

Stanisław WOŁEK

Instytut Informatyki Czasu Rzeczywistego
Politechniki Śląskiej

FIZYCZNA ORGANIZACJA DANYCH W PAMIĘCI KASETOWEJ

Streszczenie. W artykule rozwiązano zagadnienie doboru optymalnych parametrów organizacji danych w pamięci kasetowej, takich jak długość bloku oraz liczba bloków w pliku. Zakłada się dwupoziomą, sekwencyjną organizację danych oraz metodę dostępu z przydzielaniem kolejnych miejsc pamięci oraz liniowym sekwencyjnym wyszukiwaniem plików. Kryterium doboru parametrów jest czas odszukania i odczytu ostatniego zapisanego na taśmie bloku danych.

1. WPROWADZENIE

Fizyczna organizacja danych w pamięci kasetowej polega na grupowaniu danych w bloki informacyjne, a tych z kolei w pliki oraz stosowaniu dodatkowych bloków organizacyjnych, ułatwiających wyszukiwanie żadanego pliku oraz bloku w pliku. Bloki informacyjne oraz organizacyjne stanowią taką część zapisu na taśmie, która jest odczytywana i zapisywana w całości.

Istotną własnością pamięci kasetowej jest istnienie dwóch prędkości ruchu taśmy. Odczytywanie i zapisywanie bloków odbywa się z prędkością podstawową, a ponadto z prędkością podwyższoną (dwudziestokrotnie) można przewijać taśmę o odcinki zawierające zadaną liczbę plików.

Poszczególne bloki oddzielone są przerwami międzyblokowymi, będącymi wykasowanymi odcinkami taśmy. Przerwy te umożliwiają zatrzymanie ruchu taśmy między odczytami kolejnych bloków. Poszczególne pliki rozdzielane są natomiast specjalnymi mini-blokami o długości czterech bajtów, zwanymi znacznikami taśmy. Znaczniki te otoczone są długimi przerwami, które umożliwiają zatrzymanie taśmy po jej przewinięciu poza dany plik.

Parametrami fizycznej organizacji danych jest wielkość pojedynczego bloku oraz wielkość pojedynczego pliku. Wielkość pliku może być wyrażana długością lub liczbą zapisanych w nim bajtów informacji, natomiast wielkość pliku liczbą bloków znajdujących się w nim.

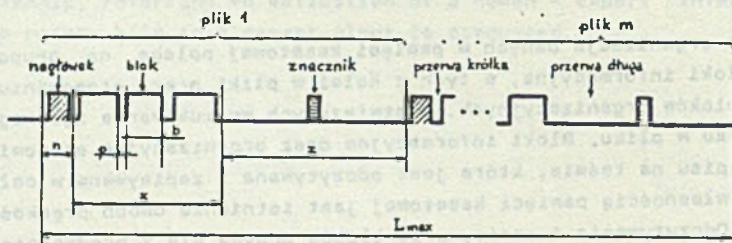
Zagadnienie doboru wielkości bloku można rozpatrywać w dwóch aspektach: stałej ich długości albo długości zmiennej (na przykład innej dla bloków organizacyjnych i innej dla informacyjnych). Doboru wielkości bloków można dokonywać niezależnie od zastosowania pamięci. Wielkość ta jest ustalana najczęściej przy realizacji oprogramowania podstawowego pamięci kasetowej.

Zagadnienie doboru wielkości pliku może być rozpatrywane tylko wtedy, gdy istnieje możliwość dowolnego podziału na pliki danych zapamiętywanych na taśmie. Pliki o dowolnej wielkości mogą być tworzone na przykład dla danych, która powstają na skutek okresowego rejestrowania pomiarów z procesu. Wyniki jednego pomiaru stanowią wtedy rekord danych, który jest ponadto opatrzony kluczem charakteryzującym go (np. numer pomiaru lub czas wykonania pomiaru).

2. SFORMUŁOWANIE PROBLEMU

Rozpatrywane jest zagadnienie doboru parametrów fizycznej organizacji danych w pamięci kasetowej przy następujących założeniach:

- 1) dane mogą być dowolnie dzielone na bloki i pliki,
- 2) bloki mogą posiadać różne długości, ustalone dla bloków organizacyjnych oraz dowolne dla informacyjnych,
- 3) stosowana jest dwupoziomowa organizacja dostępu do danych, z przydzielaniem kolejnych miejsc pamięci oraz sekwencyjnym wyszukiwaniem plików.



Rys. 1. Sekwencyjna organizacja fizyczna danych w pamięci kasetowej

Struktura fizycznej organizacji danych przedstawiona jest na rys. 1. W strukturze tej część organizacyjna pliku składa się ze znacznika taśmy oraz bloku nagłówka pliku, który zawiera informacje organizacyjne o pliku, np. jego nazwę, kolejny numer lub przedział wartości kluczy dla rekordów zapisanych w pliku. Długość bloku nagłówka "n" jest stała i zależy od ilości zawartej w nim informacji. Część informacyjna pliku o długości oznaczonej "x" składa się z "k" bloków o długości "b" oraz przerw o długości "p".

Dwupoziomowa organizacja dostępu do danych polega na odszukaniu pliku, w którym znajduje się żądany rekord danych (poziom 1), a następnie odszukaniu w danym pliku bloku, który zawiera żądany rekord (poziom 2). Zakłada się przy tym, że każdy rekord danych jest scharakteryzowany pewnym kluczem, którego wartość jest podstawą jego odszukania. W nagłówku każdego pliku podany jest natomiast przedział wartości kluczy dla danych zgromadzonych w pliku.

Algorytm wyszukania określonego kluczem rekordu składa się z następujących czynności (przy założeniu, że taśma magnetyczna jest przewinięta do fizycznego początku):

1) przejście przy ruchu z prędkością podwyższoną do początku najbliższego pliku,

2) odczyt nagłówka pliku przy ruchu z prędkością podstawową,

3) sprawdzenie, czy szukany rekord znajduje się w pliku:

- jeżeli nie, to powrót do czynności 1,

- jeżeli tak, to przejście do czynności 4,

4) odczyt kolejnego bloku danych przy ruchu z prędkością podstawową,

5) sprawdzenie, czy w odczytanym bloku znajduje się szukany rekord:

- jeżeli nie, to powrót do czynności 4,

- jeżeli tak, to koniec algorytmu.

