PERFORMANCE AND EFFICIENCY EVALUATION OF CHANNEL ALLOCATION SCHEMES FOR HSCSD IN 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

Several channel allocation schemes for High Speed Circuit Switched Data (HSCSD) over GSM are studied. The schemes differ in the way channels are packed (First Fit, Best Fit and Repacking), and in the connection admission policy. We study their blocking probability and utilisation. The performance of the schemes is studied with special focus on their simplicity/efficiency tradeoffs.

1. INTRODUCTION

High Speed Circuit Switched Data (HSCSD) over GSM service provides multi-slot high speed data service [1] using digital cellular mobile network. While the basic GSM is based on TDMA [8] and allocate one time slot for each voice connection, HSCSD allows for more than one slot per connection. In particular, several frequency carriers are allocated to a cell and each frequency carrier (or *frame*) supports eight TDMA channels henceforth referred to as time-slots. Under HSCSD, in principle, each data service can use between one and eight time-slots. In other words, a data service may occupy an entire frame. Although the standards support over four time-slots connections. Terminals supporting such connections should be more expensive as they need a separate transmitter and a separate receiver to be able to transmit and receive at the same time-slot (although in a different frequency). This may discourage use of more than four time-slot connections.

The time slots of a particular service must be consecutive within a TDMA frame. A service is not allowed to use time-slots from different frames.
We study three channel allocation schemes: First Fit, Best Fit, and Repacking. The aim is to focus on the simplicity/efficiency tradeoff. The study is based on simulations. Analytic solution is beyond the scope of this paper and readers who are interested in related analyses are 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 only accept that particular number. Flexible customers specify a range of required time slots, namely upper and lower bound. The service provider may allocate a number of time slots within that range. The flexible customers option leads to the situation whereby the service provider may choose to allocate less than the upper bound, although it has the capacity to allocate the upper bound. This way the service provider smoothens the traffic and improves utilization. This means that if the lower bound is one, the service provider may choose to allocate always one channel whereby ignoring the main premise of HSCSD. These traffic management alternatives are studied and discussed 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. Section 4 and 5 present simulation results and provide insight into the peculiarities of the different schemes and policies and their effects on network performance and Quality of Service (QoS) levels.

2. MODELING ASSUMPTIONS

As mentioned, each carrier can support several data services. In a GSM system, one channel within each cell is reserved for broadcasting, so only $8n$-1 time-slots are available for user traffic in an n carrier cell. We will consider cells with 1, 2 and 3 carriers, and hence 7, 15 and 23 available user-traffic time-slots, respectively [2].

We assume call arrivals to follow a Poisson process and to have exponential holding times. Let λ_i and $1/\mu_i$ be the Poisson arrival rate and holding time of calls that require i consecutive time-slots all in the same frequency carrier. The aim is to assign for each

0-7803-5796-5/99/$10.00 © 1999 IEEE

arrival, if admitted, the optimal available set of consecutive time-slots.

3. SERVICE OPTIONS CONSIDERED

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 upper bound and b the lower bound of acceptable capacity for the service. We consider three cases.

(1) **The inflexible customers option** – Here, $B = b$.

(2) **The flexible customers option** – Here, $B>b$ *and* $b=1$. For this option, we assume two extreme policies:

i) **Low Delay Policy (LDP)**

The network provides the highest amount of bandwidth possible not exceeding B, but may be lower than B, if B consecutive time slots in a frame are not available.

ii) **High Utilization Policy (HUP)**

The network always allocate one channel regardless of the B value. Recall that $b=1$.

4. CHANNEL ALLOCATION SCHEMES

We describe here three channel allocation schemes. The description in this section assumes the inflexible customers option. Nevertheless, it can easily apply to the flexible customers case: for the HUP, the distinction between the three channel allocation schemes is irrelevant all three options gives the same performance results. Under LDP, the implementation of the schemes will require an additional step of finding the maximal number of slots bounded by B which can be allocated.

