In *Wireless Networks*, volume 12, pages 33–43, 2006.
10. C. Li and S. Papavassilious. On the fairness and throughput trade-off of multi-user uplink scheduling in WCDMA systems. IEEE VTC '05 (Fall), Dallas, USA, 2005.
11. A. Mäder and D. Staehle. An analytical model for best-effort traffic over the UMTS enhanced uplink. IEEE VTC '06 (Fall), Montreal, Canada, 2006.
12. S. Ramakrishna and J. M. Holtzman. A scheme for throughput maximization in a dual-class CDMA system. ICUPC '97, San Diego, USA, 1997.
13. C. Rosa, J. Outes, T.B. Sorensen, J. Wigard, and P.E. Mogensen. Combined time and code division scheduling for enhanced uplink packet access in WCDMA. IEEE VTC '04 (Fall), Los Angeles, USA, 2004.