МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

МОСКОВСКИЙ ЭНЕРГЕТИЧЕСКИЙ ИНСТИТУТ
(ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

В.Н. Балашов, А.Г. Гольцов

МОДЕЛИРОВАНИЕ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ

Рекомендации к лабораторным занятиям
по курсу "Моделирование

для студентов, обучающихся по направлению
"Информатика и вычислительная техника"

Дистанционное обучение

Москва Издательство МЭИ 2016

УДК
621.398
П692
УДК:681.3

Утверждено учебным управлением МЭИ
Подготовлено на кафедре вычислительных машин, систем и сетей.

Балашов В.Н., Гольцов А.Г.
Моделирование генераторов случайных чисел
Рекомендации к лабораторным занятиям. Методическое пособие по курсу "Моделирование" / – М.: Изд-во МЭИ, 2016, 40с.

Представлены описание лабораторной работы, выполняемой на ПЭВМ типа IBM PC , по имитационному моделированию систем массового обслуживания (СМО). Лабораторная работа включает теоретический материал, рекомендации и индивидуальные задания, в соответствии с которыми проводится аналитический расчет основных характеристик СМО, а затем проводится компьютерное имитационное моделирование тог же варианта СМО с последующим сравнением результатов. Предназначен студентам, обучающимся по направлению подготовки 09.03.01 Информатика и вычислительная техника, профиль "Вычислительные машины, комплексы, системы и сети", изучающим курс "Моделирование".
Работа выполняется по индивидуальным заданиям.

1. АНАЛИТИЧЕСКИЕ МЕТОДЫ МОДЕЛИРОВАНИЯ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ.

1.1. Структура СМО

Рассмотрим в качестве примера структуру многоканальной СМО с очередью:

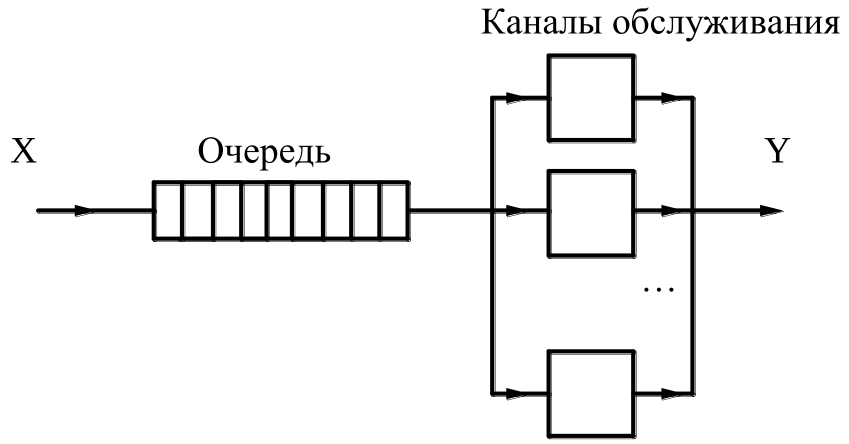

Рис. 1. Структура многоканальной СМО с очередью

СМО содержит несколько каналов обслуживания, соединенных по входам параллельно. Между входом СМО X и каналами обслуживания помещен блок "очередь", в который помещаются заявки, если все каналы обслуживания заняты.

Поток заявок. Поток заявок X, поступающий на вход реальной СМО, определяется конкретной ситуацией и в принципе неуправляем. При построении математической модели СМО необходимо построить и модель потока заявок. В качестве модели потока заявок принимается случайный процесс (последовательность случайных величин с заданным распределением), определяющий случайные моменты времени поступления заявок.

Каналы обслуживания. Заявки поступают на входы свободных каналов обслуживания. Каждая заявка обслуживается в одном канале определенное время, называемое временем обслуживания. СМО может включать один канал обслуживания (одноканальная СМО) или несколько каналов (многоканальная СМО).

Поток обслуживания. Время обслуживания каждой заявки определяется потоком обслуживания, который с математической точки зрения аналогичен потоку заявок.

Очередь. В структуру СМО может входить блок "очередь". Очередь характеризуется числом мест ожидания для заявок. В модели СМО в очереди запоминаются моменты времени поступления заявки. Очередь может быть конечной, тогда в случае переполнения заявки теряются (СМО с потерями) или бесконечной. Существуют СМО без очереди.

Структура СМО обычно характеризуется последовательностью из четырех символов {A, B, n, m}, где A и B - соответственно поток заявок и поток обслуживания, n и m - соответственно число каналов и число мест в очереди.

Состояния. Математическая модель СМО отличается от схемы, представленной на рис. 3.1. Модель СМО имеет конечное число дискретных состояний $S_0, S_1, \ldots, S_{n+m}$.

Состояние S_0 означает, что в системе нет заявок. Состояние S_1 означает, что в системе обслуживается одна заявка, очередь пуста. Состояние S_n - обслуживается n заявок, очередь пуста. Состояние S_{n+1}, - обслуживается n заявок и в очереди одна заявка. Состояние S_{n+m}, - обслуживается n заявок и в очереди m заявок. Система полностью заполнена заявками, поэтому очередная заявка будет отброшена.

Состояния $S_0, S_1, \ldots, S_{n+m}$ являются дискретной случайной величиной. Каждому состоянию S_i соответствует вероятность пребывания в этом состоянии P_i.

Замечание. Модели СМО соответствует математическая модель - Марковский процесс с непрерывным временем.

1.2 Дисциплина обслуживания

В СМО с очередью существуют различные дисциплины обслуживания (порядок обслуживания) заявок, поступивших на вход:

Обслуживание в порядке поступления заявок (FIFO, First In First Out),

Обслуживание в порядке, обратном порядку поступления заявок (LIFO, Last In First Out).

Обслуживание по приоритетам. Каждой заявке в очереди заранее присваивается приоритет (определенное число). Первой обслуживается заявка из очереди, имеющая максимальный приоритет. Приоритет бывает абсолютным или относительным. Если в очередь СМО с абсолютным приоритетом поступает заявка с приоритетом, превышающим приоритет обслуживаемой заявки, то обслуживание прекращается и начинается обслуживание поступившей заявки. Если в очередь СМО с относительным приоритетом поступает заявка с приоритетом, превышающим приоритет обслуживаемой заявки, то обслуживание поступившей заявки начинается после окончания обслуживания текущей заявки.

Случайный порядок обслуживания. Выбирается случайным образом одна из заявок в очереди.

Обслуживание в первую очередь заявок с минимальным временем обслуживания. Это очень эффективная дисциплина обслуживания, позволяющая минимизировать среднее время пребывания заявок в системе. Однако в большинстве практических случаев время обслуживания заранее не известно.

1.3. Характеристики СМО

В процессе моделирования необходимо определить количественные характеристики СМО, позволяющие оценить их работу. Поток заявок и поток обслуживания в модели СМО - последовательности случайных величин. Поэтому для определения характеристик СМО применяют методы теории вероятностей и математической статистики.

Обычно представляют интерес следующие характеристики СМО:

$P_{отк}$ - вероятность потери заявки из-за занятости очереди (вероятность отказа),

A - пропускная способность - среднее количество обслуженных заявок в единицу времени,

$k_{ср}$ - среднее количество занятых каналов обслуживания,

$r_{ср}$ - средняя длина очереди

$t_{ож}$ - среднее время ожидания заявкой обслуживания в очереди.

1.4. Потоки заявок и потоки обслуживания

В аналитических моделях СМО применяют простейшие потоки, в которых интервалы между поступлениями заявок подчинены экспоненциальному или Пуассоновскому закону.

В экспоненциальном потоке интервалы между поступлениями отдельных заявок подчиняются закону с плотностью распределения $p(t) = \lambda e^{-\lambda t}$ и функцией распределения $F(t) = 1 - e^{-\lambda t}$. Средняя длина интервала между заявками (математическое ожидание) для экспоненциального потока равно

$$M(t) = \int_0^\infty tp(t)dt = \int_0^\infty \lambda te^{-\lambda t}dt = 1/\lambda \quad ,$$

где λ - интенсивность потока.

Дисперсия для экспоненциального потока равна

$$D(t) = \int_0^\infty (t - M(t))^2 p(t)dt = 1/\lambda^2 \quad .$$

Перечислим свойства экспоненциального потока заявок.

Суммарный поток заявок, полученный объединением конечного числа экспоненциальных законов с различными интенсивностями λ_i , является экспоненциальным потоком с интенсивностью

$$\lambda = \sum_{i=1}^{N} \lambda_i ;$$

Поток заявок, полученный разрежением экспоненциального закона, при котором заявки исключаются из потока с вероятностью p, является экспоненциальным потоком с интенсивностью λp.

Важность экспоненциального потока в моделировании СМО определяется не только возможностью на его основе получать аналитические решения, но и хорошим совпадением с практикой. Многие реально наблюдаемые потоки заявок в различных СМО близки к экспоненциальному потоку.

