

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ»

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ПРОВЕДЕНИЮ ПРАКТИЧЕСКИХ (СЕМИНАРСКИХ) ЗАНЯТИЙ

по дисциплине (модулю)

Б1.В.ДВ.1.2 Моделирование систем массового обслуживания

Код и направление
подготовки

38.06.01 Экономика

Наименование профиля / программы
подготовки научно-педагогических
кадров в аспирантуре/магистерской
программы / специализация

***Математические и
инструментальные
методы экономики***

Квалификация
(степень) выпускника

***Исследователь.
Преподаватель
исследователь***

Факультет

***Прикладной
информатики***

Кафедра – разработчик

***Компьютерных
технологий и систем***

Ведущий преподаватель

Лойко В.И.

Краснодар 2015

ПЗ 1. Исследование уравнений Колмогорова (построение графа состояний системы)

Рассмотрим двухканальную систему массового обслуживания ($n = 2$) с ограниченной очередью равной шести ($m = 4$). В СМО поступает простейший поток заявок со средней интенсивностью $\lambda = 4,8$ и показательным законом распределения времени между поступлением заявок. Поток обслуживаемых в системе заявок является простейшим со средней интенсивностью $\mu = 2$ и показательным законом распределения временем обслуживания.

Данная система имеет 7 состояний, обозначим их:

S_0 – система свободная, нет заявок;

S_1 – 1 заявка на обслуживании, очередь пуста;

S_2 – 2 заявки на обслуживании, очередь пуста;

S_3 – 2 заявки на обслуживании, 1 заявка в очереди;

S_4 – 2 заявки на обслуживании, 2 заявки в очереди;

S_5 – 2 заявки на обслуживании, 3 заявки в очереди;

S_6 – 2 заявки на обслуживании, 4 заявки в очереди;

Вероятности прихода системы в состояния $S_0, S_1, S_2, \dots, S_6$ соответственно равны $P_0, P_1, P_2, \dots, P_6$.

Граф состояний системы массового обслуживания представляет собой схему гибели и размножения. Все состояния системы можно представить в виде цепочки, в которой каждое из состояний связано с предыдущим и последующим.

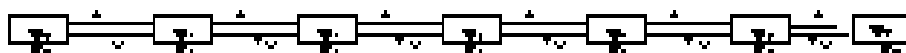


Рис. 3. Граф состояний двухканальной СМО

Для построенного графа запишем уравнения Колмогорова:

$$\begin{aligned}\frac{dP_0}{dt} &= \mu P_1 - \lambda P_0 \\ \frac{dP_1}{dt} &= \lambda P_0 + 2\mu P_2 - P_1(\mu + \lambda) \\ \frac{dP_2}{dt} &= \lambda P_1 + 2\mu P_3 - P_2(2\mu + \lambda) \\ \frac{dP_3}{dt} &= \lambda P_2 + 2\mu P_4 - P_3(2\mu + \lambda) \\ \frac{dP_4}{dt} &= \lambda P_3 + 2\mu P_5 - P_4(2\mu + \lambda) \\ \frac{dP_5}{dt} &= \lambda P_4 + 2\mu P_6 - P_5(2\mu + \lambda) \\ \frac{dP_6}{dt} &= \lambda P_5 - 2\mu P_6 \quad \text{и} \quad P_0 + P_1 + P_2 + P_3 + P_4 + P_5 + P_6 = 1\end{aligned}$$

Чтобы решить данную систему зададим начальные условия:

$$P_0(0) = 1, \quad P_5(0) = 0,$$

$$P_1(0) = 0, \quad P_6(0) = 0.$$

$$P_2(0) = 0,$$

$$P_3(0) = 0,$$

$$P_4(0) = 0.$$

Систему уравнений Колмогорова (систему дифференциальных уравнений) решим численным методом Эйлера с помощью программного пакета Maple 11.

Метод Эйлера

$$\frac{dP_i}{dt} = f(P_0, P_1, \dots, P_n), \quad i = 0, 1, \dots, n$$

где $f(P_0, P_1, \dots, P_n)$ - в нашем случае, это правые части уравнений Колмогорова, $n=6$.

$$\begin{aligned} \frac{dP_i}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{\Delta P_i}{\Delta t} \approx \frac{\Delta P_i}{\Delta t} = \frac{P_i(t + \Delta t) - P_i(t)}{\Delta t} \\ \frac{P_i(t + \Delta t) - P_i(t)}{\Delta t} &= f(P_0(t), P_1(t), \dots, P_n(t)) \\ P_i(t + \Delta t) - P_i(t) &= f(P_0(t), P_1(t), \dots, P_n(t)) \Delta t \\ P_i(t + \Delta t) &= P_i(t) + f(P_0(t), P_1(t), \dots, P_n(t)) \Delta t \quad (1) \end{aligned}$$

Выберем шаг по времени $\Delta t = 0.01$. Предположим $T = 10$, где T – это время, за которое система выходит на стационарный режим. Отсюда получаем

$$N = \frac{T}{\Delta t} = \frac{10}{0.01} = 1000$$

число шагов

Последовательно N раз вычисляя $P_i(t + \Delta t)$ по формуле (1) получим зависимости вероятностей состояний системы от времени, приведенной на рис. 4.

Значения вероятностей СМО при $T = 10$ равны:

$$P_0 = 0.04027683531$$

$$P_1 = 0.09666136713$$

$$P_2 = 0.1159892579$$

$$P_3 = 0.1391804173$$

$$P_4 = 0.1670087586$$

$$P_5 = 0.2004034991$$

$$P_6 = 0.2404798682$$

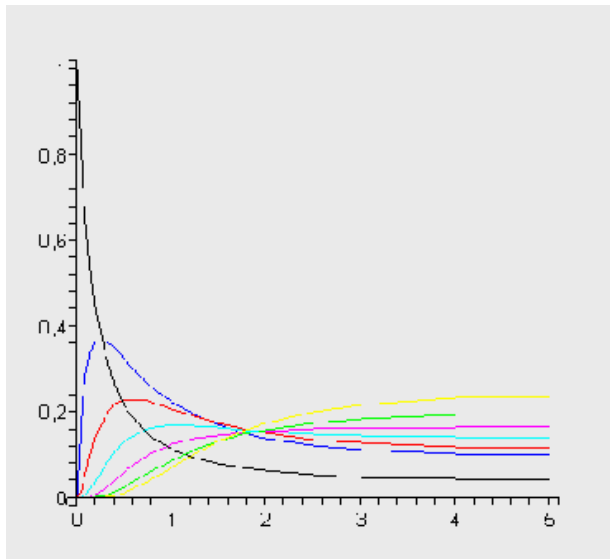


Рис. 4. Зависимости вероятностей состояний системы от времени

Финальные вероятности системы

При достаточно большом времени протекания процессов в системе ($T \gg \frac{1}{\lambda}, \frac{1}{\mu}$) могут устанавливаться вероятности состояний, не зависящие от времени, которые называются финальными вероятностями, т.е. в системе устанавливается стационарный режим. Если число состояний системы конечно, и из каждого из них за конечное число шагов можно перейти в любое другое состояние, то финальные вероятности существуют, т.е.

$$\lim_{t \rightarrow \infty} P_i(t) = P_i$$

Т.к. в стационарном состоянии производные по времени равны 0, то уравнения для финальных вероятностей получаются из уравнений Колмогорова путем приравнивания правых частей 0. Запишем уравнения для финальных вероятностей для нашей СМО.

$$\mu P_1 - \lambda P_0 = 0$$

$$2\mu P_2 - \lambda P_1 = 0$$

$$2\mu P_3 - \lambda P_2 = 0$$

$$2\mu P_4 - \lambda P_3 = 0$$

$$2\mu P_5 - \lambda P_4 = 0$$

$$2\mu P_6 - \lambda P_5 = 0$$

$$P_0 + P_1 + P_2 + P_3 + P_4 + P_5 + P_6 = 1$$

Решим данную систему линейных уравнений с помощью программного

пакета Maple 11 (см. Приложение 1).

Получим финальные вероятности системы:

$$P_{f0} := 0.04027092993$$

$$P_{f1} := 0.09665023183$$

$$P_{f2} := 0.1159802782$$

$$P_{f4} := 0.1670116006$$

$$P_{f5} := 0.2004139207$$

$$P_{f6} := 0.2404967048$$

Сравнение вероятностей, полученных из системы уравнений Колмогорова при $T = 10$, с финальными вероятностями показывает, что ошибки равны:

$$R_0 := 0.00103943599$$

$$R_1 := 0.00195995988$$

$$R_2 := 0.0015805102$$

$$R_3 := 0.0007186699$$

$$R_4 := 0.0005003733$$

$$R_5 := 0.0018345262$$

$$R_6 := 0.0029636748$$

Т.е. достаточно малы. Это подтверждает правильность полученных результатов.

Расчет показателя эффективности системы по финальным вероятностям

Найдем показатели эффективности системы массового обслуживания.

Сначала вычислим приведенную интенсивность потока заявок:

$$\rho = \frac{\lambda}{\mu} = \frac{4,8}{2} = 2,4$$

1) Вероятность отказа в обслуживании заявки, т.е. вероятность того, что заявка покидает систему не обслуженной. В нашем случае заявке отказывается в обслуживании, если все 2 канала заняты, и очередь максимально заполнена (т.е. 4 человек в очереди), это соответствует состоянию системы S_6 . Т.к. вероятность прихода системы в состояние S_6 равна P_6 , то

$$P_{\text{отказ}} = P_6 = 0,2404967048$$

2) Относительная пропускная способность – это средняя доля поступивших заявок, обслуживаемых системой.

$$Q = 1 - P_{i,i} = 1 - 0,2404967048 = 0,7595032952$$

3) Абсолютная пропускная способность – это среднее число заявок, обслуживаемых в единицу времени.

$$A = \lambda Q = 4,8 \cdot 0,7595032952 = 3,645615817$$

4) Средняя длина очереди, т.е. среднее число заявок в очереди, равна сумме произведений числа заявок в очереди на вероятность соответствующего состояния.

$$L_0 = 1 \cdot P_3 + 2 \cdot P_4 + 3 \cdot P_5 + 4 \cdot P_6 = 2,036428116$$

5) Среднее время пребывания заявки в очереди **определяется формулой Литтла:**

$$T_0 = \frac{L_0}{\lambda} = \frac{2,036428116}{4,8} = 0,4242558575(i \rightarrow i)$$

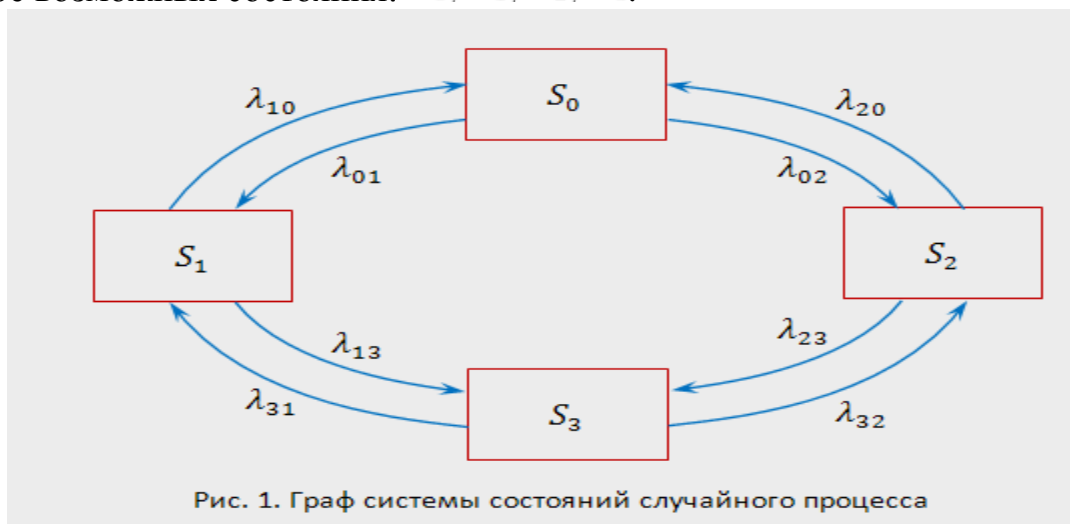
6) Среднее число занятых каналов **определяется следующим образом:**

$$k = \frac{A}{\mu} = \frac{3,645615817}{2} = 1,822807908$$

ПЗ 2. Исследование уравнений Колмогорова (создание правил составления уравнений Колмогорова)

Рассмотрим математическое описание марковского процесса с дискретными состояниями и непрерывным временем* на примере случайного процесса из примера 1, граф которого изображен на рис. 1. Будем полагать, что все переходы системы из состояния S_i в S_j происходят под воздействием простейших потоков событий с интенсивностями λ_{ij} ($i, j = 0, 1, 2, 3$); так, переход системы из состояния S_0 в S_1 будет происходить под воздействием потока отказов первого узла, а обратный переход из состояния S_1 в S_0 — под воздействием потока "окончаний ремонтов" первого узла и т.п.

Граф состояний системы с проставленными у стрелок интенсивностями будем называть **размеченным** (см. рис. 1). Рассматриваемая система S имеет четыре возможных состояния: S_0, S_1, S_2, S_3 .



Вероятностью i -го состояния называется вероятность $p_i(t)$ того, что в момент t система будет находиться в состоянии S_i . Очевидно, что для любого момента t сумма вероятностей всех состояний равна единице:

$$\sum_{i=0}^3 p_i(t) = p_0(t) + p_1(t) + p_2(t) + p_3(t) = 1.$$

8)

Рассмотрим систему в момент t , задав малый промежуток Δt , найдем вероятность $p_0(t + \Delta t)$ того, что система в момент $t + \Delta t$ будет находиться в состоянии S_0 . Это достигается разными способами.

1. Система в момент t с вероятностью $p_0(t)$ находилась в состоянии S_0 , а за время Δt не вышла из него.

Вывести систему из этого состояния (см. граф на рис. 1) можно суммарным простейшим потоком с интенсивностью $(\lambda_{01} + \lambda_{02})$, т.е. в соответствии с формулой (7), с вероятностью, приближенно равной $(\lambda_{01} + \lambda_{02})\Delta t$. А вероятность того, что система не выйдет из состояния S_0 , равна $[1 - (\lambda_{01} + \lambda_{02})\Delta t]$. Вероятность того, что система будет находиться в состоянии S_0 по первому способу (т.е. того, что находилась в состоянии S_0 и не выйдет из него за время Δt), равна по теореме умножения вероятностей:

$$p_0(t)[1 - (\lambda_{01} + \lambda_{02})\Delta t].$$

2. Система в момент t с вероятностями $p_1(t)$ (или $p_2(t)$) находилась в состоянии S_1 (или S_2) за время Δt перешла в состояние S_0 .

Потоком интенсивностью λ_{10} (или λ_{20} — см. рис. 1) система перейдет в состояние S_0 с вероятностью, приближенно равной $\lambda_{10}\Delta t$ (или $\lambda_{20}\Delta t$). Вероятность того, что система будет находиться в состоянии S_0 по этому способу, равна $1(t)\lambda_{10}\Delta t$ (или $2(t)\lambda_{20}\Delta t$).