Wprowadzenie kryterium oceny przyjętej fizycznej organizacji danych oraz metody dostępu do nich zostało przedyskutowane w artykule [1]. Przyjmuje się, że kryterium tym jest czas odczytu danej. Zakładając, że od początku taśmy do szukanego rekordu zostało napisanych P bajtów informacji, szukamy takich parametrów fizycznej organizacji danych, które minimalizują czas wyszukania i odczytu szukanego rekordu przy przyjętym algorytmie wyszukiwania.

3. DOBÓR DŁUGOŚCI PLIKU

Sformułowany problem doboru parametrów fizycznej organizacji danych sprowadza się do określenia długości pliku, jeżeli założy się, że bloki informacyjne mają ustaloną długość.

Przyjmijmy następujące oznaczenia parametrów fizycznej organizacji danych (rys. 1):

n - długość części organizacyjnej pliku (nagłówek),

x - długość części informacyjnej pliku,

b - długość bloku w części informacyjnej,

k - liczba bloków informacyjnych w pliku,

p - długość przerwy krótkiej,

z - długość znacznika taśmy z przerwami długimi,

v_1 - prędkość podstawowa,

v_2 - prędkość podwyższona,

P - liczba zapisanych bajtów danych,

r - liczba plików, potrzebnych do zapisania P bajtów danych,

g - gęstość zapisu.

Czas odczucia i odczytu bloku zawierającego P -ty bajt danych wyraża się następująco:

$$T_p(x, r) = \left(\frac{n}{v_1} + \frac{x+z}{v_2} \right) \cdot (r - 1) + \frac{n+x}{v_1}, \quad (1)$$

liczbę zapisanych bajtów danych przedstawia zależność:

$$P = g \cdot k \cdot b \cdot r, \quad (2)$$

a długość części informacyjnej zależy od parametrów pliku w sposób następujący:

$$x = k \cdot (b + p). \quad (3)$$

Podstawiając (3) do (2) oraz (2) do (1) rugujemy zmienną k i r , uzyskując zależność czasu T_p od x :

$$T_p(x) = \left(\frac{n}{v_1} + \frac{x+z}{v_2} \right) \cdot \left(\frac{P}{g \cdot x \cdot b} - 1 \right) + \frac{n+x}{v_1}, \quad (4)$$

którą można sprowadzić do następującej postaci:

$$T_p(x) = \alpha \cdot a \cdot w \cdot \frac{P}{x} + a \cdot x + T_0, \quad (5)$$

gdzie:

$$a = \frac{1}{v_1} - \frac{1}{v_2}, \quad (6)$$

$$\alpha = \frac{b + p}{b}, \quad (7)$$

$$w = \frac{n \cdot v_2 + z \cdot v_1}{g(v_1 - v_2)}, \quad (8)$$

$$T_0 = \frac{P \cdot \alpha}{g \cdot v_2} - \frac{z}{v_2}. \quad (9)$$

Rozwiązaniem zadania doboru długości pliku jest przyjęcie takiej wartości x_{opt} , która minimalizuje czas $T_p(x)$, co może zachodzić wtedy, gdy

$$\frac{d T_p(x)}{dx} = -\frac{\alpha \cdot a \cdot w \cdot P}{x^2} + a = 0 \quad (10)$$

Powyższe równanie jest spełniona dla

$$x_{\text{opt}} = \sqrt{\alpha \cdot w \cdot P} \quad \text{dla} \quad a \neq \phi \quad (11)$$

Można więc stwierdzić, że przy istnieniu dwóch różnych prędkości ruchu taśmy (warunek $a \neq 0$) powinno się dobierać długość części informacyjnej pliku za wzoru (11). Długość ta zależy istotnie od liczby zapisywanych na taśmie bajtów danych P . Współczynniki α i w są dla danej pamięci stałe, pierwszy z nich (7) uwzględnia "straty" na przerwy krótkie, a drugi (8) zależy od stałych parametrów fizycznej organizacji danych oraz parametrów technicznych pamięci.

4. DOBÓR WIELKOŚCI BLOKU I PLIKU

Jeżeli założy się, że bloki informacyjne mogą mieć dowolną wielkość, to zagadnienie doboru parametrów fizycznej organizacji danych polega na równoczesnym określeniu zarówno długości pliku, jak i wielkości bloku informacyjnego.

Przyjmując oznaczenie parametrów fizycznej organizacji danych, jak w punkcie 3 (rys. 1), czas odczytu bloku zawierającego P -ty bajt danych można wyrazić następująco:

$$T_p(k, b, r) = \left(\frac{n}{v_1} + \frac{k \cdot (b+p) + z}{v_2} \right) \cdot (r - 1) + \frac{n+k \cdot (b+p)}{v_1} \quad (12)$$

Podstawiając do (12) wyrażenie (2), rugujemy zmienną r , uzyskując zależność czasu T_p od k i b :

$$T_p(k, b) = \left(\frac{n}{v_1} + \frac{k \cdot (b+p) + z}{v_2} \right) \cdot \left(\frac{P}{g \cdot k \cdot b} - 1 \right) + \frac{n+k \cdot (b+p)}{v_1} \quad (13)$$

którę można sprowadzić do następującej postaci:

$$T_p(k, b) = a \cdot w \cdot \frac{P}{k \cdot b} + \frac{P \cdot D}{g \cdot v_2} \cdot \frac{1}{b} + a \cdot k \cdot b + a \cdot p \cdot k + T_1 \quad (14)$$

gdzie a i w określone są wyrażeniami (6), (8) oraz

$$T_1 = \frac{P}{g \cdot v_2} - \frac{z}{v_2} \quad (15)$$

Warunkiem koniecznym istnienia minimum funkcji $T_p(k, b)$ jest, aby gradient tej funkcji był równy zeru, czyli był spełniony układ równań:

$$\frac{\partial T_p(k, b)}{\partial k} = 0 \quad \text{i} \quad \frac{\partial T_p(k, b)}{\partial b} = 0. \quad (16)$$

Po wykonaniu odpowiedniego różniczkowania funkcji (14), uzyskuje się układ:

$$\left. \begin{aligned} \frac{a \cdot w \cdot P}{b} \cdot \frac{-1}{k^2} + a \cdot b + a \cdot p &= 0 \\ \frac{a \cdot w \cdot P}{k} \cdot \frac{-1}{b^2} + \frac{P \cdot p}{g \cdot v_2} \cdot \frac{-1}{b^2} + a \cdot k &= 0 \end{aligned} \right\} \quad (17)$$

