Министерство образования Республики Беларусь
БЕЛОРУССКАЯ ГОСУДАРСТВЕННАЯ ПОЛИТЕХНИЧЕСКАЯ АКАДЕМИЯ

Кафедра «Высшая математика № 2»

А.Д.Корзников
В.В.Павлов

# ЭЛЕМЕНТЫ ТЕОРИИ МАРКОВСКИХ ПРОЦЕССОВ И СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ

Методическое пособие для студентов
инженерно-экономических специальностей втузов

Минск 2001

Министерство образования Республики Беларусь
БЕЛОРУССКАЯ ГОСУДАРСТВЕННАЯ ПОЛИТЕХНИЧЕСКАЯ АКАДЕМИЯ

Кафедра «Высшая математика № 2»

А.Д.Корзников
В.В.Павлов

# ЭЛЕМЕНТЫ ТЕОРИИ МАРКОВСКИХ ПРОЦЕССОВ И СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ

Методическое пособие для студентов
инженерно-экономических специальностей втузов

Минск 2001

УДК 519.217
519.872
К66
Корзников А.Д.
Элементы теории марковских процессов и систем массового обслуживания: Метод. пособие для студ. инженерно-экономических спец. втузов / А.Д.Корзников, В.В.Павлов. – Мн.: БГПА, 2001. – 40 с.

В пособии описаны случайные марковские процессы, протекающие в вероятностных системах, – системах массового обслуживания (СМО). Помимо теоретического материала в пособии приведены примеры по основным типам СМО.

Данное пособие написано в полном соответствии с программой курса «Высшая математика» для студентов инженерно-экономических специальностей втузов.

Рецензент В.В.Веременюк

# 1. МАРКОВСКИЕ ПРОЦЕССЫ

В практической деятельности очень часто приходится сталкиваться с анализом работы своеобразных вероятностных систем, называемых системами массового обслуживания (СМО). Примерами таких систем могут служить телефонные станции, обслуживание на АЗС, обработка приходящих на железнодорожную станцию составов, ремонт неисправного оборудования и т.д. Теория массового обслуживания изучает зависимость между характером потока заявок (требований) , поступающих в случайные моменты времени на вход системы, и производительностью обслуживающих устройств, а также описывает количественные характеристики, позволяющие оценить эффективность функционирования системы.

В СМО происходит случайный процесс, что обусловлено случайным характером потока заявок и временем их обслуживания. Исследование работы СМО существенно облегчается, если случайный процесс, протекающий в СМО, является марковским. Это позволяет с помощью обыкновенных дифференциальных уравнений (в предельном случае - линейных алгебраических уравнений) найти основные количественные характеристики СМО. Поэтому вначале рассмотрим основные понятия марковских процессов.

Пусть состояние некоторой системы $S$ (АТС, АЗС, ремонтная мастерская и т.д.) изменяется в процессе ее функционирования. Если состояния системы изменяются во времени заранее непредсказуемым образом, то в системе $S$ протекает случайный процесс.

Случайный процесс, протекающий в системе $S$, называется *марковским*, если для каждого момента времени $t_0$ вероятность пребывания системы в любом из состояний в будущем $(t_1 > t_0)$ зависит только от того, в каком состоянии система находилась в настоящем $(t = t_0)$, и не зависит от того, когда и каким образом система пришла в это состояние (т.е. не зависит от того, как развивался процесс в прошлом).

Случайный процесс называется *процессом с дискретными состояниями*, если множество возможных состояний системы $\{S_1, S_2, ..., S_n\}$ конечно либо они могут быть перечислены одно за другим: $S_1$, $S_2$, ... , $S_n$. Сам процесс состоит в том, что система $S$ время от времени скачком переходит из одного состояния в другое. Граф состояний системы геометрически изображает возможные состояния системы, а стрелки указывают возможные переходы системы из состояния в состояние (рис. 1).

Если система $S$ переходит из состояния в состояние в фиксированные моменты времени $t_1$, $t_2$, ... , то случайный процесс, протекающий в системе $S$, называется *процессом с дискретным временем*; если же система $S$ переходит из состояния в состояние в случайный момент времени $t$, *то – процессом с непрерывным временем.*

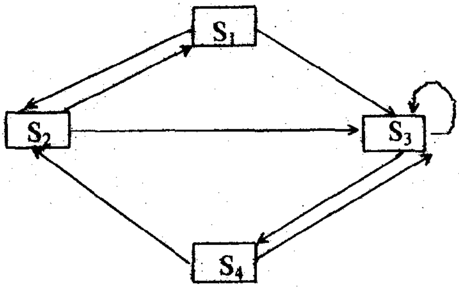


Рис. 1

Рассмотрим процесс с дискретным временем. Обозначим через $S(t_k)$ состояние системы в момент времени $t_k$, $t_k \in \{t_1, t_2, \dots\}$, $S(t_k) \in \{S_1, S_2, \dots, S_n\}$, а через $p_{ij}(k)$ – условную вероятность перехода системы $S$ из состояния $S_i$, в котором она находилась в момент времени $t_{k-1}$, в состояние $S_j$ в момент времени $t_k$:

$$p_{ij}(k) = P(S(t_k) = S_j | S(t_{k-1}) = S_i). \qquad (1)$$

Такие вероятности будем называть переходными вероятностями.

Если $p_{ij}(k)$ не зависят от того, когда и как система пришла в состояние $S_i$, то говорят, что в системе протекает марковский процесс.

Марковский процесс называется *однородным*, если вероятности перехода из состояния в состояние $p_{ij}(k)$ не зависят от времени, когда осуществляется переход, т.е. от $k$, а лишь от состояний $S_i$ и $S_j$, в противном случае процесс называется *неоднородным*.

Рассмотрим однородный марковский процесс. Обозначим через $p_{ij}$ вероятность перехода системы из состояния $S_i$ в состояние $S_j$ за один переход (шаг); через $p_{ii}$ – вероятность задержки системы $S$ в состоянии $S_i$. Полная вероятностная картина переходов будет задаваться матрицей

$$P=\begin{pmatrix} p_{11} & p_{12} \cdots p_{1n} \\ p_{21} & p_{22} \cdots p_{2n} \\ \cdots & \cdots\cdots\cdots \\ p_{n1} & p_{n2} \cdots p_{nn} \end{pmatrix}, \tag{2}$$

которая называется *матрицей переходных вероятностей*, или *матрицей вероятностей переходов.*

Элементы данной матрицы удовлетворяют условиям:

$$1)\ 0 \le p_{ij} \le 1,\ i,j=\overline{1,n};\ 2)\sum_{j=1}^{n} p_{ij} = 1,\ i=\overline{1,n}. \tag{3}$$

Зная матрицу $P$, можно построить размеченный граф состояний системы $S$:

1) количество вершин графа равно числу состояний – $n$;

2) если $p_{ij} \ne 0$, то вершины $S_i$ и $S_j$ соединяются ориентированной дугой, помеченной $p_{ij}$.

И наоборот, если дан размеченный граф состояний системы $S$, то можно выписать матрицу $P$.

*Пример 1*. Дан граф состояний системы $S$ (рис. 2). Найти матрицу переходных вероятностей.

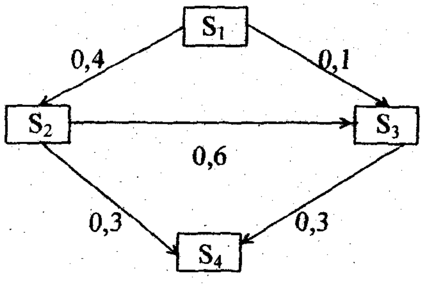


Рис. 2

*Решение*. Из графа состояний системы и свойства (3) матрицы переходных вероятностей следует:

$$p_{12} = 0{,}4;\; p_{13} = 0{,}1;\; p_{14} = 0;\; p_{11} = 1 - (p_{12} + p_{13} + p_{14}) = 0{,}5.$$
$$p_{21} = 0;\; p_{23} = 0{,}6;\; p_{24} = 0{,}3;\; p_{22} = 1 - (p_{21} + p_{23} + p_{24}) = 0{,}1.$$
$$p_{31} = 0;\; p_{32} = 0;\; p_{34} = 0{,}3;\; p_{33} = 1 - (p_{31} + p_{32} + p_{34}) = 0{,}7.$$
$$p_{41} = 0;\; p_{42} = 0;\; p_{43} = 0;\; p_{44} = 1 - (p_{41} + p_{42} + p_{43}) = 1.$$

Матрица (2) задает условные вероятности перехода из состояния в состояние за один шаг.

Однако для прогнозирования состояния, в котором будет находиться система в процессе функционирования, представляет интерес оценка вероятности каждого состояния системы на некотором $m$-м шаге.

Пусть вектор $\bar{V}(0) = (V_1(0), V_2(0), \ldots, V_n(0))$ задает вероятности состояния системы в начальный (нулевой) момент времени, то есть $V_i(0)$ – это вероятность того, что в начальный момент времени система находится в состоянии $S_i$. Если известно, что в начальный момент времени система находится в состоянии, например, $S_k$, то вектор $\bar{V}(0)$ имеет вид

$$\bar{V}(0) = (0, \ldots, 0, 1, 0, \ldots, 0), \text{ где } V_k(0) = 1,\; V_i(0) = 0 \text{ для всех } i \neq k.$$

Так как то, что система находится в одном из своих состояний – событие достоверное, следовательно, имеем $\sum\limits_{i=1}^{n} V_i(0) = 1$.