Применяя теорему сложения вероятностей, получим

откуда	$p_0(t + \Delta t) = p_1(t)\lambda_{10}\Delta t + p_2(t)\lambda_{20}\Delta t + p_0(t)[1 - (\lambda_{01} + \lambda_{02})\Delta t],$ $\frac{p_0(t + \Delta t) - p_0(t)}{\Delta t} = p_1(t)\lambda_{10} + p_2(t)\lambda_{20} - (\lambda_{01} + \lambda_{02})p_0(t).$
--------	---

Переходя к пределу при $\Delta t \rightarrow 0$ (приближенные равенства, связанные с применением формулы (7), перейдут в точные), получим в левой части уравнения производную $p'_0(t)$ (обозначим ее для простоты p'_0):

$$p'_0 = \lambda_{10}p_1 + \lambda_{20}p_2 - (\lambda_{01} + \lambda_{02})p_0.$$

Получили дифференциальное уравнение первого порядка, т.е. уравнение, содержащее как саму неизвестную функцию, так и ее производную первого порядка.

Рассуждая аналогично для других состояний системы S , можно получить систему дифференциальных уравнений Колмогорова для вероятностей состояний:

(9)

$$\begin{cases} p'_0 = \lambda_{10}p_1 + \lambda_{20}p_2 - (\lambda_{01} + \lambda_{02})p_0, \\ p'_1 = \lambda_{01}p_0 + \lambda_{31}p_3 - (\lambda_{10} + \lambda_{13})p_1, \\ p'_2 = \lambda_{02}p_0 + \lambda_{32}p_3 - (\lambda_{20} + \lambda_{23})p_2, \\ p'_3 = \lambda_{13}p_1 + \lambda_{23}p_2 - (\lambda_{31} + \lambda_{32})p_3. \end{cases}$$

Сформулируем **правило составления уравнений Колмогорова**. В левой части каждого из них стоит производная вероятности i -го состояния. В правой части — сумма произведений вероятностей всех состояний (из которых идут стрелки в данное состояние) на интенсивности соответствующих потоков событий, минус суммарная интенсивность всех потоков, выводящих систему из данного состояния, умноженная на вероятность данного (i -го состояния).

В системе (9) независимых уравнений на единицу меньше общего числа уравнений. Поэтому для решения системы необходимо добавить уравнение (8).

Особенность решения дифференциальных уравнений вообще состоит в том, что требуется задать так называемые начальные условия, т.е. в данном случае вероятности состояний системы в начальный момент $t = 0$. Так, например, систему уравнений (9) естественно решать при условии, что в начальный момент оба узла исправны и система находилась в состоянии S_0 , т.е. при начальных условиях $p_0(0) = 1, p_1(0) = p_2(0) = p_3(0) = 0$.

Уравнения Колмогорова дают возможность найти все вероятности состояний как **функции времени**. Особый интерес представляют вероятности системы $p_i(t)$ в **предельном стационарном режиме**, т.е. при $t \rightarrow \infty$, которые называются **предельными (или финальными) вероятностями** состояний.

В теории случайных процессов доказывается, что *если число состояний системы конечно и из каждого из них можно (за конечное число шагов) перейти в любое другое состояние, то предельные вероятности существуют*.

Предельная вероятность состояния S_i имеет четкий смысл: она показывает **среднее относительное время пребывания системы в этом состоянии**. Например, если предельная вероятность состояния S_0 , т.е. $p_0 = 0,5$, то это означает, что в среднем половину времени система находится в состоянии S_0 .

Так как предельные вероятности постоянны, то, заменяя в уравнениях Колмогорова их производные нулевыми значениями, получим систему линейных алгебраических уравнений, описывающих стационарный режим. Для системы Sc графом состояний, изображенном на рис. 1), такая система уравнений имеет вид:

$$\begin{cases} (\lambda_{01} + \lambda_{02})p_0 = \lambda_{10}p_1 + \lambda_{20}p_2, \\ (\lambda_{10} + \lambda_{13})p_1 = \lambda_{01}p_0 + \lambda_{31}p_3, \\ (\lambda_{20} + \lambda_{23})p_2 = \lambda_{02}p_0 + \lambda_{32}p_3, \\ (\lambda_{31} + \lambda_{32})p_3 = \lambda_{13}p_1 + \lambda_{23}p_2. \end{cases} \quad (10)$$

Систему (10) можно составить непосредственно по размеченному графу

состояний, если руководствоваться правилом, согласно которому *слева в уравнениях стоит предельная вероятность данного состояния P_i , умноженная на суммарную интенсивность всех потоков, ведущих из данного состояния, а справа — сумма произведений интенсивностей всех потоков, входящих в i -е состояние, на вероятности тех состояний, из которых эти потоки исходят*

Пример 2. Найти предельные вероятности для системы S из примера 1, граф состояний которой приведен на рис. 1, при

$$\lambda_{01} = 1, \quad \lambda_{02} = 2, \quad \lambda_{10} = 2, \quad \lambda_{13} = 2, \quad \lambda_{20} = 3, \quad \lambda_{23} = 1, \quad \lambda_{31} = 3, \quad \lambda_{32} = 2.$$

Решение. Система алгебраических уравнений, описывающих стационарный режим для данной системы, имеет вид (10) или

$$\begin{cases} 3p_0 = 2p_1 + 3p_2, \\ 4p_1 = p_0 + 3p_3, \\ 4p_2 = 2p_0 + 2p_3, \\ p_0 + p_1 + p_2 + p_3 = 1. \end{cases} \quad (11)$$

(Здесь мы вместо одного "лишнего" уравнения системы (10) записали нормировочное условие (8)).

Решив систему (11), получим $p_0 = 0,4$, $p_1 = 0,2$, $p_2 = 0,27$, $p_3 = 0,13$, т.е. в предельном, стационарном режиме система S в среднем 40% времени будет находиться в состоянии S_0 (оба узла исправны), 20% — в состоянии S_1 (первый узел ремонтируется, второй работает), 27% — в состоянии S_2 (второй узел ремонтируется, первый работает) и 13% времени — в состоянии S_3 (оба узла ремонтируются).

Пример 3. Найти средний чистый доход от эксплуатации в стационарном режиме системы S в условиях примеров 1 и 2, если известно, что в единицу времени исправная работа первого и второго узлов приносит доход соответственно в 10 и 6 ден.ед., а их ремонт требует затрат соответственно в 4 и 2 ден.ед. Оценить экономическую эффективность имеющейся возможности уменьшения вдвое среднего времени ремонта каждого из двух узлов, если при этом придется вдвое увеличить затраты на ремонт каждого узла (в единицу времени).

Решение. Из примера 2 следует, что в среднем первый узел исправно работает долю времени, равную $p_0 + p_3 = 0,4 + 0,27 = 0,67$, а второй узел — $p_0 + p_1 = 0,4 + 0,2 = 0,6$. В то же время первый узел находится в ремонте в среднем долю времени, равную $p_1 + p_3 = 0,2 + 0,13 = 0,33$, а второй узел — $p_2 + p_3 = 0,27 + 0,13 = 0,4$. Поэтому средний чистый доход \bar{D} в единицу времени от эксплуатации системы, т.е. разность между доходами и затратами, равен

$$\bar{D} = 0,67 \cdot 10 + 0,6 \cdot 6 - 0,33 \cdot 4 - 0,4 \cdot 2 = 8,18 \text{ ден. ед.}$$

Уменьшение вдвое среднего времени ремонта каждого из узлов в соответствии с (6) будет означать увеличение вдвое интенсивностей потока "окончаний ремонтов" каждого узла, т.е. теперь $\lambda_{10} = 4, \lambda_{20} = 6, \lambda_{31} = 6, \lambda_{32} = 4$ и система линейных алгебраических уравнений (10), описывающая стационарный режим системы S , вместе с нормировочным условием (8) примет вид:

$$\begin{cases} 3p_0 = 4p_1 + 6p_2, \\ 6p_1 = p_0 + 6p_3, \\ 7p_2 = 2p_0 + 4p_3, \\ p_0 + p_1 + p_2 + p_3 = 1. \end{cases}$$

Решив систему, получим $p_0 = 0,6$, $p_1 = 0,15$, $p_2 = 0,2$, $p_3 = 0,05$.

Учитывая, что $p_0 + p_2 = 0,8$, $p_0 + p_1 = 0,75$, $p_1 + p_3 = 0,12$, $p_2 + p_3 = 0,25$, а затраты на ремонт первого и второго узла составляют теперь соответственно 8 и 4 ден.ед., вычислим средний чистый доход \overline{D}_1 в единицу времени:

$$\overline{D}_1 = 0,8 \cdot 10 + 0,75 \cdot 6 - 0,2 \cdot 8 - 0,25 \cdot 4 = 9,9 \text{ ден.ед.}$$

Так как \overline{D}_1 больше \overline{D} (примерно на 20%), то экономическая целесообразность ускорения ремонтов узлов очевидна.

Процесс гибели и размножения

В теории массового обслуживания широкое распространение имеет специальный класс случайных процессов — так называемый *процесс гибели и размножения*. Название этого процесса связано с рядом биологических задач, где он является математической моделью изменения численности биологических популяций.

Граф состояний процесса гибели и размножения имеет вид, показанный на рис. 4.

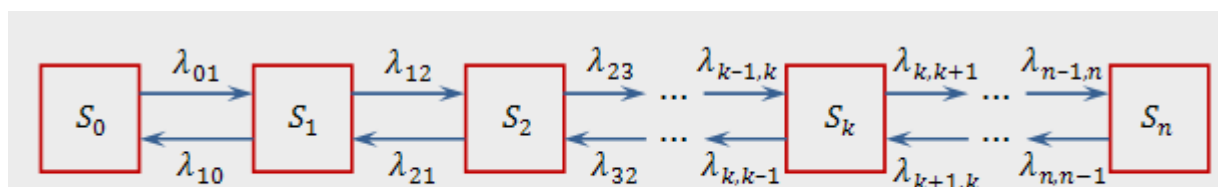


Рис. 4. Граф состояний процесса гибели и размножения

Рассмотрим упорядоченное множество состояний системы $S_0, S_1, S_2, \dots, S_k$. Переходы могут осуществляться из любого состояния только в состояния с соседними номерами, т.е. из состояния S_k возможны переходы только либо в состояние S_{k-1} , либо в состояние S_{k+1} .

Предположим, что все потоки событий, переводящие систему по стрелкам графа, простейшие с соответствующими интенсивностями $\lambda_{k,k+1}$ или $\lambda_{k+1,k}$.

По графу, представленному на рис. 4, составим и решим алгебраические уравнения для предельных вероятностей состояний (их существование вытекает из возможности перехода из каждого состояния в каждое другое и конечности числа состояний).

В соответствии с правилом составления таких уравнений (см. 13) получим:
для состояния S_0

$$\lambda_{01} \cdot p_0 = \lambda_{10} \cdot p_1,$$

12)

для состояния S_1 имеем $(\lambda_{12} + \lambda_{10})p_1 = \lambda_{01}p_0 + \lambda_{21}p_2$, которое с учетом (12) приводится к виду

$$\lambda_{12} \cdot p_1 = \lambda_{21} \cdot p_2.$$

13)

Аналогично, записывая уравнения для предельных вероятностей других состояний, можно получить следующую систему уравнений:

$$\begin{cases} \lambda_{01} \cdot p_0 = \lambda_{10} \cdot p_1, \\ \lambda_{12} \cdot p_1 = \lambda_{21} \cdot p_2, \\ \dots\dots\dots \\ \lambda_{k-1,k} \cdot p_{k-1} = \lambda_{k,k+1} \cdot p_k, \\ \dots\dots\dots \\ \lambda_{n-1,n} \cdot p_{n-1} = \lambda_{n,n+1} \cdot p_n. \end{cases} \quad (14)$$

к которой добавляется нормировочное условие

$$p_0 + p_1 + p_2 + \dots + p_n = 1.$$

15)

При анализе численности популяций считают, что состояние S_k соответствует численности популяции, равной k , и переход системы из состояния S_k в состояние S_{k+1} происходит при рождении одного члена популяции, а переход в состояние S_{k-1} — при гибели одного члена популяции.

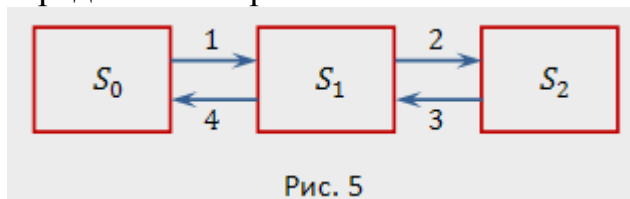
Решая систему (14), (15), можно получить

$$p_0 = \left(1 + \frac{\lambda_{01}}{\lambda_{10}} + \frac{\lambda_{12}\lambda_{01}}{\lambda_{21}\lambda_{10}} + \dots + \frac{\lambda_{n-1,n} \dots \lambda_{12}\lambda_{01}}{\lambda_{n,n-1} \dots \lambda_{21}\lambda_{10}} \right)^{-1}, \quad (16)$$

$$p_1 = \frac{\lambda_{01}}{\lambda_{10}} p_0, \quad p_2 = \frac{\lambda_{12}\lambda_{01}}{\lambda_{21}\lambda_{10}} p_0, \quad \dots, \quad p_n = \frac{\lambda_{n-1,n} \dots \lambda_{12}\lambda_{01}}{\lambda_{n,n-1} \dots \lambda_{21}\lambda_{10}} p_0. \quad (17)$$

Легко заметить, что в формулах (17) для p_1, p_2, \dots, p_n коэффициенты при p_0 есть слагаемые, стоящие после единицы в формуле (16). Числители этих коэффициентов представляют произведение всех интенсивностей, стоящих у стрелок, ведущих слева направо до данного состояния S_k ($k = 1, 2, \dots, n$), а знаменатели — произведение всех интенсивностей, стоящих у стрелок, ведущих справа налево до состояния S_k .

Пример 4. Процесс гибели и размножения представлен графом (рис. 5). Найти предельные вероятности состояний.



Решение. По формуле (16) найдем

$$p_0 = \left(1 + \frac{1}{4} + \frac{2 \cdot 1}{3 \cdot 4}\right)^{-1} = 0,706, \text{ по (17) } p_1 = \frac{1}{4} \cdot 0,706 = 0,176, \quad p_2 = \frac{2 \cdot 1}{3 \cdot 4} \cdot 0,706 = 0,118,$$

т.е. в установившемся, стационарном режиме в среднем 70,6% времени система будет находиться в состоянии S_0 , 17,6% — в состоянии S_1 и 11,8% — в состоянии S_2 .

ПЗ 3. Исследование многоканальной СМО с отказами (финальные вероятности)

Для СМО с отказами наиболее важными характеристиками являются :

1) **абсолютная пропускная способность A** – среднее число заявок, которое может обслужить система за единицу времени.

