

# TELETRAFFIC ISSUES IN HIGH SPEED CIRCUIT SWITCHED DATA SERVICE OVER GSM

Dayong Zhou and Moshe Zukerman

*Department of Electrical and Electronic Engineering*

*The University of Melbourne, Parkville, Victoria 3052, Australia*

*Email: m.zukerman@ee.mu.oz.au*

**Abstract** This paper considers a range of channel allocation schemes for High Speed Circuit Switched Data (HSCSD) over GSM and reports performance (blocking probability) and efficiency results. The channel allocation schemes studied differ in the way channels are packed (First Fit, Best Fit and Repacking), and in the connection admission policy. An overall performance comparison of the schemes is provided in order to gain insight into simplicity/efficiency tradeoffs.

**Keywords:** GSM, mobile networks, teletraffic, HSCSD, performance evaluation

## 1. INTRODUCTION

High Speed Circuit Switched Data (HSCSD) is a new GSM service that provides multi-slot high speed data service through GSM [1]. GSM supports a TDMA [8] based digital cellular mobile network. The capacity allocated to a cell is a function of frequency carriers allocated to that cell. In particular, each frequency carrier (or *frame*) supports eight TDMA channels henceforth referred to as time-slots.

Under HSCSD, each data service can obtain from one to eight time-slots. In other words, a data service may occupy an entire frame. A service is not allowed to use time-slots from different frames and the time slots of a particular service must be consecutive.

Three channel allocation schemes are studied: First Fit, Best Fit, and Repacking. Focusing on the simplicity/efficiency tradeoffs, their performance (blocking probabilities and utilisation) will be compared by simulation. Analytic solution is beyond the scope of this paper but readers who are interested in related analyses may be referred to [3][9].

We shall distinguish between two cases: (1) inflexible customers and (2) flexible customers. Inflexible customers specify the required number of time slots for their connection and will not accept any number lower than that. The flexible customers specify a lower bound which allow the service provider to allocate less than the required level (upper bound) but not less than that lower bound. The flexible customers scenario leads to an interesting situation whereby the service provider may choose to allocate less than the required level even if it has the capacity to allocate the required level. Actually, if the lower bound is one, the service provider may choose to allocate always one channel whereby ignoring the main premise of HSCSD. All these alternatives will be studied in this paper.

The remainder of the paper is organised as follows: Section 2 describes the simulation model used in the paper. Section 3 gives a description of each of the three channel allocation schemes. Sections 4 and 5 present simulation results and provide insight into the peculiarities of the different schemes and their effects on network performance and Quality of Service (QoS) levels.

## 2. THE MODEL

As mentioned, each carrier can support several data services. In a GSM system, one channel within each cell must be reserved for broadcasting, meaning that only  $8n-1$  time-slots are available for user traffic in an  $n$  carrier cell. We will consider the cases with  $n= 1, 2$  and  $3$ .

We assume call arrivals as a Poisson stream and have exponential holding times. Let  $\lambda_i$  and  $1/\mu_i$  be the Poisson arrival rate and holding time of calls that require  $i$  consecutive time-slots. The aim is to assign for each arrival the optimal available set of consecutive time-slots.

According to [9] and references therein, upon set up of an HSCSD connection, two values are specified by the user, they are denoted  $B$  and  $b$ , where  $B$  denotes the maximum acceptable capacity and  $b$  denotes the minimum acceptable capacity. In this paper, we consider three cases.

- (1)  $B = b$  This is the **inflexible customers scenario** – it is the worst for the network, and it is still a possibility that the network should consider.
- (2)  $B > b$  and  $b = 1$  This case represents the **flexible customers scenario**. We assume two extreme policies in this case:
  - i) The network provides the highest amount of bandwidth possible (not exceeding  $B$ ). This will be designated as **Low Delay Policy (LDP)**.
  - ii) The network always allocate one channel ( $b$  value in this case) regardless what the  $B$  value is. This will be designated as **High Utilization Policy (HUP)**.

### 3. CHANNEL ALLOCATION SCHEMES

In this section, we describe the three channel allocation schemes as relevant to the inflexible customers scenario. Notice that it is straightforward to apply these descriptions to the inflexible customers scenario, and notice also that in the case of the network allocating always  $b=1$  channel, all three schemes lead to the same performance.

