

**Министерство образования и науки Украины
Севастопольский национальный технический
университет**

**ОПТИМИЗАЦИЯ РАЗМЕРА СЕГМЕНТА
ДЛЯ ДАННЫХ ПЕРЕМЕННОЙ ДЛИНЫ**

Методические указания
к выполнению лабораторной работы
для студентов обучающихся по направлению
6.050101 "Компьютерные науки "
очной и заочной форм обучения

**Севастополь
2014**

Оптимизация размера сегмента для данных переменной длины: методические указания для студентов дневного и заочного отделения по направлению 6.050101 «Компьютерные науки» / Сост. Ю.В.Коваленко, В. Ю. Карлусов – Севастополь: Изд-во СевНТУ, 2014. – 12с.

Целью методических указаний является:

1. углублённое изучение основных теоретических положений дисциплины «Управление информационно-вычислительными ресурсами»;
2. выработка у студентов навыков оптимального расчёта длины сегмента, обеспечивающего минимальный расход памяти вычислительной системы.

Методические указания составлены в соответствии с требованиями программы дисциплины «Управление информационно-вычислительными ресурсами» для студентов направления 6.050101 и утверждены на заседании кафедры информационных систем, протокол № 1 от 29 августа 2014 года.

Допущено учебно-методическим центром СевНТУ в качестве методических указаний.

Рецензент: Кожаев Е.А., кандидат технических наук, доцент кафедры КВТ.

СОДЕРЖАНИЕ

	Стр.
1. Цель работы	4
2. Краткое теоретическое введение	4
3. Порядок работы	9
4. Варианты заданий	10
5. Контрольные вопросы	10
6. Содержание отчёта	10
Список литературы	11

1. Цель работы

Целью настоящей лабораторной работы является:

- углублённое изучение основных теоретических положений дисциплины «Управление информационно-вычислительными ресурсами»;
- выработка у студентов навыков оптимального расчёта длины сегмента, обеспечивающего минимальный расход памяти вычислительной системы.

2. Краткое теоретическое введение

В системах хранения и обработки данных часто возникает следующая задача.

Пусть имеется большая совокупность информационных объектов, для которых известна функция $\rho(l)$ плотности распределения вероятности того, что сообщение имеет длину l знаков. В вычислительной системе данные предполагается размещать в одном или нескольких сегментах постоянной длины $x + s$ знаков, из которых x знаков может быть использовано для хранения собственно данных, а s знаков служит для организации ссылок между сегментами, хранящими один информационный объект.

Требуется выбрать такой размер сегмента x , чтобы математическое ожидание расхода памяти для хранения всей совокупности сообщений было минимальным.

Интерпретация понятий, фигурирующих в словесной постановке задачи, в терминах виртуальной памяти тривиальна. Длина l информационного сообщения представляет собой заявку прикладной программы на выделение ей объёма оперативной памяти, адресованную супервизору. Понятие сегмента и страницы виртуального накопителя тождественны, причём для удовлетворения любой заявки выделяется целое количество неделимых страниц.

Дадим точную формулировку обозначенной задачи.

Легло видеть, что для всех сообщений, у которых

$$kx \leq l < (k + 1)x \tag{1}$$

потребуется $k + 1$ сегментов длиной $(x + s)$ знаков каждый. Количество сообщений, удовлетворяющих (1), равно:

$$\int_{kx}^{(k+1)x} \rho(l) dl \quad (2)$$

Таким образом, математическое ожидание объёма памяти, затрачиваемого на хранение одного информационного сообщения в цепочке сегментов неизвестной фиксированной длины x с указателем длины s , равно:

$$F(x) = (x + s) \sum_{k=0}^{\infty} (k + 1) \int_{kx}^{(k+1)x} \rho(l) dl. \quad (3)$$

Необходимо найти такое x_{OPT} , что

$$A(x_{OPT}) = \min. \quad (4)$$

В выражениях (3) и (4) существенным моментом является значение вида функции $\rho(l)$ плотности распределения вероятностей длин сообщений.

В работе (2) рассматриваются статистики распределения длин информационных сообщений в поисковых системах и распределение заявок программ на выделение оперативной памяти. В обоих этих случаях оказалось, что

$$\rho(l) = ble^{-rl}. \quad (5)$$

Это семейство функций зависит от параметра r , и имеет распределение с одним максимумом при $l = l_{\max}$.

Чтобы найти значение коэффициента b следует учесть, что $\int_0^{\infty} \rho(l) dl = 1$.