2) **относительная пропускная способность q** – вероятность того, что поступившее в систему требование будет принято к обслуживанию (дополняет вероятность отказа $P_{\text{отк}}$ до единицы).

3) в зависимости от задачи исследования могут быть важны и другие характеристики, например, среднее число занятых каналов \bar{N}_3 , среднее относительное время простоя системы в целом и отдельного канала и т. д.

Если требование поступает в систему в момент, когда все n каналов заняты, то оно получает отказ (покидает систему необслуженным). Если же в момент поступления требования имеется хотя бы один свободный канал, то оно принимается к обслуживанию и обслуживается до конца. Размеченный граф состояний системы с отказами представлен на рисунке.

Пусть СМО содержит n каналов, интенсивность входящего потока заявок равна λ , а интенсивность обслуживания заявки каждым каналом равна μ . Размеченный граф состояний системы изображён на рис. 1.

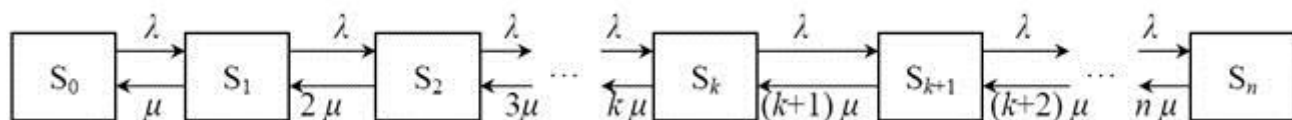


Рисунок 1 – Граф состояний многоканальной системы с отказами

Состояние S_0 означает, что все каналы свободны, состояние S_k ($k = \overline{1, n}$) означает, что обслуживанием заявок заняты k каналов, остальные свободны, S_n –

заняты все n каналов, заявка получает отказ в обслуживании.

Переход из одного состояния в другое соседнее правое происходит скачкообразно под воздействием входящего потока заявок интенсивностью λ независимо от числа работающих каналов (верхние стрелки). Для перехода системы из одного состояния в соседнее левое неважно, какой именно канал освободится. Величина $k\mu$ характеризует интенсивность обслуживания заявок при работе в СМО k каналов (нижние стрелки).

Предельные вероятности состояний системы рассчитываются по формулам Эрланга, которые для рассматриваемой СМО имеют вид:

$$\begin{cases} P_0 = \left[1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \dots + \frac{\rho^k}{k!} + \dots + \frac{\rho^n}{n!} \right]^{-1}, \\ P_k = \frac{\rho^k}{k!} P_0 \quad (1 \leq k \leq n) \end{cases} \quad (4)$$

где P_0 – вероятность свободного состояния системы, P_k – финальные вероятности состояния системы (вероятности того, что обслуживанием заявок заняты k каналов).

Сумма значений всех найденных по формулам Эрланга вероятностей должна быть равна 1.

Вероятность отказа в обслуживании $P_{\text{отк}}$ равна вероятности того, что все каналы заняты, т.е. система находится в состоянии S_n : $P_{\text{отк}} = P_n = \frac{\rho^n}{n!} P_0$.

Относительная пропускная способность – вероятность того, что заявка будет принята к обслуживанию (дополняет $P_{\text{отк}}$ до единицы): $q = 1 - P_{\text{отк}}$.

Абсолютная пропускная способность A – среднее число заявок, которое может обслужить СМО в единицу времени: $A = \lambda q = \lambda(1 - P_{\text{отк}})$.

$$\text{Среднее число занятых каналов} \quad \bar{N}_z = \frac{A}{\mu} = \frac{\lambda q}{\mu} = \rho q = \rho(1 - P_{\text{отк}}).$$

Пример 1. На вход многоканальной СМО с отказами поступает поток заявок, интенсивность которого составляет 7 заявок в час. Среднее время обслуживания одной заявки 0,25 часа. Каждая заявка приносит доход 150 ден.ед., а содержание одного канала обходится в 120 ден.ед в час. Найти оптимальное число каналов СМО.

Решение. Основные параметры системы:

- 1) число каналов n – нужно определить;
- 2) число мест в очереди $m = 1$;
- 3) среднее время обслуживания $t_{\text{обс}} = 0,25$ часа;
- 4) интенсивность потока заявок $\lambda = 7$ заявок в час;
- 5) интенсивность обслуживания $\mu = 1/t_{\text{обс}} = 1/0,25 = 4 \text{ ч}^{-1}$;
- 6) относительная нагрузка на систему $\rho = \lambda/\mu = 7/4 = 1,75$ заявок.

Из условия задачи также вытекает, что в случае, если СМО имеет n каналов,

то она приносит доход D , который можно определить по формуле $D = 150 \cdot A - 120 \cdot n$, где A – абсолютная пропускная способность СМО. Дальнейшие расчеты сведены в таблицу 12.1

Таблица 12.1 – Определение оптимального количества каналов

Число каналов	Вероятность свободного состояния системы P_0	Вероятность отказа $P_{\text{отк}} = \frac{\rho^n}{n!} P_0$	Абсолютная пропускная способность $A = \lambda(1 - P_{\text{отк}})$ заявок	Доход $D = 150 \cdot A - 120 \cdot n$, ден.ед.
$n=1$	$P_0 = \left[1 + \frac{\rho}{1!}\right]^{-1} = 0,364$	0,636	2,545	
$n=2$	$P_0 = \left[1 + \frac{\rho}{1!} + \frac{\rho^2}{2!}\right]^{-1} = 0,234$	0,204	5,569	
$n=3$	$P_0 = \left[1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \frac{\rho^3}{3!}\right]^{-1} = 0,193$	0,056	6,61	
$n=4$	$P_0 = \left[1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \frac{\rho^3}{3!} + \frac{\rho^4}{4!}\right]^{-1} = 0,18$	0,013	6,91	

Сравнивая доходы, поступающие от СМО в случаях $n=1, 2, 3, 4$, можно отметить, что при увеличении каналов от одного до трех доход растет и становится наибольшим при $n=3$. Это значение и является оптимальным количеством каналов.

Для многоканальной СМОс **неограниченным ожиданием** как абсолютная A , так и относительная пропускная способность q теряют смысл, так как каждая поступившая заявка рано или поздно будет обслужена. Зато для такой СМО весьма важными характеристиками являются:

- 1) среднее число заявок в очереди \bar{r} ,
- 2) среднее число заявок в системе \bar{k} (в очереди и под обслуживанием),
- 3) среднее время ожидания заявки в очереди $\bar{t}_{\text{ож}}$,
- 4) среднее время пребывания заявки в системе $\bar{t}_{\text{сист}}$ (в очереди и под обслуживанием), и другие характеристики ожидания.

Так как длина очереди не ограничена, то граф состояний (при $m \rightarrow \infty$) в этом случае является бесконечным.

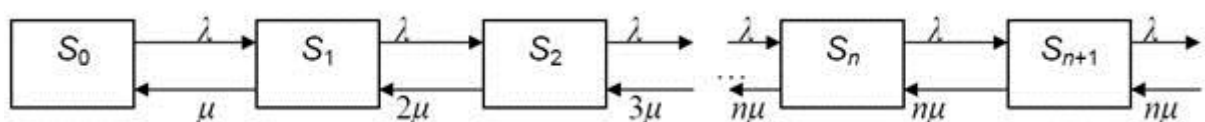


Рисунок 2 – Граф состояний многоканальной системы с неограниченным ожиданием

Далее представлены формулы для расчета характеристик СМО с неограниченной очередью.

Предельные вероятности состояний:

$$\begin{cases} P_0 = \left[1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \dots + \frac{\rho^n}{n!} + \frac{\rho^{n+1}}{n!(1-\rho)} \right]^{-1}, \\ P_k = \frac{\rho^k}{k!} P_0 \quad (1 \leq k \leq n), \\ P_{n+r} = \frac{\rho^{n+r}}{n! \cdot n!} P_0 \quad (r > 1) \end{cases} \quad (5)$$

где P_0 – вероятность свободного состояния системы, P_k – финальные вероятности состояния системы (вероятности того, что обслуживанием заявок заняты k каналов), P_{n+r} – вероятности того, что r заявок находится в очереди при занятом количестве каналов n .

Сумма значений всех найденных по формулам Эрланга вероятностей должна быть равна 1.

Использование формул (5) является корректным лишь в том случае, когда

выполняется **условие существования стационарного режима СМО**: $\frac{\rho}{n} < 1$. Смысл этого условия состоит в том, что суммарная интенсивность обслуживания, создаваемая всеми n одновременно работающими каналами, равная $n\mu$, должна быть строго больше интенсивности входного потока заявок λ .

Так как число мест в очереди не ограничено, то все требования, поступившие в систему, рано или поздно будут обслужены. Следовательно, вероятность отказа

$$P_{\text{отк}} = 0, \quad q = 1, \quad A = \lambda, \quad q = \lambda.$$

Среднее число занятых каналов

$$\bar{N}_s = \frac{A}{\mu} = \frac{\lambda}{\mu} = \rho$$

Среднее число требований, находящихся в очереди

$$\bar{r} = \frac{\rho^{n+1} \cdot P_0}{n! \cdot n! \cdot \left(1 - \frac{\rho}{n}\right)^2}.$$

Среднее число требований, находящихся в системе $\bar{k} = \bar{N}_s + \bar{r}$.

$$\bar{t}_{\text{ож}} = \frac{\rho^n \cdot P_0}{n! \cdot n! \cdot \left(1 - \frac{\rho}{n}\right)^2} = \frac{\bar{r}}{\lambda}.$$

Среднее время пребывания требования в очереди

Среднее время пребывания требования в системе $\bar{t}_{\text{сист}} = \bar{t}_{\text{ож}} + \bar{t}_{\text{обс}}$.

Пример 2. Железнодорожная касса с двумя окошками продает билеты в два пункта A и B . Интенсивность потока пассажиров, желающих купить билеты, для обоих пунктов одинакова: $\lambda_A = \lambda_B = 0,45$ пассажиров в минуту. На обслуживание пассажиров кассир тратит в среднем 2 мин. Рассматриваются два варианта продажи билетов: первый – билеты продаются в одной кассе с двумя окошками одновременно в оба пункта A и B ; второй – билеты продаются в двух специализированных кассах (по одному окошку в каждой), одна только в пункт A , другая – только в пункт B . Необходимо сравнить два варианта продажи билетов по основным характеристикам обслуживания.

ПЗ 4. Исследование многоканальной СМО с отказами (система показателей)

Под операцией в СМО понимают комплекс мероприятий по обслуживанию входящего потока заявок на интервале времени T .

В зависимости от типа системы показателями исхода операции или эффективности системы массового обслуживания являются следующие.

Для **СМО с отказами**:

- *абсолютная пропускная способность* (Q) - среднее число заявок, обслуживаемое системой за время T ;
- *относительная пропускная способность* (q) - средняя доля поступивших заявок, обслуживаемая системой (отношение среднего числа обслуженных заявок к среднему числу поступивших за время T);
- *среднее число занятых каналов* (\bar{n}_3);
- *коэффициент занятости (использования) каналов* ($K_{\text{и}} = \bar{n}_3/n$, где n - число каналов в системе);
- *коэффициент простоя каналов*, $K_{\text{п}} = 1 - K_{\text{и}}$.

Для **СМО с неограниченным ожиданием** как абсолютная, так и относительная пропускная способности теряют смысл, так как каждая поступившая заявка рано или поздно будет обслужена. Для такой СМО важными показателями являются:

- *среднее число заявок в очереди* ($\bar{l}_{\text{оч}}$);
- *среднее число заявок в системе* (в очереди и на обслуживании, \bar{l}_c);
- *среднее время ожидания заявки в очереди* ($\bar{t}_{\text{ож}}$);
- *среднее время пребывания заявки в системе* (в очереди и на обслуживании, \bar{t}_c);
- *коэффициенты использования и простоя каналов* ($K_{\text{п}}, K_{\text{и}}$);
- *среднее число свободных и занятых каналов* (\bar{n}_c, \bar{n}_3).

Для **СМО смешанного типа** используются обе группы показателей: как относительная и абсолютная пропускная способности, так и характеристики ожидания.

В зависимости от цели операции массового обслуживания любой из приведенных показателей (или совокупность показателей) может быть выбран в качестве критерия эффективности.

Аналитической моделью СМО является совокупность уравнений или формул, позволяющих определять вероятности состояний системы в процессе ее функционирования и рассчитывать показатели эффективности по известным характеристикам входящего потока и каналов обслуживания.

Всеобщей аналитической модели для произвольной СМО не существует. Аналитические модели разработаны для ограниченного числа частных случаев СМО. Аналитические модели, более или менее точно отображающие реальные системы, как правило, сложны и труднообозримы.

Аналитическое моделирование СМО существенно облегчается, если процессы, протекающие в СМО, марковские (потоки заявок простейшие, времена обслуживания распределены экспоненциально). В этом случае все процессы в СМО можно описать обыкновенными дифференциальными уравнениями, а в предельном случае, для стационарных состояний - линейными алгебраическими уравнениями и, решив их, определить выбранные показатели эффективности.

Рассмотрим примеры некоторых СМО.

2.5.1. Многоканальная СМО с отказами

Пример 2.5. Три автоинспектора проверяют путевые листы у водителей грузовых автомобилей. Если хотя бы один инспектор свободен, проезжающий грузовик останавливают. Если все инспекторы заняты, грузовик, не задерживаясь, проезжает мимо. Поток грузовиков простейший, время проверки случайное с экспоненциальным распределением.

Такую ситуацию можно моделировать трехканальной СМО с отказами (без очереди). Система разомкнутая, с однородными заявками, однофазная, с абсолютно надежными каналами.

Описание состояний:

S_0 - все инспекторы свободны;

S_1 - занят один инспектор;

S_2 - заняты два инспектора;

S_3 - заняты три инспектора.

Граф состояний системы приведен на [рис. 2.11](#).

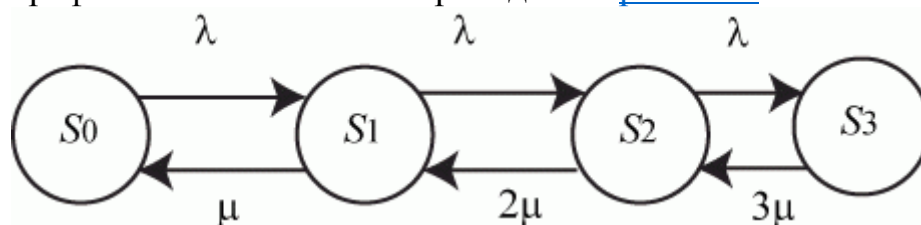


Рис. 2.11. Граф состояний трехканальной СМО с отказами

На графе: λ - интенсивность потока грузовых автомобилей; μ - интенсивность проверок документов одним автоинспектором.

Моделирование проводится с целью определения части автомобилей, которые не будут проверены.

