УДК 519.872

ДИСЦИПЛИНЫ ОБСЛУЖИВАНИЯ НА ОСНОВЕ МАТРИЦЫ ПРИОРИТЕТОВ

Т.И. Алиев[a], Э. Махаревс[b]

[a] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, aliev@cs.ifmo.ru
[b] Балтийская международная академия, Рига, LV-1019, Латвия

Аннотация. Рассматриваются дисциплины обслуживания заявок общего вида в системах массового обслуживания с неоднородной нагрузкой. Для математического описания таких дисциплин предлагается использовать матрицу приоритетов, отображающую вид приоритета (относительный, абсолютный или его отсутствие) между двумя любыми классами заявок. Такой способ описания, обладая наглядностью и простотой задания приоритетов, позволяет получить математические зависимости характеристик функционирования системы от параметров. Сформулированы требования к формированию матрицы приоритетов, введено понятие канонической матрицы приоритетов. Показано, что не всякая матрица, построенная в соответствии с этими требованиями, является корректной. Понятие некорректности матрицы приоритетов проиллюстрировано на примере; показано, что такие матрицы не обеспечивают однозначности и определенности при разработке алгоритма, реализующего соответствующие им дисциплины обслуживания. Для канонических матриц приоритетов сформулированы правила построения корректных матриц. В качестве одной из основных характеристик рассматривается время пребывания в системе заявок разных классов, которое складывается из времени ожидания начала обслуживания и времени нахождения заявки на обработке. Для этих характеристик с использованием метода введения дополнительного события получены преобразования Лапласа, на основе которых выведены математические зависимости для расчета двух первых начальных моментов соответствующих характеристик обслуживания заявок.

Ключевые слова: система обслуживания, дисциплина обслуживания, смешанные приоритеты, матрица приоритетов, канонические матрицы приоритетов, корректные и некорректные матрицы приоритетов.

QUEUEING DISCIPLINES BASED ON PRIORITY MATRIX

T.I. Aliev[a], E. Maharevs[b]

[a] ITMO University, Saint Petersburg, 197101, Russian Federation, aliev@cs.ifmo.ru
[b] Baltic International Academy, Riga, LV-1019, Latvia

Abstract. The paper deals with queueing disciplines for demands of general type in queueing systems with multivendor load. A priority matrix is proposed to be used for the purpose of mathematical description of such disciplines, which represents the priority type (preemptive priority, not preemptive priority or no priority) between any two demands classes. Having an intuitive and simple way of priority assignment, such description gives mathematical dependencies of system operation characteristics on its parameters. Requirements for priority matrix construction are formulated and the notion of canonical priority matrix is given. It is shown that not every matrix, constructed in accordance with such requirements, is correct. The notion of incorrect priority matrix is illustrated by an example, and it is shown that such matrixes do not ensure any unambiguousness and determinacy in design of algorithm, which realizes corresponding queueing discipline. Rules governing construction of correct matrixes are given for canonical priority matrixes. Residence time for demands of different classes in system, which is the sum of waiting time and service time, is considered as one of the most important characteristics. By introducing extra event method Laplace transforms for these characteristics are obtained, and mathematical dependencies are derived on their basis for calculation of two first moments for corresponding characteristics of demands queueing.

Keywords: queueing system, queueing discipline, mixed priorities, priority matrix, canonical priority matrix, correct and incorrect priority matrix.

Введение

Системы массового обслуживания (СМО) с неоднородным потоком заявок широко применяются в качестве моделей систем различного назначения, включая вычислительные системы и компьютерные сети [1–5]. Качество функционирования таких систем определяется значениями характеристик обслуживания заявок, таких как время пребывания и ожидания заявок, число заявок в системе. Для обеспечения требуемого качества функционирования используются различные стратегии управления поступающими в систему потоками заявок, задаваемые в виде дисциплин обслуживания (ДО). Важное место среди них занимают приоритетные дисциплины, в частности, с относительными (ОП) или с абсолютными (АП) приоритетами [6–8].

