М. И. Волковинский, А. Н. Кабалевский, Обслуживание со смешанным приоритетом в системах с отерями на переключение, *Автомат. и телемех.*, 1975, выпуск 11, 16–22

# *Стохастические системы*

УДК 62-501:519.283

## ОБСЛУЖИВАНИЕ СО СМЕШАННЫМ ПРИОРИТЕТОМ В СИСТЕМАХ С ПОТЕРЯМИ НА ПЕРЕКЛЮЧЕНИЕ

**М. И. ВОЛКОВИНСКИЙ, А. Н. КАБАЛЕВСКИЙ**

**(Москва)**

Рассмотрена система обслуживания со смешанным приоритетом с двумя типами требований (приоритетные и неприоритетные) и одним обслуживающим прибором с учетом времени переключения при прерывании обслуживания неприоритетного требования и при возвращении к его обслуживанию. Находится распределение виртуальных времен обслуживания в системе и производящая функция вероятностей совместного распределения длин очередей.

### 1. Введение и постановка задачи

Настоящая работа, посвященная исследованию приоритетной дисциплины обслуживания со смешанным приоритетом, является обобщением работ [1, 2]. В [1] было подробно рассмотрено обслуживание со смешанным приоритетом, однако потери на переключение при этом не учитывались. Статья [2], напротив, была посвящена именно учету потерь на переключение, но в системе с абсолютным приоритетом.

Дана однолинейная система обслуживания. На вход системы поступает два пуассоновских потока заявок с интенсивностями $\lambda_1$ и $\lambda_2$. Первый поток заявок называется приоритетным, второй — неприоритетным. Времена обслуживания заявок ($s_1$ и $s_2$), а также времена перехода от обслуживания неприоритетных заявок к приоритетным при прерывании ($s_{21}$) и от приоритетных к неприоритетным при окончании прерывания ($s_{12}$) являются непрерывными и независимыми в совокупности случайными величинами с плотностями соответственно $s_1(x)$, $s_2(x)$, $s_{21}(x)$ и $s_{12}(x)$. Дисциплина обслуживания в системе — смешанный приоритет.

Дисциплина со смешанным приоритетом является комбинацией дисциплины обслуживания с абсолютным и относительным приоритетами. Абсолютный приоритет реализуется, если время обслуживания неприоритетного требования меньше заданной величины $\theta$. В противном случае имеет место относительный приоритет. В данной работе рассматривается абсолютный приоритет с дообслуживанием.

Как и в работе [2], ниже рассматриваются потери времени на переключения, которые связаны с защитой прерываемой задачи при переключении $2\to1$ и с восстановлением прерванной задачи при окончании прерывания, когда имеет место переключение $1\to2$. В связи с этим переключение $2\to1$ не прерывается приходящими в систему приоритетными заявками, которые становятся в очередь, а переключение $1\to2$ может прерываться, что приводит к необходимости рассматривать вместо периода $s_{12}$ соответствующий цикл обслуживания.

Рассмотрение дисциплины обслуживания со смешанным приоритетом представляет известный практический интерес, поскольку такие дисциплины при оптимальном выборе $\theta$ в принципе обеспечивают не худшее обслуживание по сравнению с «чистыми» дисциплинами (с абсолютным и относительным приоритетом), так как при $\theta=\infty$ мы получаем обслуживание с абсолютным приоритетом, а при $\theta=0$ — относительный приоритет.

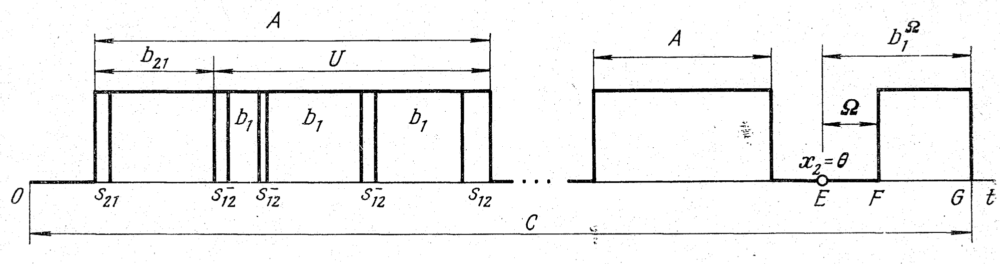
Как и в [2], требуется исследовать основные характеристики процесса обслуживания — виртуальные времена обслуживания и совместное распределение длин очередей.

