Serdica

Bulgariacae mathematicae publicationes

Сердика

Българско математическо списание

ХАРАКТЕРИСТИКИ ВОЗНИКНОВЕНИЯ ОЧЕРЕДИ В СИСТЕМАХ ОБСЛУЖИВАНИЯ С ПУАССОНОВСКИМ ВХОДЯЩИМ ПОТОКОМ

БОЯН Н. ДИМИТРОВ, СТАЙКО И. ДЕЛЕВ,
РУМЯНА Р. РАНГЕЛОВА

Рассмотрена система обслуживания типа $M/G/1$ с прибором, ненадежным в свободном состоянии. У прибора имеется очередь, если хотя бы одна заявка ожидает обслуживачия (обслуживаемая заявка не причисляется к очереди). Изучены периоды существования очереди, число заявок, становившихся в одну очередь, число заявок, обслуженных без ожидания на данном интервале времени, периоды исчезновения очереди и некоторые другие характеристики. Параллельно излагаются результаты для систем с Пуассоновским входящим потоком (бесконечный источник) и с переменным потоком Пуассоновского типа (источник с конечным числом заявок).

1. Введение и постановка задачи. Рассмотрим произвольную систему обслуживания с конечным числом обслуживающих приборов и неограниченной очередью. Будем говорить, что в момент t в системе имеется очередь, если кроме заявок, которые обслуживаются в этот момент, имеется хотя бы еще одна, ожидающая быть обслуженной. Очередь состоит только из ожидающих заявок. Если проследим время изменения состояний очереди, отметим, что периоды ее существования чередуются с периодами ее исчезновения, причем последние не совпадают с периодами отсутствия вызовов в системе.

Обозначим через θ_i длительность i-го периода существования очереди, а через ψ_i — длительность i-го периода отсутствия очереди. Если входящий поток заявок — стационарный поток пуассоновского типа, а времена обслуживания — независимые, одинаково распределенные случайные величины, тогда все θ_i (как и все ψ_i) будут независимыми между собой и одинаково распределенными случайными величинами ($i=1, 2, \ldots$). Эти величины и некоторые связанные с ними дополнительные характеристики были предметом обсуждения в работе [4] для системы $M/G/1$. В настоящей статье приводятся результаты для систем обслуживания с одним обслуживаюшим прибором, подверженным поломкам двух типов в случаях: а) с Пуассоновским входящим потоком (бесконечный источник) и б) с переменным потоком пуассоновского типа (конечный источник). Выбор этих систем обусловлен их использованием при изучении систем обслуживания с абсолютным приоритетом и ненадежным прибором. Кроме того, соответствующие результаты для систем с ненадежным или с надежным прибором получаются как частный случай здесь приведенных.

Почти все следующие результаты получены методом введения дополнительного события и на основе вероятностных толкований производящих функций распределений. При этом используется некоторая аналогия с дока-

СЕРДИКА Българско математическо списание. Том 3, 1977, с. 71—81

зательствами, приведенными в [4], и поэтому очень часто мы не будем повторять их.

2. Описание систем и обозначения. Система состоит из одного обслуживающего прибора (ОП), который обслуживает заявки по одной (порядок

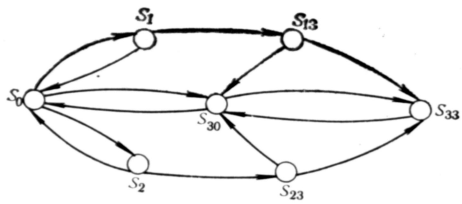

Рис. 1

обслуживания не имеет значения). Времена обслуживания предполагаются независимые одинаково распределенные случайные величины (сл. в.) с функцией распределения (ф. р.) $B(x)$. Во время обслуживания прибор с вероятностью 1 не выходит из строя. После очередного обслуживания ОП сразу переходит к следующему, если есть заявки, ожидающие в очереди. В свободном состоянии ОП может перетерпеть поломки двух типов: если прибор останется свободным в интервале времени длительностью t и в этом периоде не произойдет поломка второго (или первого) типа, то с вероятностью $F(t)$ (или с вероятностью $F_2(t)=1-e^{-ct}$, $c\geqq 0$) произойдет поломка первого (или второго) типа. После поломки i-го типа сразу начинается восстановление ОП, которое длится случайное время с ф. р. $G_i(t)$ $(i=1, 2)$. Все перечисленные сл. в. предполагаются независимыми в совокупности, а их ф. р. предположим непрерывными в нуле.

А. Система с конечным источником. В системе поступает Пуассоновский поток заявок с параметром $\lambda>0$. Параметр остается постоянным, независимо от времени и от числа заявок в системе.

Б. Система с конечным источником. Дано целое число $n>0$. Число заявок, находящихся одновременно в системе, не может быть больше n. Если в системе (в очереди и на обслуживание) имеется k заявок $(k\leqq n)$, то поступает Пуассоновский поток заявок с параметром $\lambda(n-k)$, $\lambda>0$. Когда число заявок в системе равно n, поступающего потока нет.

Среди состояний систем выделим следующие укрупненные состояния:

S_0 — ОП находится в исправности и заявок в системе нет;

S_i — ОП восстанавливается после поломки i-го типа и ожидающих заявок в системе нет;

S_{i3} — идет восстановление ОП после поломки i-го типа и имеется очередь;

S_{30} — прибор обслуживает некоторую заявку, но других ожидающих заявок нет;

S_{33} — идет обслуживание некоторой заявки и очередь не пуста.

