Информационные процессы, Том 12, № 3, 2012, стр. 159–171.

Стационарное распределение времени ожидания в системе обслуживания с отрицательными заявками и бункером для вытесненных заявок при дисциплине LAST-LIFO-LIFO[1]

А.В. Печинкин, Р.В. Разумчик

Институт проблем информатики РАН, Москва, Россия
Поступила в редколлегию 12.07.2012

Аннотация—Рассматривается система массового обслуживания с одним обслуживающим прибором и пуассоновскими потоками обычных и отрицательных заявок. Для обычных заявок имеется накопитель неограниченной ёмкости. Отрицательная заявка, поступающая в систему, выбивает обычную заявку из очереди в накопителе и перемещает её в бункер неограниченной ёмкости, откуда заявки обслуживаются с относительным приоритетом. Длительности обслуживания заявок из накопителя и из бункера имеют экспоненциальные распределения с различными параметрами. Предполагается, что отрицательная заявка выбивает в бункер последнюю заявку из очереди в накопителе, а на обслуживание из накопителя или бункера выбирается последняя заявка в очереди.
Для рассматриваемой системы найдено в терминах преобразования Лапласа–Стилтьеса стационарное распределение времени ожидания начала обслуживания поступившей в систему заявки.
КЛЮЧЕВЫЕ СЛОВА: система обслуживания, отрицательные заявки, время ожидания, период занятости.

1. ВВЕДЕНИЕ И ОПИСАНИЕ СИСТЕМЫ

Многие ситуации, возникающие в современных телекоммуникационных системах (например, перебои в работе, потеря информации и т.д.), могут быть смоделированы с помощью систем и сетей массового обслуживания с отрицательными заявками. В настоящее время продолжают появляться публикации с теоретическими исследованиями в этой области, что подтверждает её актуальность. Некоторое представление о последних результатах можно найти в [1]–[10].

В работе [11] была введена система массового обслуживания (СМО) с отрицательными заявками и бункером для вытесненных заявок, в которой, в отличие от ранее изученных моделей, выбитая из накопителя заявка обслуживается с той же интенсивностью, но только после окончания обслуживания всех невыбитых заявок. Для этой СМО были найдены стационарные распределения основных показателей функционирования, связанных с числом заявок в накопителе и бункере. Затем в работах [12] и [13] рассматривались временные характеристики данной системы и было получено в терминах преобразования Лапласа-Стильтьеса (ПЛС) стационарное распределение времени ожидания начала обслуживания заявки для 8 различных комбинаций порядков выбора заявок на обслуживание из очереди в накопителе, выбора заявок на обслуживание из очереди в бункере и выбивания заявок из накопителя в бункер.

В настоящей работе, являющейся продолжением [11], исследуются временные характеристики системы в предположении, что заявки из накопителя и бункера обслуживаются с различ-

[1] Работа выполнена при поддержке Российского фонда фундаментальных исследований (гранты № 11-07-00112 и № 12-07-00108)

ными интенсивностями. Несмотря на всю естественность данного предположения, оно существенно усложняет анализ и требует отличного от предложенного в [12] подхода к нахождению временных характеристик. Разработке такого подхода и посвящена данная статья.

Рассмотрим однолинейную СМО, в которую поступает пуассоновский поток заявок интенсивности λ (заявки этого потока будем также называть положительными). Для этих заявок имеется накопитель неограниченной ёмкости.

Помимо положительных заявок, в систему поступает пуассоновский поток отрицательных заявок интенсивности λ^-. Отрицательная заявка, поступающая в систему, вытесняет одну заявку из очереди в накопителе и перемещает её в накопитель для вытесненных заявок, или бункер, который также имеет неограниченную ёмкость. Если же в момент поступления отрицательной заявки в накопителе нет положительных заявок, а на приборе обслуживается заявка, то отрицательная заявка покидает систему, не оказывая на неё никакого воздействия. То же самое происходит и в случае, когда в момент поступления отрицательной заявки накопитель и обслуживающий прибор свободны.

После окончания обслуживания очередной заявки на прибор становится заявка из накопителя. Если же накопитель пуст, на прибор поступает заявка из бункера. Обслуживание заявок не прерывается новыми как положительными, так и отрицательными заявками.

Длительности обслуживания заявок из накопителя имеют экспоненциальное распределение с параметром μ, а из бункера — экспоненциальное распределение с параметром μ^-.

Дисциплина выбивания и обслуживания заявок заключается в том, что отрицательная заявка при поступлении выбивает в бункер последнюю заявку из очереди заявок в накопителе, а при выборе заявок на обслуживание из накопителя или бункера выбирается последняя заявка в очереди. Такую дисциплину, по аналогии с [12], будем кодировать как LAST-LIFO-LIFO.

Будем считать, что существует стационарный режим функционирования системы. Необходимое и достаточное условие для этого приведено в [14].

Основной результат работы состоит в вычислении в терминах ПЛС стационарного распределения времени ожидания начала обслуживания поступившей в систему (положительной) заявки. Для этого применяется известная техника исследования СМО на одном периоде занятости. Поскольку используемые здесь результаты по-отдельности можно найти во многих публикациях, то для удобства читателей далее будем ссылаться на учебник [16], в котором содержатся все необходимые сведения.

В следующем разделе приводятся некоторые результаты, касающиеся периода занятости СМО $M|M|1|\infty$, в разделе 3 в терминах ПЛС получены выражения для условного распределения длины периода занятости исходной СМО, а в разделе 4 показано, как с помощью полученных ранее характеристик находится ПЛС стационарного распределения времени ожидания начала обслуживания заявки.

2. ВСПОМОГАТЕЛЬНЫЕ РЕЗУЛЬТАТЫ

Рассмотрим простейшую СМО $M|M|1|\infty$ с входящим потоком интенсивности λ и интенсивностью обслуживания $\mu^* = (\lambda^- + \mu)$. Обозначим через $G_n(t)$, $n \geq 1$, совместное распределение периода занятости (ПЗ) системы и числа обслуженных на нём заявок, а через $\gamma(z,s) = \sum\limits_{n=1}^{\infty} z^n \int\limits_0^{\infty} e^{-st} dG_n(t)$ обозначим двойное преобразование — ПЛС и производящую функцию (ПФ) — этого распределения. Тогда (см. [16, с. 300]) $\gamma(z,s)$ имеет вид

$$\gamma(z,s) = \frac{(\lambda + \mu^* + s) - \sqrt{(\lambda + \mu^* + s)^2 - 4\lambda\mu^* z}}{2\lambda}\,. \tag{1}$$

Предположим теперь, что ПЗ СМО $M|M|1|\infty$ с параметрами (λ, μ^*) начинается в тот момент, когда в системе находится $k,\ \ k \geq 1$, заявок. Продолжительность этого ПЗ и число обслуженных на нём заявок имеют двойное преобразование $\gamma_k(z,s) = \gamma^k(z,s),\ \ k \geq 1$.

