ОЦЕНКА ХАРАКТЕРИСТИК ПРИОРИТЕТНОЙ МОДЕЛИ ЗВЕНА ПЕРЕДАЧИ ДАННЫХ МУЛЬТИСЕРОВИСНОЙ КОМПЬЮТЕРНОЙ СЕТИ

Л. А. Муравьева-Витковская (Санкт-Петербург)

Введение

Мультисервисные компьютерные сети (МКС) – специфические системы, предоставляющие услуги как компьютерных, так и телекоммуникационных сетей. Современные МКС характеризуются разнообразием предоставляемых пользователям услуг, увеличением числа пользователей и объемов передаваемых данных, повышением уровня требований к качеству обслуживания пользователей, неоднородностью трафика. Неоднородность трафика заключается в передаче по МКС пакетов нескольких типов (видео- и аудио-, речевых и текстовых пакетов и т.д.), к доставке которых предъявляются различные требования. Один из способов распределения сетевых ресурсов – распределение в соответствии с существующими на данный момент приоритетами. Поэтому в качестве модели звена передачи данных, состоящего из маршрутизатора и каналов связи, используется разомкнутая приоритетная сеть массового обслуживания, первый узел которой соответствует маршрутизатору, узлы с номерами 2 – n соответствуют каналам связи (рис. 1).

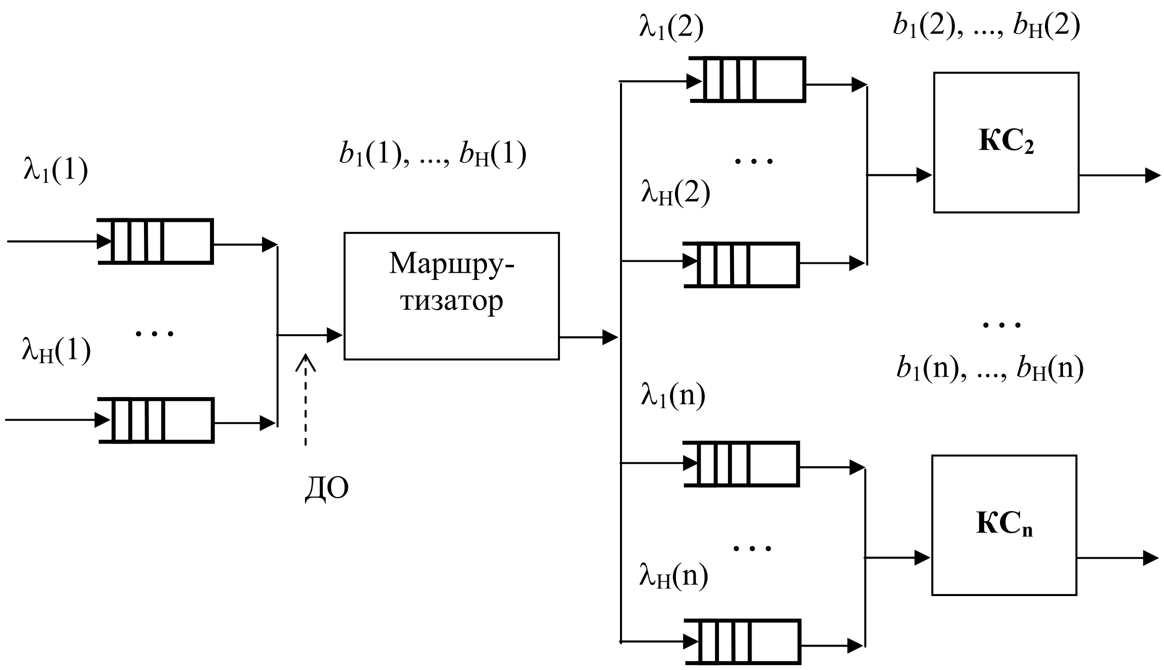

Рис. 1. Модель звена передачи данных

Неоднородность трафика, а также наличие приоритетов между пакетами различных типов резко усложняют аналитический расчет характеристик функционирования МКС на сетевых моделях. Несмотря на то, что в настоящее время имеются некоторые результаты по расчету приоритетных сетевых моделей с простой конфигурацией и частными дисциплинами обслуживания (ДО) [1], задача аналитического расчета сети с произвольной конфигурацией и ДО общего вида не получила своего решения.