4.1 First Fit (FF) [4]

We consider a cell with several frequency carriers (FC)s which are ordered and designated as FC 1, FC 2 etc. When a service which requires m time-slot arrives, under FF, we allocate to it the first m consecutive time-slots we find. We first look at FC 1, then in FC 2, etc. This channel allocation scheme is simplest to implement [5]. The time-slots are permanently allocated ID numbers as a two-dimensional array. For example, the first eight time-slots in the first carrier will have ID numbers (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 carrier may be in any one of the following possible states:

(a) no HSCSD data service in progress.

(b) any feasible combination of HSCSD data services each of which may occupy between one and eight time-slots.

The FF allocation algorithm then functions as follows: Each incoming data service which requires n time-slots will be allocated to n empty consecutive time-slots of which the first time slot has the lexicographically smallest ID-number among all EMPTY time-slots such that all n time-slots are in the same frequency carrier. The call must not be split across carriers. If such n consecutive time slots cannot be found, The HSCSD call will be blocked unless the flexible customers option allow less than n time slots allocation. No reordering of calls is performed at any time - a set of times slots remained assigned for that service until it terminates.

4.2 Best Fit (BF) [4][5]

Define a *hole* as a consecutive set of empty time-slots. Under this scheme for each incoming m time-slots service we try to find an exactly m slot hole. If such search fails, we search for an $m+1$ slot hole, etc. 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. Again, calls are not reordered.

4.3 Repacking [3][5]

Repacking starts like BF, but if a new call cannot find a hole, the call is not blocked. Instead, 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, only then 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.

5. SIMULATION RESULTS

5.1 Inflexible Customers Option

We begin with the inflexible customers option. Later, in the next subsection we present results for the flexible customers option. We shall discuss simplicity/efficiency tradeoffs, and gain insight into the bandwidth cost of providing the different services. The simulation is, in general, a discrete event simulation based on the event-by-event approach.

By simulation, we calculate, for each scheme, the blocking probabilities and the maximal utilisation subject to meeting blocking probability constraints in the case of one two and three carriers in a cell.

The arrival rate of the different services in any particular run are set to be equal (i.e. $\lambda_1 = \lambda_2 =..., =\lambda_8$). This represents a worst case scenario from the service provider point of view.

In the case of only one carrier, because one channel is reserved for broadcasting and signalling, an eight time-slot request is always blocked. Since $\lambda_1 = \lambda_2 =..., =\lambda_8$ the one carrier blocking probability must be higher than 1/8. Actually, very high blocking probability is observed also for two and three carriers under this worst case scenario. Clearly, the case with $\lambda_1 = \lambda_2 =..., =\lambda_8$ is theoretical and may not occur in practice too often. Nevertheless it signifies the worst case wastage and performance degradation.

We have already explain the high blocking probability observed in Figure 1. We also observe that, because FF and BF do not exhibit good packing compression, repacking performs better in all cases of one two or three carriers. Repacking leaves free time-slots as many as possible and efficiently utilizes them.

In the case of one carrier, in Figure 1 FF and BF exhibit very similar performance. This difference remains small (however somewhat more noticeable) in the case of 2 and 3 carriers in Figures 2 and 3. FF and BF leave holes upon departures of different calls and these holes may not be suitable for new calls. BF performs slightly better than FF as the number of carriers increases. This is because BF is trying to leave bigger holes following arrivals though it does not repack time-slots either. As the number of carriers increases, BF has greater selection to chose an optimal one.