Дисциплины обслуживания заявок с одним классом приоритетов (ОП или АП) не всегда позволяют достичь требуемого качества функционирования системы. Исходя из этого, в реальных системах часто используются ДО общего вида, в которых один и тот же класс заявок может иметь АП по отношению к одной группе классов заявок, ОП – к другой группе и не иметь приоритета к остальным классам заявок (обслуживание без приоритетов в порядке поступления, БП). Такие ДО относятся к дисциплинам со смешанными приоритетами (ДО СП) [9–12], которые при наличии ограничений на время пребывания заявок в системе позволяют наилучшим образом обеспечить выполнение этих ограничений.

При построении математических зависимостей характеристик обслуживания заявок от параметров системы при использовании ДО СП возникает задача математического описания таких дисциплин. Ранее авторами для описания ДО СП было предложено использовать матрицу приоритетов (МП), отражающую вид приоритета между двумя любыми классами заявок [10]. В отличие от других подходов к описанию ДО с несколькими уровнями приоритетов [11, 12], рассматриваемое матричное представление позволяет назначать приоритеты классам заявок в произвольном порядке, что расширяет множество возможных ДО и охватывает большое количество различных ДО даже при небольшом числе классов заявок. Отметим, что математические зависимости для расчета характеристик функционирования СМО в [10] получены только для средних значений, в частности, для среднего времени ожидания заявок разных классов.

Ниже сформулированы требования к построению МП и представлены аналитические зависимости характеристик обслуживания заявок разных классов от параметров системы на уровне преобразований Лапласа для непрерывных величин и производящих функций для дискретных величин.

Матрица приоритетов

МП представляет собой квадратную матрицу $\mathbf{Q}=[q_{ij}\ (i,j=1,\ldots,H)]$, размерность которой определяется числом классов заявок H, поступающих в систему. Элемент q_{ij} матрицы задает приоритет заявок класса i (i-заявок) по отношению к заявкам класса j (j-заявкам) и может принимать следующие значения: 0 – нет приоритета, 1 – приоритет относительный и 2 – приоритет абсолютный.

С помощью МП можно описать большое множество ДО, в том числе с одним классом приоритетов. Так, например, в случае четырех классов заявок $(H=4)$ матрицы, соответствующие традиционным ДО ОП, ДО АП и произвольной ДО СП, будут иметь следующий вид:

$$\mathbf{Q}^{\text{ОП}}=\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & 1 & 1 \\ 2 & 0 & 0 & 1 & 1 \\ 3 & 0 & 0 & 0 & 1 \\ 4 & 0 & 0 & 0 & 0 \end{array};\quad \mathbf{Q}^{\text{АП}}=\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 2 & 2 & 2 \\ 2 & 0 & 0 & 2 & 2 \\ 3 & 0 & 0 & 0 & 2 \\ 4 & 0 & 0 & 0 & 0 \end{array};\quad \mathbf{Q}^{\text{СП}}=\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & 0 & 1 \\ 2 & 0 & 0 & 0 & 0 \\ 3 & 2 & 2 & 0 & 1 \\ 4 & 0 & 1 & 0 & 0 \end{array}.$$

В отличие от традиционных ДО с одним классом приоритетов (ОП или АП), в которых обычно предполагается, что приоритеты назначены по правилу «у класса с меньшим номером – более высокий приоритет» [10], с помощью МП приоритеты могут быть назначены произвольным образом, как это показано выше для МП $\mathbf{Q}^{\text{СП}}$, где наивысший приоритет имеют заявки класса 3, а самый низкий – заявки класса 2.

Элементы МП должны удовлетворять следующим требованиям:

1. $q_{ii}=0\ \ (i=1,\ldots,H)$;
2. если $q_{ij}=1$ или 2, то $q_{ji}=0\ \ (i,j=1,\ldots,H)$.

Матрица приоритетов называется канонической, если $q_{ij}=0$ для всех $i\geq j\ \ (i,j=1,\ldots,H)$. Канонические МП описывают дисциплины, в которых заявки классов с меньшим номером имеют приоритет не ниже, чем заявки классов с бо́льшим номером.

Число вариантов заполнения канонической МП $\xi=3^{H(H-1)/2}$, где H – число классов заявок, определяющее размерность матрицы. Однако из этого числа должны быть исключены так называемые некорректные матрицы приоритетов.

