Optimal Fit of Arbitrary Sized Segments

I. P. Page
Computer Science Department, Brunel University, Uxbridge, Middlesex, UK

The frequency with which the first fit placement policy chooses the best possible hole size is investigated, together with the mean time taken to scan the free store list. These results are compared with the performance of an optimal placement policy first proposed by Campbell. The results indicate that first fit performs better than the optimal policy in both respects. This is due to the tendency of first fit to arrange the segments in size order.

INTRODUCTION

The problem of organizing a free storage area containing arbitrary sized units occurs in purely segmented virtual memory systems, Algol 68 *heap* storage and many other list processing situations. The majority of published work relating to this problem is directly concerned with the virtual memory application, where segments of differing sizes must be placed on demand. Management of such storage areas is achieved by 'placement' and 'replacement' policies. Most schemes result in the fragmentation of free storage into a series of holes interspersed between the segments. The placement scheme has to choose a hole which is at least as large as the incoming segment. When placement fails, the replacement scheme removes segments, considered least likely to be of immediate future value, until sufficient space has been liberated to facilitate placement. A good review of these policies may be found in Denning.[1]

Two placement schemes that have received much attention are best fit (BF) and first fit (FF). Knuth reported that BF was unsuitable as it created many small holes which could not be employed in future placements.[2] He also reasoned that the number of holes would on average tend to half the number of segments (Knuth's 50% rule). The rule has been shown to apply reasonably well to both BF and FF,[3,4] which suggests that the total unused memory obtained with BF will be lower than that of FF. This is true apart from unusual situations which results in Knuth's criticism of BF being a recommendation in disguise.[5] The early work on virtual memory systems contains many examples of false logic and unsupported hypotheses.

As FF scans the list of holes in address order, Denning suggested that the early part of the list would comprise very small holes,[1] thus unnecessarily increasing the mean search length. He proposed a modified form of the algorithm called next fit which commenced the search from an incrementally progressive point within a circular hole list. This was later found to destroy a most useful property of FF and therefore greatly reduce the efficiency of the algorithm.[6] It was Shore who discovered that FF tended to arrange segments in order of increasing size.[5] He found that this resulted in only a 2–3% difference in the store utilization achieved between BF and FF. Next fit does not give rise to the size ordering property and hence produces rather more unused memory.

Campbell introduced an 'optimal' placement scheme and compared its search length to that of FF.[7] His analysis was based on the false assumption that FF chose randomly from the set of suitably large holes. The size ordering later discovered by Shore enables FF to select the best fit with much higher frequency than could be expected from Campbell's assumption. Simulation studies reported here show that FF outperforms Campbell's optimal policy with respect to the frequency of best fits and the mean search length of the free store list.

THE OPTIMAL PLACEMENT POLICY

Campbell noted the similarity between the placement problem and the task of a weary cyclist attempting to select the best from a choice of N hotels. An optimal strategy was found by Dynkin and Yushkevich.[8] The cyclist should pass the first $k - 1$ hotels, where k is an integer given by the double inequality

$$\sum_{i=k}^{N-1} \frac{1}{i} \leq 1 < \sum_{i=k-1}^{N-1} \frac{1}{i}$$

The cyclist then selects the next hotel which is better than any seen so far. The probability of this being the best hotel is given by

$$p = \frac{(k-1)}{N} \sum_{m=k}^{N} \frac{1}{(m-1)}$$

This tends to $1/\varepsilon = 0.368$ as N tends to infinity. Campbell states that this is much better than the success rate of FF, which he incorrectly predicts will produce a best fit probability of $1/N$.

There are two problems with Campbell's assumptions. First, the two problems are only isomorphic if some of the hotels are fully booked, since some of the holes will be too small to accommodate the segment being placed. The more important problem is that the ordering property improves the performance of FF considerably.

SIMULATION STUDIES WITH FIRST FIT

A simulation was performed employing Knuth's model.[2] On each step of the simulation a new segment of size S and lifetime L is introduced. After L subsequent steps the segment is removed. The model was run at low levels of contention to avoid overflow when a placement could not be made. For the experiments reported here the model was run with a memory size of $M = 1024$ words

CCC–0010–4620/82/0025–0032$01.00

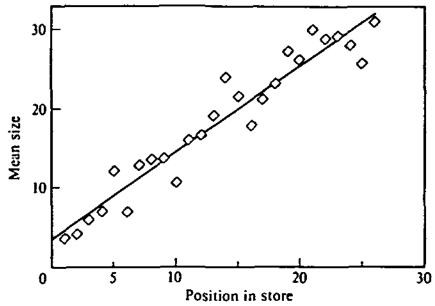

Figure 1. Mean segment size against store position with regression line slope = 1.092 correlation coefficient = 0.96.

and with uniform random segment size and lifetime, $0 < S \leq S_{\max}$, $0 < L \leq L_{\max}$. Figure 1 illustrates the segment size ordering effect of FF. The model was run for 1000 steps and snap-shots of the store were taken every 100 steps, and the mean size of segments and holes computed for each store position. Figure 1 suggests a linear relationship between segment size and store position, but this is only likely for uniform random segment size distribution and $S_{\max} \ll M$. The hole sizes do not follow such a linear trend, but the ordering property is clearly evident in Fig. 2.