1.5. Аналитическая модель системы массового обслуживания

Многие важные результаты, связанные с функционированием систем массового обслуживания (СМО), получены при помощи аналитических моделей. Наиболее распространены и изучены Марковские модели, основанные на математическом аппарате Марковских случайных процессов с непрерывным временем.

1.5.1 Состояния и интенсивность потока событий

Будем считать, что СМО имеет конечное число дискретных состояний S_1, S_2, S_3,…, S_n , при этом процесс перехода из одного состояния в другое происходит мгновенно в любой случайный момент времени (время непрерывно). Случайный процесс перехода системы из одного состояния S_i в другое состояние S_j называется Марковским, если переходная вероятность p_{ij} зависит только от текущего состояния системы и не зависит от прошлых состояний.

Для построения Марковской модели вычислительной системы необходимо определить все возможные состояния системы и показать допустимые переходы из одного состояния в другое. Для этого строится граф состояний, в котором состояния обозначаются кружками, а переходы - стрелками.

Граф, представленный на этом рисунке, содержит 5 состояний и 8 переходов. Для каждого перехода указана интенсивность потока событий λ_{ij}, (измеряется в 1/сек), определяющая переход системы из состояния S_i в состояние S_j.

$$\lambda_{ij}(t) = \lim_{\Delta t \to 0}(p_{ij}(t, t+\Delta t)) / \Delta t;$$

где p_{ij} $(t, t + \Delta t)$ - вероятность перехода из состояния i в состояние j за время $(t, t + \Delta t)$.

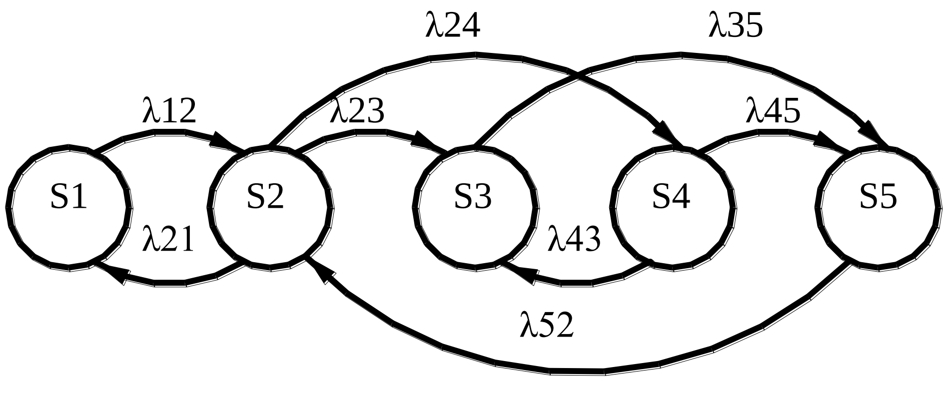

Рис. 2. Граф состояний Марковской модели

Если случайный процесс является стационарным, то интенсивность потока событий не зависит от времени и определяется соотношением

$$\lambda_{ij} = \lim_{\Delta t \to 0} p_{ij}(\Delta t) / \Delta t;$$

1.5.2. Уравнения Колмогорова

Переход Марковской модели из одного состояния в другое определяется системой дифференциальных уравнений Колмогорова. Вывод этой системы уравнений приведен в курсе лекций. Запишем эту систему на конкретном примере Марковского процесса с графом, представленным на рис. 2.

В каждый момент времени t система находится в одном из состояний S_i с вероятностью $p_i(t)$. Придадим времени малое приращение Δt. За это время система может остаться в прежнем состоянии или перейти в новое состояние. Проведем расчет вероятностей перехода для этой ситуации.

Система находится в состоянии S_1. Система останется в этом состоянии с вероятностью $p_1(t)$.

Вероятность того, что система уйдет из этого состояния в состояние S_2, равна $p_1(t)\,\lambda_{12}\,\Delta t$.

Вероятность того, что система перейдет из состояния S_2 в состояние S_1, равна $p_2(t)\,\lambda_{21}\,\Delta t$.

Объединяя эти соотношения, получим вероятность того, что система будет находиться в состоянии S_1 в момент времени $(t + \Delta t)$

$$p_1(t + \Delta t) = p_1(t) - p_1(t)\,\lambda_{12}\,\Delta t + p_2(t)\,\lambda_{21}\,\Delta t;$$

Вероятность того, что система останется в состоянии S_1 увеличивается за счет вероятности перехода из состояния S_2 (знак +) и уменьшается за счет ухода в состояние S_1 (знак -).

Перепишем это уравнение в следующем виде

$$\frac{p_1(t+\Delta t) - p_1(t)}{\Delta t} = p_2(t)\lambda_{21} - p_1(t)\lambda_{12};$$

В результате предельного перехода при $\Delta t \to 0$ в левой части этого уравнения получим производную. Тогда

$$\frac{dp_1(t)}{dt} = p_2(t)\lambda_{21} - p_1(t)\lambda_{12};$$

Это первое уравнение системы, определяющее изменение вероятности $p_1(t)$ нахождения системы в состоянии S_1 во времени.

Система находится в состоянии S_2. Вероятность того, что система останется в этом состоянии, равна $p_2(t)$.

Вероятность того, что система уйдет из этого состояния в состояния S_1, S_3 и S_4 равна $p_2(t)\lambda_{21}\Delta t + p_2(t)\lambda_{23}\Delta t + p_2(t)\lambda_{24}\Delta t$.

Вероятность того, что система перейдет из состояний S_1 и S_5 в состояние S_2 равна $p_1(t)\,\lambda_{12}\,\Delta t + p_5(t)\,\lambda_{52}\,\Delta t$.

Объединяя эти соотношения, получим, вероятность того, что система будет находиться в состоянии S_2 в момент времени $(t + \Delta t)$

$$P_2(t + \Delta t) = p_2(t) + p_1(t)\,\lambda_{12}\,\Delta t + p_5(t)\,\lambda_{52}\,\Delta t - p_2(t)\,\lambda_{21}\,\Delta t -$$
$$- p_2(t)\lambda_{23}\Delta t - p_2(t)\,\lambda_{24}\,\Delta t .$$

В результате получим второе уравнение системы уравнений Колмогорова

$$\frac{dp_2(t)}{dt} = p_1(t)\lambda_{12} + p_5(t)\lambda_{52} - p_2(t)(\lambda_{21} + \lambda_{23} + \lambda_{24});$$

Аналогично получим остальные уравнения Колмогорова.

$$\frac{dp_3(t)}{dt} = p_2(t)\lambda_{23} + p_4(t)\lambda_{43} - p_3(t)\lambda_{35};$$

$$\frac{dp_4(t)}{dt} = p_2(t)\lambda_{24} - p_4(t)(\lambda_{43} + \lambda_{45});$$

$$\frac{dp_5(t)}{dt} = p_4(t)\lambda_{45} + p_3(t)\lambda_{35} - p_5(t)\lambda_{52};$$

Система дифференциальных уравнений Колмогорова описывает динамику перехода системы из одного состояния в другое. Понятно, что для интегрирования этой системы необходимо задать начальные условия, т.е. необходимо задать вероятности нахождения системы в определенных состояниях в начальный момент времени.

Доказано, что в системах с конечным числом состояний, в которых из одного состояния в любое другое можно перейти за конечное число шагов, существует *стационарный режим работы*, в котором переходные вероятности становятся постоянными и не зависят от времени. Стационарное состояние достигается при длительной работе системы $(t \to \infty)$ и не зависит от распределения переходных вероятностей в начальный момент времени. Для исследования стационарного режима необходимо в системе уравнений Колмогорова прировнять производные в левой части к нулю.

$$\sum_{\substack{j=1\\ j\neq i}}^{n} p_j(t)\lambda_{ji} - p_i(t)\sum_{\substack{j=1\\ j\neq i}}^{n} \lambda_{ij} = 0;$$

с условием

$$\sum_{i=1}^{n} p_1 = 1;$$

Для Марковской модели с графом состояний, приведенным на рис. 2, система уравнений Колмогорова для стационарного режима имеет следующий вид

$$p_2\lambda_{21} - p_1\lambda_{12} = 0;$$

$$p_1\lambda_{12} + p_5\lambda_{52} - p_2(\lambda_{21} + \lambda_{23} + \lambda_{24}) = 0;$$

$$p_2\lambda_{23} + p_4\lambda_{43} - p_3\lambda_{35} = 0;$$
$$p_2\lambda_{24} - p_4(\lambda_{43} + \lambda_{45}) = 0;$$
$$p_4\lambda_{45} + p_3\lambda_{35} - p_5\lambda_{52} = 0;$$

с дополнительным условием

$$p_1 + p_2 + p_3 + p_4 + p_5 = 0;$$

Замечание. Эта система из 6 алгебраических уравнений с 5 неизвестными является линейно зависимой. Поэтому одно любое уравнение можно отбросить. В результате решения системы уравнений Колмогорова находится функция распределения вероятностей дискретной случайной величины, определяющей вероятности пребывания СМО в каждом состоянии.