Обозначим через $G_n(t|x),\ \ n \geq 0$, условное распределение ПЗ и числа обслуженных на нём заявок, исключая первую заявку на приборе, при условии, что длина открывающей его заявки равна x, а через $\delta(z,s|t) = \sum\limits_{n=0}^{\infty} z^n \int\limits_0^{\infty} e^{-st} dG_n(t|x)$ — двойное преобразование $G_n(t|x)$, определяемое формулой

$$\delta(z,s|t) = e^{-t(s+\lambda[1-\gamma(z,s)])}. \tag{2}$$

2.1. Характеристики на одном периоде занятости

Рассмотрим СМО $M|M|1|\infty$ с входящим потоком интенсивности λ и интенсивностью обслуживания λ^-. Пусть в начальный момент в этой системе находится $k,\ \ k \geq 1$, заявок и одна из них начинается обслуживаться. Обозначим через $B(x) = 1 - e^{-\lambda^- x}$ функцию распределения (ФР) длительности обслуживания заявки на приборе, через $b(x) = \lambda^- e^{-\lambda^- x}$ — её плотность и через $\beta(s) = \lambda^-/(s+\lambda^-)$ — ПЛС ФР $B(x)$.

Найдем $F_{km}(n,t),\ \ n \geq 1,\ \ m \geq 1,$ — вероятность одновременного выполнения следующих событий: n-я заявка будет обслужена до момента t, до момента окончания обслуживания n-й заявки система ни разу не освободится (не закончится ПЗ, начавшийся в момент 0) и сразу же после окончания обслуживания n-й заявки в системе останется $m,\ \ m \geq 1$, заявок. Естественно при этом положить $F_{km}(0,t) = 0,\ \ m \geq 1,\ \ m \neq k$, и $F_{kk}(0,t) = 1$ при всех $t > 0$. Введем обозначение $f_{km}(n,t) = F'_{km}(n,t),\ \ n \geq 0,\ \ m \geq 1$. Тогда ([16, с. 302]) при $n \geq k$ справедливо соотношение

$$f_{km}(n,t) = \sum_{i=1}^{m+1} \int\limits_0^t \frac{[\lambda\,(t-y)]^{m-i+1}}{(m-i+1)!}\, e^{-\lambda(t-y)} f_{ki}(n-1,y)\, b(t-y)\, dy, \quad n \geq k, \quad m \geq 1, \quad k \geq 1,$$

а при $n = \overline{1, k-1}$ и $k > 1$ имеет место равенство

$$f_{km}(n,t) = \sum_{i=k-(n-1)}^{m+1} \int\limits_0^t \frac{[\lambda\,(t-y)]^{m-i+1}}{(m-i+1)!}\, e^{-\lambda(t-y)} f_{ki}(n-1,y)\, b(t-y)\, dy, \quad m \geq k-n.$$

Поскольку в момент 0 в системе находится $k,\ \ k \geq 1$, заявок и одна из них начинается обслуживаться, то $F_{km}(n,t) = 0,\ \ n = \overline{1, k-2},\ \ m = \overline{1, k-n-1},\ \ t > 0$. Тогда, объединяя два последних выражения для $f_{km}(n,t)$ в одно, получаем

$$f_{km}(n,t) = \sum_{i=1}^{m+1} \int\limits_0^t \frac{[\lambda\,(t-y)]^{m-i+1}}{(m-i+1)!}\, e^{-\lambda(t-y)} f_{ki}(n-1,y)\, b(t-y)\, dy, \quad n \geq 1, \quad m \geq 1,$$

откуда, переходя к тройному преобразованию $\phi_k(z_1,s,z_2) = \sum\limits_{n=0}^{\infty} \sum\limits_{m=1}^{\infty} z_1^n z_2^m \int\limits_0^{\infty} e^{-st} f_{km}(n,t)\, dt$, находим

$$\phi_k(z_1,s,z_2) - z_2^k = \frac{z_1 \beta(\lambda + s - \lambda\, z_2)}{z_2} \phi_k(z_1,s,z_2) - z_1 \beta(s+\lambda) \frac{\partial}{\partial z_2} \phi_k(z_1,s,z_2)\Big|_{z_2=0}$$

и

$$\phi_k(z_1,s,z_2) = \frac{z_2 \big[z_2^k - z_1 \beta(s+\lambda)\, \frac{\partial}{\partial z_2} \phi_k(z_1,s,z_2)\big|_{z_2=0}\big]}{z_2 - z_1 \beta(s+\lambda-\lambda z_2)}\,. \tag{3}$$

Для определения числителя дроби в (3) заметим, что поскольку $\beta(s) = \lambda^-/(s+\lambda^-)$, то уравнение

$$z_2 - z_1\beta(s+\lambda-\lambda z_2) = 0$$

можно переписать в виде

$$u(z_2) = \lambda z_2^2 - (s+\lambda^- +\lambda)z_2 + \lambda^- z_1 = 0.$$

Функция $u(z_2)$ представляет собой квадратный трёхчлен относительно z_2, имеющий на интервале $(0,1)$ единственный корень $\tilde{z}(z_1,s)$, определяемый формулой

$$\tilde{z}(z_1,s) = \frac{s+\lambda^- +\lambda - \sqrt{(s+\lambda^- +\lambda)^2 - 4\lambda\lambda^- z_1}}{2\lambda}.$$

Используя свойство аналитичности ПФ $\phi_k(z_1,s,z_2)$ получаем, что числитель в (3) должен обращаться в нуль в точке $(z_1,s,\tilde{z}(z_1,s))$, т.е.

$$\frac{\partial}{\partial z_2}\phi_k(z_1,s,z_2)\Big|_{z_2=0} = \frac{\tilde{z}^k(z_1,s)}{z_1\beta(s+\lambda)},$$

откуда получаем окончательный вид ПФ $\phi_k(z_1,s,z_2)$:

$$\phi_k(z_1,s,z_2) = \frac{z_2\,(s+\lambda^- +\lambda-\lambda z_2)\,[z_2^k - \tilde{z}^k(z_1,s)]}{z_2(s+\lambda^- +\lambda-\lambda z_2) - \lambda^- z_1}.$$