Pierwsze z powyższych równań mnożymy obustronnie przez $b \cdot k^2$, a drugie przez $k \cdot b^2$ uzyskując nową postać układu równań:

$$\left. \begin{aligned} -a \cdot w \cdot P + a \cdot b^2 \cdot k^2 + a \cdot p \cdot b \cdot k^2 &= 0 \\ -a \cdot w \cdot P + a \cdot b^2 \cdot k^2 - \frac{P \cdot p}{g \cdot v_2} \cdot k &= 0 \end{aligned} \right\} \quad (18)$$

Aby powyższy układ miał rozwiązanie, powinno być

$$a \cdot p \cdot b \cdot k^2 = -\frac{P \cdot p}{g \cdot v_2} \cdot k,$$

czyli

$$P = -a \cdot k \cdot b \cdot g \cdot v_2,$$

co nie jest nigdy spełnione, ponieważ wszystkie występujące wielkości mają wartości dodatnie. Układ równań (16) jest więc sprzeczny, z czego wynika, że nie istnieje minimalna wartość funkcji (14).

Aby rozwiązać zadanie doboru wielkości bloku i pliku, należy wprowadzić ograniczenia na te wielkości, czyli na zmienne niezależne funkcji (14). Z charakteru tych wielkości wynika, że ograniczenia mają postać:

$$k \geq 1, \quad b_1 \leq b \leq b_2, \quad (19)$$

gdzie b_1 i b_2 oznaczają minimalną i maksymalną wartość, którą może przyjmować długość bloku.

Znalezienie minimum funkcji (14) przy ograniczeniach (19) jest zadaniem programowania nieliniowego z ograniczeniami, dla którego tworzy się funkcję Lagrange'a:

$$L(x, \lambda) = T_p(k, b) + \lambda_k(1-k) + \lambda_1(b_1-b) + \lambda_2(b-b_2), \quad (20)$$

gdzie

$$x = [k, b]^T, \quad \lambda = [\lambda_k, \lambda_1, \lambda_2]^T.$$

Ponieważ funkcje ograniczeń (19) są liniowe, to na podstawie twierdzenia Karlina [2] spełnione są warunki regularności tych ograniczeń. W konsekwencji tego warunkiem koniecznym i wystarczającym istnienia minimum funkcji (14) przy ograniczeniach (19) jest, aby były spełnione następujące warunki Kuhna-Tuckera [2]:

$$\nabla_x L(\hat{x}, \hat{\lambda}) = 0,$$

$$\langle \hat{\lambda}, \nabla_{\lambda} L(\hat{x}, \hat{\lambda}) \rangle = 0, \quad (21)$$

$$\nabla_{\lambda} L(\hat{x}, \hat{\lambda}) \leq 0,$$

$$\hat{\lambda} \geq 0$$

gdzie ∇_x oznacza gradient funkcji Lagrange'a względem x , ∇_{λ} - gradient względem λ , $\langle a, b \rangle$ - iloczyn skalarny wektorów a i b , zaś \hat{x} i $\hat{\lambda}$ są rozwiązaniami zadania.

Stosując warunki (21) do funkcji (20), uzyskuje się następujący układ 11 równań i nierówności:

$$\frac{-a \cdot w \cdot P}{b \cdot k^2} + a(b+p) - \lambda_k = 0, \quad (22)$$

$$\frac{-a \cdot w \cdot P}{k \cdot b^2} - \frac{P \cdot p}{g \cdot v_2 \cdot b^2} + a \cdot k - \lambda_1 + \lambda_2 = 0, \quad (23)$$

$$\lambda_k(1-k) = 0, \quad (24)$$

$$\lambda_1(b_1-b) = 0, \quad (25)$$

$$\lambda_2(b-b_2) = 0, \quad (26)$$

$$\lambda_1 - k \leq 0, \quad (27)$$

$$b_1 - b \leq 0, \quad (28)$$

$$b - b_2 \leq 0, \quad (29)$$

$$\lambda_k \geq 0, \quad (30)$$

$$\lambda_1 \geq 0, \quad (31)$$

$$\lambda_2 \geq 0. \quad (32)$$

Rozwiązanie powyższego układu, czyli określenie wartości \hat{x} i $\hat{\lambda}$, spełniających równania i nierówności (22) ÷ (32), jest rozwiązaniem zadania doboru wielkości bloku i pliku.

Układ (22) ÷ (32) można rozwiązać, wychodząc od równań (24) ÷ (26) i rozpatrując niezależnie 8 kombinacji wartości x i λ , które spełniają te równania.

A) $\lambda_k = \lambda_1 = \lambda_2 = 0$, dowolne - k , b :

Równania (22) i (23) przyjmują postać (17) i jak pokazano, nie mają rozwiązania.

B) $\lambda_k = \lambda_1 = 0$, $b = b_2$, dowolne - k , λ_2 :

Z równania (22) otrzymuje się:

$$k = \sqrt{\frac{w \cdot P}{b_2(b_1 + P)}}, \quad (33)$$

natomiast z nierówności (27) wynika, że musi być spełnione

$$P \geq \frac{b_2(b_2 + P)}{w}. \quad (34)$$

Z równania (23)

$$\lambda_2 = \frac{a \cdot w \cdot P}{k \cdot b_2^2} + \frac{P \cdot p}{g \cdot v_2 \cdot b_2^2} - a \cdot k,$$

a mnożąc powyższe równanie obustronnie przez $k \cdot b_2^2$ i uwzględniając (33), uzyskuje się:

$$\lambda_2 \cdot k \cdot b_2^2 = a \cdot w \cdot P + \frac{P \cdot p \cdot k}{g \cdot v_2} - a \cdot b_2^2 \frac{w \cdot P}{b_2(b_2 + P)} = a \cdot w \cdot P \left(1 - \frac{b_2}{b_2 + P}\right) + \frac{P \cdot p \cdot k}{g \cdot v_2}.$$

Prawa strona powyższego równania ma zawsze wartość dodatnią, zatem spełniona jest nierówność (32), czyli $b = b_2$ oraz k ze wzoru (33) przy warunku (34) są rozwiązaniem zadania.