Модель звена передачи данных представлена в виде разомкнутой сетевой модели, в узлах которой обслуживание заявок различных классов выполняется в соответствии с ДО общего вида из класса дисциплин обслуживания со смешанными приоритетами (ДО СП). При этом характеристики обслуживания заявок, такие как время ожидания и пребывания, число заявок в очереди, рассчитываются на уровне средних значений с учетом двух первых моментов распределений интервалов между поступающими в узлы заявками и длительностей обслуживания заявок в узлах.

Метод расчета излагается применительно к разомкнутой сетевой модели центрального обслуживания, которая наиболее подходит для отображения работы маршрутизатора с каналами связи. Однако данный метод может использоваться для расчета характеристик сетей любой конфигурации. При этом лишь увеличивается размерность задачи и становится более громоздким изложение материала.

Постановка задачи

Рассматривается модель функционирования фрагмента МКС в виде разомкнутой неоднородной сети массового обслуживания (СеМО), содержащей n одноканальных узлов.

Узел 1 сетевой модели называется центральным и отображает задержки, возникающие при маршрутизации пакетов в приоритетном маршрутизаторе, находящемся в данном узле передачи данных. Маршрутизатор в соответствии с заданным алгоритмом маршрутизации направляет пакет в один из каналов связи, выходящий из данного узла передачи. Каналы связи отображаются в модели периферийными узлами 2 – n. Задержки, возникающие в этих узлах, соответствуют времени передачи пакета по соответствующему каналу. В СеМО циркулируют H классов заявок, отображающих пакеты разных типов, передающиеся по МКС.

Заявки класса h, поступающие в систему из внешней среды, обозначенной в модели как нулевой узел 0_h, образуют случайный поток с интенсивностью $\lambda_h(0)$ и коэффициентом вариации $\alpha_h(0)$ интервалов времени между последовательно поступающими заявками. Длительность обслуживания заявок класса h в узле j описывается произвольным законом распределения с математическим ожиданием $b_h(j)$ и коэффициентом вариации $\beta_h(j)$ $(j = 0,1,...,n;\ h = 1,...,H)$.

Обслуживание заявок в центральном узле выполняется в соответствии с некоторой ДО общего вида из множества дисциплин со смешанными приоритетами.

Маршруты заявок класса h в сети описываются матрицей вероятностей передач $P_h = [p_h(i,j)]$, где $p_h(i,j)$ – вероятность того, что после обслуживания в узле i заявка класса h попадает в узел j $(i,j = 0,1,...,n;\ h = 1,...,H)$. Заявки, покидающие сеть и отображающие завершение передачи пакета по каналу связи, возвращаются во внешнюю среду, обозначенную как узлы 0_1, …, 0_h. Для рассматриваемой модели центрального обслуживания для узлов i = 0, 2, 3, n вероятности передач равны: $p_h(i,1) = 1$ и $p_h(i,j) = 0$ для всех $j \neq 1$. Вероятности передач $p_h(1,0)$, $p_h(1,2)$, …, $p_h(1,n)$ для узла 1 в общем случае отличаются от нуля и единицы, причем $\sum_{j=0}^{n} p_h(1,j) = 1 \quad (h = 1,...,H)$.

Данное условие означает, что заявки в сети не теряются и не размножаются, т.е. сетевая модель является линейной.

Задача оценки качества функционирования МКС заключается в определении характеристик сетевой модели, основными среди которых являются средние значения времени ожидания и пребывания заявок в сети. Остальные характеристики (длины очередей, число одновременно находящихся в сети заявок) могут быть получены на основе фундаментальных соотношений, справедливых для сетевых моделей общего вида. Указанные характеристики рассчитываются для отдельных классов заявок и для заявок объединенного потока как в каждом узле сети (узловые характеристики), так и для сети в целом (сетевые характеристики).

Расчет характеристик сетевой модели

Метод расчета разомкнутой СеМО с приоритетами основан на декомпозиции сети, т.е. представлении ее в виде отдельных независимых узлов с неоднородным потоком заявок различных классов, между которыми могут быть установлены приоритеты. Для этого в сетевой модели выделяются узлы композиции *К*, в которых происходит объединение нескольких потоков заявок в один, и узлы декомпозиции *Д*, в которых заявки одного потока разветвляются по нескольким направлениям.

Рассмотрим некоторый выделенный узел $j = 1, \ldots, n$ в виде СМО $G_H/G_H/1$ (рис. 2), в который поступает в общем случае H классов заявок с интенсивностями $\lambda_h(j)$ и коэффициентами вариации интервалов между заявками $\alpha_h(j)$. Длительность обслуживания заявок класса h описывается средним значением $b_h(j)$ и коэффициентом вариации $\beta_h(j)$. На выходе узла заявки класса h образуют поток, интенсивность которого равна интенсивности поступления $\lambda_h(j)$, а коэффициент вариации равен $\gamma_h(j) \quad (h = 1,\ldots,H)$.

