Comparison of Preemptive and Preserving Admission Control for the UMTS Enhanced Uplink

Andreas Mäder und Dirk Staehle

University of Würzburg, Department of Distributed Systems
Am Hubland, D-97074 Würzburg, Germany
{maeder, staehle}@informatik.uni-wuerzburg.de

Abstract The UMTS enhanced uplink provides high bit rate radio bearers with fast rate control for packet switched radio traffic. Resource Managament in UMTS networks with the enhanced uplink has to consider the requirements of the dedidated channel users and the enhanced uplink users on the shared resource, i.e. the cell load. We propose an analytical model for such a system and evaluate the impact of two resource management strategies, one with preemption for dedicated channels and one without, on key QoS-indicators like blocking and dropping probabilities as well as user and cell throughput.
Keywords: WCDMA, UMTS, Enhanced Uplink, HSUPA, radio resource management, radio network planning

1 Introduction and Related Work

The enhanced uplink (sometimes also referred to as high speed uplink packet access – HSUPA) marks the next step in the evolution process of the UMTS. Introduced with UMTS release 6 and specifically designed for the transport of packet switched data, it promises higher throughput, reduced packet delay and a more efficient radio resource utilization. A detailed overview can be found e.g. in [1] or [2]. The enhanced uplink introduces a new transport channel, the Enhanced-DCH (E-DCH) and three new signaling channels. The E-DCH can be seen as an "packet-optimized" version of the DCH. The major new features are: Hybrid ARQ, implemented similarly as in the high speed downlink packet access (HSDPA), NodeB-controlled fast scheduling, and reduced transport time intervals (TTI) of 2 ms. In a UMTS network with enhanced uplink it is expected that QoS users will use the normal dedicated channels (DCH) while best-effort users will use the new enhanced dedicated channel (E-DCH).

In [3] we propose an analytical model and a radio resource management strategy based on the notion of effective bandwidths for the UMTS enhanced uplink with a preserving admission control strategy, meaning that DCH and E-DCH connections have equal priority on call arrival. We also showed the impact of one-by-one and parallel scheduling on the system performance. The focus of this work is the comparison of admission control strategies with and without priority between DCH and E-DCH connections. The admission control without priority is called *preserving*, while the admission control with priority of DCH connections is call *preemptive* since here E-DCH connections may be dropped for the sake of DCH connections.

The analytical model considers packet data streams as a *flow*, i.e. it is seen as a continuous stream of data regardless of the underlying protocol. Related work which can be found in the literature is e.g. [4], where a queueing analysis for the CDMA uplink with best-effort services is presented. A similar approach has been taken in [5], which introduces a dynamic slow down approach for the best-effort users. Preemption for QoS-users is e.g. considered in [6] for a GPRS/HSCSD system.

The question is whether the concept of preemption is a suitable way to increase the service quality for QoS-users without degrading the service for best-effort users too strong. The answer to this question is, however, always in the hand of the operator who defines acceptable service qualities for its customers. We provide an analytical tool to calculate the blocking and dropping probabilities as well as the user bit rates of the enhanced uplink users.

The rest of this paper is organized as follows: In Sec. 2 we define the radio resource management strategy which provides the frame for our calculations. This forms the base for the interference and cell load model in Sec. 3. In Sec. 4, we describe the E-DCH rate assignment and in Sec. 5 the admission control mechanism, which is then used for a queueing model approach in Sec. 6. In Sec. 7, we show some numerical examples and finally we conclude the paper with Sec. 8.

2 Radio Resource Management for the E-DCH Best Effort Service

Radio resource management (RRM) for the E-DCH users is primarily done in the NodeBs, which control the maximum transmit power of the mobiles and therefore also the maximum user bit rate. The NodeBs send scheduling grants on the absolute or relative grant channel (AGCH and RGCH, resp.), which either set the transmit power to an absolute value or relative to the current value. The mobiles then choose the transport block size (TBS) which is most suitable to the current traffic situation and which does not exceed the maximum transmit power. The grants can be sent every TTI, i.e. every 2 ms, which enables a very fast reaction to changes of the traffic or radio conditions. Grants can be received from the serving NodeB and from non-serving NodeBs. However the latter may just send relative DOWN grants to reduce the other-cell interference in their cells. In our model, we consider grants from the serving NodeB only.