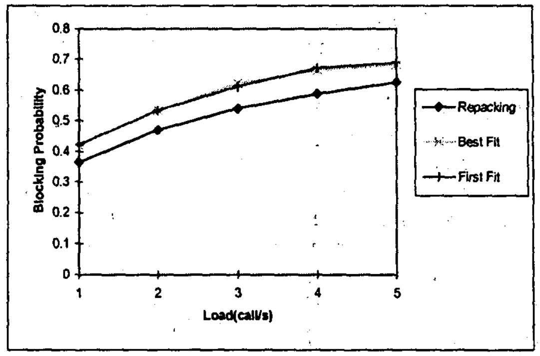

Figure 1 Blocking Probabilities of three schemes in case of one carrier

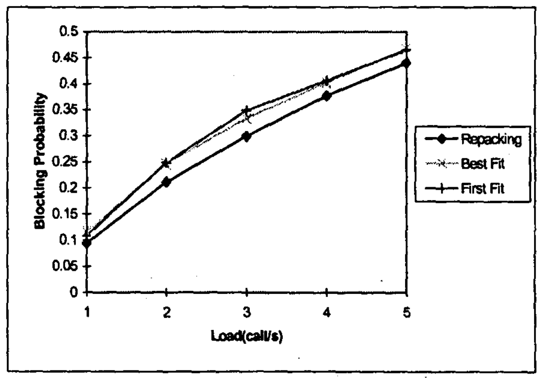

Figure 2 Blocking Probabilities of the three schemes in the case of two carriers

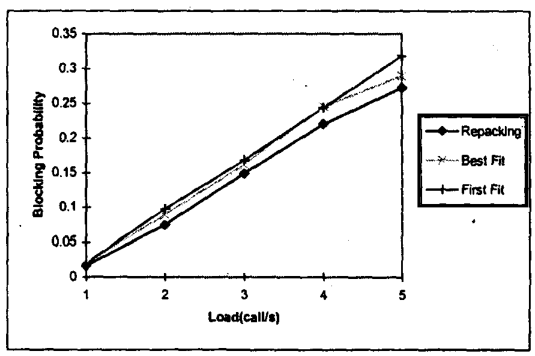

Figure 3 Blocking Probabilities of the three schemes in the case of three carriers

In Figures 4 and 5, we compare between the utilization of the three schemes again under the case of $\lambda_1 = \lambda_2 =..., =\lambda_8$. As we already know, this wide diversity of traffic leads to low performance or alternatively high wastage. This will be demonstrated in Figs 4 and 5. where we focus on *utilisation* defined as the maximal utilisation subject to meeting required blocking probability level in our case 20% or 2%. The 2% will only apply for the two and three carriers cases because as mentioned above, 1/8 of the traffic (namely the 8 time-slots service) is always blocked in one carrier case. The results presented in Figure 5 signify the enormous wastage caused by wide traffic diversity in HSCSD in the inflexible customers option.

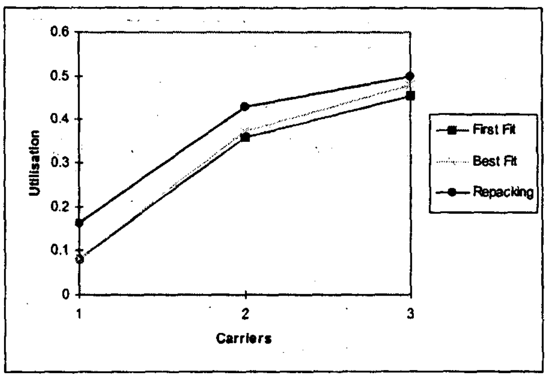

Figure 4 Utilisation of the three schemes in case of 1, 2 and 3 carriers (given 20% blocking probability)

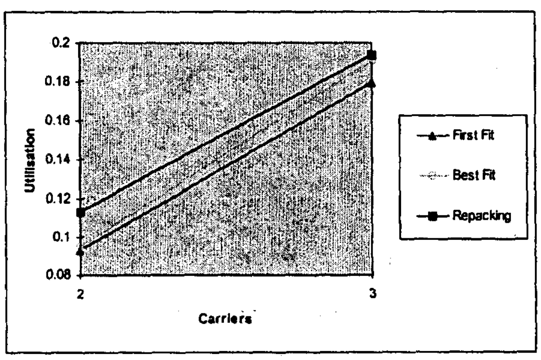

Figure 5 Comparison of Utilisation of three schemes in case of 2 and 3 carriers (given 2% blocking probability)

A clear observation of the results in Figures 4 and 5 is that allowing large diversity of services ($\lambda_1 = \lambda_2 =...,$ $=\lambda_8$), assuming inflexible customers ($B=b$), may lead to unacceptable wastage with significant cost implications. This is also true for Repacking. 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 Figures 4 and 5 it has been observed that as the number of carriers increases from 1 to 3 :

(1) The utilization of FF, BF and Repacking increases rapidly from around 10% to 60%. This is due to multiplexing gain.
(2) Utilisation of Repacking is better than that of FF and BF.
(3) Repacking is especially beneficial in case of only one carrier in the system, as multiplexing increases the benefit of efficient channel allocation is reduced.