Решение

Искомая часть вероятности P_3 - вероятности занятости всех трех инспекторов. Поскольку граф состояний представляет типовую схему "гибели и размножения", то найдем P_3 , используя зависимости (2.2).

$$P_0 = \frac{1}{\sum_{k=0}^3 \frac{\lambda^k}{k! \mu * 2\mu * 3\mu}}; P_3 = \frac{\lambda^3}{3! \mu * 2\mu * 3\mu} * P_0.$$

Пропускную способность этого поста автоинспекторов можно характеризовать относительной пропускной способностью:

$$q = 1 - P_3.$$

Пример 2.6. Для приема и обработки донесений от разведгруппы в разведотделе объединения назначена группа в составе трех офицеров. Ожидаемая интенсивность потока донесений - 15 донесений в час. Среднее время обработки одного донесения одним офицером - $t_{\text{обс}} = 12$ мин. Каждый офицер может принимать донесения от любой разведгруппы. Освободившийся офицер обрабатывает последнее из поступивших донесений. Поступающие донесения должны обрабатываться с вероятностью не менее 95 %.

Определить, достаточно ли назначенной группы из трех офицеров для выполнения поставленной задачи.

Решение

Группа офицеров работает как СМО с отказами, состоящая из трех каналов.

$$\lambda = 15 \frac{1}{\text{час}}$$

Поток донесений с интенсивностью можно считать простейшим, так как он суммарный от нескольких разведгрупп. Интенсивность

обслуживания $\mu = \frac{1}{t_{\text{обс}}} = \frac{60}{12 \text{ час}} = 5 \frac{1}{\text{час}}$. Закон распределения неизвестен, но это несущественно, так как показано, что для систем с отказами он может быть произвольным.

Описание состояний и граф состояний СМО будут аналогичны приведенным в примере 2.5.

Поскольку граф состояний - это схема "гибели и размножения", то для нее имеются готовые выражения для предельных вероятностей состояния:

$$P_1 = \frac{\rho}{1!} * P_0; P_2 = \frac{\rho^2}{2!} * P_0; \dots; P_n = \frac{\rho^n}{n!} * P_0; P_{n+1} = \frac{\rho^{n+1}}{nn!} * P_0; \dots; P_{n+m} = \frac{\rho^{n+m}}{n^m n!} * P_0;$$

$$P_0 = \left(1 + \frac{\rho}{1!} + \dots + \frac{\rho^n}{n!} + \frac{\rho^{n+1}}{nn!} + \frac{\rho^{n+2}}{n^2 n!} + \frac{\rho^{n+m}}{n^m n!} \right)^{-1}.$$

$$\rho = \frac{\lambda}{\mu}$$

Отношение называют *приведенной интенсивностью потока заявок*. Физический смысл ее следующий: величина ρ представляет собой среднее число заявок, приходящих в СМО за среднее время обслуживания одной заявки.

$$\rho = \frac{\lambda}{\mu} = \frac{15}{5} = 3$$

В примере

В рассматриваемой СМО отказ наступает при занятости всех трех каналов, то есть $P_{\text{отк}} = P_3$. Тогда:

$$P_0 = \left(1 + \frac{3}{1!} + \frac{3^2}{2!} + \frac{3^3}{3!}\right)^{-1} = 0,077; \quad P_3 = \frac{3}{3!} * P_0 = 4,5 * 0,077 = 0,346.$$

Так как вероятность отказа в обработке донесений составляет более 34 % (0,346), то необходимо увеличить личный состав группы. Увеличим состав группы в два раза, то есть СМО будет иметь теперь шесть каналов, и рассчитаем $P_{\text{отк}}$:

$$P_0 = \left(1 + \frac{3}{1!} + \frac{3^2}{2!} + \frac{3^3}{3!} + \frac{3^4}{4!} + \frac{3^5}{5!} + \frac{3^6}{6!}\right)^{-1} = \frac{1}{19,4125} = 0,051;$$

$$P_6 = \frac{3^6}{6!} * P_0 = \frac{729}{720} * \frac{1}{19,4125} = 1,012 * 0,051 = 0,052.$$

Теперь $P_{\text{обс}} = 1 - P_{\text{отк}} = 1 - 0,052 \approx 0,95$.

Таким образом, только группа из шести офицеров сможет обрабатывать поступающие донесения с вероятностью 95 %.

2.5.2. Многоканальная СМО с ожиданием

Пример 2.7. На участке форсирования реки имеются 15 однотипных переправочных средств. Поток поступления техники на переправу в среднем составляет 1 ед./мин, среднее время переправы одной единицы техники - 10 мин (с учетом возвращения назад переправочного средства).

Оценить основные характеристики переправы, в том числе вероятность в немедленной переправе сразу по прибытии единицы техники.

Решение

$\lambda = 1$ ед./мин, $\mu = 0,1$ ед./мин, $n = 15$.

$$P_0 = \left(1 + \frac{10}{1!} + \frac{10^2}{2!} + \dots + \frac{10^{15}}{15!}\right)^{-1} = 4,77 * 10^{-5};$$

$$P_{\text{отк}} = P_{15} = \frac{10^{15}}{15!} * P_0 = 0,0365;$$

$$P_{\text{обс}} = 1 - P_{\text{отк}} = 1 - 0,0365 \approx 0,9635.$$

Абсолютная пропускная способность $A = 1 - 0,9635 \approx 1$, т. е. все, что подходит к переправе, тут же практически переправляется.

Среднее число работающих переправочных средств:

$$\bar{n}_3 = \frac{A}{\mu} = \frac{1}{1/10} = 10.$$

Коэффициенты использования и простоя переправы:

$$K_{\text{и}} = \frac{\bar{n}_3}{n} = \frac{10}{15} = 0,666; K_{\text{п}} = 1 - K_{\text{и}} = 1 - 0,666 = 0,334.$$

Для решения примера была также разработана программа. Интервалы времени поступления техники на переправу, время переправы приняты распределенными по экспоненциальному закону.

Коэффициенты использования переправы после 50 прогонов практически совпадают: $K_{\text{и}} = 0,665$.

Максимальная длина очереди 15 ед., среднее время пребывания в очереди около 10 мин.

Если взять число переправочных средств 10, то коэффициент использования близок к 1 ($K_{\text{и}} = 0,997$), максимальная длина очереди - 43 единицы техники.

2.5.3. Одноканальная СМО с ограниченной очередью

Если в очереди n мест для ожидания, то система может находиться в одном из следующих $n + 2$ состояний:

S_0 - в системе нет заявок (ни в очереди, ни на обслуживании);

S_1 - в системе обслуживается одна заявка, очередь пуста;

S_2 - в системе обслуживается одна заявка, и одна заявка находится в очереди, ожидает обслуживания;

...

S_{n+1} - в системе обслуживается одна заявка и n заявок находятся в очереди, ожидают обслуживания.

Граф состояний такой системы представляет схему "гибели и размножения" (рис. 2.12).

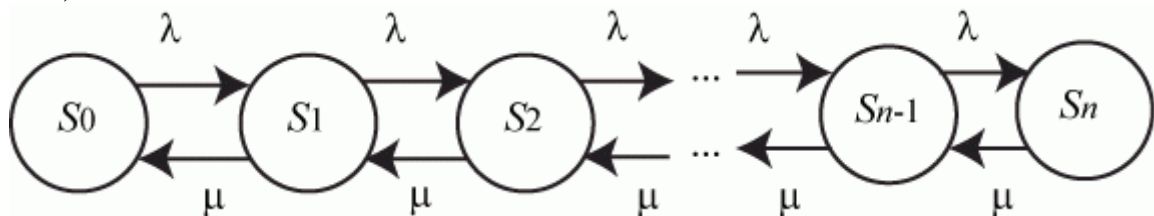


Рис. 2.12. Граф состояний одноканальной СМО с ограниченной очередью

2.5.4. Одноканальная замкнутая СМО

Опишем состояния одноканальной замкнутой СМО.

S_0 - заявок на обслуживание нет.

$S_k, k = \overline{1, N}$ - на обслуживании находится k заявок;

N - общее число заявок, циркулирующих в системе;

β - интенсивность требований на обслуживание от одной заявки.

Граф состояний одноканальной замкнутой СМО приведен на рис. 2.13. Модель данной СМО также представляет "схему гибели и размножения".

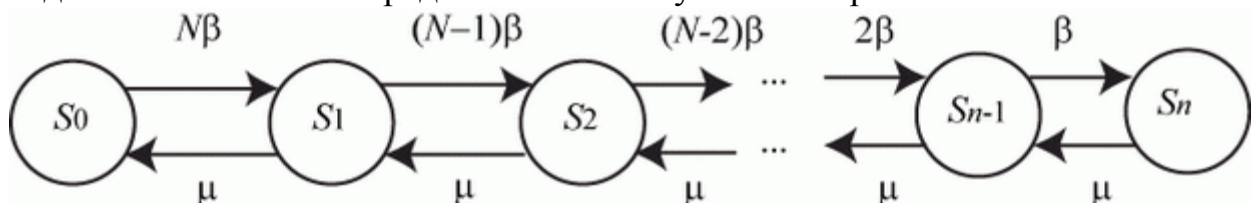


Рис. 2.13. Граф состояний одноканальной замкнутой СМО

Однако не менее часто модель СМО не сводится к схеме "гибели и размножения". Например, в СМО с конечной надежностью каналов обслуживания.

ПЗ 5. Исследование одноканальной СМО с неограниченной очередью (финальные вероятности)

ЗАДАНИЕ. Система массового обслуживания — билетная касса с одним окошком и неограниченной очередью. Касса продает билеты в пункты А и В. Пассажиры, желающих купить билет в пункт А, приходит в среднем трое за 20 мин, в пункт В — двое за 20 мин. Поток пассажиров простейший. Кассир в среднем обслуживает трех пассажиров за 10 мин. Время обслуживания — показательное. Вычислить финальные вероятности P_0, P_2, P_3 , среднее число заявок в системе и в очереди, среднее время пребывания заявки в системе, среднее время пребывания заявки в очереди.

РЕШЕНИЕ. Имеем систему массового обслуживания с одним каналом (одна касса) и неограниченной очередью. Интенсивность потока входящих заявок равна $(2+3=5$ пассажиров за 20 минут) $= (15$ пассажиров в час), то есть $I = 15$. Интенсивность потока обслуживания равна $(3$ пассажира за 10 минут) $= (18$ пассажиров за час), то есть $m = 18$.

Нагрузка системы $\rho = \lambda/\mu = 15/18 = 5/6$ нагрузка системы на один канал такая же:

$\psi = \rho = 5/6 < 1$, поэтому предельный режим работы системы существует. Рассчитаем эффективность работы СМО в предельном режиме.

Вычислим финальные вероятности:

Вероятность простоя системы: $p_0 = \psi^0(1 - \psi) = 1/6$

Вероятность того, что в системе одна заявка (один пассажир у кассы):

$$p_1 = \psi^1(1 - \psi) = 5/36$$

Вероятность того, что в системе две заявки (один пассажир у кассы и один пассажир в очереди):

$$p_2 = \psi^2(1 - \psi) = 25/216$$

Вероятность того, что в системе три заявки (один пассажир у кассы и два пассажира в очереди):

$$p_3 = \psi^3(1 - \psi) = 125/1296$$

Среднее число заявок, находящихся в очереди (пассажиров в очереди) равно

$$N_{line} = \frac{\psi^2}{(1-\psi)} = \frac{25/36}{1/6} = \frac{25}{216} \approx 0,116 \text{ человек.}$$

Среднее время ожидания в очереди равно

$$T_{line} = \frac{N_{line}}{\lambda} = \frac{0,116}{15} \approx 0,0077 \text{ часа} \approx 0,5 \text{ минут.}$$

Среднее число пассажиров, покупающих билеты, равно $N_s = \rho = \psi = 5/6 \approx 0,833$

Среднее время обслуживания равно $T_s = \frac{N_s}{\lambda} = \frac{0,833}{15} \approx 0,0555 \text{ часа} \approx 3,3 \text{ минуты}$

Тогда среднее число заявок в системе

$$N_{sys} = N_s + N_{line} = 0,116 + 0,833 = 0,949 \text{ (пассажиров).}$$

Среднее время пребывания заявки в системе $T_{sys} = T_s + T_{line} = 3,3 + 0,5 = 3,8$

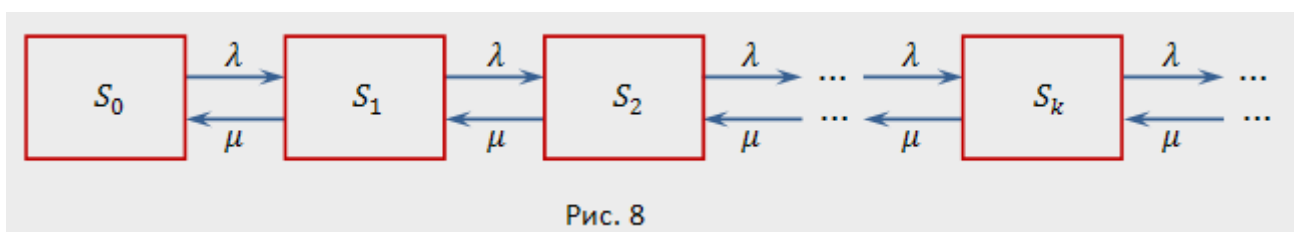
ПЗ 6. Исследование одноканальной СМО с неограниченной очередью (система показателей)

На практике часто встречаются одноканальные СМО с неограниченной очередью (например, телефон-автомат с одной будкой). Рассмотрим задачу.

Имеется одноканальная СМО с очередью, на которую не наложены никакие ограничения (ни по длине очереди, ни по времени ожидания). Поток заявок, поступающих в СМО, имеет интенсивность λ , а поток обслуживания — интенсивность μ . Необходимо найти предельные вероятности состояний и показатели эффективности СМО.

Система может находиться в одном из состояний $S_0, S_1, S_2, \dots, S_k$, по числу заявок, находящихся в СМО: S_0 — канал свободен; S_1 — канал занят (обслуживает заявку), очереди нет; S_2 — канал занят, одна заявка стоит в очереди; $\dots S_k$ — канал занят, $(k-1)$ заявок стоят в очереди и т.д.

Граф состояний СМО представлен на рис. 8.



Это процесс гибели и размножения, но с бесконечным числом состояний, в котором интенсивность потока заявок равна λ , а интенсивность потока обслуживания μ .

Прежде чем записать формулы предельных вероятностей, необходимо быть уверенным в их существовании, ведь в случае, когда время $t \rightarrow \infty$, очередь может неограниченно возрастать. Доказано, что если $\rho < 1$, т.е. среднее число приходящих заявок меньше среднего числа обслуженных заявок (в единицу времени), то предельные вероятности существуют. Если $\rho \geq 1$, очередь растет до бесконечности.