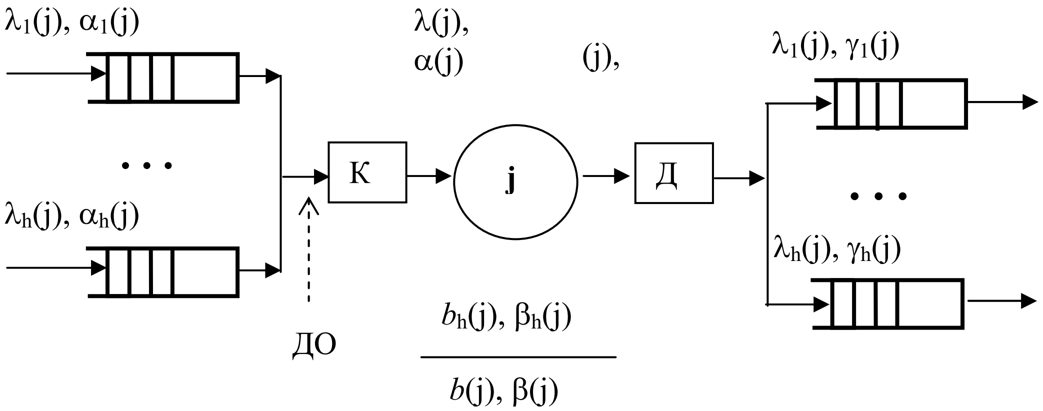

Рис. 2. Модель выделенного узла приоритетной сетевой модели

Интенсивности потоков заявок класса h в узлах линейной сети рассчитываются как $\lambda_h(j) = \xi_h(j)\,\lambda_h(0)$, где $\xi_h(j)$ – коэффициент пропорциональности, называемый коэффициентом передачи и показывающий среднее число попаданий заявки класса h в узел j за время пребывания в сети, определяется из следующей системы линейных алгебраических уравнений [2]:

$$\xi_h(j) = \sum_{i=0}^{n} p_h(i,j)\,\xi_h(i) \quad (j = 0,1,\ldots,n).$$

Для рассматриваемой модели центрального обслуживания $\xi_p(1) = 1/p_h(1,0);\quad \xi_h(j) = p_h(1,j)/p_h(1,0)$ для $j = 2, \ldots, n$.

Процессы поступления и обслуживания заявок в выделенном узле $j = 1, \ldots, n$ описываются следующими выражениями:

• интенсивность и коэффициент вариации объединенного потока заявок, поступающих в узел:

$$\lambda(j) = \sum_{h=1}^{H} \lambda_h(j); \quad \alpha^2(j) = \sum_{h=1}^{H} \pi_h(j)\,\alpha_h^2(j); \tag{1}$$

• среднее значение и коэффициент вариации длительности обслуживания заявок объединенного потока в узле:

$$b(j) = \sum_{h=1}^{H} \pi_h(j)\,b_h(j); \tag{2}$$

$$\beta^2(j) = \sum_{h=1}^{H} \pi_h(j) b_h^2(j)[1+\beta_h^2(j)]/b^2(j) - 1. \qquad (3)$$

В выражениях (1)–(3) величина $\pi_h(j) = \lambda_h(j)/\lambda(j)$ характеризует вероятность того, что поступившая в узел j заявка принадлежит классу h (h = 1, …, H).

Коэффициент вариации выходящего из узла потока заявок определяется по следующей формуле [3]:

$$\gamma^2(j) = \alpha^2(j) + 2R(j)[1-R(j)]w(j)/b(j),$$

где $R(j) = \lambda(j)b(j)$ – суммарная загрузка узла; $w(j) = \sum_{h=1}^{H} \pi_h(j) w_h(j)$ – среднее время ожидания заявок объединенного потока; $w_h(j)$ – среднее время ожидания заявок класса h в узле j при заданной дисциплине обслуживания.

Коэффициенты вариации выходящего потока заявок каждого класса определяются как $\gamma_h^2(j) = 1 + \pi_h(j)[\gamma^2(j)-1] \quad (h = 1,...,H)$.