Пусть известен вектор вероятностей состояний системы на $(m - 1)$-м шаге $\bar{V}(m-1) = (V_1(m-1), \ldots, V_n(m-1))$. Найдем вероятность того, что на $m$-м шаге система окажется в состоянии $S_j$ $(j = \overline{1,n})$. Воспользуемся *формулой полной вероятности* $V_j(m) = V_1(m - 1)p_{1j} + V_2(m - 1)p_{2j} + \ldots + V_n(m - 1)p_{nj}$, $j = \overline{1,n}$.

Поскольку таким образом вычисляется вероятность любого состояния $S_j$, $j = 1, 2, \ldots, n$, то в матричном виде это равенство имеет вид

$$\left(V_1(m), V_2(m), \ldots, V_n(m)\right) = \left(V_1(m-1), V_2(m-1), \ldots, V_n(m-1)\right) \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \quad (4)$$

или $\overline{V}(m) = \overline{V}(m-1)P$.

Но тогда получаем, что $\overline{V}(1) = \overline{V}(0)P$,

$\overline{V}(2) = \overline{V}(1)P = \overline{V}(0)P \cdot P = \overline{V}(0)P^2$,

$\overline{V}(3) = \overline{V}(2)P = \overline{V}(0)P^2 \cdot P = \overline{V}(0)P^3$ и т.д.

Продолжая таким образом, получим

$$\overline{V}(m) = \overline{V}(0)P^m. \tag{5}$$

<u>*Пример 2*</u>. В начальный момент времени система (см. пример 1) находится в состоянии $S_1$. Определить вероятности состояний системы после двух шагов.

<u>*Решение*</u>. Вектор вероятностей состояний системы $S$ в начальный момент времени $\overline{V}(0) = (1;\ 0;\ 0;\ 0)$. Найдем вектор $\overline{V}(2)$ вероятностей состояний системы после двух шагов по формуле (5): $\overline{V}(2) = \overline{V}(0)\ P^2$.

Следовательно,

$$\overline{V}(2) = (1;0;0;0)\cdot\begin{pmatrix} 0{,}5 & 0{,}4 & 0{,}1 & 0 \\ 0 & 0{,}1 & 0{,}6 & 0{,}3 \\ 0 & 0 & 0{,}7 & 0{,}3 \\ 0 & 0 & 0 & 1 \end{pmatrix}\cdot\begin{pmatrix} 0{,}5 & 0{,}4 & 0{,}1 & 0 \\ 0 & 0{,}1 & 0{,}6 & 0{,}3 \\ 0 & 0 & 0{,}7 & 0{,}3 \\ 0 & 0 & 0 & 1 \end{pmatrix} =$$

$$= (0{,}5;\ 0{,}4;\ 0{,}1;\ 0)\cdot\begin{pmatrix} 0{,}5 & 0{,}4 & 0{,}1 & 0 \\ 0 & 0.1 & 0{,}6 & 0{,}3 \\ 0 & 0 & 0{,}7 & 0{,}3 \\ 0 & 0 & 0 & 1 \end{pmatrix} = (0{,}25;\ 0{,}24;\ 0{,}36;\ 0{,}15).$$

Марковский процесс называется регулярным, если система может перейти из любого состояния $S_i$ в любое другое состояние $S_j$ за конечное число шагов. Это означает, что существует $m \in N$ такое, что все элементы матрицы $P^m$ строго положительны.

Рассмотрим марковский процесс с дискретными состояниями и непрерывным временем. Обозначим через $V_i(t)$ вероятность того, что в момент времени $t$ система будет находиться в состоянии $S_i$, $i = 1, 2, ..., n$. В любой момент времени $t$ справедливо соотношение

$$\sum_{i=1}^{n} V_i(t) = 1.$$

Обозначим через $P_{ij}(\Delta t)$ вероятность перехода системы из состояния $S_i$ в состояние $S_j$ за время $\Delta t$.

Плотностью вероятности перехода $\lambda_{ij}$ будем называть величину, равную

$$\lambda_{ij} = \lim_{\Delta t \to 0} \frac{P_{ij}(\Delta t)}{\Delta t}. \qquad (6)$$

Из (6) следует, что при достаточно малом $\Delta t$

$$P_{ij}(\Delta t) = \lambda_{ij} \cdot \Delta t + 0(\Delta t) \approx \lambda_{ij} \cdot \Delta t.$$

Если плотности вероятностей не зависят от времени ($\lambda_{ij} = const$), то марковский процесс с непрерывным временем называется *однородным*, в противном случае – *неоднородным*.

Рассмотрим однородный марковский процесс с непрерывным временем. Вероятности состояний системы $S$ $V_1(t)$, $V_2(t)$, ..., $V_n(t)$ можно найти, решая систему обыкновенных дифференциальных уравнений (уравнений Колмогорова), которая может быть получена из следующих соображений.

Найдем одну из вероятностей состояний системы, например, $V_1(t)$.

Придадим $t$ малое приращение $\Delta t$ и найдем вероятность $V_1(t + \Delta t)$ того, что в момент времени $(t + \Delta t)$ система будет находиться в состоянии $S_1$. Это событие может произойти следующими $n$ способами:

система находилась в состоянии $S_1$ и осталась в этом состоянии;

система находилась в состоянии $S_2$ и за время $\Delta t$ перешла в состояние $S_1$;

система находилась в состоянии $S_3$ и за время $\Delta t$ перешла в состояние $S_1$ и т.д.

Воспользовавшись формулой полной вероятности, получим

$$V_1(t + \Delta t) = V_1(t)P_{11}(\Delta t) + V_2(t)P_{21}(\Delta t) + \ldots + V_n(t)P_{n1}(\Delta t).$$

Так как сумма элементов любой строки матрицы вероятностей переходов равна единице, то

$$P_{11}(\Delta t) = 1 - P_{12}(\Delta t) - \ldots - P_{1n}(\Delta t).$$

Отсюда

$$V_1(t + \Delta t) = V_1(t)(1 - P_{12}(\Delta t) - \ldots - P_{1n}(\Delta t)) + V_2(t)\, P_{21}(\Delta t) + \ldots + V_n(t)P_{n1}(\Delta t) =$$
$$= V_1(t) - V_1(t)(P_{12}(\Delta t) + \ldots + P_{1n}(\Delta t)) + V_2(t)P_{21}(\Delta t) + \ldots + V_n(t)P_{n1}(\Delta t).$$

Учитывая, что $P_{ij}(\Delta t) \approx \lambda_{ij} \cdot \Delta t$, получим

$$V_1(t + \Delta t) - V_1(t) = - V_1(t)\,(\lambda_{12} + \ldots + \lambda_{1n})\,\Delta t + V_2(t)(\lambda_{21} + \ldots + V_n(t)\,\lambda_{n1})\,\Delta t\,.$$

Разделим обе части на $\Delta t$ и, устремив $\Delta t$ к нулю, перейдем к пределу:

$$\lim_{\Delta t \to 0} \frac{V_1(t + \Delta t) - V_1(t)}{\Delta t} = -V_1(t)(\lambda_{12} + \ldots + \lambda_{1n}) + V_2(t)\lambda_{21} + \ldots + V_n(t)\lambda_{n1}\,.$$

Левая часть есть не что иное, как производная функции $V_1(t)$ , то есть

$$\frac{dV_1(t)}{dt} = -V_1(t)(\lambda_{12} + \ldots + \lambda_{1n}) + V_2(t)\lambda_{21} + \ldots + V_n(t)\lambda_{n1}.$$

Таким образом, получено дифференциальное уравнение, которому должна удовлетворять функция $V_1(t)$. Аналогичные дифференциальные уравнения могут быть выведены и для остальных вероятностей состояний: $V_2(t)$, $V_3(t)$, ... , $V_n(t)$.

Получим систему

$$V_j'(t) = -V_j(t)(\lambda_{j1} + \ldots + \lambda_{j,j-1} + \lambda_{j,j+1} + \ldots + \lambda_{jn}) + V_1(t)\lambda_{1j} + \ldots +$$
$$+ V_{j-1}(t)\lambda_{j-1,j} + V_{j+1}(t)\lambda_{j+1,j} + \ldots + V_n(t)\lambda_{nj},\; j = 1,2,\ldots,n.$$

Решение этой системы дифференциальных уравнений даст нам искомые вероятности состояний как функции времени. Начальные условия берутся в зависимости от того, в каком состоянии находилась система в начальный момент времени. Например, если система в начальный момент времени ($t = 0$) находилась в состоянии $S_j$, то начальными условиями будут:

$$\text{при } t = 0 \quad V_j(0) = 1, \quad V_i(0) = 0, \quad \text{если} \quad i \neq j.$$

Если рассмотреть граф состояний системы и систему дифференциальных уравнений Колмогорова, то можно легко описать правило построения последней.

В левой части каждого уравнения стоит производная вероятности состояния системы, а правая часть содержит столько членов, сколько стрелок связано с данным состоянием в графе состояний системы. *Если стрелка направлена из состояния*, то соответствующий член имеет знак "минус", а если в состояние - знак "плюс". Каждый член равен произведению плотности вероятности перехода, соответствующей данной стрелке, умноженной на вероятность того состояния, из которого исходит стрелка.

*Пример 3*. Размеченный граф состояний системы $S$ имеет вид, приведенный на рис. 3. Составить уравнения Колмогорова, если в начальный момент времени система находится в состоянии $S_1$.