Generally, the WCDMA uplink is interference limited. Therefore, following [7], we define the load in a cell as

$$\hat{\eta} = \frac{\hat{I}_D + \hat{I}_E + \hat{I}_{oc}}{\hat{I}_0 + W\hat{N}_0}, \tag{1}$$

with $\hat{I}_D$ and $\hat{I}_E$ as received powers from the DCH and E-DCH users[1] within the cell, $\hat{I}_{oc}$ as other-cell interference from mobiles in adjacent cells, W as system chip rate, $\hat{N}_0$ as thermal noise power spectral density and $\hat{I}_0 = \hat{I}_D + \hat{I}_E + \hat{I}_{oc}$. It can be readily seen that this load definition allows the decomposition of the cell load after its origin, hence we define

$$\begin{aligned} \hat{\eta} &= \frac{\hat{I}_D}{\hat{I}_0+W\hat{N}_0} + \frac{\hat{I}_E}{\hat{I}_0+W\hat{N}_0} + \frac{\hat{I}_{oc}}{\hat{I}_0+W\hat{N}_0} \\ &= \hat{\eta}_D + \hat{\eta}_E + \hat{\eta}_{oc} \end{aligned} \tag{2}$$

subject to $\hat{\eta} < 1$. The goal of the RRM is now twofold: First, the cell load should be below a certain maximum load in order to prevent outage. Second, the RRM tries to maximize the resource utilization in the cell to provide high service qualities to the users. The second goal allows also the interpretation of the maximum load as a target load, which should be met as close as possible. Since the DCH-load and the other-cell load cannot be influenced in a satisfying way, the E-DCH load can be used as a means to reach the target cell load. The fast scheduling gives operators the means to use the E-DCH best-effort users for "water-filling" the cell[2] load at the NodeBs up to a desired target. This radio resource management strategy is illustrated in Fig. 1. The total cell load comprises the varying other-cell load, the load generated by DCH users and the E-DCH load. The received power for the E-DCH users is adapted such that the total cell load is close to the maximum load. However, due to the power control error and the other-cell interference there is always the possibility of a load "overshoot". The probability for such an event should be kept low. So, the cell

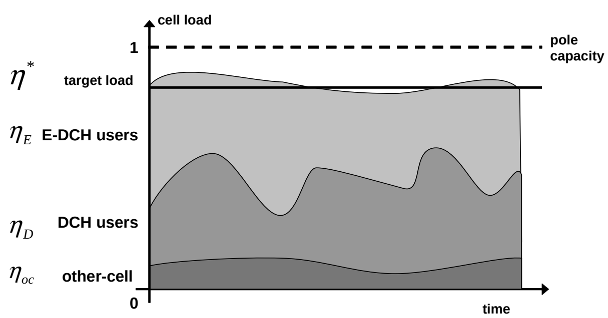

Figure 1. Illustration of the RRM principles for the E-DCH best-effort service.

load is a random variable due to fast fluctuation of the received E_b/N_0 values. We define that the goal of the RRM is to keep the probability of the total cell

[1] Note that variables $\hat{x}$ are in linear and x are in dB scale

[2] corresponding to a sector in case of multiple sectors per NodeB

load below a maximum tolerable probability p_t:

$$P\{\hat{\eta} \geq \hat{\eta}^*\} \leq p_t. \tag{3}$$

This means that the received signal power (i.e. the E-DCH interference) of the E-DCH users depends on the amount of dedicated channel and other-cell interference. More precisely, the E-DCH users are slowed down if the DCH or the other-cell load is growing, or are speed up, if more radio resources are available for the E-DCH users. If we now assume that the buffers in the mobiles of the E-DCH users are always saturated, we can use this relation to calculate the grade-of-service the E-DCH users receive depending on the scheduling strategy.

