動的配列シリンダディスクの効果分析

4P-5

岡本　正昭　　　石井　博昭
大阪ガス 情報通信部　　大阪大学 工学部

1.はじめに

高度情報化社会に進み情報処理の重点は、大量データの加工蓄積から良質な情報や付加価値を追求する段階に移行しつつある。

補助記憶システムもこの時代の要求にふさわしく、膨大な蓄積情報から、真に有効な情報を短時間に検索提供できるものでなければならない。

先に筆者は、磁気ディスクの書き替え可能特性を生かしてデータを動的に並べ替えると、ヘッドシーク時間が短縮されて既存のディスクでも高速にデータをアクセスできることを指摘した。[1]

そこで提案した「動的配列シリンダディスク(DACYD-Dynamically Arranged Cylinder Disk)」は、一種の知的記憶システムと言える。(図1参照)

今回は、DACYDの高速化効果を分析し新しい情報記憶システムとして大いに有望なことを示めす。

図　1

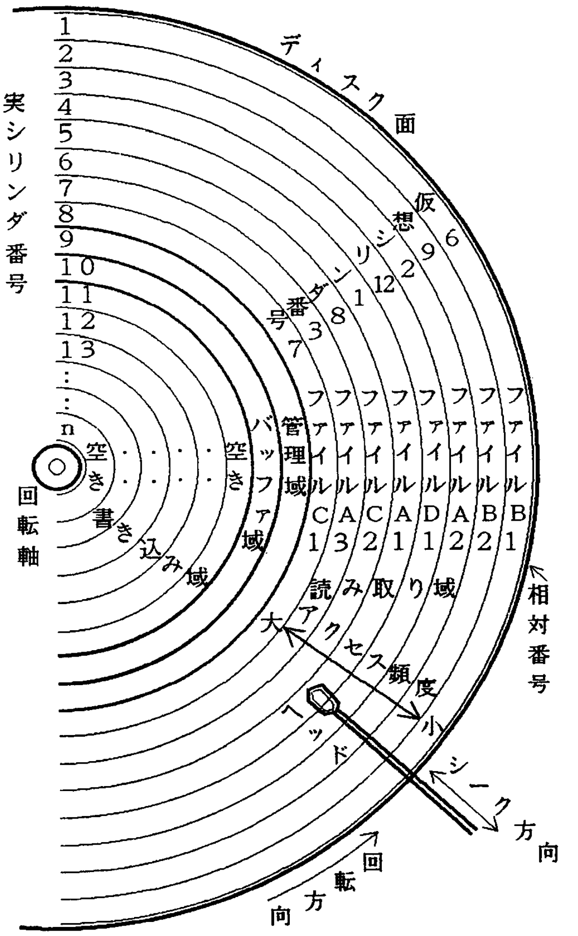


図　2

アクセス頻度分布(一般形)

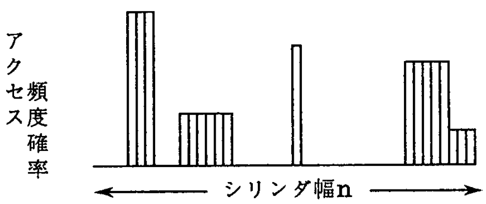


図　3

アクセス頻度分布(DACYD)

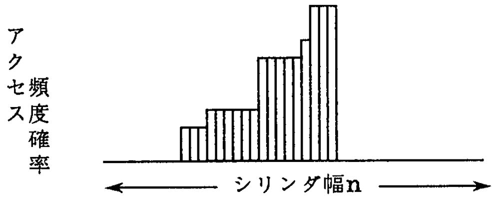
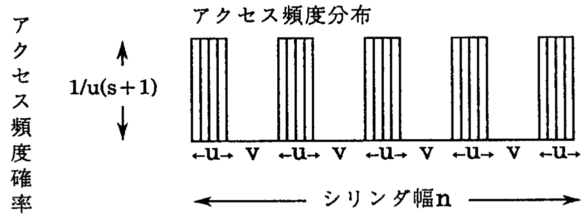


図　4

アクセス頻度分布



2.圧縮効果

ディスクのシリンダ番号を横軸にとり、各シリンダに対するアクセス頻度を縦軸にとると、一般に図2の頻度分布グラフが得られる。これをDACYDで記録すると、各シリンダのデータはアクセス頻度順に連続して詰められ図3のグラフになる。

