Minimizing Page Fetches for Permuting Information in Two-Level Storage*

Part 1. Generalization of the Floyd Model

TAKAO TSUDA**, TAKASHI SATO**·*** and TAKAAKI TATSUMI**

The Floyd model on permuting information in two-level storage is generalized in such a way that the fast memory is large enough to allow reading in w pages ($w \geqq 2$). Properties of the generalized e-function are discussed. In the second half of this paper the lower bounds of the number of page fetches are analyzed for the case of an arbitrary permutation.

1. Introduction

In case a large amount of data needs to be processed, the performance of algorithms hinges on the number of page transfers between the fast main memory and the slow but much larger memory, such as magnetic disks. Instructive results in this area are due to Floyd [1]. He assumed that the fast memory is large enough to allow fetching two pages, and discussed the problem of reading in two pages from the slow memory, permuting the records between the pages and then reading out those pages with shuffled records. This process is continued until the desired permutation of records is realized in the slow memory. He established the lower bound on the number of page transfers with which all the records in the slow memory can be redistributed among pages.

In this note we generalize the Floyd model in such a way that Floyd's assumption can be replaced by a more general one that the fast memory is large enough to allow reading in w pages ($w \geqq 2$). To make the model more realistic, an extra page of work space is provided in main memory; in this workspace a new page is constructed from those records in the resident w pages, and it is subsequently read out to the corresponding slow memory locations. Thus the model of this note can be shown schematically as in Fig. 1.

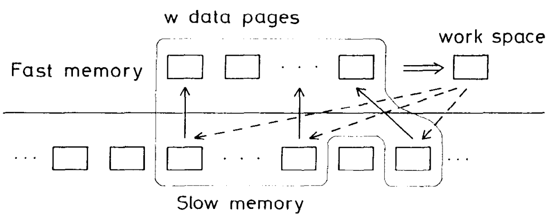

Fig. 1 The generalized Floyd Model.

This generalization may be useful when constructing algorithms for processing a very large amount of data over hierarchical memories with minimal 'use times', rather than CPU times, because the arbitrarily given integer parameter w may approximately represent the finite size of main memory available for data. Practical applications in this direction have been discussed elsewhere [2], where detailed algorithmic procedures are compared against the theoretical limits.

2. Definition of Operations

To consider distributing records among pages, suppose that without loss of generality there are p' pages in slow memory, each page containing p records. Appropriately selecting w pages out of those p' pages, these are transferred to the main memory. We call this a w-page transfer. As a subset of the union of those w pages, a new page of p (or less) records is formed in the work space, and then it is read out to the slow memory. This process is repeated w times, so that the new w pages occupy the locations of the old w pages in slow memory. With the shuffling process inclusive, the w-page transfer is defined to be w 'operations'. The problem is to find the lower bound to the number of page transfers or operations required of redistributing all the records of p' pages to realize the prescribed desirous distribution. Normally, the page-transfer time exceeds the typical CPU time by orders of magnitude, so that even in practical applications the number of page transfers can be a good complexity measure.

3. V-function and e-function

The V-function is defined as follows.

$$V = V(A|B) = \sum_i \sum_j e(X(i,j)) \tag{1}$$

where $X(i,j)$ is the number of records that are transmitted from the i^{th} page of the initial distribution of records, A, to the j^{th} page of the final distribution of records, B ($i, j = 1, 2, \cdots, p'$). The e-function in the above is defined as

*Part of Part 1 appeared in *Inf. Process. Lett.*, Vol. 16, No. 4, pp. 183–188 (1983) as a preliminary announcement of this two-part paper. Part 3 is in preparation.

**Department of Information Science, Kyoto University, Kyoto, Japan.
***On leave from the Department of Radio Engineering and Operation, Takuma Radio Technical College, Takuma-cho, Kagawa, Japan.

$$e(w^k+l)=kw^k+\left(k+\frac{w}{w-1}\right)l, \qquad (2)$$

where

$$k=0, 1, 2, \cdots, \qquad (3)$$

$$0 \leqq l<(w-1)w^k \qquad (4)$$

and

$$e(0) \triangleq 0. \qquad (5)$$

Obviously,

$$e(1)=0, \qquad (6)$$

$$e(n)=n \log_w n \quad \text{(if } n \text{ is a power of } w). \qquad (7)$$

The following holds.