Найдём

$$\int \rho(l) dl = -\frac{be^{-rl}}{r} \left(l + \frac{1}{r} \right) + C. \quad (6)$$

Вычислим несобственный интеграл

$$\int_0^{\infty} \rho(l) dl = \frac{b}{r^2}.$$

Приравняв его единице, найдём, что $b = r^2$ и, следовательно,

$$\rho(l) = r^2 le^{-rl}. \quad (7)$$

Подставляя $b = r^2$ с помощью (6), вычислим

$$\int_{kx}^{(k+1)x} \rho(l) dl = -re^{-rkx} \left\{ e^{-rx} \left[(k+1)x + \frac{1}{r} \right] - \left(kx + \frac{1}{r} \right) \right\}. \quad (8)$$

Подставляя выражение (8) в (3), получим

$$f(x) = -r(x + s) \left\{ \left[\frac{1}{r} (e^{-rx} - 1) + xe^{-rx} \right] \sum_{k=0}^{\infty} e^{-krx} + \right.$$

$$+ \left[\frac{1}{r}(e^{-rx} - 1) + xe^{-rx} + x(e^{-rx} - 1) \right] \sum_{k=0}^{\infty} ke^{-krx} + \\ + x(e^{-rx} - 1) \sum_{k=0}^{\infty} k^2 e^{-krx} \} \quad (9)$$

Теперь вычислим бесконечные суммы фигурирующие в (9). Легко видеть, что

$$\sum_{k=0}^{\infty} e^{-krx} = \frac{1}{1 - e^{-rx}}. \quad (10)$$

Для вычисления второй суммы обозначим

$$\sum_{k=0}^{\infty} ke^{-krx} = f(x). \quad (11)$$

Поскольку ряд стоящий в левой части выражения (11), абсолютно и равномерно сходящимся, его можно почленно проинтегрировать, получив выражение

$$\int f(x) dx = \int \left(\sum_{k=0}^{\infty} ke^{-krx} \right) dx = \sum_{k=0}^{\infty} \int ke^{-krx} dx = \sum_{k=0}^{\infty} -\frac{1}{r} e^{-krx} = -\frac{1}{r} \sum_{k=0}^{\infty} e^{-krx}. \quad (12)$$

Подставляя в (12) найденную сумму (11), получим

$$\int f(x) dx = \frac{1}{r(e^{-rx} - 1)}.$$

Таким образом,

$$f(x) = \sum_{k=0}^{\infty} ke^{-krx} = \left[\frac{1}{r(e^{-rx} - 1)} \right]' = \frac{e^{-rx}}{(e^{-rx} - 1)^2}. \quad (13)$$

Аналогичным образом, сводя $\sum_{k=0}^{\infty} k^2 e^{-krx}$ к сумме (13), получим, что

$$\varphi(x) = \sum_{k=0}^{\infty} k^2 e^{-krx} = \frac{e^{-rx}(e^{-rx} + 1)}{(1 - e^{-rx})^3}. \quad (14)$$

Подставляя суммы (10), (13) и (14) в выражение (9) для $F(x)$ и производя необходимые преобразования, окончательно получим

$$F(x) = (x + s) \left[\frac{1}{1 - e^{-rx}} + \frac{rxe^{-rx}}{(1 - e^{-rx})^2} \right]. \quad (15)$$

Для нахождения оптимального значения $x_{OPT} \in (0, \infty)$, при котором справедливо (4), вычислим

$$F'(x) = \frac{1}{(1 - e^{-rx})^3} \left[1 - e^{-rx} (2 - rx + r^2 x^2 + r^2 xs) + e^{-2rx} (1 - rx - r^2 x^2 - r^2 xs) \right] \quad (16)$$

и приравняем выражение (16) нулю.

При решении уравнения $F'(x) = 0$ заметим, что $1 - e^{-rx} > 0$, и отбросим в (16) член с e^{-2rx} в виду его незначительности. Тогда x_{OPT} является корнем уравнения

$$e^{rx} = 2 - rx + r^2 x^2 + r^2 sx. \quad (17)$$

Произведём замену $z = rx$, $a = 1 - rs$, тогда

$$e^z = 2 - az + z^2. \quad (18)$$

Возможен следующий путь решения уравнения (18). Первое приближение найдём посредством разложения e^z в ряд Маклорена. Полученное первое приближение уточним по методу касательных Ньютона.