Для расчета среднего времени ожидания $w_h(j)$ в узле j заявок класса h используется приближенный метод, учитывающий неэкспоненциальный характер поступающих в узел потоков заявок. Этот метод основан на пересчете среднего времени ожидания $w_h^{П}(j)$, полученного в предположении о простейшем характере потоков заявок, через средние времена ожидания при произвольном $w(j)$ и простейшем $w^{П}(j)$ объединенном потоке заявок, полученными в предположении о бесприоритетном обслуживании заявок разных классов в данном узле:

$$w_h(j) = w(j) w_h^{П}(j)/w^{П}(j),$$

где $w_h^{П}(j)$ определяется как функция элементов матрицы приоритетов [4]:

$$w_h^{П}(j) = \frac{\sum_{i=1}^{h} r_6(i,h)[j]\lambda_i(j)\, b_i^{(2)}(j)}{2(1-R_h^{(4)}(j))(1-R_h^{(5)}(j))} + \frac{R_h^{(3)}(j) b_h^{(1)}(j)}{1-R_h^{(3)}(j)}. \qquad (4)$$

В данной формуле используются следующие элементы:

$r_1(i,h)[j] = 0.5(1 - q_{ih}[j] - q_{hi}[j])(2 - q_{ih}[j] - q_{hi}[j])$; $r_2(i,h)[j] = q_{ih}[j](2 - q_{ih}[j])$;

$r_3(i,h)[j] = 0.5\, q_{ih}[j](q_{ih}[j] - 1)$; $r_4(i,h)[j] = r_2(i,h)[j] + r_3(i,h)[j]$;

$r_5(i,h)[j] = r_1(i,h)[j] + r_4(i,h)[j]$; $r_6(i,h)[j] = r_2(h,i)[j] + r_5(i,h)[j]$;

$$R_h^{(g)}(j) = \sum_{i=1}^{h} r_g(i,h)[j]\lambda_i b_i^{(1)}(j) \quad (g = 3, 4, 5);$$

$$b_i^{(l)}(j) = \int_0^{\infty} \tau^l b_i(\tau)[j] d\tau \quad (l = 1,2,...).$$

Суммарная интенсивность определяется как $\Lambda_h^{(g)}(j) = \sum_{i=1}^{h} r_g(i,h)[j]\lambda_i(j)$.

Выражение (4) можно использовать для определения $w^{П}(j)$, если все элементы матрицы приоритетов (МП) положить равными нулю.

Среднее время ожидания $w(j)$ в случае произвольного потока заявок, когда коэффициент вариации потока отличен от единицы, вычисляется по приближенной формуле, предложенной в [5]:

$$w(j) = 0{,}5b(j)\,R(j)\,(\alpha^2 + \beta^2)\,g(\alpha)[j]/(1 - R(j)). \qquad (5)$$

В последнем выражении величина $g(\alpha)[j]$ определяется в зависимости от значения коэффициента вариации $\alpha(j)$ объединенного потока заявок, поступающих в узел j:

$$g(\alpha)[j] = \begin{cases} \exp\left\{-\dfrac{2(1-R(j))(1-\alpha^2(j))^2}{3R(j)(\alpha^2(j)+\beta^2(j))}, \quad \alpha(j) < 1; \right. \\ \exp\left\{-\dfrac{(1-R(j))(\alpha^2(j)-1)}{\alpha^2(j)+4\beta^2(j)}, \quad \alpha(j) \geq 1. \right. \end{cases} \qquad (6)$$

Выбор формулы (5) для расчета $w(j)$ обусловлен тем, что, как показали исследования, это выражение позволяет с наименьшей погрешностью по сравнению с другими приближениями получить значение среднего времени ожидания в широком диапазоне изменений коэффициента вариации $\alpha(j)$ интервалов времени между поступающими заявками и коэффициента загрузки R. Так, для $R>0{,}3$ при изменении коэффициента вариации $\alpha(j)$ в диапазоне от 0 до 2 погрешность результатов расчета по формуле (5) не превышает 15% и резко уменьшается с увеличением загрузки R. В то же время другие приближения, полученные, в том числе, на основе диффузионной аппроксимации, дают удовлетворительные результаты лишь при больших значениях загрузки: $R > 0{,}5$–$0{,}7$.

Процедура расчета сетевой модели звена передачи данных МКС базируется на методе итераций, который используется для определения коэффициентов вариации $\alpha(j)$ интервалов времени между поступающими в узел $j = 1, \ldots, n$ заявками класса $h = 1, \ldots, H$. Итерационный процесс продолжается до тех пор, пока в некотором узле, в качестве которого будем использовать центральный узел, коэффициенты вариации входящих потоков заявок, вычисленные в предыдущем и последующем шагах, не совпадают с априори заданной точностью ε.