3 Interference and Load Model

Let us consider a NodeB in a UMTS network serving a single sector or cell, respectively. In the cell is a number of DCH users, each connected with a service class $s \in \mathcal{S}$. The service classes are defined by bitrate and target-E_b/N_0-value. Additionally, n_E E-DCH users are in the system. The state vector $\bar{n}$ comprises the users per DCH service class, n_s, and the E-DCH users n_E:

$$\bar{n} = (n_1, \ldots, n_{|S|}, n_E). \tag{4}$$

Each mobile power controlled by the NodeB perceives an energy-per-bit-to-noise ratio (E_b/N_0), which is given by

$$\hat{\varepsilon}_k = \frac{W}{R_k} \frac{\hat{S}_k}{W\hat{N}_0 + \hat{I}_0 - \hat{S}_k}. \tag{5}$$

In this equation, W is the chip rate of 3.84Mcps, R_k is the radio bearer information bit rate, $\hat{N}_0$ is the thermal noise power density, $\hat{S}_k$ is the received power of mobile k and $\hat{I}_0$ is the multiple-access interference (MAI) including the own- and other-cell interference. We assume imperfect power control, so the received E_b/N_0 is a lognormally distributed r.v. with the target-E_b/N_0-value ε_k^* as mean value [8] and parameters $\mu = \varepsilon_k^* \cdot \frac{\ln(10)}{10}$ and $\sigma = \mathrm{Std}[\varepsilon_k] \cdot \frac{\ln(10)}{10}$. The received power of each mobile is calculated from (5) as

$$\hat{S}_k = \hat{\omega}_k \cdot (W\hat{N}_0 + \hat{I}_0) \quad \text{with} \quad \hat{\omega}_k = \frac{\hat{\varepsilon}_k R_k}{W + \hat{\varepsilon}_k R_k}. \tag{6}$$

We define the r.v. ω_k as *service load factor* (SLF) depending on the bit rate and the E_b/N_0-value. The sum of all concurrently received powers constitutes the received own-cell interference, i.e.

$$\hat{I}_D(\bar{n}) = \sum_{s \in \mathcal{S}} \sum_{k \in n_s} \hat{S}_k \quad \text{and} \quad \hat{I}_E(\bar{n}) = \sum_{j \in n_E^a} \hat{S}_j. \tag{7}$$

$\hat{I}_D$ is the total received power of the DCH users and $\hat{I}_E$ of the E-DCH users. The substitution of $\hat{I}_D$ and $\hat{I}_E$ in Eq. (2) with Eq. (7) gives us then the load definitions depending on $\bar{n}$:

$$\hat{\eta}_D(\bar{n}) = \sum_{s\in\mathcal{S}} \sum_{k\in n_s} \hat{\omega}_k \quad \text{and} \quad \hat{\eta}_E(\bar{n}) = \sum_{j\in n_E} \hat{\omega}_j, \tag{8}$$

and the total load as

$$\hat{\eta}(\bar{n}) = \hat{\eta}_D(\bar{n}) + \hat{\eta}_E(\bar{n}) + \hat{\eta}_{oc}. \tag{9}$$

We assume the service load factors as lognormal r.v.'s with parameters μ, σ derived from the mean and variance of the E_b/N_0 distributions. These parameters depend on the service class of the users, but are equal for all users within one class. So we can write $E[\hat{\omega}_k] = E[\hat{\omega}_s]$ for all mobiles k with the same service class s. The other-cell load η_{oc} is modeled as a lognormal r.v. with constant mean and variance.

Since the total load $\hat{\eta}$ is a sum of independent lognormally distributed r.v.'s, we assume that $\hat{\eta}$ also follows a lognormal distribution [9]. We get the distribution parameters from the first moment and variance of the cell load which can be calculated directly from the moments of the SLFs:

$$E[\hat{\eta}(\bar{n})] = \sum\nolimits_{s\in\mathcal{S}} n_s \cdot E[\hat{\omega}_s] + n_E^a \cdot E[\hat{\omega}_E] + E[\hat{\eta}_{oc}]. \tag{10}$$

The variance is calculated analogously. The accuracy of this approach is validated e.g. in [10]. Another novelty of the E-DCH is Hybrid ARQ (HARQ), which combines the automatic-repeat-request protocol with code combining techniques. The effect of HARQ can be modeled as an constant gain which is included in the target-E_b/N_0 of the E-DCH and with an additional overhead on the mean data volumes of the E-DCH.