Корректность МП предполагает однозначность и определенность алгоритма, реализующего соответствующую МП. Не всякая матрица приоритетов, удовлетворяющая перечисленным выше требованиям, является корректной. МП является некорректной, если при ее реализации может возникнуть неоднозначная ситуация, причем любое принятое решение будет противоречить заданной МП.

Проиллюстрируем понятие некорректности на примере МП для трех классов заявок $(H=3)$ вида

$$\mathbf{Q}=\begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & 0 & 1 & 1 \\ 2 & 0 & 0 & 2 \\ 3 & 0 & 0 & 0 \end{array}.$$

Рассмотрим следующую ситуацию. Пусть в момент поступления в систему заявки класса 2 в приборе обслуживается заявка класса 3, а в очереди находится одна или несколько заявок класса 1, которые не могут прервать обслуживание заявки класса 3, так как в соответствии с заданной МП они имеют только ОП по отношению к заявкам класса 3 ($q_{13}=1$). При этом возникает следующая неопределенность. С одной стороны, поступившая заявка класса 2 должна прервать обслуживание заявки класса 3, поскольку по отношению к заявкам класса 3 имеет АП $(q_{23}=2)$. С другой стороны, поступившая заявка класса 2 не может начать обслуживаться раньше заявок класса 1, находящихся в очереди и имеющих ОП по отношению к заявкам класса 2 ($q_{12}=1$). Эта неопределенность приводит к неоднозначности при построении алгоритма, реализующего данную ДО. Любое решение из двух возможных (прервать и не прервать обслуживание заявки класса 3) приведет к дисциплине, не соответствующей заданной МП. Такие МП в дальнейшем исключим из рассмотрения, относя их к некорректным. При этом резко уменьшается число возможных вариантов заполнения МП (табл. 1).

Размерность МП	2×2	3×3	4×4	5×5	6×6
Число вариантов ξ заполнения канонических МП	3	27	729	59049	>14 млн
Число корректных канонических МП	3	13	75	541	4683
Оценка числа корректных канонических МП $\tilde{\xi}_H$	3	13	75	537	4644

Таблица 1. Число матриц приоритетов

Отметим, что с ростом числа классов заявок число различных ДО растет экспоненциально.

Для приближенной оценки числа корректных канонических МП при $H > 6$ экспериментальным путем с использованием метода математической индукции получено следующее выражение: $\tilde{\xi}_H \approx 1,5 \times 1,44^{H-2} H!\ \ (H \ge 2)$. Это выражение, как видно из табл. 1, дает нижнюю оценку числа корректных канонических МП. Заметим, что $\tilde{\xi}_H$ определяет только число канонических МП. Если допустить произвольные варианты заполнения МП, то число допустимых вариантов заполнения МП возрастает примерно еще в $H!$ раз.

Поскольку не всякий вариант заполнения МП является корректным, необходимо сформулировать правила, позволяющие формировать только корректные МП.

1. правило строки. После ненулевого элемента в строке не должен быть ноль, т.е. если $q_{ij} = 1$ или 2, то $q_{ik} \ne 0$ для всех $k > j\ \ (i, j, k = 1, \dots, H)$;
2. правило столбца. Элементы в пределах одного столбца должны образовывать невозрастающую последовательность: $q_{i+1\,j} \le q_{ij}\ \ (i = 1, \dots, H-1;\ \ j = 1, \dots, H)$;
3. правило БП-группы. Классы заявок, образующие БП-группу, должны иметь одинаковые приоритеты по отношению к остальным классам заявок, т.е. если $q_{i\,i+1} = 0$, то $q_{i+1\,j} = q_{ij}$ для всех $j = 1, \dots, H$.

Если не выполняется хотя бы одно из перечисленных правил, то матрица приоритетов будет некорректной. В случае неканонической МП для ее проверки на соответствие перечисленным правилам необходимо преобразовать исходную матрицу в канонический вид путем перестановки строк и столбцов в соответствии с уровнем приоритетности, который подсчитывается как сумма всех элементов исходной матрицы в пределах каждой строки. Чем больше полученное значение, тем выше уровень приоритетности у соответствующего класса заявок.