Возможные переходы между состояниями системы изображены на рис. 1. Очередь существует, когда система находится в некотором состоянии класса $U=S_{13}+S_{23}+S_{33}$. При попадении в состояние из класса U очереди нет. Сл. в. θ_i представляют собой времена пребывания системы в классе U, когда величины ψ_i являются временами пребывания в классе $\bar{U}$.

Если положим $c=0$, получим системы обслуживания с обычным ненадежным прибором; если положим $F(t)=U(t)$, $G_1(t)=U(t)$, где $U(t)$ — функция Хевисайда (несобственное распределение, сосредоточенное в нуле), получим системы с экспоненциальными выходами ОП из строя; наконец, объединением обоих частных случаев получим системы обслуживания с надежным прибором.

Когда в систему поступают заявки r типов и пуассоновские потоки $L_1, L_2, \ldots, L_r$, то случай дисциплины обслуживания с абсолютным приоритетом сводится (при изучении характеристик обслуживания заявок k-го типа) к изучению систем с одним потоком и ненадежным прибором, причем „ненадежность" ОП соответствует поломкам второго типа и занятости прибора обслуживанием заявок более высокого приоритета [2, 5]. Вот почему мы остановили свой выбор на изучением характеристик очередей систем такого типа ненадежности ОП.

Пусть $A_i(t)$, $B_j(t)$, $C_k(t), \ldots$ некоторые функции. Их преобразования Лапласа-Стилтьеса (ЛСП) будем обозначать соответственно строчными латинскими или греческими буквами $(a_i(s), \beta_j(s), \ldots)$, в отличие от обычных Лапласовых преобразований, которые будем обозначать звездочкой (напр., $a_i^*(s)$, $\beta_j^*(s), \ldots$); если это ф. р., их моменты будем снабжать дополнительно нижним индексом, обозначающим порядок момента, напр., $a_i(s)=\int_0^\infty e^{-st}dA_i(t)$; $\beta_j^*(s)=\int_0^\infty e^{-st}B_j(t)dt$; $c_{kr}=\int_0^\infty t^r\, dC_k(t)$.

3. Период существования очереди. Обозначим через θ длительность интервала времени, начинающегося с любого момента возникновения очереди и заканчивающегося ближайшим последующим моментом, когда в системе нет очереди. Назовем θ периодом существования очереди. Тогда сл. в. θ имеет одну из следующих трех возможных реализаций:

θ_i — очередь возникает в момент восстановления ОП после поломки типа i $(i=1, 2)$;

θ_3 — очередь возникает во время обслуживания некоторой заявки.

Распределения величин θ_i обозначим через $T_i(x)$, а их ЛСП — через $\theta_i(s)$. Если положим $\chi_i=1$, когда очередь начнется реализацией типа i, и $\chi_i=0$ в противоположном случае $(i=1, 2, 3)$, тогда очевидно

$$\theta=\theta_1\chi_1+\theta_2\chi_2+\theta_3\chi_3, \tag{3.1}$$

откуда для ф. р. $T(x)$ величины θ следует равенство

$$T(x)=p_1T_1(x)+p_2T_2(x)+p_3T_3(x). \tag{3.2}$$

Здесь положено $p_i=P\{\chi_i=1\}$, $(i=1, 2, 3)$.

Займемся нахождением функций $T_i(x)$ и вероятностей p_i. Во-первых, заметим, что очередь θ_i просуществует не менее времени ζ_i, равному остатку времени восстановления (при $i=1, 2$) или остатку времени текущего обслуживания (при $i=3$). Пусть $H_i(x)$ — ф. р. величины ζ_i $(i=1, 2, 3)$. Следующее утверждение имеет вспомогательное значение.

Теорема 3.1. *a) Функции $H_i(x)$ определены однозначно своими ЛСП $h_i(s)$, которые задаются выражениями* $(i=1, 2)$

$$h_i(s)=\nu[g_i(s)-g_i(\nu)]/\{[1-g_i(\nu)](\nu-s)\},\quad h_3(s)=\mu[\beta(s)-\beta(\mu)]/\{[1-\beta(\mu)](\mu-s)\}.$$

б) Если $\beta_1<\infty$, $g_{i1}<\infty$ $(i=1, 2)$, *тогда* $h_{i1}=E\zeta_i$ *всегда существуют, причем*

$$h_{i1}=g_{i1}/[1-g_i(\nu)]-1/\nu,\ i=1, 2;\ h_{31}=\beta_1/[1-\beta(\mu)]-1/\mu. \tag{3.3}$$

Здесь положено

$$\begin{aligned} &\nu=\mu=\lambda && \text{— в системе с бесконечным источником,}\\ &\nu=\lambda n,\ \mu=\lambda(n-1) && \text{— в системе с конечным источником.}\end{aligned} \tag{3.4}$$

Доказательство теоремы следует просто из определения величин ζ_i — это условное математическое ожидание остаточного времени восстановления (обслуживания), если на данном случайном интервале произошло событие $\{\chi_i=1\}$, т. е. $\zeta_i=E\{\eta_i-\xi_i/\chi_i=1\}$, где через ξ_i обозначено время до прихода заявки, а η_i — время восстановления ОП (или время обслуживания) и $\chi_i=1$ при $\eta_i\geqq\xi_i$; $\chi_i=0$ при $\eta_i<\xi_i$ $(i=1, 2, 3)$. Равенство (3.3) следует из соотношения $h_{ik}=(-1)^k h^{(k)}(0)$, связывающего моменты ф. р. с производными ее ЛСП в нуле $(k=1, 2, \dots)$. Значения параметров ν и μ получены как соответствующие значения параметра экспоненциального распределения величины ξ_i.