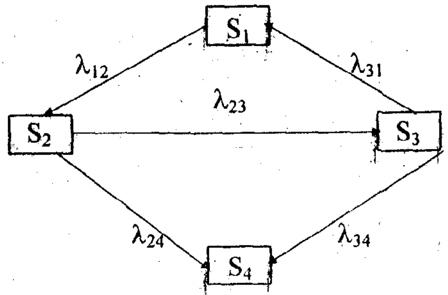


Рис. 3

*Решение*. Уравнения Колмогорова имеют вид

$$
\begin{cases}
\dfrac{dV_1}{dt} = -\lambda_{12}V_1 + \lambda_{31}V_3;\\
\dfrac{dV_2}{dt} = -\lambda_{23}V_2 - \lambda_{24}V_2 + \lambda_{12}V_1;\\
\dfrac{dV_3}{dt} = -\lambda_{31}V_3 - \lambda_{34}V_3 + \lambda_{23}V_2;\\
\dfrac{dV_4}{dt} = \lambda_{24}V_2 + \lambda_{34}V_3.
\end{cases}
$$

Начальные условия: при $t = 0$ $V_1 = 1$, $V_2 = V_3 = V_4 = 0$. Интегрирование этой системы уравнений даст нам вероятности состояний $V_1(t)$, ..., $V_n(t)$ системы $S$. Можно было бы все уравнения системы не составлять; так как $V_1 + V_2 + V_3 + V_4 = 1$, то можно составить систему уравнений, к примеру, для $V_1$, $V_2$, $V_3$ и найти $V_4 = 1 - (V_1 + V_2 + V_3)$.

Рассмотрим однородный марковский процесс с известной матрицей вероятностей переходов $P$ (см. формулу (2)). Если марковский процесс является регулярным, то при $m \to +\infty$, независимо от начального состояния, система будет функционировать в установившемся режиме, то есть вероятности состояний системы в этом режиме не зависят от начального состояния. Такие вероятности $\vec{V}_ф = (V_1, V_2, ..., V_n)$ называются *предельными* или *финальными* и могут быть найдены из следующих соображений.

Пусть $E_{n \times n}$ единичная $n \times n$ матрица. Тогда

$$\overline{V}(m) = \overline{V}(m)E = \overline{V}(m-1)P \quad \text{или} \quad \overline{V}(m)E - \overline{V}(m-1)P = 0.$$

Переходя к пределу при $m \to \infty$ и учитывая, что

$$\lim_{m \to \infty} \overline{V}(m) = \lim_{m \to \infty} \overline{V}(m-1) = V_ф,$$

получаем

$$\overline{V}_ф \cdot E - \overline{V}_ф \cdot P = 0 \text{ или} \qquad (7)$$
$$\overline{V}_ф \cdot (E - P) = 0.$$

Таким образом, для определения финальных вероятностей $\overline{V_\phi} = (V_1,...,V_n)$ необходимо решить систему линейных однородных уравнений (7) с учетом того, что $V_1 + V_2 + ... + V_n = 1$.

Пусть в системе $S$ с дискретными состояниями $S_1$, $S_2$, ..., $S_n$ протекает однородный марковский случайный процесс с непрерывным временем. Составив систему уравнений Колмогорова и проинтегрировав их при заданных начальных условиях, мы получим вероятности состояний системы $S$: $V_1(t)$, $V_2(t)$, ..., $V_n(t)$ как функции от $t$. Если существуют пределы этих функций при $t \to \infty$, то они называются предельными или финальными вероятностями состояний:

$$\lim_{t \to \infty} V_i(t) = V_i,\ i = 1,...,n. \tag{8}$$

Справедливо следующее условие: если число состояний системы $S$ конечно и из любого состояния можно перейти за конечное число шагов в любое другое состояние, то предельные вероятности состояний существуют и не зависят от начального состояния системы. На рис. 4 показан граф состояний, удовлетворяющий вышесказанному условию, а на рис. 5 – не удовлетворяющий.

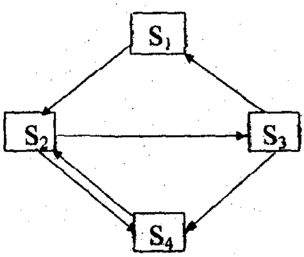


Рис. 4

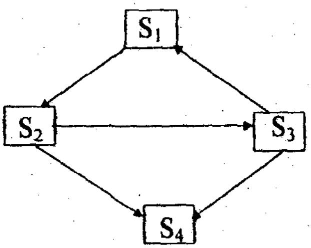


Рис. 5

При $t \to \infty$ в системе $S$ устанавливается предельный стационарный режим, причем вероятности состояний системы $V_1$, $V_2$, ... , $V_n$ представляют собой среднее относительное время пребывания системы в данном состоянии. Так как производная от постоянных предельных вероятностей равна нулю, то для того, чтобы их найти, нужно все левые части уравнений Колмогорова положить равными нулю и решить систему линейных алгебраических уравнений совместно с дополнительным условием

$$V_1 + V_2 + \ldots + V_n = 1.$$

## 2. ПРОЦЕСС «ГИБЕЛИ И РАЗМНОЖЕНИЯ»

Найдем предельные вероятности состояний системы на примере процесса "гибели и размножения". Марковский непрерывный процесс называется процессом "гибели и размножения", если граф состояний имеет вид, показанный на рис. 6.

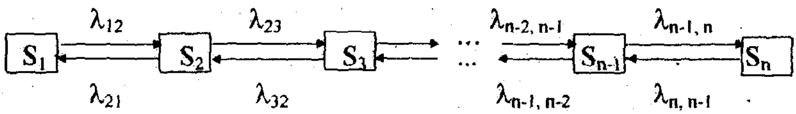


Рис. 6

Так как режим в данной системе является установившимся, то по правилу составления уравнений Колмогорова получаем систему алгебраических уравнений

$$\begin{cases} \lambda_{12}V_1 = \lambda_{21}V_2, \\ \lambda_{23}V_2 = \lambda_{32}V_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \lambda_{n-1,n}V_{n-1} = \lambda_{n,n-1}V_n, \\ V_1 + V_2 + \ldots + V_n = 1. \end{cases} \tag{9}$$

Выразим вероятности $V_2, \ldots, V_n$ через $V_1$. Из первого уравнения системы (9) имеем

$$V_2 = \frac{\lambda_{12}}{\lambda_{21}} V_1. \tag{10}$$

Из второго уравнения получаем

$$V_3 = \frac{\lambda_{23}}{\lambda_{32}} V_2 = \frac{\lambda_{23}\lambda_{12}}{\lambda_{32}\lambda_{21}} V_1. \tag{11}$$

И так далее:

$$V_n = \frac{\lambda_{n-1,n} \cdot \lambda_{n-2,n-1} \cdot \ldots \cdot \lambda_{23} \cdot \lambda_{12}}{\lambda_{n,n-1} \cdot \lambda_{n-1,n-2} \cdot \ldots \cdot \lambda_{32} \cdot \lambda_{21}} V_1. \tag{12}$$

Выразим $V_1$ из последнего уравнения системы (9), подставив в него $V_2$, $V_3$, ... , $V_n$:

$$V_1 = \left(1 + \frac{\lambda_{12}}{\lambda_{21}} + \frac{\lambda_{23} \cdot \lambda_{12}}{\lambda_{32} \cdot \lambda_{21}} + \ldots + \frac{\lambda_{n-1,n} \cdot \ldots \cdot \lambda_{23} \cdot \lambda_{12}}{\lambda_{n,n-1} \cdot \ldots \cdot \lambda_{32} \cdot \lambda_{21}}\right)^{-1}. \tag{13}$$

Тогда остальные вероятности $V_2$, ... , $V_n$ состояний системы находятся из формул (10)-(12) при подстановке в них $V_1$.

## 3. ПРОСТЕЙШИЙ ПОТОК СОБЫТИЙ

В большинстве случаев при рассмотрении процессов, протекающих в системе с дискретными состояниями и непрерывным временем, удобно представлять, что переходы системы из состояния в состояние происходят под действием потоков событий (требований), например, потока машин, подъезжающих на АЗС.

Поток требований называется *регулярным*, если требования следуют одно за другим через строго определенные промежутки времени. На практике такой поток редок, но представляет интерес как предельный случай. Чаще приходится встречаться с потоками требований, когда промежутки времени между ними являются случайными величинами. Рассмотрим потоки требований, которые обладают следующими свойствами: *стационарностью, ординарностью и отсутствием последействия.*

1. Поток требований называется *стационарным*, если вероятность появления того или иного числа требований на участок времени

длиной $\Delta t$ зависит только от его длины, но не от места его расположения на оси времени 0t (рис. 7).

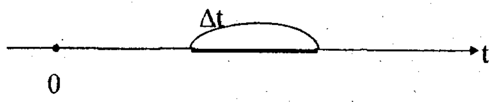


Рис. 7

Вероятностные характеристики стационарного потока требований не изменяются во времени. Интенсивность (или плотность) потока требований - среднее число требований, поступающих в единицу времени, – для стационарного потока является постоянной.

2. Поток требований называется *ординарным*, если вероятность появления двух или более требований за достаточно малый промежуток времени $\Delta t$ пренебрежимо мала по сравнению с вероятностью появления одного требования. Это означает, что требования в потоке приходят поодиночке.

3. Поток требований называется *потоком без последействия*, если для любых непересекающихся участков времени число требований, попадающих на один из них, не зависит от того, сколько требований попало на другой. Отсутствие последействия означает, что требования появляются в последовательные моменты времени независимо друг от друга.

Поток требований, обладающий всеми тремя свойствами, называется *простейшим или пуассоновским* потоком. Вероятность поступления $k$ требований в течение промежутка времени $t$ для такого потока определяется *формулой Пуассона*:

$$P(k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

Для выяснения смысла параметра $\lambda$ вычислим математическое ожидание числа требований, поступивших с СМО в течение промежутка времени (0, t):

$$M(t)=\sum_{k=1}^{\infty}kP_k(t)=\sum_{k=1}^{\infty}k\frac{(\lambda t)^k}{k!}e^{-\lambda t}=e^{-\lambda t}\lambda t\sum_{k=1}^{\infty}\frac{(\lambda t)^{k-1}}{(k-1)!}=e^{-\lambda t}\lambda t\sum_{n=0}^{\infty}\frac{(\lambda t)^n}{n!}.$$

Так как

$$\sum_{n=0}^{\infty}\frac{(\lambda t)^n}{n!}=1+\frac{\lambda t}{1!}+\frac{(\lambda t)^2}{2!}+\ldots+\frac{(\lambda t)^n}{n!}+\ldots=e^{\lambda t},$$

то получаем $M(t)=\lambda t$. При $t=1$ $M(t)=\lambda$, то есть параметр $\lambda$ равен среднему числу требований, поступающих в СМО в единицу времени. Но тогда среднее время между поступлениями требований равно $1/\lambda$.