Обозначим теперь через $R_{km}(n,t)$, $n\ge 0$, $m\ge 1$, вероятность события, заключающегося в том, что до момента t не закончится ПЗ, будет обслужено n заявок и в момент t в системе будет находится m заявок, при условии, что в момент 0 в системе было k заявок и одна уже обслуживалась на приборе. Тогда производная $r_{km}(n,t) = R'_{km}(n,t)$ имеет вид

$$r_{km}(n,t) = \sum_{i=1}^{m}\int\limits_0^t f_{ki}(n,y)\frac{[\lambda\,(t-y)]^{m-i}}{(m-i)!}\,e^{-\lambda(t-y)}\,[1-B(t-y)]\,dy, \quad n\ge 0, \quad m\ge 1.$$

Вводя тройное преобразование $\rho_k(z_1,s,z_2) = \sum\limits_{n=0}^{\infty}\sum\limits_{m=1}^{\infty} z_1^n z_2^m \int\limits_0^{\infty} e^{-st} r_{km}(n,t)\,dt$, получаем:

$$\rho_k(z_1,s,z_2) = \phi_k(z_1,s,z_2)\frac{1-\beta(s+\lambda-\lambda z_2)}{s+\lambda\,(1-z_2)} = \frac{z_2\,[z_2^k - \tilde{z}^k(z_1,s)]}{z_2\,(s+\lambda^- +\lambda-\lambda z_2) - \lambda^- z_1}. \tag{4}$$

2.2. Нестационарные характеристики

Для системы $M|M|1|\infty$ с параметрами (λ,λ^-) определим в терминах преобразований совместное нестационарное распределение числа обслуженных к моменту t заявок и числа заявок в системе в момент t при условии, что в начальный момент в системе было k, $k\ge 0$, заявок.

Пусть $W_{km}(n,t)$ — вероятность того, что к моменту t обслужено n заявок и в момент t в системе находится m заявок, при условии, что в начальный момент в системе было k, $k\ge 0$, заявок. Введём тройное преобразование

$$\omega_k(z_1,s,z_2) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} z_1^n z_2^m \int\limits_0^{\infty} e^{-st} W_{km}(n,t)\,dt. \tag{5}$$

Для нахождения $\omega_k(z_1,s,z_2)$ отдельно рассмотрим случаи, когда в начальный момент система свободна от заявок и когда в начальный момент в системе есть заявки.

Пусть в начальный момент система свободна от заявок. Тогда преобразование $\omega_0(z_1,s,z_2)$ можно представить в виде

$$\omega_0(z_1,s,z_2)=\omega_0^{(1)}(z_1,s,z_2)+\omega_0^{(2)}(z_1,s,z_2),$$

где

$$\omega_0^{(1)}(z_1,s,z_2)=\sum_{n=0}^{\infty}z_1^n\int_0^{\infty}e^{-st}W_{00}(n,t)\,dt,$$

$$\omega_0^{(2)}(z_1,s,z_2)=\sum_{n=0}^{\infty}\sum_{m=1}^{\infty}z_1^n z_2^m\int_0^{\infty}e^{-st}W_{0m}(n,t)\,dt.$$

Начнём с вычисления $\omega_0^{(1)}(z_1,s,z_2)$. Для того чтобы в момент t система была свободна, необходимо, чтобы в некоторый промежуточный момент y, $0<y<t$, система освободилась от заявок, а за оставшееся время $t-y$ в свободную систему не поступило ни одной заявки. Поэтому ([16, с. 306–307]) можно написать:

$$\omega_0^{(1)}(z_1,s,z_2)=\frac{1}{s+\lambda}\cdot\frac{1}{1-\frac{\lambda\tilde{z}(z_1,s)}{\lambda+s}}=\frac{1}{\lambda+s-\lambda\tilde{z}(z_1,s)}\,.$$

Перейдём к определению $\omega_0^{(2)}(z_1,s,z_2)$. Для того чтобы к моменту t было обслужено n заявок и в момент t в системе находилось m заявок, необходимо, чтобы в промежуточный момент y, $0<y<t$, в свободную систему поступила заявка, до момента y было обслужено $i=\overline{0,n}$ заявок, а за оставшееся время $(t-y)$ не закончился ПЗ, начавшийся в момент y, было обслужено $(n-i)$ заявок и в момент t в системе находилось m заявок. Значит ([16, с. 306–307]), можно написать:

$$\omega_0^{(2)}(z_1,s,z_2)=\rho_1(z_1,s,z_2)\cdot\frac{\frac{\lambda}{s+\lambda}}{1-\frac{\lambda\tilde{z}(z_1,s)}{\lambda+s}}=\frac{\lambda\rho_1(z_1,s,z_2)}{\lambda+s-\lambda\tilde{z}(z_1,s)}\,.$$

Общее выражение для $\omega_0(z_1,s,z_2)$ имеет вид

$$\omega_0(z_1,s,z_2)=\omega_0^{(1)}(z_1,s,z_2)+\omega_0^{(2)}(z_1,s,z_2)=\frac{1+\lambda\rho_1(z_1,s,z_2)}{\lambda+s-\lambda\tilde{z}(z_1,s)}\,. \tag{6}$$

Обратимся к случаю, когда в начальный момент в системе находится $k\geq 1$ заявок, одна из которых начинает обслуживаться. Функцию $\omega_k(z_1,s,z_2)$, $k\geq 1$, как и в случае $k=0$, можно представить в виде суммы двух слагаемых $\omega_k^{(1)}(z_1,s,z_2)$ и $\omega_k^{(2)}(z_1,s,z_2)$, причём первое слагаемое соответствует варианту, когда ПЗ, открываемый $k\geq 1$ заявками, закончится до момента t, а второе — не закончится.

