Труды ИПМИ КарНЦ РАН

Выпуск 9, 2008

УДК 519.142.2 + 519.2

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ ОЧЕРЕДЬЮ С ДВУМЯ ПРИОРИТЕТАМИ В ПАМЯТИ ОДНОГО УРОВНЯ

Е. А. Аксенова

В статье рассмотрена задача управления двухприоритетной очередью, представленой в виде двух FIFO-очередей, для трех способов организации хранения данных.

Работа выполнена при финансовой поддержке РФФИ, грант №06-01-00303.

1. Постановка задачи

Во многих приложениях требуется обработка записей с упорядоченными определенным образом ключами. Часто накапливается некоторый набор записей, после чего обрабатывается запись с максимальным значением ключа, затем, возможно, накопление записей продолжается, затем обрабатывается запись с наибольшим текущим ключом и т. д. Соответствующая структура данных, поддерживающая операции вставки нового элемента и удаления элемента с наибольшим приоритетом, называется очередью с приоритетами [1, 2].

Рассмотрим очередь с двумя приоритетами, расположенную в области памяти размера m единиц. Такую приоритетную очередь представим в виде двух FIFO-очередей. Первой FIFO-очереди присвоим приоритет 1, второй FIFO-очереди – приоритет 2. Наивысший приоритет 2. Будем считать, что время дискретно. В каждый момент времени могут произойти следующие операции:

- включение элемента с первым приоритетом с вероятностью p_1 (включение элемента в первую FIFO-очередь),
- включение элемента со вторым приоритетом с вероятностью p_2 (включение элемента во вторую FIFO-очередь),
- исключение элемента из очереди с вероятностью q,

- очередь не изменяет своей длины с вероятностью r (только чтение или отсутствие операции),

где $p_1+p_2+q+r=1$. Исключение элемента из очереди происходит по наивысшему приоритету. Это означает, что пока вторая FIFO-очередь не пуста, с вероятностью q исключение элементов происходит из этой очереди. Как только вторая FIFO-очередь станет пустой, с вероятностью q исключение элементов будет происходить из первой FIFO-очереди. При исключении элемента из пустой очереди не происходит завершение работы.

Рассмотрим последовательный, связанный и страничный способы организации FIFO-очередей. Необходимо определить, как распределить память между FIFO-очередями в последовательном способе, и какой из способов организации очередей является оптимальным. В качестве критерия оптимальности рассмотрим максимальное среднее время работы с приоритетной очередью до переполнения выделенного объема памяти.

2. Последовательное представление

В последовательном способе организации каждой FIFO-очереди выделено некоторое количество единиц памяти из данных m единиц. Пусть s – количество единиц памяти, выделенных первой FIFO-очереди, тогда $(m-s)$ – количество единиц памяти, выделенных второй FIFO-очереди (Рис. 1).

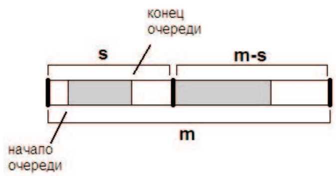

Рис. 1. Последовательное представление

2.1. Математическая модель и матрица вероятностей переходов. Обозначим текущие длины FIFO-очередей x_1 и x_2. В качестве математической модели процесса работы с приоритетной очередью рассмотрим случайное блуждание по целочисленной решетке в двухмерном пространстве $0 \leqslant x_1 < s+1$, $0 \leqslant x_2 < m-s+1$ (Рис. 2), где

$x_1 = s+1$, $x_2 = m-s+1$ – поглощающие экраны, попадание на которые характеризуется как переполнение одной из FIFO-очередей, $(-1, 0)$ и $(0, -1)$ – отражающие точки, т.к. только в состоянии $x_1 = x_2 = 0$ возможно исключение из пустой очереди. Исключение элементов из первой FIFO-очереди может происходить только при $x_2 = 0$.

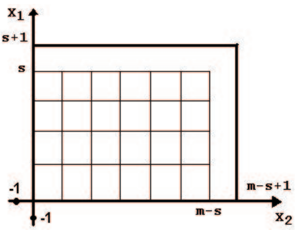

Рис. 2. Область блуждания при последовательном представлении

Блуждание начинается в точке $x_1 = x_2 = 0$. Необходимо определить такое значение $0 \leqslant s \leqslant m$, чтобы время блуждания до попадания на поглощающие экраны было максимально.

Рассмотрим случай, когда в очереди хранятся данные, длина информационной части которых равна $L > 1$ единиц памяти. Предполагаем, что m кратно L. Количество элементов в этом случае равно m/L. В качестве математической модели в этом случае будем иметь

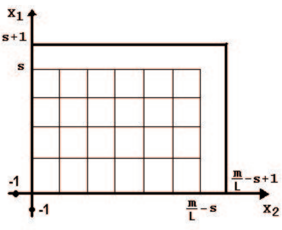

Рис. 3. Область блуждания, когда длина информационной части равна L

случайное блуждание по целочисленной решетке в двухмерном пространстве $0 \leqslant x_1 < s+1$, $0 \leqslant x_2 < m/L - s + 1$ (Рис. 3), где $x_1 = s+1$, $x_2 = m/L - s + 1$ – поглощающие экраны, $(-1, 0)$ и $(0, -1)$ – отражающие точки.