В работе приняты следующие обозначения: временные переменные обозначены буквами $t$ и $\tau$, точки комплексной плоскости в преобразовании Лапласа — $z_1$ и $z_2$, оригиналы и соответствующие изображения Лапласа различаются лишь аргументами. Величины, относящиеся к приоритетным требованиям, снабжены нижним индексом 1, а к неприоритетным — индексом 2. Верхний индекс указывает на период, к которому относится соответствующая характеристика.

Основной процесс [1, 2], сводящийся к поочередному обслуживанию приоритетных и неприоритетных требований, состоит из чередующихся периодов занятости и незанятости обслуживающего прибора. Период занятости, который подробно рассматривается в [1, 2], в свою очередь, состоит из периодов обслуживания приоритетных требований и циклов обслуживания неприоритетных требований. Как и в работах [1, 2], будем рассматривать процесс обслуживания сначала на циклах обслуживания, затем на периодах занятости и, наконец, методами теории восстановления получим характеристики основного процесса. Указанный метод был подробно рассмотрен в [1] и, применительно к системам с потерями на переключения, в [2]. Как будет показано ниже, исследование системы сведется главным образом к нахождению ее характеристик на цикле обслуживания.

## 2. Процесс на цикле обслуживания неприоритетного требования

На рисунке показана одна из выборочных реализаций процесса на цикле обслуживания. Цикл обслуживания всегда начинается с обслуживания неприоритетной заявки. Если во время ее обслуживания приходит приоритетное требование, то по правилам, которые описывались выше, может произойти прерывание. Длительность периода прерывания $A$ является случайной величиной с плотностью $A(t)$, а сам период прерывания




состоит из начального периода занятости $b_{21}$ и цикла обслуживания $U$ переключения $1\to2$. Характеристики процесса на указанных периодах были получены в [2].

Момент $E$ на рисунке соответствует моменту времени, когда суммарное время обслуживания неприоритетного требования становится равным $\theta$. После момента $E$ обслуживание неприоритетного требования уже не прерывается и заканчивается в момент времени $F$.

Поскольку в промежутке времени $EF$ прерывание задачи не имеет места, переход к обслуживанию накопившейся за время $EF$ приоритетной очереди происходит без потери времени. Момент $G$ является моментом окончания обслуживания последнего приоритетного требования и, следовательно, концом цикла обслуживания неприоритетного требования $C$. В силу сказанного выше отрезок времени $EG$ является начальным периодом занятости для изолированной приоритетной очереди $b_1^{\Omega}$. Отрезок $EF$, обозначаемый $\Omega$, является входящим в $EG$ дополнительным периодом занятости.

На цикле обслуживания $C$ естественно выделить следующие участки — $I'$ и $I''$, когда неприоритетное требование находится в системе и суммарное время его обслуживания соответственно меньше $\theta$ и больше $\theta$, и участок $L$, когда обслуживание неприоритетного требования закончилось и идет обслуживание оставшихся приоритетных требований. Суммарное время $I'$ и $I''$ будем обозначать через $I$.

Сравнивая структуру цикла обслуживания $C$ для рассматриваемого процесса и соответствующую структуру для процесса без потерь на переключение [1], нетрудно убедиться, что эти процессы отличаются только длительностью периодов прерывания в обслуживании неприоритетного требования ($A$ для рассматриваемого процесса и $b_1$ для процесса в [1]). Поэтому формулы для фазы $I'$ рассматриваемого процесса можно получить из формул, приведенных в [1], заменой характеристик на периоде $b_1$ на соответствующие характеристики на периоде $A$, а формулы для фазы $I''$ и $L$ использовать без изменений. Характеристики на периоде прерывания $A$ получены в работе [2].