Значение с.в.	1	2	3	4	5
Вероятность	p_1	p_2	p_3	p_4	p_5

Замечание. Дополнительное условие в стационарной системе уравнений Колмогорова показывают полноту значений дискретной случайной величины, определяющей вероятности пребывания СМО в каждом состоянии.

1.5.3. Одноканальная СМО с очередью

Пусть на вход одноканальной СМО (n = 1 - число каналов, m - число мест в очереди) поступает экспоненциальный поток заявок. Время обслуживания заявки в СМО - случайная величина, распределенная по экспоненциальному закону.

Интенсивность потока заявок на входе СМО равна λ (1/сек).

Среднее время обслуживания заявки обозначим - $t_{обсл}$. Обратная величина к среднему времени обслуживания определяет *интенсивность потока обслуживания* $\mu = 1/\ t_{обсл}$. Эта величина характеризует среднее число заявок, которое может обслужить система в единицу времени.

Определим состояния СМО S_i :

S_0 - система свободна, заявок нет, очереди нет;

S_1 - система обслуживает одну заявку, очереди нет;

S_2 - система обслуживает одну заявку, в очереди одна заявка;

…………………………………………

S_{m+1} - система обслуживает одну заявку, в очереди - m заявок.

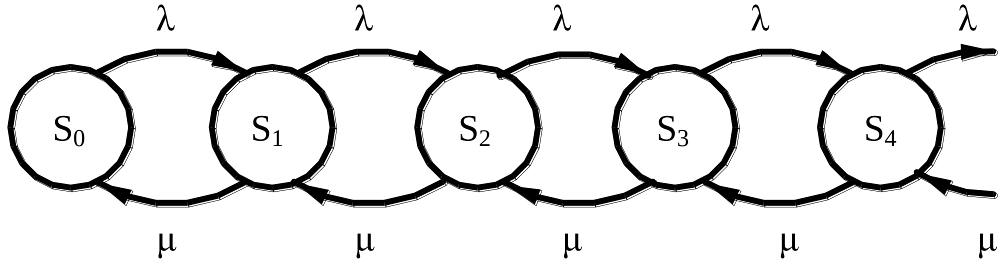

Рис. 3. Граф состояний одноканальной СМО с очередью

Верхние стрелки на этом графе показывают процесс прихода заявок, нижние стрелки - процесс обслуживания заявок.

Запишем уравнения Колмогорова для этой системы

$$
\begin{aligned}
&-\lambda p_0 + \mu p_1 = 0;\\
&-\lambda p_1 - \mu p_1 + \lambda p_0 + \mu p_2 = 0;\\
&-\lambda p_2 - \mu p_2 + \lambda p_1 + \mu p_3 = 0;\\
&-\lambda p_3 - \mu p_3 + \lambda p_2 + \mu p_4 = 0;\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&p_0 + p_1 + p_2 + p_3 + p_4 + \cdots = 1;
\end{aligned}
$$

После решения первых уравнений системы, получим

$$p_k = \left[\frac{\lambda}{\mu}\right]^k p_0;$$

Подставим эти вероятности в последнее условие, получим

$$p_0 + \left(\frac{\lambda}{\mu}\right)p_0 + \left(\frac{\lambda}{\mu}\right)^2 p_0 + \cdots + \left(\frac{\lambda}{\mu}\right)^{m-1} p_0 = 1;$$

Из этой формулы находим p0

$$p_0 = 1/\left(1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2 + \cdots + \left(\frac{\lambda}{\mu}\right)^{m-1}\right);$$

Коэффициент загрузки и коэффициент простоя СМО. Вероятность p_0 соответствует состоянию S_0, а это по определению вероятность того, что система свободна и не обслуживает ни одной заявки. Тогда p_0 - вероятность простоя СМО, или коэффициент простоя. Поэтому величина $p_{заг} = 1 - p_0$ - вероятность загрузки СМО или коэффициент загрузки.

СМО с бесконечной очередью.

Пусть очередь в СМО бесконечна (m = 8). Тогда знаменатель в предыдущей формуле для p0 является геометрической прогрессией. Если $(\lambda/\mu) < 1$, то ряд сходится и его сумма равна

$$S = 1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2 + \cdots + \left(\frac{\lambda}{\mu}\right)^m + \cdots = \frac{1}{1-\frac{\lambda}{\mu}};$$

Это позволяет записать выражение для вероятности p_0. Это коэффициент простоя.

$$p_0 = 1/S = 1 - \frac{\lambda}{\mu};$$

Тогда коэффициент загрузки равен

$$p_{заг} = 1 - p_0 = \frac{\lambda}{\mu} < 1;$$

Вероятность отказа. В СМО с *конечной очередью* существует вероятность состояния, в котором очередная пришедшая заявка не может быть обслужена. Это последнее состояние S_{m+1}. Поэтому вероятность отказа $p_{отк}$ равен вероятности пребывания СМО в этом состоянии.

$$p_{отк} = p_{m+1} = (m+1) \cdot \left(\frac{\lambda}{\mu}\right)^{m+1};$$

В СМО с *бесконечной очередью* вероятность отказа равна нулю, если

$$p_{заг} = \frac{\lambda}{\mu} < 1.$$

Замечание. Условие $p_{заг} < 1$ является фундаментальным в теории СМО. Если принять, что $\lambda = \mu$ или $p_{заг} = 1$, то вероятность отказа

$$p_{отк} = (m+1) \cdot 1 > 1,$$

превышает единицу, что делает СМО неработоспособной.

Пропускная способность СМО. Для СМО с любой структурой *пропускная способность* А - количество обслуженных заявок, покидающих систему, в единицу времени. Пропускная способность равна количеству входящих заявок в единицу времени, умноженное на вероятность того, что заявка будет обслужена

$$A = \lambda \cdot (1 - p_{отк});$$

Среднее число заявок, находящихся в СМО.

В системе с конечной очередью в каждый момент времени может быть одно из следующих состояний:

S_0: в системе 0 заявок;

S_1: в системе 1 заявка в канале обслуживания и 0 заявок в очереди;

S_2: в системе 1 заявка в канале обслуживания и 1 заявка в очереди;

S_3: в системе 1 заявка в канале обслуживания и 2 заявки в очереди;

…………………………………………………………………

S_{m+1}: в системе 1 заявка в канале обслуживания и m заявок в очереди;

Номер состояния S_k соответствует числу k заявок, находящихся в СМО. Вероятность этого состояния равна p_k.

Число заявок, находящихся в СМО, является дискретной случайной величиной со следующей функцией распределения вероятностей

Значение с.в.	0	1	2	…	m + 1
Вероятность p_k	p_0	p_1	p_2		p_{n+1}

Среднее число заявок в СМО $n_{ср}$ равно математическому ожиданию этой случайной величины

$$n_{ср} = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + \cdots + (m+1) \cdot p_{m+1} = \sum_{k=1}^{m+1} k \cdot p_k;$$

или

$$n_{ср} = 1 \cdot \frac{\lambda}{\mu} + 2 \cdot \left(\frac{\lambda}{\mu}\right)^2 + 3 \cdot \left(\frac{\lambda}{\mu}\right)^3 + \cdots (m+1) \cdot \left(\frac{\lambda}{\mu}\right)^{m+1} = \sum_{k=1}^{m+1} k \cdot \left(\frac{\lambda}{\mu}\right)^k;$$

Для *системы с бесконечной очередью* ($m \to \infty$) эта формула превращается в убывающую геометрическую прогрессию. В этом случае среднее число заявок в СМО равно

$$n_{ср} = n_\infty = \frac{\lambda}{\mu} / (1 - \frac{\lambda}{\mu});$$

Среднее число обслуживаемых заявок. В канале обслуживания одноканальной СМО может обслуживаться одна заявка или ни одной заявки. Этому соответствует следующая дискретная случайная величина

Значение с.в.	0	1
Вероятность p_k	p_0	p_1

Среднее число обслуживаемых заявок $n_{оз}$ равно математическому ожиданию этой случайной величины

$$n_{оз} = 0 \cdot p_0 + 1 \cdot p_1 = \frac{\lambda}{\mu};$$

Средняя длина очереди. Средняя длина очереди $r_{ср}$ равна среднему числу заявок в системе $n_{ср}$ за вычетом среднего числа обслуживаемых заявок.

$$r_{ср} = n_{ср} - n_{оз} = 2 \cdot \left(\frac{\lambda}{\mu}\right)^2 + 3 \cdot \left(\frac{\lambda}{\mu}\right)^3 + \cdots (m+1) \cdot \left(\frac{\lambda}{\mu}\right)^{m+1} = \sum_{k=2}^{m+1} k \cdot \left(\frac{\lambda}{\mu}\right)^k;$$

В одноканальной СМО с *бесконечной очередью* средняя длина очереди равна

$$r_{ср} = \left(\frac{\lambda}{\mu}\right)^2 / (1 - \frac{\lambda}{\mu});$$

Средняя длина очереди резко возрастает при стремлении коэффициента загрузки $p_{заг}$ к единице.

$$p_{заг} = \frac{\lambda}{\mu} \to 1;$$

Для $p_{заг}$ больше единицы работа СМО невозможна.

