Dynamic half-rate connections in GSM

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

TU/e technische universiteit eindhoven

SPOR-Report 2004-06

Dynamic half-rate connections in GSM

E.M.M. Winands
J. Wieland
B. Sanders

/department of mathematics and computing science

Dynamic Half-rate Connections in GSM

E.M.M. Winands[1], J. Wieland[2] & B. Sanders[2]

[1]Department of Mathematics and Computer Science
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
e.m.m.winands@tue.nl

[2]Vodafone Group Research and Development
P.O. Box 1500, 6201 BM Maastricht, The Netherlands
{jeroen.wieland,bart.sanders}@vodafone.com

February 2004

Abstract

Dynamic Half-rate is an optional feature that allows a GSM cell to switch new incoming half-rate capable calls to half-rate speech coding, when the cell is nearly congested. Since two half-rate speech calls can be put together in one single time slot, Dynamic Half-rate has the potential to double the radio capacity of the cell. We develop a model to analyze the performance of this feature. This model is based on a reduction of the state space, which makes an efficient approximation possible. The developed approximation is a modification of the well-known Kaufman-Roberts recursion. It turns out to be extremely accurate, computationally efficient and approximately insensitive with respect to the holding time distribution. Finally, with the help of this approximation the benefits of Dynamic Half-rate are shown.

keywords: Dynamic Half-rate, GSM, congestion, approximation, Kaufman-Roberts recursion

1 Introduction

Global System for Mobile communication (GSM) uses a combination of *Frequency Division Multiple Access* (FDMA) and *Time Division Multiple Access* (TDMA) as multiple access scheme. The FDMA scheme divides the GSM frequency band into a number of carrier frequencies, which in their turn are split up into time slots by means of a TDMA scheme. A frame consists of a number of consecutive time slots. The time slots in a frame are then assigned to individual users (for a more detailed description see e.g. Rappaport [11]). Throughout this paper, the assignment of time slots is often referred to as the allocation of channels. The present paper aims to study *Dynamic Half-rate* (DHR), an optional feature in GSM.

DHR allows a GSM cell to switch new incoming half-rate capable calls to half-rate speech coding, when the cell is nearly congested. This means that the half-rate call only requires its allocated time slot every other frame. Hence, a channel is capable of supporting one full-rate
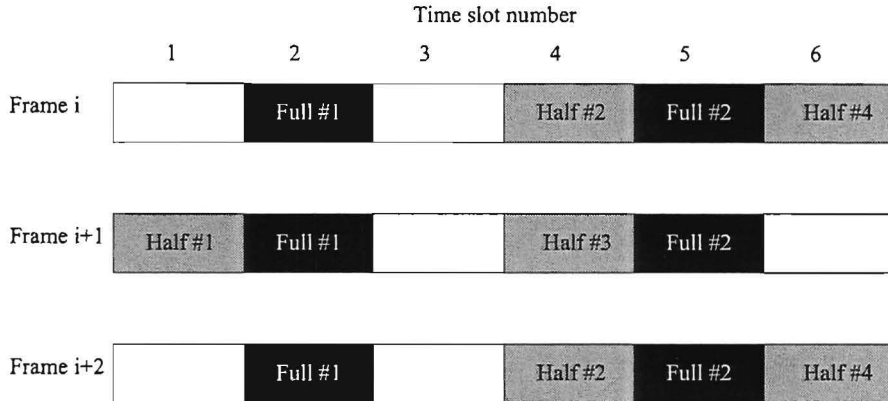
Figure 1: System state at arbitrary time.

call or two half-rate calls. The feature DHR allocates half-rate capable mobiles to full-rate or half-rate channels according to the existing traffic situation in the cell. When the number of idle traffic channels in the cell is above a pre-defined threshold, half-rate capable mobiles are allocated to full-rate channels. Otherwise, half-rate capable mobiles will be given a half-rate channel. Mobiles that are not capable of half-rate will always be allocated to a full-rate channel. This selection will be done both at call set-up and at handovers into the cell. Once a mobile has been allocated to a full/half-rate channel, the mobile will operate in this mode until the call is terminated. Since half-rate calls have a lower voice-quality due to the half-rate speech coding, it is not advisable that DHR is completely applied all the time.

This selection procedure can create so-called *partially allocated* time slots, which are time slots occupied by only one half-rate call. To make an optimal use of resources, a *re-packing* mechanism is applied. Firstly, when no idle time slots exist and there are two or more partially allocated time slots, these time slots are merged together. Secondly, when a half-rate call arrives, the system will always try to allocate it to a partially allocated time slot. Figure 1 depicts an example of the system at an arbitrary time. We note that although DHR clearly has the potential to double the radio capacity of the GSM radio network, the actual benefit of DHR depends on the amount of half-rate capable mobiles in the network.

Recently, a few papers (see e.g. Ivanovich *et al.* [6] or Lin and Lin [9]) have studied half-rate allocation schemes. However, the schemes studied so far assume that half-rate capable mobiles are always given half-rate channels independently of the traffic situation in the cell. They only decide which time slot is allocated to a newly arriving call. The DHR allocation scheme studied in the present paper is an extension with an additional dynamic aspect that decides when a half-rate capable mobile will use half-rate. The goal of our research is to get more insight into the performance of a single isolated GSM cell deploying DHR in terms of blocking probabilities. Furthermore, we want to measure the impact of aspects like the size of the cell, the offered load and the holding time distribution.