Блуждание начинается в точке $x_1 = x_2 = 0$. Необходимо определить такое значение $0 \leqslant s \leqslant m/L$, чтобы время блуждания до попадания на поглощающие экраны было максимально.

Если в произвольный момент времени процесс блуждания находится в состоянии (x_1, x_2), то в следующий момент времени он переходит в состояние (x'_1, x'_2), которое определяется правилами:

$$(x_1, x_2) \xrightarrow{p_1} (x'_1, x'_2) = \{(x_1 + 1, x_2), \quad 0 \leqslant x_1 \leqslant s,\ 0 \leqslant x_2 \leqslant m - s;$$

$$(x_1, x_2) \xrightarrow{p_2} (x'_1, x'_2) = \{(x_1, x_2 + 1), \quad 0 \leqslant x_1 \leqslant s,\ 0 \leqslant x_2 \leqslant m - s;$$

$$(x_1, x_2) \xrightarrow{q} (x'_1, x'_2) = \left\{ \begin{array}{l} (x_1 - 1, x_2), \quad 0 < x_1 \leqslant s,\ x_2 = 0; \\ (x_1, x_2 - 1), \quad 0 \leqslant x_1 \leqslant s,\ 0 < x_2 \leqslant m - s; \\ (x_1, x_2), \quad x_1 = 0,\ x_2 = 0; \end{array} \right.$$

$$(x_1, x_2) \xrightarrow{r} (x'_1, x'_2) = \{(x_1, x_2), \quad 0 \leqslant x_1 \leqslant s,\ 0 \leqslant x_2 \leqslant m - s.$$

Если длина информационной части элемента в очереди равна L, то вместо m в этом случае будем рассматривать m/L.

Случайное блуждание будем рассматривать как конечную однородную поглощающую цепь Маркова [3]. Пронумеруем невозвратные состояния цепи так, как показано графически (Рис. 4). Количество невозвратных состояний в такой цепи равно $(s+1)(m-s+1)$. Например, для $m = 6, s = 4$ количество невозвратных состояний равно 15 (Рис. 4).

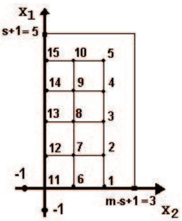

Рис. 4. Область блуждания и нумерация невозвратных состояний при $m = 6, s = 4$

Рассмотрим матрицу Q вероятностей переходов из невозвратных состояний в невозвратные. Для размера памяти m и заданного s размерность матрицы Q будет $(s+1)(m-s+1) \times (s+1)(m-s+1)$.

Теорема 1. *Матрица Q при заданной нумерации, размере памяти m и заданной величине $s > 0$ имеет следующий вид:*

$$Q = \begin{pmatrix} A_{(s+1)\times(s+1)} & B_{(s+1)\times(s+1)} & & & 0 \\ C_{(s+1)\times(s+1)} & A_{(s+1)\times(s+1)} & \dots & & \\ & C_{(s+1)\times(s+1)} & \dots & B_{(s+1)\times(s+1)} & \\ & & \dots & A_{(s+1)\times(s+1)} & B_{(s+1)\times(s+1)} \\ 0 & & & C_{(s+1)\times(s+1)} & A'_{(s+1)\times(s+1)} \end{pmatrix},$$

где

$$A_{(s+1)\times(s+1)} = \begin{pmatrix} r & p_1 & & & 0 \\ & r & p_1 & & \\ & & \dots & \dots & \\ & & & r & p_1 \\ 0 & & & & r \end{pmatrix},$$

$$A'_{(s+1)\times(s+1)} = \begin{pmatrix} r+q & p_1 & & & 0 \\ q & r & p_1 & & \\ & q & \dots & \dots & \\ & & \dots & \dots & p_1 \\ 0 & & & q & r \end{pmatrix},$$

$$B_{(s+1)\times(s+1)} = \begin{pmatrix} q & & & & 0 \\ & q & & & \\ & & \dots & & \\ & & & q & \\ 0 & & & & q \end{pmatrix},$$

$$C_{(s+1)\times(s+1)} = \begin{pmatrix} p_2 & & & & 0 \\ & p_2 & & & \\ & & \dots & & \\ & & & p_2 & \\ 0 & & & & p_2 \end{pmatrix}.$$

Доказательство. Индукция проводится по размеру памяти m.