Замечание. Мы рассматриваем только практически значимый *стационарный режим* работы СМО, в котором переходные вероятности перестают зависеть от времени и система может работать сколь угодно долго. Для исследования *переходного режима* необходимо численно решить систему дифференциальных уравнений Колмогорова и проанализировать графики изменения переходных вероятностей во времени, как это было показано выше.

Среднее время пребывания заявки в СМО. В стационарном режиме работы СМО ($p_{заг} < 1$) интенсивности входного и выходного потоков равны между собой и равны λ. Пусть в стационарном режиме среднее число заявок в системе в единицу времени равно $n_{ср}$. Тогда согласно теореме Литтла среднее время пребывания заявки в СМО - $T_{сист}$ равно

$$T_{сист} = n_{ср} / \lambda;$$

Теорема Литтла доказана для любого распределения потока заявок и при любом распределении времени обслуживания.

Для СМО с бесконечной очередью эту формулу можно переписать в следующем виде

$$T_{сист} = \frac{1}{\mu(1-\lambda/\mu)};$$

Аналогично можно записать формулу для среднего времени пребывания заявки в очереди $T_{оч}$. В общем случае

$$T_{оч} = r_{ср} / \lambda;$$

Для СМО с бесконечной очередью получим

$$T_{оч} = \frac{\lambda}{\mu^2(1-\lambda/\mu)};$$

Эти формулы позволяют определить среднее время обслуживания заявки СМО - $t_{обсл}$ из следующего соотношения

$$T_{сист} = t_{обсл} + T_{оч};$$

1.5.4. Многоканальная СМО с очередью

Пусть на вход многоканальной СМО (n - число каналов, m - число мест в очереди) поступает экспоненциальный поток заявок. Длительность обслуживания заявки в каждом канале СМО - случайная величина, также распределенная по экспоненциальному закону. Интенсивность потока заявок одинакова для каждого канала и равна λ (1/сек), среднее время обслуживания заявки - $t_{обсл}$. Обратная величина к среднему времени обслуживания определяет интенсивность потока обслуживания $\mu = 1/\ t_{обсл}$, характеризующее среднее число заявок, которое может обслужить система в единицу времени.

При наличии свободного канала приходящая заявка сразу отправляется на обслуживание. Если все каналы заняты, то заявка ставится в очередь. Если очередь заполнена, то заявка получает отказ. При освобождении канала и наличии заявок в очереди в освободившийся канал сразу помещается первая заявка из очереди. Обслуженные заявки покидают систему.

Определим состояния S_i для многоканальной СМО.

S_0 - система свободна, заявок нет, очереди нет;

S_1 - система обслуживает заявку в одном канале, очереди нет;

S_2 - система обслуживает заявки в двух каналах, очереди нет;

…………………………………………

S_n - система обслуживает заявки в n каналах, очереди нет;

S_{n+1} -система обслуживает заявки в n каналах, одна заявка в очереди;

…………………………………………

S_{n+m} - система обслуживает заявки в n каналах, в очереди находятся m заявок.

Длина очереди m может быть неограниченной.

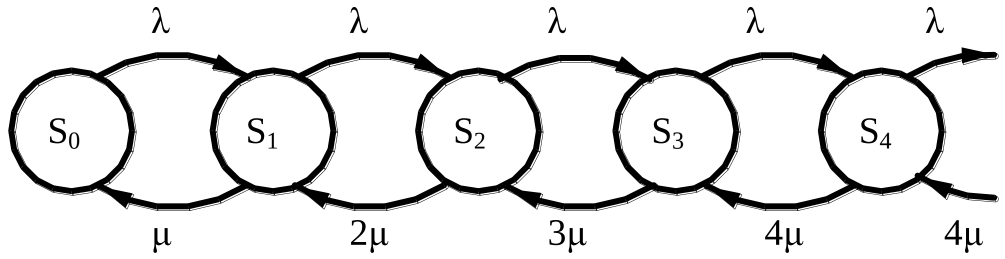

Рис. 4. Граф состояний четырехканальной СМО с очередью

Граф состояний многоканальной СМО подобен графу СМО (рис. 3) и отличается только интенсивностями потока обслуживания (нижние стрелки на графе). При поступлении очередной заявки (если есть свободные каналы) система переходит в соседнее правое состояние. Интенсивность перехода вправо определяется интенсивностью входного потока λ. По-другому обстоит дело с потоком обслуживания. Если система находится в состоянии S_1, то работает один канал и интенсивность обслуживания равна μ. В состоянии S_2 параллельно работают 2 канала и поэтому интенсивность обслуживания равна 2μ. Рост интенсивности обслуживания прекращается после того, как все каналы оказываются занятыми. Для системы с n каналами максимальное значение интенсивности обслуживания равно $n\mu$.

Запишем уравнения Колмогорова для стационарного режима работы 4-х канальной СМО.

$$
\begin{aligned}
&-\lambda p_0 + \mu p_1 = 0;\\
&-\lambda p_1 - \mu p_1 + \lambda p_0 + 2\mu p_2 = 0;\\
&-\lambda p_2 - 2\mu p_2 + \lambda p_1 + 3\mu p_3 = 0;\\
&-\lambda p_3 - 3\mu p_3 + \lambda p_2 + 4\mu p_4 = 0;\\
&-\lambda p_4 - 4\mu p_4 + \lambda p_3 + 4\mu p_5 = 0;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&p_0 + p_1 + p_2 + p_3 + p_4 + \cdots = 1;
\end{aligned}
$$

В этом примере видно, что в первых 4-х уравнениях, связанных с процессом заполнения каналов обслуживания, прослеживается одна закономерность, а в остальных уравнениях - другая.

Выражая последовательно вероятности p_1, p_2, … , p_{n+m} через p_0, получим решение алгебраической системы уравнений Колмогорова

$$
p_1 = \left[\frac{\lambda}{\mu}\right]^1 \frac{1}{1!} p_0; \qquad p_k = \left[\frac{\lambda}{\mu}\right]^k \frac{1}{k!} p_0; \qquad p_n = \left[\frac{\lambda}{\mu}\right]^n \frac{1}{n!} p_0;
$$

для $k < n$,

$$
p_{n+1} = \left[\frac{\lambda}{\mu}\right]^{n+1} \frac{1}{n^1 n!} p_0; \qquad p_{n+i} = \left[\frac{\lambda}{\mu}\right]^k \frac{1}{n^{k-n} n!} p_0; \qquad p_{n+m} = \left[\frac{\lambda}{\mu}\right]^{n+m} \frac{1}{n^m n!} p_0;
$$

для $k > n$,

Вероятность p_0 находится из последнего уравнения после подстановки в него предыдущих формул

$$
p_0\left(1 + \left(\frac{\lambda}{\mu}\right) + \left(\frac{\lambda}{\mu}\right)^2 \frac{1}{2!} + \cdots + \left(\frac{\lambda}{\mu}\right)^n \frac{1}{n!} + \left(\frac{\lambda}{\mu}\right)^{n+1} \frac{1}{nn!} + \cdots \left(\frac{\lambda}{\mu}\right)^{n+m} \frac{1}{n^m n!}\right) = 1;
$$

На следующем шаге находим вероятности p_1, p_2, … p_{n+m}.

Проведем на основе этих выражений анализ основных вариантов функционирования многоканальной СМО.

Система без очереди (с отказами) ($m = 0$).

Вероятность отказа. Очередная заявка, поступающая в СМО, не обслуживается, если все каналы обслуживания заняты. Это значит, что система находится в состоянии S_n. Вероятность нахождения системы в этом состоянии p_n является вероятностью "отказа" в обслуживании.

$$p_{отк} = p_n = \left[\frac{\lambda}{\mu}\right]^n \frac{1}{n!} p_0;$$

Вероятность p_0 соответствует состоянию S_0, означающему, что в СМО нет заявок. Это *коэффициент простоя* p_0.

Величина (1 - p_0) характеризует факт, что СМО занята обслуживанием заявок. Это *коэффициент загрузки* $p_{заг} = 1 - p_0$.

$$A = \lambda \cdot (1 - p_{отк});$$

Пропускная способность A. Это количество обслуженных заявок, покидающих систему в единицу времени. Пропускная способность равна количеству входящих заявок в единицу времени, умноженное на вероятность того, что заявка будет обслужена

Среднее число заявок $n_{ср}$, находящихся в СМО. В текущий момент времени в системе может быть 0, 1, 2, …, n заявок, при этом все заявки обслуживаются. Вероятность того, что в системе находится k заявок, равна p_k. Среднее число заявок $n_{ср}$ равно математическому ожиданию дискретной случайной величины со следующей функцией распределения вероятностей

Значение с.в.	0	1	2	…	n
Вероятность p_k	p_0	p_1	p_2	…	p_n

$$n_{ср} = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + \cdots + n \cdot p_n = \sum_{k=1}^{n} k \cdot p_k;$$

Среднее число обслуживаемых заявок $n_{оз}$ - математическое ожидание дискретной случайной величины "количество занятых каналов" со следующей функцией распределения вероятности

Значение с.в.	0	1	2	…	n
Вероятность p_k	p_0	p_1	p_2	…	p_n

$$n_{оз} = 0 \cdot p_0 + 1 \cdot p_1 + \cdots + n \cdot p_n;$$

В рассматриваемом случае в СМО очереди нет, $n_{ср} = n_{оз}$. Средняя длина очереди равна нулю.