両者を比較すると、DACYDは隙間のシリンダを取り除く圧縮効果とアクセス頻度順に並べる傾斜効果でヘッドシーク距離が削減できることがわかる。

文献[1]から図4の分布モデルでは平均シーク長は

$$L=\{u(n-u)(s+2)+(u+1)(u-1)\}/3u(s+1) \quad (1)$$

となるが、

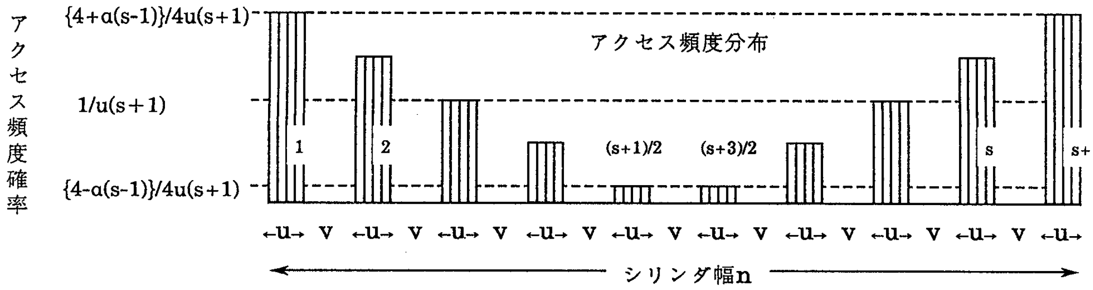
$$n=u(s+1)+vs \quad (2)$$

の関係から

Efficiency of Dynamically Arranged Cylinder Disk

Masaaki OKAMOTO[1], Hiroaki ISHII[2]
1 OSAKA GAS, Ltd. 2 OSAKA University

図　5



$$L=n/3+(v/3)\times(1-1/uvs)/(1+1/s) \quad (3)$$

と変形し、$u\geqq1$、$v\geqq1$、$s\geqq1$を考慮すれば

$$n/3<L<n/3+v/3 \quad (4)$$

となるので、Lはアクセスされるシリンダの両端の距離で決まる。すなわち平均シーク長は、シリンダの圧縮割合に比例する。

3.傾斜効果

図5の様にアクセス頻度分布が2つの一次式で変化する場合にシリンダ x からシリンダ t に移動する平均シーク長を求める。

各アクセスシリンダ群の幅 u は一定で $s+1$ 個存在する。隙間の幅 v も一定で s 個存在する。全シリンダ幅 n は図4と同じく(2)式の関係がある。p_i を各群のアクセス頻度確率とする。

この場合の平均シーク長は

$$\begin{aligned}L=&\sum_{y=1}^{s}\sum_{z=0}^{y-1}\sum_{x=y(u+v)+1}^{y(u+v)+u}\sum_{t=z(u+v)+1}^{z(u+v)+u}p_{y+1}p_{z+1}(x-t)\\&+\sum_{y=0}^{s-1}\sum_{z=y+1}^{s}\sum_{x=y(u+v)+1}^{y(u+v)+u}\sum_{t=z(u+v)+1}^{z(u+v)+u}p_{y+1}p_{z+1}(t-x)\\&+\sum_{y=0}^{s}\sum_{x=1}^{u}\sum_{t=1}^{x}p_{y+1}{}^2(x-t)\\&+\sum_{y=0}^{s}\sum_{x=1}^{u}\sum_{t=x+1}^{u}p_{y+1}{}^2(t-x)\end{aligned} \quad (5)$$

で求まる。結果は

$$\begin{aligned}L=&n/3+(v/3)\times(1-1/uvs)/(1+1/s)\\&+(\alpha/192)(u+v)\times(4s^2+11s-12)\\&-(\alpha^2/960)(u+v)\times(s-1)(s+3)(s^2+2s+5)/(s+1)\\&+(\alpha^2/144)(u+1)(u-1)\times(s-1)(s+3)/\{u(s+1)\}\end{aligned}$$

ただし、$-4/(s-1)\leqq\alpha\leqq4/(s-1)$ (6)

である。s が十分大きい場合は近似的に、$\alpha=4/(s-1)$ の時に最大値

$$L_{MAX}\approx(2/5)(u+v)(s+1.2) \quad (7)$$

となり、$\alpha=-4/(s-1)$ の時に最小値

$$L_{MIN}\approx(7/30)(u+v)(s+0.1) \quad (8)$$

となる。(8)式で、シリンダを圧縮し $v=0$ とすると

$$L_{MIN}\approx(7/30)u(s+0.1) \quad (9)$$

である。(9)式と(7)式から平均シーク長の短縮比は

$$L_{MIN}/L_{MAX}\approx(7/12)\times\{u/(u+v)\} \quad (10)$$

になる。第1項が傾斜効果、第2項が前述の圧縮効果である。しかし、DACYDの場合は頻度分布が片流れなので平均シーク長は、(9)式の代わりに

$$L_{DACYD}\approx(4/15)u(s+2) \quad (11)$$

となり、(10)式の代わりに

$$L_{DACYD}/L_{MAX}\approx(2/3)\times\{u/(u+v)\} \quad (12)$$

が得られる。

なお、図5の分布で各シリンダのアクセス頻度がランダムに入れかわって、平均シーク長が(7)式と(8)式の平均程度と見なせる様な一般的な場合は

$$L_{MEAN}\approx(19/60)(u+v)(s+0.8) \quad (13)$$

となり

$$L_{MIN}/L_{MEAN}\approx(14/19)\times\{u/(u+v)\} \quad (14)$$

$$L_{DACYD}/L_{MEAN}\approx(16/19)\times\{u/(u+v)\} \quad (15)$$

が得られる。$u=v$ の場合について各ケースの平均シーク長の短縮比を試算すると下表になる。

	傾斜効果	圧縮効果	相乗効果
L_{MIN}/L_{MAX}	0.58	0.50	0.29
L_{DACYD}/L_{MAX}	0.67	0.50	0.33
L_{MIN}/L_{MEAN}	0.74	0.50	0.37
L_{DACYD}/L_{MEAN}	0.84	0.50	0.42

4.おわりに

アクセス頻度分布が一次式で変化する場合に、圧縮効果と傾斜効果が相乗的に効くことがわかった。従って、一般的なディスク使用環境ではDACYDにより平均シーク長はほぼ半減すると見込める。

なお、変化が一次式でない場合や、アクセスされるシリンダ位置が等間隔でない場合も、ここで検討した方法を拡張して考えることができる。

(参考文献)

[1]岡本:「動的配列シリンダディスクの構成法」,コンピュータアーキテクチャシンポジウム論文集,pp183-189,May,1988.