1) Пусть размер памяти $m = 2, s = 1$ (Рис. 5). Размерность матри-

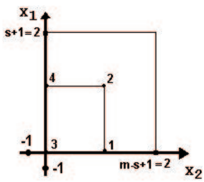

Рис. 5. Область блуждания и нумерация невозвратных состояний при $m = 2, s = 1$

цы Q будет 4×4, и матрица имеет вид

$$Q = \left(\begin{array}{cc|cc} r & p_1 & q & 0 \\ 0 & r & 0 & q \\ \hline p_2 & 0 & r+q & p_1 \\ 0 & p_2 & q & r \end{array}\right).$$

Пусть $m = 6, s = 4$ (Рис. 4). Размерность матрицы Q будет 15×15, и матрица имеет вид

$$Q = \left(\begin{array}{ccccc|ccccc|ccccc}
r & p_1 & 0 & 0 & 0 & q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & r & p_1 & 0 & 0 & 0 & q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & r & p_1 & 0 & 0 & 0 & q & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & r & p_1 & 0 & 0 & 0 & q & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & r & 0 & 0 & 0 & 0 & q & 0 & 0 & 0 & 0 & 0 \\
\hline
p_2 & 0 & 0 & 0 & 0 & r & p_1 & 0 & 0 & 0 & q & 0 & 0 & 0 & 0 \\
0 & p_2 & 0 & 0 & 0 & 0 & r & p_1 & 0 & 0 & 0 & q & 0 & 0 & 0 \\
0 & 0 & p_2 & 0 & 0 & 0 & 0 & r & p_1 & 0 & 0 & 0 & q & 0 & 0 \\
0 & 0 & 0 & p_2 & 0 & 0 & 0 & 0 & r & p_1 & 0 & 0 & 0 & q & 0 \\
0 & 0 & 0 & 0 & p_2 & 0 & 0 & 0 & 0 & r & 0 & 0 & 0 & 0 & q \\
\hline
0 & 0 & 0 & 0 & 0 & p_2 & 0 & 0 & 0 & 0 & r+q & p_1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & p_2 & 0 & 0 & 0 & q & r & p_1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & p_2 & 0 & 0 & 0 & q & r & p_1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_2 & 0 & 0 & 0 & q & r & p_1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_2 & 0 & 0 & 0 & q & r
\end{array}\right).$$

2) Предположим, что для размера памяти $m = d - 1$ и заданного значения $s > 0$ выполняется утверждение теоремы. Количество невозвратных состояний в цепи равно $(s+1)(d-1-s+1) = (s+1)(d-s)$. Размерность матрицы Q будет $(s+1)(d-s) \times (s+1)(d-s)$ и матрица имеет вид

$$Q = \begin{pmatrix} A_{(s+1)\times(s+1)} & B_{(s+1)\times(s+1)} & & & 0 \\ C_{(s+1)\times(s+1)} & A_{(s+1)\times(s+1)} & \dots & & \\ & C_{(s+1)\times(s+1)} & \dots & B_{(s+1)\times(s+1)} & \\ & & \dots & A_{(s+1)\times(s+1)} & B_{(s+1)\times(s+1)} \\ 0 & & & C_{(s+1)\times(s+1)} & A'_{(s+1)\times(s+1)} \end{pmatrix}.$$

3) Проверим, что при размере памяти $m = d$ выполняется утверждение теоремы. Если добавить еще одну единицу памяти, то при фиксированном s поглощающая граница области блуждания при $m = d - 1$ $x_2 = d - s$ сдвинется на единицу вправо и будет $x_2 = d + 1$. Тогда получим $s + 1$ новых невозвратных состояний. Согласно введенной нумерации, эти состояния будут иметь младшие номера. Количество невозвратных состояний будет $(s+1)(d-s) + (s+1) = (s+1)(d-s+1)$. Тогда в матрицу Q при $m = d - 1$ добавятся подматрицы: $A_{(s+1)\times(s+1)}$, $B_{(s+1)\times(s+1)}$, $C_{(s+1)\times(s+1)}$. Получили матрицу

Q размерности $(s+1)(d-s+1) \times (s+1)(d-s+1)$:

$$Q = \left(\begin{array}{c|cccc} A_{(s+1)\times(s+1)} & B_{(s+1)\times(s+1)} & & & \\ \hline C_{(s+1)\times(s+1)} & A_{(s+1)\times(s+1)} & B_{(s+1)\times(s+1)} & & 0 \\ & C_{(s+1)\times(s+1)} & \cdots & \cdots & \\ & & \cdots & A_{(s+1)\times(s+1)} & B_{(s+1)\times(s+1)} \\ & 0 & & C_{(s+1)\times(s+1)} & A'_{(s+1)\times(s+1)} \end{array}\right).$$

Теорема доказана.

Подматрицы матрицы Q характеризуют определенные операции над очередями. Подматрица A характеризует операции, при которых не изменяется длина очереди (например, чтение), и операции включения элементов в очередь с приоритетом 1, подматрица B – операции исключения элементов из очереди с приоритетом 2, подматрица C – операции включения элементов в очереди, подматрица A' – операции, при которых не изменяется длина очереди и операции включения и исключения элементов в очередь с приоритетом 1.

3. Связанное представление

В связанном способе организации FIFO-очередь представлена в виде связанного списка, каждый элемент которого состоит из двух единиц памяти, одна из которых используется для хранения информации, вторая содержит указатель на следующий элемент в списке (Рис. 6).