Lemma 1. Given X_t $(t=1, 2, \cdots, w)$ such that

$$X_t=w^{k_t}+l_t, \qquad (8)$$

where

$$k_t=0, 1, 2, \cdots, \qquad (9)$$

and

$$0 \leqq l_t<(w-1)w^{k_t}, \qquad (10)$$

then

$$e\left(\sum_{t=1}^{w} X_t\right) \leqq \sum_{t=1}^{w} \{e(X_t)+X_t\}. \qquad (11)$$

Proof. Without loss of generality, we can assume

$$X_1 \geqq X_2 \geqq \cdots \geqq X_w. \qquad (12)$$

$w^{k_t} \leqq X_t<w^{k_t+1}$ by (8), so that if $X_{t'} \leqq X_{t''}$ $(t' \neq t'')$, then $k_{t'} \leqq k_{t''}$. It follows therefore from (12) that

$$k_1 \geqq k_2 \geqq \cdots \geqq k_w. \qquad (13)$$

Moreover,

$$\sum_{t=1}^{w} \{e(X_t)+X_t\} = \sum_{t=1}^{w} (k_t+1)w^{k_t}+ \sum_{t=1}^{w} \left(k_t+1+\frac{w}{w-1}\right)l_t. \qquad (14)$$

Proof proceeds by dividing into two cases.

Case A: $\sum_t X_t \geqq w^{k_1+1}$.

Since $w^{k_1+1} \leqq \sum_t X_t<w^{k_1+2}$, we have

$$e(\sum_t X_t)=e(w^{k_1+1}+(\sum_t X_t-w^{k_1+1}))$$
$$=-\frac{w}{w-1}w^{k_1+1}+\left(k_1+1+\frac{w}{w-1}\right)(\sum_t (w^{k_t}+l_t)). \qquad (15)$$

It follows from (14) and (15) that

$$\Delta_A \triangleq \sum_t \{e(X_t)+X_t\}-e(\sum_t X_t)$$
$$=\frac{w}{w-1}w^{k_1+1}-\sum_t \left(k_1-k_t+\frac{w}{w-1}\right)w^{k_t}-\sum_t (k_1-k_t)l_t.$$

Using the relation

$$\sum_t (k_1-k_t)l_t \leqq \sum_t (k_1-k_t)(w-1)w^{k_t}$$

(where equality holds for $k_1=k_2=\cdots=k_w$), we have

$$\Delta_A \geqq \frac{w}{w-1}[w^{k_1+1}-\sum_t ((w-1)(k_1-k_t)+1)w^{k_t}]. \qquad (16)$$

It is not difficult to show that

$$w^v \geqq (w-1)v+1 \qquad (17)$$

(where $v=0, 1, 2, \cdots$; equality holds for $v=0, 1$). An application of this inequality gives $w^{k_1-k_t} \geqq (w-1) \times (k_1-k_t)+1$, so that

$$w^{k_1} \geqq ((w-1)(k_1-k_t)+1)w^{k_t}, \qquad (18)$$

the equality being either for $k_1=k_t$ or $k_1=k_t+1$. It follows from (16) and (18) that

$$\Delta_A \geqq \frac{w}{w-1}\left(w^{k_1+1}-\sum_{t=1}^{w} w^{k_1}\right)=0,$$

which means

$$e(\sum_t X_t) \leqq \sum_t (e(X_t)+X_t),$$

where equality is for $k_1=k_2=\cdots=k_w$.

Case B: $\sum_t X_t<w^{k_1+1}$.

$$e(\sum_t X_t)=e(w^{k_1}+(\sum_t X_t-w^{k_1}))$$
$$=-\frac{w}{w-1}w^{k_1}+\left(k_1+\frac{w}{w-1}\right)\sum_t w^{k_t}$$
$$+\left(k_1+\frac{w}{w-1}\right)\sum_t l_t.$$

This, together with (14), yields

$$\Delta_B \triangleq \sum_t \{e(X_t)+X_t\}-e(\sum_t X_t)$$
$$=\frac{w}{w-1}w^{k_1}-\sum_t \left(k_1-k_t-1+\frac{w}{w-1}\right)w^{k_t}$$
$$-\sum_t (k_1-k_t-1)l_t.$$