В [1] (стр. 185) получено выражение для преобразования Лапласа плотности виртуального времени

$$W_1^c(z_1, z_2) = \int_0^\theta f^I(0, x_2, z_1)\, dx_2 \,[1 + \lambda_1 W_1^{b_1}(z_1, z_2)] + \\ + f^I(0, \theta, z_1)\, W_1^\Omega(z_1, z_2). \tag{1}$$

С учетом времени переключения $s_{21}$ и с упомянутой выше заменой характеристик, относящихся к периоду $b_1$, получаем

$$W_1^c(z_1, z_2) = \int_0^\theta f^I(0, x_2, z_1)\, [s_{21}(z_1) + \lambda_1 W_1^A(z_1, z_2)]\, dx_2 + \\ + f^I(0, \theta, z_1)\, W_1^\Omega(z_1, z_2), \tag{2}$$

где $f^I(0, x_2, z_1)$ есть преобразование Лапласа плотности вероятности того, что на обслуживании находится неприоритетное требование, которое уже обслуживается время $x_2$. Для $f^I(0, x_2, z_1)$ из [1] с заменой характеристик на $b_1$ имеем

$$f^I(0, x_2, z_1) = \\ = \exp\left(-\{\lambda_1[1 - A(z_1)] + z_1\}\, x_2 - \int_0^{x_2} \eta_2(x)\, dx\right), \quad x_2 \leqslant \theta. \tag{3}$$

Из [2] находим

$$A(z_1) = b_{21}(z_1)\, U(z_1), \tag{4}$$

$$b_{21}(z_1) = b_1(z_1)\, s_{21}(\lambda_1[1 - b_1(z_1)] + z_1), \tag{5}$$

$$b_1(z_1) = s_1(\lambda_1[1 - b_1(z_1)] + z_1). \tag{6}$$

Здесь $b_1(z_1)$ — преобразование Лапласа плотности вероятности $b_1(t)$ периода занятости $b_1$ для изолированного приоритетного процесса.

В [1] (стр. 103) приводится выражение для длительности цикла обслуживания переключения $1 \to 2$:

$$U(z_1) = \int_0^\infty \frac{\exp(-[\lambda_1 + z_1]\, y)\, s_{12}(y)\, dy}{1 - \dfrac{\lambda_1}{\lambda_1 + z_1}\, b_1(z_1)\, [1 - \exp(-[\lambda_1 + z_1]\, y)]}. \tag{7}$$

Входящая в формулу (3) $\eta_2(x)$ есть условная плотность окончания обслуживания неприоритетного требования в момент времени $x$, при условии, что за время $x$ требование не обслуживалось, т. е.

$$\eta_2(x) = \frac{s_2(x)}{1 - \int\limits_0^x s_2(y)\,dy}. \tag{8}$$

Виртуальное время $W_1^c$ в (2) выражается через $W_1^A$ и $W_1^\Omega$. Выражения, определяющие преобразование Лапласа плотности виртуального времени $W_1^A$, были получены в [2]:

$$W_1^A(z_1, z_2) = W_1^{b_{21}}(z_1, z_2) + b_{21}(z_1)\,W_1^u(z_1, z_2), \tag{9}$$

где

$$W_1^{b_{21}}(z_1, z_2) = \frac{s_1(z_2)\,s_{21}(z_2) - b_1(z_1)\,s_{21}(\lambda_1[1 - b_1(z_1)] + z_1)}{z_1 - z_2 - \lambda_1[1 - s_1(z_2)]}, \tag{10}$$

$$W_1^u(z_1, z_2) = \int\limits_0^\infty \frac{[1 - \exp(-[\lambda_1 + z_1]\,y)]\,s_{12}(y)\,dy}{\lambda_1 + z_1 - \lambda_1 b_1(z_1)\,[1 - \exp(-[\lambda_1 + z_1]\,y)]} \times \tag{11}$$

$$\times\,[1 + \lambda_1 W_1^{b_1}(z_1, z_2)],$$

где $W_1^{b_1}$ — виртуальное время для изолированной приоритетной очереди на периоде занятости $b_1$,

$$W_1^{b_1}(z_1, z_2) = \frac{b_1(z_1) - s_1(z_2)}{z_2 - z_1 - \lambda_1[1 - s_1(z_2)]}.$$

Выражение для преобразования плотности виртуального времени $W_1^\Omega$, относящееся к периоду $L$, находим в [1] (стр. 32):

$$W_1^\Omega(z_1, z_2) = \frac{\Omega(z_1 + \lambda_1[1 - b_1(z_1)]) - \Omega(z_2)}{z_2 - z_1 - \lambda_1[1 - s_1(z_2)]}. \tag{12}$$

Таким образом, $W_1^c$ полностью определено.