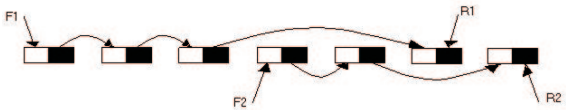

Рис. 6. Связанное представление

F_1 и F_2 – указатели на первые элементы FIFO-очередей, R_1 и R_2 – указатели на последние элементы FIFO-очередей. Предполагаем, что m кратно 2. Для хранения указателей используется $m/2$ единиц памяти, для хранения информации используется $m/2$ единиц памяти. Количество элементов равно $m/2$.

3.1. **Математическая модель.** Обозначим текущие длины FIFO-очередей x_1 и x_2. В качестве математической модели процесса работы с приоритетной очередью рассмотрим случайное блуждание по целочисленной решетке в треугольной области $x_1 \geqslant 0$, $x_2 \geqslant 0$, $x_1+x_2 < m/2+1$ (Рис. 7), где $x_1+x_2 = m/2+1$ – поглощающий экран, попадание на который характеризуется как переполнение выделенного объема памяти, $(-1, 0)$ и $(0, -1)$ – отражающие точки, т.к. только в состоянии $x_1 = x_2 = 0$ возможно исключение из пустой очереди. Исключение элементов из первой FIFO-очереди может происходить только при $x_2 = 0$.

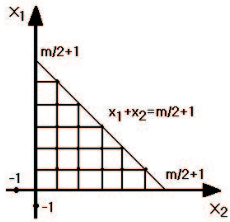

Рис. 7. Область блуждания для связанного представления

Блуждание начинается в точке $x_1 = x_2 = 0$. Необходимо найти среднее время блуждания T до попадания на поглощающий экран.

Рассмотрим случай, когда в очереди хранятся данные, длина информационной части которых равна $L > 1$ единиц памяти. Тогда длина каждого элемента очереди будет $L + 1$, где одна единица памяти содержит указатель на следующий элемент. Предполагаем, что m кратно $L+1$. Количество элементов в этом случае будет $m/(L+1)$. Для хранения указателей используется $m/(L+1)$ единиц памяти. В качестве математической модели будем иметь случайное блуждание по целочисленной решетке в треугольной области $x_1 \geqslant 0$, $x_2 \geqslant 0$, $x_1 + x_2 < m/(L+1) + 1$ (Рис. 8), где $x_1 + x_2 = m/(L+1) + 1$ – поглощающий экран, $(-1, 0)$ и $(0, -1)$ – отражающие точки.

Блуждание начинается в точке $x_1 = x_2 = 0$. Необходимо найти среднее время блуждания T до попадания на поглощающий экран и сравнить со средним временем блуждания в случае последовательного представления, когда длина информационной части равна L.

4. Страничное представление

В страничном способе организации FIFO-очередь представлена в виде связанного списка страниц размера x единиц памяти. Если одна FIFO-очередь заняла страницу, то другая FIFO-очередь не может ее использовать (Рис. 9). F_1 и F_2 – указатели на первые элементы FIFO-очередей, R_1 и R_2 – указатели на последние элементы FIFO-очередей.

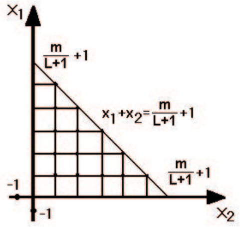

Рис. 8. Область блуждания, когда длина информационной части равна L

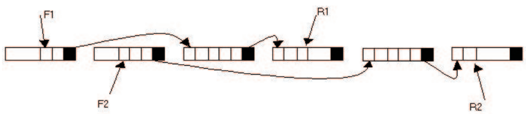

Рис. 9. Страничное представление

Предполагаем, что m кратно x. Для хранения указателей используется m/x единиц памяти, для хранения информации используется $m - m/x$ единиц памяти. Количество страниц равно m/x. Для хранения информации у каждой страницы используется $x - 1$ единиц памяти. При $x = 2$ получаем связанное представление очередей.

Предполагаем, что существует механизм работы со списком свободных страниц, который позволяет при заполнении страницы предоставить очереди новую страницу. Если в очереди освободилась страница, то она возвращается в список свободных страниц. Переполнение наступает тогда, когда список свободных страниц пуст и какая-то из FIFO-очередей переполнила страницу.

4.1. **Математическая модель и матрица вероятностей переходов.** Обозначим текущие длины FIFO-очередей x_1 и x_2. В качестве математической модели процесса работы с приоритетной очередью рассмотрим случайное блуждание по целочисленной решетке в треугольной области $x_1 \geqslant 0$, $x_2 \geqslant 0$, $x_1 + x_2 < m - m/x + 1$ (Рис. 10), где $x_1+x_2 = m-m/x+1$ – поглощающий экран, попадание на который характеризуется как переполнение выделенного объема памяти, $(-1, 0)$ и $(0, -1)$ – отражающие точки, т.к. только в состоянии $x_1 = x_2 = 0$

возможно исключение из пустой очереди. Исключение элементов из первой FIFO-очереди может происходить только при $x_2 = 0$.