4 Rate Assignement

The available E-DCH load depends on the DCH and other-cell load. The task of the RRM is to assign each E-DCH mobile a service load factor ω such that the E-DCH load is completely utilized if possible. Generally, the user bit rate depends on the E-DCH cell load which may be generated without violating the RRM target in (3). The channel bit rate of the E-DCH is defined by the amount of information bits which can be transported within one TTI. This quantity is defined in [11] by the set of transport block sizes $\mathcal{TBS}$. With a TTI of 2ms, the information bit rate per second follows as $R_{i,E} = TBS_i \cdot 1/\text{TTI}$, where $i = 1, \ldots, |\mathcal{TBS}|$ indicates the index of the TBS. We further define $R_{0,E} = 0$. With this interpretation we can map the E-DCH bit rate to a service load factor according to Eq. (6) as

$$\hat{\omega}_{i,E} = \frac{\hat{\varepsilon}_E R_{i,E}}{\hat{\varepsilon}_E R_{i,E} + W}, \tag{11}$$

where $\hat{\varepsilon}_E$ is the E_b/N_0 for the E-DCH RB. Note that here we assume that the target-E_b/N_0-values are equal for all rates. However, this restriction can be easily

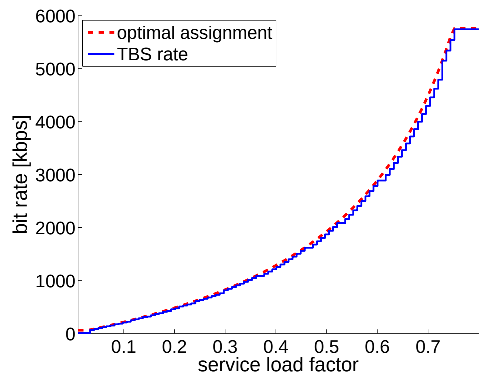

Figure 2. Mapping of service load factors to bit rates

avoided by introducing individual target-E_b/N_0-values for each rate (and ω), if they are available. The next step is to select the information bit rate such that (3) is fulfilled:

$$R_E(\bar{n}) = \max\{R_{i,E} | P(\hat{\eta}_D(\bar{n}) + n_E \cdot \hat{\omega}_{i,E} + \eta_{oc} \geq \hat{\eta}^*) \leq p_t\} \qquad (12)$$

Figure 2 shows the mapping of the service load factors to information bit rates in case of a target-E_b/N_0 of 3 dB. The optimal case indicated by the dashed line is calculated from the definition of the service load factors as $R_{\text{opt}} = \frac{\omega \cdot W}{\hat{\varepsilon}^*(1-\omega)}$. The solid line shows the corresponding rate calculated from the TBS. Both curves are very close to each other, and we see that for high SLFs, a small change means a large change on the bit rate. The non-linear dependency between bit rate and SLF is the basis for the argument that a slow-down (in terms of bit rate) of the users leads to an increased system capacity in terms of admissible sessions if an admission control based on the cell load is used ([4] and [5]). However, if we define capacity as the cumulated bit rate per cell, the capacity shrinks with the number of parallel transmitting users due to the increased interference. This is the argument for the throughput gain with one-by-one scheduling in [3].

5 Admission Control

The admission control (AC) is responsible for keeping the cell load below the maximum load. Generally, we model the AC on basis of the RRM target condition. We distinguish between two RRM policies for incoming QoS users: The first, which we call *preserving* treats E-DCH and QoS equally, which means that an incoming connection of either class is blocked if there are not enough resources available. The second, which we call *preemptive*, gives priority to QoS users, which means that eventually active best effort connections may be dropped from the system in order to make room for the incoming QoS user. In both policies existing E-DCH connections are slowed-down if the number of QoS-connections increases. However, with the preserving strategy incoming QoS-calls are blocked if the RRM cannot slow-down the E-DCH connections any more. With the preemptive strategy, one or more E-DCH connections are dropped from the system in this case, meaning that blocking for the QoS users occurs only if nearly all resources are occupied by QoS connetions, cf. Fig. 3.