Для пуассоновского потока с параметром $\lambda$ промежутки времени между поступлениями требований распределены по экспоненциальному закону с функцией распределения $F(t)=1-e^{-\lambda t}$. Действительно, $F(t)=P(T<t)$. Но вероятность $P(T<t)$ того, что промежуток времени между поступлениями требований меньше $t$, равна вероятности того, что за этот промежуток поступило хотя бы одно требование, т.е.

$$P(T<t)=\sum_{k=1}^{\infty}P_k(t)=\sum_{k=1}\frac{(\lambda t)^k}{k!}e^{-\lambda t}=$$
$$=e^{-\lambda t}\left(-1+1+\frac{\lambda t}{1!}+\frac{(\lambda t)^2}{2!}+\ldots+\frac{(\lambda t)^n}{n!}+\ldots\right)=$$
$$=e^{-\lambda t}(-1+e^{\lambda t})=1-e^{-\lambda t}.$$

Отметим еще одно важное свойство, которым обладают пуассоновские потоки. Объединение нескольких пуассоновских потоков дает снова пуассоновский поток, а вероятностное разделение пуассоновского потока на несколько дает независимые пуассоновские потоки. Если поток требований формируется $n$ независимыми источниками интенсивности $\lambda_i$, $i=1, 2, \ldots, n$, потоки соизмеримы и оказывают на сумму равномерно малое влияние, то объединение этих потоков в один дает пуассоновский поток интенсивности $\lambda=\lambda_1+\lambda_2+\ldots+\lambda_n$.

## 4. КЛАССИФИКАЦИЯ СМО

При исследовании СМО предполагаем, что потоки требований, переводящие систему из одного состояния в другое, являются простейшими. Это обусловлено следующим:

1) простейшие потоки часто встречаются на практике;

2) для простейшего потока сравнительно просто получить формулы для вычисления характеристик работы различных СМО;

3) если потоки требований не являются простейшими, то при анализе такой СМО могут быть получены приближенные, достаточно удовлетворительные результаты, если этот поток заменить простейшим с такой же интенсивностью.

В СМО выделяют следующие основные компоненты: *входящий поток требований*, *дисциплина очереди*, *узел обслуживания*, *выходящий поток требований*. Будем считать, что все потоки требований, переводящие СМО из состояния в состояние, являются простейшими.

*Входящий поток* представляет процесс поступления в систему требований, нуждающихся в обслуживании. Если требования поступают в систему по одному (например, прибытие автомобилей на АЗС, поступление заказов в ателье и т.д.), то такие СМО называются системами с единичным поступлением требований. Если же требования поступают в систему группами (посетители ЗАГСа, прибытие вагонов на станцию под погрузку и разгрузку и т.д.), то такие СМО называют *системами с групповым поступлением требований*.

СМО, у которых требование, поступившее в тот момент, когда все узлы обслуживания заняты, получает отказ и покидает систему, называются *системами с отказами (потерями)*.

СМО, у которых требование, поступившее в тот момент, когда все узлы обслуживания заняты, становится в очередь и ожидает, пока не освободится один из узлов, называются *системами с ожиданием (очередью)*. Системы с ожиданием (очередью) делятся на системы с неограниченным ожиданием и системы с ограниченным ожиданием.

В системах с неограниченным ожиданием требование, поступившее в тот момент, когда все узлы обслуживания заняты, становится в очередь и ждет обслуживания. В таких системах поступившее требование всегда будет обслужено.

В системах с ограниченным ожиданием на пребывание требования (заявки) в очереди могут накладываться ограничения: длина очереди – число одновременно находящихся в очереди требований; ограниченное время пребывания каждого требования в системе.

Процесс выбора требований из очереди на обслуживание характеризуется дисциплиной очереди. Чаще всего на практике встречается *дисциплина*: первым поступил – первым обслужен. В этом случае обслуживание в СМО с ожиданием является упорядоченным, требования обслуживаются в порядке поступления. Кроме того, в некоторых СМО применяются приоритетные дисциплины очереди – *приоритетные СМО*. В таких СМО некоторые требования обслуживаются в первую очередь на основании каких-то признаков.

После выбора из очереди требования поступают в узел обслуживания, который характеризуется продолжительностью и порядком выполнения процедуры обслуживания. Узел обслуживания может состоять из одного обслуживающего устройства (канала), в этом случае система называется *одноканальной СМО*, или из нескольких – *многоканальная СМО*.

Время обслуживания является одной из самых важных характеристик обслуживающего устройства. Как правило, это случайная величина, зависящая в каждом конкретном случае от ряда причин. Так например, время обслуживания автомобиля, поступившего на станцию технического обслуживания для ремонта, будет зависеть от диагностики, серьезности неисправности, наличия запасных частей и т.д. Поэтому случайная величина – время обслуживания, $T_{обсл.}$, может быть полностью охарактеризовано функцией распределения

$$F(t) = P(T_{обсл.} < t),$$

где $P(T_{обсл.} < t)$ – вероятность того, что время обслуживания не превосходит величины $t$.

Законы распределения времени обслуживания могут быть различными, однако наиболее часто встречающимся в практических приложениях является *экспоненциальный* закон. В этом случае функция распределения имеет вид

$$F(t) = 1 - e^{-\mu t},$$

где, как это было установлено выше, положительный параметр $\mu$ определяет интенсивность обслуживания, то есть среднее число требований, обслуживаемых устройством в единицу времени, а величина $1/\mu$ равна среднему времени обслуживания ($\frac{M}{T_{обсл.}} = \frac{1}{\mu}$). Кроме того, поток обслуженных требований, покидающих СМО, будет простейшим (пуассоновским).

Пусть узел обслуживания СМО состоит из $n$ параллельных обслуживающих каналов. Время обслуживания каждым каналом распределено по экспоненциальному закону с параметрами $\mu_i, i = \overline{1,n}$. Требования обслуживаются в порядке поступления, а при наличии одновременно нескольких свободных каналов с одинаковой вероятностью выбирается любой. В этом случае закон распределения времени обслуживания всеми каналами также будет экспоненциальным с функцией распределения

$$F(t) = 1 - e^{-(\mu_1 + \mu_2 + \ldots + \mu_n)t}$$

Это означает, что интенсивность обслуживания требований всеми каналами равна сумме интенсивностей каждого канала $\mu = \mu_1 + \ldots + \mu_n$. Если же каналы идентичны и их интенсивности равны между собой и равны $\mu$, то $\mu = n\mu$, то есть интенсивность обслуживания увеличивается в $n$ раз, а среднее время обслуживания уменьшается в $n$ раз по сравнению с временем обслуживания одним каналом.

После обслуживания требования покидают обслуживающее устройство. Если обслуженные требования вновь возвращаются в систему (например, отремонтированные станки на некотором предприятии), то такие системы называются *замкнутыми* системами массового обслуживания.

Состояние СМО характеризуется количеством требований, находящихся в ней, причем система может переходить из состояния в состояние под воздействием входного потока требований (в систему поступило новое требование) и выходящего потока (требование обслужено и покинуло систему). Поскольку оба потока мы считаем пуассоновскими, то вероятность того, что в систему поступит или систему покинет более одного требования за достаточно малый промежуток времени $\Delta t$, есть величина бесконечно малая по сравнению с $\Delta t$. Поэтому процесс, протекающий в такой СМО, является процессом "*гибели и размножения*", то есть если через $S_j$ обозначить состояние системы, когда в ней находится $j$ требований, то за небольшой промежуток времени возможны ее переходы лишь в соседние состояния $S_{j-1}$ и $S_{j+1}$ и граф состояний имеет вид, изображенный на рис. 8.

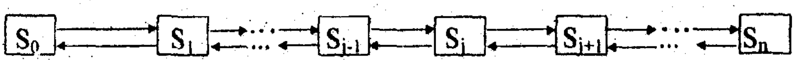


*Рис. 8*

Далее рассмотрим наиболее часто встречающиеся и хорошо изученные системы массового обслуживания.

## 5. СМО С ОТКАЗАМИ

Рассмотрим $n$-канальную СМО с отказами, которая может находиться в одном из следующих $(n + 1)$ состояний:

$S_0$ – все обслуживающие устройства свободны;

$S_1$ – занято одно устройство, остальные свободны;

..........................................................................................

$S_k$ – заняты $k$ устройств, остальные свободны;

..........................................................................................

$S_n$ – все $n$ устройств заняты.

Заявка, поступившая в тот момент, когда все обслуживающие устройства заняты, получает отказ и покидает систему.

Один и тот же поток требований с интенсивностью $\lambda$ переводит систему из состояния в состояние последовательно от начального $S_0$ в конечное $S_n$. Интенсивность потока обслуженных требований, переводящего систему из состояния в состояние, будет определяться числом каналов, занятых обслуживаем, т.е. если занято $k$ каналов (состояние $S_k$), то интенсивность потока обслуженных требований равна $k\mu$. Тогда граф состояний СМО имеет вид, изображенный на рис. 9.