Аналогичным образом находим на цикле $C$ производящую функцию $\Pi^c(\alpha_1, \alpha_2, z_1)$ совместного распределения вероятностей длин очередей приоритетных и неприоритетных требований. Выражение для $\Pi^c(\alpha_1, \alpha_2, z_1)$ получено в [1] (стр. 182):

$$\Pi^c(\alpha_1, \alpha_2, z_1) = \alpha_2\Pi^I(\alpha_1, z_1) + \Pi^L(\alpha_1, z_1). \tag{13}$$

Входящая в это соотношение производящая функция $\Pi^I$ на периоде $I$ получается из [1] с заменой, как и выше, характеристик, относящихся к периоду $b_1$,

$$\Pi^I(\alpha_1, z_1) = \int\limits_0^\theta f^I(0, x_2, z_1)\,dx_2\,[1 + \lambda_1\Pi^A(\alpha_1, z_1)] + \tag{14}$$

$$+ \frac{f^I(0, \theta, z_1)}{F_2^c(\theta)} \int\limits_\theta^\infty F_2^c(x_2)\exp(-[\lambda_1(1 - \alpha_1) + z_1]\,[x_2 - \theta])\,dx_2,$$

где

$$F_2^c(x) = \int\limits_x^\infty s_2(x_2)\,dx_2. \tag{15}$$

Производящая функция $\Pi^A$ на периоде $A$ определена в [2]:

$$\Pi^A(\alpha_1, z_1) = \Pi^{b_{12}}(\alpha_1, z_1) + b_{21}(z_1)\,\Pi^U(\alpha_1, z_1), \tag{16}$$

где

(17) $$\Pi^{b_{21}}(\alpha_1, z_1) = \\ = \alpha_1 \frac{\{\alpha_1 s_{21}(\lambda_1[1-\alpha_1]+z_1) - b_1(z_1) s_{21}'(\lambda_1[1-b_1(z_1)]+z_1)\} \times \\ \times \{1 - s_1(\lambda_1[1-\alpha_1]+z_1)\}}{[\alpha_1 - s_1'(\lambda_1[1-\alpha_1]+z_1)]\lambda_1(1-\alpha_1)+z_1} + \\ + \alpha_1 \frac{1 - s_{21}(\lambda_1[1-\alpha_1]+z_1)}{\lambda_1(1-\alpha_1)+z_1},$$

(18) $$\Pi^u(\alpha_1, z_1') = \int_0^\infty \frac{[1-\exp(-[\lambda_1+z_1]y)]s_{12}(y)\,dy}{\lambda_1+z_1-\lambda_1 b_1(z_1)[1-\exp(-[\lambda_1+z_1']y)]} \times \\ \times [1+\lambda_1\Pi^{b_1}(\alpha_1, z_1)],$$

а производящая функция $\Pi^{b_1}$ на периоде занятости $b_1$ изолированной приоритетной очереди приведена в [1] (стр. 19—20):

(19) $$\Pi^{b_1}(\alpha_1, z_1) = \frac{\alpha_1 - b_1(z_1)}{1 - \frac{1}{\alpha_1} s_1(\lambda_1[1-\alpha_1]+z_1)} \cdot \frac{1 - s_1(\lambda_1[1-\alpha_1]+z_1)}{\lambda_1(1-\alpha_1)+z_1}.$$

Выражение для производящей функции $\Pi^L$ на периоде $L$ находим в [1] (стр. 182):

(20) $$\Pi^L(\alpha_1, z_1) = \\ = f^I(0, \theta, z_1) \frac{\Omega(\lambda_1[1-\alpha_1]+z_1) - \Omega(\lambda_1[1-b_1(z_1)]+z_1)}{1 - \frac{1}{\alpha_1} s_1(\lambda_1[1-\alpha_1]+z_1)} \times \\ \times \frac{1 - s_1'(\lambda_1[1-\alpha_1]+z_1')}{\lambda_1(1-\alpha_1)+z_1},$$

где

(21) $$\Omega(x) = s_2(\theta+x) / F_2^c(\theta),$$

$$\Omega(z_1) = \frac{\exp(z_1\theta)}{F_2^c(\theta)} \left[ s_2(z_1) - \int_0^\theta s_2(x)\exp(-z_1 x)\,dx \right].$$

Наконец, определим плотность вероятности длительности цикла $C$. Выражение для $C(z_1)$ приводится в [1] (стр. 181). Заменяя, как и ранее, функции, относящиеся к периоду $b_1$, и производя замену переменных под знаком интеграла, получаем

(22) $$C(z_1) = \int_0^\theta s_2(x_2)\exp(-\{\lambda_1[1-A(z_1)]+z_1\}x_2)\,dx_2 + \exp([A(z_1) - \\ - b_1(z_1)]\theta\lambda_1) \int_\theta^\infty s_2(x_2)\exp(-\{\lambda_1[1-b_1(z_1)]+z_1\}x_2)\,dx_2.$$

## 3. Период занятости для основного процесса

Будем предполагать, что если требование прибывает в систему, когда обслуживающий прибор свободен, оно немедленно поступает на обслуживание, независимо от того, приоритетное оно или неприоритетное.