If a new connection is to be established to the network, the AC is done in two steps: At first, the amount of resources ω which the incoming connection will occupy is identified. In case of a QoS-connection, this is simply ω_s. In case of an E-DCH connection, incoming connections are admitted if a minimum bit rate $R_{min,E}$ can be guaranteed. The corresponding SLF is denoted with $\omega_{min,E}$. Let us further denote with $\bar{n}^+$ the state vector $\bar{n}$ plus the incoming connection with service class s or with an additional E-DCH connection. The second step is then to estimate the probability for exceeding the maximum load with the new connection included. This step depends on the implemented policy:

Preserving Policy: In the preserving case, we calculate the parameters for the distribution of the expected cell load η_{AC} as in (10), but with $\omega_{\min,E}$ for the E-DCH users:

$$\eta_{AC}(\bar{n}^+) = \eta_D(\bar{n}^+) + n_E^+ \cdot \hat{\omega}_{\min,E} + \eta_{oc}, \tag{13}$$

where n_E^+ is the number of E-DCH mobiles with the incoming mobile included, if any. So, if the probability $P(\eta_{AC} \geq \eta^*)$ is higher than the target probability p_t, the connection is rejected, otherwise the connection is admitted.

Preemptive Policy: With preemption, the incoming call is admitted if enough resources are available such that $P(\eta_{AC} \geq \eta^*) \leq p_t$, as in the preserving case. However, if the resources are insufficient, we distinguish two cases: If the incoming call belongs to an E-DCH user, the call is blocked. If the incoming call belongs a QoS user, the RRM calculates from the service requirement ω_s the number of E-DCH connections with minimum rate $R_{E,min}$ which must be dropped from the system such that the incoming call can be admitted. The number of E-DCH connections $n_d(\bar{n}, s)$ which must be dropped depends on the current state and on the SLF of the incoming QoS-connection. It is given by the following rule:

$$n_d(\bar{n}, s) = \min\{n | P(\eta_D(\bar{n}^+) + (n_E - n) \cdot \omega_{\min,E} + \eta_{oc} \geq \eta^*) \leq p_t\}. \tag{14}$$

Note that $0 \leq n_d \leq \lceil \frac{\omega_s}{\omega_{\min,E}} \rceil$. Blocking for QoS-users occurs if the number of E-DCH connections is too low to meet the requirements of the service class, i.e. if $n_d(\bar{n}, s) > n_E$. Blocking for E-DCH users occurs if the existing connections cannot be slowed down any further, due to the constraint on the minimum bit rate.

After admission control, the RRM executes the rate assignment as in Eq. (12) to adjust the bit rate of the E-DCH users to the new situation. Figure 4 illustrates the principle of admission control and rate selection. It shows the mean and the $(1 - p_t)$-quantile (here $p_t = 5\%$) of the cell load distribution for 5 DCH users and an increasing number of E-DCH users. The target load is

$\hat{\eta}^* = 0.85$. Due to the discretization of the available rates, the $(1-p_t)$-quantile does not exactly meet the target-load, but stays just below. Since the coefficient of variation of the cell load is decreasing with the number of users in the system, the mean load comes closer to the target load with an increasing number of E-DCH users.

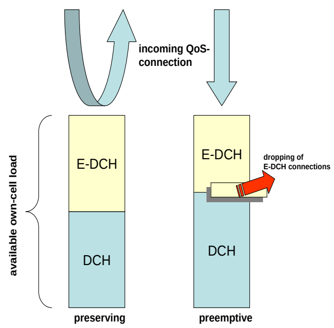

Figure 3. Principle of the preserving and preemptive policy.

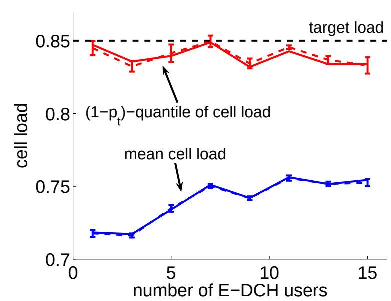

Figure 4. Mean cell load and 95%-quantiles.

6 Performance Evaluation

Now we assume that all calls arrive with exponentially distributed interarrival times with mean $\frac{1}{\lambda}$. The users choose a DCH service class or the E-DCH with probability $p_{s|E}$, hence the arrival rates per class are $\lambda_{s|E} = p_{s|E} \cdot \lambda$. The holding times for the DCH calls are also exponentially distributed with mean $\frac{1}{\mu_s}$. For the E-DCH users we assume a volume based user traffic model [12]. With exponentially distributed data volumes, the state-dependent departure rates of the E-DCH users are then given by

$$\mu_E(\bar{n}) = n_E \cdot \frac{R_E(\bar{n})}{E[V_E]}, \tag{15}$$

where $E[V_E]$ is the mean traffic volume of the E-DCH users.