Для определения предельных вероятностей состояний воспользуемся формулами (16), (17) для процесса гибели и размножения (здесь мы допускаем известную нестрогость, так как ранее эти формулы были получены для случая конечного числа состояний системы). Получим:

$$p_0 = \left[1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu} \right)^2 + \dots + \left(\frac{\lambda}{\mu} \right)^k + \dots \right]^{-1} = (1 + \rho + \rho^2 + \dots + \rho^k + \dots)^{-1}.$$

32)

Так как предельные вероятности существуют лишь при $\rho < 1$, то геометрический ряд со знаменателем $\rho < 1$, записанный в скобках в формуле (32), сходится к сумме, равной $\frac{1}{1 - \rho}$. Поэтому

$$p_0 = 1 - \rho,$$

33}

и с учетом соотношений (17)

$$p_1 = \rho \cdot p_0, \quad p_2 = \rho^2 \cdot p_0, \quad \dots, \quad p_k = \rho^k \cdot p_0, \quad \dots$$

найдем предельные вероятности других состояний

$$p_1 = \rho(1 - \rho), \quad p_2 = \rho^2(1 - \rho), \quad \dots, \quad p_k = \rho^k(1 - \rho), \quad \dots$$

34)

Предельные вероятности $p_0, p_1, p_2, \dots, p_k, \dots$ образуют убывающую геометрическую прогрессию со знаменателем $\rho < 1$, следовательно, вероятность p_0 — наибольшая. Это означает, что если СМО справляется с потоком заявок (при $\rho < 1$), то наиболее вероятным будет отсутствие заявок в системе.

Среднее число заявок в системе $L_{\text{сист.}}$ определим по формуле математического ожидания, которая с учетом (34) примет вид

$$L_{\text{сист.}} = \sum_{k=1}^{\infty} k p_k = (1 - \rho) \sum_{k=1}^{\infty} k \rho^k$$

35)

(суммирование от 1 до ∞ , так как нулевой член $0 \cdot p_0 = 0$).
Можно показать, что формула (35) преобразуется (при $\rho < 1$) к виду

$$L_{\text{сист.}} = \frac{\rho}{1 - \rho}.$$

36)

Найдем среднее число заявок в очереди $L_{\text{оч.}}$. Очевидно, что

$$L_{\text{оч.}} = L_{\text{сист.}} - L_{\text{об.}},$$

37)

где $L_{\text{об.}}$ — среднее число заявок, находящихся под обслуживанием.

Среднее число заявок под обслуживанием определим по формуле математического ожидания числа заявок под обслуживанием, принимающего значения 0 (если канал свободен) либо 1 (если канал занят):

$$L_{\text{об.}} = 0 \cdot p_0 + 1 \cdot (1 - p_0),$$

т.е. среднее число заявок под обслуживанием равно вероятности того, что канал занят:

$$L_{\text{об.}} = P_{\text{зан.}} = 1 - p_0.$$

38)

В силу (33)

$$L_{\text{об.}} = P_{\text{зан.}} = \rho.$$

39)

Теперь по формуле (37) с учетом (36) и (39)

$$L_{\text{оч.}} = \frac{\rho^2}{1 - \rho}.$$

40)

Доказано, что при любом характере потока заявок, при любом распределении времени обслуживания, при любой дисциплине обслуживания среднее время пребывания заявки в системе (очереди) равно среднему числу заявок в системе (в очереди), деленному на интенсивность потока заявок, т.е.

$$T_{\text{сист.}} = \frac{1}{\lambda} \cdot L_{\text{сист.}},$$

41)

$$T_{\text{оч.}} = \frac{1}{\lambda} \cdot L_{\text{оч.}}.$$

42)

формулы (41) и (42) называются формулами Литтла. Они вытекают из того, что в предельном, стационарном режиме среднее число заявок, прибывающих в систему, равно среднему числу заявок, покидающих ее: оба потока заявок имеют одну и ту же интенсивность λ .

На основании формул (41) и (42) с учетом (36) и (40) среднее время пребывания заявки в системе определится по формуле:

$$T_{\text{сист.}} = \frac{\rho}{\lambda(1 - \rho)},$$

43)

а среднее время пребывания заявки в очереди —

$$T_{\text{оч.}} = \frac{\rho^2}{\lambda(1 - \rho)}.$$

44)

Пример 8. В порту имеется один причал для разгрузки судов. Интенсивность потока судов равна 0,4 (судов в сутки). Среднее время разгрузки одного судна составляет 2 суток. Предполагается, что очередь может быть неограниченной длины. Найти показатели эффективности работы причала, а также вероятность того, что ожидают разгрузки не более чем 2 судна.

Решение. Имеем $\rho = \frac{\lambda}{\mu} = \lambda \bar{t}_{\text{об.}} = 0,4 \cdot 2 = 0,8$. Так как $\rho = 0,8 < 1$, то очередь на разгрузку не может бесконечно возрастать и предельные вероятности существуют. Найдем их.

Вероятность того, что причал свободен, по (33) $p_0 = 1 - 0,8 = 0,2$, а вероятность того, что он занят, $P_{\text{зан.}} = 1 - 0,2 = 0,8$. По формуле (34) вероятности того, что у причала находятся 1, 2, 3 судна (т.е. ожидают разгрузки 0, 1, 2 судна), равны

$$p_1 = 0,8(1 - 0,8) = 0,16; \quad p_2 = 0,8^2(1 - 0,8) = 0,128; \quad p_3 = 0,8^3(1 - 0,8) = 0,1024.$$

Вероятность того, что ожидают разгрузку не более чем 2 судна, равна

$$P = p_1 + p_2 + p_3 = 0,16 + 0,128 + 0,1024 = 0,3904.$$

По формуле (40) среднее число судов, ожидающих разгрузки, $L_{\text{och.}} = \frac{0,8^2}{1 - 0,8} = 3,2$ среднее время ожидания разгрузки по формуле (42) $T_{\text{och.}} = \frac{3,2}{0,8} = 4$ (сутки).

По формуле (36) среднее число судов, находящихся у причала, $L_{\text{sist.}} = \frac{0,8}{1 - 0,8} = 4$ (сутки) (или проще по (37) $L_{\text{sist.}} = 3,2 + 0,8 = 4$ (сутки), а среднее время пребывания судна у причала по формуле (41) $T_{\text{sist.}} = \frac{4}{0,8} = 5$ (сутки).

Очевидно, что эффективность разгрузки судов невысокая. Для ее повышения необходимо уменьшение среднего времени разгрузки судна $\bar{t}_{\text{об.}}$ либо увеличение числа n причалов.

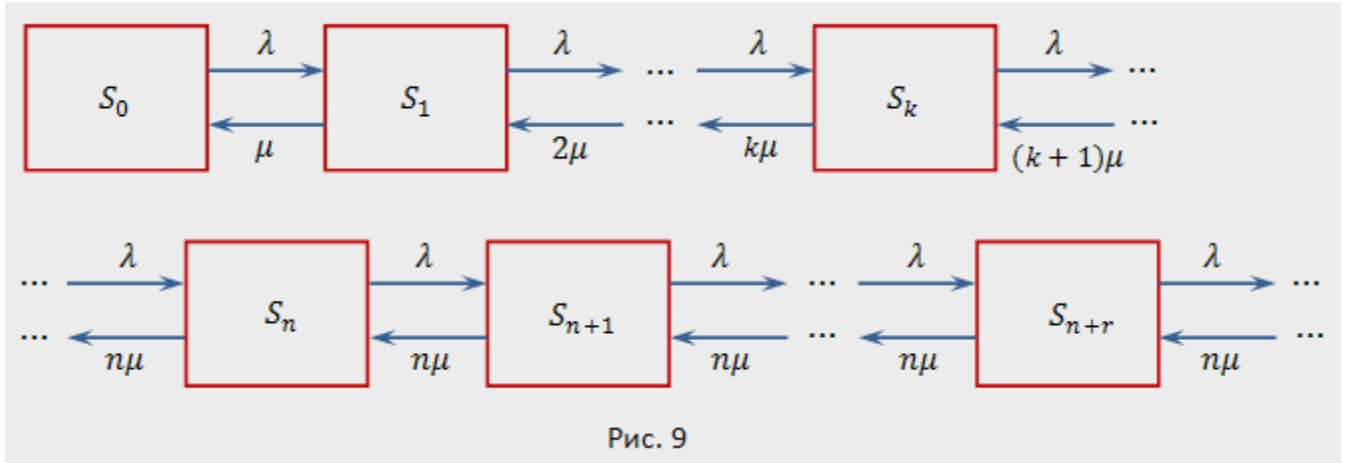
ПЗ 7. Исследование многоканальной СМО с неограниченной очередью (финальные вероятности)

Рассмотрим задачу. Имеется n -канальная СМО с неограниченной очередью. Поток заявок, поступающих в СМО, имеет интенсивность λ , а поток обслуживания — интенсивность μ . Необходимо найти предельные вероятности состояний СМО и показатели ее эффективности.

Система может находиться в одном из состояний $S_0, S_1, S_2, \dots, S_k, \dots, S_n, \dots$ нумеруемых по числу заявок, находящихся в СМО: S_0 — в системе нет заявок (все каналы свободны); S_1 — занят один канал, остальные свободны; S_2 — заняты два канала, остальные свободны; $\dots S_k$ — занято k каналов, остальные свободны; $\dots S_n$ — заняты все n каналов (очереди нет); S_{n+1} — заняты все n каналов, в очереди одна заявка; $\dots S_{n+r}$ — заняты все n каналов, r заявок стоит в очереди, и т.д.

Граф состояний системы показан на рис. 9. Обратим внимание на то, что в отличие от предыдущей СМО, интенсивность потока обслуживания

(переводящего систему из одного состояния в другое справа налево) не остается постоянной, а по мере увеличения числа заявок в СМО от 0 до n увеличивается от величины μ до $n\mu$, так как соответственно увеличивается число каналов обслуживания. При числе заявок в СМО большем, чем n , интенсивность потока обслуживания сохраняется равной $n\mu$.



Можно показать, что при $\frac{\rho}{n} < 1$ предельные вероятности существуют. Если $\frac{\rho}{n} \geq 1$, очередь растет до бесконечности. Используя формулы (16) и (17) для процесса гибели и размножения, можно получить следующие формулы для предельных вероятностей состояний n -канальной СМО с неограниченной очередью

$$p_0 = \left(1 + \frac{\rho^1}{1!} + \frac{\rho^2}{2!} + \dots + \frac{\rho^n}{n!} + \frac{\rho^{n+1}}{n!(n-\rho)} \right)^{-1},$$

45)

$$p_1 = \frac{\rho^1}{1!} \cdot p_0, \dots, p_k = \frac{\rho^k}{k!} \cdot p_0, \dots, p_n = \frac{\rho^n}{n!} \cdot p_0,$$

46)

$$p_{n+1} = \frac{\rho^{n+1}}{n \cdot n!} \cdot p_0, \dots, p_{n+r} = \frac{\rho^{n+r}}{n^r \cdot n!} \cdot p_0, \dots$$

47)

Вероятность того, что заявка окажется в очереди,

$$P_{\text{оч.}} = \frac{\rho^{n+1}}{n!(n-\rho)} \cdot p_0.$$

48)

Для n -канальной СМО с неограниченной очередью, используя прежние приемы, можно найти:

среднее число занятых каналов

$$\bar{k} = \frac{\lambda}{\mu} = \rho,$$

49)

среднее число заявок в очереди

$$L_{\text{оч.}} = \frac{\rho^{n+1} \cdot p_0}{n \cdot n!} \cdot \left(1 - \frac{\rho}{n}\right)^{-2},$$

50)

среднее число заявок в системе

$$L_{\text{сист}} = L_{\text{оч.}} + \rho.$$

1

Среднее время пребывания заявки в очереди и среднее время пребывания заявки в системе, как и ранее, находятся по формулам Литтла (42) и (41).

Замечание. Для СМО с неограниченной очередью при $\rho < 1$ любая заявка, пришедшая в систему, будет обслужена, т.е. вероятность отказа $P_{\text{отк.}} = 0$, относительная пропускная способность $Q = 1$, а абсолютная пропускная способность равна интенсивности входящего потока заявок, т.е. $A = \lambda$.

Пример 9. В универсаме к узлу расчета поступает поток покупателей с интенсивностью $\lambda = 81$ чел. в час. Средняя продолжительность обслуживания контролером-кассиром одного покупателя $\bar{t}_{\text{об.}} = 2$ мин. Определить:

а. Минимальное количество контролеров-кассиров n_{min} , при котором очередь не будет расти до бесконечности, и соответствующие характеристики обслуживания при $n = n_{\text{min}}$.

б. Оптимальное количество $n_{\text{opt.}}$ контролеров-кассиров, при котором относительная величина затрат $C_{\text{отн.}}$, связанная с издержками на содержание каналов обслуживания и с пребыванием в очереди покупателей, задаваемая, например, как $C_{\text{отн.}} = \frac{n}{\lambda} + 3T_{\text{оч.}}$, будет минимальна, и сравнить характеристики обслуживания при $n = n_{\text{min}}$ и $n = n_{\text{opt.}}$.

в. Вероятность того, что в очереди будет не более трех покупателей.

Решение. а. По условию $\lambda = 81$ (1/ч) $= \frac{81}{60} = 1,35$ (1/мин.). По формуле

(24) $\rho = \frac{\lambda}{\mu} = \lambda \bar{t}_{\text{об.}} = 1,35 \cdot 2 = 2,7$. Очередь не будет возрастать до бесконечности при условии $\frac{\rho}{n} < 1$, т.е. при $n > \rho = 2,7$. Таким образом, минимальное количество контролеров-кассиров $n_{\min} = 3$.

Найдем характеристики обслуживания СМО при $n = 3$.

(45) Вероятность того, что в узле расчета отсутствуют покупатели, по формуле

$$p_0 = \left(1 + 2,7 + \frac{2,7^2}{2!} + \frac{2,7^3}{3!} + \frac{2,7^4}{3!(3-2,7)} \right)^{-1} = 0,025$$

т.е. в среднем 2,5% времени контролеры-кассиры будут простаивать. Вероятность того, что в узле расчета будет очередь, по (48)

$$P_{\text{оч.}} = \frac{2,7^4}{3! \cdot (3 - 2,7)} \cdot 0,025 = 0,735.$$

Среднее число покупателей, находящихся в очереди, по (50)

$$L_{\text{оч.}} = \frac{2,7^4}{3 \cdot 3! \cdot (1 - 2,7/3)^2} \cdot 0,025 = 7,35.$$

Среднее время ожидания в очереди по (42)

$$T_{\text{оч.}} = \frac{7,35}{1,35} = 5,44 \text{ (мин).}$$

Среднее число покупателей в узле расчета по (51)

$$L_{\text{сист.}} = 7,35 + 2,7 = 10,05.$$

Среднее время нахождения покупателей в узле расчета по (41)

$$T_{\text{сист.}} = \frac{10,05}{1,35} \approx 7,44 \text{ (мин).}$$

Среднее число контролеров-кассиров, занятых обслуживанием покупателей, по (49) $\bar{k} = 2,7$.