In this paper we will develop an approximate one-dimensional recursion for computing the blocking probabilities, which is comparable to the recursion used by Kaufman [8] for the more general single-threshold model. Our recursion turns out to be extremely accurate, computationally efficient and approximately insensitive with respect to the holding time distribution. With the help of this approximation, we will see that DHR has the potential to considerably

increase the capacity of a GSM cell without confronting many users with half-rate voice quality. With the expected increase of *general packet radio service* (GPRS) traffic in the coming years, which is offered over the same bearer as GSM, DHR will therefore be a very useful feature to free up time slots. Hence, the results of the present paper will be relevant for network operators.

The rest of the paper is organized as follows. In Section 2, a two-dimensional model is developed for the feature DHR based on a reduction of the state space. In Section 3, the equilibrium distribution of this model is computed by the already mentioned approximate one-dimensional recursion. The developed approximation is a modification of the well-known Kaufman-Roberts recursion (see Kaufman [7] and Roberts [12]). Section 4 investigates the accuracy of the developed approximation, the sensitivity of the system with respect to the holding time distribution and the effect of the feature DHR on a single isolated GSM cell. In particular, this section contains the results of a field study in the Vodafone-Netherlands network. Section 5 presents an extension of the studied model, in which the arrival rates depend on the state of the system. Finally, the last section contains the main conclusions of our research.

2 Model

We start this section with the development of a three-dimensional model of DHR. Due to the enormous state space of this model, numerically solving the balance equations causes both calculation and memory problems for large cell configurations. Therefore, Subsection 2.2 develops an approximate two-dimensional model based on a reduction of the state space. This latter model will be the basis of the performance analysis of Section 3.

2.1 Original re-packing

We consider a model of DHR with N parallel identical traffic channels and no waiting room. An arriving call is admitted into the system when there is sufficient room and blocked otherwise. Calls arrive according to a Poisson process with rate λ. π_h denotes the probability that a call is made by a half-rate capable mobile. The holding times are assumed to be independent, identically distributed exponential random variables with parameter μ for both half and full-rate calls. In Section 4, we investigate the sensitivity of the system with respect to this exponential holding time assumption. The traffic for ordinary (ρ_f) and half-rate capable (ρ_h) mobiles is defined by:

$$\rho_f \;=\; \frac{\lambda_f}{\mu} = \frac{\lambda(1-\pi_h)}{\mu}, \tag{1}$$

$$\rho_h \;=\; \frac{\lambda_h}{\mu} = \frac{\lambda\pi_h}{\mu}, \tag{2}$$

and the total traffic (ρ) is given by:

$$\rho = \rho_f + \rho_h = \frac{\lambda}{\mu}. \tag{3}$$

The variable K represents the threshold set by the operator, which decides when a half-rate capable mobile will use half-rate or full-rate.

The state of the system $\mathbf{X}(t)$ at time t can be described by the following three variables:

3

- $X_f(t)$: the total number of full-rate calls;

- $X_h(t)$: the total number of half-rate calls;

- $Y(t)$: the number of partially allocated time slots.

The last variable is needed to distinguish between the state in which there are two partially allocated time slots and the state in which one time slot supports two half-rate calls. It is easily verified that the associated stochastic process $\{(X_f(t), X_h(t), Y(t)), t \geq 0\}$ is a finite, aperiodic and irreducible three-dimensional Markov process. The utilization $U(t)$ at time t is given by:

$$
\begin{aligned}
U(t) &= X_f(t) + Y(t) + \frac{1}{2}\big(X_h(t) - Y(t)\big) \\
&= X_f(t) + \frac{1}{2}\big(X_h(t) + Y(t)\big).
\end{aligned}
\tag{4}
$$

Half-rate capable mobiles are allocated to a half-rate channel only when $U(t) \geq N - K$. The state space S is subject to the following conditions:

$$
\begin{aligned}
U(t) &\leq N, & (5) \\
Y(t) &\leq X_h(t), & (6) \\
Y(t) \bmod 2 &= X_h(t) \bmod 2, & (7) \\
\neg\big(U(t) = N & \wedge Y(t) \geq 2\big). & (8)
\end{aligned}
$$

We define the equilibrium state probability $p(\mathbf{n})$, $\mathbf{n} = (n_f, n_h, n_p)$, for this finite-state Markov process and the stationary utilization probability $\pi(u)$ as follows:

$$
\lim_{t\to\infty} P[(X_f(t), X_h(t), Y(t)) = \mathbf{n}] = p(\mathbf{n}), \qquad \mathbf{n} \in S,
\tag{9}
$$

$$
\lim_{t\to\infty} P[U(t) = u] = \pi(u), \qquad u = \frac{1}{2}, 1, \frac{3}{2}, \ldots, N.
\tag{10}
$$

Of course, we can write down the balance equations of the Markov process $\{(X_f(t), X_h(t), Y(t)), t \geq 0\}$. However, these equations provide little insight into the structure of the process. Instead, Table 1 sums up all the possible events starting from state $\mathbf{n}$ with the corresponding conditions for the old state, the new state reached and the transition intensity. The corresponding balance equations can in principle be solved numerically. However, the lack of a product-form structure and the size of the state space for realistic values of N make an explicit determination of the equilibrium distribution usually prohibitive. Therefore, the next subsection proposes an approximate two-dimensional model.