The resulting system is a multi-service $M/M/n-0$ loss system with state dependent departure rates for the E-DCH users. We are now interested in calculating the steady-state distribution of the number of users in the system. Since the joint Markov process is not time-reversible which can be instantly verified with Kolomogorov's reversibility criterion, no product form solution exists. The steady state probabilities follow by solving

$$Q \cdot \bar{\pi} = 0 \quad \text{s.t.} \quad \sum \pi = 1 \tag{16}$$

for $\bar{\pi}$, where Q is the transition rate matrix. The rate matrix Q is defined with help of the bijective index function $\phi(\bar{n}) : \Omega \rightarrow N$, which maps the state vector $\bar{n}$ to a single index number. The transition rate $q(\phi(\bar{n}), \phi(\bar{n} \pm \bar{1}))$ in the rate matrix between states $\bar{n}$ and $\bar{n} \pm \bar{1}$ is then

$$q(\phi(\bar{n}), \phi(\bar{n} + \bar{1}_s)) = \lambda_s \tag{17}$$
$$q(\phi(\bar{n}), \phi(\bar{n} + \bar{1}_E)) = \lambda_E \tag{18}$$
$$q(\phi(\bar{n}), \phi(\bar{n} - \bar{1}_s)) = n_s \cdot \mu_s \tag{19}$$
$$q(\phi(\bar{n}), \phi(\bar{n} - \bar{1}_E)) = \mu_E(\bar{n}) \tag{20}$$

for all valid states in the state space Ω and $q(\phi(\bar{n}), \phi(\bar{n} \pm \bar{1})) = 0$ otherwise. The sets of $\Omega^{+}_{ps,b}$ states where blocking occurs in the preserving case are defined by the condition $P(\eta(\bar{n}^{+}) \geq \eta^{*}) > p_t$, i.e. they form the 'edges' of the state space. With preemption, an E-DCH connection is dropped if $P(\eta(\bar{n}^{+s}) \geq \eta^{*}) > p_t$ and $n_d(\bar{n}, s) \geq \lceil \frac{\omega_s}{\omega_{\min,E}} \rceil$, i.e. in case of an incoming QoS connection. We define this set as $\Omega^{+s}_{pe,d}$. Blocking occurs then in the set $\Omega^{+s}_{pe,b} = \Omega^{+s}_{ps,b} \setminus \Omega^{+s}_{pe,d}$. The set of blocking states for E-DCH connections is the same for both policies. For the preemptive policy, an additional entry in the transition rate matrix is generated for states where preemption may occur:

$$q(\phi(\bar{n}), \phi(\bar{n} + \bar{1}_s - \bar{n}_d(\bar{n}, s))) = \lambda_s. \tag{21}$$

As performance measures we choose the service-dependent call blocking probabilities P_s, the call dropping probability P_d which applies only in the case of the preemptive strategy and the mean user bit rate $E[R_U]$ achieved by the E-DCH users. The call blocking probabilities are easily calculated as the sum of all states probabilities in which blocking may occur:

$$P_s = \sum_{\bar{n} | \bar{n} \in \Omega^{+s}_{b}} \pi(\bar{n}). \tag{22}$$

Note that we omit the qualifier for the admission control policy. We define the call dropping probability in our analysis as the probability that an E-DCH connection is dropped if a QoS-call is arriving in the system. This probability is given by

$$P_d = \sum_{\bar{n} | \Omega^{+s}_{pe,d}} \frac{\pi(\bar{n}) \cdot \sum_{s' \in \mathcal{S}} P^{a}_{s'} \cdot P^{\text{sel}}_{s'}}{\sum_{\bar{n}' | n_E > 0} \pi(\bar{n}')}, \tag{23}$$

where P^{a}_{s} is the probability that the incoming connection is of class s and P^{sel}_{s} is the probability that an active E-DCH connection is selected for dropping:

$$P^{a}_{s} = \frac{\lambda_s}{\sum_{s' \in \mathcal{S}} \lambda_{s'}}, \quad \text{and} \quad P^{\text{sel}}_{s} = \frac{n_d(\bar{n}, s)}{n_E}. \tag{24}$$