Поскольку в дальнейшем некоторые выражения симметричны по отношению индекса i, положим $g_3(s)\equiv\beta(s)$, $s>0$; $g_{3k}=\beta_k$, $k=1, 2, \dots$

Теорема 3.2. А. *а) Для системы с бесконечным источником ф. р.* $T_i(x)$ *определены однозначно своими ЛСП, задающиеся равенствами*

$$\theta_i(s)=\{\lambda/[1-g_i(\lambda)]\}\,[\pi^{(i)}(s)-g_i(\lambda)]/[\lambda\pi(s)-s],$$

где функции $\pi^{(i)}(s)$ *определяются функциональными уравнениями* $\pi^{(i)}(s)=g_i(s+\lambda-\lambda\pi(s))$, $i=1, 2, 3$; $\pi(s)=\pi^{(3)}(s)$;

б) Если $\varrho=\lambda\beta_1\leqq 1$, *то* $\lim_{x\to\infty}T_i(x)=1$, *а в случае* $\varrho>1$

$$\lim_{x\to\infty}T_i(x)=[g_i(\lambda-\lambda\sigma)-g_i(\lambda)]/\{\sigma[1-g_i(\lambda)\},\ i=1, 2, 3,$$

где σ *— единственный корень уравнения* $\sigma=\beta(\lambda-\lambda\sigma)$, *находящийся внутри интервала* $(0, 1)$;

в) Если $\varrho<1$ *и* $g_{i1}>\infty$ $(i=1, 2)$, *то* $E\theta_i<\infty$ *и* $\theta_{i1}=E\theta_i=[g_i(\lambda)-\lambda g_{i1}-\varrho]/\{\lambda(1-\varrho)[1-g_i(\lambda)]\}$, $i=1, 2, 3$. *Существуют все моменты* θ_i *до того порядка, до которого конечны моменты распределений* $B(x)$ *и* $G_i(x)$.

Доказательство теоремы аналогично доказательству теоремы 2 в работе [4] и мы его опустим.

Теорема 3.2. Б. *а) Для системы с конечным источником функции* $T_i(x)$ *определены однозначно своими ЛСП* $\theta_i(s)$ *по формулам*

$$\theta_i(s)=\nu\,\{\Sigma_{k=0}^{n-\varepsilon}\binom{n-\varepsilon}{k}v_{k-1}(s)[g_i(\lambda k+s)-g_i(\nu)]/\mu_k(\lambda, s)\}/\{[1-g_i(\nu)]\,\Sigma_{k=0}^{n-1}\binom{n-1}{k}v_{k-1}(s)\},$$

где положено $\varepsilon=2$; $\nu=\lambda(n-1)$; $\mu_k(\lambda, s)=\beta(\lambda k+s)[\lambda(n-k-1)-s]$, *когда* $i=3$ *и* $\varepsilon=1$; $\nu=\lambda n$; $\mu_k(\lambda, s)=\lambda(n-k)-s$, *когда* $i=1, 2$, *а функции* $v_{k-1}(s)$ *заданы равенствами*

$$v_{k-1}(s)=\begin{cases}\Pi_{j=0}^{k-1}[1-\beta(j\lambda+s)]/\beta(j\lambda+s), & k>0,\\ 1, & k=0.\end{cases} \tag{3.5}$$

б) Если $g_{ik}<\infty$ $(i=1, 2, 3)$, *то существуют моменты* $\theta_{ik}=E\theta_i^k$ *того же порядка. В частности,*

$$\theta_{i1} = h_{i1} + \beta_1 \Sigma_{k=0}^{n-\varepsilon} \binom{n-\varepsilon}{k} \bar{v}_{k-1} [1 - h_i(k\lambda)],$$

где $h_i(s)$ *и* h_{i1} $(i=1, 2, 3)$ *определены согласно теореме 3.1;* $\varepsilon=1$, *когда* $i=1, 2$, *и* $\varepsilon=2$ *при* $i=3$. *Кроме того, положено*

$$\bar{v}_{k-1} = \begin{cases} \Pi_{j=1}^{k-1} [1-\beta(j\lambda)]/\beta(j\lambda), & \text{если } k>1; \\ 1 \qquad\qquad\qquad , & \text{если } k=1. \end{cases}$$

Доказательство проводится методом введения дополнительного события [1, 6] и рассуждениями, подобными тем в [5], или в [3, гл. II, § 2.2]. Заметим, что периоды θ_i напоминают периоды занятости с задержкой ζ_i, но в некоторой редуцированной системе с конечным источником, что используется в доказательстве теоремы.