Если до момента t закончится ПЗ, открываемый $k\geq 1$ заявками, то для того чтобы в момент t в системе было $m\geq 0$ заявок и за время t было обслужено n заявок, необходимо, чтобы до некоторого промежуточного момента $0<y<t$ система освободилась (т.е. закончились k одинаково распределенных ПЗ системы $M|M|1$ с параметрами (λ,λ^-)), за время y было обслужено $i\geq k$ заявок, за оставшееся время $t-y$ было обслужено $n-i$ заявок и в момент $t-y$ в системе оказалось $m\geq 0$ заявок. Учитывая, что двойное преобразование ПЗ системы

$M|M|1$ с параметрами (λ, λ^-) и числа обслуженных на нём заявок равно $\tilde{z}(z_1, s)$ ([16, с. 300]), преобразование $\omega_k^{(1)}(z_1, s, z_2)$ можно записать в виде

$$\omega_k^{(1)}(z_1, s, z_2) = \tilde{z}^k(z_1, s)\,\omega_0(z_1, s, z_2), \quad k \geq 1.$$

Если до момента t ПЗ, открываемый $k \geq 1$ заявками, не закончится, то преобразование $\omega_k^{(2)}(z_1, s, z_2)$ имеет вид

$$\omega_k^{(2)}(z_1, s, z_2) = \rho_k(z_1, s, z_2), \quad k \geq 1.$$

В итоге общее тройное преобразование числа обслуженных к моменту t заявок и числа заявок в момент t при условии, что в начальный момент в системе было $k \geq 1$ заявок, имеет вид

$$\omega_k(z_1, s, z_2) = \rho_k(z_1, s, z_2) + \tilde{z}^k(z_1, s)\,\omega_0(z_1, s, z_2), \quad k \geq 1. \tag{7}$$

3. ПЕРИОД ЗАНЯТОСТИ

Вернёмся к рассмотрению исходной системы. Для нахождения стационарного распределения времени ожидания начала обслуживания при рассматриваемой дисциплине обслуживания и выбивания заявок из накопителя необходимо предварительно найти распределение ПЗ исходной системы. Покажем, как это можно сделать.

Пусть $R_k(t, n)$, $k \geq 0$, $n \geq 0$, — вероятность того, что все заявки из накопителя будут обслужены до момента времени t и за это время в бункер перейдёт n заявок, вытесненных из накопителя, при условии, что в начальный момент на приборе начинает обслуживаться заявка из накопителя, а в накопителе остаётся k других заявок. Обозначим через $\delta_k(z, s) = \sum\limits_{n=0}^{\infty} z^n \int\limits_0^{\infty} e^{-st} dR_k(t, n)$ двойное преобразование распределения $R_k(t, n)$. Положим также $R(t, n) = R_0(t, n)$ и $\delta(z, s) = \delta_0(z, s)$.

Для нахождения $R_k(t, n)$ заметим, что заявки поступают в накопитель с интенсивностью λ, а пока в накопителе имеется хотя бы одна заявка, уходят из него (сразу на прибор или сначала в бункер, а потом уже когда-нибудь на прибор) с интенсивностью $\mu^* = \lambda^- + \mu$. При этом вероятность того, что заявка уйдёт из накопителя на прибор, минуя бункер, равна μ/μ^*, а вероятность того, что заявка сначала попадёт в бункер (будет "убита"), а затем на прибор — λ^-/μ^*. Таким образом, процесс прихода заявок в накопитель и ухода из него можно представить как процесс обслуживания заявок в СМО $M|M|1|\infty$ с параметрами (λ, μ^*), в которой обслуженная заявка с вероятностью λ^-/μ^* может быть "убита". Для такой системы совместное распределение ПЗ и числа "убитых" после обслуживания на нём заявок в терминах двойного преобразования выражается через преобразование $\gamma(z, s)$, определяемое формулой (1), следующим образом:

$$\sum_{n=1}^{\infty} \sum_{i=0}^{n} C_n^i z^i \left(\frac{\lambda^-}{\mu^*}\right)^i \left(\frac{\mu}{\mu^*}\right)^{n-i} \int\limits_0^{\infty} e^{-st} dG_n(t) = \gamma\Big(\frac{\lambda^-}{\mu^*} z + \frac{\mu}{\mu^*}, s\Big). \tag{8}$$

Заметим теперь, что распределение $R(t, n)$ совпадает с совместным распределением ПЗ системы $M|M|1|\infty$ с параметрами (λ, μ^*), открываемого единственной экспоненциально распределённой с параметром μ заявкой, и числа "убитых" после обслуживания на нём заявок и с учётом (2) имеет двойное преобразование

$$\delta(z, s) = \int\limits_0^{\infty} \mu e^{-\mu t} \delta\Big(\frac{\lambda^-}{\mu^*} z + \frac{\mu}{\mu^*}, s \Big| t\Big)\, dt = \frac{\mu}{\mu + s + \lambda\big[1 - \gamma\big(\frac{\lambda^-}{\mu^*} z + \frac{\mu}{\mu^*}, s\big)\big]}\,. \tag{9}$$

Если же в начальный момент в накопителе имеется хотя бы одна заявка, то время до того момента, когда в накопителе впервые останется ровно на одну заявку меньше, и число “убитых” за это время заявок совпадают с ПЗ, открываемым экспоненциально распределённой с параметром μ^* заявкой, и числом “убитых” на нём заявок для той же самой системы $M|M|1|\infty$ с параметрами (λ, μ^*). Поэтому

$$\delta_k(z,s) = \gamma_k\Big(\frac{\lambda^-}{\mu^*}z + \frac{\mu}{\mu^*}, s\Big)\delta(z,s) = \gamma^k\Big(\frac{\lambda^-}{\mu^*}z + \frac{\mu}{\mu^*}, s\Big)\delta(z,s). \tag{10}$$

Обозначим через $R_{km}(n,t)$, $k,m,n \ge 0$, вероятность того, что в СМО $M|M|1|\infty$ с параметрами (λ, λ^-) за промежуток времени t, распределенный по экспоненциальному закону с параметром μ^-, будет обслужено n заявок и в момент t в системе будет находиться m заявок, при условии, что в начальный момент в системе было k заявок. Тогда двойное преобразование $\delta_{km}(z,s) = \sum\limits_{n=0}^{\infty} z^n \int\limits_0^{\infty} e^{-st} dR_{km}(n,t)$, $k,m \ge 0$, можно записать в виде

$$\delta_{km}(z,s) = \sum_{n=0}^{\infty} z^n \int\limits_0^{\infty} e^{-st}\mu^- e^{-\mu^- t} W_{km}(n,t)\,dt, \quad m \ge 0, \quad k \ge 0.$$