C) $\lambda_k = \lambda_2 = 0$, $b = b_1$, dowolne - k , λ_1 :

Z równania (22) otrzymuje się:

$$k = \sqrt{\frac{w \cdot P}{b_1(b_1 + p)}} \quad (35)$$

natomiast z (23):

$$\begin{aligned} \lambda_1 \cdot k \cdot b_1^2 &= a \cdot b_1^2 \cdot k^2 - a \cdot w \cdot P - \frac{P \cdot p \cdot k}{g \cdot v_2} = \\ &= a \cdot w \cdot P \left(\frac{b_1}{b_1 + p} - 1 \right) - \frac{P \cdot p \cdot k}{g \cdot v_2}. \end{aligned}$$

Prawa strona powyższego równania ma zawsze wartość ujemną, zatem nie jest spełniona nierówność (31), czyli $b = b_1$ oraz k z (35) nie spełniają układu (22) - (32).

D) $\lambda_k = 0$, $b = b_1$, $b = b_2$, dowolne - k , λ_1 , λ_2 :

Brak rozwiązania, ponieważ nie może być $b = b_1 = b_2$.

E) $k = 1$, $\lambda_1 = 0$, $b = b_2$, dowolne - λ_k , λ_2 :

Z równania (22) i nierówności (30) uzyskuje się warunek:

$$P \leq \frac{b_2(b_2 + p)}{w} \quad (36)$$

natomiast z (23) i (32)

$$P \geq \frac{b_2^2}{w_1} \quad (37)$$

gdzie

$$w_1 = w + \frac{p}{a \cdot g \cdot v_2} \quad (38)$$

Aby sprawdzić, czy warunki (36) i (37) mogą być spełnione równocześnie, dokonamy przekształceń dla różnicy

$$\Delta_1 = \frac{b_2(b_2 + p)}{w} - \frac{b_2^2}{w_1} \quad (39)$$

podstawiając z (38) za w_1 :

$$\Delta_1 = \frac{b_2(b_2 + p)}{w} - \frac{b_2^2 \cdot a \cdot g \cdot v_2}{p + a \cdot g \cdot v_2 \cdot w}$$

$$\Delta_1 \cdot \frac{w(p+a.g.v_2.w)}{b_2} = (b_2+p) \cdot (p+a.g.v_2.w) - b_2 \cdot a.g.v_2.w =$$

$$= p(b_2+p+a.g.v_2.w).$$

Ponieważ wszystkie wyrażenia w powyższej równości mają wartości dodatnie, Δ_1 także przyjmuje wartość dodatnią. Wartości $k = 1$ i $b = b_2$ stanowią więc rozwiązanie zadania przy warunku

$$\frac{b_2^2}{w_1} \leq p \leq \frac{b_2(b_2 + p)}{w}, \quad (40)$$

F) $k = 1$, $b = b_1$, $\lambda_2 = 0$, dowolne $-\lambda_k, \lambda_1$:

Z równania (22) i nierówności (30) uzyskuje się warunek:

$$p \leq \frac{b_1(b_1 + p)}{w}, \quad (41)$$

natomiast z równania (23) i (31)

$$p \leq \frac{b_1^2}{w_1}. \quad (42)$$

Aby określić wspólny warunek, można wprowadzić różnicę

$$\Delta_2 = \frac{b_1(b_1 + p)}{w} - \frac{b_1^2}{w_1},$$

która ma postać podobną do Δ_1 z (39), posiada zatem także wartość dodatnią, z czego wynika, że przy spełnionym warunku (42) spełniony jest zawsze (41). Zatem wartości $k = 1$ i $b = b_1$ stanowią rozwiązanie zadania przy warunku (42).

G) $k = 1$, $\lambda_1 = \lambda_2 = 0$, dowolne $-\lambda_k, b$:

Z równania (23) otrzymuje się:

$$b = \sqrt{w_1 \cdot p}. \quad (43)$$

natomiast z nierówności (28) i (29) wynika, że musi być spełnione

$$\frac{b_1^2}{w_1} \leq p \leq \frac{b_2^2}{w_1}. \quad (44)$$

Z równania (22)

$$\lambda_k = a(b + p) - \frac{a \cdot w \cdot P}{b},$$

a mnożąc powyższe równanie dwustronnie przez $\frac{b}{a}$ i uwzględniając (43) oraz (38) uzyskuje się:

$$\lambda_k \cdot \frac{b}{a} = b^2 + b \cdot p - w \cdot P =$$

$$= \left(w + \frac{p}{a \cdot g \cdot v_2} \right) \cdot P + p \cdot \sqrt{w_1 \cdot P} - w \cdot P = \frac{p \cdot P}{a \cdot g \cdot v_2} + p \cdot \sqrt{w_1 \cdot P}.$$

Prawa strona powyższego równania ma wartość dodatnią, zatem spełniona jest nierówność (30), czyli $k = 1$ oraz $b = \sqrt{w_1 \cdot P}$ przy warunku (44) są rozwiązaniem zadania.

H) $k = 1$, $b = b_1$, $b = b_2$ - brak rozwiązania, podobnie jak w D).

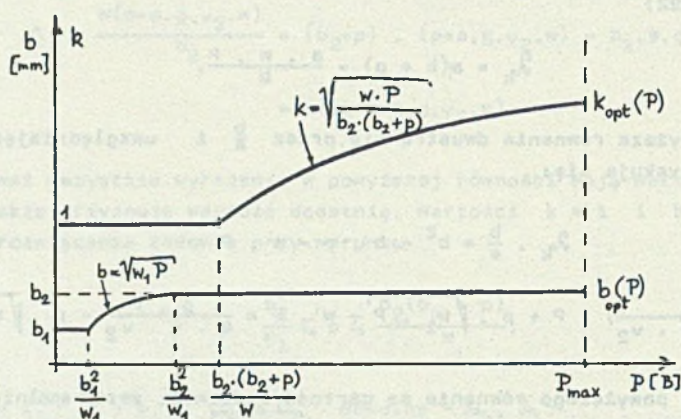
Uzyskane rozwiązania zadania doboru wielkości pliku i bloku można zebrać w następującej tabeli:

Tabela A

| Warunek | Rozwiązanie | Przypadek |
|---|--|-----------|
| $P < \frac{b_1^2}{w_1}$ | $k = 1, \quad b = b_1$ | F |
| $\frac{b_1^2}{w_1} < P < \frac{b_2^2}{w_1}$ | $k = 1, \quad b = \sqrt{w_1 \cdot P}$ | G |
| $\frac{b_2^2}{w_1} \leq P < \frac{b_2(b_2 + p)}{w}$ | $k = 1, \quad b = b_2$ | E |
| $P > \frac{b_2(b_2 + p)}{w}$ | $k = \sqrt{\frac{w \cdot P}{b_2(b_2 + p)}}, \quad b = b_2$ | B |