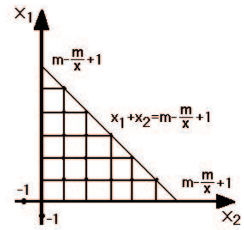

Рис. 10. Область блуждания для страничного представления

Блуждание начинается в точке $x_1 = x_2 = 0$. Необходимо найти среднее время блуждания T до попадания на поглощающий экран.

Обозначим n – количество единиц памяти, которые используются для хранения информации. Для страничного представления FIFO-очередей $n = m - m/x$. Для связанного представления FIFO-очередей $n = m/2$, если длина информационной части равна 1, и $n = m/(L+1)$, если длина информационной части равна $L > 1$.

Если в произвольный момент времени процесс блуждания находится в состоянии (x_1, x_2), то в следующий момент времени он переходит в состояние (x_1', x_2'), которое определяется правилами:

$$(x_1, x_2) \xrightarrow{p_1} (x_1', x_2') = \{(x_1 + 1, x_2), \quad x_1 \geqslant 0,\ x_2 \geqslant 0,\ 0 \leqslant x_1 + x_2 \leqslant n;$$

$$(x_1, x_2) \xrightarrow{p_2} (x_1', x_2') = \{(x_1, x_2 + 1), \quad x_1 \geqslant 0,\ x_2 \geqslant 0,\ 0 \leqslant x_1 + x_2 \leqslant n;$$

$$(x_1, x_2) \xrightarrow{q} (x_1', x_2') = \begin{cases} (x_1 - 1, x_2), \quad 0 < x_1 \leqslant n,\ x_2 = 0; \\ (x_1, x_2 - 1), \quad x_1 \geqslant 0,\ x_2 > 0,\ 0 < x_1 + x_2 \leqslant n; \\ (x_1, x_2), \quad x_1 = 0,\ x_2 = 0; \end{cases}$$

$$(x_1, x_2) \xrightarrow{r} (x_1', x_2') = \{(x_1, x_2), \quad x_1 \geqslant 0,\ x_2 \geqslant 0,\ 0 \leqslant x_1 + x_2 \leqslant n.$$

Случайное блуждание будем рассматривать как конечную однородную поглощающую цепь Маркова. Пронумеруем невозвратные состояния цепи так, как показано на рисунке 11, т.е. снизу вверх, справа налево. Количество невозвратных состояний вычисляется с помощью формулы суммы арифметической прогрессии и равно $(n+1)(n+2)/2$.

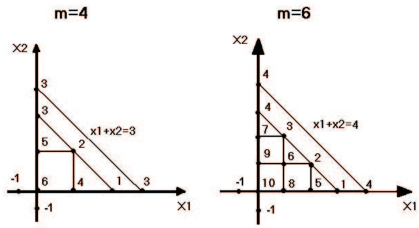

Рис. 11. Нумерация невозвратных состояний при $x = 2$ для $m = 4$ и $m = 6$

Рассмотрим примеры построения области блуждания для $m = 4$ и $m = 6$ при $x = 2$ (Рис. 11). Для $m = 4$ $n = 2$, количество невозвратных состояний равно $3 \cdot 4/2 = 6$, поглощающий экран $x_1 + x_2 = n + 1 = 3$. Для $m = 6$ $n = 3$, количество невозвратных состояний равно $4 \cdot \cdot 5/2 = 10$, поглощающий экран $x_1 + x_2 = n + 1 = 4$. Размеры памяти меньше 4 рассматривать не будем, т.к. в этом случае не имеет смысла рассматривать страничный способ.

Рассмотрим матрицу Q вероятностей переходов из невозвратных состояний в невозвратные.

Размерность матрицы Q будет $(n + 1)(n + 2)/2 \times (n + 1)(n + 2)/2$.

Теорема 2. *Матрица Q при заданной нумерации и величине n имеет следующий вид:*

$$Q = \begin{pmatrix} A_{n+1} & B_n & & & & 0 \\ C_{n+1} & A_n & B_{n-1} & & & \\ & C_n & A_{n-1} & \dots & & \\ & & C_{n-1} & \dots & B_2 & \\ & & & \dots & A_2 & B_1 \\ 0 & & & & C_2 & A_1 \end{pmatrix},$$

где подматрица A_k размера $k \times k$

$$A_k = \begin{pmatrix} r & & & & 0 \\ & r & & & \\ & & \dots & & \\ & & & r & \\ 0 & & & & r \end{pmatrix}, \quad k = n + 1, n, \dots, 3, 2;$$

$$A_1 = (r + q),$$

подматрица B_k размера $(k+1) \times k$

$$B_k = \begin{pmatrix} q & & & 0 \\ & q & & \\ & & \dots & \\ & & & q \\ 0 & & & q \end{pmatrix}, \quad k = n, n-1, \dots, 2, 1;$$

подматрица C_k размера $(k-1) \times k$

$$C_k = \begin{pmatrix} p_2 & p_1 & & & 0 \\ & p_2 & p_1 & & \\ & & \dots & \dots & \\ 0 & & & p_2 & p_1 \end{pmatrix}, \quad k = n+1, n, \dots, 3, 2.$$

Доказательство. Индукция проводится по количеству единиц памяти n, которое используется для хранения информации.