Система работоспособна, если $\frac{\lambda}{n\mu} < 1.$

Система с ограниченной очередью ($m \neq 0$.).

Вероятность отказа. Очередная заявка, поступающая в СМО, не обслуживается, если все каналы обслуживания и все места в очереди заняты. Это значит, что система находится в состоянии S_{n+m}. Вероятность нахождения системы в этом состоянии p_{n+m} является вероятностью "отказа" в обслуживании.

$$p_{отк} = p_{n+m} = \left[\frac{\lambda}{\mu}\right]^{n+m} \frac{1}{n^m n!} p_0;$$

Вероятность p_0 соответствует состоянию S_0, означающему, что в СМО нет заявок. Это *коэффициент простоя* p_0.

Величина $(1 - p_0)$ характеризует факт, что СМО занята обслуживанием заявок. Это *коэффициент загрузки* $p_{заг} = 1 - p_0$.

Пропускная способность А. Это количество обслуженных заявок, покидающих систему в единицу времени. Пропускная способность равна количеству входящих заявок в единицу времени, умноженное на вероятность того, что заявка будет обслужена.

$$A = \lambda \cdot (1 - p_{отк});$$

Среднее число заявок, находящихся в СМО. В текущий момент времени в системе может быть 0, 1, 2, …, n + m заявок, при этом заявки 1, … n могут обслуживаться, а n + 1, … n + m могут находиться в очереди. Вероятность того, что в системе находится k заявок, равна p_k. Среднее число заявок $n_{ср}$ равно математическому ожиданию дискретной случайной величины со следующей функцией распределения вероятности

Значение с.в.	0	1	2	…	n + m
Вероятность p_k	p_0	p_1	p_2	…	p_{n+m}

$$n_{ср} = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + \cdots + (n+m) \cdot p_{n+m} = \sum_{k=1}^{n+m} k \cdot p_k;$$

Среднее число обслуживаемых заявок $n_{оз}$ - математическое ожидание дискретной случайной величины "количество занятых каналов" со следующей функцией распределения вероятности (в состояниях S_n, … S_{n+m} обслуживается n заявок)

Значение с.в.	0	1	…	n	…	n
Вероятность p_k	p_0	p_1	…	p_n	…	p_{n+m}

$$n_{оз} = 0 \cdot p_0 + 1 \cdot p_1 + \cdots + n(p_n + p_{n+1} + \cdots + p_{n+m});$$

Средняя длина очереди $r_{оч}$. - математическое ожидание дискретной случайной величины со следующей функцией распределения (в первых состояниях S_0, … S_n очереди нет)

Значение с.в.	1	2	…	m
Вероятность p_k	p_{n+1}	p_{n+2}	…	p_{n+m}

$$r_{оч} = 1 \cdot p_{n+1} + 2 \cdot p_{n+2} + \cdots + m \cdot p_{n+m};$$

Средняя длина очереди резко возрастает при стремлении коэффициента загрузки $p_{заг}$ к единице.

$$p_{заг} = 1 - p_0 \to 1;$$

Для $p_{заг}$ больше единицы работа СМО в стационарном режиме невозможна.

Система с бесконечной очередью ($m = \infty$).

Вероятность отказа. В бесконечной очереди могут разместиться все пришедшие заявки. Поэтому вероятность отказа равна нулю.

Вероятность p_0 соответствует состоянию S_0, означающему, что в СМО нет заявок. Это *коэффициент простоя* p_0.

Величина p_0 находится из последнего уравнения системы уравнений Колмогорова, в которое входит сумма бесконечного числа слагаемых

$$p_0\left(1 + \left(\frac{\lambda}{\mu}\right) + \left(\frac{\lambda}{\mu}\right)^2 \frac{1}{2!} + \cdots + \left(\frac{\lambda}{\mu}\right)^n \frac{1}{n!} + \left(\frac{\lambda}{\mu}\right)^{n+1} \frac{1}{nn!} + \cdots \left(\frac{\lambda}{\mu}\right)^{n+m} \frac{1}{n^m n!} + \cdots\right) = 1;$$

В этом уравнении конечное число первых членов суммы со степенями 1,… n соответствуют n каналам обслуживания СМО. Остальные члены со степенями n + 1, …n + m, … соответствуют очереди и образуют убывающую геометрическую прогрессию со знаменателем $q = [\frac{\lambda}{\mu}]\frac{1}{n} < 1$ и первым членом $a_0 = \lambda / n\mu$.

Подставим в предыдущую формулу выражение для суммы членов геометрической прогрессии и получим формулу для вычисления p_0.

$$p_0\left(1 + \left(\frac{\lambda}{\mu}\right) + \left(\frac{\lambda}{\mu}\right)^2 \frac{1}{2!} + \cdots + \left(\frac{\lambda}{\mu}\right)^n \frac{1}{n!} + \left(\frac{\lambda}{\mu}\right)^{n+1} \frac{1}{nn!}\left(\frac{1}{1-[\frac{\lambda}{\mu}]\frac{1}{n}}\right)\right) = 1;$$

Эта формула позволяет найти вероятность p_0. Затем находятся остальные вероятности p_1, p_2, … p_{n+m}, … по формулам, приведенным выше. В системе с бесконечной очередью бесконечное число состояний S_i и соответственно бесконечная последовательность вероятностей { p_i }. Последовательность вероятностей { p_i } для состояний, соответствующих очереди, монотонно убывает. Поэтому продолжаем вычисление этих вероятностей до тех пор, пока соответствующее значение p_N окажется меньше ε, где ε - принятая погрешность.

Величина (1 - p_0) характеризует факт, что СМО занята обслуживанием заявок. Это *коэффициент загрузки* $p_{заг} = 1 - p_0$.

Пропускная способность A. Это количество обслуженных заявок, покидающих систему в единицу времени. В системе с бесконечной очередью все поступившие заявки будут обслужены. Поэтому

$$A = \lambda;$$

Среднее число обслуживаемых заявок $n_{оз}$ - математическое ожидание дискретной случайной величины "количество занятых каналов" со следующей функцией распределения вероятности (в состояниях Sn, … Sn+m обслуживается n заявок)

Значение с.в.	0	1	…	n	n	…	n	…
Вероятность p_k	p_0	p_1	…	p_n	p_{n+1}	…	p_{n+m}	…

$$n_{оз} = 0 \cdot p_0 + 1 \cdot p_1 + \cdots n(p_n + p_{n+1} + \cdots p_{n+m} + \cdots);$$

В этой формуле в скобках записана бесконечно убывающая геометрическая прогрессия с первым членом a_1

$$a_1 = p_0 \cdot \frac{n}{n!} \cdot \left(\frac{\lambda}{\mu}\right)^n \left(\frac{1}{1-\lambda/n\mu}\right)$$

и знаменателем

$$q = \lambda / n\mu.$$

В результате получим формулу для вычисления среднего числа обслуживаемых заявок в СМО.

$$n_{оз} = p_0\left(\left(\frac{\lambda}{\mu}\right) + \left(\frac{\lambda}{\mu}\right)^2 + \left(\frac{\lambda}{\mu}\right)^3 \frac{1}{2} + \left(\frac{\lambda}{\mu}\right)^4 \frac{1}{2\cdot 3} + \cdots \left(\frac{\lambda}{\mu}\right)^{n-1} \frac{1}{(n-2)!} + \left(\frac{\lambda}{\mu}\right)^n \frac{1}{(n-1)!}\left(\frac{1}{1-\lambda/n\mu}\right)\right);$$

Средняя длина очереди $r_{оч}$. - математическое ожидание дискретной случайной величины со следующей функцией распределения (в первых состояниях S_0, … S_n очереди нет)

Значение с.в.	1	2	…	m	…
Вероятность p_k	p_{n+1}	p_{n+2}	…	p_{n+m}	…

Очередь в рассматриваемом случае бесконечна, однако, средняя длина очереди - конечное число.

$$r_{оч} = 1 \cdot p_{n+1} + 2 \cdot p_{n+2} + \cdots k \cdot p_{n+k} + \cdots$$

Для вычисления $r_{оч}$ преобразуем бесконечный ряд следующим образом

$$r_{оч} = 2 \cdot p_{n+1} + 3 \cdot p_{n+2} + \cdots (k+1) \cdot p_k + \cdots - p_{n+1} - p_{n+2} - \cdots p_k - \cdots$$