Коэффициент (доля) занятых обслуживанием контролеров-кассиров $\bar{k}_{\text{отн.}} = \frac{\rho}{n} = \frac{2,7}{3} = 0,9$.

Абсолютная пропускная способность узла расчета $A = 1,35$ (1/мин), или 81 (1/ч), т.е. 81 покупатель в час.

Анализ характеристик обслуживания свидетельствует о значительной перегрузке узла расчета при наличии трех контролеров-кассиров.

б. Относительная величина затрат при $n = 3$

$$C_{\text{отн.}} = \frac{n}{\lambda} + 3T_{\text{оч.}} = \frac{3}{1,35} + 3 \cdot 5,44 = 5,44.$$

2). Рассчитаем относительную величину затрат при других значениях n (табл. 2).

Таблица 2					
Характеристика обслуживания	Число контролеров-кассиров				
	3	4	5	6	7
Вероятность простоя контролеров-кассиров p_0	0,025	0,057	0,065	0,067	0,067
Среднее число покупателей в очереди $T_{och.}$	5,44	0,60	0,15	0,03	0,01
Относительная величина затрат $C_{отн.}$	18,54	4,77	4,14	4,53	5,22

Как видно из табл. 2, минимальные затраты получены при $n = n_{opt.} = 5$ контролерах-кассирах.

Определим характеристики обслуживания узла расчета при $n = n_{opt.} = 5$. Получим

$$P_{och.} = 0,091; L_{och.} = 0,198; T_{och.} = 0,146; L_{sist.} = 2,9; T_{sist.} = 2,15; \bar{k} = 2,7; k_3 =$$

Как видим, при $n = 5$ по сравнению с $n = 3$ существенно уменьшились вероятность возникновения очереди $P_{och.}$, длина очереди $L_{och.}$ и среднее время пребывания в очереди $T_{och.}$, и соответственно среднее число покупателей $L_{sist.}$ и среднее время нахождения в узле расчета $T_{sist.}$, а также доля занятых обслуживанием контролеров k_3 . Но среднее число занятых обслуживанием контролеров-кассиров \bar{k} и абсолютная пропускная способность узла расчета A естественно не изменились.

в. Вероятность того, что в очереди будет не более 3 покупателей, определится как

$$P\{r \leq 3\} = p_1 + p_2 + p_3 + p_4 + p_5 + p_{5+1} + p_{5+2} + p_{5+3} = 1 - P_{och.} + p_{5+1} + p_{5+2} +$$

где каждое слагаемое найдем по формулам (45)–(48). Получим при $n = 5$:

$$P\{r \leq 3\} = 1 - \frac{2,7^6}{5!(5-2,3)} \cdot 0,065 + \frac{2,7^6}{5 \cdot 5!} \cdot 0,065 + \frac{2,7^7}{5^2 \cdot 5!} \cdot 0,065 + \frac{2,7^8}{5^3 \cdot 5!} \cdot 0,065 = 0,986.$$

(Заметим, что в случае $n = 3$ контролеров-кассиров та же вероятность существенно меньше: $P\{r \leq 3\} = 0,464$).

ПЗ 8. Исследование многоканальной СМО с неограниченной очередью (система показателей)

Пример 10. Железнодорожная касса с двумя окошками продает билеты в

два пункта A и B . Интенсивность потока пассажиров, желающих купить билеты, для обоих пунктов одинакова: $\lambda_A = \lambda_B = 0,45$ (пассажиров в минуту). На обслуживание пассажиров кассир тратит в среднем 2 мин. Рассматриваются два варианта продажи билетов: первый — билеты продаются в одной кассе с двумя окошками одновременно в оба пункта A и B , второй — билеты продаются в двух специализированных кассах (по одному окошку в каждой), одна только в пункт A , другая — только в пункт B . Необходимо:

а. Сравнить два варианта продажи билетов по основным характеристикам обслуживания.

б. Определить, как надо изменить среднее время обслуживания одного пассажира, чтобы по второму варианту продажи пассажиры затрачивали на приобретение билетов в среднем меньше времени, чем по первому варианту.

Решение.

а. По первому варианту имеем двухканальную СМО, на которую поступает поток заявок интенсивностью $\lambda = 0,45 + 0,45 = 0,9$; интенсивность потока обслуживания $\mu = \frac{1}{2} = 0,5$; $\rho = \frac{\lambda}{\mu} = 1,8$. Так как $\frac{\rho}{n} = \frac{1,8}{2} = 0,9 < 1$, то предельные вероятности существуют.

Вероятность простоя двух кассиров по (45)

$$p_0 = \left(1 + \frac{1,8}{1!} + \frac{1,8^2}{2!} + \frac{1,8^3}{2! \cdot (2 - 1,8)} \right)^{-1} \approx 0,0526.$$

Среднее число пассажиров в очереди по (50)

$$L_{\text{оч.}} = \frac{1,8^3}{2 \cdot 2! \cdot (1 - 1,8/2)^2} \cdot 0,0526 = 7,67.$$

Среднее число пассажиров у кассы по (51)

$$L_{\text{сист.}} = 7,67 + 1,8 = 9,47.$$

Среднее время на ожидание в очереди и покупку билетов равно соответственно (по формулам (42) и (41)):

$$T_{\text{оч.}} = \frac{7,67}{0,9} = 8,52 \text{ (мин)} \text{ и } T_{\text{сист.}} = \frac{9,47}{0,9} = 10,5 \text{ (мин)}.$$

По второму варианту имеем две одноканальные СМО (два специализированных окошка); на каждую поступает поток заявок с интенсивностью $\lambda = 0,45$. По-прежнему $\mu = 0,5$; $\rho = \frac{\lambda}{\mu} = 0,9 < 1$, предельные вероятности существуют. По формулам (40), (36), (42), (41)

$$L_{\text{оч.}} = \frac{0,9^2}{1 - 0,9} = 8,1; \quad L_{\text{сист.}} = \frac{0,9}{1 - 0,9} = 9; \quad T_{\text{оч.}} = \frac{8,1}{0,45} = 18; \quad T_{\text{сист.}} = \frac{9}{0,45} = 20.$$

Итак, по второму варианту увеличились и длина очереди, и среднее время ожидания в ней и в целом на покупку билетов. Такое различие объясняется тем, что в первом варианте (двухканальная СМО) меньше средняя доля времени, которую простаивает каждый из двух кассиров: если он не занят обслуживанием пассажира, покупающего билет в пункт A , он может заняться обслуживанием пассажира,

покупающего билет в пункт B , и наоборот. Во втором варианте такой взаимозаменяемости нет.

Можно заметить, что среднее время на покупку билетов по второму варианту увеличилось более чем в 2 раза. Такое значительное увеличение связано с тем, что СМО работает на пределе своих возможностей ($\rho = 0,9$): достаточно незначительно увеличить среднее время обслуживания $\bar{t}_{ob.}$, т.е. уменьшить μ , и ρ превзойдет 1, т.е. очередь начнет неограниченно возрастать.

б. Выше было получено, что по первому варианту продажи билетов при среднем времени обслуживания одного пассажира $\bar{t}_{ob.} = 2$ (мин) среднее время на покупку билетов составит $T_{sist.1} = 10,5$ (мин). По условию для второго варианта

продажи $T_{sist.2} < T_{sist.1}$, или с учетом (36) и (41): $\frac{1}{\lambda} \frac{\rho}{1 - \rho} < T_{sist.1}$.

Полагая $\rho = \frac{\lambda}{\mu} = \lambda \bar{t}_{ob.}$, получим $-\frac{\bar{t}_{ob.}}{1 - \lambda \bar{t}_{ob.}} < T_{sist.1}$, откуда найдем $\bar{t}_{ob.} < \frac{T_{sist.1}}{1 + \lambda T_{sist.1}}$ или $\bar{t}_{ob.} < \frac{10,5}{1 + 0,45 \cdot 10,5} = 1,83$ (мин).

Итак, средние затраты времени на покупку билетов по второму варианту продажи уменьшатся, если среднее время обслуживания одного пассажира уменьшится более чем на 0,17 мин, или более чем на 8,5%

ПЗ 9. Исследование методов «Алоха»

Требуется разработать программу, моделирующую работу методов случайного доступа к каналу передачи данных в локальных вычислительных сетях: чистая и синхронная Алоха. Обработать наборы из 100, 1000 и 10000 кадров.

Исходными данными являются:

величины нормированной пропущенной нагрузки 0,075; 0,25 и 1,6.

величины временных интервалов 2, 5, 16.

Выходными данными являются:

Время, необходимое для передачи всего набора кадров;

среднее время, необходимое для передачи кадра;

нормированная производительность протокола передачи;

количество коллизий.

Двумя основными способами доступа к общей среде передачи являются управляемый доступ с применением опроса и случайный доступ. В свою очередь существуют различные типы стратегий случайного доступа.

Методы случайного доступа полностью децентрализованы. Пользователь может передавать когда угодно, лишь с незначительными ограничениями, зависящими от метода доступа.

Из-за случайности моментов времени, в которые пользователи могут решить начать передачу, независимо от метода не исключена возможность того, что два или несколько пользователей могут выйти на связь в пересекающиеся промежутки времени. Это приводит к столкновениям (коллизиям), которые сначала должны быть распознаны, а затем разрешены. При увеличении нагрузки увеличивается и вероятность коллизий, что приводит к возможной неустойчивости работы

рассматриваемых механизмов.

В результате производительность ограничивается некоторым максимальным значением, меньшим пропускной способности канала, и это значение в каждом случае зависит от первоначального механизма доступа и алгоритма разрешения коллизий.

Сначала методы случайного доступа были предложены для случаев, когда большое число пользователей пытаются довольно редко передать пачки сообщений или, когда друг с другом связываются небольшое число ЭВМ. Но применимо к производственным процессам, которые требуют строгого управления задержкой доступа, более предпочтителен управляемый доступ. Рассмотрим два простейших типа стратегии случайного доступа: чистую Алоху и синхронную Алоху.

ЧИСТАЯ АЛОХА

Эта схема сначала была применена для доступа к общему каналу сотрудниками Гавайского университета в начале 1970-х годов. По этой схеме пользователь, желающий передать сообщение, делает это когда угодно. В результате могут накладываться во времени два или несколько сообщений, вызвав столкновение (коллизию).

Распознавание коллизий и сообщение о них пострадавшим пользователям в первоначальной системе Алоха направлялись по радио на центральный пункт. Также это могло осуществляться путем применения положительных подтверждений в сочетании с перерывом. При обнаружении столкновения пострадавшие станции предпринимают попытки повторной передачи потерянного сообщения, но они должны распределять время попыток случайным образом, следуя некоторому алгоритму столкновения нового конфликта.

Стратегия доступа типа Чистой Алохи позволяет добиться производительности самое большее $1/2e \approx 0,18$ пропускной способности канала. Введем сначала некоторые определения. За доступ к каналу состязаются N станций. Станция передает, в среднем, 1 пакетов в секунду (интенсивность обращений к сети). Величина $1/m$ представляет собой пропускную способность канала (m) в передаваемых пакетах в секунду. Рассмотрим теперь частный случай, при котором все передаваемые сообщения (пакеты) имеют среднюю длину t , соответствующую m единицам времени передачи. Будем считать, что интенсивность нагрузки S (эквивалентно ρ - нормированной по m нагрузке) характеризует использование канала вновь поступающими пакетами

$$S \equiv \rho = N\lambda m_{(1)}$$

Величина $1/t$, которая обозначается μ , представляет собой пропускную способность канала в передаваемых пакетах в секунду. Таким образом, $N\lambda/\mu = N\lambda m$ - относительное использование канала, или производительность, нормированная относительно каждого компьютера одинакова. Общая интенсивность пакетов, передаваемых в канал, включая вновь генерируемые и передаваемые повторно, имеет некоторое значение $\lambda' > \lambda$. Это происходит, потому что из-за коллизий от

каждого компьютера будет передаваться больше сообщений из-за необходимости возобновлять поток. Тогда фактическая интенсивность нагрузки, или использование канала, является параметром G , который равен:

$$G = N\lambda'\tau \quad (2)$$

Рассмотрим типичное сообщение длительностью s , показанное на рис. 2.2.

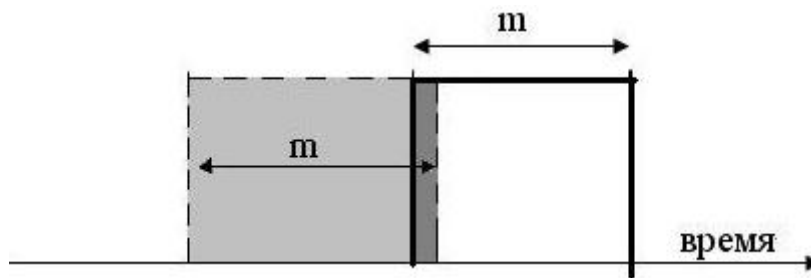


Рисунок 2.2. Столкновение двух сообщений

Оно подвергается столкновению с другим сообщением, если эти два сообщения будут наложены одно на другое в любой точке. Легко заметить, передвигая пунктирное сообщение во времени, что столкновение может произойти в промежутке времени продолжительностью 2τ . Вероятность того, что в промежутке 2τ с не произойдет столкновения, равна

$$e^{-2N\lambda'm} = e^{-2G} \quad (3)$$

Отношение S/G представляет долю сообщений из числа передаваемых в канал, которые проходят успешно. Это число должно быть равно вероятности отсутствия столкновений. Таким образом, уравнение производительности для чистой Алохи:

$$S = Ge^{-2G} \quad (4)$$

Здесь S - нормированная производительность (средняя скорость поступления пакетов, деленная на максимальную производительность $1/m$), а G - нормированная пропущенная нагрузка. Таким образом, S - независимая переменная, а G - ее функция. График зависимости G от S имеет вид двузначной кривой (рисунок 2.3).