1) Пусть размер памяти $m = 4, x = 2, n = 2$ (Рис. 11). Размерность матрицы Q будет 6×6.

Рассмотрим состояние 2. Из него можно перейти в состояние 5 с вероятностью q исключения элемента из очереди с приоритетом 2, можно остаться в состоянии 2 с вероятностью r. Остальные элементы в этой строке будут нулями, т.к. переходы из состояния 2 в состояния 1,3,4,6 невозможны.

Рассмотрим состояние 5. Из него можно перейти в состояние 3 с вероятностью p_1 включения элемента в очередь с приоритетом 1, в состояние 2 с вероятностью p_2 включения элемента в очередь с приоритетом 2, в состояние 6 с вероятностью q исключения элемента из очереди с приоритетом 1 (исключения из очереди с приоритетом 1 могут происходить только при $x_2 = 0$, т.е. если очередь с приоритетом 2 пуста), можно остаться в состоянии 5 с вероятностью r. Остальные элементы в этой строке будут нулями, т.к. переходы из состояния 5 в состояния 1,4 невозможны.

Рассмотрим состояние 6. Из него можно перейти в состояние 4 с вероятностью p_2 включения элемента в первую очередь, можно перейти в состояние 5 с вероятностью p_1 включения элемента во вторую очередь, можно остаться в состоянии 6 с вероятностью $r+q$ (q прибавляется потому, что в этом состоянии обе очереди пусты, а исключение из пустой очереди не меняет ее длины). Остальные элементы в этой строке будут нулями, т.к. переходы из состояния 6 в состояния 1,2,3 невозможны. Другие состояния рассматриваются аналогично. Матрица Q имеет вид:

$$Q = \left(\begin{array}{ccc|cc|c} r & 0 & 0 & q & 0 & 0 \\ 0 & r & 0 & 0 & q & 0 \\ 0 & 0 & r & 0 & q & 0 \\ \hline p_2 & p_1 & 0 & r & 0 & q \\ 0 & p_2 & p_1 & 0 & r & q \\ \hline 0 & 0 & 0 & p_2 & p_1 & r+q \end{array}\right).$$

Если рассмотреть $m = 6, x = 3, n = 2$, то область блуждания и матрица Q будут такие же, как при $m = 4, x = 2, n = 2$.

Пусть $m = 6, x = 2, n = 3$ (Рис. 11). Размерность матрицы Q будет 10×10, и матрица имеет вид:

$$Q = \left(\begin{array}{cccc|ccc|cc|c} r & 0 & 0 & 0 & q & 0 & 0 & 0 & 0 & 0 \\ 0 & r & 0 & 0 & 0 & q & 0 & 0 & 0 & 0 \\ 0 & 0 & r & 0 & 0 & 0 & q & 0 & 0 & 0 \\ 0 & 0 & 0 & r & 0 & 0 & q & 0 & 0 & 0 \\ \hline p_2 & p_1 & 0 & 0 & r & 0 & 0 & q & 0 & 0 \\ 0 & p_2 & p_1 & 0 & 0 & r & 0 & 0 & q & 0 \\ 0 & 0 & p_2 & p_1 & 0 & 0 & r & 0 & q & 0 \\ \hline 0 & 0 & 0 & 0 & p_2 & p_1 & 0 & r & 0 & q \\ 0 & 0 & 0 & 0 & 0 & p_2 & p_1 & 0 & r & q \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_2 & p_1 & r+q \end{array}\right).$$

2) Предположим, что для размера памяти $n = d - 1$ выполняется утверждение теоремы. Количество невозвратных состояний в цепи будет $d(d+1)/2$. Матрица Q имеет размерность $d(d+1)/2 \times d(d+1)/2$ и блочную структуру:

$$Q = \begin{pmatrix} A_d & B_{d-1} & & & & 0 \\ C_d & A_{d-1} & B_{d-2} & & & \\ & C_{d-1} & A_{d-2} & \dots & & \\ & & C_{d-2} & \dots & B_2 & \\ & & & \dots & A_2 & B_1 \\ 0 & & & & C_2 & A_1 \end{pmatrix}.$$

3) Проверим, что при размере памяти $n = d$ выполняется утверждение теоремы. Если добавить еще одну единицу памяти, тогда поглощающая граница сдвинется вправо на единицу и будет $x_1 + x_2 = d+1$. Тогда в цепи появятся $d+1$ новых невозвратных состояний. Количество невозвратных состояний в цепи будет равно $d(d+1)/2+(d+1) = (d+1)(d+2)2$. Количество строк и столбцов в матрице Q при $n = d-1$ увеличится на $d+1$. Согласно введенной нумерации, новые состояния будут иметь младшие номера, а состояния цепи при $n = d - 1$ будут

следовать за ними. Тогда в матрицу Q добавятся подматрицы: A_{d+1} размерности $(d+1)\times(d+1)$, B_d размерности $(d+1)\times d$, C_{d+1} размерности $d\times(d+1)$, которые будут описывать переходы из новых $d+1$ состояний. Получили матрицу Q размерности $(d+1)(d+2)/2\times$ $\times(d+1)(d+2)/2$:

$$Q=\left(\begin{array}{c|ccccc} A_{d+1} & B_d & & & & \\ \hline C_{d+1} & A_d & B_{d-1} & & & 0 \\ & C_d & A_{d-1} & \dots & & \\ & & C_{d-1} & \dots & B_2 & \\ & & & \dots & A_2 & B_1 \\ & 0 & & & C_2 & A_1 \end{array}\right).$$

Теорема доказана.