Разобьем этот ряд на два. Сумму положительных членов обозначим S_1, а сумму отрицательных - S_2,

$$r_{оч} = S_1 - S_2.$$

Вычислим сумму первого ряда S_1. Подставим в ряд формулы для p_k, получим

$$S_1(x) = \frac{\lambda^n p_0}{\mu^n n!}(2 \cdot x + 3 \cdot x^2 + 4 \cdot x^3 + \cdots (k+1) \cdot x^k + \cdots),$$

где $x = \lambda / n\mu.$

Ряд $S_1(x)$ сходится равномерно и поэтому его сумма является непрерывной функцией от x. Этот ряд можно почленно интегрировать и дифференцировать. Вычислим интеграл от суммы ряда

$$\int S_1(x)dx = \frac{\lambda^n p_0}{\mu^n n!}(x^2 + x^3 + x^4 + \cdots x^{k+1} + \cdots);$$

В этой функцией в скобках стоят члены геометрической прогрессией с первым членом $x^2 = (\lambda / n\mu)^2$ и знаменателем $x = \lambda / n\mu.$

Запишем сумму этого ряда

$$\int S_1(x)dx = \frac{\lambda^n p_0}{\mu^n n!}\frac{x^2}{1-x};$$

Продифференцируем эту формулу по x и получим сумму ряда S_1(x).

$$S_1(x) = \frac{\lambda^n p_0}{\mu^n n!}\frac{2x-x^2}{(1-x)^2};$$

Второй ряд S_2(x) - геометрическая прогрессия с первым членом и знаменателем $x = \lambda / n\mu.$

Сумма ряда S_2 равна

$$S_2 = \frac{\lambda^n p_0}{\mu^n n!}\frac{x}{1-x} = \frac{\lambda}{n\mu(1-\lambda / n\mu)};$$

Объединяя формулы для S_1 и S_2, получим

$$r_{оч} = \frac{\lambda^{n+1}}{\mu^{n+1} nn!(1-\lambda / n\mu)};$$

Суммируя среднее число обслуживаемых заявок и среднее число заявок в очереди, получим среднее число заявок, находящихся в СМО $n_{ср}$.

$$n_{ср} = n_{оз} + r_{оч}.$$

Среднее время пребывания заявки в СМО. В стационарном режиме работы СМО ($p_{заг}$ <1) интенсивности входного и выходного потоков равны между собой и равны λ. Пусть в стационарном режиме среднее число заявок в системе в единицу времени равно $n_{ср}$. Тогда согласно теореме Литтла среднее время пребывания заявки в СМО - $T_{сист}$ равно

$$T_{сист} = n_{ср} / \lambda;$$

Аналогично можно записать формулу для среднего времени пребывания заявки в очереди $T_{оч}$. В общем случае

$$T_{оч} = r_{ср} / \lambda;$$

Эти формулы позволяют определить среднее время обслуживания заявки СМО - $t_{обсл}$ из следующего соотношения

$$T_{сист} = t_{обсл} + T_{оч};$$

В заключение отметим, что простые аналитические решения получены только для простейших систем массового обслуживания с входными потоками заявок экспоненциального или Пуассоновского типов. Поэтому эти результаты необходимо рассматривать как приближенные и оценочные. В аналитических моделях не удается учесть приоритеты заявок, реальные законы распределений, процессы обработки заявок в системе и многое другое. Поэтому в настоящее время основным методом анализа СМО является имитационное моделирование, позволяющее в численном виде получить характеристики системы, близкие к реальным.

2. ЧИСЛЕННЫЕ МЕТОДЫ МОДЕЛИРОВАНИЯ СМО.

2.1. Моделирование СМО как Марковского процесса.

Вероятность состояний СМО можно найти в процессе численного моделирования Марковского процесса. Рассматриваемый далее алгоритм может применяться для любых Марковских процессов с дискретными состояниями и непрерывным временем. Моделирование возможно, если случайное время перехода системы из одного состояния в другое распределено по экспоненциальному закону. В результате численного моделирования можно получить только вероятности состояний СМО. Остальные характеристики СМО определяются дополнительно по формулам, рассмотренным выше.

Графу состояний трехканальной СМО можно сопоставить следующую матрицу интенсивностей переходов Марковского процесса.

$$\Lambda = \left\| \lambda_{ij} \right\| = \begin{Vmatrix} 0 & \lambda & 0 & 0 \\ \mu & 0 & \lambda & 0 \\ 0 & 2\mu & 0 & \lambda \\ 0 & 0 & 3\mu & 0 \end{Vmatrix}$$

Рис. 5. Матрица интенсивностей переходов для СМО

В общем случае моделирования Марковского процесса размерность квадратной матрицы равна количеству состояний m и она не является трехдиагональной.

Будем для определенности считать, что строки и столбцы матрицы нумеруются с 0. Тогда λ_{ij} — интенсивность перехода в состояние *j* при условии, что система находится в состоянии *i*. Логично задать начальное состояние системы *i*=0 (нет заявок). Тогда, если *i* — номер текущего состояния, алгоритм моделирования произвольного Марковского процесса работает следующим образом:

1. Установить текущее состояние i = 0, обнулить счетчик модельного времени T и счетчики времени пребывания в k-том состоянии T_k, k = 0, …m-1, а также счетчик числа обслуженных заявок N.
2. Для каждого состояния j = 0, … m – 1 генерируется случайная величина – время до перехода из текущего состояния i в состояние j. Это время рассчитывается по формуле
$$t_{ij} = -\frac{1}{\lambda_{ij}} \ln u$$
где u – случайное число с равномерным законом распределения на интервале [0, 1). Это генератор случайных чисел с экспоненциальным законом распределения с интенсивностью λ_{ij}, построенный по методу обратной функции.
3. Среди полученных времен t_{ij} находим минимальное. Другими словами находим индекс j такой, что для него время t_{ij} минимально при изменении $0 \le j \le$ m-1. Найденное минимальное время определяет действительное время

пребывания системы в текущем состоянии i и показывает новое действительное состояние системы j.

4. Увеличиваем счетчик модельного времени T и счетчик времени пребывания системы в состоянии i на величину t_{ij} .

$$T = T + t_{ij}: \qquad T_i = T_i + t_{ij};$$

5. Если $j < i$, то увеличить число в счетчике обслуженных заявок N на единицу.

6. Если N меньше заданного числа обслуженных заявок, то перейти в следующее состояние j. (i = j), а затем перейти к п.2.

Если N равно заданному числу заявок – СТОП.

Алгоритм определяет новое состояние, время перехода в которое минимально, и определяет, находится ли новое состояние выше или ниже главной диагонали. Для первого случая принимается решение, что пришла очередная заявка, для второго – обслуженная заявка покинула систему.

На каждом шаге алгоритма расчет времени до ближайшего перехода производится заново. Это допустимо в рамках модели Марковского процесса, в которой условная вероятность перехода из одного состояния в другое равно безусловной вероятности. Предысторию в Марковском процессе учитывать не нужно.

Вероятности состояний рассчитываются как отношение времени, в течение которого система находилась в данном состоянии, к общему времени моделирования, т.е.

$$p_i = \frac{T_i}{T}$$

Для случаев с конечным количеством состояний m накапливать модельное время необязательно — ведь T равно сумме всех T_i.

Очевидно, что для случая СМО возможен переход процесса из текущего состояния i только в два соседних состояния: $j = i + 1$ и $j = i - 1$. Это позволяет упростить алгоритм моделирования. Рассмотрим два более простых алгоритма моделирования на примере СМО с бесконечной очередью.

При моделировании СМО с бесконечной очередью возникает ряд трудностей, связанных с бесконечной размерностью матрицы интенсивностей переходов.

Решение 1.

Заметим, что матрица интенсивностей переходов для СМО – трехдиагональная (рис. 4). В такой матрице все отличные от нуля элементы размещены на двух диагоналях, прилегающих к главной. Таким образом, если $j = i + 1$, то переход происходит к элементу j, лежащему выше главной диагонали (интенсивность равна λ). Если $j = i - 1$, то переход происходит к элементу j , лежащему ниже главной диагонали.

Эти простые условия позволяют описать матрицу интенсивностей переходов функцией языка программирования.

Необходимо вычислять случайное время перехода из текущего состояния i только в два соседних состояния: $j = i + 1$ и $j = i - 1$. Переходы в другие состояния невозможны.

Необходимо подсчитывать время пребывания системы только в тех состояниях, которые достигались в действительности. Суммарное время пребывания системы во всех действительных состояниях равно модельному времени.

Решение 2. Ввести искусственные ограничения на длину очереди и решить конечную задачу, подобрав в эксперименте достаточно большую (например, 40 или 100) максимальную длину очереди. В процессе моделирования необходимо проверять, что бы очередь ограниченной длины полностью не заполнялась. Тогда результаты моделирования СМО с конечной очередью будут полностью соответствовать системе с бесконечной очередью. Если последнее состояние в очереди достигается, необходимо увеличить ее длину.