В терминах преобразования $u_k(z,s,z_2) = \sum\limits_{m=0}^{\infty} z_2^m \delta_{km}(z,s)$ функция $\delta_{km}(z,s)$ с учётом (5) имеет вид

$$u_k(z,s,z_2) = \mu^-\omega_k(z, s+\mu^-, z_2), \quad k \ge 0. \tag{11}$$

Пусть $D_k(t,n)$, $k,n \ge 0$, — вероятность того, что все заявки из накопителя будут обслужены до момента времени t и за это время в бункер перейдёт n заявок, вытесненных из накопителя, при условии, что в начальный момент на приборе начинает обслуживаться заявка из бункера, а в накопителе остаётся k других заявок. Обозначим через $\tilde{\delta}_k(z,s) = \sum\limits_{n=0}^{\infty} z^n \int\limits_0^{\infty} e^{-st} dD_k(t,n)$ двойное преобразование вероятности $D_k(t,n)$. Напомним, что после окончания обслуживания очередной заявки на прибор сначала поступает заявка из накопителя (если он не пуст). Тогда

$$\tilde{\delta}_k(z,s) = \delta_{k0}(z,s) + \sum_{i=1}^{\infty}\delta_{ki}(z,s)\delta_{i-1}(z,s), \ k \ge 0,$$

что с учётом (4), (6), (7) и (9)–(11) после несложных, но утомительных преобразований приводится к виду

$$\begin{aligned}\tilde{\delta}_k(z,s) = \tilde{z}^k(z, s+\mu^-)\Bigg[&\frac{\mu^-}{\lambda + s + \mu^- - \lambda\tilde{z}(z, s+\mu^-)} - \\ &- \frac{\mu^-\delta(z,s)}{\big[s+\mu^-+\lambda^-+\lambda-\lambda\gamma\big(\frac{\lambda^-}{\mu^*}z+\frac{\mu}{\mu^*},s\big)\big]\gamma\big(\frac{\lambda^-}{\mu^*}z+\frac{\mu}{\mu^*},s\big)-\lambda^- z} + \\ &+ \frac{\lambda\mu^-\delta(z,s)\rho_1\Big(z_1, s+\mu^-, \gamma\Big(\frac{\lambda^-}{\mu^*}z+\frac{\mu}{\mu^*},s\Big)\Big)}{\gamma\big(\frac{\lambda^-}{\mu^*}z+\frac{\mu}{\mu^*},s\big)\big[\lambda+s+\mu^- - \lambda\tilde{z}(z,s+\mu^-)\big]}\Bigg] + \\ &+ \frac{\mu^-\delta(z,s)\big[\gamma\big(\frac{\lambda^-}{\mu^*}z+\frac{\mu}{\mu^*},s\big)\big]^k}{\big[s+\mu^-+\lambda^-+\lambda-\lambda\gamma\big(\frac{\lambda^-}{\mu^*}z+\frac{\mu}{\mu^*},s\big)\big]\gamma\big(\frac{\lambda^-}{\mu^*}z+\frac{\mu}{\mu^*},s\big)-\lambda^- z}, \quad k \ge 0.\end{aligned} \tag{12}$$

Наконец, после нахождения вспомогательных вероятностей можно выписать выражения для распределения длительности ПЗ исходной СМО при различных условиях, которые обусловлены тем, что заявки из накопителя и бункера обслуживаются с различными интенсивностями.

Поскольку рассматривается дисциплина, при которой отрицательная заявка выбивает последнюю заявку из очереди в накопителе, а на прибор выбирается последняя заявка из очереди в накопителе или в бункере, то время ожидания начала обслуживания заявки из бункера зависит только от числа заявок в накопителе и на приборе и числа заявок, стоящих в очереди в бункере позади неё, но не зависит от числа заявок, стоящих в бункере впереди. Значит, время от момента перехода заявки в бункер до момента поступления её на прибор совпадает с ПЗ исходной системы, открываемого всеми теми заявками, которые находятся в момент перехода только в накопителе и на приборе.

Итак, пусть $F(x)$ — ФР длительности ПЗ исходной СМО при условии, что в начальный момент на приборе начинает обслуживаться заявка из бункера, а накопитель и бункер пусты. Обозначим через $\tilde{\gamma}(s)$ её ПЛС. Продолжительность рассматриваемого ПЗ является суммой $\nu + 1$ случайных величин, первая из которых — сумма времени обслуживания заявки на приборе и времени до полного освобождения накопителя от заявок, а остальные описывают ПЗ, каждый из которых имеет распределение $F(x)$. Случайная величина ν равна числу заявок, поступивших в бункер за время до опустошения накопителя. Совместное распределение времени до освобождения накопителя от заявок и числа заявок, перешедших за это время в бункер, при условии, что в начальный момент на приборе начинает обслуживаться заявка из бункера, а в накопителе нет других заявок, задаётся тройным преобразованием $\tilde{\tilde{\delta}}_0(z, s)$ по формуле (12). Поэтому ПЛС $\tilde{\gamma}(s)$ удовлетворяет функциональному уравнению

$$\tilde{\gamma}(s) = \tilde{\tilde{\delta}}_0(\tilde{\gamma}(s), s). \tag{13}$$

Обозначим через $F_k(x)$ ФР длительности ПЗ исходной СМО при условии, что в начальный момент на приборе начинает обслуживаться заявка из бункера, в накопителе остаётся $k \geq 1$ других заявок и бункер пуст, а через $\tilde{\gamma}_k(s)$ — ее ПЛС. Используя аналогичные рассуждения, приходим к выражению

$$\tilde{\gamma}_k(s) = \tilde{\tilde{\delta}}_k(\tilde{\gamma}(s), s), \quad k \geq 1. \tag{14}$$

Наконец, пусть $\gamma_k(s)$ — ПЛС длительности ПЗ исходной СМО при условии, что в начальный момент на приборе начинает обслуживаться заявка из накопителя, в накопителе остается $k \geq 0$ других заявок, а бункер пуст. Тогда имеет место равенство

$$\gamma_k(s) = \delta_k(\tilde{\gamma}(s), s), \quad k \geq 0. \tag{15}$$

Поскольку ПЗ исходной системы открывается приходящей в пустую систему заявкой, которая, очевидно, поступает на прибор из накопителя, то именно последнее равенство при $k = 0$ задаёт ПЛС ПЗ исходной системы. Ввиду того, что $\tilde{\gamma}(s)$ находится из функционального уравнения (13), не удаётся обратить (15) и получить распределение ПЗ в явном виде. Однако с помощью (15) можно вычислять моменты ПЗ.

4. СТАЦИОНАРНОЕ РАСПРЕДЕЛЕНИЕ ВРЕМЕНИ ОЖИДАНИЯ НАЧАЛА ОБСЛУЖИВАНИЯ

Перейдём к нахождению стационарного распределения времени ожидания начала обслуживания заявки. Далее будут использованы результатами работы [14].