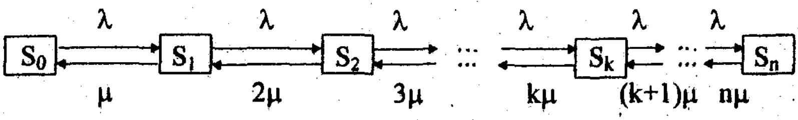


Рис. 9

Процесс, протекающий в СМО, является частным случаем процесса "гибели и размножения". Тогда из формул (10)-(13) следует:

$$\begin{cases} V_k = \dfrac{\left(\dfrac{\lambda}{\mu}\right)^k}{k!} V_0,\ k = 1,2,\dots,n; \\ V_0 = \left(1 + \dfrac{\left(\dfrac{\lambda}{\mu}\right)}{1!} + \dfrac{\left(\dfrac{\lambda}{\mu}\right)^2}{2!} + \dots + \dfrac{\left(\dfrac{\lambda}{\mu}\right)^n}{n!}\right)^{-1}. \end{cases} \tag{14}$$

Обозначим $\lambda / \mu = \rho$ и назовем эту величину "приведенной интенсивностью" потока требований. Она выражает среднее число требований, приходящих в СМО за среднее время обслуживания одного требования. Формулы (14) после этого обозначения примут вид

$$\begin{cases} V_k = \dfrac{\rho^k}{k!} \cdot V_0,\ k = 1,2,\dots,n; \\ V_o = \left(1 + \dfrac{\rho}{1!} + \dfrac{\rho^2}{2!} + \dots + \dfrac{\rho^n}{n!}\right)^{-1}. \end{cases} \tag{15}$$

Вероятность того, что требование получит отказ, равна

$$P_{отк.} = V_n = \frac{\rho^n}{n!} \cdot V_o. \tag{16}$$

Относительная пропускная способность $q$ – средняя доля поступивших требований, обслуживаемых системой (отношение среднего числа требований, обслуживаемых системой в единицу времени, к среднему числу поступающих за это время требований), находится как вероятность того, что требование будет принято к обслуживанию. Эта вероятность дополняет $P_{отк.}$ до единицы, следовательно:

$$q = 1 - V_n. \tag{17}$$

Абсолютная пропускная способность $A$ – среднее число требований, которое может обслужить система за единицу времени, равна

$$A = \lambda \cdot q = \lambda \cdot (1 - V_n). \tag{18}$$

Среднее число занятых обслуживающих устройств (обозначим эту величину через $\bar{k}$), что для данной СМО это то же, что и среднее число требований, находящихся в системе, можно вычислить как математическое ожидание дискретной случайной величины, принимающей значения 0, 1, 2, ..., $n$ с соответствующими вероятностями $V_0$, $V_1$, ..., $V_n$ :

$$\bar{k} = V_1 + 2V_2 + \ldots + n \cdot V_n. \tag{19}$$

Однако значительно проще эту характеристику можно получить из следующих соображений. Абсолютная пропускная способность $A$ характеризует среднее число заявок, обслуживаемых за единицу времени. Один канал обслуживает в среднем за единицу времени $\mu$ заявок, поэтому среднее число занятых каналов равно

$$\bar{k} = \frac{A}{\mu} = \frac{\lambda \cdot (1 - V_n)}{\mu} = \rho \cdot (1 - V_n). \tag{20}$$

Тогда среднее число простаивающих каналов равно $\bar{n}_{пр.} = n - \bar{k}$ .

Все характеристики одноканальной СМО с отказами получаются из формул (16)-(18) при $n = 1$.

_Пример 4_. Трехканальная СМО с отказами представляет собой АТС с тремя телефонными линиями. Заявка-вызов, пришедший в момент, когда все три линии заняты, получает отказ (соединения не происходит). Интенсивность потока вызовов $\lambda = 0,8$ вызовов в минуту. Средняя продолжительность разговора равна $\bar{t}_{обс.}$ = 1,5 минут. Найти вероятности состояний, абсолютную и относительную пропускные способности, вероятность отказа и среднее число занятых и простаивающих каналов.

*Решение.* Интенсивность обслуживания одной линией (среднее число разговоров, обслуженных одной линией за единицу времени) равна

$$\mu = \frac{1}{\bar{t}_0} \approx 0,667.$$

Приведенная интенсивность потока заявок

$$\rho = \frac{\lambda}{\mu} = \frac{0,8}{0,667} = 1,2.$$

По формулам (15) получаем:

$$V_1 = \frac{\rho}{1!} V_0 = 1,2 \cdot V_0;$$

$$V_2 = \frac{\rho^2}{2!} V_0 = 0,72 \cdot V_0;$$

$$V_3 = \frac{\rho^3}{3!} V_0 = 0,288 \cdot V_0.$$

$$V_0 = \frac{1}{1+1,2+0,72+0,288} \approx 0,312.$$

$V_1$ =1,2·0,312=0,374; $V_2$ =0,72·0,312≈0,224; $V_3$=0,288·0,312≈0,090.

Вероятность отказа $P_{отк.} = V_3 = 0,090$.

Относительная и абсолютная пропускные способности равны

$$q = 1 - V_3 = 0,910; \quad A = \lambda \cdot q = 0,8 \cdot 0,910 = 0,728.$$

Среднее число занятых каналов $\bar{k} = \rho(1 - V_3) = 1,2 \cdot 0,91 \approx 1,09$.

Таким образом, в установившемся режиме работы СМО в среднем будет занято чуть больше одного канала, остальные будут проставать. Этой ценой достигается относительно высокий уровень эффективности обслуживания – около 91% поступивших вызовов будет обслужено.

## 6. СМО С ОЖИДАНИЕМ

Рассмотрим СМО с $n$ обслуживающими каналами, в которую поступает поток требований (заявок на обслуживание) с интенсивностью $\lambda$. Интенсивность обслуживания заявок у всех каналов одинакова и равна $\mu$. Заявка, поступившая на обслуживание в СМО, когда все каналы заняты, становится в очередь для ожидания обслуживания. Число мест в очереди равно $m$. Заявка, поступившая в СМО, когда все каналы и места для ожидания в очереди заняты, получает отказ (покидает СМО).

Состояния системы будем нумеровать по числу заявок, находящихся в СМО:

$$\begin{cases} S_0 - \textit{все обслуживающие каналы свободны;} \\ S_1 - \textit{занят один канал, остальные свободны;} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ S_k - \textit{заняты } k \textit{ каналов, остальные свободны;} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ S_n - \textit{заняты все } n \textit{ обслуживающих каналов;} \end{cases}$$

$S_{n+1}$ – *заняты все $n$ каналов, одно требование стоит в очереди;*

…………………………………………………

$S_{n+r}$ – *заняты все $n$ каналов, $r$ требований стоит в очереди;*

…………………………………………………

$S_{n+m}$ – *заняты все $n$ каналов, $m$ требований стоит в очереди.*

Поток требований поступает на СМО с интенсивностью $\lambda$, интенсивность обслуживания для одного устройства равна $\mu$ . Тогда граф состояний системы имеет вид, изображенный на рис. 10

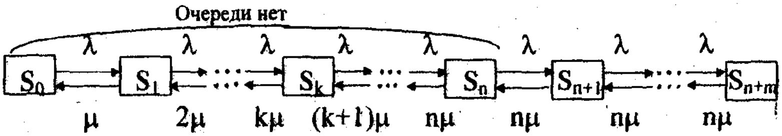


Рис. 10

Как видим, процесс, протекающий в данной СМО, является частным случаем процесса "гибели и размножения". Введя "приведенную интенсивность" $\lambda / \mu = \rho$ и используя формулы (10)-(13), мы получаем выражения для предельных вероятностей состояний:

$$\begin{cases} V_1 = \dfrac{\rho}{1!}V_0;\ V_2 = \dfrac{\rho^2}{2!}V_0;\ \ldots;\ V_n = \dfrac{\rho^n}{n!}V_0; \\ V_{n+1} = \dfrac{\rho^{n+1}}{n \cdot n!}V_0;\ V_{n+2} = \dfrac{\rho^{n+2}}{n^2 \cdot n!}V_0;\ \ldots;\ V_{n+m} = \dfrac{\rho^{n+m}}{n^m \cdot n!}V_0; \\ V_0 = \left[1 + \dfrac{\rho}{1!} + \dfrac{\rho^2}{2!} + \ldots + \dfrac{\rho^n}{n!} + \left(\dfrac{\rho^{n+1}}{n \cdot n!} + \dfrac{\rho^{n+2}}{n^2 \cdot n!} + \ldots + \dfrac{\rho^{n+m}}{n^m \cdot n!}\right)\right]^{-1} \end{cases} \tag{21}$$

В круглых скобках в выражении для $V_0$ стоит геометрическая прогрессия со знаменателем $\rho/n$. Суммируя ее, мы получаем выражение для $V_0$:

$$V_0 = \left(1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \ldots + \frac{\rho^n}{n!} + \frac{\rho^n}{n!} \cdot \frac{\left(\dfrac{\rho}{n}\right)^{m+1} - \dfrac{\rho}{n}}{\left(\dfrac{\rho}{n}\right) - 1}\right)^{-1}. \tag{22}$$

Вероятность того, что поступившее требование получает отказ, равна

$$P_{отк.} = V_{n+m} = \frac{\rho^{n+m}}{n^m \cdot n!} \cdot V_0. \tag{23}$$

Тогда относительная пропускная способность как дополнение вероятности отказа до единицы равна

$$q = 1 - V_{n+m} = 1 - \frac{\rho^{n+m}}{n^m \cdot n!} \cdot V_0. \qquad (24)$$

Абсолютная пропускная способность равна

$$A = \lambda \cdot q = \lambda \left(1 - \frac{\rho^{n+m}}{n^m \cdot n!} \cdot V_0\right). \qquad (25)$$