Характеристики обслуживания заявок

Рассмотрим характеристики одноканальной СМО, в которую поступают H классов заявок, образующих простейшие потоки с интенсивностями $\lambda_1, \dots, \lambda_H$. Длительность τ_{b_k} обслуживания заявок класса k распределена по произвольному закону с плотностью распределения вероятностей $b_k(\tau)$. Выбор заявок из очереди на обслуживание осуществляется в соответствии с ДО СП, заданной с помощью МП.

В качестве основной характеристики, описывающей эффективность функционирования системы, будем рассматривать время пребывания в системе τ_{u_k} заявок класса $k = 1, \dots, H$, которое складывается из времени ожидания начала обслуживания τ_{x_k} и времени нахождения заявки на обработке τ_{v_k}, включающего в себя время ожидания заявки в прерванном состоянии:

$$\tau_{u_k} = \tau_{x_k} + \tau_{v_k}\,. \qquad (1)$$

Ниже при записи выражений для определения характеристик обслуживания заявок используются следующие обозначения.

1. $F_k(\tau) = \Pr(\tau_{f_k} < \tau)$ – функция распределения непрерывной случайной величины $\tau_{f_k} > 0$, причем f_k – один из следующего набора символов: b_k, x_k, z_k, v_k, w_k, u_k.
2. $f_k(\tau) = F_k'(\tau)$ – плотность распределения случайной величины τ_{f_k}.
3. $f_k^{(n)} = \int_0^\infty \tau^n f_k(\tau)\, d\tau$ – начальный момент порядка $n\ \ (n = 1, 2, \dots)$ случайной величины τ_{f_k}, причем для простоты при записи первого начального момента (математического ожидания) верхний индекс будем опускать: $f_k = f_k^{(1)}$.

4. $F_k^*(s) = \int_0^\infty e^{-s\tau} f_k(\tau)\, d\tau \quad (s > 0)$ – преобразование Лапласа плотности $f_k(\tau)$.
5. $M_k^*(z) = \sum_{m=0}^{\infty} z^m P_k(m)$ – производящая функция дискретной случайной величины – числа заявок класса k, находящихся в системе, где $P_k(m)$ – вероятность того, что в системе находится m заявок класса k.
6. $R = \sum_{i=1}^{H} \rho_i$ – суммарная загрузка системы, где $\rho_i = \lambda_i b_i$ – загрузка, создаваемая i-заявками, причем предполагается, что $R < 1$, т.е. система работает без перегрузок.
7. $r_g(i,k)$ – коэффициенты, принимающие значения 0 или 1 в зависимости от значений элементов q_{ik} и q_{ki} матрицы приоритетов и позволяющие выделить классы заявок i и k, имеющие между собой один и тот же вид приоритета (ОП, АП, БП или любое их сочетание):

- $r_0(i,k) = 1$, если между заявками классов i и k нет приоритетов;
- $r_1(i,k) = 1$, если i-заявки имеют ОП по отношению к k-заявкам;
- $r_2(i,k) = 1$, если i-заявки имеют АП по отношению к k-заявкам.

Формулы для расчета коэффициентов $r_g(i,k)$ и их значения приведены в табл. 2.

Коэффициенты $r_0(i,k)$, $r_1(i,k)$ и $r_2(i,k)$ являются основными. На их основе формируются дополнительные коэффициенты $r_3(i,k)$, $r_4(i,k)$, $r_5(i,k)$ и $r_6(i,k)$, представляющие собой комбинацию основных коэффициентов. Отметим, что на значения коэффициентов оказывает влияние местоположение номеров классов в круглых скобках, и в общем случае $r_g(i,k) \neq r_g(k,i)$.