2.2 Complete re-packing

We modify the model by assuming that whenever there are two partially allocated time slots, these are merged into one time slot (*complete re-packing*). Since the number of half-rate and full-rate calls will now give us all information about the allocation of the time slots, the state of the system $\mathbf{X}(t)$ at time t is described by the two-dimensional vector $(X_f(t), X_h(t))$. We now approximate the utilization $U(t)$ at time t as follows:

$$
U(t) = X_f(t) + \frac{1}{2}X_h(t).
\tag{11}
$$

4

Events			
Events	Condition	New state	Intensity
Arrival ordinary mobile			
Re-packing	$(u = N - 1) \wedge (n_p \geq 2)$	$(n_f + 1, n_h, n_p - 2)$	$(1 - \pi_h)\lambda$
No re-packing	$(u \leq N - 2) \vee (u = N - 1 \wedge n_p \leq 1)$	$(n_f + 1, n_h, n_p)$	$(1 - \pi_h)\lambda$
Arrival half-rate mobile			
Half-rate allocation			
Re-packing	$(u \geq N - K) \wedge (n_p \geq 0)$	$(n_f, n_h + 1, n_p - 1)$	$\pi_h \lambda$
No re-packing	$(N - K \leq u \leq N - 1) \wedge (n_p = 0)$	$(n_f, n_h + 1, n_p + 1)$	$\pi_h \lambda$
Full-rate allocation	$u < N - K$	$(n_f + 1, n_h, n_p)$	$\pi_h \lambda$
Departure full-rate call			
	$n_f \geq 0$	$(n_f - 1, n_h, n_p)$	$n_f \mu$
Departure half-rate call			
Channel was fully allocated			
Re-packing	$(u = N) \wedge (n_p = 1)$	$(n_f, n_h - 1, 0)$	$(n_f - 1)\mu$
No re-packing	$(n_p \neq 1) \vee (u = N)$	$(n_f, n_h - 1, n_p + 1)$	$(n_f - n_p)\mu$
Channel was partly allocated	$n_p > 0$	$(n_f, n_h - 1, n_p - 1)$	$n_p \mu$

Table 1: Possible events starting from state $\mathbf{n} = (n_f, n_h, n_p)$.

The balance equations of the finite, aperiodic and irreducible Markov process $\{(X_f(t), X_h(t)), t \geq 0\}$ are given by:

$$(\lambda + (n_f + n_h)\mu)p(n_f, n_h) = \lambda p(n_f - 1, n_h) + (n_f + 1)\mu p(n_f + 1, n_h) + (n_h + 1)\mu p(n_f, n_h + 1),$$
$$u \leq N - K, \qquad (12)$$

$$(\lambda + (n_f + n_h)\mu)p(n_f, n_h) = \lambda p(n_f - 1, n_h) + \pi_h \lambda p(n_f, n_h - 1) + (n_f + 1)\mu p(n_f + 1, n_h) +$$
$$(n_h + 1)\mu p(n_f, n_h + 1), \qquad u = N - K + \frac{1}{2}, \quad (13)$$

$$(\lambda + (n_f + n_h)\mu)p(n_f, n_h) = (1 - \pi_h)\lambda p(n_f - 1, n_h) + \pi_h \lambda p(n_f, n_h - 1) + (n_f + 1)\mu p(n_f + 1, n_h) +$$
$$(n_h + 1)\mu p(n_f, n_h + 1), \qquad N - K + 1 \leq u \leq N - 1, \quad (14)$$

$$(\pi_h \lambda + (n_f + n_h)\mu)p(n_f, n_h) = (1 - \pi_h)\lambda p(n_f - 1, n_h) + \pi_h \lambda p(n_f, n_h - 1) + (n_h + 1)\mu p(n_f, n_h + 1),$$
$$u = N - \frac{1}{2}, \qquad (15)$$

$$(n_f + n_h)\mu p(n_f, n_h) = (1 - \pi_h)\lambda p(n_f - 1, n_h) + \pi_h \lambda p(n_f, n_h - 1), \qquad u = N,$$

where the terms $p(-1, n_f)$ and $p(n_h, -1)$ are zero.

The complete re-packing model also has in general no product-form structure. Nevertheless, it is possible to derive exact analytic expressions in two extreme cases. Firstly, if the operator decides not to use DHR ($K = 0$), the system reduces to the well-known Erlang loss system (see Subsection 3.1). Secondly, if DHR is completely activated ($K = N$), the model is a multidimensional generalization of this Erlang loss system (see Subsection 3.2). If DHR is activated only partially ($0 < K < N$), an algorithmic procedure appears to be necessary. Fortunately, the reduction of the state space due to complete re-packing makes an efficient approximation possible (see Subsection 3.3). We close this section with two remarks.