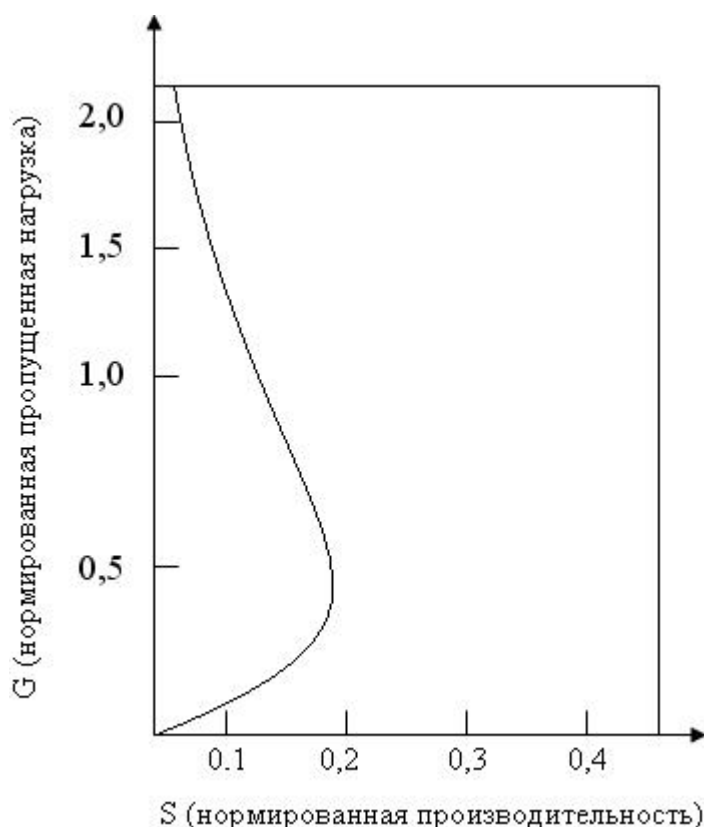


Рисунок 2.3. Характеристика производительности. Чистая Алоха

Отметим, что S имеет максимум $S = 0,5e^{-1} \approx 0,18$ при $G = 0.5$. Судя по формуле (4) или кривой при малой поступающей нагрузке S столкновения происходят редко, и $G \approx S$. Когда S начинает расти, приближаясь к максимальному значению 0.18, число столкновений быстро увеличивается, что ведет в свою очередь к росту вероятности столкновения. Система теряет устойчивость, S падает, а G увеличивается до больших значений.

СИНХРОННАЯ АЛОХА

Максимально возможная производительность схемы чистой Алохи может быть удвоена с помощью простого приема разметки шкалы времени, и разрешения пользователям начинать попытки передачи сообщений только в начале каждого временного интервала m (равного длительности сообщения). Эта схема требует, чтобы работа всех пользователей системы была синхронизирована во времени. Пример работы такой системы показан на рисунке 2.4, на котором одно сообщение передано успешно, а с другим произошло столкновение.

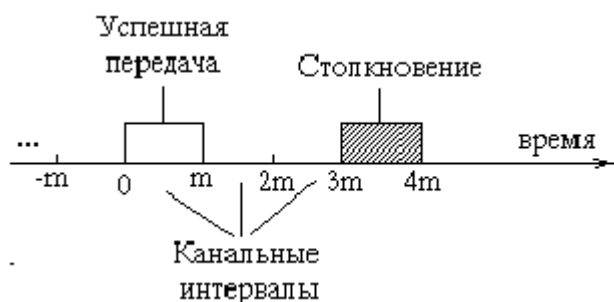


Рис 2.4. Передача при синхронной Алохе

Поскольку сообщения могут быть переданы только в размеченные промежутки времени, столкновения происходят лишь, когда одна или несколько попыток передачи совершаются в том же промежутке.

Вероятность успешной передачи задается в виде формулы:

$$e^{-G} \quad (5)$$

а производительность для синхронной Алохи рассчитывается по формуле:

$$S = Ge^{-G} \quad (6)$$

Нормированная производительность S достигает максимального значения $1/e \approx 0,368$ при $G = 1$. Зависимость пропущенной нагрузки от производительности для синхронной Алохи показана на рисунке 2.5, где она сравнивается с соответствующей зависимостью для чистой Алохи.

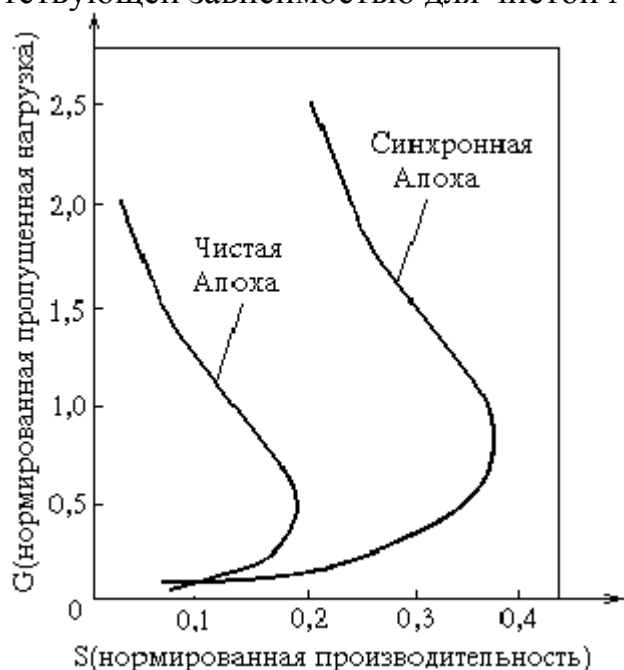


Рис. 2.5. Характеристика производительности. Синхронная Алоха

Из приведенной характеристики, очевидно, что ввиду двух возможных значений G при заданной производительности S , для этой системы доступа также характерна неустойчивость.

ПЗ10. Исследование методов МДПН/ОС

Этот метод является развитием метода чистой Алохи и позволяет улучшить ее характеристики. Метод МДПН/ОС (в оригинале CSMA/CD) входит в протокол сети Ethernet и является одним из стандартных методов в ЛВС (стандарт IEEE 802.3).

Рассмотрим основную идею метода CSMA/CD.

Все станции прослушивают передачу по линии. Станция, желающая

передать сообщение, выходит на связь только после обнаружения свободного состояния канала. Эта процедура называется проверкой несущей (ПН), а соответствующая стратегия доступа - МДПН.

Очевидно, столкновение все же может возникнуть из-за того, что станции физически разнесены одна от другой, а время распространения сигнала t - конечно. Если столкновение все же происходит и станции его обнаруживают (ОС), то они передают сигнал о помехе, и отменяют свои передачи.

Был предложен и проанализирован ряд методов МДПН, они различаются тем, как происходит управление передачей, если канал окажется занятым.

Например, в схеме с p -настойчивостью станция, обнаружившая занятым канал, осуществляет передачу после того, как канал освободится с вероятностью p , и с вероятностью $1-p$, отказывается от передачи на промежутке t , где t - время распространения сигнала в линии.

При схеме с 1-настойчивостью станция осуществляет передачу сразу, как только канал окажется свободным.

При ненастойчивой схеме станция переносит передачу на другое случайное время (в соответствии с некоторым распределением задержек передачи).

Все эти схемы рассчитаны на то, что станция сможет обнаружить конец передачи вскоре после ее завершения. Т.е. они требуют, чтобы $t \ll m$. Это условие обычно описывается параметром

$$a = t/m \ll 1.$$

Протокол МДПН/ОС (CSMA/CD), работающий по схеме с не настойчивостью, был принят в качестве стандартного протокола в ЛВС Ethernet. Если обнаруживается столкновение, и передача прекращается, то попытка повторной передачи предпринимается через случайный промежуток времени, как в схемах Алохи. Этот случайный промежуток удваивается каждый раз после обнаружения нового столкновения до некоторой максимальной величины, при которой станция выходит из строя и извещает вышестоящие условия о нарушении связи.

Расчет характеристик метода CSMA/CD.

Рассмотрим структуру с общей шиной и пусть А и В - 2 наиболее удаленные станции (рис. 6.14). Рассчитаем среднее время для успешной передачи сообщения.

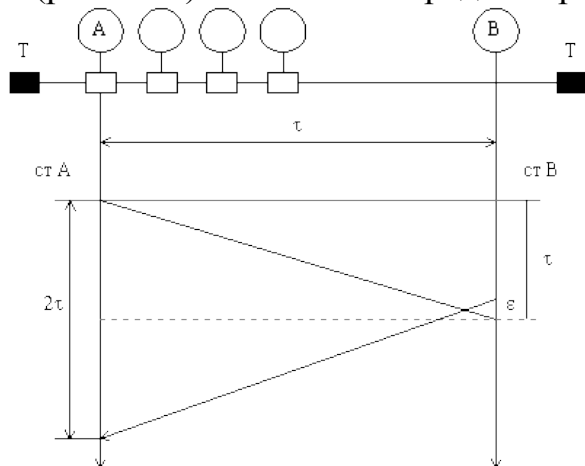


Рис. 6.14.

Сохраняя прежние обозначения назовем время до успешного завершения передачи сообщения - виртуальным временем t_v . Оно включает время передачи кадра m , время t - необходимое для проверки завершения передачи и время кратное $2t$ для разрешения столкновений, если они возникнут.

Итак, пусть возникло столкновение между сигналами, передаваемыми станциями А и В. В худшем случае обнаружение столкновения займет на станциях А и В время $2t$ сек, после чего передача будет выключена.



Рис. 6.15. Расчет величины t_v в методе CSMA/CD.

Если произошло столкновение, то предположим, что нам требуется еще $2tI$ ед. времени для его разрешения, где I - среднее число повторных передач, после того как столкновение произошло. Оно сопоставимо с параметром $p = G/S - 1$, который был введен для чистой Алохи. Тогда

$$t_v = m + t + 2t * I = m[1 + a(1 + 2I)], \quad (6.19)$$

где $a = t/m$.

Теперь необходимо найти величину I . Она зависит от стратегии повторных передач.

Предположим, что длительность интервала столкновений описывается геометрическим распределением единиц, кратных $2t$, с параметром p . В частности, этот интервал равен $2t$ с вероятностью p , $4t$ с вероятностью $p(1-p)$, $6t$ с вероятностью $p(1-p)^2$. Т.о. p является вероятностью успеха, а $(1-p)$ - вероятностью неуспеха, т.е. столкновения.

Тогда среднее число повторных попыток передач

$$I = \sum_{k=1}^{\infty} k p (1-p)^{k-1} = \frac{1}{p} \quad (6.20)$$

Теперь необходимо найти p . Пусть в передаче может участвовать n станций ($n \gg 1$). Пусть вероятность того, что 1 станция намеревается передавать в промежутке $2t$ равна p . Тогда вероятность, что передает точно 1 станция и эта передача успешна (столкновения нет), равна

$$p = np(1-p)^{n-1}. \quad (6.21)$$

Будем искать $\max_p p$. Легко проверить, что величина $p_0 = 1/n$ максимизирует $p(p)$. Учитывая, что $n \gg 1$, в пределе получим

$$p_{\max} = \lim_{n \rightarrow \infty} \left(1 - \frac{1}{n}\right)^{n-1} = e^{-1} \quad (6.22)$$

Т.о. величина $p_{\max} = e^{-1}$ и $I = 1/p = e$.

В результате найдем время успешной передачи

$$t_v = m[1 + a(1 + 2e)], \quad (6.23)$$

где $a = t/m$.

Тогда максимальная производительность протокола

$$l_{\max} = 1/t_v,$$

$$a \quad r_{\max} = l_{\max} \times m = \frac{1}{1 + a(1 + 2e)} \quad (6.24)$$

Обозначим через l среднюю общую интенсивность передачи по каналу сообщений от всех станций, найдем нормированную пропускную способность (ПС) протокола CSMA/CD

$$r = l \times m < \frac{1}{1 + a(1 + 2e)} \cong \frac{1}{1 + 6,44a}.$$

Пример.

Пусть $a = 0,1$. Для этого значения имеем $r \approx 0,6$, что существенно лучше, чем для чистой Алохи ($r = 0,184$) и синхронной Алохи ($r = 0,368$).

Стандарты IEEE на ЛВС.

Стандарты на ЛВС 802 комиссией института IEEE. Эти стандарты рассматривают физический уровень и частично уровень канала, и согласуются с эталонной моделью ВОС/МОС.

В настоящее время наиболее известны следующие стандарты на ЛВС:

- ☐ IEEE 802.3 на ЛВС с общей шиной и методом доступа CSMA/CD,
- ☐ IEEE 802.4 на ЛВС с общей шиной и передачей маркера,
- ☐ IEEE 802.5 на ЛВС кольцевой структуры с передачей маркера.

Стандарт IEEE 802.3.

Первая работа по созданию метода случайного доступа CSMA/CD была выполнена компанией XEROX, которая создала сеть Ethernet. Она в дальнейшем получила развитие, и совместными усилиями компаний DEC, Intel и XEROX была разработана подробная спецификация, которая в дальнейшем была утверждена как стандарт IEEE 802.3.

Стандарт IEEE 802.5 на кольцо с передачей маркера основан на разработках Цюрихской ХИЛ фирмы IBM. Стандарты методов доступа, входящие в семейство IEEE 802. были разработаны в соответствии с эталонной моделью ВОС. Они охватывают физический и часть уровня канала этой модели.

Эти стандарты описывают форматы и протоколы, применяемые на подуровне управления доступом к передающей среде уровня канала, а также на физическом уровне. В свою очередь все стандарты доступа связаны с вышестоящим уровнем. ВОС/МОС через стандарты логического управления каналом IEEE 802.2, который составляет верхнюю часть уровня канала.

Взаимосвязь между эталонной моделью ВОС/МОС и стандартами IEEE 802. для ЛВС приводятся на рис. 6.16.



Рис. 6.16.

Дадим краткую характеристику стандартов IEEE 802. Канальный уровень ВЛВС делится на 2 подуровня: подуровень управления логическим звеном (УЛЗ) и подуровень управления доступом к передающей среде (УДС).

Подуровень УЛЗ с точки зрения стандартной модели ВОС пользуется услугами подуровня управления доступом к передающей среде (УДС) для предоставления услуг сетевому и другим вышестоящим уровням. УЛЗ выполняет также функции канала передачи данных, которые не зависят от типа передающей среды и методы доступа описаны 2 типа процедур управления ЛЗ.

Первый тип, называемый услугами соединения, выполняет функции восстановления сбоев, основанные на асинхронном балансном режиме (LAPB).

Второй тип, называемый услугой без соединения, не предоставляет возможности восстановления сбоев. Эта возможность для распределенных сетей предоставляет транспортный протокол 4 класса - ТП-4.

Одним из применений услуги без соединения подуровня УЛЗ может быть соединение ЛВС с широкополосной передающей средой, например по оптоволоконному каналу.

Локальная сеть Ethernet с методом CSMA/CD. Функции физического уровня сети Ethernet (стандарт IEEE 802.3) показаны на рис. 6.17.



Рис. 6.17.

Физический уровень осуществляет проверку наличия передачи в передающей среде, путем прослушивания канала и передает сигнал проверки несущей подуровню управления доступом к передающей среде.

Этот же уровень сравнивает сигнал в среде с сигналом, генерированным при передаче и выдает сигнал об обнаружении столкновения, если оно возникло. Эти функции выполняют блок доступа к каналу передачи. Для генерирования сигнала проверки несущей служат блоки доступа к каналу передачи и канала приема. Эти сигналы генерируются подуровнем УДС канального уровня.

Кроме проверки двух сигналов блоки доступа к каналу передают символы в кабель и принимают их из кабеля. Блок кодирования передаваемых символов кодирует передаваемую информацию в двоичные сигналы с помощью так называемого <манчестерского кода>. (см. рис. 6.18).