Теперь заметим, что реализация θ_i периода существования очереди θ существенно зависит от состояния, в котором находилась система перед этим. В процессе обслуживания очередь будет исчезать переходом системы в состояние S_{30}, поэтому все очередные реализации θ, за исключением, может быть, первой, будут иметь одно и то же распределение $T(x)$, заданное равенством (3.2). Предположим, что в начальный момент $t_0=0$ система находится в состоянии s_0 и при этом условии обозначим вероятности реализации θ_i через $\widetilde{p}_i$. Тогда имеет место

Теорема 3.3. *а) Вероятности* $\widetilde{p}_i$ *задаются равенствами*

$$\widetilde{p}_1 = \varphi(c+\nu)[1-g_1(\nu)]/[1-k(c, \nu, \mu)];$$

$$\widetilde{p}_2 = [c/(c+\nu)]\ [1-\varphi(c+\nu)]\ [1-g_2(\nu)]/[1-k(c, \nu, \mu)];$$

$$\widetilde{p}_3 = [\nu/(c+\nu)]\ [1-\varphi(c+\nu)]\ [1-\beta(\mu)]/[1-k(c, \nu, \mu)],$$

где положено

$$k(c, \nu, \mu) = \varphi(c+\nu)[1-g_1(\nu)] + [1-\varphi(c+\nu)]\{c[1-g_2(\nu)] + \nu[1-\beta(\mu)]\}/(c+\nu), \tag{3.6}$$

а параметры ν *и* μ *определяются правилом* (3.4);

б) Вероятности p_i *и* $\widetilde{p}_i$ *связаны между собой равенствами* $p_i = \beta(\mu) \times \widetilde{p}_i$, *для* $i=1, 2$; $p_3 = 1-\beta(\mu)+\beta(\mu)\widetilde{p}_3$, *где* μ *определено правилом* (3.4);

в) Распределение $T(x)$ *определено своим ЛСП* $\theta(s) = p_1\theta_1(s) + p_2\theta_2(s) + p_3\theta_3(s)$, *где функции* $\theta_i(s)$ *определены теоремой 3.2. А. (или теоремой 3.2. Б.).*

Доказательство теоремы 3.3 легче всего произвести методом, использующим вероятностные толкования ЛСП функций распределения и теореме о полной вероятности, учитывая возможные переходы между состояниями графа на рис. 1 и их вероятности.

4. Период отсутствия очереди. Обозначим через ψ длительность интервала времени, начинающегося с момента исчезновения очереди и кончающегося первым последующим моментом, когда вновь появится очередь. Назовем ψ периодом отсутствия очереди. Соответствующую ф. р. обозначим через $\Psi(x)$, а ее ЛСП — через $\psi(s)$.

Ясно, что если в момент $t_0=0$ система свободна от заявок, а ОП — исправен, то первый период отсутствия очереди $\widetilde{\psi}$ имеет иное распределение $\widetilde{\Psi}(x)$, а все последующие интервалы времени, когда очереди нет, суть независимые между собою сл. в., все с распределением величины ψ.

Теорема 4.1. *а) Функция* $\widetilde{\Psi}(x)$ *определена своим ЛСП* $\widetilde{\psi}(s)$ *через равенство* $\widetilde{\psi}(s)=\{\nu[1-\varphi(\nu_1)][1-\beta(\mu+s)]\ \mu/(\mu+s)+\varphi(\nu_1)\ [1-g_1(\nu+s)]\,\nu/(\nu+s)$ $+c[1-\varphi(\nu_1)][1-g_2(\nu+s)]\,\nu/(\nu+s)\}/\{\nu_1-\nu[1-\varphi(\nu_1)]\ \beta(\mu+s)-\nu_1\varphi(\nu_1)\ g_1(\nu+s)$ $-c[1-\varphi(\nu_1)]\ g_2(\nu+s)\}$, *где положено* $\nu_1=\nu+c+s$ *и еще*

$$\nu=\lambda,\ \mu=\lambda \quad \text{— для системы с бесконечным источником;} \tag{4.1}$$
$$\nu=\lambda n,\ \mu=\lambda(n-1) \quad \text{— для системы с конечным источником;}$$

б) Функции $\psi(s)$ *и* $\widetilde{\psi}(s)$ *связаны равенством* $\psi(s)=\mu[1-\beta(\mu+s)]/(\mu+s)$ $+\beta(\mu+s)\widetilde{\psi}(s)$, *где величина* μ *определена правилом* (4.1);

в) Средняя длительность периода отсутствия очереди существует во всех случаях и задается равенствами

$$\widetilde{\psi}_1=\{[1+(\nu\,[1-\beta(\mu)]/\mu+c[1-g_2(\nu)]/\nu]\ [1-\varphi(\nu+c)]/(\nu+c)$$
$$+\varphi(\nu+c)\ ([1-g_1(\nu)]/\nu-1/(\nu+c))\}\{1-\varphi(\nu+c)\ g_1(\nu)$$
$$-[1-\varphi(\nu+c)]\ [\nu\beta(\mu)+cg_2(\nu)]/(\nu+c)\}^{-1};$$
$$\psi_1=[1-\beta(\mu)]/\mu+\beta(\mu)\,\widetilde{\psi}_1.$$

Участвующие здесь величины ν *и* μ *определены соотношением* (4.1).

В доказательстве теоремы использованы вероятностные рассуждения, учитывающие вероятности переходов по стрелкам ориентированного графа на рис. 1 и то, что моменты перехода в состояние S_{30} являются моментами регенерации для процесса обслуживания. Вероятностное толкование ЛСП при $s>0$ намного упрощает вывод приведенных результатов.