Коэффициенты $r_g(i,k)$		Значения элементов МП				
	q_{ik}	0	0	0	1	2
	q_{ki}	0	1	2	0	0
$r_0(i,k) = 0,5(1 - q_{ik} - q_{ki})(2 - q_{ik} - q_{ki})$		1	0	0	0	0
$r_1(i,k) = q_{ik}(2 - q_{ik})$		0	0	0	1	0
$r_2(i,k) = 0,5\, q_{ik}(q_{ik} - 1)$		0	0	0	0	1
$r_3(i,k) = r_0(i,k) + r_1(i,k)$		1	0	0	1	0
$r_4(i,k) = r_1(i,k) + r_2(i,k)$		0	0	0	1	1
$r_5(i,k) = r_0(i,k) + r_4(i,k)$		1	0	0	1	1
$r_6(i,k) = r_5(i,k) + r_1(k,i)$		1	1	0	1	1

Таблица 2. Значения коэффициентов $r_g(i,k)$

8. $\Lambda_k^{(g)} = \sum_{i=1}^{H} r_g(i,k)\lambda_i$ – частичные суммарные интенсивности потоков заявок.
9. $R_k^{(g)} = \sum_{i=1}^{H} r_g(i,k)\rho_i$ – частичные суммарные загрузки.

Величины $\Lambda_k^{(g)}$ и $R_k^{(g)}$ имеют также простое физическое толкование:

- $\Lambda_k^{(0)}$ и $R_k^{(0)}$ представляют собой суммарную интенсивность потоков и суммарную загрузку, создаваемые заявками всех классов, имеющих такой же приоритет, что и k-заявки, т.е. заявками тех классов, которые вместе с заявками класса k образуют БП-группу;
- $\Lambda_k^{(1)}$, $R_k^{(1)}$ и $\Lambda_k^{(2)}$, $R_k^{(2)}$ – суммарные интенсивности и загрузки, создаваемые заявками всех классов, которые обладают по отношению к k-заявкам более высоким ОП и АП соответственно, и т.д.

Для одноканальной системы с ДО СП, в которую поступают простейшие потоки заявок разных классов, аналитические зависимости для расчета характеристик функционирования системы получены с использованием метода введения дополнительного события [13] на уровне преобразований Лапласа и производящих функций.

Из (1) следует, что плотность распределения времени пребывания $u_k(\tau)$ заявок класса k представляет собой свертку плотностей $x_k(\tau)$ и $v_k(\tau)$. Тогда преобразование Лапласа плотности распределения $u_k(\tau)$: $U_k^*(s) = X_k^*(s)V_k^*(s)$.

Преобразование Лапласа плотности распределения $x_k(\tau)$ времени ожидания начала обслуживания имеет вид

$$X_k^*(s) = \frac{R_k^{(6)}\sigma_k + \sum_{i=1}^{H} r_1(k,i)\lambda_i[1 - B_k^*(\sigma_k)]}{s - \Lambda_k^{(0)} + \sum_{i=1}^{H} r_0(i,k)\lambda_i\, B_i^*(\sigma_k)]}, \qquad (2)$$

где $\sigma_k = s + \Lambda_k^{(4)} - \Lambda_k^{(4)} D_k^*(s)$, а $D_k^*(s)$ определяется из уравнения:

$$\Lambda_k^{(4)} D_k^*(s) = \sum_{i=1}^{H} r_4(i,k)\lambda_i B_i^*(s + \Lambda_k^{(4)} - \Lambda_k^{(4)} D_k^*(s)). \qquad (3)$$

Преобразование Лапласа плотности распределения $v_k(\tau)$ времени нахождения k-заявки на обработке

$$V_k^*(s) = B_k^*(s + \Lambda_k^{(2)} - \Lambda_k^{(2)} C_k^*(s)), \qquad (4)$$

где $C_k^*(s)$ определяется из уравнения

$$\Lambda_k^{(2)} C_k^*(s) = \sum_{i=1}^{H} r_2(i,k)\lambda_i B_i^*(s + \Lambda_k^{(2)} - \Lambda_k^{(2)} C_k^*(s)). \qquad (5)$$

В выражениях (2)–(5) используются следующие обозначения:

– $C_k^*(s)$ – преобразование Лапласа плотности распределения $c_k(\tau)$ периода занятости τ_{c_k} прибора заявками с более высоким АП, чем рассматриваемая k-заявка:

$$C_k^*(s) = \int_0^\infty e^{-s\tau} c_k(\tau)\, d\tau; \qquad (6)$$

– $D_k^*(s)$ – преобразование Лапласа плотности распределения $d_k(\tau)$ периода занятости τ_{d_k} прибора заявками с более высоким ОП и АП, чем рассматриваемая k-заявка:

$$D_k^*(s) = \int_0^\infty e^{-s\tau} d_k(\tau)\, d\tau. \qquad (7)$$

Выражения (2)–(7) позволяют определить все основные характеристики обслуживания заявок в системе, в частности, их начальные моменты.