(45)

Z tabeli wynika, że wartości optymalne długości bloku oraz liczby bloków w pliku zależą od liczby zapisywanych na taśmie bajtów danych (rys. 2). Przy zapamiętywaniu odpowiednio małej ilości informacji należy stosować pliki jednoblokowe, dobierając odpowiednio długość bloku z przedziału ograniczeń od b_1 do b_2 . Natomiast dla większej ilości informacji należy przyjąć maksymalną dopuszczalną długość bloku, a liczbę bloków w pliku wyznaczyć ze wzoru (33):



Rys. 2. Zależność wartości optymalnych k i b od liczby P bajtów danych zapisywanych na taśmie

$$k_{\text{opt}}(P) = \sqrt{\frac{w \cdot P}{b_2 \cdot (b_2 + p)}} \quad (46)$$

Podstawiając wyrażenie (46) do wzoru (14) na czas odczytania i odczytu bloku zawierającego P -ty bajt uzyskuje się:

$$T_{P,\text{min}}(b_2) = T_P(k_{\text{opt}}, b_2) = 2 \cdot a \sqrt{w \cdot P \left(\frac{P + b_2}{b_2} \right)} + \frac{P}{g \cdot v_2} \cdot \frac{P + b_2}{b_2} - \frac{z}{v_2} \quad (47)$$

Ponieważ w powyższym wyrażeniu b_2 występuje dwukrotnie w czynniku $\frac{P + b_2}{b_2}$, aby osiągnąć minimalną wartość T , powinno się przyjmować jak największą dopuszczalną długość bloku.

Na ograniczenie długości bloku może wpływać kilka czynników. Programy korzystające z pamięci kasetowej muszą posiadać w pamięci operacyjnej bufor, do którego wczytuje się treść bloku. Dostępność pamięci operacyjnej narzuca więc ograniczenie na wielkość takiego bufora. Podobne bufory mogą posiadać sterowniki łączące sprzętowo pamięć kasetową z komputerem, wielkość bloku rzutuje wtedy na wielkość i koszt sterownika. Informacje zgromadzone w bloku są najczęściej podzielone na mniejsze jednostki logiczne zwane rekordami. Odszukanie żądanego rekordu w zbyt długim bloku może być czasochłonne. W przeprowadzonej dotychczas analizie założono, że operacje odczytu przebiegają zawsze poprawnie. W praktyce mogą występować przekłamanie, stwarzające konieczność powtórnego lub nawet wielokrotnego odczytywania tego samego bloku. Przy zbyt długich blokach ich wielokrotne odczytywanie może w sposób istotny zwiększyć wartość przyjętego kryterium oceny organizacji danych.

Należy dodać, że z wyrażenia (46) uzyskuje się wartości liczby bloków w pliku będące liczbami rzeczywistymi, natomiast liczba bloków w pliku musi mieć wartość naturalną. Istnieje więc konieczność zaokrąglenia uzyskanej z (46) wartości do liczby naturalnej lub wybrania jednej z dwóch sąsiednich liczb naturalnych na podstawie uzyskiwanych dla nich wartości kryterium danego wzorem (14).

Wzór (46) odpowiada wynikowi uzyskanemu w punkcie 3. Uwzględniając go w wyrażeniu (3) uzyskuje się bowiem

$$x = k(b_2 + p) = \sqrt{\frac{(b_2 + p) \cdot w \cdot P}{b_2}} = \sqrt{\alpha \cdot w \cdot P},$$

czyli wyrażenie (11).

5. PRZYKŁAD PRAKTYCZNY DLA PAMIĘCI KASETOWEJ PK-1

Pamięć kasetowa PK-1 posiada następujące parametry techniczne:

- prędkość ruchu taśmy $v_1 = 127$ [mm/s], $v_2 = 1500$ [mm/s],
- gęstość zapisu $g = 4$ [B/mm].

Przyjęto następujące wartości stałych parametrów fizycznej organizacji danych:

- długość przerwy krótkiej $p = 20$ [mm],
- długość nagłówka pliku $n = 24$ [mm],
- długość znacznika taśmy $z = 181$ [mm].

Dla powyższych danych współczynniki występujące we wzorach (45) posiadają następujące wartości:

$$w = 10.74 \text{ [mm}^2/\text{B]},$$

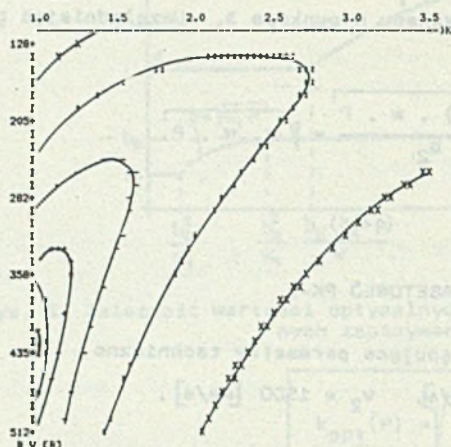
$$w_1 = 11.13 \text{ [mm}^2/\text{B]}.$$

W tabeli B przedstawione są wartości liczby P zapamiętywanych bajtów danych dla charakterystycznych punktów z rysunku 2, dla dwóch wartości ograniczeń b_2 na długość bloku. Obok liczby bajtów danych podana jest długość L taśmy potrzebna do ich zapisania.