#### 3.1 First Fit [4]

In this scheme, the frames are ordered and designated as Frame 1, Frame 2 and Frame 3 etc. When a service which requires  $m$  time-slot arrives, we allocate it  $m$  consecutive time-slots based on availability first in frame 1, then in frame 2, etc. In a particular frame we allocate the first available  $m$  time-slots. This channel allocation scheme is simplest to implement [5].

The eight time-slots in the frame are permanently allocated ID numbers as a two-dimensional array. For example, the first eight time-slots in the first carrier will have ID number (1,1), (1,2), ..., (1,8), the 8 time-slots in the second carrier will have ID number (2,1), (2,2), ..., (2,8).

Each frequency carrier may be in any one of the following two states :

- (i) no HSCSD data service in progress.
- (ii) any feasible combination of HSCSD data services each of which may occupy between one and eight time-slots.

The allocation algorithm then functions as follows :

- (a) Each incoming HSCSD call which needs one time slot channel is allocated one EMPTY time slot whose ID-number is the lexicographically smallest among all time-slots that are currently EMPTY. (Henceforth, we shall use the word smallest to mean lexicographically smallest ID numbers of time-slots.)
- (b) Each incoming data service which requires  $n$  time-slots will be allocated  $n$  empty consecutive time-slots of which the first time slot has the smallest ID-number among all EMPTY time-slots.
- (c) When there is no possibility to fit that call into one carrier, try the next carrier. The call must not be split across carriers.
- (d) The HSCSD call will be blocked after channel allocation in all carriers in the cell have been tried and failed.
- (e) No reordering of calls is performed at any time a set of times slots remained assigned for that service until it terminates.

### 3.2 Best Fit [4][5]

Let a *hole* be a consecutive set of empty time-slots. Under Best Fit for each incoming  $m$  time-slots service we try to find an  $m$  slot hole. If such search fails, we search for an  $m+1$  slot hole. The aim is to keep the allocated time-slots close together. If more than one hole of the same size is available we select based on the smallest ID number. The HSCSD call will be blocked if no such area exists. No action is taken upon call departure and call arrival.

### 3.3 Repacking [3][5]

Here, for new calls we implement the Best Fit approach. Unlike Best Fit, under this approach if a new call arrives and cannot find a suitable hole, the time-slots allocated to the calls in progress are rearranged to find a suitable hole for the new call. This rearrangement is implemented by solving the bin packing problem [6] using Branch and Bound algorithm [7].

If such suitable hole cannot be found even with rearrangement of time-slots, the new call is blocked.

Implementation of the Repacking strategy makes use of intracell handover including Repacking across different radio frequency carriers within the same cell. A large number of intracell handovers during a call may have a negative effect on the QoS. It is therefore important to limit this number.

## 4. SIMULATION RESULTS: INFLEXIBLE CUSTOMERS

In this section, we present performance results for the inflexible customers scenario. In Section 5 we shall present results for the flexible customers scenario, and compare them with the results of this section. The following performance measures will be considered:

- a. Blocking Probabilities of each scheme in the case of three carriers, two carriers and one carrier in a cell.
- b. Maximal utilisation of each scheme subject to meeting blocking probability constraints in case of 1, 2, 3 carriers in a cell.

In the simulations presented in this paper, we assume that the arrival rate of the different services in any particular run are equal (i.e.  $\lambda_1 = \lambda_2 = \dots = \lambda_8$ ). This is a worst case scenario from the service provider point of view. We shall demonstrate the significant wastage occurs with such traffic scenario.

## 4.1 Blocking Probability

It is easily noticed that in the case of only one carrier, because one channel is reserved for broadcasting and signalling, the services of eight time-slots is always blocked. Since we assumed equal arrival rate, the blocking probability in this scenario must be higher than 1/8. Actually, very high blocking probability is observed also for two and three carriers under this worst case scenario. Of course, the worst case scenarios we are presenting here are theoretical and may not occur in practice too often. They nevertheless signifies the worst case wastage and performance degradation. Based on above considerations, sufficiently long simulations were run such that 95% confidence intervals are as presented in the following tables :

*Table 1* Blocking probabilities of three schemes in case of one carrier (7 time-slots for HSCSD service, 1 for signalling)