Пусть p_0 — стационарная вероятность того, что система свободна, а $p_{i,j,m}$, $i, j \geq 0$, $m = 1, 2$, — стационарная вероятность того, что в накопителе находится i заявок, в бункере ожидают j заявок, вытесненных из накопителя, и на приборе обслуживается заявка либо из накопителя ($m = 1$), либо из бункера ($m = 2$). В [14] данные вероятности найдены в терминах двойных производящих функций

$$P(u,v) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} p_{i,j,1}u^i v^j. \quad N(u,v) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} p_{i,j,2}u^i v^j. \tag{16}$$

Конечные формулы для производящих функций $P(u,v)$ и $N(u,v)$ из [14] здесь не приводятся, поскольку они являются достаточно громоздкими и для понимания сути дальнейших рассуждений и выводов значения не имеют.

Ясно, что время ожидания начала обслуживания заявки зависит от того, какая заявка была на приборе в момент её поступления в систему.

Найдём сначала стационарную вероятность $V_1(x)$ того, что поступившая в систему заявка, заставшая на приборе заявку из накопителя, будет ожидать начала обслуживания время меньше x. Заметим, что независимо от числа заявок в накопителе перед поступившей заявкой и времени их обслуживания, а также от времени обслуживания заявки на приборе и всех заявок, поступивших в систему после выделенной, но обслуженных или выбитых в бункер раньше момента перехода выделенной заявки на прибор или в бункер, вероятность выделенной заявке поступить сразу из накопителя на прибор равна μ/μ^*, а сначала в бункер и затем уже из бункера на прибор — λ^-/μ^*. Однако, если поступающая в систему заявка не попадает сразу же на прибор, то за то время, пока она находится в накопителе, обслужатся или перейдут в бункер все заявки, пришедшие за время обслуживания заявки на приборе, и все их потомки (напомним, что заявки, находящиеся в накопителе до прихода выделенной, на время её пребывания в накопителе не влияют). Можно показать (см., например, [12, раздел 5]), что распределение времени от момента прихода заявки в (непустую) систему и до момента поступления её на прибор или перехода в бункер имеет ПЛС $\gamma(1,s)$. Тогда, учитывая (15), выражение для стационарной вероятности $V_1(x)$ в терминах ПЛС $\varphi_1(s) = \int\limits_0^\infty e^{-sx}dV_1(x)$ примет вид

$$\varphi_1(s) = \gamma(1,s)\frac{\mu}{\mu^*}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} p_{1,i,\cdot} + \gamma(1,s)\frac{\lambda^-}{\mu^*}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \gamma_i(s)p_{1,i,\cdot},$$

что с учётом (15) и (16) можно записать следующим образом:

$$\varphi_1(s) = \frac{\gamma(1,s)}{\mu^*}\Big[\mu P(1,1) + \lambda^-\delta(\tilde\gamma(s),s)\,P\Big(\gamma\Big(\frac{\lambda^-\tilde\gamma(s)+\mu}{\mu^*},s\Big),1\Big)\Big]. \tag{17}$$

Теперь найдём стационарную вероятность $V_2(x)$ того, что поступившая в систему заявка и заставшая на приборе заявку из бункера, будет ожидать начала обслуживания время меньше x. Здесь нужно рассмотреть два случая: когда заявка из бункера, находящаяся на приборе в момент поступления выделенной заявки, не успеет обслужиться до момента ухода выделенной заявки в бункер, и когда успеет.

Начнём с первого случая, т.е. с нахождения стационарной вероятности $V_{2,1}(x)$ того, что поступившая в систему заявка и заставшая на приборе заявку из бункера, которая не обслужилась до момента “убийства” поступившей заявки, будет ожидать начала обслуживания время меньше x. Обозначим через $Q(t)$ распределение ПЗ системы $M|M|1|\infty$ с параметрами (λ,λ^-). ПЛС распределения $Q(t)$ имеет вид

$$\gamma(s) = \tilde z(1,s) = \frac{\lambda+\lambda^-+s-\sqrt{(\lambda+\lambda^-+s)^2-4\lambda\lambda^-}}{2\lambda}.$$

Будем считать, что в момент поступления в систему заявка открывает ПЗ СМО $M|M|1|\infty$ с параметрами (λ, λ^-), с помощью которой можно описать процесс поступления положительных и отрицательных заявок в исходную систему до того момента, когда выделенная заявка перейдет либо в бункер, либо на прибор. Если этот ПЗ закончится в момент t и до этого момента не произойдёт ни одного обслуживания заявки на приборе (плотность вероятностей этого события равна $e^{-\mu^- t}Q'(t)$), то это будет означать, что выделенная заявка перешла в бункер и оставила после своего ухода в накопителе и на приборе те же самые заявки, которые были в момент её поступления в систему. Поэтому выражение для стационарной вероятности $V_{2,1}(x)$ в терминах ПЛС можно записать в виде

$$\varphi_{2,1}(s) = \int\limits_0^\infty e^{-st} e^{-\mu^- t} dQ(t) \sum_{i=0}^{\infty} \tilde{\gamma}_i(s)\, p_{2,i,\cdot} = \tilde{z}(1, s+\mu^-) \sum_{i=0}^{\infty} \tilde{\delta}_i(\tilde{\gamma}(s), s) p_{2,i,\cdot},$$

что с учётом (12), (13) и (16) даёт

$$\varphi_{2,1}(s) = \left[\tilde{\gamma}(s) - \frac{\mu^- \delta(\tilde{\gamma}(s), s)}{\left[s+\mu^- + \lambda^- + \lambda - \lambda\gamma\left(\frac{\lambda^- \tilde{\gamma}(s)+\mu}{\mu^*}, s\right)\right] \gamma\left(\frac{\lambda^- \tilde{\gamma}(s)+\mu}{\mu^*}, s\right) - \lambda^- \tilde{\gamma}(s)}\right] \times$$
$$\times \tilde{z}(1, s+\mu^-)\, N(\tilde{z}(\tilde{\gamma}(s), s+\mu^-), 1) +$$
$$+ \frac{\mu^- \delta(\tilde{\gamma}(s), s)\, \tilde{z}(1, s+\mu^-)\, N\left(\gamma\left(\frac{\lambda^- \tilde{\gamma}(s)+\mu}{\mu^*}, s\right), 1\right)}{\left[s+\mu^- + \lambda^- + \lambda - \lambda\gamma\left(\frac{\lambda^- \tilde{\gamma}(s)+\mu}{\mu^*}, s\right)\right] \gamma\left(\frac{\lambda^- \tilde{\gamma}(s)+\mu}{\mu^*}, s\right) - \lambda^- \tilde{\gamma}(s)}. \quad (18)$$