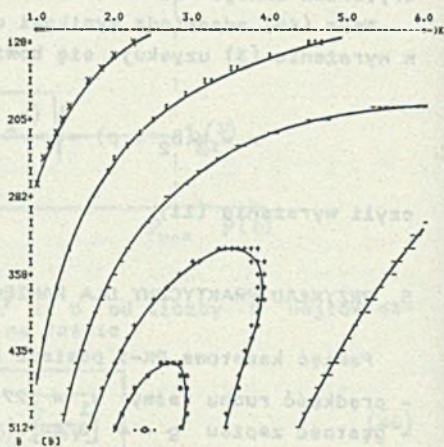
Tabela B

| b_1 [B] | b_2 [B] | | $\frac{b_1^2}{w_1}$ | $\frac{b_2^2}{w_1}$ | $\frac{b_2(b_2+p)}{w}$ | P_{max} | $k_{opt}(P_{max})$ |
|-----------|-----------|--------|---------------------|---------------------|------------------------|-----------|--------------------|
| 128 | 512 | P [kB] | 0.09 | 1.47 | 1.59 | 280 | 13 |
| | | L [m] | 0.18 | 1.01 | 1.09 | 89.84 | |
| | 1024 | P [kB] | 0.09 | 5.89 | 6.22 | 280 | 7 |
| | | L [m] | 0.18 | 2.77 | 2.92 | 84.06 | |

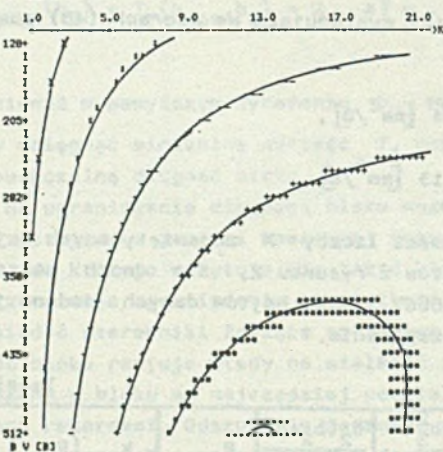
| ZNAK | T | PGJ | DLUGOSC | KOPT | BOPT | TRIN |
|------|-----|-------|---------|------|--------|--------|
| . | 1.7 | 11KB3 | 0.78CM3 | 1.0 | 422CB3 | 1.7CS3 |
| * | 1.7 | | | | | |
| + | 1.8 | | | | | |
| - | 1.8 | | | | | |
| | 2.0 | | | | | |
| X | 2.5 | | | | | |



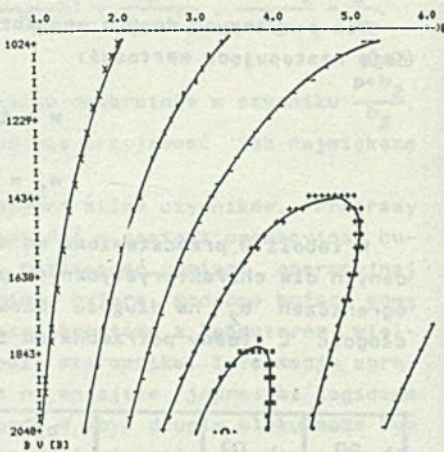
| ZNAK | T | PGJ | DL. GOSC | KOPT | BOPT | TRIN |
|------|------|-------|----------|------|--------|--------|
| . | 4.9 | 10CB3 | 4.55CM3 | 2.4 | 512CB3 | 4.9ES3 |
| * | 7.0 | | | | | |
| + | 7.2 | | | | | |
| - | 7.8 | | | | | |
| | 9.6 | | | | | |
| X | 14.3 | | | | | |



| ZNAK | T | PGJ | DLUGOSC | KOPT | BOPT | TRIN |
|------|-------|---------|----------|------|--------|---------|
| . | 80.7 | 280CKB3 | 89.79CM3 | 12.7 | 512CB3 | 80.7CS3 |
| * | 83.4 | | | | | |
| + | 91.4 | | | | | |
| - | 112.4 | | | | | |
| | 170.3 | | | | | |
| X | 327.1 | | | | | |



| ZNAK | T | PGJ | DLUGOSC | KOPT | BOPT | TRIN |
|------|------|---------|----------|------|---------|---------|
| . | 73.9 | 280CKB3 | 81.14CM3 | 3.3 | 2048CB3 | 73.9ES3 |
| * | 74.3 | | | | | |
| + | 74.9 | | | | | |
| - | 76.8 | | | | | |
| | 82.2 | | | | | |
| X | 96.9 | | | | | |



Rys. 3. Przebiegi stałych wartości kryterium (14) w funkcji zmiennych parametrów organizacji danych

Z powyższych zależności wynika, że przy wykorzystaniu taśmy powyżej 1,2% lub 3,5% jej maksymalnej długości należy stosować największą dopuszczalną długość bloku i dobierać do niej liczbę bloków w pliku ze wzoru (46).

Na rysunku 3 przedstawione są przebiegi stałych wartości czasu odszukania i odczytu bloku zawierającego P -ty bajt danych w funkcji zmiennych k i b , charakteryzujących przyjętą fizyczną organizację danych. Wykresy uzyskane są za pomocą programu przedstawionego w załączniku. Kolejne krzywe odpowiadają wartościom czasu odszukania i odczytu zmieniającym się wykładniczo. Dla wygody w miejsce długości bloku w milimetrach stosowano wielkość bloku wyrażoną liczbę zapisanych w nim bajtów.

Przedstawione przebiegi odpowiadają uzyskanym wcześniej wynikiom teoretycznym. Minimalne wartości czasu T osiągnąć się dla zmiennych k i b przyjmujących wartości graniczne. Dla małej liczby P zapamiętywanych bajtów danych wynoszącej 1 kB (rys. 3a) należy przyjąć długość bloku z przedziału ograniczeń oraz tylko jeden blok w pliku. Dla zwiększonej liczby bajtów P do wartości 10 kB oraz 280 kB (rys. 3b i 3c) należy przyjąć maksymalną dopuszczalną długość bloku oraz odpowiednio rosnącą liczbę bloków w pliku (2 bloki dla 10 kB i 12 dla 280 kB). Wpływ wartości górnego ograniczenia długości bloku na optymalne wartości parametrów organizacji danych prezentują przebiegi 3c i 3d. Przyjęcie większej wartości dopuszczalnej długości bloku powoduje zmniejszenie optymalnej liczby bloków w pliku.

Z charakteru przebiegów wynika, że na uzyskiwaną wartość czasu T większy wpływ posiada parametr k , natomiast mniejszą parametr b . Dwukrotne powiększenie maksymalnej wielkości bloku z 512 [B] na 1024 [B] (przy stosowaniu optymalnej liczby bloków w pliku wynoszącej 13 i 7) powoduje zmianę minimalnego czasu T z wartości 80,7 [a] na 76,2 [a], czyli o 5,6%. Dwukrotne powiększenie liczby bloków w pliku dla wielkości bloku 512 [B] z liczby 13 na 26 powoduje zmianę czasu T z 80,7 [a] na 88,1 [a], czyli o 9,2%, a dwukrotne zmniejszenie z 13 na 6 zmianę T z 80,7 [a] na 88,5 [a], czyli o 9,7%.