In inequality (13), we consider the first subscript i such that

$$k_1=k_2=\cdots=k_i>k_{i+1} \geqq k_{i+2} \geqq \cdots \geqq k_w. \qquad (19)$$

Note that in the absence of such an i the case is that $k_1=k_2=\cdots=k_w$, which reduces to Case A, because

$$\sum_t X_t \geqq w^{k_1+1}$$

holds. Therefore,

$$\Delta_B=\frac{w}{w-1}w^{k_1}-\frac{i}{w-1}w^{k_1}+\sum_{t=1}^{i} l_t$$
$$-\sum_{t=i+1}^{w} \left(k_1-k_t-1+\frac{w}{w-1}\right)w^{k_t}$$
$$-\sum_{t=i+1}^{w} (k_1-k_t-1)l_t.$$

Noting that

$$\sum_{t=1}^{i} l_t \geqq 0$$

(where equality is for $l_1=l_2=\cdots=l_i=0$) and that

$$\sum_{t=i+1}^{w} (k_1-k_t-1)l_t \leqq \sum_{t=i+1}^{w} (k_1-k_t-1)(w-1)w^{k_t}$$

(where equality holds for $k_{i+1}=k_{i+2}=\cdots=k_w=k_1-1$), we have

$$\Delta_B \geqq \frac{1}{w-1} \times \left[(w-i)w^{k_1}-\sum_{t=i+1}^{w} \{(w-1)(k_1-k_t-1)+1\}w^{k_t+1}\right]. \qquad (20)$$

Again, using (17), we have

$$w^{k_1} \geqq \{(w-1)(k_1-k_t-1)+1\} w^{k_t+1},$$

which, when combined with (20), results in

$$\Delta_B \geqq \frac{1}{w-1}\left[(w-i)w^{k_1} - \sum_{t=i+1}^{w} w^{k_t}\right]=0.$$

Namely, we have arrived at:

$$e\left(\sum_{t=1}^{w} X_t\right) \leqq \sum_{t=1}^{w} \{e(X_t)+X_t\},$$

where the equality sign holds for the case that $k_1=k_2=\cdots=k_i=k_{i+1}+1=k_{i+2}+1=\cdots=k_w+1$ and $l_1=l_2=\cdots=l_i=0$. (*Q.E.D.*)

4. Lower Bound of the Number of Page Transfers

The following theorem follows from the preceding lemma.

Theorem 1. Selecting w pages from the set of p' pages in the slow memory (where each page contains p records), we permute the records on main memory. The new w pages thus obtained then replace the old w pages in the slow memory. By repeating this process each with a w-page transfer, we want to realize a final distribution of records, B, among the p' pages of slow memory. In this process a change of distribution from A to A' resulting from one w-page transfer corresponds to a change of the V-function such that

$$V(A'|B)-V(A|B) \leqq wp; \tag{21}$$

in other words, the increase of V due to a one-page transfer (i.e. an operation) is at most p.

Proof. Let all the p' pages in slow memory be identified by integers $1, 2, \cdots, p'$. Also let those w pages selected be referred to by $i_1, i_2, \cdots, i_w$; therefore, $\{i_1, i_2, \cdots, i_w\} \subset \{1, 2, \cdots, p'\}$. In the i^{th} page of the w pages selected, we denote the set of records that should be in the j^{th} page of the final distribution B by $X(i,j)$ $(i=i_1, i_2, \cdots, i_w; j=1, 2, \cdots, p')$. For example, $X(i_1, 1)$ indicates the set of those records of page i_1 now in main memory that should have been transferred to page 1 when the final distribution B is established. For simplicity we understand that $X(i,j)$ also denotes the number of the records that make the set $X(i,j)$ depending on the context of discussions. We move the records between the w pages fetched by a w-page transfer in such a way that those records with a common destination page get together in some of the new w pages. The new w pages are then pushed to slow memory. More exactly, we need at least one page for work space, in which a new page is formed and then copied back in slow memory by a one-page transfer. This process is repeated w times, while the w pages in main memory remain intact. Thus the change of distribution from A to A' by a w-page transfer (or w operations) causes the change of V-function such that