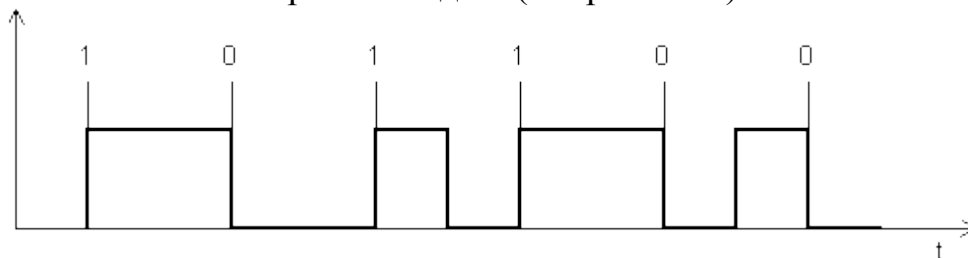


Рис. 6.18.Манчестерский код.

Функции подуровня УДС.

Структура подуровня УДС приведена на рис. 6.19.

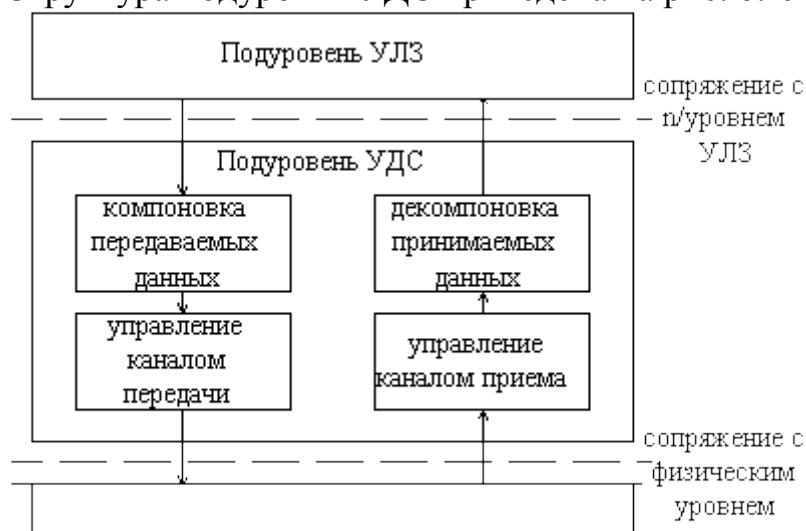


Рис. 6.19.

Блоки управления каналом предотвращают столкновение, используя сигналы проверки несущей, и разрешают столкновения, используя сигналы обнаружения столкновений (ОС).

Формирование кадров, адресование и обнаружение ошибок выполняется блоками компоновки и декомпоновки кадров.

Информация по каналу передается в виде кадров. Формат кадра передается в виде кадров. Формат кадра сети Ethernet представлен на рис. 6.20.

Преамбула	нач. огр.	Адрес получателя	Адрес источника	Тип	Данные	ЗАП	ППК
7 октет	1 окт	6 октет	6 октет	2 октет	46-1500		4 октет

Рис. 6.20.

Преамбула (7 октет) используются для целей синхронизации. Эти же функции выполняет начальный ограничитель (1 октет). Поле типа резервируется для вышестоящих уровней. В поле Данные приводится информация пользователя, размер его изменяется от 46 до 1500 октет.

В случае, если объем передаваемых данных меньше минимального размера (46 октет), то используется вспомогательное поле ЗАП (запас), для дополнения информационной части до 46 октет.

Поле ППК - проверочная последовательность кадра используется для контроля достоверности передаваемой информации и обнаружения ошибок.

Таким образом, размер кадра колеблется от 72 до 1526 октет.

Спецификация сети Ethernet.

Спецификации сети Ethernet предусматривают физический канал в виде коаксиального кабеля со скоростью передачи 10 Мбит/с и 100 Мбит/с.

При такой скорости (10Мбит/с) общая длина коаксиального кабеля между двумя любыми приемопередатчиками (ПП) не превышает 1500м. Скорость распространения сигнала в кабеле 4,33 мкс/км. При этой скорости время распространения сигнала $t \approx 6,5$ мкс. При самой короткой длине кадра 72 октет нижний предел время передачи кадра $m=57,6$ мкс, это дает величину $a=t/m > 0$. И, что достаточно велико. Если поле данных увеличивается до 200 октет, то параметр a уменьшается до величины $a=6,5/180=0,036$, что гораздо лучше.

Кроме того, для обеспечения надлежащего времени восстановления управляющих устройств и физического канала принимается минимальное межкадровое время 9,6 мкс. При этом условии станция, желающая передать сообщение, должна задержать передачу по крайней мере на 9,6 мкс, после обнаружения конца кадра в канале, что снижает вероятность столкновений.

Станция сети Ethernet не обязательно должны быть подключены к одному и тому же кабелю. Отрезки кабеля могут быть соединены через т.н. регенераторы (или репитеры).

Типичные примеры конфигураций приведены на рис. 21. Отдельный сегмент кабеля не должен превышать $l=500$ м. В тракте между двумя РС не может быть более 3^x регенераторов.

Максимальное число РС в любой части сети Ethernet равно 1024, причем в одном сегменте не может быть более 100 РС. Соединительный канал между регенераторами от одного конца до другого не может превышать 1000 м.

В сети Ethernet для управления повторными передачами применяется усеченный алгоритм двоичного замедления. После обнаружения столкновений задержка перед повторной передачей определяется как случайное число e канальных интервалов (по 512 разрядов), равномерно распределенных в диапазоне от 1 до 2^n , где n - номер повторной попытки ($n \leq 10$). В случае, когда

число k повторных попыток превышает 10, то интервал задержки фиксируется на уровне 2^{10} до достижения 15 попытки.

После этого событие столкновения интерпретируется как сбой, станция отключается от сети, и эта ситуация разрешается на вышестоящем уровне.

ЛВС с передачей маркера.

Рассмотрим организацию работы ЛВС с передачей маркера. В настоящее время наиболее известными являются следующие стандарты IEEE 802.4 - на шину с передачей маркера и IEEE 802.5 - на кольцо с передачей маркера (Token Ring).

Архитектуры этих ЛВС приведены на рис. 6.21 и 6.22.

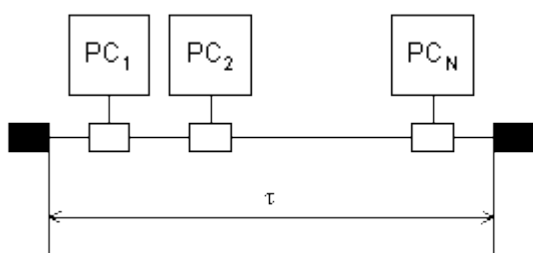


Рис. 6.21 Шина с передачей маркера

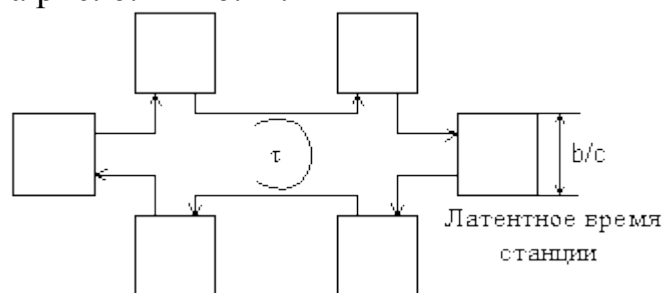


Рис. 6.22 Кольцо с передачей маркера

Все станции работают в децентрализованном режиме. Все сообщения (кадры) передаются по кольцу и активно повторяются каждой станцией, через которую они проходят. Станция, на которой в качестве адреса получателя считывается ее собственный адрес, копирует этот кадр во время его поразрядной передачи на следующую станцию кольца. Затем информация, содержащаяся в кадре, передается в терминалы или контроллеры, включенные на этой станции.

Циркулирующий кадр удаляется из кольца станцией, которая его передала, после его прохождения по кольцу и поступлению в источник. Время, требуемое для организации повторения кадра на станции, т.е. его задержки на данной станции называется латентным временем станции.

Латентный период кольца - это величина, эквивалентна L , она равна сумме задержек t распространения сигнала по кольцу и латентного времени b/c станций $L = t + Nb/c$.

Рассмотрим теперь технологию передачи маркера.

Пусть по кольцу циркулирует свободный маркер (специальный кадр).

Первая станция, имеющая данные для передачи, захватывает маркер и переходит к передаче, а по ее окончании передает маркер дальше.

Существует ряд стратегий выполнения этой операции.

Одна стратегия (подобная передаче управления) состоит в выдаче маркера в составе передаваемого кадра, т. о. следующая станция получает разрешение на передачу как только на рассматриваемой станции завершится передача последнего разряда кадра.

Другая стратегия состоит в передаче маркера только после того, как передающая станция получит обратно свое сообщение, прошедшее по кольцу.

Здесь возможны 2 версии:

1) ожидание того момента, пока не будет принято (и стерто) все сообщение;

2) выдача нового маркера, как только будет принят предыдущий.

В обоих случаях в кольце допускается в любой момент времени только 1 маркер. Определим среднее время t_f необходимое для передачи данных от станции-источника до станции-назначения в кольце.

Оно в точности равно ранее приведенной формуле для циклического опроса с добавлением времени передачи кадра m и величины $L/2$ для учета задержки при распространении сигнала до получателя на пути, который в среднем равен половине длины кольца:

$$t_f = \frac{L(1 - \rho/N)}{2(1 - \rho)} + \frac{\lambda \bar{m}^2}{2(1 - \rho)} + m + \frac{L}{2} \quad (6.25)$$

где $r = N \lambda \bar{m}$, m - длительность передачи кадра,
 $L = t + Nb/c$, b/c - латентное время станции.

Сравнение характеристик методов доступа CSMA/CD и кольца с передачей маркера. На рис.6.24 и 6.25 приводятся зависимости среднего времени задержки в доставке t_f/m для методов доступа CSMA/CD и маркерного метода доступа в зависимости от загрузки $r = \lambda m$, для различных скоростей передачи $V_{\text{пер}} = 1$ Мбит/с и $V_{\text{пер}} = 10$ Мбит/с. Число станций $N = 50$. Латентное время станции - 1 бит/ст.

Длина пакета распределена показательно со средней $l_{\text{cp}} = 1000$ бит, заголовок - 24 разряда.

На рис. 6.26 приводится зависимость нормированной средней задержки для метода доступа CSMA/CD в зависимости от значений $a = t/m$ для различных значений $r = \lambda m$. В пределе при $a \rightarrow 0$

$$t_f = m + \frac{\lambda \bar{m}^2}{2(1 - \rho)}$$

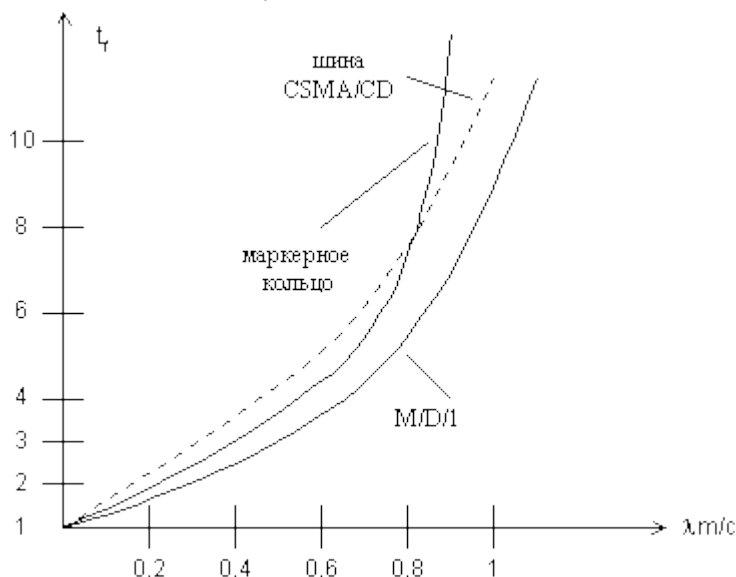


Рис. 6.24. Зависимость задержки передачи от производительности.

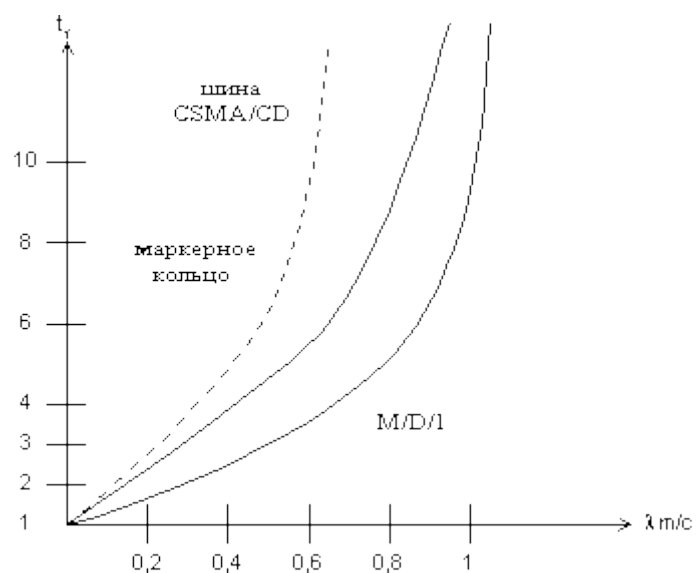


Рис.6.25.

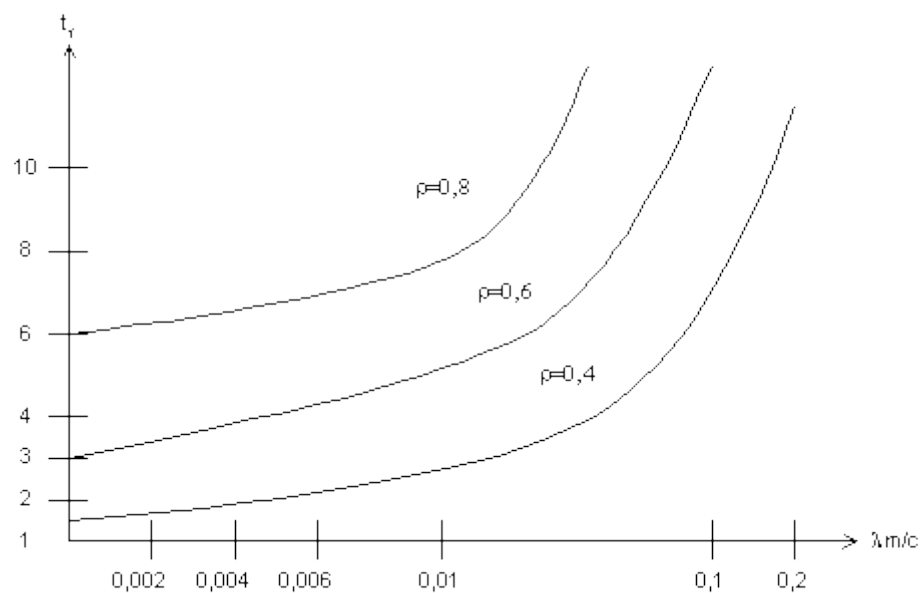


Рис.6.26. Зависимость нормированной средней задержки от $a=t/m$.