Найдем $\bar{k}$ – среднее число занятых обслуживанием каналов. Для данной СМО оно уже не будет совпадать со средним числом требований, находящихся в системе. $\bar{k}$ можно найти как математическое ожидание дискретной случайной величины:

$$\bar{k} = \sum_{k=1}^{n} k \cdot V_k + \sum_{j=1}^{m} n \cdot V_{n+j}. \qquad (26)$$

Так как каждый канал обслуживает в среднем $\mu$ заявок в единицу времени, а вся СМО обслуживает в среднем $A$ заявок в единицу времени, то среднее число занятых обслуживанием каналов можно найти следующим образом:

$$\bar{k} = A / \mu = \rho \left(1 - \frac{\rho^{n+m}}{n^m \cdot n!} \cdot V_0\right). \qquad (27)$$

Легко получить среднее число простаивающих устройств:

$$\bar{n}_{np.} = n - \bar{k}. \qquad (28)$$

Тогда можно определить коэффициент простоя обслуживающих каналов и коэффициент занятости

$$k_{пр.} = \frac{\bar{n}_{пр.}}{n}; \quad k_{зан.} = \frac{\bar{k}}{n} = 1 - k_{пр.} \qquad (29)$$

Среднее число требований в очереди обозначим $\bar{r}_{ож.}$ и вычислим как математическое ожидание дискретной случайной величины:

$$\bar{r}_{ож.} = 1 \cdot V_{n+1} + 2 \cdot V_{n+2} + ... + m \cdot V_{n+m}.$$

Учитывая формулы (21), получаем:

$$\bar{r}_{ож.} = \frac{\rho^{n+1}}{n \cdot n!} V_0 \left[ 1 + 2\frac{\rho}{n} + 3\left(\frac{\rho}{n}\right)^2 + ... + m\left(\frac{\rho}{n}\right)^{m-1} \right]. \qquad (30)$$

Введем обозначение $\rho / n = \alpha$. Тогда сумма в формуле (30) имеет вид

$$\bar{r}_{ож.} = 1 + 2\alpha + 3\alpha^2 + ... + m \cdot \alpha^{m-1}.$$

Если найти сумму геометрической прогрессии $\alpha + \alpha^2 + \alpha^3 + ... + \alpha^m$ со знаменателем $\alpha$ и продифференцировать затем обе части, мы получим, что

$$\bar{r}_{ож.} = 1 + 2\alpha + 3\alpha^2 + ... + m \cdot \alpha^{m-1} = \frac{1 - \alpha^m (m + 1 - m\alpha)}{(1 - \alpha)^2}. \qquad (31)$$

Среднее число требований, находящихся в системе, равно

$$\bar{r} = \bar{r}_{ож.} + \bar{k}. \qquad (32)$$

Найдем среднее время $\bar{t}_{ож.}$ ожидания требования в очереди. Поступающее в систему требование будет ожидать в очереди только тогда, когда будут заняты все обслуживающие устройства, при этом в очереди могут быть 0, 1, 2, ..., $m$ - 1 требований.

Если в очереди 0 требований, то поступающее требование будет ждать в среднем время, равное $1 / n\mu$, т.к. поток обслуженных требований $n$ каналами имеет интенсивность $n\mu$. Если в очереди 1 требование, то поступающее требование ожидает время $2 / n\mu$ (по $1 / n\mu$ на каждое впереди стоящее требование) и т.д. Если в очереди $k$ требований ($k < m$), то поступившее требование ожидает в среднем время $(k + 1) / n\mu$. Тогда

$$\bar{t}_{ож.} = \frac{1}{n\mu}V_n + \frac{2}{n\mu}V_{n+1} + \ldots + \frac{m}{n\mu}V_{n+m-1} = \frac{\rho^n}{n \cdot n!\mu}V_0\left(1 + \frac{2\rho}{n} + \frac{3\rho^2}{n^2} + \ldots + \frac{m\rho^{m-1}}{n^{m-1}}\right).$$

Можно заметить, что это выражение отличается от формулы (30) только множителем $1 / \rho\mu = 1/\lambda$ .Тогда получаем, что

$$\bar{t}_{ож.} = \frac{\bar{r}_{ож.}}{\lambda}. \qquad (33)$$

Обозначим через $T_{сист.}$ случайную величину – время нахождения требования в системе, и через $\Theta$ – случайную величину, равную времени обслуживания. Причем $\Theta = T_{обс.}$, если требование обслуживается, и $\Theta = 0$ в противном случае:

$$T_{сист.} = T_{ож.} + \Theta.$$

По теореме сложения математических ожиданий получаем

$$M(T_{сист.}) = M(T_{ож.}) + M(\Theta).$$

Отсюда среднее время $\bar{t}_{сист.}$ нахождения требования в системе, с учетом того, что $M(\Theta) = q$, $\bar{t}_{обс.} = q/\mu$, равно

$$\bar{t}_{сист.} = \bar{t}_{ож.} + q/\mu. \qquad (34)$$

*Пример 5.* Автозаправочная станция (АЗС) имеет две колонки ($n = 2$). Поток машин, прибывающих на заправку, имеет интенсивность $\lambda = 2$ (машины в минуту). Среднее время обслуживания (заправки) одной машины $\bar{t}_{обс.} = 2$ мин. Площадка у АЗС может вместить не более трех машин ($m$ = 3). Машина, прибывшая в момент, когда все три места в очереди заняты, покидает АЗС (получает отказ). Найти вероятность отказа, абсолютную и относительную пропускные способности, среднее число машин в очереди, среднее время ожидания и пребывания машины на АЗС.

*Решение.* Имеем: $n = 2$, $m = 3$, $\lambda = 2$, $\mu = 1/\bar{t}_{обс.} = 0{,}5$, $\rho = 4$, $\alpha = \rho/n = 2$.

По формулам (21)-(22) находим

$$V_0 = \frac{1}{1+\frac{4}{1}+\frac{4^2}{2}+\frac{4^2}{2}\cdot\frac{2-2^4}{1-2}} = \frac{1}{125} = 0{,}008.$$

Вероятность отказа

$$P_{отк.} = V_{n+m} = \frac{4^5}{2^3\cdot 2}V_0 = 64V_0 = 0{,}512.$$

Относительная пропускная способность $q = 1 - P_{отк.} = 0{,}488$.

Таким образом, более половины машин, прибывших на АЗС, не могут быть обслужены.

Абсолютная пропускная способность $A = q \cdot \lambda = 0{,}976$ машин в минуту.

Среднее число занятых колонок $\bar{k} = \dfrac{A}{\mu} = \dfrac{0{,}976}{0{,}5} = 1{,}952$, т.е. обе колонки почти все время заняты.

Среднее число машин в очереди

$$\bar{r}_{ож.} = \frac{4^3}{2\cdot 2\cdot 125}\cdot\frac{1-4\cdot 2^3+3\cdot 2^4}{(1-2)^3} = \frac{16}{125}\cdot 17 = 2{,}18.$$

Среднее время ожидания в очереди: $\bar{t}_{ож.} = \dfrac{\bar{r}_{ож.}}{\lambda} = \dfrac{2{,}18}{2} \approx 1{,}09$ минуты.

Среднее время пребывания машины на АЗС (включая время обслуживания)

$$\bar{t}_{сист.} = \bar{t}_{ож.} + q \cdot \bar{t}_{обс.} = 1{,}09 + 0{,}976 = 2{,}07 \text{ минуты.}$$

Выше была рассмотрена СМО, когда длина очереди ограничена некоторым числом $m$. Посмотрим, что будет происходить, если длина очереди не ограничена, а может быть сколь угодно большой.

Рассмотрим многоканальную СМО с неограниченной очередью. Граф состояний системы имеет вид, изображенный на рис. 11.

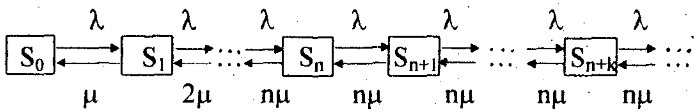


Рис. 11

Вероятности состояний системы найдем предельным переходом (при $m \to \infty$) из формул (21)-(22). Пусть $\rho/n < 1$ (иначе при $\rho/n \geq 1$ очередь будет бесконечно возрастать), тогда сумма соответствующей геометрической прогрессии сходится. Отсюда получаем

$$\begin{cases} V_1 = \dfrac{\rho}{1!}V_0;\ V_2 = \dfrac{\rho^2}{2!}V_0;\ \ldots;\ V_n = \dfrac{\rho^n}{n!}V_0; \\ V_{n+1} = \dfrac{\rho^{n+1}}{n \cdot n!}V_0;\ V_{n+2} = \dfrac{\rho^{n+2}}{n^2 \cdot n!}V_0;\ \ldots;\ V_{n+k} = \dfrac{\rho^{n+k}}{n^k \cdot n!}V_0;\ldots; \\ V_0 = \left[1 + \dfrac{\rho}{1!} + \dfrac{\rho^2}{2!} + \ldots + \dfrac{\rho^n}{n!} + \dfrac{\rho^{n+1}}{n!(n-\rho)}\right]^{-1}. \end{cases} \tag{35}$$

Так как каждое поступающее в систему требование будет обслужено в любом случае, то $P_{отк.} = 0$, $q = 1$, $A = \lambda \cdot q = \lambda$.

Среднее число занятых обслуживанием устройств найдем из (27):

$$\bar{n}_{з.} = \frac{A}{\mu} = \frac{\lambda}{\mu} = \rho. \tag{36}$$

Среднее число требований в очереди получим из формулы (30) при $m \to \infty$:

$$\bar{r}_{ож.} = \frac{\rho^{n+1}}{n \cdot n!} \cdot \frac{1}{\left(1 - \frac{\rho}{n}\right)^2} \cdot V_0. \qquad (37)$$

Среднее время ожидания требования в очереди и среднее число требований, находящихся в системе, будем находить по формулам (32)-(33) соответственно.