5. Число заявок, становившихся в одну очередь. Пусть $\varkappa$ — число заявок, ждавших в очереди в течение одного периода ее существования. Аналогично равенству (3.1) получим представление $\varkappa=\varkappa_1\chi_1+\varkappa_2\chi_2+\varkappa_3\chi_3$, где $\varkappa_i$ — число заявок, ждавших в очереди при реализации типа i для периода ее существования. Пусть $r_m=P\{\varkappa=m\}$ и пусть $r(z)=r_1z+r_2z^2+r_3z^3+\cdots$ производящая функция вероятностей $\{r_m\}$. Верны следующие утверждения.

Теорема 5.1. А. *а) В системе с бесконечным источником функция* $r(z)$ *задается равенством*

$$r(z)=p_1zh_1(\lambda-\lambda q(z))+p_2zh_2(\lambda-\lambda q(z))+p_3zh_3(\lambda-\lambda q(z)), \tag{5.1}$$

где функция $q(z)$ *определена функциональным уравнением*

$$q(z)=z\beta(\lambda-\lambda q(z)). \tag{5.2}$$

Функции $h_i(s)$ *определены теоремой* 3.1. *а), а вероятности* p_i *определяются теоремой* 3.3. $(i=1,2,3)$;

б) Равенства (5.1), (5.2) *определяют единственную функцию* $r(z)$, *которая при* $\lambda\beta_1\leq r$ *есть производящая функция некоторого вероятностного распределения* $\{r_m\}$;

в) Если $\beta_1<\infty$, $g_{i1}<\infty$ $(i=1,2)$ *и* $\lambda\beta_1<1$, *то существует* $\varkappa_1=E\varkappa$ *и* $\varkappa_1=1+\{p_1\lambda g_{11}[1-g_1(\lambda)]^{-1}+p_2\lambda g_{21}[1-g_2(\lambda)]^{-1}+\lambda\beta_1[1-\beta(\lambda)]^{-1}-1\}/(1-\lambda\beta_1)$, *где вероятности* p_i *определены теоремой* 3.3.

Теорема 5.1. Б. *а) Для системы с конечным источником функция* $r(z)$ *определена выражением*

$$r(z)=\{\Sigma_{i=1}^{2}p_i\Sigma_{k=1}^{n-1}\binom{n-1}{k}u_{k-1}(z)\,h_i(k\lambda)+p_3\Sigma_{k=1}^{n-2}\binom{n-2}{k}u_{k-1}(z)h_3(k\lambda)\}/$$

$$\{\Sigma_{k=1}^{n-1}\binom{n-1}{k}u_{k-1}(z)\},$$

где функции $u_{k-1}(z)$ *заданы равенствами*

$$u_{k-1}(z)=\begin{cases}\Pi_{j=0}^{k-1}[1-z\beta(j\lambda)]/[z\beta(j\lambda)], & \textit{если } k>0;\\ 1, & \textit{если } k=0,\end{cases}$$

а функции $h_i(s)$ *и вероятности* p_i $(i=1, 2, 3)$ *определены соответственно теоремами* 3.1 *и* 3.3;

б) Среднее значение $\varkappa_1=E\varkappa$ *всегда существует, когда* $\beta_1>\infty$ *и*

$$\varkappa_1=1+\Sigma_{i=1}^{2}p_i\Sigma_{k=1}^{n-1}\binom{n-1}{k}\overline{v}_{k-1}[1-h_i(k\lambda)]+p_3\Sigma_{k=1}^{n-2}\binom{n-2}{k}\overline{v}_{k-1}[1-h_3(k\lambda)],$$

где величины $\overline{v}_{k-1}$ *определены равенствами* (3.5).

При доказательстве этих теорем тоже используется метод введения дополнительного события и вероятностное толкование производящих функций $r(z)$ и $q(z)$ при $z\in[0, 1]$. В доказательстве теоремы 5.1. б) еще используется дискретная трансформация, аналогичная примененной в [5].

6. Число заявок, обслуженных без ожидания. Рассмотрим число заявок N, которые будут обслужены в одном периоде отсутствия очереди. Их обслуживание начнется сразу в момент их прихода в систему. Поскольку каждый период исчезновения очереди начинается с обслуживания заявки, которая ждала последней в очереди, то в таком периоде может и не быть ни одной заявки, обслуженной без предварительного ожидания.

Теорема 6.1. *Случайная величина* N *имеет геометрическое распределение*

$$P\{N=k\}=q^{k-1}(1-q),\ k=0, 1, 2, \ldots, \tag{6.1}$$

где параметр q *определен равенством*

$$q=\nu\beta(\mu)[1-\varphi(c+\nu)]/\{1-\varphi(\nu+c)g_1(\nu)-[1-\varphi(\nu+c)]g_2(\nu)\,c/(\nu+c)\}, \tag{6.2}$$

причем

$$\begin{cases}\nu=\lambda,\ \mu=\lambda & \text{— для системы с бесконечным источником;}\\ \nu=\lambda n,\ \mu=\lambda(n-1) & \text{— для системы с конечным источником.}\end{cases} \tag{6.3}$$

Доказательство. Сл. в. N примет значение 0, если

а) либо во время обслуживания заявки, с началом обслуживания которой исчезает очередь, придет на обслуживание следующая заявка и положит начало новой очереди (вероятность чего есть $1-\beta(\mu)$);

б) либо это первое обслуживание кончится без прихода новой заявки, система попадет в состояние S_0 (см. фиг. 1) и приход следующей заявки случится во время восстановления ОП (т. е. система попадет в состояние класса $S_{13}+S_{23}$ прежде, чем попасть в состояние S_{30}). Вероятность этого случая равна $\beta(\mu)(\widetilde{q}_1+\widetilde{q}_2)$, где $\widetilde{q}_i$ — вероятность, что первая заявка (после попадания системы в состояние (S_0) придет во время ремонта ОП типа i $(i=1, 2)$.