Load (calls/s)	Blocking_Prob of Repacking	Blocking_Prob of Best Fit	Blocking_Prob of First Fit
1	0.364±0.0069	0.403±0.0022	0.403±0.0022
2	0.470±0.0063	0.534±0.0029	0.534±0.0029
3	0.540±0.0060	0.621±0.0032	0.612±0.0030
4	0.588±0.0057	0.633±0.0032	0.671±0.0038
5	0.625±0.0054	0.681±0.0039	0.689±0.0040

*Table 2* Blocking probabilities of three schemes in case of two carriers (15 time-slots for HSCSD service, 1 for signalling)

Load (calls/s)	Blocking_Prob of Repacking	Blocking_Prob of Best Fit	Blocking_Prob of First Fit
1	0.094±0.0041	0.116±0.0047	0.108±0.0044
2	0.211±0.0055	0.246±0.0052	0.248±0.0049
3	0.299±0.0072	0.333±0.0057	0.348±0.0053
4	0.377±0.0088	0.402±0.0066	0.407±0.0062
5	0.440±0.0100	0.468±0.0069	0.465±0.0066

*Table 3* Blocking probabilities of three schemes in case of three carriers (23 time-slots for HSCSD service, 1 for signalling)

Load (calls/s)	Blocking_Prob of Repacking	Blocking_Prob of Best Fit	Blocking_Prob of First Fit
1	0.016±0.0010	0.019±0.0022	0.022±0.0018
2	0.074±0.0018	0.089±0.0023	0.097±0.0024
3	0.149±0.0043	0.162±0.0041	0.168±0.0047
4	0.220±0.0061	0.245±0.0062	0.244±0.0061
5	0.273±0.0082	0.290±0.0069	0.318±0.0087

## Discussion on the Blocking Probability Results

Table 1 shows that, in the case of one carrier, for arrival rates between 1 and 5 calls/s there is no significant difference between the blocking probabilities of First Fit and Best Fit. This difference remains small (however somewhat more noticeable) in the case of 2 and 3 carriers in Tables 2 and 3. This is because these schemes do not exhibit good packing compression. There will be a group of free time-slots be created upon departures of different calls. This is even worse in First Fit because it puts calls into carriers without considering following arrivals. Best Fit performs better than First Fit as available carrier goes from 1 to 3. This is because Best Fit is trying to leave as many time-slots as it can for following arrivals though it does not repack time-slots either. As the number of carriers increases, Best Fit has greater selection to chose an optimal one. But in case of one carrier, the selection is limited.

Best Fit does not show the same advantage in this simulation as it in a simulation which only has full rate and half rate calls [2]. The reason is that HSCSD provides a large diversity of channel (1 to 8 time-slots) which adversely affects Best Fit performance. Best Fit does not give significant benefit in case of one carrier system. But we can get higher benefit when there are more carriers.

From Table 1, Table 2 and Table 3, it is evident that Repacking has the best performance in any case. This is because Repacking leaves free time-slots as many as possible and fully uses them.

## 4.2 Utilisation

As mentioned above, this wide diversity of traffic leads to low performance or alternatively high wastage. This will be demonstrated in this section. We will seek to find what the maximal utilisation is subject to meeting blocking probability requirements. We will start with blocking probability level of 20% and then we will consider a more realistic blocking probability of 2% and we shall demonstrate the enormous wastage required to provide such low blocking probability. As discussed above, the 2% will only apply for the two and three carriers cases because 1/8 of the traffic (namely the 8 time-slots service) is always blocked in one carrier case.

In this paper, utilisation is defined as the average number of occupied time-slots divided by the total number of time-slots. For every given blocking probability, we will find by using bisection the maximal utilisation subject to meeting required blocking probability level (in our case 20% or 2%).