<u>*Пример 6.*</u> Автозаправочная станция с двумя колонками ($n$ = 2) обслуживает поток машин с интенсивностью $\lambda = 0{,}8$ машин в минуту. Среднее время обслуживания одной машины $\bar{t}_{обс.} = \frac{1}{\mu} = 2$ мин.

В данном районе нет другой АЗС, так что очередь машин перед АЗС может расти практически неограниченно. Найти характеристики СМО.

<u>*Решение.*</u> Имеем: $n = 2$, $\lambda = 0{,}8$, $\mu = \frac{1}{\bar{t}_{обс.}} = 0{,}5, \rho = 1{,}6, \alpha = \frac{\rho}{n} = 0{,}8$.

Поскольку $\alpha < 1$, очередь не растет безгранично и имеет смысл говорить о предельном стационарном режиме работы СМО. По формулам (35) находим вероятности состояний:

$$V_0 = \left[1 + 1{,}6 + 1{,}28 + \frac{4{,}09}{2 \cdot 0{,}4}\right]^{-1} \approx 0{,}111;$$

$$V_1 = 1{,}6 V_0 \approx 0{,}178; V_2 = 1{,}28 V_0 \approx 0{,}142;$$

$$V_3 = \frac{1{,}6^3}{2 \cdot 2!} V_0 \approx 0{,}114; V_4 = \frac{1{,}6^4}{2^2 \cdot 2!} V_0 \approx 0{,}091$$

и т.д.

Среднее число занятых каналов найдем, разделив абсолютную пропускную способность СМО $A = \lambda = 0{,}8$ на интенсивность обслуживания $\mu = 0{,}5$:

$$\bar{k} = \frac{0{,}8}{0{,}5} = 1{,}6.$$

Вероятность отсутствия очереди у АЗС будет

$$V_0 + V_1 + V_2 \approx 0{,}431.$$

Среднее число машин в очереди

$$\bar{r}_{ож.} \approx \frac{1{,}6^3 \cdot 0{,}111}{2 \cdot 2 \cdot 0{,}42^2} \approx 0{,}71.$$

Среднее число машин на АЗС

$$\bar{r} = \bar{k} + \bar{r}_{ож.} \approx 0{,}71 + 1{,}6 = 2{,}31.$$

Среднее время ожидания в очереди

$$\bar{t}_{ож.} = \frac{\bar{r}_{ож.}}{\lambda} \approx 0{,}89 \text{ мин.}$$

Среднее время пребывания машины на АЗС

$$\bar{t}_{сист.} = \bar{t}_{ож.} + \bar{t}_{обс.} \approx 0{,}89 + 2 = 2{,}89 \text{ мин.}$$

*Замечание.* Для нахождения характеристик работы одноканальных СМО с ограниченной очередью и неограниченной очередью нужно в вышеприведенных расчетных формулах для многоканальной СМО положить $n = 1$.

## 7. ЗАМКНУТЫЕ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ

В вышерассмотренных СМО заявки на обслуживание приходили извне и интенсивность потока заявок не зависела от состояния самой системы.

СМО, в которых интенсивность общего потока заявок зависит от того, сколько заявок связано с процессом обслуживания, т.е. от состояния самой системы, называется *замкнутой*. Элементами замкнутой СМО являются как обслуживающие каналы, так и источники заявок. Для замкнутой СМО характерно ограниченное число источников заявок.

Рассмотрим задачу обслуживания рабочим n станков. Интенсивность потока неисправностей каждого станка равна $\lambda$. Интенсивность общего потока заявок зависит от того, сколько имеется неисправных станков, т.к. при выходе из строя станок перестает быть источником заявок. Таким образом, интенсивность потока станков, требующих обслуживания, зависит от состояния самой системы, следовательно, это замкнутая СМО. На наладку станка рабочий тратит в среднем $\bar{t}_{обс.} = \dfrac{1}{\mu}$, где $\mu$ – интенсивность потока отремонтированных станков. Число станков $n$ – ограничено.

Наша система имеет следующие состояния:

$S_0$ – все станки исправны (рабочий свободен);

$S_1$ – один станок неисправен (рабочий занят его наладкой);

$S_2$ – два станка неисправны, один налаживается, другой находится в очереди;

.............................................................................

$S_n$ – все $n$ станков неисправны, один налаживается, ($n$ - 1) станков находятся в очереди.

Граф состояний имеет вид, показанный на рис. 12.

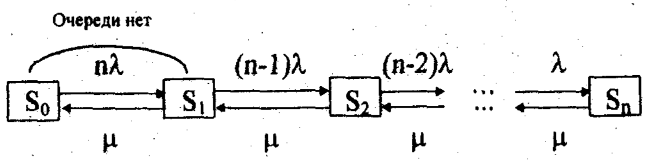


Рис. 12

Как видим, процесс, протекающий в данной СМО, является частным случаем процесса "гибели и размножения". Введя обозначение $\dfrac{\lambda}{\mu} = \rho$, из формул (10)-(13) получаем

$$V_1 = \frac{\lambda_{01}}{\lambda_{10}} V_0 = \frac{n\lambda}{\mu} V_0 = n\rho V_0 ;$$

$$V_2 = \frac{\lambda_{01}\lambda_{12}}{\lambda_{21}\lambda_{10}} = \frac{n\lambda \cdot (n-1)}{\mu^2} V_0 = n(n-1)\rho^2 V_0 ;$$

...............................................

$$V_n = \frac{\lambda_{01}\lambda_{12}...\lambda_{n-1,n}}{\lambda_{n,n-1}...\lambda_{21}\lambda_{10}} V_0 = \frac{n(n-1)...1 \cdot \lambda^n}{\mu^n} V_0 = n(n-1)...1 \cdot \rho^n V_0 ; \quad (38)$$

$$V_0 = \left( 1 + \frac{\lambda_{01}}{\lambda_{10}} + \frac{\lambda_{01}\lambda_{12}}{\lambda_{21}\lambda_{10}} + ... + \frac{\lambda_{01}\lambda_{12}...\lambda_{n-1,n}}{\lambda_{n,n-1}...\lambda_{21}\lambda_{10}} \right)^{-1} =$$

$$= \left( 1 + n\rho + n(n-1)\rho^2 + ... + n(n-1)...1 \cdot \rho^n \right)^{-1}.$$

Вероятность того, что рабочий будет свободен (соответствует вероятности состояния $S_0$), равна $V_0$. Поэтому вероятность того, что рабочий занят наладкой станка, равна:

$$P_{зан.} = 1 - V_0. \quad (39)$$

Если рабочий обслуживает $\mu$ станков в единицу времени, то абсолютная пропускная способность системы (среднее число станков, обслуженных рабочим в единицу времени) равна

$$A = P_{зан.} \cdot \mu = (1 - V_0) \cdot \mu. \quad (40)$$

Поскольку каждая заявка (станок) будет обслужена, то относительная пропускная способность $q = 1$.

Среднее число станков, связанных с процессом обслуживания (математическое ожидание числа неисправных станков), будет равно

$$\overline{\omega} = 1 \cdot V_1 + 2 \cdot V_2 + ... + n \cdot V_n. \quad (41)$$

Общее число $W$ станков, связанных с обслуживанием, равно $W = R + U$, где $R$ – число станков в очереди; $U$ – число станков, которые обслуживаются (ремонтируются), $U = \{0, 1\}$. Тогда среднее значение $U$ (ее математическое ожидание) равно

$$\bar{u} = 0 \cdot V_0 + 1 \cdot V_{зан.} = 1 - V_0.$$

Вычитая эту величину из $\bar{w}$, получим среднее число станков $\bar{r}$, находящихся в очереди в ожидании обслуживания:

$$\bar{r} = \bar{w} - \bar{u}. \qquad (42)$$

Важной характеристикой данной СМО является производительность группы станков, обслуживаемых рабочим.

Средняя относительная потеря производительности за счет неисправностей равна $L = \bar{w} \cdot l$, где $\bar{w}$ – среднее число неисправных станков, а $l$ – производительность исправного станка в единицу времени.

*Пример 7*. Рабочий обслуживает три станка. Каждый станок останавливается в среднем 2 раза в час. Процесс наладки занимает у рабочего в среднем 10 минут. Определить характеристики замкнутой СМО: вероятность занятости рабочего; абсолютную пропускную способность, среднюю относительную потерю производительности станков за счет неисправностей.

*Решение*. По условию имеем: $n = 3$, $\lambda = 2$, $\mu = \frac{1}{t_{обс.}} = \frac{1}{1/6}$, $\rho = \frac{\lambda}{\mu} = \frac{1}{3}$. По формулам (38) находим:

$$V_0 = \frac{1}{1 + 3 \cdot \frac{1}{3} + 3 \cdot 2 \cdot \frac{1}{3^2} + 3 \cdot 2 \cdot 1 \cdot \frac{1}{3^3}} \approx 0{,}346.$$

$$V_1 = 3 \cdot \frac{1}{3} V_0 \approx 0{,}346.$$

$$V_2 = 3 \cdot 2 \cdot \left(\frac{1}{3}\right)^2 \cdot 0{,}346 \approx 0{,}231.$$

$$V_3 = 3 \cdot 2 \cdot 1 \left(\frac{1}{3}\right)^3 \cdot 0{,}346 \approx 0{,}077.$$

Вероятность занятости рабочего $P_{зан} = 1 - V_0 \approx 0{,}654$

Абсолютная пропускная способность рабочего (среднее число неисправностей, которое он ликвидирует за час) равна $A = 0{,}654 \cdot 6 = 3{,}94$.