Перед началом решения проясним смысл параметра r в определении (7). Для этого заметим, что $\rho(l)$ достигает максимума при

$$\rho'(l) = r^2 e^{-rl} (1 - rl) = 0, \text{ т.е. } l_{\max} = \frac{1}{r}. \text{ В практических ситуациях наиболее}$$

вероятная длина информационного сообщения $l_{\max}$ даже в информационно-поисковых системах обычно превосходит 20-30 знаков и зачастую достигает

100-200 знаков, тогда $r = \frac{1}{l_{\max}} < 0,05$. Для задачи касательно виртуальной

памяти, $l_{\max}$ значительно больше 100-150 байт, отсюда $r < 0,01$.

В современных вычислительных машинах для организации ссылки s достаточно 4-х байт. Отсюда в самом худшем случае, для поисковых систем, $rs < 0,2$, а в задаче относительно размера страницы виртуальной памяти даже $rs < 0,04$, так что постоянная a в уравнении (18) удовлетворяет неравенству

$$0,8 < a < 1 \quad (19)$$

или более сильному неравенству

$$0,8 < a < 1. \quad (19a)$$

С учётом соотношения (19) исследуем количество корней в уравнении (18).

Рассматривая производную функции $P(z) = e^z - z^2 + az - 2$, легко

установить, что при $z > 0$ $p'(z) > 0$. Т.е. $p(z)$ монотонно возрастает и,

следовательно, может иметь не более одного положительного корня. Поскольку $p(0) = -1$, а $p(1) = e + a - 3 > 0$, то то искомый единственный положительный

корень $z_{ОПТ}$ существует и находится в интервале $[0, 1]$. Уточняя границы, найдём, что $P\left(\frac{1}{2}\right) \approx \frac{a}{2} - 0,6 < 0$, а $p\left(\frac{3}{4}\right) \approx \frac{3}{4}a - 0,44 > 0$, так что

$$z_{ОПТ} \in \left(\frac{1}{2}, \frac{3}{4}\right). \quad (20)$$

Теперь обратимся непосредственно к решению уравнения (18). Так как первое приближение будет уточняться с помощью метода Ньютона, ограничимся при разложении e^z тремя членами. Тогда уравнение (18) примет вид

$$1 + z + \frac{z^2}{2} = 2 - az + z^2.$$

Решая это квадратное уравнение, заменяя в решении $\sqrt{1-t} \approx 1 - \frac{t}{2}$ и отбрасывая один из корней в соответствии с (20), получим

$$z_{опт}^{прибл.} \approx \frac{1}{a+1}. \quad (21)$$

Заметим, что с учётом (19) найденное первое приближение $z_{ОПТ}$ удовлетворяет (20) и подтверждает правомочность такого разложения e^z .

Уточним его с помощью метода Ньютона:

$$z_{опт}^{прибл2} = z_{опт}^{прибл1} - \frac{p(z_{опт}^{прибл1})}{p'(z_{опт}^{прибл1})}.$$

Подставляя в это выражение (21) и выполняя необходимые преобразования, получим

$$z_{ОПТ} = \frac{\frac{1}{ae^{\frac{1}{a+1}}} - \frac{1}{(a+1)^2}}{a + e^{\frac{1}{a+1}} - \frac{2}{a+1}}. \quad (22)$$

Можно показать, что это значение $z_{ОПТ}$ также удовлетворяет ограничениям (20).

Так как формула (22) неудобна для практических вычислений и учитывая диапазон изменения a , задаваемый в (19), заменим $e^{\frac{1}{a+1}}$ на первые 4 члена разложения в ряд Маклорена. Оценив погрешность разложения (опустим эти преобразования), и учтя, что сама функция $f(a) = e^{\frac{1}{a+1}}$ с учётом (19) меняется

в пределах от 1.74 до 1.65, определим, что относительная погрешность такой замены не превосходит 3,6%.

Тогда (22) примет вид

$$z_{OPT} = \frac{6a^4 + 24a^3 + 33a^2 + 20a + 6}{6a^5 + 30a^4 + 54a^3 + 45a^2 + 19a + 4}. \quad (23)$$

Вернувшись к старым переменным с учётом того, что $r = \frac{1}{l_{\max}}$, и заменяя

$b = \frac{s}{l_{\max}}$, приведём выражение (23) к следующему виду:

$$x_{OPT} = l_{\max} \frac{89 - 182b + 141b^2 - 48b^3 + 6b^4}{158 - 421b + 447b^2 - 234b^3 + 60b^4 - 6b^5}.$$

Выполним здесь деление и отбросим члены с b^2 и далее. Это вполне допустимо, поскольку при оценке погрешности исследование остаточного члена

показывает, что он не превосходит 0,82 $b^2 = \frac{0,82 \cdot 16}{l_{\max}^2}$. Так как