Итерационная процедура реализуется по следующей схеме.

1. Выделяется центральный узел 1, и на основе выражений (1)–(6) рассчитываются коэффициенты вариации выходящего потока заявок $\gamma_h(1)$ в предположении, что $\alpha_h(1) = \alpha_h(0)$.

2. Определяются коэффициенты вариации $\alpha(j)$ потоков заявок, поступающих в узлы $j = 2, \ldots, n$ по формуле (1), где $\alpha_h^2(j) = 1 + p_h(1, j)[\gamma_h^2(1) - 1]$.

3. Последовательно для узлов $j = 2, \ldots, n$ рассчитываются коэффициенты вариации выходящего потока заявок $\gamma_h(j)$.

4. Рассчитываются новые значения коэффициентов вариации $\tilde{\alpha}_h(1)$ поступающих в центральный узел потоков заявок: $\tilde{\alpha}_h^2(1) = \sum_{\substack{j=0 \\ j\neq 1}}^{n} p_h(1, j)\gamma_h^2(j)$.

5. Пункты 2–4 выполняются до тех пор, пока $\left|\tilde{\alpha}_h(1) - \alpha_h(1)\right|$ для всех $h = 1, \ldots, H$ не станет меньше ε.

Для полученных значений коэффициентов вариации $\alpha_h(j)$ потоков заявок с использованием формул (4)–(6) рассчитываются средние времена ожидания $w_h(j)$ заявок класса $h = 1, \ldots, H$ в узлах $j = 1, \ldots, n$ сетевой модели.

Далее определяются узловые и сетевые характеристики по следующим формулам.

Узловые характеристики $j = 1, \ldots, n$ для заявок класса $h = 1, \ldots, H$:

• среднее время пребывания заявок $u_h(j) = w_h(j) + b_h(j)$;

• средние значения длины очереди и числа заявок в узле $l_h(j)=\lambda_h(j)w_h(j);\quad m_h(j)=\lambda_h(j)u_h(j)$.

Аналогичные характеристики для объединенного потока заявок:

$$u(j)=\sum_{h=1}^{H}\pi_h(j)u_h(j);\quad w(j)=\sum_{h=1}^{H}\pi_h(j)w_h(j);\quad l(j)=\lambda(j)w(j);\quad m(j)=\lambda(j)u(j).$$

Сетевые характеристики для заявок класса $h = 1, \ldots, H$:

• средние времена пребывания и ожидания заявок в сети $U_h=\sum_{j=1}^{n}\xi_h(j)u_h(j);\quad W_h=\sum_{j=1}^{n}\xi_h(j)w_h(j);$

• среднее число заявок, находящихся в сети и в очередях $M_h=\sum_{j=1}^{n}m_h(j);\quad L_h=\sum_{j=1}^{n}l_h(j).$

Аналогичные сетевые характеристики для объединенного потока заявок:

$$U=\sum_{h=1}^{H}\pi_h(0)U_h;\quad W=\sum_{h=1}^{H}\pi_h(0)W_h;$$

$$M=\sum_{h=1}^{H}M_h=\sum_{j=1}^{n}m(j);\quad L=\sum_{h=1}^{H}L_h=\sum_{j=1}^{n}l(j).$$

Погрешности результатов расчетов, как показало имитационное моделирование, в большинстве случаев составляют не более 20–30% для сетевых характеристик, причем с увеличением загрузки узлов точность выполняемых расчетов повышается. Наличие методической погрешности обусловлено использованием приближенных зависимостей для определения среднего времени ожидания заявок в узлах сети с неэкспоненциальными входящими потоками заявок, коэффициентов вариации в узлах композиции и выходящих из узлов обслуживания потоков заявок, а также использованием итерационной процедуры для определения коэффициентов вариации входящих потоков заявок в узлах сетевой модели.

Предложенный метод позволяет получить точные значения характеристик обслуживания заявок в следующих случаях:

• для однородной разомкнутой сетевой модели, в которую поступает простейший поток заявок, при длительностях обслуживания заявок в узлах, распределенных по экспоненциальному закону;

• для неоднородной разомкнутой сетевой модели с приоритетами, в которую поступают простейшие потоки заявок, при длительностях обслуживания заявок разных классов в пределах одного узла, распределенных по экспоненциальному закону с одним и тем же средним значением.