Характеристики СМО, полученные в результате моделирования Марковского процесса необходимо сравнить с характеристиками, полученными в результате аналитического расчета.

2.2. Моделирование СМО по событиям.

Компьютерное моделирование СМО на основе Марковского процесса возможно, если потоки заявок и потоки обслуживания распределены по экспоненциальному закону. В общем случае для произвольных распределений этот метод не применим.

Для моделирования СМО с произвольными распределениями потоков заявок и потоков обслуживания применяется метод имитационного моделирования по событиям. Событиями в СМО называются моменты времени, когда очередная заявка приходит в систему и когда обслуженная заявка покидает систему. Другими словами, это моменты изменения состояния системы. Эти моменты образуют список событий, в котором можно выделить текущее событие и списки прошедших и будущих событий. Метод имитационного моделирования включает последовательное формирование списка событий и продвижение процесса моделирования от одного события к другому. Модельное время при имитационном моделировании является зависимой переменной и изменяется скачками, равными интервалам времени между последовательными событиями.

Схема алгоритма имитационного моделирования по событиям приведена на рис. 6.

Принятые обозначения:

Входные параметры:

N — максимальное число входящих заявок (условие окончания моделирования).

NK — количество каналов.

LMAX — максимально допустимая длина очереди.

Накапливаемая статистика:

T — текущее модельное время, изменяющееся скачком между моментами возникновения событий в системе.

TA — момент прихода очередной заявки (время, прошедшее с начала моделирования).

K — счетчик пришедших заявок.

KS — счетчик обслуженных заявок.

L — текущая длина очереди.

LOS — счетчик отказов (заявок, поучивших отказ в обслуживании).

Массивы:

OCP[i] — признак занятости i-го канала (0 — канал свободен, 1 — канал занят).

TD[i] — ожидаемый момент выхода заявки из i-го канала (время, прошедшее с начала моделирования).

TOS[i] — счетчик суммарного времени занятости i-го канала — сколько единиц модельного времени в течение всего процесса моделирования канал был занят.

TL[M] — суммарное время пребывания системы в состоянии, когда в ней ровно M заявок.

Вспомогательные величины:

MIN — ближайший момент выхода обслуженной заявки из канала, считая от текущего модельного времени.

S — номер канала, который в текущем состоянии системы освободится первым (в момент времени MIN).

DTA — время между приходами заявок (генерируемая в процессе моделирования случайная величина).

DTS — время обслуживания заявки в канале (генерируемая в процессе моделирования случайная величина).

Для рассматриваемой СМО

$$DTA = -\frac{1}{\lambda}\ln u_1,$$

$$DTS = -\frac{1}{\mu}\ln u_2 = t_{обсл.}\ln u_2$$

где u_1, u_2 — равномерно распределенные на интервале [0, 1) случайные величины (рекомендуется использовать для их генерации два независимых датчика ррсч[0,1)).

В процессе моделирования в любой момент времени известны моменты наступления событий для всех возможных источников (момент прихода очередной заявки и моменты освобождения каждого из каналов). Если канал свободен, то считается (с целью упрощения алгоритма), что ожидаемое время его освобождения бесконечно велико; в предлагаемой схеме алгоритма в качестве заведомо большого значения взято число 999999.

Установка начальных значений (блок 2) состоит в заполнении массива TD большими числами (999999), а всех переменных и других массивов — нулевыми значениями. Это соответствует ситуации, когда все каналы свободны и модельное время равно 0.

Блок 3 определяет момент прихода очередной заявки, при этом увеличивается счетчик пришедших заявок K, модельное время не изменяется.

Блок 4 — проверка условия окончания моделирования.

Блоки 7-11 реализуют определение номера канала, который освобождается первым, и момента его освобождения. При начальном запуске системы все каналы свободны, поэтому номер канала S не важен (вообще-то это будет последний канал), а в качестве момента его освобождения MIN будет обнаружено заведомо большое число.

Далее происходит проверка: какое из событий наступит раньше, приход заявки в момент времени TA или освобождение канала в момент времени MIN (блок 12)? При начальном запуске ближайшим событием окажется приход заявки (так как все каналы свободны) и произойдет переход на блок 19.

Блок 19 осуществляет накопление статистики по пребыванию системы в состоянии с M заявками. Очевидно, что M равно сумме количества занятых каналов (суммы всех элементов массива OCP[i]) и количества заявок в очереди L. Если текущее модельное время T, а переход в новое состояние произойдет в момент времени TA по приходу заявки, то в состоянии M система пробудет (TA□T) единиц модельного времени.

Блок 20 изменяет модельное время. Теперь состояние системы можно сформулировать как "очередная заявка только что пришла в систему".

Блоки 21-24 организуют цикл поиска свободного канала. Если свободный канал найден, то происходит переход на блок 25 и досрочный выход из цикла поиска.

Блоки 25-26 реализуют действия алгоритма моделирования в случае, когда в системе есть свободный канал. В этом случае канал объявляется занятым (блок 25) и сразу же разыгрывается время обслуживания заявки в канале DTS и момент его освобождения TD[i], равный текущему модельному времени T плюс DTS. Сразу же в массиве TOS[i] учитывается и время, которое занимаемый канал потратит на обслуживание заявки.

Если в системе нет свободных каналов, то заявка помещается либо в очередь, либо, если очередь уже имеет максимально допустимую длину, получает отказ. Эта логика реализуется блоками 27-29.

После обработки прихода очередной заявки в блоках 19-29 происходит переход на блок 3, где происходит определение момента прихода очередной заявки.

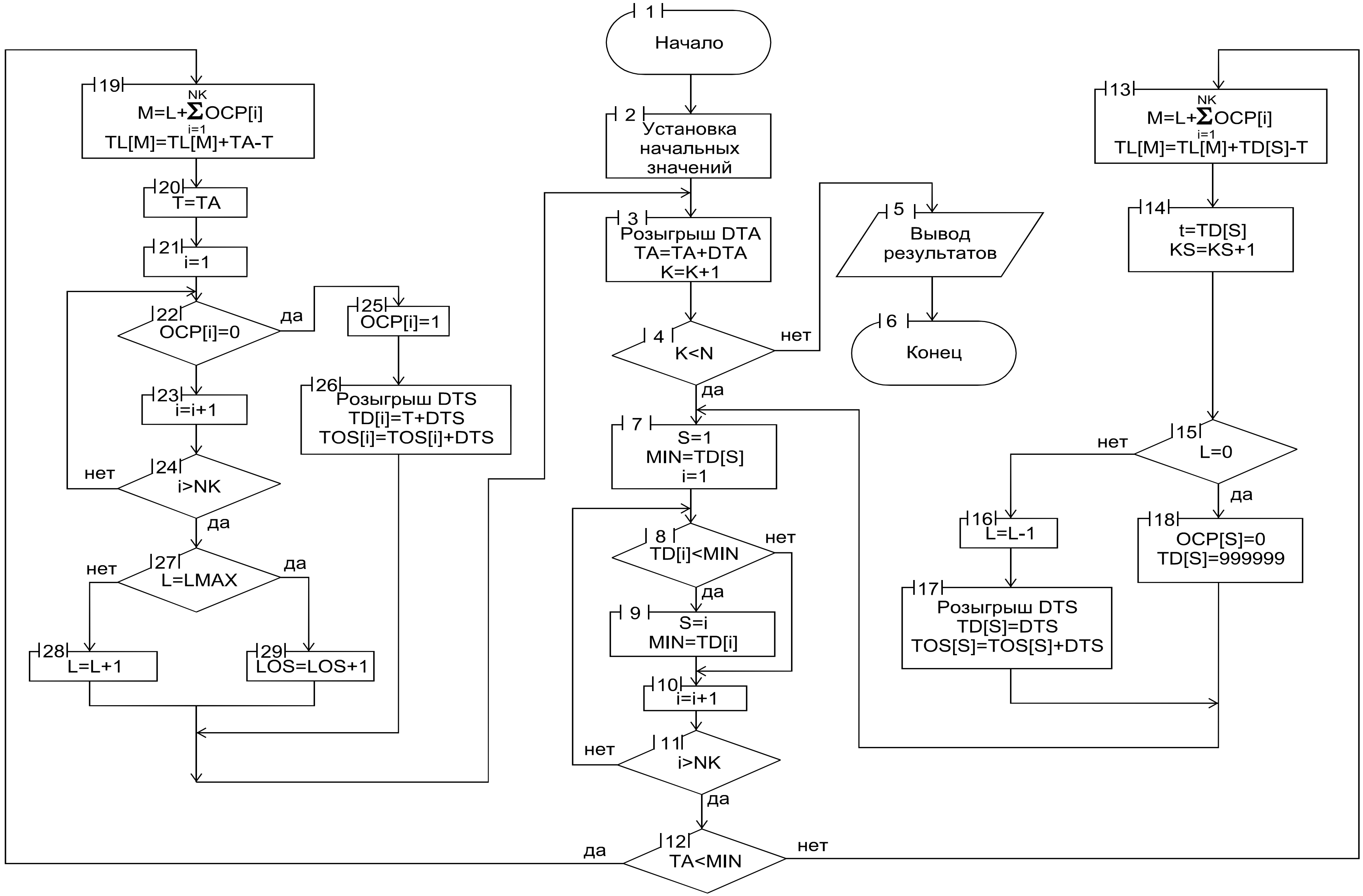

Рис. 6. Схема алгоритма моделирования многоканальной СМО по событиям

Рассмотрим теперь вариант, когда текущее состояние системы таково, что приход очередной заявки случится позже, чем освобождение одного из каналов. При этом в результате выполнения поиска первого освобождающегося канала в блоках 7-11 будет определен номер S канала, освобождающегося первым, и момент его освобождения MIN. Переход из блока 12 произойдет на блок 13.