$$V(A'|B)-V(A|B)=\sum_{j=1}^{p'} e\left(\sum_{r=1}^{w} X(i_r,j)\right)-\sum_{j=1}^{p'}\sum_{r=1}^{w} e(X(i_r,j)) \leqq \sum_{j=1}^{p'}\sum_{r=1}^{w} X(i_r,j),$$

where the previously-shown lemma has been used. The number of records to be transmitted to the p' pages of the final distribution from an arbitrarily selected set of w pages is wp, so that the above equation gives

$$V(A'|B)-V(A|B) \leqq wp. \tag{21'}$$

Next, consider pages i_1 and i_2 of the w pages. Writing a and b instead of $X(i_1,j)$ and $X(i_2,j)$ for short, respectively, we can easily show that

$$e(a+b) \geqq e(a+b-1)+e(1) \geqq e(a+b-2)+e(2) \geqq \cdots \geqq e\left(\left\lceil\frac{a+b}{2}\right\rceil\right)+e\left(\left\lfloor\frac{a+b}{2}\right\rfloor\right). \tag{22}$$

This endorses the fact (already used in the present proof) that in order to increase the e-function maximally it suffices to have those records with a common destination page get together in the same page while forming a new set of w pages by inter-page shuffling. If, therefore, the change of distribution from A to A' is realized by *whatever* inter-page shuffling, inequality (21′) holds. □

The following corollary is the direct consequence of the above theorem.

Corollary 1. There are p' pages in slow memory each of which contains p records. The problem is to redistribute the records among p' pages without loss of any information. The distribution of records can be represented by a $p' \times p'$ matrix. The lower bound of the number of one-page transfers to have the final distribution B from the initial distribution A is given by

$$[V(B|B)-V(A|B)]/p. \quad \square$$

As an example, let A be a $p \times p$ matrix where p is some power of w, with each row per page, and B its transposed form. Then,

$$V(B|B)=\sum_i \sum_j e(X(i,j))=\sum_i e(X(i,i))=p\, e(p)=p^2 \log_w p,$$

$$V(A|B)=\sum_i \sum_j e(1)=0,$$

so that the lower bound for a matrix transposition is $p \log_w p$.

Remark. It can be demonstrated that Eq. (22) is generalized as follows:

$$e\left(\sum_{t=1}^{w} X_t\right) \geqq \sum_{t=1}^{w} e(X_t), \tag{23}$$

where X_t and w are those notations defined in Lemma 1.

Lemma 2. For a fixed value of $x>0$, let the following function of y be defined:

$$f_x(y) \triangleq x \log_w (x-y)+\frac{w}{w-1}y, \tag{24}$$

where

$$0 \leqq y < \frac{w}{w-1}x. \tag{25}$$

It then follows that

$$f_x(y) \leqq x \log_w\left(\frac{w-1}{w \log w}x\right)+\frac{w}{w-1}\left(1-\frac{w-1}{w \log w}\right)x \tag{26}$$

where $f_x(y)$ assumes the maximum value shown on the right side of (26) at

$$y=\left(1-\frac{w-1}{w\log w}\right)x. \qquad (27)$$

Lemma 3. If $w \geqq 2$, $0 \leqq q_i \leqq 1$ ($i=1, 2, \cdots, n$) and

$$\sum_{i=1}^{n} q_i = 1,$$

then the following inequality holds:

$$\sum_{i=1}^{n} q_i \log_w q_i \leqq 0,$$

where $0 \log_w 0 \triangleq 0$ and equality holds when $q_i = \delta_{i,i_0}$ ($i=1, 2, \cdots, n$; i_0 is one of these n integers). □

Denoting X_{ij} in place of $X(i, j)$, one has

$$0 \leqq l_{ij} < \frac{w-1}{w} X_{ij}, \qquad (28)$$

hence (25), since by definition the argument of the e-function is subject to relations

$$X_{ij} = w^{k_{ij}} + l_{ij},$$
$$k_{ij} = 0, 1, 2, \cdots,$$
$$0 \leqq l_{ij} < (w-1) w^{k_{ij}}.$$