Поскольку моменты попадания в состояние S_0 являются точками регенерации для процесса обслуживания, то легко получить (учитывая только благоприятствующие скачки системы по графу на рис. 1) для вероятностей $\widetilde{q}_i$ следующие уравнения

$$\widetilde{q}_1=\varphi(c+\nu)[1-g_1(\nu)]+\{\varphi(c+\nu)\,g_1(\nu)+[1-\varphi(c+\nu)\,g_2(\nu)]c/(c+\nu)\}\ \widetilde{q}_1;$$

$$\widetilde{q}_2=[1-\varphi(c+\nu)]\ [1-g_2(\nu)]\ c/(c+\nu)+\{\varphi(c+\nu)\,g_1(\nu)+[1-\varphi(c+\nu)]\,g_2(\nu)c/(c+\nu)\}\widetilde{q}_2. \tag{6.4}$$

Отсюда, после несложных преобразований, получим выражение для вероятности

$$1-q=1-\beta(\mu)+\beta(\mu)\ (\widetilde{q}_1+\widetilde{q}_2), \tag{6.5}$$

решение которого по отношению вероятности q приведет к формуле (6.2)

Событие $\{N=1\}$ осуществится, если первое обслуживание (с которого начинается период отсутствия очереди) закончится без прихода новой заявки (вероятность — $\beta(\mu)$), после попадания системы в состояние S_0 приход первой следующей заявки осуществится в момент, когда прибор находится в свободном состоянии (вероятность — $1-\widetilde{q}_1-\widetilde{q}_2$, где $\widetilde{q}_i$ определены равенствами (6.4)) и после начала ее обслуживания следующая заявка положит начало новой очереди (вероятность — $1-q$). По теореме умножения вероятностей получаем $P\{N=1\}=\beta(\mu)\ (1-\widetilde{q}_1-\widetilde{q}_2)\ (1-q)$. Но ввиду (6.5) $\beta(\mu)\ (1-\widetilde{q}_1-\widetilde{q}_2)=q$, откуда $P\{N=1\}=q(1-q)$. Вообще, приняв во внимание, что моменты начала обслуживания заявки, когда очередь пуста, являются моментами регенерации для процесса обслуживания, формулу (6.1) можно вывести методом математической индукции.

Значения (6.3) параметров μ и ν следуют из соответствующих значений параметра входящего потока в системе при данном ее состоянии.

Производящая функция распределения величины N, как известно, равна

$$N(z)=(1-q)/(1-qz), \tag{6.6}$$

а ее среднее значение —

$$EN=q/(1-q). \tag{6.7}$$

Займемся теперь числом заявок N_t, обслуженных без ожидания в интервале $(0, t)$. Для простоты предположим, что в момент $t_0=0$ начинается период отсутствия очереди. Через $\widetilde{N}_t$ обозначим число заявок, принятых на обслуживание без ожидания до момента t при условии, что в $(0, t)$ очередь не появлялась. Положим $\widetilde{P}(z, t)=Ez^{\widetilde{N}_t}$, $P(z, t)=Ez^{N_t}$, $|z|\leq 1$, а их преобразования Лапласа обозначим соответственно через $\widetilde{p}^*(z, s)$ и $p^*(z, s)$. Тогда имеет место следующая теорема:

Теорема 6.2. *а) Функция $\widetilde{p}^*(z, s)$ определена равенством*

$$p^*(z, s)=\{[1-\varphi(\nu_1)]+c[1-\varphi(\nu_1)]\ [1-g_2(\nu+s)]/(\nu+s)+\varphi(\nu_1)[1-g_1(\nu+s)]\nu_1/(\nu+s)$$
$$+\nu z[1-\beta(\mu+s)]/(\mu+s)\}/\{\nu_1-c[1-\varphi(\nu_1)]g_2(\nu+s)-\nu_1\varphi(\nu_1)g_1(\nu+s)-\nu z\beta(\mu+s)\},$$

где $\nu_1=\nu+c+s$, а величины ν и μ определены (6.3);

б) Функции $p^*(z, s)$ и $\widetilde{p}^*(z, s)$ связаны соотношением

$$p^*(z, s)=\widetilde{p}^*(z, s)/\{1-\theta(s)\psi(s)\,N(z)\},$$

где функции $\theta(s)$ и $\psi(s)$ определены соответственно в пунктах 2 и 3, а $N(z)$ задана равенствами (6.6) *и* (6.2).

Теорема доказывается методом введения дополнительного события, позволяющего найти вероятностное толкование для преобразований Лапласа $sp^*(z, s)$ и $s\widetilde{p}^*(z, s)$ производящих функций распределения величин N_t и $\widetilde{N}_t$ (подобное доказательство см., например, в [1, гл. I, §5]). Еще используется тот факт, что моменты исчезновения очереди образуют процесс восстановления с периодом $\theta+\psi$.