We define further the mean throughput per user at a random time instance as

$$E[R_U] = \sum_{R_E > 0} R_E \cdot \frac{\sum_{\bar{n} | R_E(\bar{n}) = R_E} n_E \cdot \pi(\bar{n})}{\sum_{\bar{n}' | n_E > 0} n'_E \cdot \pi(\bar{n}')}, \tag{25}$$

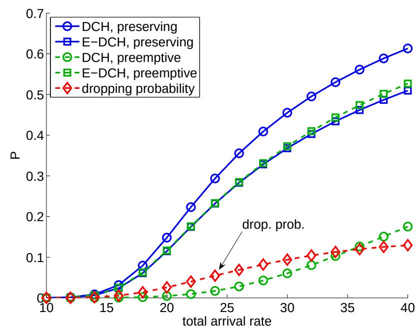

Figure 5. Blocking and dropping probabilities for QoS and E-DCH users

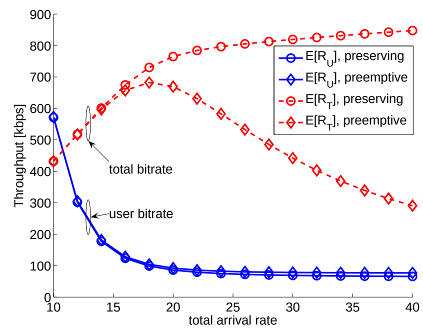

Figure 6. Mean user and cell bit rates.

which is conditioned with the probability that at least one E-DCH user is in the system, because otherwise the mean does not exist.

7 Numerical Results

In this section we give some numerical examples for our model. Our scenarios, if not stated otherwise, consist of two service classes: 64 kbps QoS-users (i.e. DCH users) with a target-E_b/N_0 of 4 dB and the E-DCH best effort users with a target-E_b/N_0 of 3 dB. The service probabilities are $p_1 = 0.4$ and $p_E = 0.6$. In Fig. 5, blocking and dropping probabilities for both admission policies are shown. The curves with a circle marker indicate the blocking probabilities for the 64 kbps QoS users, while the curves with a square marker show the blocking probabilities for the E-DCH users. The dashed line with diamond markers shows the dropping probabilities in case of preemption. Although a system with such high blocking probabilities would be considered as heavily overloaded, we show these results for a better understanding of the effect of preemption. It can be stated that preemption leads to an enormous performance gain for the QoS users, which is caused by the substantially smaller sets of states where blocking can occur at all. The blocking probabilities for the E-DCH users, however, are nearly identical and only begin to differ from each other under very high load. The dropping probabilities do not exceed approx. 10% because in high load regions the system is nearly fully occupied by QoS users.

The impact of preemption on the user and cell bit rates (defined as the cumulated bit rates of all users at any time) is shown in Fig. 6. The user throughputs have solid lines, while the cell throughputs have dashed lines. The expected user throughputs in both cases $E[R_U]$ are nearly identical with a slight advantage for the preemptive case. However, due to the dropping of users the total cell throughput $E[R_T]$ in the preemptive case is significantly lower than in the preserving case. Since the cell throughputs also consider the case if no E-DCH user at all is in the system, the curves are first increasing and then decreasing. In

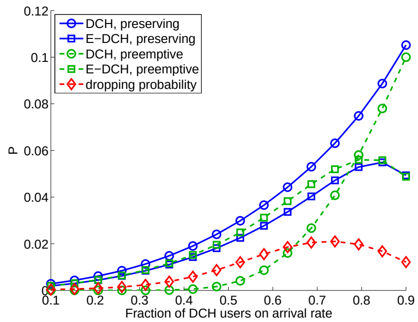

Figure 7. Impact of preemption depends on ratio between DCH and E-DCH users.

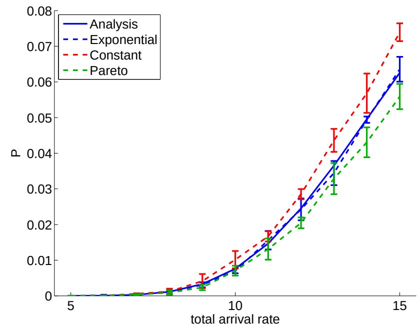

Figure 8. Sensitivity of the dropping probabilities against volume size distribution.

the next scenario we fix the total arrival rate to 15 and vary the ratio between DCH and E-DCH arrivals from 10%/90% to 90%/10%. The results are shown in Fig. 7. They show that in situations with a high fraction of best-effort traffic preemption leads to a substantial decrease of the blocking probabilities for the QoS users with still acceptable dropping probabilities. However, if the ratio is shifted to the QoS side, the decreasing load available to the E-DCH users leads to increased dropping probabilities.