Remark 1. It is noted here that there exists a small difference in performance between the original system and the complete re-packing system. In the latter, the threshold is reached later, because all time slots are fully utilized all the time. Hence, we allocate fewer half-rate channels and this leads to a slightly higher blocking probability. $\square$

Remark 2. For simplicity, we have assumed that the mean holding times are equal for both half-rate and full-rate calls. However, half-rate calls have a lower bit rate and are thus expected to have a lower speech quality. In its turn, this would probably lead to smaller call lengths for half-rate calls. Fortunately, the analysis in this paper fully applies for the case, in which every type has its own temporal requirements, *mutatis mutandis.* □

3 Performance analysis

This section will first analyze the two extremes cases of no and full use of DHR, for which there exist explicit expressions of the blocking probabilities. In Subsection 3.3, we study an approximate one-dimensional modification of the Kaufman-Roberts recursion, that can be used to predict the blocking probabilities in a cell in the case of partial use of DHR.

3.1 DHR not used.

If one decides not to use DHR ($K = 0$), a half-rate capable mobile will always be allocated to a full-rate channel ($X_h = 0$). Because there is now no difference between ordinary and half-rate capable mobiles - they have the same spatial and temporal service requirements - we are only interested in the utilization at an arbitrary time and not in the exact state vector. The system reduces to the Erlang loss system with the following equilibrium distribution (see e.g. Ross [14]):

$$\pi(u) = \frac{\frac{\rho^u}{u!}}{\sum_{i=0}^N \frac{\rho^i}{i!}}, \qquad u = 0, 1, \ldots, N. \tag{16}$$

The blocking probability Π_B is given by the Erlang loss formula (see again e.g. Ross [14]):

$$\Pi_B = \pi(N) = \frac{\frac{\rho^N}{N!}}{\sum_{i=0}^N \frac{\rho^i}{i!}}. \tag{17}$$

3.2 Full use DHR.

If DHR is completely activated in a cell ($K = N$), a half-rate capable mobile will always be given a half-rate channel. Obviously, the system is a product-form type network (see e.g. Ross [14]) and the equilibrium distribution is given by:

$$p(\mathbf{n}) = \frac{\frac{\rho_f^{n_f} \rho_h^{n_h}}{n_f! \, n_h!}}{\sum_{\mathbf{n} \in S} \frac{\rho_f^{n_f} \rho_h^{n_h}}{n_f! \, n_h!}}, \qquad \mathbf{n} \in S. \tag{18}$$

An exact one-dimensional recursion for calculating the utilization probabilities $\pi(u)$ was independently published by Kaufman [7] and Roberts [12]. This recursion was originally developed for the multidimensional generalization of the classical Erlang loss model, of which our model is a special case. The modified version of their recursion runs as follows (cf. Kaufman [7] and Roberts [12]):

$$u\pi(u) = \rho_f \pi(u-1) + \frac{1}{2}\rho_h \pi(u - \frac{1}{2}), \qquad u = \frac{1}{2}, 1, \frac{3}{2}, \ldots, N. \tag{19}$$

6

This relation can be proved by using the following local balance equations (cf. Kaufman [7]):

$$\lambda_f \pi(u-1) = E(X_f|u)\mu\pi(u), \qquad u = \frac{1}{2}, 1, \frac{3}{2}, \ldots, N, \tag{20}$$

$$\lambda_h \pi(u-\frac{1}{2}) = E(X_h|u)\mu\pi(u), \qquad u = \frac{1}{2}, 1, \frac{3}{2}, \ldots, N, \tag{21}$$

where $E(X_f|u)$ and $E(X_h|u)$ represent the conditional expected value of the number of full-rate and half-rate calls, respectively, given that the utilization U is equal to u. Multiplying Equation (21) with $\frac{1}{2}$, dividing both equations by μ and summing them yields the desired Equation (19). In this derivation we use the following obvious identity:

$$E(X_f + \frac{1}{2}X_h|u) = u, \qquad u = \frac{1}{2}, 1, \frac{3}{2}, \ldots, N. \tag{22}$$

Now, we can give an explicit expression for the total blocking probability Π_B (cf. Ross [14]):

$$\begin{aligned} \Pi_B &= (1-\pi_h)[\pi(N-\frac{1}{2}) + \pi(N)] + \pi_h\pi(N) \\ &= (1-\pi_h)\pi(N-\frac{1}{2}) + \pi(N). \end{aligned} \tag{23}$$

The local balance equations defined by Equations (20) and (21) will be, in a modified form, the basis of the approximation developed in the next subsection.

Remark 3. It is noticed here that several other authors have used the local balance equations defined in Equations (20) and (21) in order to approximate utilization probabilities in non product-form variants of the generalized Erlang loss model (see e.g. Kaufman [8], Moscholios et al. [10], Roberts [13] or Chapter 3 of Ross [14]). □

3.3 Partial use DHR.

As mentioned before, in the case of partial use of DHR ($0 < K < N$) there exists no closed-form expression for the equilibrium distribution. Hence, we approximate the complete re-packing model. Thereto, we make the following assumptions:

1. *Below the threshold*, we consider the original system.