Среднее число неисправных станков находим по формуле (41):

$$\bar{w} = 1 \cdot 0{,}346 + 2 \cdot 0{,}231 + 3 \cdot 0{,}077 \approx 1{,}04.$$

Средняя относительная потеря производительности станков за счет неисправностей $\bar{w}/n = 0{,}347$. Таким образом, за счет неисправностей станков теряется около 35% их общей производительности.

Рассмотрим теперь более общую замкнутую СМО. Предположим, что имеется $n$ устройств, которые могут потребовать обслуживания, и $m$ обслуживающих каналов ($m < n$).

Перечислим возможные состояния системы.

*$S_0$ – все устройства работают, обслуживающие каналы свободны;*

*$S_1$ – одно устройство вышло из строя, один канал занят обслуживанием;*

……………………………………………………………………………

*$S_m$ – $m$ устройств вышло из строя, все каналы заняты обслуживанием;*

*$S_{m+1}$ – $(m + 1)$ устройство вышло из строя, $m$ из них налаживаются, одно ждет в очереди;*

*$S_n$ – все $n$ устройств вышли из строя, $m$ из них налаживаются, $(n - m)$ ждут в очереди.*

Если обозначить, как обычно, через $\lambda$ – интенсивность, с которой одно устройство выходит из строя, а через $\mu$ – интенсивность потока обслуживаний одним каналом, и учитывая, что процессы, протекающие в СМО, являются марковскими, можно построить граф состояний СМО. Он имеет вид, изображенный на рис. 13.

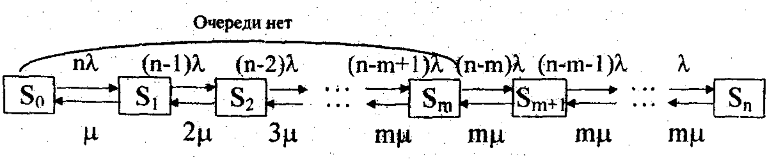


Рис. 13

Как обычно, интенсивности потоков событий, переводящих систему из состояния в состояние, проставлены у стрелок. Воспользовавшись правилом составления уравнений Колмогорова в установившемся режиме (формулы (9)-(13)), находим предельные вероятности:

$$
\left.
\begin{aligned}
&V_1 = \frac{n}{1}\cdot\frac{\lambda}{\mu}V_0,\\
&V_2 = \frac{n(n-1)}{1\cdot 2}\cdot\left(\frac{\lambda}{\mu}\right)^2 V_0,\\
&V_3 = \frac{n(n-1)(n-2)}{1\cdot 2\cdot 3}\cdot\left(\frac{\lambda}{\mu}\right)^3 V_0,\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&V_m = \frac{n(n-1)(n-2)\dots(n-m+1)}{1\cdot 2\cdot 3\dots m}\cdot\left(\frac{\lambda}{\mu}\right)^m V_0,\\
&V_{m+1} = \frac{n(n-1)\dots(n-m)}{1\cdot 2\dots m\cdot m}\cdot\left(\frac{\lambda}{\mu}\right)^{m+1} V_0,\\
&V_{m+2} = \frac{n(n-1)\dots(n-m-1)}{1\cdot 2\dots m\cdot m^2}\cdot\left(\frac{\lambda}{\mu}\right)^{m+2} V_0,\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&V_n = \frac{n(n-1)\dots 1}{1\cdot 2\dots m\cdot m^{n-m}}\cdot\left(\frac{\lambda}{\mu}\right)^n \cdot V_0,\\
&V_0 = \left[1+\frac{n}{1}\cdot\frac{\lambda}{\mu}+\frac{n(n-1)}{1\cdot 2}\cdot\left(\frac{\lambda}{\mu}\right)^2+\dots+\frac{n(n-1)\dots(n-m+1)}{1\cdot 2\dots m}\cdot\left(\frac{\lambda}{\mu}\right)^m+\right.\\
&\left.+\frac{n(n-1)\dots(n-m)}{1\cdot 2\dots m\cdot m}\cdot\left(\frac{\lambda}{\mu}\right)^{m+1}+\dots+\frac{n(n-1)\dots 1}{1\cdot 2\dots m\cdot m^{n-m}}\left(\frac{\lambda}{\mu}\right)^n\right]^{-1}
\end{aligned}
\right\}
\quad (43)
$$

Обозначая, как всегда, $\frac{\lambda}{\mu}=\rho$, приведем формулы к виду

$$\left.\begin{aligned}
V_0 &= \left[1 + \frac{n}{1!}\rho + \frac{n(n-1)}{2!}\rho^2 + \ldots + \frac{n(n-1)\ldots(n-m+1)}{m!}\rho^m + \right. \\
&\quad \left. + \frac{n(n-1)\ldots(n-m)}{m!\cdot m}\rho^{m+1} + \ldots + \frac{n(n-1)\ldots 1}{m!\cdot m^{n-m}}\rho^n\right]^{-1}; \\
V_1 &= \frac{n}{1!}\rho V_0; \\
V_2 &= \frac{n(n-1)}{2!}\rho^2 V_0; \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
V_m &= \frac{n(n-1)\ldots(n-m+1)}{m!}\rho^m V_0; \\
V_{m+1} &= \frac{n(n-1)\ldots(n-m)}{m!\cdot m}\rho^{m+1} V_0; \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
V_n &= \frac{n(n-1)\ldots 1}{m!\cdot m^{n-m}}\rho^n V_0.
\end{aligned}\right\} \qquad (44)$$

Через эти вероятности выражается среднее число $\bar{z}$ занятых рабочих:

$$\begin{aligned}
\bar{z} &= 0 \cdot V_0 + 1 \cdot V_1 + 2 \cdot V_2 + \ldots + m\,(V_m + V_{m+1} + \ldots + V_n) = \\
&= V_1 + 2 \cdot V_2 + \ldots + (m-1)\,V_{m-1} + m\,(1 - V_0 - V_1 - \ldots - V_{m-1}).
\end{aligned} \qquad (45)$$

Через $\bar{z}$ выражается, в свою очередь, среднее число станков, обслуживаемых бригадой в единицу времени (абсолютная пропускная способность):

$$A = \bar{z} \cdot \mu, \qquad (46)$$

а также среднее число неисправных станков

$$\bar{\omega} = n - \frac{\bar{z} \cdot \mu}{\lambda} = n - \frac{\bar{z}}{\rho}. \qquad (47)$$

Отсюда же находится и средняя потеря производительности группы станков в единицу времени за счет неисправностей: нужно умножить среднее число неисправных станков $\overline{w}$ на производительность $l$ одного станка в единицу времени.

<u>*Пример 8.*</u> Два рабочих обслуживают группу из шести станков. Остановки каждого (работающего) станка случаются, в среднем, через каждые полчаса. Процесс наладки занимает у рабочего в среднем 10 минут. Определить характеристики замкнутой СМО: среднее число занятых рабочих; абсолютную пропускную способность; среднее количество неисправных станков.

<u>*Решение.*</u> Имеем: $n = 6$, $m = 2$, $\lambda = 2$, $\mu = \frac{1}{\bar{t}_{обс.}} = 6$, $\rho = \frac{\lambda}{\mu} = \frac{1}{3}$.

По формулам (44)

$$V_0 = \left(1 + \frac{6}{1}\cdot\frac{1}{3} + \frac{6\cdot5}{1\cdot2}\cdot\frac{1}{3^2} + \frac{6\cdot5\cdot4}{1\cdot2\cdot2}\cdot\frac{1}{3^3} + \frac{6\cdot5\cdot4\cdot3}{1\cdot2\cdot2}\cdot\frac{1}{3^3} + \frac{6\cdot5\cdot4\cdot3}{1\cdot2\cdot2^2}\cdot\frac{1}{3^4} + \right.$$

$$\left. + \frac{6\cdot5\cdot4\cdot3\cdot2}{1\cdot2\cdot2^3}\cdot\frac{1}{3^5} + \frac{6\cdot5\cdot4\cdot3\cdot2\cdot1}{1\cdot2\cdot2^4}\cdot\frac{1}{3^6}\right)^{-1} = \frac{1}{6{,}549} \approx 0{,}153;$$

$$V_1 = \frac{6}{1}\cdot\frac{1}{3}\cdot 0{,}153 \approx 0{,}306.$$

Вероятность того, что два рабочих заняты обслуживанием, равна $(1 \cdot V_0 - V_1)$.

$$\bar{z} = 1 \cdot V_1 + 2\,(1 - V_0 - V_1) = 1 \cdot 0{,}153 + 2 \cdot 0{,}541 \approx 1{,}235.$$

По формуле (46) находим абсолютную пропускную способность

$$A = 1{,}235 \cdot 6 = 7{,}41.$$

По формуле (47) находим среднее число неисправных станков

$$\overline{w} = 6 - \frac{7{,}41}{2} = 2{,}295.$$

## Содержание



Учебное издание

КОРЗНИКОВ Александр Дмитриевич
ПАВЛОВ Валерий Валентинович

ЭЛЕМЕНТЫ ТЕОРИИ МАРКОВСКИХ ПРОЦЕССОВ И СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ

Методическое пособие для студентов инженерно-экономических специальностей втузов

Редактор Т.Н.Микулик
Компьютерная верстка Н.А.Школьниковой

Подписано в печать 15.03.2001.
Формат 60x84 1/16. Бумага типографская №2.
Печать офсетная. Гарнитура книжно-журнальная.
Усл.печ. л. 2,3. Уч.-изд. л. 1,8. Тираж 200. Заказ 406.

Издатель и полиграфическое исполнение:
Белорусская государственная политехническая академия.
Лицензия ЛВ №155 от 30.01.98.
220027, Минск, проспект Ф.Скорины, 65.