Fig. 8 shows the sensitivity of the system to different volume size distributions for the E-DCH users. The results are calculated with an event-based simulation which was also used for the validation of the analytical results. Three cases are presented: Constant volume size, exponentially and Pareto distributed volume sizes (with parameters $k = 1.5$ and $x_m = 2.4 \cdot 10^4$), all with the same mean. As expected (see e.g. [13]), a higher variance leads to lower dropping probabilities, although in this very low load regions the differences are quite small, which may lead to the conclusion that the exponential assumption may be a sufficient approximation in these cases. Of course, this should be carefully validated.

8 Conclusion

We presented an analytical model for QoS and best-effort traffic over the enhanced uplink under the assumption of two admission control policies, *preserving* and *preemptive*. The model includes the effects of imperfect power control and lognormal distributed other-cell interference. The model of the admission control uses a load-based approach, i.e. connects the primarily limiting shared resource, which is the multiple access interference in the uplink, to the blocking and dropping probabilities. The evaluation of the two admission control policies showed that preemption can lead to a substantial decrease in blocking probabilities for the QoS users, but it should be generally carefully used since it can also lead to high dropping probabilities in scenarios with low quantities of best-effort traffic. A possible solution to this would be to reserve a certain amount of load to best-effort users only.

Acknowledgments:

The authors thank Prof. Phuoc Tran-Gia and Tobias Hoßfeld, University of Würzburg, Dr. Hans Barth, T-Mobile International, and Tuo Liu, University of Sidney for the fruitful discussions.

References

1. 3GPP: 3GPP TS 25.309 V6.4.0 FDD enhanced uplink; Overall description; Stage 2. Technical report, 3GPP (2005)
2. Parkvall, S., Peisa, J., Torsner, J., Sågfors, M., Malm, P.: WCDMA Enhanced Uplink – Principles and Basic Operation. In: Proc. of VTC Spring '05, Stockholm, Sweden (2005)
3. Mäder, A., Staehle, D.: An Analytical Model for Best-Effort Traffic over the UMTS Enhanced Uplink. In: Proc. of IEEE VTC Fall '06. (2006)
4. Altman, E.: Capacity of Multi-Service Cellular Networks with Transmission-Rate Control: A Queueing Analysis. In: Proc. of MobiCom '02, Atlanta, Georgia, USA (2002) 205–214
5. Fodor, G., Telek, M.: Performance Analysis of the Uplink of a CDMA Cell Supporting Elastic Services. In: Proc of NETWORKING 2005, Waterloo, Canada (2005) 205–216
6. Litjens, R., Boucherie, R.J.: Performance Analysis of Fair Channel Sharing Policies in an Integrated Cellular Voice/Data Network. Telecommunication Systems **19**(2) (2002) 147–186
7. Holma, H., Toskala, A.: WCDMA for UMTS. John Wiley & Sons, Ltd. (2001)
8. Viterbi, A., Viterbi, A.: Erlang Capacity of a Power Controlled CDMA System. IEEE Journal on Selected Areas in Communications **11**(6) (1993) 892–900
9. Fenton, L.F.: The sum of log-normal probability distributions in scatter transmission systems. IRE Transactions on Communication Systems **8**(1) (1960) 57–67
10. Staehle, D., Leibnitz, K., Heck, K., Tran-Gia, P., Schröder, B., Weller, A.: Analytic Approximation of the Effective Bandwidth for Best-Effort Services in UMTS Networks. In: Proc. of VTC Spring '03, Jeju, South Korea (2003)
11. 3GPP: 3gpp ts 25.321 v6.6.0 medium access control (mac) protocol specification. Technical report, 3GPP (2005)
12. Hossfeld, T., Mäder, A., Staehle, D.: When do we need Rate Control for Dedicated Channels in UMTS? In: Proc. of VTC Spring '06. (2006)
13. Litjens, R., Boucherie, R.J.: Elastic calls in an integrated services network: the greater the call size variability the better the QoS. Performance Evaluation **52**(4) (2003) 193–220