5.2 Flexible Customers Option

In Figure 6, we observe that in case of flexible customers Repacking is the worst performer and leads to the highest blocking probability. This is explained by the observation that 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.

Figure 7 presents results on comparison between LDP and HUP under First Fit. Recall that under HUP, all three schemes give the same performance results.
Comparing the blocking probability results presented in Figure 7 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.

This means that HUP gives significant increase in utilization over LDP. Indeed, our simulation have shown gain of 50% in utilization in favor of HUP over LDP. This is intuitively clear and it is consistent with the objective of the High Utilization Policy (HUP); however, the significant gain demonstrated here of 50% is important to notice.

On the other hand, LDP gives better delay. Clearly, if a connection receives four time slots the data transmission will be completed four times quicker that if the same connection is assigned only one time slot. Now the tradeoff between high utilization and low blocking and cheaper calls under HUP versus low delay and more expensive calls under LDP is clear.

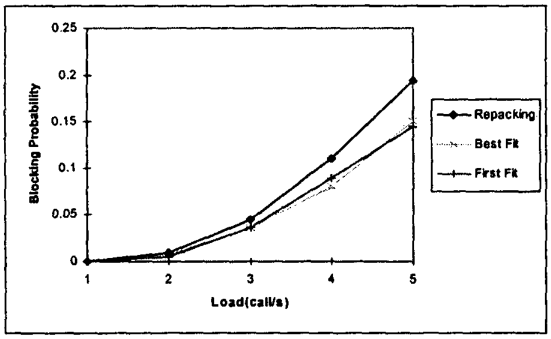

Figure 6 Blocking Probability of the three schemes under LDP with three frequency carriers

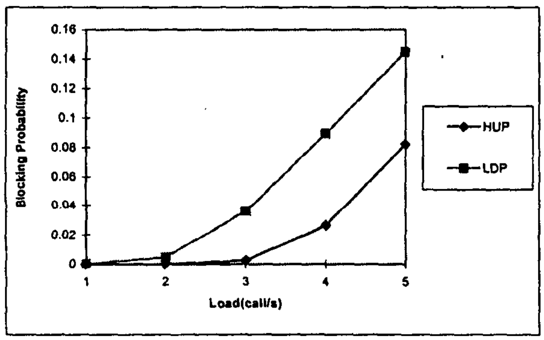

Figure 7 Blocking Probability for LDP and HUP under First Fit

6. CONCLUSIONS

The most important conclusion of this paper from a practical point of view is that if a telecommunications provider provides HSCSD, and if cell capacity is limited to a small number of single slot channels, then provision of HSCSD service may be very costly especially if more than four time slots are provided. The increase in cost of provision multiple time slot channels increases far more than linearly with the connection rate provided. We have described three channel allocation schemes and several connection admission policies related to HSCSD over GSM. From simulations 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 and repacking is found to be more efficient than First Fit and Best Fit but it is more complex to implement and requires more processing. Best Fit is slightly better than First Fit. First Fit has been the worst performer. On the other hand, under the flexible customers option with LDP, the results have been reversed. Repacking became the worst performer. HUP exhibits higher utilization lower blocking and cheaper calls but longer delay than LDP.

Acknowledgement

The authors would like to thank M. Gitlits for fruitful discussions.

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 ICC '96*, 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.*, Dec. 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. Zeghlache, Performance Analysis of High Speed Circuit Switched Data (HSCSD) over GSM, *Proceedings of ICC'98*, Atlanta, Georgia, June 1998.