6. PARAMETRY ORGANIZACJI DANYCH PRZY PEŁNYM WYKORZYSTANIU TAŚMY

Przyjmijmy, że dane jest górne ograniczenie b_2 długości bloku oraz że liczbę bloków w pliku wyznacza się ze wzoru (46), w którym uwzględnia się taką liczbę bajtów danych, którą można zapisać na całej długości L_{\max} taśmy. Długość L_{\max} zależy następująco od parametrów organizacji danych

$$L_{\max} = r(n + z + k(b_2 + p)). \quad (48)$$

Wyznaczając z powyższego równania r i podstawiając to do (2), uzyskuje się równanie

$$P = g \cdot k \cdot b_2 \frac{L_{\max}}{n + z + k(b_2 + p)} \quad (49)$$

Jednocześnie przekształcając wyrażenie (46), uzyskuje się

$$P = \frac{k^2 \cdot b_2(b_2 + p)}{w} \quad (50)$$

Porównanie prawych stron powyższych równań prowadzi do równania kwadratowego

$$(b_2 + p) \cdot k^2 + (n + z) \cdot k - \frac{w \cdot g \cdot L_{\max}}{b_2 + p} = 0 \quad (51)$$

Wprowadźmy oznaczenia

$$c = n + z, \quad d = \sqrt{4 \cdot w \cdot g \cdot L_{\max} + c^2} \quad (52)$$

Równania (51) posiada rozwiązanie

$$k_{\text{opt}} = \frac{d - c}{2(b_2 + p)} \quad (53)$$

Z wyrażen (48) i (53) uzyskuje się zależność na liczbę plików n na taśmie

$$n_{\text{opt}} = \frac{2 \cdot L_{\max}}{d + c} \quad (54)$$

oraz na długość pliku

$$h_{\text{opt}} = \frac{d + c}{2} \quad (55)$$

Podstawiając (53) do (49) uzyskuje się wzór określający uzyskaną pojemność taśmy

$$P_{\max} = g \cdot L_{\max} \cdot \frac{b_2}{b_2 + p} \cdot \frac{d - c}{d + c} \quad (56)$$

Następująco można przedstawić wyrażenie określające minimalny czas odzyskania i odczytu ostatniego bloku danych, zapisanych na taśmie o długości L_{\max}

$$T_{\min} = \frac{4 \cdot a \cdot w \cdot g \cdot L_{\max}}{d + c} + \frac{L_{\max}}{v_2} \cdot \frac{d - c}{d + c} - \frac{z}{v_2} \quad (57)$$

Istotne jest, że we wzorach (54), (55) i (57) występują tylko stałe parametry fizycznej organizacji danych, a liczba plików na taśmie, długość pliku i uzyskana wartość czasu T_{\min} nie zależą od parametrów zmiennych, czyli długości bloku i liczby bloków w pliku. Długość bloku b_2 wpływa na uzyskiwaną maksymalną pojemność taśmy (56), natomiast dla przyjętej długości należy przyjąć liczbę bloków w pliku ze wzoru (53), uzyskując w efekcie stałą długość pliku (54) oraz stałą liczbę plików na taśmie (53).

Na podstawie wyprowadzonych wzorów można przedstawić następujące rozwiązanie sformułowanego w punkcie 2 problemu:

Twierdzenie

Minimalny czas odszukania i odczytu ostatniego na taśmie bloku danych nie zależy od długości bloku i liczby bloków w pliku, a uzyskuje się go przy podziale taśmy na pliki o długości podanej wzorem (55), w liczbie określonej wzorem (54). Maksymalną pojemność taśmy uzyskuje się dla największej dopuszczalnej długości bloku, dla której należy przyjąć podaną w (53) liczbę bloków w pliku.

W pamięci kasetowej PK-1 stosuje się taśmę o długości $L_{\max} = 85$ [m]. Poniższa tabela przedstawia teoretyczne wartości parametrów fizycznej organizacji danych uzyskiwane dla różnych wartości dopuszczalnej długości bloku.

Tabela C

| b_2 [B] | k_{opt} | m_{opt} | h_{opt} [mm] | P_{\max} [kB] | T_{\min} [s] |
|-----------|------------------|------------------|-----------------------|-----------------|----------------|
| 128 | 34,8 | 42,2 | 2016 | 188 | 76,9 |
| 256 | 21,6 | 42,2 | 2016 | 233 | 76,9 |
| 512 | 15,6 | 42,2 | 2016 | 253 | 76,9 |
| 1024 | 12,2 | 42,2 | 2016 | 264 | 76,9 |
| 2048 | 10,0 | 42,2 | 2016 | 271 | 76,9 |
| 4096 | 8,6 | 42,2 | 2016 | 277 | 76,9 |

Dla zbyt małej długości bloku należy stosować stosunkowo dużą liczbę bloków w pliku i wtedy dość znacznie traci się na uzyskiwanej pojemności taśmy. Dla większych długości bloków zmiany pojemności są mniejsze.

W tabeli C podane zostały wartości k_{opt} i m_{opt} jako liczby rzeczywiste, natomiast w praktyce muszą być one przybliżone do liczb natural-

nych. Przybliżenie można zrealizować zaczynając od parametru k lub od m , według następujących algorytmów:

a)

$$k = \text{round}(k_{\text{opt}})$$

$$m = \text{int}\left(\frac{L_{\text{max}}}{c + k \cdot (b_2 + p)}\right)$$

b)

$$m = \text{round}(m_{\text{opt}})$$

$$k = \text{int}\left(\frac{L_{\text{max}}}{b_2 + p} - \frac{c}{m}\right),$$

w których zapisie round oznacza funkcję zaokrąglenia, a int funkcję oznaczającą część całkowitą z argumentu.

W tabeli D zebrane są wartości parametrów dla pamięci PK-1 wyznaczone według algorytmu a), a w tabeli E według b).