$x_{OPT} \approx 0,564l_{\max}$, то относительная погрешность не превосходит

$$\frac{0,82 \cdot 16}{0,564(l_{\max})^3} = \frac{23}{l_{\max}^3}.$$

Поскольку, как мы указывали ранее, даже в информационно поисковых-системах $l_{\max}$ обычно превосходит 20-30, то

относительная погрешность из-за отбрасывания членов, начиная с b^2 , заметно меньше 0,5%. В результате этих действий получим

$$x_{OPT} = l_{\max} (0,564 + 0,352b) = 0,564l_{\max} + 0,352s.$$

Исследуя знак $F''(x_{OPT})$, можно показать, что в этой точке действительно достигается минимальное значение $F(x)$.

Если учесть, что в сегменте должен ещё размещаться указатель длиной s знаков, получим окончательно, что полная длина оптимального сегмента составляет с погрешностью не более 4%:

$$x_{OPT} = 0,564l_{\max} + 1,352s. \quad (24)$$

3. Порядок работы

1. Получить у преподавателя вариант задания.
2. Изучить теоретический раздел.

3. Используя выражение (7), для параметров указанных в таблице 1, построить функции распределения плотности вероятности длин сообщений (для трех различных r при заданном l). Графики представить в отчёте.
3. По графикам функций распределения определить величину $l_{\max}$.
4. Используя выражение (24), для трёх пар значений параметров $l_{\max}$ и соответствующему варианту s , вычислить оптимальную длину сегмента памяти $X_{\text{ОПТ}}$.
5. Для полученных трёх значений $X_{\text{ОПТ}}$ построить график функции $X_{\text{ОПТ}}(r)$.
6. Проанализировать результаты. Сделать выводы по работе. Написать отчёт.
7. Все расчёты и построения производить в системе MATLAB.

4. Варианты заданий

Таблица 1 – Варианты заданий

№	1	2	3	4	5	6	7	8	9	10
l	100	150	200	250	300	350	400	450	500	550
s	2	2	2	2	2	2	2	2	2	2
r_1	1/25	1/38	1/50	1/63	1/75	1/88	1/100	1/113	1/125	1/138
r_2	1/50	1/75	1/100	1/125	1/150	1/175	1/200	1/225	1/250	1/275
r_3	1/75	1/112	1/150	1/188	1/225	1/263	1/300	1/338	1/375	1/413
№	11	12	13	14	25	16	17	18	19	20
l	600	650	700	750	800	850	900	950	1000	1100
s	3	3	3	3	3	3	3	3	3	3
r_1	1/150	1/163	1/175	1/188	1/200	1/213	1/225	1/238	1/250	1/275
r_2	1/300	1/325	1/350	1/375	1/400	1/425	1/500	1/475	1/500	1/550
r_3	1/450	1/488	1/525	1/563	1/600	1/638	1/725	1/713	1/750	1/825
№	21	22	23	24	25	26	27	28	29	30
l	1200	1300	1400	1500	1600	1700	1800	1900	2000	2100
s	4	4	4	4	4	4	4	4	4	4
r_1	1/300	1/325	1/350	1/375	1/400	1/425	1/450	1/475	1/500	1/525
r_2	1/600	1/650	1/700	1/750	1/800	1/850	1/900	1/950	1/1000	1/1050
r_3	1/900	1/975	1/1050	1/1125	1/1200	1/1275	1/1350	1/1425	1/1500	1/1575

5. Контрольные вопросы

1. Что такое плотность вероятности?

2. Какую роль играют параметры x , s , l , k ?
3. Какова основная идея решения задачи по выбору оптимального размера сегмента?
4. Каков смысл параметра r ?
5. Что подразумевается под страничной организацией памяти?
6. Что показывает параметр $l_{\max}$?
7. Какова зависимость величины сегмента памяти от размера исходного сообщения и величины r ?

6. Содержание отчёта

Отчёт по лабораторной работе оформляется каждым студентом индивидуально. Отчёт должен включать: название и номер лабораторной работы; цель работы; постановку задачи; основные математические выражения, при помощи которых проводились требуемые расчёты; текст программы, реализующей поставленную задачу; результаты выполнения программы; выводы по работе.

Список литературы

1. Гендель Е. Г., Левин Н. А. Оптимизация технологии обработки информации в АСУ. М., «Статистика», 1977.
2. Белоногов Г. Г., Богатырёв В. И. Автоматизированные информационные системы. Б. «Сов. Радио», 1973.