Theorem 2. Let $p' \geqq p$, $w \geqq 2$ and the following function be defined:

$$V(X) \triangleq \sum_{i=1}^{p'} \sum_{j=1}^{p'} \left\{ X_{ij} \log_w (X_{ij} - l_{ij}) + \frac{w}{w-1} l_{ij} \right\}, \qquad (29)$$

where

$$\sum_{i=1}^{p'} X_{ij} = \sum_{j=1}^{p'} X_{ij} = p, \qquad (30)$$

$$0 \leqq l_{ij} < \frac{w}{w-1} X_{ij}. \qquad (28')$$

It then follows that

$$0 \leqq V(X) \leqq p'p\left(\log_w \frac{p(w-1)}{ew\log w} + \frac{w}{w-1}\right), \qquad (31)$$

where the right equality holds when

$$l_{ij} = \left(1 - \frac{w-1}{w \log w}\right) X_{ij} \qquad (27')$$

and the $p' \times p'$ matrix, (X_{ij}), is either

$$\begin{pmatrix} p & & & \\ & p & & 0 \\ & & \cdot & \\ & & & \cdot \\ 0 & & & & \cdot \\ & & & & & p \end{pmatrix}$$

or that matrix given by an arbitrary permutation of the columns of this diagonal matrix.

Proof.

$0 \leqq V(X)$ (obvious)

$$\leqq \sum_{i=1}^{p'} \sum_{j=1}^{p'} \left\{ X_{ij} \log_w \left(\frac{w-1}{w \log w} X_{ij}\right) + \frac{w}{w-1}\left(1 - \frac{w-1}{w\log w}\right) X_{ij} \right\} \qquad \text{(by lemma 2)}$$

$$= \sum_{i=1}^{p'} p \sum_{j=1}^{p'} \frac{X_{ij}}{p} \log_w \frac{X_{ij}}{p} + \left(\log_w \frac{p(w-1)}{ew\log w} + \frac{w}{w-1}\right) \times \sum_{i=1}^{p'} \sum_{j=1}^{p'} X_{ij}$$

$$\leqq p'p\left(\log_w \frac{p(w-1)}{ew \log w} + \frac{w}{w-1}\right). \qquad \text{(by lemma 3)} \ \square$$

Table 1. Values of function $\delta(w)$

w	$\delta(w)$
2	0.086
5	0.194
10	0.269
30	0.371
50	0.411
100	0.459
∞	1

Writing

$$\log_w \frac{p(w-1)}{ew \log w} + \frac{w}{w-1} \triangleq \log_w p + \delta(w),$$

one then has

$$\delta(w) = \log_w \frac{w-1}{ew\log w} + \frac{w}{w-1}, \qquad (31)$$

a monotonously increasing function of w, as shown in Table 1. The function $\delta(w)$ does not exceed unity for any finite value of w.

Corollary 2. Given an initial and final distributions of records, A and B, respectively, the following holds:

$$[V(B|B) - V(A|B)]/p \leqq [\underset{ij}{\text{Max}}\ V(X_{ij}) - \underset{ij}{\text{Min}}\ V(X_{ij})]/p$$
$$= p'[\log_w p + \delta(w)]; \qquad (32)$$

namely, the supremum of the lower bounds to the number of page fetches, required of realizing an arbitrarily given permutation of records, does not exceed the value given by the right-hand side of (32). In rare cases (see theorem 2) the equality in the middle of (32) holds approximately.* □

Remark. The supremum value of lower bounds given in corollary 2 is that number of page fetches which should hopefully be attained by some near best possible algorithm.

Acknowledgment

We thank Prof. T. Ichinose of the Mathematics Department of Hokkaido University for the derivation of theorem 2.

References

1. Floyd, R. W., Permuting information in idealized two-level storage in *Complexity of Computer Computations* (R. Miller and J. Thatcher, editors), pp. 105–109, Plenum Press, New York (1972).
2. Tsuda, T. and Sato, T., Transposition of Large Tabular Data Structures with Applications to Physical Database Organization. Part 1. Transposition of Tabular Data Structures, *Acta Inf.*, Vol. 19, pp. 13–33 (1983); Tsuda, T., Urano, A. and Sato, T., Transposition of Large Tabular Data Structures with Applications to Physical Database Organization. Part 2. Applications to Physical Database Organization, *Acta Inf.*, Vol. 19, pp. 167–182 (1983).

(Received April 1, 1982)

*Note that l_{ij} (and X_{ij} as well) is not a continuous variable, but assumes zero or integer values only.