Перейдём ко второму случаю, т.е. к нахождению стационарной вероятности $V_{2,2}(x)$ того, что поступившая в систему заявка и заставшая на приборе заявку из бункера, которая обслужилась до момента “убийства” поступившей заявки, будет ожидать начала обслуживания время меньше x. Введём вспомогательную величину $h_k(t),\ k \geq 1$, — вероятность того, что ПЗ СМО $M|M|1|\infty$ с параметрами (λ, λ^-) не закончится до момента t и в этот момент в системе будет k заявок. Двойное преобразование (ПЛС и ПФ) данной вероятности с учётом (4) имеет вид

$$h(s,z) = \sum_{k=1}^{\infty} z^k \int\limits_0^\infty e^{-st} h_k(t)\, dt = \rho_1(1, s, z).$$

Как и в предыдущем случае, поступающая в систему заявка и застающая на приборе заявку из бункера и в накопителе i других заявок (с вероятностью $p_{2,i,\cdot}$), открывает ПЗ системы $M|M|1|\infty$ с параметрами (λ, λ^-). Поскольку теперь заявка, которую застала на приборе выделенная заявка в момент своего прихода, успевает обслужиться до момента “убийства” выделенной заявки, то за это время открытый ПЗ не должен закончиться, и в момент окончания обслуживания заявки на приборе, за выделенной заявкой может оказаться $k \geq 0$ заявок. Если за выделенной окажется 0 заявок, то (в силу дисциплины обслуживания) она сразу же поступит на прибор. Если же за выделенной заявкой окажется $k \geq 1$ других заявок, то прежде чем выделенная заявка перейдет на прибор (с вероятностью μ/μ^*) или в бункер (с вероятностью λ^-/μ^*), должны окончиться $(k-1)$ ПЗ системы $M|M|1|\infty$ с параметрами (λ, μ^*). Наконец, если заявка попала в бункер, то будет ожидать начала обслуживания дополнительное время, равное длительности ПЗ исходной СМО при условии, что в начальный момент на приборе начинает обслуживаться заявка из накопителя, в накопителе остаётся $i \geq 0$ других заявок, а бункер пуст, ПЛС которого задается равенством (15). Теперь, используя свойства ПЛС и ПФ и применяя формулу полной вероятности, вероятность $V_{2,2}(x)$ в терминах ПЛС $\varphi_{2,2}(s) = \int\limits_0^\infty e^{-sx} dV_{2,2}(x)$

можно записать в виде

$$w_{2,2}(s) = \sum_{i=0}^{\infty} p_{2,i,\cdot} \int_0^{\infty} e^{-st}\mu^- e^{-\mu^- t}\left(h_1(t) + \sum_{k=2}^{\infty} h_k(t)\,\gamma_{k-1}(1,s)\left[\frac{\mu}{\mu^*} + \frac{\lambda^-}{\mu^*}\gamma_i(s)\right]\right)dt.$$

Введём следующие обозначения:

$$\tilde{h}_1(s) = \int_0^{\infty} e^{-st}h_1(t)\,dt, \qquad \tilde{h}(z,s) = \sum_{k=2}^{\infty} z^{k-1}\int_0^{\infty} e^{-st}h_k(t)\,dt.$$

Нетрудно видеть, что

$$\tilde{h}_1(s) = \frac{1}{z}\rho_1(1,s,z)\Big|_{z=0} = \frac{\gamma(s)}{\lambda^-}, \qquad \tilde{h}(z,s) = \frac{1}{z}h(s,z) - \tilde{h}_1(s).$$

Тогда с учетом введённых обозначений выражение для $w_{2,2}(s)$ можно записать в виде

$$w_{2,2}(s) = \mu^-\left[\tilde{h}_1(s+\mu^-) + \frac{\mu}{\mu^*}\tilde{h}(\gamma(1,s), s+\mu^-)\right]N(1,1) + \\ + \frac{\lambda^-\mu^-}{\mu^*}\tilde{h}(\gamma(1,s), s+\mu^-)\,\delta(\tilde{\gamma}(s),s)\,N\Big(\gamma\Big(\frac{\lambda^-\tilde{\gamma}(s)+\mu}{\mu^*}, s\Big), 1\Big).$$

Теперь можно получить формулу для стационарного распределения $V(x)$ общего времени ожидания начала обслуживания поступившей в систему заявки. Учитывая, что, приходя в свободную систему, заявка сразу же попадает на прибор (время ожидания начала её обслуживания равно нулю), в терминах ПЛС $\varphi(s) = \int_0^{\infty} e^{-sx}dV(x)$ это распределение имеет вид

$$\varphi(s) = p_0 + w_1(s) + w_{2,1}(s) + w_{2,2}(s).$$

Дифференцируя последнюю формулу соответствующее число раз, можно получить выражения для моментов любого порядка стационарного распределения времени ожидания начала обслуживания заявки.

5. ЗАКЛЮЧЕНИЕ

В настоящей статье рассматривается неклассическая марковская система массового обслуживания с отрицательными заявками и бункером для вытесненных заявок, в которой заявки из накопителя и бункера обслуживаются с различной интенсивностью. Предложен метод нахождения (в терминах ПЛС) стационарного распределения времени ожидания начала обслуживания поступившей в систему заявки при условии, что отрицательная заявка выбивает последнюю заявку из очереди в накопителе, а на прибор после окончания обслуживания очередной заявки выбирается последняя заявка из очереди в накопителе или последняя заявка из очереди в бункере (если накопитель пуст). Показано, как с помощью предложенного метода можно также найти (в терминах ПЛС) распределение ПЗ исходной системы. Оказывается, что, как и для систем типа $M|G|1$, ПЛС ПЗ не выражается в явном виде, а задаётся функциональным уравнением.

Все полученные аналитические результаты были проверены путём сравнения с результатами работы имитационной модели, разработанной на языке GPSS.