Из теоремы 6.2 в качестве следствия можно получить преобразование Лапласа $m^*(s)$ среднего числа $M(t)=EN_t$ заявок, обслуженных без ожидания в интервале $(0, t)$, а именно:

$$m^*(s)=\left.\frac{\partial p^*(z, s)}{\partial z}\right|_{z=1}=\left.\frac{\partial \widetilde{p}^*(z, s)}{\partial z}\right|_{z=1}/[1-\theta(s)\psi(s)]+s^{-1}\, EN\,\theta(s)\psi(s)/[1-\theta(s)\psi(s)],$$

где EN определено (6.7) и (6.2). Используя это представление, можно получить следующее асимптотическое равенство $sm^*(s)=[EN+o(s)]/[s(\theta_1+\psi_1)+o(s)]$ при $s\to 0$, которое имеет место, если конечны первые моменты θ_1 и ψ_1 периода существования и периода отсутствия очереди. Применив теорему Таубера, приходим к выводу, что справедливо асимптотическое соотношение $M(t)\sim qt/[(1-q)(\theta_1+\psi_1)]$ при $t\to\infty$, дающее асимптотическое приближение числа заявок, обслуженных без ожидания на интервале большой длительности t.

7. Число появления очередей в одном периоде занятости. Периодом занятости ОП назовем интервал времени, начинающийся с момента прихода заявки в систему и кончающийся первым последующим моментом, когда в системе заявок не будет. Такой период может начаться либо в момент восстановления ОП типа i $(i=1, 2)$, либо в момент, когда ОП находится в состоянии S_0. Эту реализацию назовем периодом занятости типа 3. Ясно, что в первых двух случаях всегда будет хоть одно (первое) появление очередей, но в третьем очереди может и не быть (если обслужится только первая заявка и за это время не появятся другие заявки).

Пусть q_i — вероятность того, что осуществится реализация типа i для периода занятости, и пусть δ — число появлений очередей в одном периоде занятости. Тогда справедливо следующее утверждение:

Теорема 7.1. *а) Вероятности* q_i $(i=1, 2, 3)$ *определены соотношениями* $q_i=\widetilde{p}_i$, $i=1, 2$, $q_3=\nu[1-\varphi(c+\nu)]/\{(c+\nu)[1-k(c, \nu)]\}$, *где* $\widetilde{p}_i$ *заданы равенствами* (3.6) *и положено*

$$k(c, \nu)=\varphi(c+\nu)[1-g_1(\nu)]+[1-\varphi(c+\nu)]\ [1-g_2(\nu)]c/(c+\nu),$$

причем $\nu=\lambda$ *— для системы с бесконечным источником и* $\nu=\lambda n$ *для системы с конечным источником;*

б) Распределение величины δ *задается формулами*

$$P\{\delta=0\}=q_3\,\beta(\mu);$$
$$P\{\delta=k\}=\beta(\mu)[1-q_3\beta(\mu)]\ [1-\beta(\mu)]^{k-1},\ k=1, 2, 3, \ldots, \tag{7.1}$$

где $\mu=\lambda$ *— если источник заявок бесконечен и* $\mu=\lambda(n-1)$, *когда источник конечен;*

в) Всегда существует среднее $E\delta$ *и* $E\delta=[1-q_3\beta(\mu)]/\beta(\mu)$.

В доказательстве теоремы отметим, что вероятности q_i $(i=1, 2)$ играют лишь вспомогательную роль. Распределение величины δ является смесью трех геометрических распределений, т. е.

$$P\{\delta=k\}=q_1P\{\delta_1=k\}+q_2P\{\delta_2=k\}+q_3P\{\delta_3=k\},\ k=0, 1, 2, \ldots,$$

где δ_i — число появления очередей при i-той реализации периода занятости типа i. Но легко найти, что $P\{\delta_3=0\}=\beta(\mu)$; $P\{\delta_i=0\}=0$, $i=1, 2$ и еще $P\{\delta_3=k\}=[1-\beta(\mu)]^k\beta(\mu)$; $P\{\delta_i=k\}=[1-\beta(\mu)]^{k-1}\beta(\mu)$, $i=1$, 2, откуда $P\{\delta=k\}=\{q_3[1-\beta(\mu)]+q_1+q_2\}\ [1-\beta(\mu)]^{k-1}\beta(\mu)$, $k>1$.

Так как $q_1+q_2+q_3=1$, то из последнего равенства следует (7.1).

8. Вероятность существования очереди. В процессе обслуживания в системе очередь то появляется, то исчезает. Предположим, что момент $t_0=0$ является началом периода отсутствия очереди и через $P_0(t)$ и $P_1(t)$ обозначим соответственно вероятности отсутствия и существования очереди в момент $t>0$. Пусть $p_i^*(s)$ $(i=0,\ 1)$ — преобразования Лапласа введенных вероятностей.