Tabela D

| b_2 [B] | k | m | h [mm] | P [kB] | T [s] | L [m] |
|-----------|-----|-----|----------|----------|---------|---------|
| 128 | 35 | 41 | 2025 | 184 | 75.4 | 83.0 |
| 256 | 22 | 41 | 2053 | 231 | 76.4 | 84.2 |
| 512 | 16 | 41 | 2061 | 252 | 76.7 | 84.5 |
| 1024 | 12 | 42 | 1981 | 258 | 75.4 | 83.2 |
| 2048 | 10 | 42 | 2005 | 269 | 76.2 | 84.2 |
| 4096 | 9 | 40 | 2113 | 276 | 76.9 | 84.5 |

Tabela E

| b_2 [B] | k | m | h [mm] | P [kB] | T [s] | L [m] |
|-----------|-----|-----|----------|----------|---------|---------|
| 128 | 34 | 42 | 1973 | 183 | 75.1 | 82.9 |
| 256 | 21 | 42 | 1969 | 226 | 75.0 | 82.7 |
| 512 | 15 | 42 | 1945 | 242 | 74.1 | 81.7 |
| 1024 | 12 | 42 | 1981 | 258 | 75.4 | 83.2 |
| 2048 | 10 | 42 | 2005 | 269 | 76.2 | 84.2 |
| 4096 | 8 | 42 | 1901 | 258 | 72.6 | 79.9 |

Po porównaniu tych danych wynika, że wartości bliższe teoretycznym oraz dające lepsze wykorzystanie taśmy uzyskuje się dla algorytmu a). Parametry zebrane w tabeli D stanowią więc praktyczne rozwiązanie zadania doboru parametrów fizycznej organizacji danych w pamięci kasetowej PK-1 przy dwupoziomowej organizacji dostępu do danych i przy założeniach podanych w punkcie 2.

7. ZAŁĄCZNIK

```

PROGRAM WYKRES;
CONST
  V1=127.0; (* [MM/S] *)
  V2=1500.0; (* [MM/S] *)
  N=24.0; (* [MM] *)
  Z=181.0; (* [MM] *)
  P=20.0; (* [MM] *)
  G=4.0; (* [B/MM] *)
  SKB=30; (* WIERSZY *)
  SKK=60; (* KOLUMNY *)
VAR
  I,WSZ,KL: INTEGER;
  K,KO,K1,DK,B,BO,B1,DB,Y,TO,T1,TI,POJ,EPS,E0,DE,KOPT,BOPT,FI: REAL;
  FUN: ARRAY [0..SKK, 0..SKB] OF REAL;
  WIERSZ: ARRAY [0..SKK] OF CHAR;
  ZN: ARRAY[0..5] OF CHAR;
VALUE ZN= ('.', '*', '+', '-', ':', 'X');
FUNCTION F ( K,B: REAL ): REAL;
BEGIN
  F := ( N/V1 + ( Z+K*(D/G+P))/V2 ) * ( POJ/(K*B)-1 ) +
        ( N + K*(B/G+P) ) / V1;
END;
BEGIN
  READLN ( POJ, KO, K1, BO, B1 );
  DK := (K1-KO)/SKK; DB := (B1-BO)/SKB;
  TO := F(KO,BO); T1:=10; KOPT:=KO; BOPT:=BO;
  FOR W SZ:=0 TO SKB DO
    BEGIN
      B:=BO+WSZ*DB;
      FOR KL:=0 TO SKK DO
        BEGIN
          K:=KO+KL*DK;
          T:=F(K,B);
          FUNCL,WSZ:=T;
          IF T<TO THEN BEGIN
            KOPT:=K; BOPT:=B; TO:=T
          END
        ELSE
          IF T>T1 THEN T1:=T;
        END
      END
    END
  END;
  DT := (T1-TO)/(EXP(6)-1);
  E0 := (T1-TO)/2E3;
  DE := DT/40;
  WRITELN (' ZNAK T POJ DŁUGOSC KOPT BOPT TMIN' );
  WRITELN(ZNCO:=4, TO:=7:1, POJ/1000:=6:0, 'EKB');
  POJ*(N+Z+(P+BOPT/G)*KOPT)/KOPT/BOPT/1000:=7:2, 'EMJ',
  KOPT:=5:1, BOPT:=6:0, 'EBJ', TO:=6:1, 'ESJ');
  FOR I:=1 TO 5 DO WRITELN ( ZNII:=4, TO+DT*(EXP(I)-1):7:1 );
  WRITELN;
  WRITE ( KO:=8:1 ); FOR I:=1 TO 5 DO WRITE(KO+I*12*DK:=12:1); WRITELN;
  WRITE (' ':6 ); FOR I:=1 TO 5 DO WRITE('+-');
  WRITELN('+-');
  FOR WSZ:=0 TO SKB DO
    BEGIN
      FOR KL:=0 TO SKK DO WIERSZ[KL]:= ' ';
      FOR I:=5 DOWNTO 0 DO
        BEGIN
          FI := TO + DT*(EXP(I)-1);
          EPS := E0+DE*(EXP(I)-1);
          FOR KL:=0 TO SKK DO
            IF ABS(FUNCL,WSZ)-FI>EPS THEN WIERSZ[KL]:=ZNII;
          END;
        END
      IF WSZ MOD 6 = 0 THEN WRITE ( BO+WSZ*DB:=5:0, '*')
      ELSE WRITE ( 'I':6 );
      FOR KL:=0 TO SKK DO WRITE ( WIERSZ[KL] ); WRITELN;
    END;
  WRITELN('B U [B]':10);
END.

```

LITERATURA

- [1] Wołek S.: Dobór parametrów fizycznej organizacji plików danych w pamięci kasetowej przy zmiennej długości bloków. Podatowy Sterowania Tom 13 (1983), z. 3.
- [2] Findelsen W., Szymanowski J., Wierzbicki A.: Teoria i metody obliczeniowe optymalizacji. PWN, Warszawa 1977.

Recenzent: Doc. dr hab. inż. Adam Wołisz

Wpłynęło do Redakcji: 30.03.1984 r.

ФИЗИЧЕСКАЯ ОРГАНИЗАЦИЯ ДАННЫХ В КАСЕТНОЙ ПАМЯТИ

Р е з ю м е

В работе решена проблема подбора оптимальных параметров организации данных в кассетной памяти таких как: длина блока а также количество блоков в связке. Предполагается двухуровневая последовательная организация данных а также метод доступа к данным приписывающий очередные места памяти с линейным последовательным поиском связок. Критерий оценки подбора параметров - это время поиска и считывания последнего блока, записанного на кассетной ленте.

DATA PHYSICAL ARRANGEMENT IN CASSETTE MEMORY

S u m m a r y

The paper deals with the solution of optimum parameters selection for data arrangement in cassette memory, such as blocks length and a number of blocks in a file. Two level sequential data organisation and data division into files and blocks are considered. The data access method using sequential memory space allocation and linear sequential files searching is also considered. The time of retrieval and reading out the last writing block is the criterion for the selection of parameters.