В [1] были получены выражения для основных характеристик процесса с двумя классами требований. Так как полученные там формулы нигде не содержат характеристик процесса на цикле обслуживания в явном виде, то они справедливы и для процесса, рассматриваемого в настоящей работе, поэтому мы здесь определим основные величины и приведем без вывода соотношения, из которых они могут быть найдены.

$B_2$, $B_2^{b_1}$ — длительность периода занятости и начального периода занятости с дополнительным интервалом занятости $b_1$ в изолированной неприоритетной очереди, когда в качестве времени обслуживания выступает длительность цикла обслуживания.

$\gamma$ — длительность периода занятости в основном процессе с двумя классами требований.

$H(m_2, t)$ — вероятность того, что в момент времени $t$ неприоритетное требование поступает на обслуживание при $(m_2-1)$ неприоритетных требованиях в очереди.

$I(\alpha_2, t)$ — производящая функция вероятностей $H(m_2, t)$.

Очевидно, что $I(1, t)$ является вероятностью того, что в момент времени $t$ неприоритетное требование поступает на обслуживание.

$W_i^\gamma(t, \tau)$ — соответственно плотность вероятности виртуального времени на периоде занятости для приоритетных $(i=1)$ и неприоритетных $(i=2)$ требований.

$W_1^{b_1}(t, \tau)$ — плотность вероятности виртуального времени для изолированной приоритетной очереди на периоде занятости $b_1$.

$m_i(t)$ — длина приоритетной $(i=1)$ и неприоритетной $(i=2)$ очереди на периоде занятости $\gamma$.

$\Pi^\gamma(\alpha_1, \alpha_2, t)$ — производящая функция вероятностей совместного распределения длин очередей $m_i(t)$ $(i=1, 2)$ на периоде занятости.

Для преобразований Лапласа, определенных выше величин и их плотностей вероятности, справедливы соотношения

$$\gamma(z_1) = \frac{\lambda_1}{\lambda} B_2^{b_1}(z_1) + \frac{\lambda_2}{\lambda} B_2(z_1), \quad \text{где } \lambda = \lambda_1 + \lambda_2, \tag{23}$$

$$B_2(z_1) = C(\lambda_2[1-B_2(z_1)]+z_1), \tag{24}$$

$$B_2^{b_1}(z_1) = b_1(\lambda_1[1-B_2(z_1)]+z_1), \tag{25}$$

$$I(\alpha_2, z_1) = \tag{26}$$
$$= \alpha_2 \frac{\lambda_2[\alpha_2 - B_2(z_1)] + \lambda_1\{b_1(\lambda_2[1-\alpha_2]+z_1) - b_1(\lambda_2[1-B_2(z_1)]+z_1)\}}{\lambda[\alpha_2 - C(\lambda_2[1-\alpha_2]+z_1)]},$$

$$I(1, z_1) = \frac{\lambda_2[1-B_2(z_1)] + \lambda_1\{b_1(z_1) - b_1(\lambda_2[1-B_2(z_1)]+z_1)\}}{\lambda[1-C(z_1)]}, \tag{27}$$

$$W_1^\gamma(z_1, z_2) = \frac{\lambda_1}{\lambda} W_1^{b_1}(z_1, z_2) + I(1, z_1)\, W_1^c(z_1, z_2), \tag{28}$$

$$W_2^\gamma(z_1, z_2) = \frac{\lambda_2[B_2(z_1) - C(z_2)] + \lambda_1[b_1(z_1 + \lambda_2[1-B_2(z_1)]) - b_1(z_2)]}{\lambda[z_2 - z_1 - \lambda_2\{1 - C(z_2)\}]}, \tag{29}$$

$$\Pi^\gamma(\alpha_1, \alpha_2, z_1) = \frac{1}{\alpha^2} I(\alpha_2, z_1)\, \Pi^c(\alpha_1, \alpha_2, z_1 + \lambda_2[1-\alpha_2]) + \tag{30}$$
$$+ \frac{\lambda_1}{\lambda} \Pi^{b_1}(\alpha_1, z_1 + \lambda_2[1-\alpha_2]).$$