Блок 13 идентичен блоку 19 и осуществляет накопление статистики по пребыванию системы в состоянии с M заявками.

Блок 14 имитирует окончание обслуживания заявки в канале S. Счетчик обслуженных заявок увеличивается на единицу, а модельное время становится равным времени MIN выхода заявки из канала S.

В блоке 15 осуществляется проверка наличия заявок, ожидающих в очереди.

Если заявки в очереди есть, то ожидающая заявка сразу направляется в только что освободившийся канал S, генерируется время ее обслуживания DTS и определяется момент окончания обслуживания TD[S]. Сгенерированное время DTS учитывается также в статистике времени занятости канала TOS[S] (блоки 16-17).

Если очередь пуста, то канал S объявляется свободным и в качестве момента его освобождения TD[S] устанавливается заведомо большое число.

Для систем с бесконечной этот алгоритм обычно тоже применим, при этом в LMAX должно быть установлено достаточно большое число (например 100).

Вычисление характеристик СМО. По окончании моделирования основные характеристики СМО могут быть получены следующим образом:

1. Вероятности пребывания в состоянии, когда в системе ровно *i* заявок, определяется как доля времени, в течение которого в системе было *i* заявок, в общем времени моделирования:

$$p_i = \frac{TL[i]}{T}$$

2. Вероятность отказа — доля заявок, получивших отказ, в общем количестве заявок:

$$p_{отк.} = \frac{LOS}{N}$$

3. Пропускная способность — количество обслуживаемых в единицу времени заявок:

$$A = \frac{KS}{T}$$

4. Среднее количество занятых каналов можно определить как отношения суммарного времени занятости всех каналов к общему времени моделирования:

$$k_{ср.} = \frac{\sum_{i=1}^{NK} TOS[i]}{T}$$

5. Среднее время ожидания в очереди можно определить как суммарное время ожидания заявок в очереди, отнесенное к количеству входящих заявок. Если B — это время, проведенное всеми заявками в системе, и C — время, проведенное всеми заявками в каналах, то их разность — время, проведенное всеми заявками в очереди. Тогда

$$B = \sum_{m=0}^{LMAX+NK} m \cdot TL[m], \quad ; \quad C = \sum_{i=1}^{NK} TOS[i], \quad t_{ож.} = \frac{B - C}{N}$$

6. Средняя длина очереди может быть вычислена как математическое ожидание количества заявок в очереди в процессе моделирования:

$$r_{ср.} = \sum_{i=NK+1}^{NK+LMAX} \frac{TL[i]}{T} \cdot (i - NK) = \frac{\sum_{i=NK+1}^{NK+LMAX} TL[i] \cdot (i - NK)}{T}$$

Полученные по этим формулам характеристики СМО необходимо сравнить с характеристиками, полученными в результате аналитического расчета.

3. ПОРЯДОК ВЫПОЛНЕНИЯ ЛАБОРАТОРНЫХ РАБОТ

Лабораторная работа состоит из двух частей. :

Первая часть:

Составление и аналитическое решение стационарного варианта уравнений Колмогорова для двух вариантов индивидуального задания:

- Одноканальная СМО (пункт 12 календарного плана) и
- Многоканальная СМО (пункт 13 календарного плана).

Для каждого варианта необходимо:

- Составить список состояний СМО;
- Нарисовать граф состояний СМО;
- Составить и решить систему уравнений Колмогорова;
- Рассчитать следующие основные характеристики СМО:

1. Вероятности нахождения системы в каждом состоянии;
2. Вероятность отказа;
3. Коэффициент простоя;
4. Коэффициент загрузки;
5. Пропускная способность;
6. Среднее число заявок, обслуживаемое в СМО;
7. Средняя длина очереди;
8. Среднее количество занятых каналов;
9. Среднее время пребывания заявки в СМО;
10. Среднее время пребывания заявки в очереди;

По результатам аналитического решения составляется краткий отчет, который направляется на проверку преподавателю.

После проверки правильности расчетов и исправления ошибок проводится тест на знание теоретического материала.

Вторая часть:

Численное моделировании СМО для двух вариантов индивидуального задания, рассмотренных в первой части задания:

- Одноканальная СМО (пункт 21 календарного плана) и
- Многоканальная СМО (пункт 22 календарного плана).

Моделирование проводится по программам, предоставленным преподавателем.

Необходимо составить таблицу, в которой сравнить характеристики СМО, полученные в результате аналитического и численного расчета.

По результатам численного решения составляется краткий отчет, который направляется на проверку преподавателю.

После проверки правильности расчетов и исправления ошибок проводится тест на знание теоретического материала, в частности на знание алгоритмов численного моделирования СМО.

Отчеты должны содержать:

1. Исходные данные для моделирования СМО в соответствии с вариантом задания;
2. Список состояний СМО;
3. Граф состояний СМО;
4. Уравнения Колмогорова;
5. Вероятности нахождения СМО в каждом состоянии;
6. Основные характеристики СМО, полученные аналитически,
7. Основные характеристики СМО, полученные численно,
8. Анализ полученных результатов.

ЗАДАНИЯ
к лабораторной работе «Моделирование одноканальной СМО» (часть 1)

№	Длина очереди	Интенсивность потока заявок λ	Среднее время обслуживания 1/μ	Метод численного моделирования
1	3	1.5	0.5	Марковский пр-с
2	5	1	0.5	Марковский пр-с
3	4	2	0.3	Марковский пр-с
4	3	1.5	0.3	Марковский пр-с
5	4	2	0.25	Марковский пр-с
6	5	2.5	0.2	Марковский пр-с
7	3	1.5	0.25	Марковский пр-с

ЗАДАНИЯ
к лабораторной работе «Моделирование многоканальной СМО» (часть 2)

№	Число каналов	Длина очереди	Интенсивность потока заявок λ	Среднее время обслуживания 1/μ	Метод численного моделирования
1	2	≤ 3	2	1	По событиям
2	3	≤ 2	1	2.5	По событиям
3	3	0	3	1.2	По событиям
4	4	0	2	2	По событиям
5	2	∞	2	0.8	По событиям
6	3	∞	1.5	1.6	По событиям
7	2	≤ 3	1.5	1.5	По событиям

Контрольные вопросы.

1. Дайте определение Марковского процесса с дискретным временем.
2. Составьте матрицу переходных вероятностей размерности 4 Х 4 для Марковского процесса с дискретным временем. Численные значения переходных вероятностей определите самостоятельно.
3. Каким свойством обладает сумма переходных вероятностей, записанных в строках матрицы переходных вероятностей ?
4. Каков математический (физический) смысл этого свойства с точки зрения теории вероятностей?
5. Нарисуйте граф переходов для составленной в п. 2 матрицы переходных вероятностей.
6. Дайте определение Марковского процесса с непрерывным временем.
7. Составьте матрицу переходных интенсивностей размерности 4 Х 4 для Марковского процесса с непрерывным временем. Численные значения переходных интенсивностей определите самостоятельно.
8. Каковы размерности вероятности и интенсивности потока событий?
9. Нарисуйте граф переходов для Марковского процесса с матрицей интенсивностей, составленной в п. 7.
10. Нарисуйте структурную схему одноканальной СМО с очередью.
11. Составьте список состояний СМО, имеющей один канал обслуживания и три места в очереди.
12. Нарисуйте граф переходов для СМО, имеющей один канал обслуживания и три места в очереди.
13. Запишите систему дифференциальных уравнений Колмогорова для этой СМО.
14. Задайте и поясните смысл начальных условий для этой системы.
15. Разберите понятия переходного и стационарного режимов работы СМО.
16. Запишите вариант стационарной системы уравнений Колмогорова для этой СМО. Какое условие заменяет начальные условия в стационарной системе. Какой смысл имеет это условие с точки зрения теории вероятности.
17. Решите стационарную систему Колмогорова для этой СМО.
18. Рассчитайте основные характеристики СМО.
19. Проведите численный расчет этого варианта СМО по одной из программ численного моделирования. Сравните результаты аналитического и численного расчетов .