Из (1) следует, что математическое ожидание u_k и второй начальный момент $u_k^{(2)}$ времени пребывания в системе k-заявок ($k = \overline{1,H}$) равны соответственно

$$u_k = x_k + v_k; \qquad u_k^{(2)} = x_k^{(2)} + 2x_k v_k + v_k^{(2)}.$$

Значения x_k, $x_k^{(2)}$ и v_k, $v_k^{(2)}$ определяются путем дифференцирования соответствующих выражений (2) и (4) в точке $s = 0$:

$$\left.\begin{aligned}
& x_k = \frac{\sum_{i=1}^{H} r_6(i,k)\lambda_i\, b_i^{(2)}}{2(1-R_k^{(4)})(1-R_k^{(5)})}; \qquad v_k = \frac{b_k}{1-R_k^{(2)}};\\
& x_k^{(2)} = \frac{\sum_{i=1}^{H} r_6(i,k)\lambda_i\, b_i^{(3)}}{3(1-R_k^{(4)})^2(1-R_k^{(5)})} + \frac{\sum_{i=1}^{H} r_5(i,k)\lambda_i\, b_i^{(2)} \sum_{i=1}^{H} r_6(i,k)\lambda_i\, b_i^{(2)}}{2(1-R_k^{(4)})^2(1-R_k^{(5)})^2} +\\
& + \frac{\sum_{i=1}^{H} r_4(i,k)\lambda_i\, b_i^{(2)} \sum_{i=1}^{H} r_6(i,k)\lambda_i\, b_i^{(2)}}{2(1-R_k^{(4)})^3(1-R_k^{(5)})};\\
& v_k^{(2)} = \frac{b_k^{(2)}}{(1-R_k^{(2)})^2} + \frac{\sum_{i=1}^{H} r_2(i,k)\lambda_i\, b_i^{(2)}}{(1-R_k^{(2)})^3},
\end{aligned}\right\} \qquad (8)$$

где $b_i^{(n)}$ – n-й начальный момент длительности обслуживания i-заявок ($i = \overline{1,H}$; $n = 1, 2, \ldots$).

На основе выражений (8) могут быть вычислены дисперсии и коэффициенты вариации соответствующих временных характеристик. При необходимости по двум моментам можно выполнить аппроксимацию вероятностных распределений [14].

Выражения (1)–(8) позволяют определить время ожидания в прерванном состоянии и полное время ожидания заявок каждого класса. В частности, их средние значения соответственно равны $(k=\overline{1,H})$

$$\left.\begin{aligned} z_k &= v_k - b_k = \frac{R_k^{(2)} b_k}{1-R_k^{(2)}};\\ w_k &= x_k + z_k = \frac{\sum_{i=1}^{H} r_6(i,k)\lambda_i b_i^{(2)}}{2(1-R_k^{(4)})(1-R_k^{(5)})} + \frac{R_k^{(2)} b_k}{1-R_k^{(2)}}. \end{aligned}\right\} \qquad (9)$$

Характеристики обслуживания заявок ДО с одним классом приоритетов (ОП и АП) и ДО БП, как частные случаи ДО СП, могут быть получены на основе выражений (2)–(9).

При решении задач синтеза реальных систем, связанных, например, с оценкой емкости накопителя (буферной памяти), требуется определять число заявок, находящихся в системе.

Производящая функция числа заявок класса $k=\overline{1,H}$, находящихся в системе, связана с преобразованием Лапласа времени пребывания следующей зависимостью [15]:

$$M_k^*(z) = U_k^*(\lambda_k - \lambda_k z) \qquad (k=\overline{1,H}). \qquad (10)$$

Дифференцируя соответствующее число раз выражение (10) по z в точке $z = 1$, получим зависимости, связывающие начальные моменты распределений числа заявок и времени их пребывания в системе, в частности, для двух первых моментов:

$$m_k = \lambda_k u_k; \qquad m_k^{(2)} = \lambda_k^2 u_k^{(2)} + m_k \qquad (k=\overline{1,H}). \qquad (11)$$

Первое выражение в (11) представляет собой формулу Литтла [15], связывающую среднее число заявок в системе со средним временем их пребывания.