2. *Above or at the threshold*, we assume that, unlike the original system, half-rate capable mobiles are still allocated to half-rate channels. We correct this by multiplying the mean holding times for both half-rate and full-rate calls with a correction factor c.

The key approximation idea is that we adjust the original system by decreasing the spatial requirements of the half-rate capable customers and compensate this by an increment of the holding times. On average customers require less channels, but they will occupy them for a longer time.

All half-rate capable mobiles are now allocated to half-rate channels and we can define the following state-dependent parameters:

$$\mu(u) = \begin{cases} \frac{\mu}{c}, & \frac{1}{2} \le u \le N - K, \\ \mu, & N - K + \frac{1}{2} \le u \le N, \end{cases} \tag{24}$$

$$\rho_f(u) = \frac{\lambda_f}{\mu(u)}, \qquad u = \frac{1}{2}, 1, \frac{3}{2}, \ldots, N, \tag{25}$$

$$\rho_h(u) = \frac{\lambda_h}{\mu(u)}, \qquad u = \frac{1}{2}, 1, \frac{3}{2}, \ldots, N. \tag{26}$$

It remains to compute an appropriate value of the correction factor c. Therefore, we take a closer look at the subsystem above the threshold in the original system. In isolation, this subsystem is an Erlang loss system with $N - K$ servers with blocking probability as given by the Erlang loss formula (see Equation 17). On the other hand, the subsystem above the threshold of the approximation is a two-dimensional Erlang loss system, of which the blocking probability can be obtained with the help of Recursion (19). Now, we numerically compute that value of the correction factor c, for which the blocking probabilities of the original subsystem and the approximation are equal.

Although the system is not product-form, we will use the following local balance equations:

$$\lambda_f \pi(u - 1) = E(X_f|u)\mu(u)\pi(u), \qquad u = \frac{1}{2}, 1, \frac{3}{2}, \ldots, N, \tag{27}$$

$$\lambda_h \pi(u - \frac{1}{2}) = E(X_h|u)\mu(u)\pi(u), \qquad u = \frac{1}{2}, 1, \frac{3}{2}, \ldots, N. \tag{28}$$

Similar to the derivation of Recursion (19), we can obtain the following one-dimensional recursion:

$$u\pi(u) = \rho_f(u)g(u - 1) + \frac{1}{2}\rho_h(u)g(u - \frac{1}{2}), \qquad u = \frac{1}{2}, 1, \frac{3}{2}, \ldots, N. \tag{29}$$

The total blocking probability can be computed with the help of Equation (23).

Notice that the exact probabilities are computed in the extreme cases of no and full use of DHR (cf. Subsections 3.1 and 3.2). In the former, there exists no subsystem below the threshold. For the subsystem above the threshold, we compute that value of c, for which the blocking probability and the Erlang loss probability are equal. If DHR is completely activated, there exists no subsystem above the threshold and we do not need the correction factor. The subsystem below the threshold is equal to the system with full use of DHR. In Subsection 4.1, we shall study the accuracy of this approximation. We close this section with a remark.

Remark 4. In order to evaluate large cell configurations, the time required to compute the blocking probabilities in a cell is of high importance. Furthermore, to make full use of the dynamical aspect offered by DHR, it is favorable to dynamically adjust the value of the threshold parameter to the (expected) traffic situation in the GSM cell. One can think, for example, of a more aggressive use of DHR in high traffic situations such as events and traffic jams. This online monitoring and scheduling also requires an efficient algorithm to predict the blocking probabilities. It is easily verified (see e.g. Ross [14]) that the computational complexity, the number of executions needed to evaluate a problem instance, of Recursion

Cell configurations	
TRUs	Traffic channels (N)
1	6
2	14
3	21
4	29

Table 2: Cell configurations.

(29) is $O(N)$ (the number of steps to compute the correction factor c is negligible). This should be compared with the $O(N^9)$ effort to numerically solve all balance equations of the original three-dimensional Markov process described in Subsection 2.1. □

4 Numerical results

In this section, we present numerical results in order to investigate the accuracy of the developed approximation. Furthermore, we study the sensitivity of the system with respect to the holding time distribution and the effect of the feature DHR on a single isolated GSM cell. A GSM cell consists of a number of *transceiver units* (TRUs). Each TRU has eight channels. However, a certain number of these channels cannot be used for traffic, but is reserved for signaling. In Table 2, the most common cell configurations are summarized.