Теорема 8.1. *а) Функция $p_0^*(s)$ определена выражением*

$$p_0^*(s)=[1-\psi(s)]s^{-1}\,[1-\theta(s)\psi(s)]^{-1}, \tag{8.1}$$

где функции $\theta(s)$ и $\psi(s)$ определены соответственно в пунктах 3 и 4;

б) Функции $p_0^*(s)$ и $p_1^*(s)$ связаны соотношением

$$p_1^*(s)=s^{-1}-p_0^*(s)=\psi(s)[1-\theta(s)]s^{-1}[1-\theta(s)\psi(s)]^{-1}; \tag{8.2}$$

в) Если выполнены условия теоремы 3.2. А. в) (соотв. теоремы 3.2. Б. в), то существует предел

$$P_1=\lim\nolimits_{t\to\infty}P_1(t)=\theta_1/(\theta_1+\psi_1), \tag{8.3}$$

где θ_1 и ψ_1 — средние значения периодов существования и отсутствия очереди соответственно.

Доказательство теоремы легко провести, если заметить, что периоды отсутствия очереди и периоды ее существования независимы между собой и образуют альтернирующий процесс восстановления. Очередь в момент t существует, когда этот момент „накрыт" неким периодом существования очереди. Пользуясь методом вероятностного толкования функций $sp_0^*(s)$ и $[1-\psi(s)]$, $\psi(s)$, $\theta(s)$, получим для $p_0^*(s)$ уравнение $sp_0^*(s)=[1-\psi(s)]+\psi(s)\theta(s)\,sp_0^*(s)$, откуда следует (8.1). Равенство (8.2) следует из равенства $P_0(t)+P_1(t)=1$, а (8.3) получается с помощью тауберовой теоремы, согласно которой $\lim_{t\to\infty}P_1(t)=\lim_{s\downarrow 0}sp_1^*(s)$.

9. Очередь в многолинейной системе обслуживания типа $GI|M|m$. Рассмотрим систему, состоящуюся из m обслуживающих приборов, в которую поступает рекуррентный поток заявок с распределением $A(x)$. Времена обслуживания i-го прибора распределены экспоненциально с параметром λ_i $(i=1, 2, \ldots, m)$. Очередь в системе предполагается неограниченной. Она возникает, когда в систему придет заявка и застанет все приборы занятыми обслуживанием.

Задача нахождения периода существования очереди θ и периода ее исчезновения ψ сводится к уже известным задачам. Для определения θ следует заметить, что во время существования очереди обслуженные заявки уходят из системы, формируя поток Пуассона с параметром $\lambda=\lambda_1+\cdots+\lambda_m$. Если предположим, что заявка, с приходом которой появится очередь в системе, пропускает перед собой все пришедшие после нее заявки, заставшие ее ждавшей, тогда очередь кончится тем, что эта заявка поступит в обслуживание. Таким образом период существования очереди не изменится, но это предположение поможет понять, что изучаемый период по существу равен периоду занятости ОП в однолинейной системе обслуживания типа $G|M|1$. Входящий поток рекуррентен с ф. р. $A(x)$, время обслуживания — экспоненциально с параметром λ, а период занятости начинается с приходом заявки, сложившую начало очереди. Поэтому для ф. р. $T(x)$ периода сушествования очереди имеем [6, § 13, теорема 2]:

Теорема 8.1. *а) Функция $T(x)$ определяется однозначно своим ЛСП $\theta(s)=\lambda[1-\gamma(s)]/\{s+\lambda[1-\gamma(s)]\}$, где $\lambda=\lambda_1+\cdots+\lambda_m$, а $\gamma(s)$ определена как решение функционального уравнения $\gamma(s)=a(s+\lambda-\lambda\gamma(s))$. Здесь $\alpha(s)$ — ЛСП функция $A(x)$. При этом $T(0)=0$ и $\lim_{x\to\infty} T(x)=\min(1, \lambda\alpha_1)$;*

б) *Если $\alpha_1(\lambda_1+\cdots+\lambda_m)>1$, то существует среднее $\theta_1=(\lambda_1+\cdots+\lambda_m)^{-1}\cdot(1-\sigma)^{-1}$, где σ — единственный корень уравнения $\sigma=\alpha((\lambda_1+\cdots+\lambda_m)(1-\sigma))$, лежащий внутри интервала $(0, 1)$. Если $\alpha_1(\lambda_1+\cdots+\lambda_m)\leqq 1$, то $\theta_1=\infty$.*

Определение же периода отсутствия очереди ψ сводится к задаче Пальма для системы $GI|M|m$ с потерями. Если считать, что заявки, заставшие все линии занятыми, теряются, то период ψ будет равняться интервалу времени между моментами двух очередных потерь заявок. Принципиальное решение этой задачи известно [6, § 23].

ЛИТЕРАТУРА

1. Б. В. Гнеденко, Э. А. Даниелян, Б. Димитров, Г. Климов и В. Матвеев. Приоритетные системы обслуживания. Москва, 1973.
2. Э. Даниелян, Б. Димитров. О системе $M|G|1$ с ожиданием и двумя типами отказов. *Mathem. Balkanica*, **2**, 1972, 21—37.
3. Н. Джейсуол. Очереди с приоритетами. Москва, 1973.
4. Б. Димитров. Някои нови характеристики на процесите на обслужване. Математика и математическо образование. Доклади на Трета пролетна конференция на БМД — Бургас — 1974. София, 1976, 160—168.
5. Б. Димитров, Хр. Карапенев. О системах обслуживания с конечным источником и ненадежным прибором. *Годишник на Соф. унив.*, **67**, 1974, 483—505.
6. Г. П. Климов. Стохастические системы обслуживания. Москва, 1966.

Единый центр науки и подготовки кадров по математике и механике
1000 София П. Я. 373

Поступила 16.2. 1976