Пример реализации метода

Рассмотрим разомкнутую сетевую модель звена передачи данных МКС, содержащую три узла: центральный узел 1, отображающий работу приоритетного маршрутизатора, и периферийные узлы 2 и 3, отображающие каналы связи, выходящие из узла передачи данных. Пусть в МКС передаются пакеты трех классов, которым в модели соответствуют заявки трех классов, поступающие в сеть с интенсивностями $\lambda_1(0)=0{,}01c^{-1}$ и $\lambda_2(0)=\lambda_3(0)=0{,}05c^{-1}$. Заявки первого класса в процессе обработки в сети после обслуживания в узле 1 могут перейти только в узел 2, а заявки второго и третьего классов – в узлы 2 и 3, причем вероятности передач равны: $p_1(1,0)$= 0,005; $p_1(1,2)$= 0,995; $p_2(1,0)$ = 0,02; $p_2(1,0)$ = 0,02; $p_2(1,2)$ = 0,58; $p_2(1,3)$ = 0,40; $p_3(1,0)$ = 0,01; $p_3(1,2)$ = 0,09; $p_3(1,3)$ = 0,90. Средние длительности обслуживания заявок в узлах сети задаются следующими значениями: $b_1(1) = 0{,}05$ с; $b_2(1) = 0{,}2$ с; $b_3(1) = 0{,}04$ с; $b_h(2) = b_h(3) = 0{,}1$ с ($h = 1, 2, 3$).

В таблице приведены средние значения времени пребывания U_1, U_2, U_3 пакетов классов 1, 2, 3 в сети для различных значений коэффициентов вариации потоков $\alpha_h(0)$ и длительностей обслуживания $\beta_h(j)$ пакетов в узлах $j = 1, \ldots, n$ ($h = 1, 2, 3$) при использовании ДО без приоритетов (ДО БП), ДО с абсолютными приоритетами (ДО АП) и ДО СП, заданной МП вида

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

Средние значения времени пребывания (доставки) пакетов [с]

ДО	$\alpha_h(0)=1; \beta_h(j)=1$			$\alpha_h(0)=0{,}5; \beta_h(j)=0{,}5$			$\alpha_h(0)=0; \beta_h(j)=0$		
	U_1	U_2	U_3	U_1	U_2	U_3	U_1	U_2	U_3
БП	163,0	49,5	86,7	102,1	33,6	53,2	81,1	28,1	41,5
АП	43,2	34,5	169,7	37,6	25,6	98,1	35,7	22,6	73,9
СП	65,2	34,5	164,5	49,3	25,4	94,2	43,9	22,3	70,4

Рассчитанные средние значения времени пребывания в сетевой модели соответствуют средним значениям времени доставки пакетов от одного выделенного узла МКС до другого.

Заключение

Разработанный приближенный аналитический метод расчета моделей МКС, представляемых в виде приоритетной СеМО, позволяет лишь оценить значения характеристик функционирования сети, при этом результаты могут иметь большую погрешность. Оценка погрешности разработанного метода расчета с использованием имитационного моделирования показала, что погрешность характеристик в области больших нагрузок (свыше 0,75) не превышает 20%, а в области нагрузок от 0,2 до 0,5 может достигать 45%. В последнем случае для повышения достоверности результатов следует применять комбинированный метод моделирования, основанный на сочетании аналитических и имитационных методов [6].

Литература

1. **Клейнрок Л.** Вычислительные системы с очередями. М.: Мир, 1979. 600 с.
2. Основы теории вычислительных систем / **С.А. Майоров, Г.И. Новиков, Т.И. Алиев, Э.И. Махарев, Б.Д. Тимченко.** М.: Высшая школа, 1978. 408 с.
3. **Marshall K. T.** Some inequalities in queuing. Operations Research. 1968. Vol. 16. P. 651–655.
4. **Алиев Т. И.** Характеристики дисциплины обслуживания заявок со смешанными приоритетами // Изв. вузов СССР. Приборостроение. 1981. № 11. С. 36–40.
5. **Kramer W., Langebach-Belz M.** Approximate formulae for the delay in the queueing system GI/G/1. ITC-8, Melbourn, 1976. P. 121–132.
6. **Алиев Т. И., Никульский И. Е., Пяттаев В. О.** Моделирование ядра мультисервисной сети с относительной приоритезацией неоднородного трафика // Научно-технический вестник СПбГУ ИТМО, выпуск 04(62). СПб, 2009. С. 88–96.