We have obtained the following results:

*Table 4* Utilisation of three schemes in case of 1, 2 and 3 carriers (given 20% blocking probability)

Number of Carriers in a cell	Utilisation of First Fit	Utilisation of Best Fit	Utilisation of Repacking
1	0.081±0.0053	0.082±0.0059	0.1619±0.0350
2	0.3574±0.0161	0.3729±0.0287	0.4280±0.0316
3	0.4538±0.0290	0.4800±0.0277	0.5030±0.0251

*Table 5* Utilisation of three schemes in case of 2 and 3 carriers (given 2% blocking probability)

Number of Carriers in a cell	Utilisation of First Fit	Utilisation of Best Fit	Utilisation of Repacking
2	0.09338±0.0041	0.09486±0.038	0.11316±0.0051
3	0.17924±0.028	0.19064±0.037	0.19351±0.026

The results presented in Table 5 signify the enormous wastage caused by wide traffic diversity in HSCSD in the inflexible customers scenario with large diversity of services.

### Discussion On Utilization Results

A clear observation of the results is that allowing large diversity of services, in other word, allowing multiple time-slots (from 1 to 8), assuming inflexible customers ( $B=b$ ), may lead to unacceptable wastage with significant cost implications. This is also true for Repacking - the most efficient scheme. The problem can be resolved by either not allowing  $b>4$  service or by charging heavily inflexible customers who insist on having more than four channels. Optimal charging scheme for HSCSD in GSM is a topic for further research. From tables 4 and 5 it has been observed that as carrier goes from 1 to 3 :

1. The utilization of First Fit, Best Fit and Repacking increases rapidly from around 10% to 60%. This is due to the fact that for a given blocking probability more carriers means the system can have higher arrival rate and thus provide more free time-slots for incoming calls. In another word, the efficiency of usage of time-slots increases while the number of carriers increase.
2. Utilisation of Repacking is better than that of the other two schemes.
3. In case of only one carrier in the system, Table 4 shows that utilisation of Repacking is much better than First Fit and Best Fit. We observe that when the system has very limited resources, the allocation scheme turns out to be very important. As the number of carriers goes up, Repacking

does not seem to be much better than Best Fit especially in case of three carriers in the system. Best Fit works well when we have enough carriers.

## 5. SIMULATION RESULTS: FLEXIBLE CUSTOMERS

We now focus on the flexible customers case. Figure 1 demonstrates that in this case, Repacking is the worst performer. This is because under Repacking, because of the creation of large holes, more customers get larger chunks of capacity leaving no space for others. First Fit and Best Fit, on the other hand, have smaller holes (but more holes) and are forcing the network to allocate less than the required capacity to customers leaving more holes for others. Figures 2, 3 and 4 present results on comparison between LDP and HUP. In all three Figures the First Fit scheme is implemented. (Recall that under HUP, all three schemes give the same performance results). Comparing the blocking probability results presented in Figure 2 between LDP versus HUP clearly show higher blocking in the case of LDP because HUP accepts more calls at lower rate and hence the lower blocking.

In Figure 3 we present the utilization of HUP and LDP. The utilization values are calculated by fixing blocking probability levels and finding by simulation the maximal load which gives the required (allowable) blocking probability. We clearly observe a significant increase in utilization in favor of HUP for all levels of blocking probability. This is intuitively clear and it is consistent with the objective of HUP; however, the significant gain demonstrated here of 50% is important to notice.

The results of Figure 4 which demonstrate the significantly lower delay of the Low Delay Policy (LDP) over HUP are also important to notice. The tradeoff between high utilization and low blocking and cheaper calls under HUP versus low delay and more expensive calls under LDP.

## 6. CONCLUSIONS

This paper has studied by simulation the average blocking probabilities and utilisation for a range of channel allocation schemes for HSCSD over GSM. Systems with 1, 2, and 3 carriers were considered. We have observed that allowing large diversity of multiple time-slots (1 to 8), and assuming inflexible customers ( $B=b$ ), can lead to significant decrease in efficiency. Under the inflexible customers scenario, for HSCSD traffic, Repacking is found to incur the lowest blocking probability and highest utilisation, but it introduces higher complexity and excessive processing costs. Best Fit is better than First Fit and has a competitive performance in a system with three



carriers, but not as efficient as Repacking when the system resource is limited ( for example, one carrier in a cell ).

First Fit has been the worst performer. On the other hand, under the flexible customers scenario with Low Delay Policy, the results have been reversed. Repacking became the worst performer. The tradeoff between high utilization and low blocking and cheaper calls under HUP versus low delay and more expensive calls under LDP.