4.1 Accuracy of the approximation.

Tables 3 and 4 show the total blocking probabilities for a variety of traffic values. The exact blocking probabilities are marked E. These numerically computed values are those for the system under the original re-packing scheme (see Subsection 2.1). The approximate values are denoted by A. As can be seen, the approximate values closely approximate the exact values for all cell configurations. It is seen that the approximation becomes less accurate, when the fraction of half-rate capable mobiles in the network increases. This decrement in accuracy can be explained as follows. When the fraction of half-rate capable mobiles in the network is small, the system approximates the Erlang loss system, for which the exact blocking probabilities are known (see Subsection 3.1). As mentioned before, the exact blocking

Blocking probability (%)					
	$\pi_h = 0.2$		$\pi_h = 0.4$		$\pi_h = 0.8$	
K	E	A	E	A	E	A
0	5.22	5.22	5.22	5.22	5.22	5.22
1	4.87	4.73	4.10	3.95	2.12	2.04
2	4.49	4.36	3.20	3.10	0.94	0.88
3	4.17	4.09	2.59	2.55	0.46	0.44
4	3.96	3.93	2.23	2.23	0.26	0.26
5	3.86	3.85	2.07	2.08	0.20	0.20
6	3.84	3.84	2.04	2.04	0.18	0.18

Blocking probability (%)					
	$\pi_h = 0.2$		$\pi_h = 0.4$		$\pi_h = 0.8$	
K	E	A	E	A	E	A
0	8.56	8.56	8.56	8.56	8.56	8.56
3	6.38	6.42	4.01	4.25	0.89	0.97
6	5.30	5.36	2.50	2.64	0.17	0.18
9	4.90	4.93	1.99	2.05	0.05	0.06
12	4.81	4.81	1.86	1.88	0.03	0.03
15	4.80	4.80	1.85	1.85	0.03	0.03
18	4.80	4.80	1.85	1.85	0.03	0.03

Table 3: Blocking probability ($N = 6$, $\rho = 3$). Table 4: Blocking probability ($N = 21$, $\rho = 18$).

9

Blocking probability (%)		
$\pi_h = 0.2$		
K	exponential(10)	deterministic(10)	uniform(0,20)
---	---	---	---
3	4.29	4.33 ± 0.08	4.31 ± 0.04
6	3.52	3.53 ± 0.09	3.55 ± 0.04
9	3.30	3.30 ± 0.08	3.32 ± 0.04
12	3.28	3.28 ± 0.09	3.30 ± 0.05
$\pi_h = 0.4$		
K	exponential(10)	deterministic(10)	uniform(0,20)
3	2.64	2.71 ± 0.08	2.70 ± 0.04
6	1.60	1.61 ± 0.06	1.62 ± 0.03
9	1.33	1.30 ± 0.05	1.36 ± 0.03
12	1.30	1.28 ± 0.05	1.33 ± 0.03
$\pi_h = 0.8$		
K	exponential(10)	deterministic(10)	uniform(0,20)
3	0.54	0.65 ± 0.05	0.59 ± 0.04
6	0.09	0.10 ± 0.02	0.10 ± 0.01
9	0.04	0.04 ± 0.01	0.04 ± 0.01
12	0.03	0.03 ± 0.01	0.03 ± 0.01

Table 5: Blocking probability and 5% confidence intervals ($N = 14$, $\lambda = 1$).

probabilities are computed in the two extreme cases of no and full use of DHR. Finally, there exists no indication that the errors increase as the number of channels grows.

4.2 Insensitivity property.

So far, we have assumed that the holding times are exponentially distributed. However, it is well-known that the system possesses an *insensitivity property* in the two extreme cases of no and full use of DHR, i.e. the corresponding equilibrium distributions, and thus the blocking probabilities, depend only on the mean holding times and not on their actual distribution (see e.g. Chapter 2 of Ross [14]). Table 5 shows the blocking probabilities corresponding to *exponentially*, *deterministic* and *uniformly* distributed holding times, respectively. The values for the latter two distributions are obtained via a *discrete-event* simulation. Therefore, these values are shown with an approximate 95% confidence interval. As can be seen, the insensitivity property remains approximately valid for all threshold values. In turn, this observation implies that the accuracy of the developed approximation is hardly dependent on

Statistics of the sample	
Quantity	Value
Sample size	904964288
Range (seconds)	[0, 22246]
Mean (seconds)	104.5
Mode (seconds)	2.0
Standard deviation (seconds)	220.7
Coefficient of variation	2.1

Table 6: Statistics of the sample.

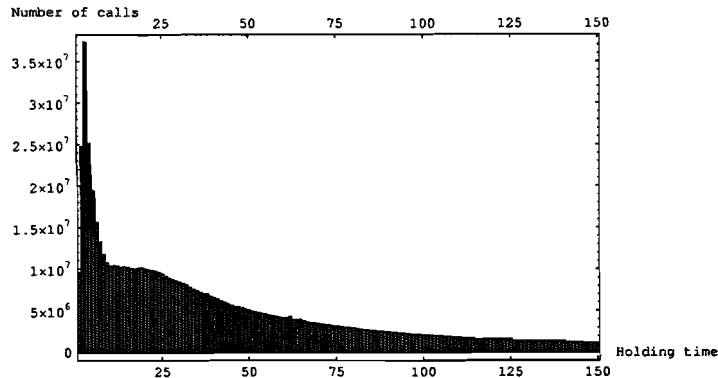| Accumulated probability (%)		
Time (seconds)	Sample	Exponential
15	26.4	13.4
30	42.0	25.0
60	61.0	43.7
120	78.4	68.3
240	89.7	90.0
480	96.1	99.0

Table 7: Accumulated probability (%).

10

Figure 2: Histogram of the holding times.

the actual holding time distribution.