These experiments also revealed that FF tends to choose the best fit hole much more frequently than predicted by Campbell. Table 1 gives the mean search length and the best fit frequency of FF. It is seen that the best fit frequency of FF is considerably higher than $1/N$, and rather better than that of the optimal scheme.

SEARCH LENGTH

The optimal placement scheme passes $k - 1$ holes before beginning the selection process. The total scan length may be estimated as follows. If there are N distinct hole sizes, all of which are sufficiently large to accommodate the incoming segment, then the probability that the k_{th} is as good or better than any of the previous $k - 1$ is

$$p_k = \frac{1}{N^k} \sum_{i=k}^{N-1} i^{k-1}$$

The probability of a search of d extra holes is then given by the geometric distribution, $p_{k+d} = q_k^{(d-1)} \cdot p_k$ where $q_k = 1 - p_k$. The mean extra search length is then $1/p_k$. For 10 different hole sizes, $k \geq 4$ and the mean extra search length is 4.94, giving a total search length of at least 8.

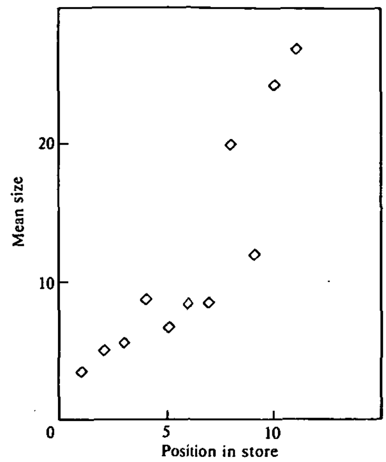

Figure 2. Mean hole size against store position.

If a set of N differently sized holes is considered, all larger than the incoming segment, then $p_k = 1/k$ and the total search length is $2k - 1$. For 10 holes the total search length would be 7. Neither of these cases applies to FF, since the ordering property greatly improves the search time. If the hole size distribution exactly matched that of the segments and the N holes were completely ordered then the mean search length of FF would be $N/2$. The results in Table 1 show that the mean search length of FF approaches $N/2$ as N increases. Applying Knuth's 50% rule, we note that the mean search length of FF tends to one quarter of the number of segments. A further exposition of the behaviour of first fit and best fit may be found in Ref. 3.

Table 1. Search length and best fit frequency of first fit

$S_{\max}$	$L_{\max}$	Mean search length	Correct choice frequency	Mean number of holes	Optimal scheme's best fit rate
120	10	1.82	0.91	2.72	0.5
100	17	2.56	0.80	4.15	0.45
80	25	3.37	0.69	6.15	0.43
40	55	6.27	0.59	12.2	0.39
40	59	6.54	0.59	13.1	0.39

CONCLUSIONS

Simulation experiments with first fit have indicated that the frequency of best fit placements and the search length of the free storage list are better than those obtained by the optimal policy proposed by Campbell. The advantage is due to the segment size ordering effect of first fit noted by Shore.

REFERENCES

1. P. J. Denning, Virtual memory systems. *ACM Computing Surveys* **2** (No. 3), (1970).
2. D. Knuth, *Fundamental Algorithms*, Addison Wesley, Reading, Massachusetts (1972).
3. I. P. Page, An analysis of certain storage allocation policies, PhD Thesis, Brunel University, Uxbridge, Middlesex (1979).
4. J. E. Shore, Anomalous behaviour of the fifty percent rule. *Communications of the ACM* **20** (No. 11), (1977).
5. J. E. Shore, On the external storage fragmentation produced by first fit and best fit allocation policies. *Communications of the ACM* **18** (No. 8), (1975).
6. C. Bays, A comparison of next fit, first fit and best fit. *Communications of the ACM* **20** (No. 3), (1977).
7. J. A. Campbell, A note on an optimal fit method for dynamic storage allocation. *Computer Journal* **14** (No. 1), (1971).
8. E. B. Dynkin and A. A. Yushkevich, *Markov Processes: Theory and Practice*, Plenum Press, New York (1969).

Received July 1980