Заключение

Рассмотренные дисциплины обслуживания, построенные на основе матрицы приоритетов, обобщают традиционные дисциплины с одним классом приоритетов и расширяют множество возможных дисциплин обслуживания. Наибольший эффект от применения этих дисциплин состоит в возможности наилучшим образом обеспечить требуемое качество функционирования системы при наличии ограничений на времена пребывания в системе заявок разных классов. При этом задача сводится к определению значений элементов матрицы приоритетов, при которых выполняются заданные ограничения.

Литература

1. Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы, технологии, протоколы: Учебник для вузов. 3-е изд. СПб.: Питер, 2006. 944 с.
2. Aliev T.I., Nikulsky I.Y., Pyattaev V.O. Modeling of packet switching network with relative prioritization for different traffic types // Proc. 10th International Conference on Advanced Communication Technology (ICACT-2008). Phoenix Park, South Korea, 2008. Art. 4494220. P. 2174–2176.
3. Bogatyrev V.A. An interval signal method of dynamic interrupt handling with load balancing // Automatic Control and Computer Sciences. 2000. V. 34. N 6. P. 51–57.
4. Bogatyrev V.A. On interconnection control in redundancy of local network buses with limited availability // Engineering Simulation. 1999. V. 16. N 4. P. 463–469.
5. Вишневский В.М., Семенова О.В. Системы поллинга: теория и применение в широкополосных беспроводных сетях. М.: Техносфера, 2007. 312 с.
6. Муравьева-Витковская Л.А. Обеспечение качества обслуживания в мультисервисных компьютерных сетях за счет приоритетного управления // Изв. вузов. Приборостроение. 2012. Т. 55. № 10. С. 64–68.
7. Алиев Т.И. Проектирование систем с приоритетами // Изв. вузов. Приборостроение. 2014. Т. 57. № 4. С. 30–35.
8. Алиев Т.И., Муравьева Л.А. Система с динамически изменяющимися смешанными приоритетами и ненадежным прибором // Автоматика и телемеханика. 1988. Т. 49. № 7. С. 99–106.
9. Alfa A.S. Matrix-geometric solution of discrete time MAP/PH/1 priority queue // Naval Research Logistics. 1998. V. 45. N 1. P. 23–50.
10. Основы теории вычислительных систем / Под ред. С.А. Майорова. М.: Высшая школа, 1978. 408 с.
11. Гольдштейн Б.С. Об оптимальном приоритетном обслуживании в программном обеспечении ЭАТС. В кн. Системы управления сетями. М.: Наука, 1980. С. 73–78.
12. Zhao J.-A., Li B., Cao X.-R., Ahmad I. A matrix-analytic solution for the DBMAP/PH/1 priority queue // Queueing Systems. 2006. V. 53. N 3. P. 127–145.

13. Климов Г.П. Стохастические системы обслуживания. М.: Наука, 1966. 242 с.
14. Алиев Т.И. Аппроксимация вероятностных распределений в моделях массового обслуживания // Научно-технический вестник информационных технологий, механики и оптики. 2013. № 2 (84). С. 88–93.
15. Клейнрок Л. Вычислительные системы с очередями: Пер. с англ. М.: Мир, 1979. 600 с.

Алиев Тауфик Измайлович – доктор технических наук, профессор, заведующий кафедрой, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, aliev@cs.ifmo.ru

Махаревс Эдуардс – доктор технических наук, хабилитированный доктор инженерных наук Латвийской Республики, профессор, профессор, Балтийская международная академия, Рига, LV-1019, Латвия, eduard@rostourism.lv

Taufik I. Aliev – D.Sc., Professor, Department head, ITMO University, Saint Petersburg, 197101, Russian Federation, aliev@cs.ifmo.ru

Eduards Maharevs – D.Sc., Professor, Baltic International Academy, Riga, LV-1019, Latvia, eduard@rostourism.lv

Принято к печати 11.09.14
Accepted 11.09.14