## 4. Основной процесс

Обозначим через $e(t)$ вероятность того, что в момент времени $t$ обслуживающий прибор свободен. Тогда для преобразования Лапласа вероятности $e(t)$ справедлива формула, приведенная в [1] (стр. 82):

$$e(z_1) = \{z_1 + \lambda[1-\gamma(z_1)]\}^{-1}. \tag{31}$$

Через $\Pi(\alpha_1, \alpha_2, t)$ обозначим производящую функцию вероятностей совместного распределения длин очередей $m_1(t)$ и $m_2(t)$ в основном процессе. Тогда

$$\Pi(\alpha_1, \alpha_2, t) = \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \alpha_1^{m_1} \alpha_2^{m_2} P\{m_1(t) = m_1, m_2(t) = m_2\}. \tag{32}$$

Через $W_i(t, \tau)$ $(i=1, 2)$ обозначим плотности вероятностей виртуальных времен для приоритетных и неприоритетных требований в основном процессе.

Пользуясь теорией восстановления [1] для преобразований Лапласа определенных выше функций, получаем

$$\Pi(\alpha_1, \alpha_2, z_1) = e(z_1)[1+\lambda\Pi^\gamma(\alpha_1, \alpha_2, z_1)], \tag{33}$$

$$W_i(z_1, z_2) = e(z_1)[1+\lambda W_i^\gamma(z_1, z_2)], \tag{34}$$

где $\Pi^\gamma(\alpha_1, \alpha_2, z_1)$ определяется из (30), $W_1^\gamma(z_1, z_2)$ из (28) и $W_2^\gamma(z_1, z_2)$ из (29).

## 5. Стационарный режим для основного процесса

Аналогично [2] для стационарной вероятности того, что прибор свободен, имеем

$$e = [1+\lambda M\{\gamma\}]^{-1}. \tag{35}$$

Обозначим через $\Pi(\alpha_1, \alpha_2)$ производящую функцию стационарных вероятностей длин очередей $m_1$ и $m_2$:

$$\Pi(\alpha_1, \alpha_2) = \lim_{z_1\to 0} z_1 \Pi(\alpha_1, \alpha_2, z_1) = e[1 + \lambda\Pi^\gamma(\alpha_1, \alpha_2)], \tag{36}$$

где $\Pi^\gamma(\alpha_1, \alpha_2) = \Pi^\gamma(\alpha_1, \alpha_2, 0)$.

Обозначим, наконец, через $W_i(\tau)$ плотности вероятностей виртуальных времен для приоритетных и неприоритетных требований в стационарном режиме. Для соответствующих преобразований Лапласа получаем

$$W_i(z_2) = e[1+\lambda W_i^\gamma(z_2)] \quad (i=1, 2), \tag{37}$$

где $W_i^\gamma(z_2) = W_i^\gamma(0, z_2)$.

## 6. Заключение

Полученные соотношения позволяют найти числовые характеристики процесса обслуживания. Для этого можно воспользоваться соотношением

$$M(\xi^n) = (-1)^n \left.\frac{d^n f_\xi(z)}{dz^n}\right|_{z=0}, \tag{38}$$

где $\xi$ — случайная величина, а $f_\xi(z)$ — преобразование Лапласа ее плотности вероятности.

При заданном критерии качества процесса обслуживания может быть также решена задача оптимального выбора величины $\theta$.

Поступила в редакцию
29 августа 1974 г.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. *Джейсуол Н.* Очереди с приоритетами. «Мир», 1973.
2. *Волковинский М. И., Кабалевский А. Н.* Обслуживание с абсолютным приоритетом в системах с потерями на переключение. Автоматика и телемеханика, № 10, стр. 35–42, 1975.

---

## SERVICE WITH MIXED PRIORITY IN SYSTEMS WITH LOSSES IN SWITCHING

M. I. VOLKOVINSKY, A. N. KABALEVSKY

The paper is concerned with a service system with mixed priority and two types of entries (priority and non-priority) and one servicing unit with due regard for the switching time lost when a non-priority entry service is interrupted and resumed. The distribution of virtual service times and the generating function of probabilities of simultaneous distribution of queue lengths are found.

---