To further investigate this insensitivity property, data for the holding times collected in the Vodafone-Netherlands network are used as input for our simulation. Table 6 lists the statistics of these data. It is noticed here that the coefficient of variation is significantly larger than 1. This observation, together with the values in Table 7, suggests that the sample deviates considerably from the exponential distribution, which is used in the development of the approximate recursion. The histogram of the holding times in the range 0 − 150 seconds is shown in Figure 2. Our data follow on the whole the same patterns as the samples for the holding times in (mobile) telephony of Bolotin [2], Chlebus [4], Barceló and Jordán [1], and Bolotin *et al.* [3]. For approaches to fit probability distributions on holding time data we refer to these papers and the references therein. Moreover, Bolotin [2] gives a psychophysical explanation for the shapes of the fitted distribution functions based on Weber's Law.

Table 8 lists the blocking probabilities corresponding to the exponentially distributed holding times with mean 104.5 and the empirical holding times collected in the Vodafone-Netherlands network, respectively. The latter values are shown with an approximate 95% confidence interval. Although Table 7 and Figure 2 show that the exponential distribution severely underestimates the actual holding time distribution for both short and long holding times, it can be concluded from Table 8 that via this distribution one still can extremely accurately predict the empirical blocking probabilities. The developed approximation can, therefore, be directly used by network operators in the performance evaluation and in the determination

Blocking probability (%)					
$\pi_h = 0.2$		$\pi_h = 0.4$		$\pi_h = 0.8$	
K	Exponential	Data	Exponential	Data	Exponential	Data
3	4.29	4.30 ± 0.03	2.64	2.63 ± 0.02	0.54	0.50 ± 0.01
6	3.52	3.54 ± 0.03	1.60	1.60 ± 0.02	0.09	0.09 ± 0.00
9	3.30	3.32 ± 0.03	1.33	1.35 ± 0.02	0.04	0.04 ± 0.00
12	3.28	3.30 ± 0.03	1.30	1.32 ± 0.02	0.03	0.03 ± 0.00

Table 8: Blocking probability and 5% confidence intervals ($N = 14$, $\rho = 10$).
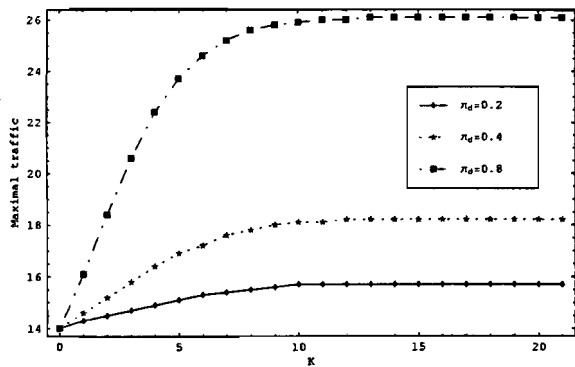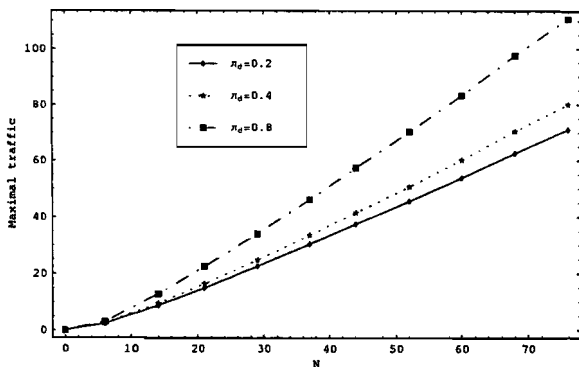
Figure 3: Maximal traffic ($K = 0.15N$, $\Pi_B = 0.02$). **Figure 4**: Maximal traffic ($N = 21$, $\Pi_B = 0.02$).

of the required parameter values of a GSM cell.

4.3 Benefits of DHR.

Figure 3 shows the maximal traffic as a function of the number of traffic channels N, when a blocking rate of 2% is maintained. For each N, the threshold K is set equal to $0.15N$. As can be seen, an increment of the fraction of half-rate capable mobiles in the network leads to a more pronounced effect of DHR. The same conclusion can be drawn from Figure 4, in which the maximal traffic is shown as a function of the threshold K in a cell consisting of 3 TRUs, when again a blocking probability of 2% is allowed. In the case of an increase of the fraction of half-rate capable mobiles, the same amount of traffic, while maintaining a blocking probability of 2%, requires a lower threshold. On the other hand, it is seen that increasing the threshold makes not much sense, when the threshold is almost equal to the number of channels in the cell.

Finally, Table 9 shows the possible increment in capacity if the threshold is increased from zero to one traffic channel under the condition that the blocking probability does not exceed 2%. In computing these values we assume that the penetration degree of half-rate capable mobiles equals 100%. The last column of this table shows the fraction of half-rate capable users that are actually allocated to half-rate channels. We can conclude that by deploying DHR large increases in radio capacity are possible without confronting many users with the lower half-rate voice quality.

Effect of DHR		
N	Capacity increase (%) (compared to $K = 0$)	Half-rate allocation (%)
6	63.6%	22.4%
14	34.0%	17.4%
21	29.3%	16.9%

Table 9: Effect of DHR ($K = 1$, $\pi_h = 1$, $\Pi_B = 0.02$).