Подматрицы матрицы Q описывают определенные операции над очередями. Подматрица A характеризует операции, при которых не изменяется длина очереди (например, чтение), подматрица B – операции исключения информации из очередей, подматрица C – операции включения информации в очереди.

5. Решение задачи

Для решения задачи вычислим фундаментальную матрицу $N=(I-Q)^{-1}$. В случае последовательного представления необходимо построить матрицу Q, вычислить матрицу N для всевозможных значений s. Для вычисления среднего времени блуждания просуммируем элементы матрицы N в строке, которая соответствует начальному состоянию $x_1=x_2=0$. Согласно введенной нумерации это будет строка с номером $(s+1)(m-s+1)-s$. Затем сравним полученное время для разных значений s и выберем максимальное. Соответствующее максимальному времени значение s будет оптимальным разбиением памяти для последовательного представления.

В случае связанного и страничного представления необходимо построить матрицу Q, вычислить матрицу N и просуммировать элементы матрицы N в строке, которая соответствует начальному состоянию $x_1=x_2=0$. Согласно введенной нумерации это будет последняя строка матрицы N.

В случае страничного представления вычисленное среднее время блуждания будет оценкой сверху для реального среднего времени. Обозначим эту оценку T_{max}. Также, можно вычислить оценку снизу

T_{min} для среднего времени блуждания. Для оценки сверху рассмотрим блуждание в области $x_1 + x_2 < m - m/x + 1$, для оценки снизу блуждание в области $x_1 + x_2 < m - m/x - 3(x-2) + 1$. При связанном представлении вычисленное время T – это среднее время блуждания.

Для страничного представления оптимальный размер страницы равен $x = \sqrt{6 \cdot m}/3$. Рассматривая, для какого размера памяти m область блуждания в случае связанного представления может быть меньше, чем область блуждания в случае страничного представления для минимальной оценки среднего времени, получаем, что для $m > 24$ при $x = \sqrt{6 \cdot m}/3$ и для любого m при $2 < x < m/6$ среднее время блуждания до поглощения в случае связанного представления будет меньше, чем минимальная оценка среднего времени в случае страничного представления.

6. Результаты численных экспериментов

Для реализации алгоритма решения задачи разработаны программы для ЭВМ, которые генерируют матрицу Q, вычисляют матрицу N и среднее время работы для каждого из рассмотренных способов представления.

В таблицах 1, 2, 3, 4 представлены результаты вычислений. В первых четырех столбцах находятся вероятностные характеристики приоритетной очереди, в столбце с названием m содержится размер выделенной памяти.

Столбец с названием A_1 соответствует последовательному представлению для размера памяти m, столбец с названием A_1' соответствует последовательному представлению для размера памяти $m-2$, т.е. когда учитываются две единицы памяти при реализации двух циклических FIFO-очередей. В столбце с названием s находится оптимальное количество единиц памяти, выделенных первой FIFO-очереди, в столбце с названием T находится среднее время работы.

В столбце с названием A_2 находится среднее время работы T для связанного представления. В столбце с названием A_3 находятся минимальная T_{min} и максимальная T_{max} оценки среднего времени для страничного представления и оптимальный размер страницы x.

Из представленных результатов в таблицах 1 и 2 видно, что при $m = 20$ $(m < 24)$ среднее время для связанного представления находится между оценками T_{min} и T_{max} для страничного представления. При $m = 32$ $(m > 24)$ видно, что среднее время для связанного

Таблица 1. Среднее время работы с очередью по приоритетам для $m = 20$

r	p_1	p_2	q	m	A_1		A'_1		A_2	A_3		
					s	T	s	T	T	x	T_{min}	T_{max}
0.25	0.25	0.25	0.25	20	13	62.99	12	56.38	40.0	4	36.0	60.0
0.1	0.4	0.4	0.1	20	11	25.4	10	22.82	15.51	4	14.08	22.65
0	0.1	0.6	0.3	20	5	43.02	5	38.76	25.63	4	23.13	38.13
0.2	0.6	0.1	0.1	20	17	30.47	15	27.49	18.06	4	16.39	26.39
0	0.7	0.3	0	20	13	18.4	12	16.6	11.0	4	10.0	16.0
0	0.3	0.7	0	20	7	18.4	6	16.6	11.0	4	10.0	16.0
0.1	0.3	0.2	0.4	20	15	128.69	13	111.61	73.44	4	64.3	121.13
0.1	0.2	0.3	0.4	20	13	121.58	11	106.2	73.44	4	64.3	121.13
0.5	0.2	0.2	0.1	20	12	57.09	10	51.2	35.56	4	32.22	52.22