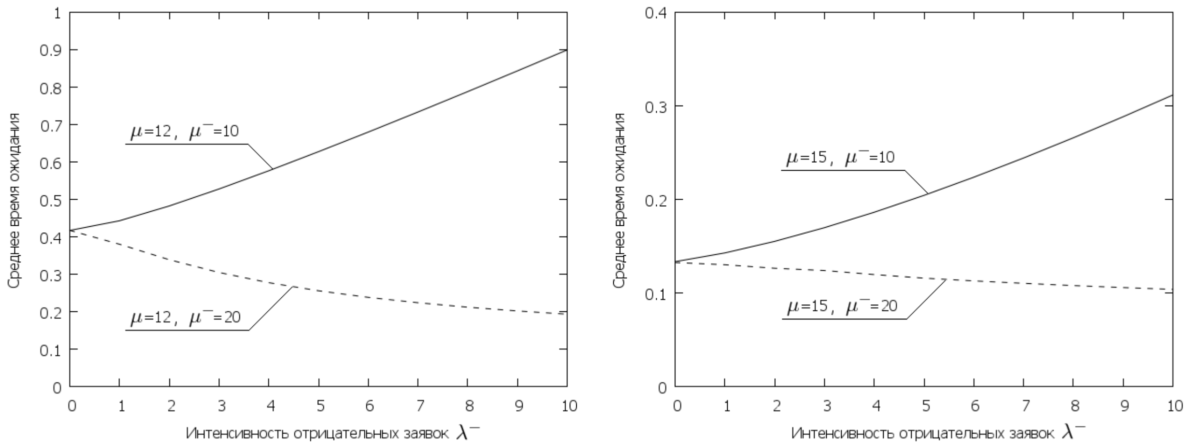

Рис. 1. Зависимость среднего времени ожидания от интенсивности λ^- потока отрицательных заявок.

В качестве иллюстрации на рис. 1 приведены результаты расчёта среднего времени ожидания начала обслуживания заявки при различных значениях начальных параметров. Во всех расчётах интенсивность потока положительных заявок принималась равной $\lambda = 10$.

В заключении авторы хотели бы отметить, что интересным, в первую очередь с теоретической точки зрения, предметом дальнейших исследований являются задачи нахождения стационарного распределения времени ожидания начала обслуживания при других дисциплинах выбивания и обслуживания заявок из накопителя и бункера, а также сравнение между собой временных характеристик обслуживания заявок при различных дисциплинах.

СПИСОК ЛИТЕРАТУРЫ

1. Klimenok V., Dudin A. A BMAP/PH/N Queue with Negative Customers and Partial Protection of Service. *Special Issue: Sixth St. Petersburg Workshop on Simulation*, 2012, vol. 41, issue 7, pp. 1062–1082.
2. Печинкин А.В., Разумчик Р.В. Об одном методе расчёта стационарного распределения очереди в системе массового обслуживания с потоками обычных и отрицательных заявок и бункером для выбитых заявок. *Информационные процессы*, т. 12, № 1, 2012, стр. 53–67.
3. Tien Van Do. Bibliography on G-networks, negative customers and applications. *Mathematical and Computer Modelling*, 2011, vol. 53, issues 1–2, pp. 205–212.
4. Van Dijk N. Queueing Networks - A Fundamental Approach. International Series in Operations Research and Management Science (Vol. 154). Ed. R.J. Boucherie and N.M. van Dijk: Springer, 2011.
5. Dimitriou I. A mixed priority retrial queue with negative arrivals, unreliable server and multiple vacations. *Applied Mathematical Modelling*, 2012.
6. Dimitriou I. A preemptive resume priority retrial queue with state dependent arrivals, unreliable server and negative customers. *TOP: An Official Journal of the Spanish Society of Statistics and Operations Research*, 2011, pp. 1–30.
7. Krishna Kumar B., Pavai Madheswari S., Anantha Lakshmi S. R. An M/G/1 Bernoulli feedback retrial queueing system with negative customers. *Journal of Operational Research*, 2011, pp. 1–24.
8. Wang J., Huang Y., Dai Z. A discrete-time on-off source queueing system with negative customers. *Journal Computers and Industrial Engineering*, 2011, vol. 61, issue 4, pp. 1226–1232.
9. Jia S., Chen Y. A discrete time queueing system with negative customers and single working vacation. *3rd International Conference on Computer Research and Development (ICCRD)*, 2011, vol. 4, pp. 15–19.

10. Ayyappan G., Muthu Ganapathi Subramanian, Sekar G. Study of influences of negative arrival for single server retrial queueing system with Coxian phase type service. *Proceedings of the 6th International Conference on Queueing Theory and Network Applications*, 2011, pp. 53–59.

11. Мандзо Р., Касконе Н., Разумчик Р.В. Экспоненциальная система массового обслуживания с отрицательными заявками и бункером для вытесненных заявок. *Автоматика и телемеханика*, 2008, № 9, стр. 103–113.

12. Печинкин А.В., Разумчик Р.В. О временных характеристиках в экспоненциальной системе массового обслуживания с отрицательными заявками и бункером для вытесненных заявок. *Автоматика и телемеханика*, 2011, № 12, стр. 75–90.

13. Печинкин А.В., Разумчик Р.В. О времени ожидания при некоторых дисциплинах обслуживания в системе с отрицательными заявками и бункером. *Proceedings of the International Conference "Modern Probabilistic Methods for Analysis and Optimization of Information and Telecommunication Networks"*, 2011, pp. 207–212.

14. Разумчик Р.В. Система массового обслуживания с отрицательными заявками, бункером для вытесненных заявок и различными интенсивностями обслуживания. *Информатика и ее применения*, 2011, том 5, вып. 3, стр. 39–43.

15. Bocharov P.P., D'Apice C., Pechinkin A.V., Salerno S. *Queueing Theory.* Utrecht: VSP Publishing, 2003.

16. Бочаров П.П., Печинкин А.В. *Теория массового обслуживания.* М.: Изд-во РУДН, 1995.

Stationary waiting time distribution in queueing system with negative customers and bunker for ousted customers under LAST-LIFO-LIFO service discipline

Pechinkin A.V., Razumchik R.V.

Consideration is given to the single server queueing system with Poisson input flows of ordinary and negative customers. An arriving ordinary customer occupies one place in an infinite buffer. Negative customer upon arrival pushes out one ordinary customer from the queue in the buffer to another queue (bunker) and leaves the system. Customers from bunker are served with relative priority. Service times of customers from buffer and bunker are both exponentially distributed but with different rates. It is assumed that negative customer always pushes out the last customer in the queue in the buffer and after service completion the next customer to enter server is the last customer in the queue in the buffer (or, if it empty, the last customer in the queue in the bunker).

It is shown how to obtain stationary waiting time distribution of an arriving ordinary customer in terms of Laplace-Stieltjes transform.

KEYWORDS: queueing system, negative customers, waiting time, busy period.