12

5 Extension

Throughout this paper we have assumed that calls arrive according to a Poisson process. This means, among other things, that the arrival process is independent of the state of the system. A natural extension would be to let the arrival rates depend on the state of the system. We assume that these state-dependent arrival rates for ordinary and half-rate capable mobiles, respectively, admit the following linear form (cf. Delbrouck [5]):

$$\lambda_f(\mathbf{n}) = \alpha_f + \beta_f n_f, \tag{30}$$

$$\lambda_h(\mathbf{n}) = \alpha_h + \beta_h n_h. \tag{31}$$

Besides the Poisson process, examples of arrival processes that satisfy Conditions (30) and (31) are the Bernoulli and the Pascal process (see Delbrouck [5]). The former represents a finite-source arrival process, whereas the latter is an accurate approximation of overflow traffic (see Roberts [13]).

Now, the following recursion can be derived in the system with full use of DHR (cf. Delbrouck [5]):

$$u\pi(u) = \frac{\alpha_f}{\mu} \sum_{i=1}^{N} (\frac{\beta_f}{\mu})^{i-1} \pi(u-i) + \frac{\alpha_h}{2\mu} \sum_{i=1}^{2N} (\frac{\beta_h}{\mu})^{i-1} \pi(u - \frac{1}{2}i), \qquad u = \frac{1}{2}, 1, \ldots, N. \tag{32}$$

In the case of partial use of DHR, it is natural to try approximating the corresponding blocking probabilities by modifying Recursion (32) along the lines of Subsection 3.3. This more general model is a current research topic of the authors.

6 Conclusions

In this paper, an analytic model for the feature DHR has been created. The equilibrium distribution of this model has been computed by an approximate one-dimensional recursion. The benefit of this approximation is that it can handle large cell configurations, while the differences between the exact and approximate values are small within a reasonable margin. Furthermore, it turned out that the accuracy of the developed approximation is almost independent of the actual holding time distribution. With the help of this approximation, we have studied the impact of DHR on the performance of a GSM cell. We can conclude that DHR has the potential to considerably increase the capacity of a GSM cell without confronting many users with half-rate voice quality. Therefore, in the near future it will be a useful feature to improve GPRS throughput.

Acknowledgement

13

References

[1] Barceló, F., Jordán, J., (1999). *Channel Holding Time Distribution in Public Cellular Telephony* (Proceedings of the 16th International Teletraffic Congress, Edinburgh, pp. 107-116).

[2] Bolotin, V.A., (1994). *Telephone Circuit Holding Time Distributions* (Proceedings of the 14th International Teletraffic Congress, Antibes Juan-les-Pins, pp. 125-134).

[3] Bolotin, V.A., Levy, Y., Liu, D., (1999). *Characterizing Data Connection and Messages by Mixtures of Distributions on Logarithmic Scale* (Proceedings of the 16th International Teletraffic Congress, Edinburgh, pp. 887-894).

[4] Chlebus, E., (1997). *Empirical Validation of Call Holding Time Distribution in Cellular Communications Systems* (Proceedings of the 15th International Teletraffic Congress, Washington, pp. 1179-1188).

[5] Delbrouck, L.E.N., (1983). *On the Steady-State Distribution in a Service Facility Carrying Mixtures of Traffic with Different Peakedness Factors and Capacity Requirements* (IEEE Transactions on Communications, vol. 31, pp. 1209-1211).

[6] Ivanovich, M., Zukerman, M., Fitzpatrick, P., Gitlits, M., (1998). *Performance Between Circuit Allocation Schemes for Half- and Full-Rate Connections in GSM* (IEEE Transactions on Vehicular Technology, vol. 47 (3), pp. 790-797).

[7] Kaufman, J.S., (1981). *Blocking in a Shared Resource Environment* (IEEE Transactions on Communications, vol. 29, pp. 1474-1481).

[8] Kaufman, J.S., (1992). *Blocking with Retrials in a Completely Shared Resource Environment* (Performance Evaluation, vol. 15, pp. 99-113).

[9] Lin, P., Lin, Y.B., (2000). *Channel Assignment for GSM Half-Rate and Full-Rate Traffic* (Computer Communications, vol. 23 (5-6), pp. 476-482).

[10] Moscholios, I., Logothetis, M., Kokkinakis, G., (2002). *Connection Dependent Threshold Model: A Generalization of the Erlang Multiple Service Loss Model* (Performance Evaluation, vol. 48, pp. 177-200).

[11] Rappaport, T.S., (2002). *Wireless Communications: Principles and Practice, second edition* (Prentice Hall, New Jersey).

[12] Roberts, J.W., (1981). *A Service System with Heterogeneous User Requirements* (Performance of Data Communications Systems and Their Applications, G. Pujolle (ed.), North Holland, pp. 423-431).

[13] Roberts, J.W., (1983). *Teletraffic Models for the Telecom 1 Integrated Service Networks* (Proceedings of the 10th International Teletraffic Congress, Montreal, session 1.1, paper 2).

[14] Ross, K.W., (1995). *Multiservice Loss Models for Broadband Telecommunication Networks* (Springer-Verlag, Berlin).