Таблица 2. Среднее время работы с очередью по приоритетам для $m = 32$

r	p_1	p_2	q	m	A_1		A'_1		A_2	A_3		
					s	T	s	T	T	x	T_{min}	T_{max}
0.25	0.25	0.25	0.25	32	22	102.93	21	96.33	64.0	4	72.0	96.0
0.1	0.4	0.4	0.1	32	18	40.9	16	38.29	24.08	4	26.94	35.51
0	0.1	0.6	0.3	32	9	70.41	8	65.77	40.63	4	45.63	60.63
0.2	0.6	0.1	0.1	32	27	48.63	26	45.6	28.06	4	31.39	41.39
0	0.7	0.3	0	32	21	29.5	20	27.7	17.0	4	18.99	25.0
0	0.3	0.7	0	32	11	29.5	10	27.7	17.0	4	19.0	25.0
0.1	0.3	0.2	0.4	32	26	232.42	24	215.1	130.9	4	150.58	210.15
0.1	0.2	0.3	0.4	32	22	217.51	21	201.28	130.9	4	150.58	210.15
0.5	0.2	0.2	0.1	32	19	92.45	18	86.55	55.56	4	62.22	82.22

представления меньше минимальной оценки среднего времени T_{min} для страничного представления, оптимальный размер страницы $x = 4$ удовлетворяет неравенству $2 < x < m/6$.

Сравнивая среднее время работы для последовательного и страничного представления, видно, что для некоторых наборов вероятностей среднее время для последовательного представления меньше оценки сверху T_{max} для страничного представления. Тем не менее, видно, что для некоторых наборов вероятностей среднее время для

Таблица 3. Среднее время работы с очередью по приоритетам для $m = 60$ при длине информационной части $L > 1$

r	p_1	p_2	q	m	$L=2$			$L=3$		
					A_1		A_2	A_1		A_2
					s	T	T	s	T	T
0.25	0.25	0.25	0.25	60	21	96.33	80.0	13	62.99	60.0
0.1	0.4	0.4	0.1	60	16	38.29	29.8	11	25.4	22.65
0	0.1	0.6	0.3	60	8	65.77	50.63	5	43.02	38.13
0.2	0.6	0.1	0.1	60	26	45.6	34.72	17	30.47	26.39
0	0.7	0.3	0	60	20	27.69	21.0	13	18.4	16.0
0	0.3	0.7	0	60	10	27.69	21.0	7	18.4	16.0
0.1	0.3	0.2	0.4	60	24	215.1	170.37	15	128.69	121.13
0.1	0.2	0.3	0.4	60	21	201.3	170.37	13	121.58	121.13
0.5	0.2	0.2	0.1	60	18	86.55	68.89	12	57.09	52.22

Таблица 4. Среднее время работы с очередью по приоритетам для $m = 60$ при длине информационной части $L > 1$

r	p_1	p_2	q	m	$L=4$			$L=5$		
					A_1		A_2	A_1		A_2
					s	T	T	s	T	T
0.25	0.25	0.25	0.25	60	10	46.37	48.0	7	36.73	40.0
0.1	0.4	0.4	0.1	60	8	19.05	18.37	6	15.2	15.51
0	0.1	0.6	0.3	60	4	32.14	30.63	3	25.55	25.63
0.2	0.6	0.1	0.1	60	12	22.79	21.39	10	18.52	18.06
0	0.7	0.3	0	60	10	13.87	13.0	8	11.17	11.0
0	0.3	0.7	0	60	5	13.87	13.0	4	11.17	11.0
0.1	0.3	0.2	0.4	60	11	88.04	92.2	8	64.66	73.44
0.1	0.2	0.3	0.4	60	9	84.22	92.2	7	63.0	73.44
0.5	0.2	0.2	0.1	60	9	42.62	42.22	7	34.23	35.56

последовательного представления больше, чем максимальная оценка T_{max} для страничного представления.

Сравнивая связанное и последовательное представление, видно, что среднее время для связанного представления всегда меньше среднего времени для последовательного представления.

Сравнивая результаты в случае, когда длина информационной части равна $L > 1$, видно, что с ростом L среднее время для связанного представления становится больше среднего времени для последовательного представления. Например, при $L = 5$ среднее время для связанного представления в основном больше среднего времени для последовательного представления.

Литература

1. Седжвик Р. Фундаментальные алгоритмы на C++. Анализ /Структуры данных /Сортировка/ Поиск: Пер. с англ./ Роберт Седжвик. К.: Издательство "ДиаСофт", 2001.
2. Боллапаргада В., Мэрфи К., Расс У. Структура операционной системы Cisco IOS. М.: Вильямс, 2002, 163–189.
3. Кемени Дж., Снелл Дж. Конечные цепи Маркова. М.: Наука, 1970.