## References

- [1] Nokia's Vision for a Service Platform Supporting Circuit Switched Applications, White Paper  
<http://www.nokia.com/products/networks/cellular/gsm/person1.htm>.
- [2] M. Ivanovich, M.Zukerman, P. Fitzpatrick and M. Gitlits, "Channel allocation schemes for half and full rate connections in GSM", *Proceedings of IEEE International Conf. on Comm.*, Dallas, 1996.
- [3] M. Zukerman, "Circuit allocation and overload control in a hybrid switching system", *Computer Networks and ISDN Systems*, vol. 16, pp. 281-298, 1988/1989.
- [4] D. E. Knuth, *The Art of Computer Programming*, Volume 1: Fundamental Algorithms, NY: Addison Wesley, 1973.
- [5] M. Zukerman, "Bandwidth allocation for bursty isochronous traffic in a hybrid switching system", *IEEE Trans. on Comm.*, December 1989.
- [6] M. R. Garey and David S. Johnson, *Computers and Interactability: A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company, San Fransico, 1979.
- [7] R. S. Garfinkel, George L. Nemhauser, *Integer Programming*, USA, John Wiley & Sons, 1972.
- [8] ETSI GSM Specifications, Series 01-12.
- [9] D. Calin and D. Zeglache, Performance Analysis of High Speed Circuit Switched Data (HSCSD) over GSM, *Proceedings of ICC'98*, Atlanta, Georgia, June 1998.

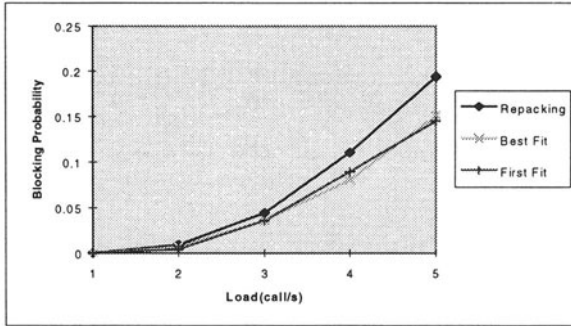


Figure. 1 Comparison of Blocking Probability of the three schemes under LDP with three frequency carriers

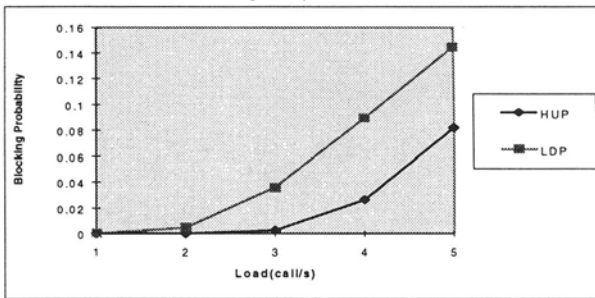


Figure. 2 Comparison of Blocking Probability between LDP versus HUP

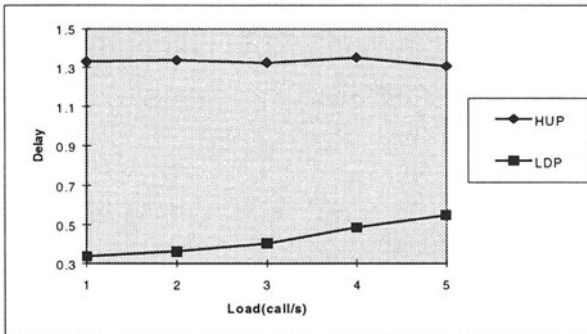


Figure. 3 Comparison of Utilisation achieved by LDP versus HUP

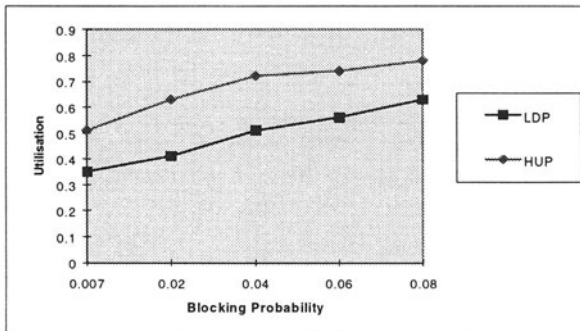


Figure 4 Comparison of Delay between LDP and HUP under First Fit